Addressing Agricultural Salinity in the American West: Harnessing Behavioral Diversity to Institutional Design

By

Beaudry E. Kock

S.M. Civil and Environmental Engineering
Massachusetts Institute of Technology
Cambridge, MA (2008)

MSci. Environmental Geology
Imperial College London
London, U.K. (2005)

Submitted to the Department of Urban Studies and Planning
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Urban and Regional Planning

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2010

Author  _____
Department of Urban Studies and Planning
May 5, 2010

Certified by  _____
Professor Michael Flaxman
Department of Urban Studies and Planning
Dissertation Supervisor

Accepted by  _____
Professor Eran Ben-Joseph
Chair, PhD Committee
Department of Urban Studies and Planning

Addressing Agricultural Salinity in the American West: Harnessing Behavioral Diversity to Institutional Design

by

Beaudry E. Kock

Submitted to the Department of Urban Studies and Planning
on May 7, 2010 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy in
Urban and Regional Planning

ABSTRACT

Salinity accumulation in the Lower Arkansas Basin (LAB) of Colorado threatens environmental quality, the agricultural economy and the potential for efficient reuse of water. Salinity is a threat to "hydraulic sustainability", since it will affect societal and environmental sustainability in a system heavily dependent on engineered structures for its water supply. Institutional solutions are preferable, being usually cheaper, quicker, and more reversible than infrastructure. Market institutions — water quality trading markets — have been often applied in the past to deal with salinity problems, but have been largely ineffective despite theoretical promise. Explanations for such institutional failure typically assume that stakeholders are boundedly–rational economic actors, but I review evidence that this is empirically unjustified, may be insufficiently explanatory, and precludes consideration of more innovative behavioral change solutions. Through collaborative work with basin stakeholders, I developed an agent-based model — "ArkAgent" — which simulates a water quality trading market; the water use and market interactions of basin actors; and basin hydrology. I conduct experiments to show that a simulated neoclassical market institution is less effective at reducing salinity when we make more realistic provisions for attitudinal and behavioral heterogeneity among resource users. I show that the use of post-hoc informational feedbacks as alternative non-monetary institutional incentives can address this performance issue, even in the face of conflicting economic pressures. I further demonstrate that exploiting social networks in non-economic incentive design can go even further in improving sustainability benefits. This work makes new theoretical contributions by showing how our models of institutional performance are critically dependent on behavioral assumptions; and that consequently our institutions for addressing hydraulic sustainability challenges may have incentives poorly matched to real behavioral complexity. This work also shows how an appropriately designed market institutional intervention in the LAB could achieve salinity reduction benefits over an 8 year period. Many of the model's practical insights are also relevant to large salinity-threatened basins across the western United States. The ArkAgent model provides an example of how we can use collaborative systems modeling and empirically-based behavioral assumptions to develop more robust institutions for sustainability.

Thesis Supervisor: Michael Flaxman
Title: Assistant Professor of Urban Technologies and Information Systems

# Acknowledgements

# Contents

# Chapter 1

# Introduction

the human mind is something of an embarrassment to certain disciplines, notably economics, that have found the model of the rational consumer to be powerfully productive (Schelling, 1984, 342)

## 1.1 Problem statement

In this thesis I tackle the problem of institutional failure, nested within the broader issue of hydraulic sustainability.

The lower Arkansas Basin (LAB) of Colorado is a quintessential example of a "hydraulic society": a social-hydrologic complex characterized by an extensive and highly engineered water collection and delivery infrastructure (Worster, 1982).The LAB, like many other hydraulic societies globally, faces problems of deteriorating environmental conditions, particularly with regard to salt levels in basin waters and soils (Miles, 1977; Goff et al., 1998). High levels of salt lead to reduced crop yield, challenging water supply conditions, and extensive ecological damage (Childs & Hanks, 1975). Reducing salinity levels in the LAB is a critical problem that has not yet found a clear solution. "Hydraulic sustainability" is a term I use to characterize the broader challenges of moving hydraulic societies along the path towards longer term social, economic and environmental sustainability. Salinity reductions in the LAB are clearly a component of achieving hydraulic sustainability in the basin.

One approach to addressing this problem could involve making infrastructural modifications to remove the excess salt from the system, such as constructing desalination plants or adding extra reservoir storage. However, infrastructural modifications are expensive and ecologically disruptive, increasingly rendering their viability in question. An alternative to such options is institutional intervention (institutions are defined as the rules by which human actions and interactions are governed, Ostrom (2005)). The capacity to use institutions to foment change in resource-based systems has been well explored in fields as diverse as common pool resource studies (Huerta, 2008), economics (Nishibe & Uemura, 2005) and sustainable development (Veeman & Polityло, 2003). Assuming that institutions do have some capacity to address complex social-environmental problems within large systems, the important question to ask is which institutional form will be most effective given the nature of the problem and its social-environmental context.

The success of the 1990 sulphur dioxide emissions permit trading program led to widespread calls for the use of market-based instruments to reduce water pollution (Stavins, 1998). Despite the implementation of water quality trading markets (WQTMs) in multiple and diverse settings across the United States in the last decade, these institutional interventions have so far failed to achieve their theoretical benefits (King, 2005). The institutional failure of water quality trading markets is particularly problematic: other institutional forms have been similarly unsuccessful, while WQTMs are, in principle, ideally suited to handling the thorny issue of non-point source pollution, which is the primary cause of the salinization problem in the LAB. Non-point source pollution problems present considerable monitoring and enforcement challenges, due to the diffuse spatial distribution of sources and the relatively minor contribution of any one source. In theory, WQTMs allow the information requirements in regulating pollution reductions to be decentralized, and provide for behavioral and technological flexibility in how individual polluters achieve their reductions.

A WQTM has never been implemented in the LAB, and the broader experience with water markets in the basin has been dismal. In the late 1980s, a formal water quantity trading market was introduced in the LAB, but was closed after failing to complete a single transaction. Water trading activity in the area has since been limited to infrequent temporary leases between

major water uses (Brown, 2006), and a small number of very large water sales. Neither are characteristics of a dynamic market, and indeed, limited liquidity has been a reality of water quantity and quality markets across the United States.

In this thesis I tackle the following core problem: *the singular failure of a particular institutional form, the water quality trading market, to succeed in furthering the cause of hydraulic sustainability by reducing non-point source pollution.* Specifically, I focus on the puzzle of why the WQTM institution has largely failed to incentivize participation on the part of water users. My broader concern is not proving or disproving the worth of market-based mechanisms in supporting hydraulic sustainability. Consequently, my focus is on *institutional failure*, not market failure. I also avoid addressing every possible source of failure in what are highly contextual institutions with complex mechanisms. I intend instead to test alternative explanations that have hitherto received little attention, and in so doing shed practical light on possible solutions for the real salinity problems of the LAB.

## 1.2   Theoretical frame

Water is a common pool resource (CPR), an ecosystem service whose "yield is subtractable and the exclusion from which is nontrivial" (Ostrom et al., 1992, 404). As with other common pool resources, water resource management questions are often framed as "social dilemmas" (Ostrom, 1998): situations where the maximization of self-interest by individuals leaves everyone else worse off. Collective action theories, and studies of institutions, have all sought to address the questions in these terms.

The common problem of salinization in urban-agricultural water systems is a concrete example of a social dilemma. It is in the self-interest of individual polluters to continue polluting, because the costs of their pollution are borne by downstream users. The LAB appears to provide a classic example of just such a social dilemma. The most upstream irrigators along the Arkansas River can make use of extremely high quality water, with a very low dissolved salt content. Irrigators at the downstream end of the system, on the other hand, make use of water that is so salty that certain crops will fail entirely. The problem of salinity in the LAB is also

one of long term sustainability: by threatening crop yields, salinity depresses an already fragile agricultural economy and contributes to a demographic decline among farmers and the rural communities dependent upon farming; by reducing biodiversity in riparian settings and in the High Plains soils, salinity reduces the resilience of the system to invasion by exotic species, and reduces the potential for more diverse land uses.

If we take the assumptions of the social dilemma formulation to be correct, achieving sustainability becomes a matter of finding the right institutional design/mix which ensures that the propensity for individuals to pursue their self-interest is mitigated, and that (dis)incentives are deployed to encourage individual action in the cause of the common good. The debate then often boils down to market-based versus regulatory systems: markets leverage self-interest to minimize the social cost of sustainability, while regulations curb self-interest for the common good. Arguments for and against either institutional option (or a mix of the two) are predicated on the assumptions inherent in the rational choice theorem (Geanakoplos, 2004) (or its principal variant, the theory of bounded rationality Simon (1990)). It is important to emphasize that the space of conceivable institutions is constrained by these behavioral assumptions.

In this thesis, I implicitly pose the question: what if the social dilemma is no dilemma at all? If we dispense with self-interest as being *necessarily* the prime mover of resource use, the concept of a "social dilemma" needs to be reformulated. Instead of starting from the presumption that all individuals act self-interestedly, and so need to be enticed or cudgeled into cooperative behavior, we could start from the presumption that some individuals will act in ways that do not a) necessarily attempt to maximize a single or even multi-utility function, and b) are not necessarily solely motivated by self-interest, either in the short or long term. By relaxing our assumptions on the homogeneity of behavioral strategies across a population, we open up the CPR debate to more inventive and flexible institutional forms.

This is not an outlandish proposal. A wealth of empirical studies suggest we should move beyond traditional assumptions of either rational or boundedly rational choice as constituting the behavioral models of resource users. Such studies indicate that pro-social, longer term, and dynamic behavioral strategies really do exist among resource users (Norton et al., 1998; Bowles & Gintis, 2004; Guagnano, 2001; van den Bergh et al., 2000). This empirical evidence is

directly antithetical to neoclassical assumptions of self-interest, rigidly short term discounting and non-transitive preferences. Such studies also indicate that resource users may not even behave *as if* they are following these strategies, regardless of what their actual motivations are (Gowdy, 2008). Theoretical critiques of the foundations of rational and boundedly rational choice theories point out that such theories are based on untestable and tautological assumptions: the assumption of self-interest can be extended to any potential human motivation (in this framing, even non-self-interested motivations are self-interested), and so essentially contributes nothing substantive to our understanding of how resource users come to decisions (Chouinard et al., 2008). In Sen's words, this is a "robust piece of evasion" (Sen, 1977, 323). Many rational choice theoreticians have argued that even in the absence of empirical evidence to support their choice of behavioral model, the self-interest hypothesis is justified by methodological parsimony (Tirole, 2002). But even acknowledging Friedmans tenuous argument that the plausibility of assumptions is less important than their predictive power (Friedman, 1953), we are left with the reality that rational and boundedly rational choice theories consistently fail to predict resource use behavior when their restrictive governing assumptions are relaxed. Furthermore, following such traditional assumptions can have damaging effects on the way we pursue stable and sustainable societies (Sanchez-Cuenca, 2008). Note that the focus of my critique is not on the normative component of the neoclassical hypothesis, which suggests that resource users *should* behave in an economically rational manner because economic rationality is more natural or simply "better" in some dimension (Coase, 1960; Parisi, 2004). My concern is with what it takes to incentivize sustainable behavior; whether or not the behavior is economically rational is beside the point. I simply argue that all discussions of appropriate incentives must be rooted in a proper understanding of how resource users *actually* behave.

Assuming more behavioral heterogeneity among resource users has implications for how we study and design institutions for sustainability. By "behavioral heterogeneity" I mean a departure from rational and boundedly rational choice theories, and a more eclectic attitude towards what actually drives resource user decision making in a given setting. Exploring the institutional implications of this departure constitutes the central theoretical theme of this thesis.

## 1.3 Practical approach

The Lower Arkansas Basin of Colorado is a complex social-environmental system, incorporating a diversity of actors and biophysical processes. In pursuing the theoretical goals outlined above, I also seek to explore the practical benefits of institutional intervention in a complex system for the benefit of local and global sustainability. I develop an agent-based model and integrate it with a hydrogeological simulation tool, so that I can represent both social dynamics - the decisions and actions of farmers, ditch managers, municipal and federal water officials - and environmental dynamics - the movement of water in surface and subsurface, and the fate and transport of dissolved salts. I apply this integrated modeling tool to explore variations on one potential solution to the salinity issues in the LAB, an institutional intervention relying on market mechanisms. While my hypotheses are principally designed to develop my theoretical arguments, an important benefit of their formulation is to demonstrate the potential environmental and social benefits to an institutional intervention in the salinity problems of the LAB.

"ArkAgent" is an agent-based model developed primarily using collaborative modeling techniques, eliciting stakeholder knowledge and opinion on the social systems in the LAB. Agent-based modeling is a computational technique allowing more accessible and intuitive representation of human actors through the use of sophisticated behavioral algorithms and a one-agent-one-actor philosophy. ArkAgent simulates a range of actors from the real world who would likely have critical roles to play in a salinity solution: the farmers, whose irrigation activities drive salt loading; the ditch managers, who help maintain the infrastructure through which irrigation is possible; the municipalities, whose interest in basin water has been shifting its primary use away from irrigation for several decades; and the federal water manager, who controls the inflow of water to the system, and may have statutory authority to support efforts to combat salinity. I discuss these actors, why they are included, and more importantly why others are not included, in Chapter 2. I also simulate major pre-existing institutions in the basin, in particular the system of prior appropriations in the form of ditch and farmer water rights. Into this complex social mix I introduce a water quality trading market (WQTM), with the focus not on what it would take to establish such an institution, but what would happen if

it were.

The social component of ArkAgent is linked to the GeoDSS, a simulation tool developed by researchers at Colorado State University based on over a decade of work in the LAB (Triana & Labadie, 2007). The GeoDSS simulates the movement of water within the LAB (including delivery of water to ditch and lateral headgates), the return flows of water after irrigation, and any subsequent changes in salt content of the return flow and river water. Linking the GeoDSS to ArkAgent allows the realistic simulation of water availability for agent use, and the credible simulation of system responses to changed agent behavior. This is critical, since ArkAgent is intended to be used to test alternative behavioral assumptions and institutional strategies. ArkAgent runs over an 8 year historic period (1999-2007), at a weekly time step.

I apply ArkAgent in 13 separate main experiments (often referred to as "scenarios" in later chapters), each one exploring a unique set of actor assumptions and parameters for my chosen institutional intervention, the WQTM. ArkAgent returns detailed results, week by week over the simulation period, on actor decision making as well as broader social and environmental conditions. Social results returned include economic conditions (agent incomes, assets, debts), social trends (perception of conflict, equity of resource distribution) and environmental responses (salt load from the various irrigated regions of the basin, volumetric flow in the river). I use the diverse set of response factors ArkAgent provides to rank the various experimental scenarios in terms of overall sustainability, and also to dig deeper into the underlying agent and environmental processes. I apply these rankings and analyses to proving or refuting the 5 main hypotheses.

### 1.3.1 Chapter Structure

In Chapter 2 (Literature Review), I review the literature to support my chosen approach, as well as provide more detail on the nature of the LAB and its salinity problem. In Chapter 3 (Hypotheses), I outline my hypotheses, reviewing additional literature where necessary. In Chapter 4 (Methodology), I articulate in comprehensive detail the various techniques I am using to test my hypotheses. In Chapter 5 (Results), I outline the results of the simulation experiments and what they suggest for my five main hypotheses. In Chapter 6 (Discussion), I review the

simulation and empirical results in light of the hypotheses I posited in this chapter, suggesting the theoretical and practical implications of the results and any new questions or further work this raises. Finally, in Chapter 7 (Conclusions), I reflect on how well I affirmed or refuted the hypotheses, answered the questions, and contributed to the resolution of the central problem. The Appendix materials contain the model code; the model platform; detailed instructions for running and collecting data from the model; model results from all 13 experiments plus additional sensitivity testing; Stata do-files for generating analyses and plots; and any additional text or graphic material referenced in the main chapters.

# Chapter 2

# Literature Review

In this chapter, I conduct a targeted review of the literature associated with some of the major theoretical and practical themes raised in Chapter 1. I also lay the groundwork for the hypotheses, which are articulated in Chapter 3. I outline the broad problem of hydraulic sustainability, and the specific nature of the salinity challenge in the lower Arkansas Basin. Based on a review of existing practical and theoretical approaches to solving these problems, I argue that we are insufficiently inventive in our studies of institutional solutions for hydraulic sustainability, and part of the reason for this is that we are excessively bound by mainstream theories of human cognition and behavior. Based on a variety of empirical studies, I build an alternate conceptualization of water resource systems, which recognizes the diversity in resource user goals, attitudes and behaviors. I apply this alternate conceptualization in a simulation model, Ark-Agent, which explores the real world implications for hydraulic sustainability - the potential benefits as well as costs - of changing our theoretical assumptions on human behavior.

## 2.1 Hydraulic sustainability: definition, precedent, importance and analysis

I define a "hydraulic society" as a modern social-hydrologic complex, the result of large populations becoming dependent on extensive infrastructure for their water supply needs. This definition has roots in studies of ancient irrigation civilizations, but has renewed applicabil-

ity. Many of the large population centers of the western United States face challenges derived largely from their own over-exploitation of regional hydrologic systems. I define "hydraulic sustainability" as meeting the water supply needs of large social-hydrologic systems in a manner which does not damage underlying hydrologic systems, while minimizing supply inequity and associated social conflict. Simulation modeling of such systems is not yet advanced, despite the considerable potential for new insights. I develop ArkAgent as an agent-based simulation tool to address this.

### 2.1.1  Introducing the concept of the "hydraulic society"

The term "hydraulic society" was first coined by Steward in 1953, and added to the lexicon of anthropology by Wittfogel in 1957 (Wittfogel, 1957; Steward, 1953). Both authors applied the term in the context of studying differential survival rates among the great irrigation-based civilizations of the past, and it was briefly in vogue: Herman explored hydraulic societies in China (Herman, 1959), and Leach did the same for Sri Lanka, then Ceylon (Leach, 1959). An article in 1959 reviewing Wittfogel's work characterized "hydraulic society" as referring to a social system "based on large-scale irrigation with its demand for large-scale managerial organization" (Spate, 1959, 90). The term in the late 1950s is in every instance used to refer to ancient or "foreign" civilizations, and in an anthropological light. Wittfogel has been critiqued on several grounds, including the evidence that irrigation-dependent societies do not necessarily require the top-heavy bureaucracy Wittfogel posits (Hunt, 1988) and so presumably are not necessarily as vulnerable to wholesale collapse. Soon after his initial publication, Wittfogel was also heavily criticized for his interpretation of east Asian history, upon which much of his analysis was based (Meisner, 1963). Nevertheless, the idea was sown, and Worster revived the term for application to modern, western US populations highly dependent on water delivered through extensive infrastructure. Worster argues that California during the 20th century exemplified a modern hydraulic society, "a social order founded on the intensive management of water" (Worster, 1982, 504). Worster goes on to say: "I take this to be the essence of the hydraulic thesis: the domination of nature is an ambition that first appears stark and unchecked in the archaic desert empires" (Worster, 1982, 506).

20

### 2.1.2 Introducing the concept of "hydraulic sustainability"

The issue of sustainability was one that Worster explicitly raised: "most ancient hydraulic empires...had fallen into the infrastructure trap, building a bigger and bigger water system until they could no longer keep pace with the ecological backlash they were creating...The State of California, some recent evidence indicates, may be approaching a similar alkalinity fate" (Worster, 1982, 514). Worster later re-visited this idea in his 1985 work on western water (Worster, 1985) and has since been followed by Postel's thesis that the "irrigation miracle" can be pursued too far, along with extensive scientific study into the perils of salinization (Khan et al., 2006). The idea that technologically-dependent societies risk their own downfall by ignoring environmental systems is well explored elsewhere (Mumford, 1967). Out of this broad narrative comes the idea of "hydraulic sustainability": melding the broader concept of sustainability with the context of populous and structurally complex water systems. The World Commission on Environment and Development defined sustainability as pertaining to any human activity which "meets the needs of the present without compromising the ability of future generations to meet their own needs" (Brundtland, 1987, 43). The sources of such compromise can be environmental degradation (Ehrlich & Holdren, 1971; Swart et al., 2004; Tilman et al., 2002; Turner, 2008), but also conflict and other social dimensions (Sneddon et al., 2002; Turner, 2005; Wils et al., 1998). Historically, the study of hydraulic societies was examining what ancient irrigation civilizations did wrong. The contemporary study of hydraulic sustainability focuses on what modern water-dependent societies can do to avoid both old and new problems that will otherwise preclude their long term survival. A sampling of the literature identifies some of the problems facing western US hydraulic societies: salinity (Bouwer, 2002), shrinking snowpack (Barnett et al., 2005), collapsing biodiversityBelsky1999, emptying aquifers (Bolin et al., 2008), valley subsidence (Zekster et al., 2005), economic disaster (Barbier, 2004; Booker et al., 2005) and water conflict (Boronkay & Abbott, 1997).

### 2.1.3 Critiquing and refining the concept of hydraulic sustainability

The perspectives of Wittfogel and Worster are only one among many, and I do not adopt the concept of hydraulic sustainability wholesale and uncritically. Wittfogel's analysis is archaic in the context of modern urban and agricultural systems. California does not resemble ancient cases, and in more ways than just having technological tools for dealing with ecological collapse: the governing institutions are considerably different, and have (at least theoretically) considerably more adaptive capacity than ancient feudal hierarchies ever had. Thinking of modern western US communities as Wittfogel-type hydraulic societies is attractive in its simplicity, but cannot account for the true complexity in social and environmental dynamics we see in these systems. Sherow argues that Worster was reacting to the utilitarian philosophies advanced by the Harvard Water Program (HWP), and in so doing becomes excessively dogmatic (Sherow, 1990). The members of this research initiative argued that the single most important objective of water resources management was the maximization of economic benefit, and that objectives falling under this effort should be achieved through efficient and effective engineering and economic study (Maass et al., 1962). The work of Maass and the HWP was pioneering in establish benefit-cost analyses as an important tool in water project planning, and along with Maass' critique of massive inefficiencies and unprofessional conduct among the "Army Engineers" (Maass, 1951), laid the basis for the widespread use of benefit-cost analysis by the US Army Corps of Engineers (Hird, 1991; Tarlock, 2003). While there is plenty of evidence the principles advanced in Maass' work were occasionally used by both the Army Corps and the Bureau of Reclamation to justify projects with devastating social and environmental consequences (Reisner & Bates, 1990; Reisner, 1993), Worster appears to respond by implying grand conspiracy between the economic, engineering and governing disciplines exemplified by the Maass school. Sherow, Ingram and others have developed more sophisticated accounts of the development of western water. For Sherow, the historical narrative speaks to shortsighted, exploitative but largely parochial development of water to meet perceived economic and cultural needs. For Ingram, the special status of water in society meant communities small and large in the western US were driven to seek control of and exploit their water resources, often to the exclusion of competing needs, and often at great environmental cost. She outlines, by way

of example, the extremely complex process by which one of the last big federal Reclamation projects was created, which suggests if nothing else that conspiracy would have required such a mammoth feat of coordination among conflicting interests that we can discount its possibility (Ingram, 1967).

A further critique of Worster's definition of hydraulic sustainability is that neo-Malthusian warnings of impending disaster have a history of failing to come true. For example, "The Limits to Growth" report by Meadows and others (Meadows, 1972), which predicted global crisis as a result of population growth and environmental degradation, has mostly not been borne out by real events. Worster's arguments are excessively simplistic, and do not specify exactly how "ecological backlash" will succeed in bringing down modern as opposed to ancient civilizations. The Anasazi, for example, were rather more vulnerable to problems of salinization and scarcity in the absence of appropriate technologies to reverse technological decline, of which modern hydraulic societies can now avail themselves. Linner recently advanced a thesis of the neo-Malthusian "discourse of crisis", in which prominent researchers have warned of impending crises of various descriptions due to population growth reaching resource limits (Linner, 2004). There are plenty of dissenting voices to this discourse, of course. Some authors have argued that such narratives are overly negative, not supported by evidence, and damage our ability to properly prioritize issues of concern to society (Lomborg, 2001).

On this basis I could argue that "hydraulic sustainability" — in Worster's definition — is unfounded because existing systems are not particularly threatened, and in any case what matters most is that social benefits (typically economic) are maximized. I reject these premises, however. In the first place, it is factually indisputable that the major water development initiatives of the west have caused significant damage to many ecosystems across the region. The California Bay Delta; the Colorado River Delta; the Grand Canyon; and the Columbia and Snake basins have all yielded extensive and empirically documented evidence that large engineered water systems are ecologically unsustainable. There is also clear evidence that ecological damage can be directly linked to societal damage in various economic and non-economic forms (Lichatowich, 1999; Presser & Ohlendorf, 1987; Johnson et al., 2006), without even considering the evidence for indirect linkages through reductions in ecological resilience (Folke, 2003). In

the second place, as Ingram, Gleick and others have pointed out, there is far from global agreement that single or even multiple economic objectives are appropriate grounds for assessing the benefits of a given project (Ingram, 2006; Gleick, 2000).

Consequently, I critically refine my use of the terms "hydraulic society" and "hydraulic sustainability". I follow Ingram, Sherow, Reisner and others in imparting more complexity to the way I define hydraulic societies and discuss their sustainability. Western US water systems are not the "Orientally despotic" feudal hierarchies of 1950s anthropological lore. Nor are they the product of a sinister plot by economists and water engineers to concrete over the west, as exciting as that sounds. But they do possess remarkable dependence on very extensive engineering structures and such considerable sums are expended in maintaining these systems, that we would be unwise to discount the problems they might face in future. While Worster, Postel and others do a good job of articulating concern and pointing out risks hydraulic societies face, the future of the modern hydraulic society is not immediately threatened. My adoption of "hydraulic sustainability" acknowledges that these systems have a level of resilience born of the sophistication of the technology involved, and the complex institutional structures developed to manage them. I focus not on system-wide breakdown, since the chances of this happening (in the short term, at least) are small. The slim likelihood of Worster-style ecological collapse does not mean, however, that these systems will not face urgent problems that will incur considerable costs and consume substantial resources in their resolution if not dealt with soon. Salinity is one such problem. It simply is not worth ignoring, since along with climate change and invasive species, salinity can cause creeping problems that affect regional ecosystems, economies, land use and settlement patterns, and the usability of large quantities of freshwater.

### 2.1.4 Precedent in the treatment of hydraulic sustainability

The concept of hydraulic sustainability as I define it here has not received much attention in the literature. There has been a tendency to treat issues of sustainability and water separately. There are plenty of studies focusing on urban systems, rural systems and networks of infrastructure, but in a segregated fashion rather than gathering these systems under the hydraulic

society umbrella (Delleur, 2003; Vincent, 2003; Kandiah & Rao, 2004). Integrated Water Resources Management (IWRM; GWP (2000)) and social-ecological resilience (Carpenter et al., 2001; Gunderson & Holling, 2002) are two research areas which have got closest to the concept of hydraulic sustainability. However, IWRM misses Worster's point that hydraulic societies exist at the grace of a hydrologic system: IWRM literature persists in the hubris dominating the hydraulic and hydrologic engineering disciplines, emphasizing the importance of integrated environmental "management" (Born & Sonzogni, 1995). Social-ecological resilience focuses more on the biological dimensions of sustainability, and often fails to conceptualize water systems as complexes of social and environmental forces. Consequently, my use of the term "hydraulic sustainability" is novel. The lack of a hydraulic society framing for sustainability questions has led to a dearth of theory and few studies which explore system dynamics at the scales relevant to western US water systems.

The use of simulation models to study complex social-environmental systems is common (An et al., 2005; Barthel et al., 2005; Gimblett et al., 2001; Hansen, 2007), and their potential to ease analysis and the search for sustainability solutions is widely acknowledged (Costanza & Wainger, 1991; Pahl-Wostl, 2007a; Rauch et al., 2005). However, few such models have been developed for large-scale hydrologic systems, largely because of the technological challenges in developing and integrating simulation modules for the large variety of physical and societal processes in such systems (Parker et al., 2008). Still fewer models address such systems in the context of hydraulic sustainability, largely because the concept lacks an established body of theory. Virtually no models embody the sophistication needed to examine the consequences of more complex assumptions, as we explore these largely unknown systems (Janssen & Ostrom, 2006). In answer, I develop ArkAgent, an agent-based simulation tool which integrates social and hydrologic models. ArkAgent tackles the technological challenges of model integration, builds theory surrounding the analysis and resolution of hydraulic sustainability issues, and contributes policy analysis to a practical instance of hydraulic "un-sustainability" (Sugden, 2000; Maki, 2002).

## 2.2 Problem focus: the lower Arkansas Basin

The practical context for ArkAgent, the lower Arkansas Basin (LAB), suffers from a problem of irrigation-induced salinization which presents a fundamental challenge for hydraulic sustainability. This particular context is especially appropriate for a study of hydraulic societies and sustainability, not just because of the severity of the problem and the chance to contribute meaningfully to a solution.

First, a major focus of my study is exploring the systemic implications of changing some of our fundamental assumptions on the cognition and behavior of resource users within hydraulic societies. The LAB provides empirical material particularly appropriate to this effort. The basin embodies the confluence of a rich history (Sherow, 1990), complex sociology (Lepper, 2008) and an unusually extreme physical environment.

Second, the LAB provides a blank slate for the kind of institutional solution I focus on, the water quality trading market. It is a solution that has not been tried in this context, and its complexities provide an opportunity to test out alternative designs for the institution that might address failings which have emerged elsewhere. This prospective approach is distinguished from the wealth of post mortem examinations of the failure of water quality markets (Pharino, 2007).

Finally, as a candidate system for model development, the LAB greatly benefits from already being the subject of a lengthy study by local scientists (Gates et al., 2006). This has resulted in sophisticated regional ground and surface water quantity and quality models, available for use by and integration with ArkAgent. This provides an opportunity for more sophisticated use of social-environmental modeling than has hitherto been possible.

### 2.2.1 The lower Arkansas Basin: environmental and social context

The lower Arkansas Basin is a classic western hydraulic society: semi-arid; dependent on extensive infrastructure for all aspects of water management; and situated within a geographically extensive, crowded and growing metropolitan region. Product of a long history of development, water resources in the system are over-allocated. Existing water institutions are highly

structured, complex and resistant to change. Agriculture dominates water use in the basin, but municipal demands are growing.

The lower Arkansas Basin is defined as that portion of the Arkansas River Basin lying between Pueblo Dam and the Colorado-Kansas state line. The Arkansas River has its headwaters (and derives most of its flow) from the snowpack in the Rocky Mountains near Leadville, Colorado, flowing 2,364 km through Colorado, Kansas and Oklahoma before joining the Mississippi-Missouri river system in Arkansas. Once it leaves the mountains, the Arkansas River is a semi-arid (Goff et al., 1998) Great Plains river characterized by gentle meanders, shallow canyons and low average flow volumes (Schumm, 1963; Webb, 1981). The geomorphology and vegetation of the river have been heavily altered by over a century of irrigation diversions and other engineered structures (Friedman et al., 1998; Culver et al., 2007).

The LAB supports a population of several million in one major city (Pueblo) and several smaller cities strung out along the course of the Arkansas River (see Figures 2.1 and 2.2). It is also the primary conduit for water supplying an extensive system of irrigated agriculture, consisting of six major and 28 minor ditches. Historically up to 400,000 acres were irrigated by water from the mainstem Arkansas, but due to two decades of water transfers to urban areas, around 311,000 acres remain (Woodka, 2007). Dominant crops in the valley include wheat, sorghum, alfalfa and corn. Other economic forces within the LAB include federal and state prisons, healthcare facilities, wind farms, manufacturing and the BNSF railroad.

LAB water also contributes to the water supplies of Colorado Springs, located within the basin, and Aurora, located outside of the LAB, with combined populations of over 680,000. The LAB consequently takes a role in one of the largest hydraulic societies in the United States: the Front Range Urban Corridor (Colorado-Wyoming) includes over 4 million people, 24 major cities, and dependency on three major river systems (the Colorado, the Platte and the Arkansas).

Water in the Arkansas River is fully appropriated through water rights that date back to the mid 19th century; it is also over-allocated, meaning that even in the best hydrologic years, water rights with lower priority do not receive all the water they need. The Colorado water rights system (the "Colorado doctrine"; Hobbs (2007)) operates on the basis of prior appropriation
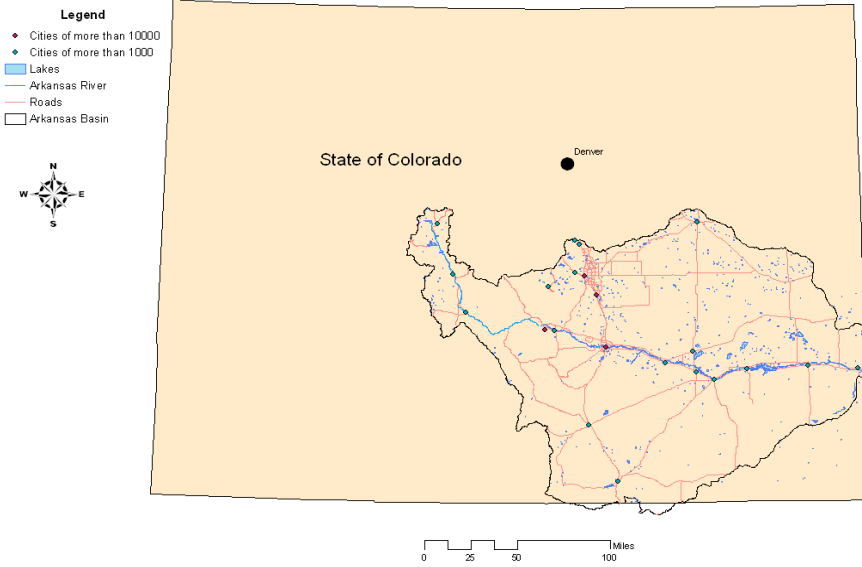
Figure 2.1: Location of the Arkansas Basin within the state of Colorado
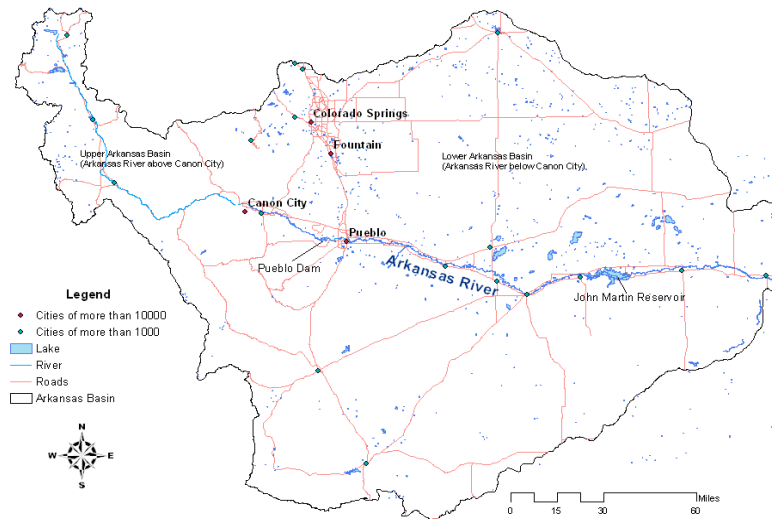


Figure 2.2: Detail of the Arkansas Basin, with upper and lower portions labeled. The focus of the dissertation is the lower portion of the basin.

28

law (Lueck, 1995; Grigg, 1996; Hobbs, 1997): the first appropriator to arrive at the river and make a diversion for beneficial use receives a water right with a priority date and flow rate at that diversion point. All subsequent appropriators also receive priority dates and flow rates, with descending seniority. The Colorado doctrine has largely resisted any substantial change for over a century (Kenney, 2005; Nichols et al., 2001), although there have been modifications to allow municipalities to condemn water rights that usurp domestic needs; to allow instream flows to count as beneficial use; to enable the state to acquire rights to meet minimum flow requirements; to allow cities, counties and water districts to appropriate water for recreational needs; and to address the different management needs of groundwater (Hobbs, 2003). The system remains in tension, however, as exemplified by the contrasting perspectives of State Supreme Court Justice Gregory Hobbs Jr (advocating the preservation of traditional priority doctrine and against the use of the public trust to justify changes in use; Hobbs (1985, 1994, 2002)) and Charles Wilkinson (advocating for a reconceptualization of the priority doctrine into a more comprehensive, watershed based approach; Wilkinson (1989, 1991)).

Despite these rumblings of change, irrigated agriculture remains the primary user of Arkansas River water, with local municipalities also holding some very senior rights to some of the flow. Irrigated agriculture is supplied by diversions into canal systems branching off from the main-stem river. The majority of these ditches formed between 1860 and 1910 (Lepper, 2008), and some of them undertook to build storage reservoirs and transmountain diversion tunnels in the upper and lower basin (Milenski, 1990), adding up to a total capacity of over 600,000 acre-feet. The size and influence of the lower Arkansas Basin ditch companies has declined in recent years, as a considerable number of farmers have sold up land or water rights or both, leading to the closure of several canals. Municipalities either hold rights to LAB flows (such as for the city of Pueblo), or rights to flows from the upper basin (such as for the city of Colorado Springs), or both.

An important development in the social history of water in the LAB was the creation of the Fryingpan-Arkansas Project in 1953 (USBR, 1973). This is an extensive and federally-funded transmountain diversion project, consisting of five storage dams/reservoirs and sixteen diversion structures, of which nine are tunnels under the continental divide. Pueblo Reservoir is

the major project storage within the lower basin, and delivers water to Colorado Springs via the Fountain Valley conduit and to Pueblo via river flow. Water diverted from the western slope and added to Arkansas River flows provides supplemental irrigation for over 280,600 acres in the LAB, although significantly less than those total acres are actually irrigated (USBR, 1973). More recently, the Southern Delivery System pipeline project has been proposed to allow Colorado Springs Utilities to pump water out of Pueblo Reservoir, taking advantage of water rights the municipality already owns in the river.

While the LAB has all the properties most commonly associated with a hydraulic society, it is much smaller in all dimensions than the classic example of a modern hydraulic society, California. Yet the basin suffers from similar problems of hydraulic sustainability, and disproportionately so. The LAB has been suffering from salinization issues for longer and more seriously than much of the Californian agricultural system (Schoups et al., 2005; Gates et al., 2006).

### 2.2.2 The salinity problem in the Arkansas River

The Arkansas River in the LAB is one of the most saline rivers in the United States (Miles, 1977). The high levels of salts in basin soils and water contribute to serious economic and environmental damage, and are the result of a long struggle to cultivate marginal areas through heavy use of water infrastructure. Existing studies of the problem have failed to develop solutions that are workable in the complex social and environmental conditions within the basin. Many other hydraulic societies around the world are just beginning to grapple with similar problems, and so studying the LAB may derive lessons applicable both locally and globally.

The severe problem of salinity in the Arkansas River within the lower Arkansas Basin is derived from the concentration and re-concentration of salts in return flows from canal-based irrigation. Salts in these waters are moved close to the surface by upflux to a high groundwater surface, then concentrated by evapotranspiration. A percentage of water extracted from canals then returns to the river via surface flow and throughflow in the subsurface. Naturally occurring salts in local shale beds further add to the salt concentration, and the cycle begins again further downstream when water is drawn out of the river once more. Average salinity in the

river ranges from 500 mg/L near Canon City to over 3,500 mg/L near the Colorado-Kansas state line (Goff et al., 1998). Miles (Miles (1977)) reported that over 81,000 hectares within the lower Valley were classified as C4, the US Salinity Laboratory's highest classification for inland waters. More recent work (Gates et al., 2002) has shown that the average salinity concentration of ground water is above 3,100 mg/L. All six major irrigation canals in the valley have low irrigation efficiencies and high rates of salt loading (ibid). Average soil salt concentrations up to 2005 ranged between 4,800 to 5,600 mg/L throughout the valley, a significant contribution to reduced crop yield and other environmental damage (ibid). Selenium (Se) concentrations are of particular concern: in the recent work by Gates et al. (2006), all but two of the samples taken of Arkansas river water from across the LAB exceeded state and federal standards for Se.

Salinity is a problem for a number of reasons. The high levels of salts in the soils and water of the LAB mean reductions in crop yields (Childs & Hanks, 1975), damaging the primary economic base of the LAB community. Also important is the variety of seriously damaging effects high levels of aquatic salinity can have on riparian and aquatic ecosystems (Rysgaard et al., 1999; Nielsen et al., 2003; Vandersande et al., 2001; Magdych, 1984; Nielsen et al., 2003). Finally, under Article IV of the Arkansas River Compact, Kansas has the legal sanction to disrupt agriculture in the Colorado portion of the Arkansas Valley by demanding that water delivered across the state line be "usable". While it is arguable that such sanction might actually motivate change in the system, this is not generally how it is viewed by the agricultural community (pers. comm. Valliant 2007).

Salinity is a classic problem facing hydraulic societies globally, particularly those incorporating extensive irrigated agriculture (Schoups et al., 2005). It is a unique and compelling challenge to sustainability because it threatens the very substrate of a productive food system, the soil. Salinity has a certain resonance as an environmental issue, because there is strong evidence that past irrigation-dependent civilizations have collapsed at least in part because of salinity (Hillel, 1992; Binford et al., 1997). Modern hydraulic societies are predisposed to salinity problems because of their tendency to over-abstract ground water and over-apply surface water, but salinity is also a major problem because it so difficult to avoid in the arid climates that are the setting of most hydraulic societies. Salinization is also tremendously difficult to fix

once it has begun, without causing considerable disruption to agricultural economies and communities. Salinity works itself into all elements of a hydraulic society by affecting the viability of rural communities and also threatening urban water supplies: it is a true existential threat.

Existing studies of the salinization problem in the LAB have focused exclusively on characterizing the nature of the problem and exploring technological solutions at the field level (Miles, 1977; Burkhalter & Gates, 2005; Gates et al., 2002, 2006; Goff et al., 1998). Efforts to implement these solutions have run up against what appear to be a mix of economic, institutional and behavioral obstacles (pers. comm. Gates 2009). Such efforts have lacked grounding in a deeper understanding of how technological and other "pipe-end" solutions can be embedded within an institutional context, and there is considerable scope for work that would demonstrate the possibilities of using more sophisticated interventions to reduce salinity.

### 2.2.3   Principal Actors, Responsibilities and Behaviors

To provide material for populating the artificial agent populations in ArkAgent, I identify the principal actors in the basin, and review evidence for their responsibilities and behaviors. The principal actor in the LAB at a national level is the US Bureau of Reclamation. The Bureau is primarily responsible for operating the Federal supply project in the basin, and is typically conservative in interpreting its mission. At a regional level, the major municipal utilities are important actors: Colorado Springs, Aurora and Pueblo. These utilities are primarily responsible for meeting their own water supply needs, but display complex behavior that belies easy categorization. Finally, at a local level, ditch companies and farmers represent the principal users of water in the LAB. I focus on these actors most closely in my simulation model.

At a national level, the **US Bureau of Reclamation** (USBR; Department of the Interior bureau) operates the Fryingpan-Arkansas Project, which delivers extra water from the Leadville region of the Rocky Mountains into the Arkansas River. USBR is responsible for water delivery as part of this project, and for flood control through manipulation of water levels in the Pueblo Dam, but is statutorily limited in getting further involved in water management decisions downstream of the dam. The USBR has long been known as a conservative force in water across the western United States, with an overwhelming focus on supply management.

The legendarily brash technocracy (Reisner, 1993) and supply-side "ideology of Reclamation" (Pisani, 2002, 13) has long been diluted by a collapse in funding, an exodus of staff, no major new storage or diversion projects in more than two decades, and the devolution of many of their existing facilities to private hands (Simon, 2002). Within the Arkansas Basin, the USBR has been most recently involved in providing space in its Federal reservoirs for new water destined to supply Colorado Springs and Aurora. All recent actions by the USBR in the region demonstrate its primary strategy of meeting existing needs through supply-side solutions, and relatively little interest in expanding the scope of its operations. I simulate the Federal water manager in solely a supply-side role,.

In Colorado, the state water quantity/quality apparatus is important. While the state does not have the same powers that other western US states have in terms of reassigning allocations (Wescoat, 1986), the state engineer's office is broadly empowered to shut down wells or headgates where the appropriation is deemed to be out of order or in exceedance of historical rights. The state relies on an elaborate legal system of water courts for resolving most disputes between rights holders. On the quality side, the Colorado Water Control Commission (CWCC), in partnership with the state Board of Health, has the power to develop specific water quality standards in line with the Colorado Water Quality Control Act (CDPHE, 2002). This Act argues that the beneficial use of water (a central tenet of the prior appropriation doctrine, which governs water allocation in Colorado) is maximized when water quality is improved, but does not specify how this squares with the more traditional legal definition of "beneficial use" in Colorado, which does not explicitly include water quality. The Act provides for basic standards in an array of salts affecting the LAB, including selenium and magnesium, but these standards are based on public health concerns and not riparian health. Furthermore, actual control efforts particularly in agricultural systems are subject to economic reasonability tests, and are frequently subordinate (in the absence of a pressing public health concern) to concerns over preserving extant water rights. Water quality control in Colorado for non-point source agricultural systems is a tangle of conflicted interests, and the CWCC has not played a particularly decisive role so far. See the specific discussion of Colorado water quantity and quality law for further treatment of these and other issues (section 2.2.4.1, below). In ArkAgent, I do not

simulate the CWCC or the Water Control Division more generally, because over the historical period of simulation these agencies played very little role; and because of the technical challenge arising from the necessity to implement the water courts system along with these agents. Whether or not they will play a larger role in future remains to be seen, although it must be assumed that they would be involved in implementing a water quality trading market were it to come into existence. But *for the purposes of the simulation and the questions/hypotheses I pose*, it is not necessary to simulate these agencies.

The regional municipalities play a critical role in the current and future state of water management in the LAB. **Colorado Springs Utilities** (CSU) has been a major customer of farmers selling their water rights in the LAB, and is currently working on the Southern Delivery System, a proposed pipeline from Pueblo Dam up to Colorado Springs. CSU holds many senior water rights, as well as a storage and delivery infrastructure of its own. The **Board of Water Works of Pueblo** has less volumetric punch, but retains some senior water rights and has a priority call on Pueblo Dam for domestic supply. **Aurora Water**, serving the city of Aurora, has increasingly been involved in water supply discussions in the LAB; it is a fast growing city, with only limited domestic supply support from the Denver Water Board, and so has been involved in numerous water purchase deals in the past. Inferring from their past behavior, all the regional municipal utilities appear to be exploring options for additional water from the LAB agricultural community, although in response to public controversy the utilities have become a great deal less aggressive. In reality, however, the true objectives and strategies of the utilities are substantively unknown. Naively, the actions utilities take are fundamentally constrained by the legal responsibilities they bear to their respective citizen bodies: they must meet actual present and perceived future needs. Realistically, however, this might limit but does not determine utility behavior. For one, "future needs" is a flexible concept, and utilities have in the past used growth projections to justify additional water supply (Anderson & Hill, 1975; Worster, 1985; Ellison, 1995). Second, urban water politics are complex, with effects often hidden by the technocracy of modern utilities (Oshio, 1997). Third, motivations for bureaucratic behavior are not always predictable (Moe, 1997), particularly so for managers within public utilities (Perry, 1996). Fourth, each utility has a different income base, a different portfolio

of water supply, different demand profiles and different operational strategies (Kenney et al., 2004). I develop a model of utility behavior which is fundamentally constrained to meeting urban water needs. However, in keeping with the uncertainty and apparent variation in utility strategies, I assign variable degrees of aggressiveness with which utility managers pursue new supply, and variable philosophies towards demand management.

Key stakeholders at a more local level, but also regionally when they collaborate, are the **ditch companies**. These are cooperative entities designed to spread the considerable cost of constructing and maintaining diversion ditches, and some of the larger ditches are over a century old. The largest of the ditches have very senior rights on the Arkansas River, and supply a large proportion of the 311,000 acres of total irrigated agriculture in the LAB. The ditches are responsible for supplying their shareholders (farmers up and down the lower Arkansas valley) and maintaining their infrastructure, but they also take an active interest in water use and management issues in the valley. It is important to note that ditches are non-profit cooperatives with strong identities. They have a long history of both fiercely defending and shrewdly using their water resources (Milenski, 1990). The motivations of these institutions may extend well beyond the economic calculus: historically such ditches were viewed (and viewed themselves) in the light of a hydrologic manifest destiny (Smythe, 1969), a means to "reclaim" the prairies of the Front Range for productive use (Reisner, 1993; Sherow, 1990). Sherow highlights how different ditch companies in the LAB had very different complexions in their early days: "contentment with their circumstances" for the Rocky Ford Ditch versus a "litigious tendency" for the Fort Lyon Canal and a "willingness to extend their control over nature" (Sherow, 1990, 27). Mutual irrigation companies differed then, and continue to differ in their canal length, the size and priority of their water right, the physical conditions of their ditches, their links to other industries, the various costs they have to meet (Wilkins-Wells et al., 1999) and the nature of the personalities running the canals. Given this diversity, it would be absurd to characterize ditches as possessing common motivations and behavioral strategies: all are clearly interested in surviving and fulfilling their water supply duty to shareholders, but the diverse courses the various companies have charted since the early 1900s indicates great complexity in their decision making (Sherow, 1990). For the present study, ditches are treated as sharing a common

goal — fulfilling water supply obligations to shareholders — but are equipped with variable attitudes towards their own and the basin's water management. Variation in these attitudes is connected as far as possible to their physical properties, since there exists no modern accounts of the bureaucratic personalities of each ditch.

The **farmers** of the LAB are the core focus of salinity concerns. They are both the primary source and the primary victims of high levels of salinity. Recent work suggests that farmers are losing an average of \$232/hectare/year due to waterlogging and salinity, 10-15% of annual crop yield (Gates et al., 2006). Farmers certainly have the most power to alter water quality trajectories in the LAB. It is their daily and annual business and agricultural decisions that largely determine salt loadings in the river. While the potential impact they could have is clear, understanding what would motivate them to participate in such efforts is less so. No specific behavioral studies of irrigators in the LAB have been conducted, although there have been a variety of studies examining economic dimensions of farming in the basin (Gorelick & Lefkoff, 1990; Howe et al., 1990; Wilkins-Wells & Epley, 2003). Nevertheless, various authors have their theories on the irrigators of the LAB: Sherow argues for a contradictory mix of cooperative tendencies and competitive market culture, in which irrigators have always struggled to find the right attitude towards their harsh environment, and the right set of behaviors for successful farming (Sherow, 1990). Lepper argues that cooperative behavior dominates (Lepper, 2008), while others suggest irrigators are driven by economic self-interest (Gorelick & Lefkoff, 1990; Gates et al., 2006). These are positions advanced largely without empirical support. Fortunately, there is a very extensive empirical literature on farmer decision making, particularly in relation to pro-environmental behavior (see section 2.3.12). In developing the detailed models of farmer agents, I rely heavily on a review of this material, as well as the empirical data gathered in interviewing, surveying and collaborative modeling workshops.

### 2.2.4 Salinity Reduction: Potential Solutions and Stakeholder Roles

In the absence of major desalination infrastructure, salinity reduction in the LAB must be operationalized through changes to behavior and technology at the on-farm and on-ditch level. For farmers, available options include upgrades to irrigation equipment, changes to farming

practice, and installation of drainage tiles. Ditch managers can choose from various means to seal or otherwise reduce seepage through the bottom of their ditches. Municipalities also have a role in reducing salinity, through their water purchase efforts and their considerable financial muscle.

Salinity reduction in the LAB can only happen through changes to practice and technology on the farm and on the ditch. I focus on the institutional framework within which such changes would be embedded, but I also need to simulate the on-the-ground modifications: a sophisticated simulation of the institutional framework would lack credibility if the physical changes to practice and technology were not also scientifically valid. The physical mechanisms for salt buildup involve a flux of salt to the surface, and high ground water levels facilitate this. The literature indicates that over-irrigation contributes the most to high ground water levels, and so excess application of water is the primary mechanism for salt concentration in the basin (Goff et al., 1998; Gates et al., 2006). Over-irrigation largely results from the persistence of highly inefficient furrow (flood) irrigation (Pereira et al., 2002), and the elimination of these practices is a primary goal of salinity reduction efforts. An additional contributor to high groundwater levels is seepage through thousands of miles of unlined canals and ditches. A secondary option, then, is to reduce this seepage. Some solutions can be achieved through behavioral, some through technological change, although in practice it is artificial to segregate the two. Many technological changes require significant adjustments to farmer and ditch manager attitudes and behavior. Figures 2.3 and 2.4 below describe some of the more feasible options for farmers and ditch managers in the LAB.

ArkAgent does not simulate all options, on the basis of what regional experts expect to be feasible and acceptable in the agricultural community. Installation of drainage tiles, for example, is relatively expensive and less well known in the community (pers. comm. Gates 2009). Targeted pumping can be logistically demanding, while removal of vegetation has very limited effects. Through their on-farm decision making, farmers have a central role in any solutions to reduce salinity. But they can also have effects as shareholders on mutual irrigation ditches. Ditch managers must be responsive to shareholder wishes, but also have some autonomy to push through canal lining/sealing efforts. Assuming maximum uptake of irrigation efficiency

| Option | Description | Change in technology | Change in behavior |
|---|---|---|---|
| Adopt surge irrigation | Modified furrow irrigation: pulses of water sent down furrow | Minimal: gated pipe and electronic controls | Moderate: changing long-practiced irrigation technique |
| Change field dimensions | Shorter fields are more uniformly infiltrated | None | Minimal: change in cropping patterns |
| Re-grade fields | Steeper, uniformly graded fields are more uniformly infiltrated | None | None |
| Remove phreatophyte vegetation | Cottonwoods and Tamarisk enhance upflux of salts | None | None |
| Targeted pumping | Groundwater pumps lower water table | Investment in pumps | None |
| Install drainage tiles | Lowers water table by draining water away from fields | Investment in under-field drains | None |
| Adopt drip irrigation | Reduces water waste and reduces water needs | Investment in drip technology | Shift towards drip-irrigable crops |

Figure 2.3: Description of the more feasible options for reducing over-irrigation by farmers in the LAB

| Option | Description | Change in technology | Change in behavior |
|--------|-------------|---------------------|-------------------|
| Seal canals | Use polyacrylamide or polysaccharide to cause flocculation of sediment | Minimal: purchase of cheap spreading equipment | Moderate: requires semi-annual re-application |
| Line canals with concrete | Add a concrete bed to canals | Large: very expensive and disruptive | Minimal: once installed, lasts for several decades |
| Line canals with geotextile | Add an impermeable membrane to canals | Moderate: reasonably expensive, but easy to apply | Moderate: requires maintenance |

Figure 2.4: Description of the more feasible options for reducing canal seepage by ditches in the LAB

technologies and practices, and extensive sealing of ditches, up to a 90% reduction in ground water contributions could be expected (pers. comm. Temeepattanapongsa 2009).

Actors of secondary importance are the municipalities, who can effect dramatic reductions in over-irrigation (and salinity) simply through buying up water rights. The municipalities have not, in the past, been willing to maintain farms after assuming control of their water rights, and so some 60,000 acres of farm land went out of production between 1970 and 2000 (Howe, 1998). Increasingly, however, new transfers are occurring as leases, with some land remaining in cultivation. Expected irrigation reductions from transfers will be less in these cases. Additionally, municipalities have the financial resources and bureaucratic flexibility to support and participate in valley water management directly. A municipality may have an interest in ensuring a ditch remains in operation, to safeguard any rights it has bought under that ditch, for example. ArkAgent includes municipalities, mainly acting as sources of investment for the basin in the form of water transfer agreements and potentially in supporting irrigation equipment upgrades. Municipalities do also contribute salt load through their water treatment plant discharges: Pueblo, Fountain Creek and Colorado Springs are principal contributors in this manner. Technical limitations in the physical simulation model mean that changes to this contribution cannot be simulated, however.

Water is allocated and consumed in the Arkansas Basin within the prior appropriation framework, constituted by a century of accumulated Colorado and western water law. Quality dimensions of water management fall mostly within state-level regulations, as outlined earlier. Pertinent regulatory or otherwise empowered actors therefore include the state of Colorado (in the form of the state engineer, the Governor's Office, the Legislature and other governmental and quasi non-governmental agencies), the federal Environmental Protection Agency (EPA), the State of Kansas and the U.S. Supreme Court. Of all these actors, the Supreme Court has the greatest capacity to inflict change on the Arkansas Basin, but the likelihood of a dramatic shift in the existing Colorado-Kansas compact is slim. The EPA lacks any significant regulatory power to improve sustainability in the system. The many state-level agencies, with support from the legislature, could theoretically contribute a great deal to the sustainability of the system. For example, the Colorado Water Control Commission could raise the salt standards and make them less susceptible to compromise by the water courts and more in line with evidence on riparian damage. However, given that state agencies have long shown a conspicuous lack of enthusiasm for regulating salinity in the non-point source agricultural context, state-level regulatory action is an unlikely source of sustainability change in the Arkansas basin. State-level actors other than the state engineer (with a limited policing role) are, consequently, not represented in ArkAgent. Note, however, that despite the lack of these state actors, the allocation of water in ArkAgent is still governed by the existing institutional frame of Colorado water law, and agent actions in relation to water use are circumscribed by what is customarily possible under Colorado water law. See the sub-section 2.2.4.1 below for a more detailed discussion of the interplay between water quantity and quality in Colorado.

In many other U.S. settings, non-governmental environmental and agricultural organizations often have significant roles to play in fomenting sustainability change. The Arkansas Basin, devoid of any particularly unique ecosystems or endangered species, has not attracted much attention from national or local environmental organizations. The Nature Conservancy has had a recent role in promoting conservation easements as a way to protect and enhance riparian lands in the Arkansas Basin (Woodka, 2009), but is not a major player in salinity reduction efforts generally. Similarly, non-governmental agricultural interests in this family-farm

dominated system are not major institutional players. There is no doubt that, in the event of a renewed focus on salinity in the basin, more state and national non-governmental attention might be drawn to the basin. But at present, it is difficult to see where this would come from, and how to simulate it. For these reasons, non-government organizations of all kinds are excluded from ArkAgent.

### 2.2.4.1 Water quality and quantity laws in Colorado, and impact on potential salinity solutions

A brief review of the prior appropriation framework is in order, since this provides the backbone to water use and management in Colorado. The concept of "first in time, first in right" allows for the ordered allocation of water among users when the total availability is exceeded by demand: the first appropriator on the watercourse is granted a decree to use the water, and all appropriators after this time receive priority dates and appropriation amounts in deference to the more senior appropriators. Central to the smooth operation of the system is the "call", which describes a request by a senior appropriator — made to the District Commissioner or Division Engineer — that another appropriator or appropriators modify their appropriations until the decreed rights of the senior appropriator are met. Junior appropriators may also make "calls" to ensure that senior appropriators only divert their decreed amount. While all water in Colorado is considered to be owned by the state, the prior appropriation system allows for "beneficial use" of the water resources for private gain. The definition of beneficial use has been shaped by what originally bore social and economic value in the formative years of the state, namely, irrigated agriculture (Hobbs, 2002). More recently, instream flows for ecosystem health and recreational use have become institutionalized, but rarely to the extent that it has subverted clearly defined existing appropriative rights. An elaborate system of water courts with over a century of experience presides over disputes between rights holders and indeed any change to a water right whatsoever (Nichols et al., 2001).

The water rights system also provides for the trading of appropriation decrees, as long as the trade does not injure other water rights holders and is limited to the historically-defined beneficial use (ibid). This is one way in which the water rights system might support reduced

salt loads, and municipalities in the Front Range have been seeking additional water sources in this way for decades (Tarlock & de Wetering, 1999). Front Range municipalities have not (until relatively recently: USBR (2010)) had the means to physically move Arkansas River water out of its basin, and many leases and purchases are hedges against future needs. Consequently, the effect of leasing water from farmers is often to cause fallowing and reduced diversions. In principle, because much of the salt load arises from return flow dissolution, reduced diversions may result in reduced salt loads. Also, any unappropriated water will remain in the river, diluting existing salt loads and reducing point concentrations. However, since the Arkansas Basin is over-appropriated, there is the possibility that other appropriators will soak up this extra water, mitigating any potential benefits (pers. comm. Wescoat 2010).

The legal system for water quality in Colorado is separate from and supplemental to the water rights framework, and evolved through the Clean Water Act (CWA) of 1972. A later court case (A-B Cattle Co. v. United States) underlined the superiority of beneficial use provisions over water quality concerns (Brown, 1979). Given its genesis in and dependence on the CWA, Colorado's water quality laws have been focused on point source discharges that the CWA handles relatively well. More complex non-point sources have not been well integrated into Colorado law, particularly since Total Maximum Daily Load (TMDL) levels — one of the principal instruments for combatting point and non-point sources — have been repeatedly delayed in implementation (GPO, 2001). The Lower Arkansas River has been added to the 303(d) list for impaired streams by the EPA, an action which is supposed to precipitate TMDL development, but this has not yet happened. Through a combination of state inertia, federal delays and community resistance, the Arkansas salinity problem has remained largely unaddressed.

The late arrival of water quality laws in Colorado mean that water quality issues tend to struggle in relation to water quantity. Both federal (U.S.C. §1251(g) 2000) and state regulations place water rights above water quality in relative importance (Nichols et al., 2001), and greatly limit what the state can do to require water rights holders to comply with controls (Colo. Rev. Stat. §25-8-104). Improvements to water quality can force changes in diversion volumes and patterns, which limits what quality controls can be instituted (Wescoat, 1986). Of particular importance to the Arkansas Basin is what happens to water which is gained through increases

in irrigation efficiency, and whether or not irrigation efficiency in fact drives increased use of water and injury to other appropriators. These issues have so far served to complicate and slow the search for solutions.

In summary, the legal framework for water quality in Colorado is fairly weak in relation to the powerful political and legal forces supporting prior appropriations. While the EPA continues to push for better integration, the status quo prevails for the Lower Arkansas Basin. I have noted earlier that ArkAgent simulates the water rights system as accurately as historical records allow, but ArkAgent does not explicitly simulate the water quality legal framework nor the state and federal agencies associated with it. It does so implicitly, however, in the simulation of a water quality trading market, which would by necessity include TMDLs. In the case of the simulated WQTM, the TMDLs are annual rather than daily, and the enforcing authority is the state of Colorado. Given the tangled state of water quality laws in Colorado, this is something of a suspension of disbelief. But I argue that the stagnant state of water quality/water quantity relations in Colorado is even more reason to demonstrate how improved water quality regulations can work in support of the prior appropriation system, rather than against it. Changed perceptions among appropriators will be necessary before state legislators or even the EPA feel confident in mandating more significant water quality improvements.

### 2.2.4.2 Previous salinity control efforts in Colorado

Salinity control in Colorado has hitherto been limited to direct regulation of point sources, including the establishment of trading programs. There have been no major state-driven attempts to regulate or otherwise address agricultural salt loading, while the federal government has invested considerable resources in improving irrigation efficiency in the Grand Valley of the upper Colorado River basin.

Major attempts to control salinity in Colorado are divided into two broad groups: isolated market-based point-source control efforts, like the Lake Dillon Trading Program; and broader non-point source, non-trading efforts, of which there is in effect only one of significance, the Colorado River Basin Salinity Control Program. With regard to the first group, a recent survey of WQTMs in the United States identifies six active and one nascent effort to establish point

source/non point source trading programs. Five of the six are located in the Front Range, and all primarily address urban and industrial effluents. According to Pharino's account, trades within these programs have been sparse, and some of the greatest water quality benefits have come from wastewater treatment plant upgrades prompted by increased regulatory attention (Pharino, 2007).

The Colorado River basin suffers from serious agriculturally-derived salinity issues (Glennon & Culp, 2002). Agriculture in Colorado, Wyoming and Utah is thought to contribute at least 30% of the excess salt in the system (Gardner & Young, 1988), and so has received particular attention in Colorado. Much of the on-farm efforts to improve efficiency and reduce salt loading have come through federal efforts originating in the 1974 Colorado River Basin Salinity Control Act (Wescoat, 1986), and in particular through the efforts of the NRCS and the Bureau of Reclamation (Gardner & Young, 1985, 1988). The state role in relation to agriculture has been limited to supporting existing NRCS and Bureau of Reclamation efforts, and taking a seat on the Salinity Control Forum for the basin.

The implication for the present study is that there remains considerable room — in both theoretical and practical terms — for the study of salinity control in Colorado. Existing salinity control institutions are either exercises in federal statutory authority or simple trading arrangements between cities and regulators; the potential for institutional innovation, particularly involving farming communities, is considerable. This is particularly true in eastern Colorado, where the state and federal governments have shown very little interest in tackling the salinity problem. In the next section, I discuss what general role institutions have in the search for solutions to sustainability challenges like salinity, laying the foundations for the present study.

## 2.3 The role of institutions in the search for sustainability solutions

Following Ostrom (Ostrom, 2005), I define institutions as any sort of prescription, implicit or explicit, used to organize repeated and structured interactions between individuals and groups. In water resources management, institutions have long been of secondary interest relative to infrastructural solutions. However, institutions have potential advantages over infrastructure,

as solutions for issues of hydraulic sustainability. These include scope, flexibility, reversibility and reduced cost.

Institutions, webs of legal, cultural and behavioral rules, norms and traditions, provide the context for the technological and behavioral solutions. Needless to say, if it were simply a matter of getting technology to the farmers, the salinity problem in the LAB would likely already be resolved. However, just the single task of rolling out improved irrigation technologies, decades after the invention of the center-pivot sprinkler (Zybach, 1960), has barely begun: less than 5% of all farms in the basin have sprinkler irrigation, persisting in using a system of flood irrigation which dates back to the 1860s (pers. comm. Gates 2009). The real challenge is encountered in modifying behavior, in working against century-old water law and tradition. Institutions can be significant obstacles to change, but institutions also provide some of the most flexible, efficient and reversible options for sustainable change. Unfortunately, in water management science and practice, institutions have long taken a back seat relative to infrastructure.

"Infrastructure" is used to mean physical structures designed to store, move, process or otherwise handle water. Water management in the western US has long been dominated by the concept of "water development". Such development consisted of the development of dams and pipelines to facilitate water supply, flood control, irrigation supply and hydroelectric power generation (Barrow, 1998). Supply-side infrastructure in the western US has undoubtedly provided great public and private goods, yet at great cost: massive damage to riparian and aquatic ecology, physical change to entire landscapes, as well as considerable social and economic disruption (Belsky et al., 1999; Libecap, 2008; Reisner & Bates, 1990). Gleick and others argue that a combination of rising costs and changing public values have led to significant questioning of traditional approaches to water development: Gleick terms this the "changing water paradigm", while Torrecilla and Gil describe a "New Culture of Water" (Gleick, 2000; Torrecilla & Martinez-Gill, 2005). These authors describe a rising interest in using non-development approaches to meeting water challenges, most of which can be loosely described as new or changed institutions.

"Institutions" are a major topic of study in several disciplines. Herbert Simon was one of the first to examine rules of behavior within and between organizations (Simon, 1945), and stud-

ies of institutions have since flourished in economics, political science, development studies, public administration and resource management (Eggertsson, 1990; Moe, 1984; Ostrom et al., 1988; Nabli & Nugent, 1989). In view of the great diversity of meanings for the word (see, for example: Plott (1967); Riker (1982); Schotter (1981), I adopt a definition offered by Ostrom in 2005: institutions are "the prescriptions that humans use to organize all forms of repetitive and structured interactions including those within families, neighborhoods, markets, firms, sports leagues, churches, private associations, and governments at all scales." (Ostrom, 2005, 3).

The institutions Gleick and others describe involve water resource management (WRM). Broadly defined, a WRM institution is any collection of rules that is used to manage water resources (Fox, 1976; Adger, 2003; Blomquist, 2006; Blomquist et al., 2004b; Boissau & Castella, 2003; Bretsen & Hill, 2006). The "management" of water resources is, simply, any manipulation of the hydrosphere by and for society. Water institutions mediate the interactions between humans and the hydrologic environment, and compared to infrastructure may have considerable benefits in supporting hydraulic sustainability. Institutions may have far smaller social, economic and environmental cost (Gleick, 2000). Infrastructure may lack resilience in the face of rapid or large-scale social and environmental change (Anderies et al., 2006), while institutions retain considerable flexibility (Gunderson, 1999). In any case, institutions already have powerful effects on water management. Examples include the prior appropriation doctrine, the interstate water compacts and the Endangered Species Act (Tarlock, 2001; Muys, 1977; Sax, 2000).

### 2.3.1   Finding the right institutional form

Much research into water resources management institutions emphasizes the analysis of existing institutions at the macroscale. At more local scales, considerable effort is devoted to exploring institutional solutions for addressing problems of scarcity, pollution and conflict. Common Pool Resources studies have focused heavily on the institutions which can fix water allocation problems. These studies frame the central problem as using institutional means to encourage otherwise self-interested, rational individuals to cooperate in managing non-excludable, subtractable resources.

Much institutional research on western water resources has emphasized existing institutions in the west, particularly those operating at the macroscale: prior appropriation, the Endangered Species Act, interstate water compacts, and the Reclamation Act. These have exercised many researchers who seek to critique, applaud, modify or replace these institutions (Schlager, 2006; Ferge, 1999; Sax, 2000; Muys, 1977). At regional and local scales, there is also much debate over what alternative models can replace these existing institutions: water quantity and quality trading markets; public trust doctrines; adaptive management; integrated management; in-stream allocation and so on (Ryan, 2001; Griffin & Hsu, 1993; Norton & Steinemann, 2001; Durham et al., 2002). Many of these solutions focus on addressing enduring problems in hydraulic societies: over-allocation of scarce resources, degraded water quality and damage to riparian environments, and social conflict through the emergence of competing uses. Some of the most theoretically and empirically rigorous approaches to institutional analyses have come from the Common Pool Resources (CPR) literature. Studies in this genre have paid extensive attention to local scale water resource management, particularly irrigation schemes in the developing world (Ostrom, 1990), attempting to diagnose both reasons for failure and providing prescriptions for success.

Within the debate over new institutional forms in western water resources, there has long been a tension between two families of institutional solutions: government and market. The tension has often been manifested as opposing camps arguing for, on the one hand, privatized access to water resources (e.g. Rogers et al. (2002)), and on the other, increased roles for government in managing those resources (e.g. Richter et al. (2003)). Many others argue for hybrid institutional forms, and in practice most water resource institutions in the western US take this form: most typically, the state government provides a legal framework of allocation and enforcement, while diversion rights are owned by individual appropriators. Public incursions into this regulated private market have been seen in the form of the public trust doctrine (Ryan, 2001), and in-stream flow allocation requirements. This tension has also been seen within the CPR literature: the early proponents of Hardin's "tragedy of the commons" theory (Hardin, 1968) were dominated by microeconomic studies arguing that private ownership and clearing markets were the most efficient institutional forms for resource management. Critics were soon

to emerge, however, pointing out that the conditions under which these forms were theoretically efficient rarely existed in real situations, and that the resource economists typically failed to acknowledge that resources can be valued by humans in many different ways, which might motivate more cooperative behavior.

The CPR literature has, indeed, been the venue for the most focused debate on which institutional forms work best to manage resources that are subtractable but not excludable, i.e. one person's use diminishes the resource base for other users, and yet that use (or misuse) cannot easily be excluded (Feeny et al., 1990). The traditional hydrologic example for this kind of resource is ground water (Holland & Moore, 2003; Ostrom & Ostrom, 1999), and water quality is also generally treated as non-excludable (Sarker et al., 2008). According to this frame, the water quality problem is that individually rational appropriators continue to pollute because it is in their self-interest to do so. Their diminishment of the "quality pool" is not excludable, and so no individual has any incentive to stop polluting. The institutional problem, in this case, is finding systems of rules or norms by which individuals can be incentivized to collective action, internalizing the costs of pollution to every appropriator (Ostrom, 1990). Studies have recognized that collective management of common pool resources does occur in many settings, and have focused heavily on identifying why (Ostrom et al., 1999).

Out of this work, the CPR community has developed competing profiles of institutions that work, in general, to address what the literature sees as the central problem of collective action (Adams et al., 2003). Several sets of design principles were reviewed in Ostrom et al (Ostrom et al., 2002), with a synthesis grouped under several themes: resource system characteristics; group characteristics; institutional arrangements; and external environment. These principles were typically drawn either from the meta-analysis of existing case studies of local resource management efforts, or from new empirical studies. The theoretical framework on which such work hinges, Ostrom's Institutional Analysis and Design framework (Ostrom, 2005), has come to dominate the study of institutions and natural resources management.

### 2.3.2 The Institutional Problem: Theory Failing in Practice

Sustainable water resources management through institutions, particularly regarding quality, has not yet been achieved. Institutional study has failed to develop the predictive capacity to determine which solutions will work where, despite considerable effort. This is no more true than in the case of market-based institutional solutions for water quality, which promise much yet have delivered little.

Despite several decades of extensive theoretical and practical debate over institutional forms, we appear to be no closer to sustainable management of water resources through appropriate institutions. The idealized role of the institution in ensuring sustainable resource management remains an ideal. A number of studies have documented dwindling ground water stocks (Konikow & Kendy, 2005), extensively damaged riparian environments (Belsky et al., 1999), and the collapse of major freshwater fisheries (McClure et al., 2003), among other problems.

Existing theory appears to be far better at post-hoc explanation than offering clear predictions for policy makers as to what will work. Economic theory provides plenty of reasons why institutions fail, but these are usually variations on the theme that the real world makes inconvenient departures from the assumptions of economic theory. Such explanations are unhelpful in formulating institutional designs which work in the real world. Similarly, CPR researchers have labored mightily to uncover cases where collective action succeeded, and continues to focus on comparative analysis to produce generalizable principles (Blomquist et al., 2004a; Schlager, 45) and more coherent theory (Ostrom, 1995). The dizzying array and ambiguous formulation of such principles (e.g. Baland & Platteau (1996); Ostrom (1990); Ostrom et al. (2002); Wade (1988)) can be difficult to translate into practice.

The most disappointing aspect of this failure is the considerable benefit to be derived, in theory, from the success of institutional forms. In particular, market-based instruments promise greater efficiency, precision and speed in pursuing a sustainability goal. Theoretically, market-based instruments can generate a correct price for a resource, ensure that allocation of this resource is pareto-optimal, and achieve all this through the very simple mechanism of an open marketplace (Daly & Farley, 2003). While trading institutions have met with some success in areas as diverse as air pollution reduction and ground water management, they have singularly

failed in the area of water quality.

### 2.3.3 Market-based Institutions: Examining the Evidence

On the strength of a very successful sulphur dioxide trading program in the early 1990s, water quality trading markets (WQTMs) have been repeatedly instituted across the United States. This solution is often advanced on the bases that a market can lead to the lowest marginal abatement costs, incentivize technological innovation, and handle informational costs that can otherwise be prohibitive. WQTMs have, in practice, largely failed. Explanations for this failure are diverse, but generally emphasize the details of the institution itself, and rarely questioning the underlying assumptions on individual resource user behavior.

The most commonly cited example of a successful market institution is the sulphur dioxide (SO2) permit trading program implemented in the US in the early 1990s. Implemented under Title IV of the Clean Air Act Amendments 1990, the Environmental Protection Agency (EPA) administered a SO2 allowance trading system. Under the program, allowance holders were able to trade their excess allowance with other companies, or purchase additional allowance to avoid violating their maximum. The theoretical outcome is that the emitters with lowest emissions reductions costs will do so, and sell their allowances to those for whom emissions reductions are more expensive, leading to equal marginal abatement costs and minimal total abatement cost (Stavins, 1998). Further, because the informational requirements are distributed among polluters, a considerable informational load is removed from the government (Burtraw et al., 2005). The program has exceeded emissions reductions targets, built up an allowance excess, saved up to $1 billion annually (Burtraw et al., 1997), and come in cheaper than comparable purely regulatory programs (USEPA, 1997; Gomez-Ibanez & Kalt, 1986). Studies have found the market to be generally efficient, to achieve cost reductions relative to alternatives, and deliver considerable environmental and health benefits (Burtraw & Mansur, 1999; Burtraw et al., 2005).

In the wake of the relative success of the SO2 allowance trading system, water quality trading markets (WQTMs) were seen as highly viable mechanisms for reducing water pollution. By 2005 there were more than 70 WQTMs active in the US (King, 2005), supported by favorable

1996 and 2003 directives from the EPA (USEPA, 1996, 2003). Most of these programs did not exist prior to 1989 (Woodward & Kaiser, 2002), and have focused exclusively on surface water pollution. Woodward and Kaiser argue that the driver for most programs was the increasingly widespread adoption of total maximum daily load (TMDL) programs, which despite being originally required in the Clean Water Act had not been extensively implemented until the late 1990s. Water quality trading markets have been established in states as hydrologically diverse as Colorado, Wisconsin, New Jersey, Florida and California, with most markets focused on trading between point sources. As with air quality permit trading, theoretical benefits of WQTMs are cost-effectiveness and incentives for technology innovation and diffusion (Stavins, 2000; Atkinson & Lewis, 1974; Hoag & Hughes-Popp, 1997). Other theoretical benefits include lower conflict with regulators, since individual polluters are free to meet their quota in any way they see fit (Wolf, 1998).

Despite the considerable theoretical benefits of WQTMs, actual implementations have singularly failed to meet expectations (Pharino, 2007). Most established trading markets have seen dismal trading activity among point sources, and no trading involving agricultural non-point sources (King, 2005). A few programs have achieved their pollution reduction goals, but this typically occurred without much trading activity (Woodward, 2003). Fang et al report on only one relatively successful experimental program implemented on the Minnesota River, which in three years generated five major transactions (Fang et al., 2005). Nelson and Keeler describe how most of the 40 water quality trading initiatives reviewed in 2003 have failed to generate any significant cost savings over traditional regulatory programs (Nelson & Keeler, 2005).

Within the frames of both CPR and resource economics, there is a clear emphasis on institutional details as explanations for this widespread failure. CPR studies emphasize how the "rules of the game" (Ostrom et al., 1992, 404) may fail to adequately incentivize sustainable behavior (Ostrom & Ostrom, 1999), or how they may fail to appropriately restrict access to the common pool (Schlager, 2006). The economics literature is similarly focused on details of implementation: transaction costs (Fang et al., 2005; Stavins, 1995), definition and enforcement issues (Fang et al., 2005; King, 2005; Horan & Ribaudo, 1999), handling of uncertainty (Nguyen et al., 2006) and other inappropriate institutional design choices. To be clear, this approach

is not in itself a problem. Addressing these concerns is an important area of work, particularly in the Colorado basin setting: transaction costs within the prior appropriation framework are high; enforcement is a critical problem when state agencies are underfunded and under-resourced; and there is considerable uncertainty in water use and management derived from complex climatic signals, poorly understood geology, and a complex legal system. My argument here, however, is that a neglected area of concern is the implicit assumptions regarding resource user behavior. It is possible that even if we succeeded in reducing transaction costs, achieving perfect monitoring and enforcement, and eliminating uncertainty, we would still fail to appropriately incentivize institutional participation and sustainability gains *without the right understanding of resource user behavior*.

### 2.3.4    Implicit Behavioral Assumptions and Analytical Limitations

Studies of institutional failure typically make unquestioning use of assumptions regarding the behavioral characteristics of resource users. This is particularly true for economics, but also for more nuanced common pool resources studies. Empirical work, however, is beginning to question these assumptions, and develop an alternative body of theory on resource user behavior. This alternative body critiques rational self-interest as the necessary foundation of institutional analysis.

Implicit in many studies of the failure of markets is the assumption that individual traders conform to a particular behavioral model. In economics, critiques generally assume that traders are rational actors who are seeking to minimize their pollution abatement costs, not to minimize the pollution itself (I am concerned here with the positive assertions of many economists that this is how traders *do* behave, rather than the similar but separate and normative assertion that this is how they *should* behave; see also section 2.3.9). For example, Fang blames transaction costs for market failure: high transaction costs mean costs for traders are higher, which according to rational actor logic, means the traders will demur (Fang et al., 2005). This, however, is self-evidential reasoning. The hypothesis on the damaging effects of transaction costs is self-evidently true once one finds evidence that transaction costs exist, if the whole analysis is conducted on the basis that self-interested actors react aversely to transaction costs (see

similar arguments from King and Nguyen for uncertainty and monitoring/enforcement; King (2005); Nguyen et al. (2006)). Again, this is not to dispute the existence of transaction costs (or uncertainty, or monitoring/enforcement), or their potential negative impact on trading, particularly in the LAB setting. But because these approaches hold fixed, narrow assumptions on behavior, they may overestimate the impact of transaction costs (or uncertainty, or monitoring/enforcement), and will certainly ignore behavioral change as a way to address such problems if they do exist.

The behavioral assumptions of the CPR literature, while proceeding from economics, allow for considerably more sophistication. CPR now allows for dimensions of reciprocity, reputation, trust, altruism, collective rationality through communication, and bounded cognitive resources (Ostrom, 1998; Schlager, 45). But even this more nuanced set of assumptions is still based on a basic model of self-interested rationality. This extends back to the original CPR study, Hardin's "Tragedy of the Commons" (Hardin, 1968), which did not question the assumption that the behavioral foundation of resource over-exploitation was rational self-interest.

Unfortunately for these critiques, empirical evidence indicates that the rational, self-interested model (as developed in rational choice, and boundedly rational choice theories) is far from a fundamental element of human decision making. It is likely, instead, to be a special case of a broad set of alternative behavioral regularities (Gintis, 2000). Studies from behavioral economics and cognitive science document that behaviors departing from the RCT/B-RCT stable are not just anomalous supplements to rational actor theory (Ostrom, 1998; Schlager, 45; Smith, 1991). The effect on analysis is subtle, but important. With different behavioral models, the problem still remains one of creating incentives for sustainable action (Ostrom & Ostrom, 1999), since it is clear that individuals do not always behave in sustainable ways. But the range of solutions changes, because we are able to explore systems of rules which work to handle (and even exploit) far more than just self-interested, rational action. It also opens the way for explanations of institutional failure to go beyond examining the rules, to focus more attention on the actors.

### 2.3.5 Reviewing the Evidence Against Traditional Behavioral Assumptions

The classical model of the self-interested rational actor (rational choice theory, RCT) was developed as part of Walrasian economics. This model holds that self-interested and rational actors engage competitively in market institutions, leading to equilibrium conditions. This model has held up extremely poorly to empirical testing, and has also faced severe theoretical critique. Some even argue that the popularization of its fundamentally normative assertions has led to more self-interested behavior, and consequent damage to society.

Adam Smith and Bernard de Mandeville (de Mandeville, 1924; Smith, 1937) were among the first to argue that self-interest was the driving force for human behavior. Later, a procession of economists — Walras, Coase, Friedman — worked this assumption into general equilibrium theory, which argues that markets evolve out of individual agents exercising their self-interest through economic preferences (Friedman, 1955). Such preferences are usually said to be transitive, complete, sovereign and static, in that the preferences are ordered, stable across possible combinations of options, independent of external conditions, and invariant with time. "Walrasian economics" describes self-interested, fundamentally rational actors engaged competitively in an equilibrium market (Gowdy, 2008). A "rational" actor in Walrasian economics is one who makes the most preferred choice among a set of alternatives ranked by some preference metric (Richter, 1971; Sen, 1977).

This model suffers greatly when subjected to empirical testing, yet still holds powerful sway over many economists and decision makers (Altman, 2006). As a brief (and non-exhaustive) selection, empirical studies have shown that: humans do not always behave in ways resembling the self-interested rational actor, nor do they always act "as if" they obey the laws of rational choice theory (Haigh & List, 2005; Gowdy, 2008); individuals often display behavior directly contradicting their self-interest (Clore & Huntsinger, 2007; Guagnano, 2001); preferences can evolve over time (Norton et al., 1998); individuals may have fundamental social interests, above and beyond self interest (Bowles & Gintis, 2002); choices can be inconsistent (DeShazo, 2002); individuals do not and cannot optimize (Thaler & Benartzi, 2004); and individuals often select strategies inappropriate to context (Barber & Odean, 2008; Odean, 1999).

Importantly, the RCT has been shown to fail to predict behavior in a variety of more com-

plex situations (Fowler & Smirnov, 2005; Dent et al., 1995), exactly the kind of situations rarely explored in more traditional economic and game theoretic realms where the RCT appears to work well (Gintis, 2000). Given that empirical evidence may be moot to a discipline which is little concerned with measuring its theoretical positions against reality (Sen, 1977), it is fortunate that Walrasian economics has faced some severe theoretical critiques. "Self-interest" as a motivation is tautological and, in fact, explains nothing (Sen, 1977; Chouinard et al., 2008). "Utility" is similarly meaningless: if a decision maker is consistent, then the decision maker is always maximizing personal utility, which tells us nothing about what motivates the clearly different behaviors of egoists, altruists and terrorists (Sanchez-Cuenca, 2008). In its formulation and application, RCT prejudges the correctness of the rationality assumption (Hay, 2004), stifling paradigmatic debate. There is also the risk of RCT becoming self-fulfilling, since the RCT will do best when it describes humans who have internalized the tenets of the RCT (ibid). Studies from behavioral economics have shown that framing a problem as a matter of rational self-interest can lead to more self-interested strategies (Rodriguez-Sickert et al., 2008; Reeson, 2008; Miller, 1999; Lynne, 1995). Indeed, this may be intentional given the debate over whether humans should be rational (Broome, 2007; Kolodny, 2005), but note that this is not the focus of the present discussion.

### 2.3.6   Evolving beyond the rational choice theory: imposing bounds

Herbert Simon popularized what seemed to be a reasonable modification to rational choice theory: bounded rationality. This model imposes resource bounds on human decision making: individuals cannot optimize, so they "satisfice". This model has come to form the foundations of many disciplines studying resource management, but particularly common pool resources studies.

One of the first to seriously challenge the RCT was Herbert Simon. On the basis of empirically-supported flaws in the RCT behavioral model, Simon articulated a modified rational actor model, the "boundedly rational" decision maker (Simon, 1990). A boundedly rational actor remains fundamentally self-interested and rational, but has limited information, memory, computing power and other resources with which to make the decision. In essence, Simon's "satis-

ficing" agent retained a simplistic behavioral model but dispensed with the ideal of optimality (Goodrich et al., 2000). The work of Kahneman and Tversky, in particular, provided apparent empirical confirmation of Simon's theory. Through empirical testing, these authors identified a large number of what they term "heuristics and biases", (Tversky & Kahneman, 1974). Tversky and Kahneman argued that most people used simple heuristic solutions instead of full optimization, and used their empirical work to show that such heuristics could lead to non-rational choices (Tversky & Kahneman, 1981; Tversky et al., 1990).

Two dominant schools of thought have developed within natural resources management, since Simon's early critique. One school, mainstream environmental economics, continues to assert the pedigree of RCT and dismiss empirical critiques of the theory as "anomalous" (Gowdy, 2008). Another school has embraced the spirit of Simon's work, and sought to explore the further implications of bounded rationality. This is particularly true for CPR studies, which has embraced the B-RCT as a framework for analyzing problems of resource over-exploitation, while retaining RCT as a foundation for analysis (Ostrom et al., 2002). Many contemporary CPR studies have been focused on adding more complex "attributes of human behavior" on top of the fundamental B-RCT model (Jager et al., 2000). Ostrom describes this effort as developing a second-generation model of rationality (Ostrom, 1998), layering satisficing behavior, heuristics, learning and various norms (reciprocity, reputation, trust) on top of the basic assumptions of rationality and self-interest. There is, nonetheless, a strong sense within CPR studies that the B-RCT has good explanatory power in common pool resource cases (Schlager, 45). As Ostrom argues, "the clear and unambiguous predictions stemming from complete rational choice theories will continue to serve as a critical benchmark in conducting empirical studies and for measuring the success or failure of any other explanation offered for observed behavior" (Ostrom, 1998, 16).

### 2.3.7 Why the B-RCT is insufficient...

Herbert Simon's revolutionary concept of cognition being resource-limited is generally agreed to be empirically valid. Yet the B-RCT is still susceptible to many of the empirical critiques of RCT, indicating that it is not sufficiently valid for uncritical inclusion in institutional analysis.

Dogmatic and exclusionary application of B-RCT ignores the potential of alternative theories, and philosophy of science can be used to reveal contradictions in justifications for the B-RCT. This indicates that B-RCT is insufficient for the purpose of achieving a more complete and useful understanding of human behavior, particularly in resource management settings. However, evidence that the B-RCT can hold under certain conditions, suggests it is in fact a possible (albeit partial) theory. B-RCT has some validity, as long as it is subsumed within a broader theory of behavioral heterogeneity. In the following section, I argue why the B-RCT is not sufficient, and in section 2.3.8, I outline the evidence that supports keeping B-RCT in the stable of possible behavioral models while no longer assigning it a privileged role.

The empirical and theoretical literature generally agrees on the reasonable and helpful nature of Simon's innovation (Klaes & Sent, 2005). The more fundamental problem with B-RCT is the extent to which it still retains core tenets of RCT, and the extent to which this embeds empirically unfounded behavioral theory at the core of many analyses of resource management problems. Contra Ostrom, a theory which has had its limitations so comprehensively demonstrated by empirical study is hardly a good candidate for a "critical benchmark" (Ostrom, 1998, 16). Economists and CPR theorists who argue that the RCT is simply the least bad option for describing human decision making are guilty of subverting their own paradigm (Simon, 1979). This argues against the B-RCT as being sufficiently explanatory, and against its implicit usage in studies of water resources institutions.

Some defenses of the B-RCT argue that it should be retained because no single theory has emerged to comprehensively subvert its provisions. This is a philosophically-dubious argument from necessity. The prosaic reality of scientific inquiry is that a new theory may be weak initially, but show promise of becoming more justified with time. Similarly, a new theory may address failures of the old theory, even thought it does not cover everything the old theory explained, and in time may lead to more comprehensive theories (Maxwell, 1972; Feyerabend, 1970; Lakatos, 1968; Kuhn, 1962). The argument may be sufficient to retain B-RCT as a possible explanation, but it does not justify its exclusive use. A theory with so many empirical holes cannot hope to exclude other theories, however incomplete they may be.

Other defenses of the B-RCT argue that in the simplicity of the B-RCT lies its strength:

any more complex explanation is somehow less valid. Occam's Razor is a folk theorem, and informally applied can be helpful, but it has been repeatedly challenged by contradictions often found in natural science. For example, Newtonian mechanics provide excellent and simple explanations at scales of greater than one angstrom and considerably less than the speed of light, but fail outside these bounds (Eisenbud, 2007; Rudenko, 1991). Newtonian mechanics is in fact explicable as a special condition within the context of the theory of relativity, providing a parallel to arguments that B-RCT is explicable within the context of broader theories of human cognition. The simplicity of the B-RCT is attractive, but may effectively remove every cognitive sophistication which makes humans uniquely human (Nelson, 2005). As Kahneman puts it, "the alternative to simple and precise models is not chaos" (Kahneman, 2003, 1449).

### 2.3.8 ...and why the B-RCT is still necessary: arguments for behavioral heterogeneity

I have gone to considerable lengths so far to highlight the empirical and theoretical work showing the B-RCT to be a limited and often incorrect means of describing human behavior. I have also argued it to be inadequate to the task of achieving a deeper understanding of resource user behavior. However, I also noted that no theory or empirical evidence has emerged to suggest that the B-RCT should be discarded entirely. The empirical and theoretical work I have discussed goes so far as to dismiss the universal use of the B-RCT, but not its use altogether.

I do not question the mountainous evidence for the existence of limits to human cognition, temporal or otherwise. I also acknowledge work showing that modifying the conceptions of rationality with cultural assumptions (so-called plural rationality; Grauer et al. (1984)) can be effective at improving the explanatory power of the traditional RCT model. Most importantly, though, I note the ample evidence to suggest that B-RCT can adequately explain human behavior under certain, narrowly defined conditions (Sontheimer, 2006; Gowdy, 2008). Empirical studies indicate that B-RCT can be effective in predicting behavior relating to foraging, perceptual trade-offs, motor movements and other simple behavioral cases (Camerer et al., 2005). In addition, some work has shown that in the case of organizations and institutions, individually non-rational strategies may aggregate to give outcomes far closer to the predictions of B-RCT

(Camerer & Fehr, 2006).

The B-RCT does indeed have a "clear and unambiguous" nature, as Ostrom suggests. This lack of ambiguity clearly sets it apart as an attractive choice for research. In the development of theory, it is far easier to find exceptions to a general rule, then to find a general model which handles all specific exceptions. Once the theory is established, it then becomes intellectually safer to use the same model (even with misgivings), particularly if the primary purpose of the study is not to elucidate the underlying behavioral model. This is one contributory reason for the widespread use of an otherwise empirically-discredited theory.

The rationality and self-regarding preferences typically assumed by economic analyses make mathematical analyses of human systems far more tractable than more complex alternatives (Camerer & Fehr, 2006). Until the last decade, sophisticated social simulation tools like agent-based models were largely out of the reach of the average researcher. These extremely powerful computational tools are steadily reducing logistical restrictions on the complexity of behavioral models. Even if this is the case, however, there is a tension between the complexity that new computers make possible and the need to fulfill the optimization/satisficing requirements of the B-RCT. Reconciling this tension can result in either vastly overcomplicated models with many special exceptions in specific cases, or excessively simple models that allow speedy solution (Gigerenzer, 2006).

Finally, the use of B-RCT as opposed to more complex models satisfies a preference in most sciences for the simpler solution to a complex problem. The proliferation of such "fast and frugal" heuristics in computational economics and psychology is understandable given the appeal of "simple, psychologically plausible algorithms that, counterintuitively, perform as well as — or sometimes even better than — more complex ones" (Newell, 2005, 11).

There are additional arguments for and reasons why the B-RCT has seen such wide use across many literatures. But the four reasons cited are enough to suggest that the B-RCT should be a part of any considerations of human behavior. My critique of the B-RCT is focused less on its existence altogether than on its frequently dogmatic application in a manner that pre-judges its validity, particularly in more complex cases. I disagree with Ostrom that a theory with such limited apparent real world application should be a benchmark; the evidence sug-

gests it should be an end-member. But rather than offer my own particular brand of behavioral model to replace the B-RCT, I prefer to follow Guth and others in suggesting that there may never be "convincing truly general theories of human behavior" (Guth & Kliemt, 2004, 364). With ArkAgent I am concerned instead with building new and integrative theories of human behavior using the B-RCT *and* other theories from the behavioral sciences (Newell, 2005; Pesendorfer, 2006; Rieskamp et al., 2006; Gowdy, 2008). I accept and operationalize the lack of a unitary behavioral model, through the idea of *behavioral heterogeneity*. Just as Herbert Simon's original modification of the RCT may have been to "criticize neoclassical economists for their lack of interest in the formal foundations of rationality" (Klaes & Sent, 2005, 38), my thesis is organized so as to criticize many resource management studies for their apparent lack of interest in properly understanding the implications of alternatives to the B-RCT.

### 2.3.9 Side note: positive, normative and functional perspectives on economic rationality

Before going further, it is worth underlining this important point: I am not making a normative case for any particular brand of behavior or particular intervention strategy. As a field, economics has long distinguished between the Chicago school, which holds that economically efficient behavior is a natural and ultimate objective in and of itself (Coase, 1960; Ehrlich & Posner, 1974); the Yale school, skeptical of the rational market and advocating for government intervention to correct market failures (MacKaay, 2000); and the Virginia school, which is skeptical of both the Chicago and Yale schools, and emphasizes a methodologically-neutral approach to studying the interaction between individual choice and institutional outcomes (Parisi, 2004). My approach is that all schools are missing the point. I am not concerned with whether or not individuals should be encouraged to behave in an economically rational manner; or, whether we should be intervening with government or market tools. Instead, I am concerned with what it takes to achieve sustainable social-hydrological systems, and all indications are that this will involve encouraging *sustainable behavior*. This kind of behavior may or may not be economically rational, and I am agnostic as to whether a particular behavioral model has more or less normative value than any other (but I am clearly against any theory which makes universal

assumptions of rational or boundedly-rational behavior). My concern is instead with incentivizing sustainable behavior, whatever form that behavior might take, and to do so we need a more empirically-based understanding of how individuals relate to sustainability institutions, and what implications this has for institutional analysis and design.

### 2.3.10 Behavioral heterogeneity: concept and role in institutional analysis

Behavioral heterogeneity is rooted in differences in cognitive abilities, material conditions, and developmental paths among individuals. It is a recognized phenomenon, and has explanatory power at the macroscale. In economic studies of institutions, behavioral heterogeneity is recognized, but often ascribed to shallow differences in material conditions and rational preferences. CPR studies recognize behavioral heterogeneity as a factor in diminishing or increasing the chance for cooperation, but are limited in the sophistication with which they treat this concept.

The general concept of behavioral heterogeneity is well supported by an enormous empirical and theoretical literature (Bowles & Gintis, 2002; Chouinard et al., 2008; Olander & Thogersen, 1995), and some argue it is the defining characteristic of human society (Kirman, 1992). I characterize behavioral heterogeneity as fundamental differences in behavior between individuals, explicable partly through variation in their material circumstances (Meghir & Pistaferri, 2004), partly through their varied cognition (Gomez & Wilson, 2006) and partly through their unique path to a construction of self (Harter, 2001). Cognitive and behavioral heterogeneity has considerable power in explaining how real world social systems stay robust in the face of shock and high error rates (Gintis, 2004), and the diversity of institutional forms and dynamics we see in almost every realm of society (Kirman, 1992). Still, the concept of behavioral heterogeneity remains more a methodological (Janssen & Ostrom, 2006) approach and an analytical puzzle (Jager & Mosler, 2007) than a coherent explanatory theory.

Within economics (and even neoclassical economics), behavioral heterogeneity is increasingly seen as having important effects on institutional performance, and as responsible for much of the "anomalous" performance of institutions relative to the ideal case (Amilon, 2008). Authors have explored forms of heterogeneity as diverse as cultural predispositions to individualist versus collectivist behavior (Cox et al., 1991), varieties of auction bidding behavior

(Gerber & Lewis, 2004), the effects of variable trader intelligence (Yeh, 2008), and variable preferences in the context of water quality trading (Peterson et al., 2007). A common theme in most economic studies, however, is the "cognitive biases" approach, where behavioral heterogeneity is manifested on the basis of different beliefs and some preferences, without deviating from self-interest and rationality as fundamental driving mechanisms of cognition and behavior (Coval & Shumway, 2005). This is a shallow concept of heterogeneity, consistent with Walrasian economic assertions on endogenous preferences, and little recognition of the evidence that behavioral strategies may depart from limits imposed by the RCT.

Resource user heterogeneity has also been extensively explored in the CPR literature. The concept is defined variously as social, economic and/or cultural differences within or between resource user groups, although economic and material dimensions tend to dominate. Its importance is clearly recognized in several empirical, theoretical and review studies (Janssen & Ostrom, 2007; Poteete & Ostrom, 2004; Kopelman et al., 2002; Lange & Liebrand, 1991; Lange & Kuhlman, 1994; Roch & Samuelson, 1997). Nevertheless, the literature still tends to be limited to a shallow concept of behavioral heterogeneity. Discussions of behavioral heterogeneity tend to focus only on whether it increases or decreases the likelihood of cooperation in creating or operating within institutions (Bardhan & Dayton-Johnson, 2002; Kopelman et al., 2002). Baden articulates this presupposition in CPR theory: "it is a fundamental truth that people in a commons have an incentive to ignore the social impact of private behavior" (Baden, 1998, 51). This immediately precludes a deeper form of behavioral heterogeneity, since it asserts that all individuals have this incentive, not just those with a behavioral strategy that is self-interested.

### 2.3.11 The implications of "deep" behavioral heterogeneity for institutional dynamics

An excessively simple treatment of cognitive and behavioral heterogeneity has implications. One is that we may be incorrect in our explanations for institutional failure. Another is that we may end up promoting forms of behavior which we in fact seek to minimize. A third is that we may miss opportunities to improve the effectiveness of our institutional designs.

By not examining the implications of cognitive and behavioral complexity, we may risk

incomplete explanations for failure. One fact distinguishing the SO2 market from nitrogen, phosphate and other water pollution trading regimes is that the SO2 market is exclusively comprised of large corporate entities. Each such entity, particularly public ones, can be characterized as having a single, simple motivation: profit maximization (Jensen, 2002; Kaysen, 1957; Mueller, 1994; Sheppard, 1994). It is no surprise, then, that permit trading in this context performs so well: a perfectly efficient trading program relies on the assumption that firms minimize their total production costs (Hahn, 1989). But when a permit trading program is translated from corporations to individuals or smaller organizations, the theory is not modified to account for their much more heterogeneous preferences and behaviors (Wiklund et al., 2003). Indeed, most laboratory studies of trading markets treat individuals as adequate proxies for corporate entities (Cason & Plott, 1996; Murphy & Stranlund, 2007), despite evidence that individuals do not behave like corporations (Estes, 1995; Werhane, 1985).

Incomplete and incorrect theoretical explanations have real implications for institutional performance. What I describe above is effectively a naive downscaling of a market instrument designed originally for the classical concept of a firm. Not only does this downscaling inadequately account for more behavioral diversity at the small business and individual level, but it may actively undermine extant motivations for pro-environmental behavior. "Crowding-out" theory holds that extrinsic motivations (e.g. monetary incentives) can destroy intrinsic motivations (regarding an activity as valuable in and of itself), and has significant empirical support from field and laboratory trials (Frey, 1999), including in the context of trading regimes. Frey argues that trading regimes result in contradictory effects: while they might reduce pollution by making violation more costly, they may destroy any pre-existing interest in reducing pollution out of a desire to safeguard the environment.

More sophisticated behavioral analysis can provide all sorts of opportunities for improved institutional design. By ignoring real behavioral diversity, market institutions like WQTMs miss out on opportunities to fundamentally address failings in their operation, because they either do not identify the problem correctly or use the wrong set of incentives. For example, a traditional analysis would ignore the fact that individuals sometimes protect the environment for reasons entirely unrelated to profit maximization (Dietz et al., 2005). Such pro-environmental

values can be leveraged for improved institutional performance, as the record on recycling programs demonstrates (Hopper & Nielsen, 1991).

In the next section, I marshal empirical evidence to suggest that the irrigation communities of the LAB possess greater behavioral complexity and diversity than might traditionally be expected (particularly given the tendency of agricultural economics studies to treat farmers as monolithic profit maximizers). In view of the foregoing discussion, I will argue that unexplored potential exists for leveraging such behavioral complexity and diversity in the cause of hydraulic sustainability.

### 2.3.12   Farmers as complex and varied decision makers

It is a core argument in this thesis that real water users show far greater variety in the attitudes and behaviors they display than traditional economic and other decision support models allow. Farmers are perhaps the most important stakeholders in the LAB, both from the perspective of the volumes of water they use and the potential they have to help address the salinity problem. What general and specific empirical evidence is there that farmers are heterogeneous decision makers?

Bennett provides an early and convincing account of the many non-economic values farmers in different communities bring to bear on their objectives; and the variety of farmer strategies born of complex adaptation rather than simple wealth maximization (Bennett, 1969). Similarly, the storied history of Mormon irrigation suggests behavioral complexity well beyond the explanatory power of neoclassical theory (Arrington & May, 1975; Arrington, 1984; Gardner, 1917). More recent studies of farmer decision making have increasingly recognized farmers as having multiple, often competing objectives to achieve (Sumpsi et al., 1996; Ban, 1998; Wallace & Moss, 2002), that the traditional single-minded utility maximization model is insufficient (Willock et al., 1999a; Dent et al., 1995), and that there is considerable heterogeneity in the factors influencing farmer decisions (Pennings & Leuthold, 2000) as well as attitudes farmers bring to their decision making (Kuehne et al., 2007). Beyond evidence that farmers make complex decisions (Roncoli, 2006), there is little agreement on a single behavioral model (Beedell & Rehman, 2000), and many studies suggesting complex antecedent variables still re-

tain an emphasis on a single outcome variable (profit or production maximization) (Willock et al., 1999b). A number of studies suggest that farmer decision making can be characterized as quick, incremental, adaptive, qualitative, and tailored to particular locales (Ohlmer et al., 1998; Pannell et al., 2006; Wilson et al., 1988; Goetz & Debertin, 2001). Willock and others review studies of farmer decision making, distinguishing between farmer attitudes and objectives, and farmer behavior. These authors review evidence that while farmers clearly place emphasis on generating and increasing their on-farm income, other important and often competing values include social (prestige, continuing farming traditions), expressive (pride in ownership and meeting challenges), and intrinsic (enjoyment of work, independence, environmental conservation) (Willock et al., 1999b). Other studies characterize the value sets as conservation, lifestyle and economic (Maybery et al., 2005).

Concordant with broader evidence that farmers are not simple maximizers of their production function (Lin et al., 1974), there are plenty of studies identifying both instrinsic and extrinsic/instrumental attitudes towards conservation (Vukina et al., 2008), and ample evidence of pro-environmental behavior among farmers. Some authors suggest that financial factors are the strongest determinants of pro-environmental behavior (Norris & Batie, 1987). Other authors point to evidence that psychological factors can be as important or more so than financial factors (Willock et al., 1999a; Ervin & Ervin, 1982; Lynne et al., 1988). Still others find that physical factors can overwhelm attitudes for or against conservation behavior (Ervin & Ervin, 1982). Michel-Guillou and Moser identify strong correlations between social factors and pro-environmental behavior, but also argue that most farmers have a fundamental respect for the natural system on which their livelihood depends (Michel-Guillou & Moser, 2006). Perhaps the most widespread conclusion is that the relationship farmers have with their environment is complex and mediated by social, physical and psychological factors, implying that farmers may be responsive to all sorts of incentives well beyond simple economic instruments (Burton et al., 2008). This is supported by studies exploring the effectiveness of various inducements to pro-environmental behavior, finding that farmers can be influenced by education, environmental subsidy, social pressures (Traore et al., 1998), economic incentives, regulation and other factors.

There is a considerable volume of work exploring farmer attitudes and behavior specifically in relation to water quality issues. Such research tends to have similar findings: diverse attitudes towards recognizing and dealing with the problem, different a priori behaviors contributing to or reducing the problem, and a range of responses to pro-conservation inducements (Lichtenberg & Zimmerman, 1999). Studies focusing exclusively on salinity issues have found that farmers can often be acutely aware of salinity problems, but lack either the personal knowledge, resources or collective institutional support to adopt changes to either technology or practice (Ritzema et al., 2008; Haw et al., 2000). In a departure from mainstream analysis, Pannell and others argue that the traditional economic explanation for salinity problems, as the result of externality relations among farmers located upstream or downstream of one another, is over-emphasized. They suggest that, in many cases, on-farm salinity can be dramatically reduced by a farmer's efforts regardless of whether upstream farmers are doing the same, and so efforts should be devoted to communicating this fact (Pannell et al., 2001). This is echoed anecdotally in the lower Arkansas Basin, where considerable reductions in local soil salinity can be achieved on an individual basis by the use of appropriate technologies and practices (pers. comm. Gates 2009). Farmer attitudes towards salinity, and readiness to deal with salinity, may be treated too simplistically by government authorities. Existing inducements, which emphasize distributive fairness and business security, may be less effective than incentives which leverage trust between farmers and awareness of broader community benefits to salinity reduction (Marshall, 2004). In the only published study of the economic dimensions of salinity reduction in the LAB, Gorelick and Lefkoff provide a classic example of how the salinity problem is often conceptualized, and by extension how the search for solutions is framed: these authors treat farmer willingness to reduce irrigation applications as an inverse function of their willingness to forgo short term profits on the farm (Gorelick & Lefkoff, 1990).

From what little information exists on irrigator attitudes and behaviors from the LAB itself, there is nothing to suggest that heterogeneous behavioral models would be unsuitable for this population. A review of the Arkansas River Water Bank Pilot Program suggests, indeed, that water users in the basin have complex constructions of self, varied cognitive abilities, and highly variable material conditions. This evidence indicates that the RCT and B-RCT are insuf-

ficiently explanatory in the case of LAB water users, and that there are empirical grounds for a study of behavioral heterogeneity in the system. The Arkansas River Water Bank Pilot Program was instituted by the state of Colorado in 1991, with the purpose of facilitating the movement of surplus water from rights holders to those who had a water shortfall, through the provision of an infrastructure enabling lease, sale or option of water rights. It managed to attract four deposits totaling 123.25 acre-feet, and three registered bidders, but failed to complete a single water transaction (Lepper, 2008). Shortly afterwards, the water bank was closed for lack of interest. On the basis of interviews with stakeholders, Lepper argues that the Pilot Program failed because it came up against powerful resistance to trading water outside of the basin: while the Pilot Program was intended to mitigate the effects of out-of-basin transfers, it did not explicitly prohibit them, and it was these rules which Lepper reports as causing significant controversy (ibid).

The response to the Pilot Program indicates that farmer motivations transcend self-interest and embody complex historical factors. Source-protectionist attitudes among farming communities are common in the West (e.g. Hanak (2005)), regardless of the degree to which communities could benefit from organized markets. In the LAB, the long and arduous history of cultivation, and farmer expressions regarding their work as continuing an effort to reclaim the prairie for productive use (Milenski, 1990) suggests that many members of the community have a deep attachment to the land. This is not uncommon in the agricultural community at large (e.g. Ryan et al. (2003); Zube (1987)), but in rational choice terms, such behavior is inexplicable.

A second insight from the failure of the Pilot Program is that irrigators differ considerably in their attitudes, goals and strategies. Some irrigators had little apparent concern for the community impact of selling their water rights, and justified doing so on a purely self-interested basis (Howe et al., 1990). But many other farmers went out of their way to resist the Pilot Program in the face of the clear potential for economic benefit. This is clear evidence that behavioral heterogeneity exists within the LAB community of irrigators, and the B-RCT cannot, on its own, account for such variation.

Finally, while there is no published study of how LAB farmers might respond to various salinity reduction inducements, anecdotal evidence suggests that some farmers are already

proactive regarding salinity reduction, and that there have been periods since the early 1900s when farmers took active steps to address the problem (pers. comm. Gates 2009)

The main implication of this review is that I am well supported in my portrayal in Ark-Agent of farmer agents as heterogeneous entities, varied in the goals they adopt, the degree of rationality with which they pursue these goals, the social and environmental context they make decisions within, and their predispositions to pro-sustainable versus pro-self behavior. But the review also provides strong support for my choice of the LAB as a suitable setting for testing the potential institutional implications of behavioral heterogeneity. Not only do we have a real salinity problem, but the actors in this system appear to display some of the behavioral complexity I argue is far more widespread than researchers and policy makers typically acknowledge. It provides a great testing ground for a new approach to formulating more sustainable institutions: exploratory behavioral and institutional simulation with complex systems models. I review the precedent for this approach in the next section.

## 2.4 Testing the implications of behavioral heterogeneity through simulation studies

I have sought to construct a case that existing studies of institutional failure, particularly with regard to market instruments, make inadequate characterizations of human behavior and embed these assumptions uncritically. This case makes a number of assertions: that behavioral heterogeneity can have significant implications for institutional performance; and that behavioral heterogeneity can provide opportunities for improving institutional performance. Ark-Agent is an agent-based modeling tool applied to test the implications of these assertions for resolving salinization, a problem of serious practical concern in hydraulic societies more generally, and in the lower Arkansas Basin in particular. The hypotheses laid out in Chapter 3 systematically explore the concept of behavioral heterogeneity in this context, each hypothesis building on the last. In the following section, I review the precedent for using a simulation model in this way.

### 2.4.1 Agent-based modeling, behavioral heterogeneity and salinity

Simulation tools, particularly agent-based models, are clearly suitable for exploring the implications of behavioral heterogeneity and the challenge of addressing salinity concerns in hydraulic societies. Few studies have yet applied agent-based modeling to the study of water quality trading markets, but none do so incorporating anything other than standard rational choice theory assumptions, or in the context of the salinity problem. The CPR literature dominates the simulation of behavioral heterogeneity, but is still limited by adherence to the B-RCT as an exclusive foundation for behavioral models. Common practice in ABM is to overly simplify agent behavioral models, and uncritically adopt bounded rationality as a core behavioral assumption.

Simulation modeling has long been known as a means to make the analysis of complex human-environmental systems more tractable (Axelrod, 1997; Costanza et al., 1993). Simulation modeling is appropriate for exploring theories relating to real world systems, since it provides safe "sandboxes" for conducting otherwise logistically impossible or ethically dubious experiments (Epstein & Axtell, 1996; Epstein, 1999; Jager & Mosler, 2007). Agent-based models have an individual focus, but can facilitate the study of emergent phenomena (Axelrod, 1997; Doran, 1995). Agent models handle heterogeneity and human-environmental dynamics well (Bonabeau et al., 1999; Brown & Robinson, 2006; Axtell, 2000; Epstein, 1999). They are also methodologically neutral with regard to behavioral models, finding application in diverse fields (Smith & Conrey, 2007; Goldstone & Janssen, 2005). Agent-based models represent a good choice for simulating the institutional implications of behavioral heterogeneity.

This is also true for the study of hydraulic societies. The use of agent-based models in agricultural as well as urban systems is widespread (Benenson & Torrens, 2004; Berger, 2001; Happe et al., 2006; Mathevet et al., 2003; Matthews et al., 2007), as it is in the study of water resources management institutions (Lempert, 2002; Pahl-Wostl, 2002). Irrigated agricultural systems have also attracted considerable attention (Schluter & Pahl-Wostl, 2007; Gurung et al., 2006), particularly among the simulation and gaming/companion modeling community (Barreteau & Bousquet, 2000; Barreteau et al., 2003; Becu et al., 2002; Dare & Barreteau, 2003). While there have been a number of equation-based and other "traditional" approaches to integrated

economic-hydrologic studies of agricultural systems (e.g. Weaver et al. (1996)), agent-based models of water quality issues associated with agriculture have attracted few detailed efforts. Carpenter's 1999 study was one of the earliest to apply ABM to a nonpoint source pollution problem (Carpenter et al. (1999); see also Janssen & Carpenter (1999)) but the model was highly stylized and the few that have followed have employed simplistic behavioral models to explore nitrogen and phosphate-derived runoff pollution. No studies use ABM to explicitly address the agricultural salinity issue, and certainly none have examined salinity issues in the lower Arkansas Basin. Similarly, while there have been studies using ABM to study trading markets, there have been no studies of WQTMs in the context of salinity reduction.

Agent-based simulation studies of natural resources are dominated by approaches which are framed with CPR theory (e.g. Dead & Schlager (2002); Purnomo et al. (2005); Schluter & Pahl-Wostl (2007)) and/or adopt RCT/B-RCT (e.g. Ligtenberg et al. (2004); Thebaud & Locatelli (2001)). Many such models do, in fact, develop sophisticated models of human behavior (Goldstone & Janssen, 2005), but these largely remain based on B-RCT foundations. Applications of ABM in other disciplines are also pioneering the use of more sophisticated and empirically-based conceptualizations of human cognition and behavior (Lustick, 2000; Sun, 2008; Turner & Penn, 2002), but these have not yet seen widespread application in water resources studies. There remains a broad tendency in ABM, particularly within resource management applications, to simplify behavioral models and embed uncritical assumptions (O'Sullivan & Haklay, 2000).

### 2.4.2 Why simulate the WQTM?

So far I have drawn extensively on critiques of rational choice theorems. These theorems, particularly the assumption of self-interest, have long been used to uncritically justify the implementation of market mechanisms in the context of sustainability (as well as in other contexts, see Stiglitz (2010)). I persist in experimenting with the WQTM concept for the following reasons:

1. Of all possible institutional solutions to salinity in the LAB context, the water quality trading market is one of the most reasonable practical solutions. This is true independent

of behavioral assumptions. WQTMs have clear informational and transactional benefits in geographically extensive systems embodying considerable physical uncertainty Eheart & Ng (2004). WQTMs have little resemblance to traditional government regulation, a real benefit in the western United States (Dowdle, 1984). They are also a familiar format for farmers in the LAB, accustomed to "market culture" (Lepper, 2008; Sherow, 1990).

2. Despite the best efforts of economists, it is possible to conceive of market institutions like the WQTM as more complex than a simple clearing house for rational traders. Market transactions, often viewed as purely utilitarian interactions by disinterested consumers seeking to maximize their own utility, should instead be considered as moral, ethical and even spiritual arenas for human expression and interaction (Bruni & Sugden, 2008; Fourcade & Healy, 2007). My argument for more diverse and nuanced incentive structures fits well within this frame.

3. The WQTM provides a particularly good vehicle for testing the particular institutional-behavioral dynamics I focus on in this thesis. WQTMs already incorporate incentives of one kind, and it is relatively easy to conceive of additional incentive structures.

4. Finally, as discussed in Chapter 1, the broad failure of WQTMs to achieve many water quality improvements in the United States presents intriguing theoretical and practical problems. Not only are these problems challenging and complex, but they have not been explored in the way I am proposing, which adds potential theoretical and practical benefits to the outcomes of my study.

### 2.4.3 Laying the groundwork for the hypotheses

In this chapter I have sought to touch on some of the key literatures which provide a foundation for the hypotheses: the social and physical context of the salinity problem in the LAB; the role and relative benefits of institutions in finding solutions; the potential drawbacks to the status quo assumptions of boundedly rational actors in such systems when we attempt to model and find policy solutions to salinity and other hydraulic sustainability challenges; what new outcomes may be possible if we change those assumptions; and finally, the use of simulation

modeling in pursuing this form of analysis. In the next chapter, I lay out the core questions and hypotheses which guide the development and application of ArkAgent, and the presentation and analysis of the results.

# Chapter 3

# Questions and Hypotheses

As discussed in the introductory chapter, the problem of salinity in the lower Arkansas Basin resembles a "social dilemma", if we accept the assumptions of rational choice theory and its variants. It is no surprise, then, that in most past work exploring the question of how to design institutions to handle common pool resource issues, the typical assumption is that any given institution must handle universally self-interested populations of resource users. For example, the benefit of a water quality trading market is often ascribed to its lower social cost relative to regulatory mechanisms: in theory, when individual resource users are maximizing self-interest, the most efficient way to solve a social dilemma is to find some institutional mechanism whereby that self-interest can be harnessed for pollution reduction (Fang et al., 2005). In other words, the WQTM resolves the social dilemma by providing trading opportunities and financial incentives whereby self-interested actors can gain through pollution reductions, while incidentally contributing to the public good. For many theoreticians in common pool resources, the principal problem for institutional design is achieving the alignment of individual and collective interest. The traditional formulation of the social dilemma concept implies that these are, a priori, never aligned.

Based on this assumption, the debate over the failure of WQTMs typically ranges over causal questions regarding the impacts of: transaction costs; the limitations of infrastructure; economically unhelpful qualities of water including mobility, predispositions to monopoly economies, supply uncertainty, use diversity, and use interdependency (Young, 1986); trading

institutions features; and government regulations on various scales (Pharino, 2007). A wealth of studies have been conducted arguing for and against particular explanations for WQTM failure.

My central argument in formulating and answering the following questions is that, while all these factors may play some role, the institutional debate is unnecessarily limited: I argue that we should look also at the fundamental behavioral assumptions we make when analyzing institutional performance.

## 3.1 Question and Hypothesis 1

Consequently, the first question I will tackle in the thesis relates to the impact of foundational behavioral assumptions on institutional performance. Specifically:

**Question 1** *What is/are the potential impact(s) of assuming more heterogeneous behavioral models on the success or failure of water quality trading markets?*

Note that in the following discussion, terms like institutional "performance", "success" or "failure" all refer to the sustainability benefits that an institution provides. It may or may not necessarily be correlated with high rates of participation, or any dimension of institutional operation. If I wish to refer explicitly to these dimensions, I will make this explicit.

To help answer this question, I hypothesize that the reasons for institutional failure of WQTMs may have less to do with transaction costs, infrastructural limitations, the nature of water or government regulations (although they all may play some role), and a great deal more to do with the behavioral assumptions underlying institutional analysis and design. Water quality trading markets, just like other market-based institutions, are designed on the assumption that participants in the institution will always try to maximize their own self-interest, through minimizing the costs they incur by participating (or not participating) and maximizing the financial benefits they receive. WQTM success is theoretically founded on the 1st and 2nd welfare theorems: that competitive markets can achieve pareto-optimal resource allocations; and that assuming a competitive market, resource transfer by government intervention can

74

achieve pareto-optimal conditions (Geanakoplos, 2004). The explanatory power of both theorems is critically dependent on assumptions of neoclassical rationality, and maximization of self-interest as a logical consequence of that rationality. Consequently, I develop a hypothesis which tweaks not only the core assumptions on actor rationality, but also their driving motivations. There have been extensive criticisms of the rationality-as-primitive assumption, and many alternative behavioral models proposed. In the common pool resource literature, Simons theory of bounded rationality has predominated (Simon, 1990). This theory makes concessions to empirical evidence, by dispensing with the assumption that individuals have perfect information and infinite computing capacity. Yet, the theory still assumes that users rationalize their self-interest, even if they lack the best cognitive equipment for doing so. Even attempts to broaden this behavioral model to include the concept of altruism still come back to the fundamental assumption that altruism has some benefit to self (Gintis, 2004). On this basis, many studies suggest that altruism evolved as a means to stabilize social systems, as a response to the "warm glow" individuals receive when undertaking altruistic acts (Chouinard et al., 2008; Crumpler & Grossman, 2008), or simply that the behavior is irrational and characterized as a sacrifice (Kaplan, 2000). The tendency of economists to attribute empirical results deviating from rational choice theory as "anomalous" is widespread (Gowdy, 2008).

Instead of being content with the hedged bet of bounded rationality, I choose to base Hypothesis 1 on the ample empirical work which goes significantly beyond Simons critique of the rational choice theorem (Norton et al., 1998; Bowles & Gintis, 2004; Guagnano, 2001; van den Bergh et al., 2000; Gowdy, 2008). Such work allows me to construct a sound case that the assumption of self-interested maximizing behavior, even constrained by cognitive limitations, is applied too often and too dogmatically. The empirical evidence, while not supporting a particular alternative to boundedly rational behavior, does support an assertion of behavioral heterogeneity within any resource user population.

**Hypothesis 1** *Self-interested maximizers versus varied and complex resource users. The assumption of universal self-interested, maximizing behavior among resource users is incorrect. A more realistic assumption of behavioral heterogeneity among resource users can help explain why the theoretical benefits of water quality trading markets have not, in general, materialized.*

With Hypothesis 1, I begin by testing a behavioral model conforming to the tenets of the boundedly-rational choice theory (B-RCT). These agents (Deadman, 1999) are heterogeneous only to the extent that their material conditions vary, and otherwise have identical (unitary) goals and self-interested strategies. This conforms to the typical assumptions made by economic analysis of trading markets, and many CPR studies (Dyner & Franco, 2004; Fenichel, 2009; Feuillette et al., 2003; Kunc & Morecroft, 2007; Sethi & Franke, 1995; Smith et al., 2009). I then compare this "orthodox" model with a "heterodox" model of behavioral heterogeneity, which includes varied goals, varied degrees of self versus social interest (where "social" interest is coincident with concern for social, economic and environmental sustainability; Casey & Lynne (1999); Chouinard et al. (2008); Cory (1999); Cutforth et al. (2001); Lynne et al. (1998)), and varied consistency in behavior between time steps and decisions. Both models build on the standard assumptions of purposeful behavior and bounded access to resources, which are the least controversial elements of simulated human cognition (Sontheimer, 2006). Note that, for the orthodox model this is a relaxation of Walrasian assumptions, consistent with contemporary trends in computational economics (Hommes, 2006). Each behavioral model is tested in a separate experiment in which a neoclassical WQTM is implemented (Pharino, 2007; Woodward & Kaiser, 2002).

My hypothesis is that the WQTM will perform better, and generate greater sustainability gains, when we assume the orthodox model. The WQTM will not perform as well as a truly orthodox model (i.e. one which assumed perfect rationality Sargent (1993)), since we assume bounded rationality. Rational agents lead to efficient markets, since no trade opportunity goes unexploited (Fama, 1965). Consequently, any departure from rationality, even just to resource limitation, leads to market inefficiency. But relatively speaking, because all agents are consistently seeking to maximize profit/wealth/net economic benefits, the orthodox model will still perform better than the heterodox model. In the heterodox model, because there is greater diversity in the motivations of agents and the extent to which they pursue self-interest, there is far less likelihood of market efficiency and a higher chance of market volatility (Frankel & Froot, 1986; Shiller, 1989; Smith et al., 1988). The single species of incentives the WQTM provides (financial gain and avoidance of financial loss) are not well matched to the variety of different

attitudes and behavioral strategies encountered under the heterodox model.

To re-emphasize: while I clearly do not give much shrift to the B-RCT model, I am not dismissing it. In Chapter 3 I outlined reasons why the B-RCT theorem can be useful in particular contexts, and why it remains in widespread use for both theoretical and practical reasons. In this sense, this hypothesis is primarily intended to widen the range of behavioral models that are considered as explanatory and useful in both theoretical and decision-support contexts.

## 3.2   Question and Hypothesis 2

The implication of behavioral heterogeneity is that the WQTM, an institution designed purely to exploit self-interested resource users, will underperform relative to its theoretical capacity when faced with some self-interested and some less self-interested users. I posit that this is because its incentive structures are not matched to the behavioral models of less self-interested users, and so are less effective at encouraging their participation. This raises an important question:

**Question 2** *Given a population of resource users with different behavioral models distributed in space, what happens to a WQTM institution designed with incentives that work best with a behaviorally-homogeneous population? In other words, how does the performance of this kind of institution compare with a similar institutional form incorporating more diverse incentives?*

Incentives, in their most basic definition, are "any stimulus positively influencing the willingness and/or potential of an individual or organization to undertake a desired action, or to abandon an undesired action" (Hesseling, 1996). Various authors offer different classification systems for incentive types: economic or non-economic (Campen, 1992); internal (to the community) or external (Smith, 1994); and positive (motivating and reinforcing), or negative (alerting against or deterring) (Sargent, 1994). Consonant with the literature, I assume that some form of incentive is required to move an individual from one set of behaviors to another, or to enhance an existing behavior. People with different behavioral strategies respond to different incentive types (Fischer & Bliss, 2008), and so I argue that more diverse incentives will better engage a more behaviorally heterogeneous population.

Individuals and groups do respond to incentives, both in their actions and at a cognitive level (Lea & Webley, 2006). The form of incentive can be an important determinant of success (Encinosa et al., 2007; Lindbeck et al., 1999). Most studies have focused on the power of positive or negative financial incentives (e.g. Gurerk et al. (2009); Jeffrey & Shaffer (2007)), but there is also increasing research on the use and effectiveness of non-monetary incentives to influence behavioral change, particularly in the area of energy conservation. This is starting to erode the traditional, narrow view of incentives as either economic or regulatory (e.g. McNeely (1993)): education, informational campaigns, appealing to values, modifying institutional structures, use of moral and ethical symbols, and changing the decision environment can all act as behavioral change incentives (Stern, 1999; Glanz & Mullis, 1988; Fischer & Bliss, 2008). Furthermore, there are suggestions that the way incentives are delivered can have important effects on what behaviors the incentives engender (McMakin et al., 2002).

Among non-monetary incentive types, informational incentives have seen considerable study. Rationalist theory regarding information holds that individuals can be attitudinally disposed to a particular behavior, but lack the appropriate information and/or understanding with which to operationalize that behavior (Nelson, 1970). Information provision is one of the principal interventions that is practicable on the scales likely to be necessary to achieve widespread sustainable change (Grodzinka-Jurczak et al., 2006). It is certainly the most practicable for hydraulic societies, large and complex as they are. However, basic informational incentives alone tend to be unreliable (Lutzenhiser, 1993; Stern, 1999). Targeted post-hoc informational feedback, on the other hand, can be highly effective in fostering sustainable behavior (Abrahamse et al., 2005; Darby, 2006; Seligman et al., 1981).

In other words, empirical evidence suggests that (a) incentives are important institutional instruments; that (b) individuals can vary in their sensitivity to different forms of incentives; and that (c) information presented in feedback form can be particularly effective at incentivizing more sustainable behavior. This is an appropriate empirical foundation on which to develop a hypothesis:

**Hypothesis 2** *Adopting non-economic incentives to exploit behavioral variation. A pro-social WQTM is a trading institution that incorporates informational incentives, providing feedback to WQTM*

*participants on the sustainability benefits of their WQTM actions. This package of incentives is better matched to a broader spectrum of self- and socially interested resource users, and so will be more robust to behavioral heterogeneity than a neoclassical, "pro-self" WQTM.*

With Hypothesis 2 I am arguing that more diverse institutional incentives can better leverage behavioral heterogeneity than traditional incentives. I argue that individuals participating in a WQTM will bring variable preferences and behavioral strategies, and so will adopt different approaches to the WQTM depending on the match between the institutional incentives and their individual constitution. The implication is that institutions incorporating such diverse incentives will result in more sustainable outcomes within the LAB. Hypothesis 2 is similar to Hypothesis 1 in that it is intended to allow the demonstration of how system-wide outcomes can change when we shift our behavioral assumptions. But Hypothesis 2 goes further by allowing me to test what might be appropriate institutional means to exploit those changed behavioral assumptions. The core theoretical benefit of testing Hypothesis 2, then, is to develop a new simulation approach towards institutional design. This brings with it clear practical benefits, if I can show that the overall sustainability of the LAB system could benefit from a "pro-social" WQTM design.

## 3.3   Question and Hypothesis 3

In order to acknowledge the possibility of behavioral heterogeneity in water resource systems, I discard the assumption that all resource users can be described cognitively using the theory of bounded rationality. It follows, then, that I also reject the idea of universally static preferences. These theories assume that preferences for particular actions are exogenous to the market, i.e. they do not vary with time and are not susceptible to direct or indirect influence (Saari, 1998). Once again, empirical studies (as well as common sense) are solidly in favor of rejecting this assumption (van den Bergh et al., 2000) in the form it is traditionally applied, that is, as descriptive for entire populations of resource users.

Substantial effort has been devoted to exploring what factors affect individual discount rates, across realms as diverse as healthcare, natural hazards assessment and resource con-

sumption (Ferster & Skinner, 1957; Finke & Huston, 2003; Akpalu, 2008; Frederick, 2006; Luckert & Adamowicz, 1993). Factors vary with time, and so discount rates must also. In addition, there is evidence that individual identity changes with time regardless of changes in factors (Parfit, 1971), supported by neuroscience findings that brain structure and function can certainly change on both the short and long term (Doidge, 2007; Resnick et al., 2000). Preferences, behavioral strategies and other artifacts of cognition change as factors affecting cognition, and the brain itself, change with time. Arguments that variation in discount rates with time and other factors would even out at the macroscale (Fisher, 1930) assume linearities in change which are not strongly supported by empirical studies (Nyhus & Webley, 2006). Critics argue that static preferences are nothing more than "theoretical entities that cannot be regarded as anything more than hypothetical constructs" (Norton et al., 1998, 198). They marshal theoretical and empirical evidence to show that "consumer sovereignty" is a methodological convenience often justified on the sole basis that it simplifies and adds weight to economic theory. They outline how Tversky and others have provided numerous examples of preference modification and reversal under varying but certainly not implausible conditions (Tversky et al., 1988, 1990; Fischoff, 1991; Irwin et al., 1993); how preferences can be constructed under certain social pressures (Gregory et al., 1993; Slovic, 1995); and how social psychology and anthropology provide evidence for preference changes in social settings (Cross & Guyer, 1980; Harris, 1979).

This evidence indicates that true preferences do not resemble the preferences of Walrasian economics, and one implication is that if preferences are at least partially exogenous, they are vulnerable to influence. If preferences can be influenced, then so can behavioral strategies. The power of institutional framing in engendering both attitudinal and concrete behavioral change is indicated by diverse empirical studies (Bohnet, 2007; Dolfsma, 2002; Johnston, 2002; Rindfuss et al., 2003). Given the sound empirical evidence that attitudinal change is possible and vulnerable to influence, this allows me to pose Question 3:

**Question 3** *What are the implications for institutional performance of assuming a population of behaviorally heterogeneous agents, where behavioral strategies are responsive to institutional incentives (as per Question 2) but also dynamic with time due to attitudinal/preference change?*

In this question and in later discussion, I refer to attitudes and preferences interchangeably. I define both in general as the set of cognitive conditions in an agent which affect the outcome of a decision process and lead to a specific behavior. I use "behavior", then, to describe the outcome of applying a decision process (driven by attitudes/preferences) and taking an action. "Behavioral change" refers to a change in behavior relative to previous choices, and may or may not be driven by underlying attitudinal change.

A large literature in the social and environmental sciences, particularly in environmental psychology, explores the question of what it takes to motivate pro-environmental behavior among those who do not display such behavior or are not predisposed to (Kollmuss & Agyeman, 2002). There is good empirical evidence that behavioral strategies can be changed from anti- to pro-environmental through some kind of external intervention. However, it is less clear exactly what is most effective in accomplishing that transition, and an extensive literature explores this question. For my part, I am less concerned with what actually works, and more with the idea that a less self-interested behavior can be encouraged which benefits efforts to set resource consumption on a more sustainable footing. To begin answering Question 3, I construct a hypothesis around the use of post-hoc feedback information within a WQTM institutional design.

**Hypothesis 3** *Capacity for attitudinal change and effects on institutional performance. Assuming heterogeneous capacity for attitudinal change within a resource user population, a WQTM variant that incorporates non-economic incentives intended to promote behavioral change (S-WQTM) will perform better than a neoclassical WQTM (N-WQTM) missing such incentives. The S-WQTM will also perform better under the behavioral assumptions of Hypothesis 3 than under the behavioral assumptions of Hypothesis 2.*

Hypothesis 2 assumed static preferences, and under such assumptions, a pro-social WQTM variant may be reasonably successful by leveraging pre-existing pro-sustainability attitudes in some resource users. However, without any change in the preference makeup of resource users, success will be limited to the numbers and potential contributions of the pro-sustainability resource users. In Hypothesis 3, I posit the same WQTM variant, but now assume that the

same package of incentives that leverages the pro-sustainability users can also convert at least some pro-self users to pro-social attitudes and behavior. Because Hypothesis 3 allows both leveraging *and* preference shift, the same pro-social WQTM variant seen in Hypothesis 2 will perform even better under the behavioral assumptions of Hypothesis 3. The S-WQTM will still perform better than an N-WQTM under set of behavioral assumptions, since the N-WQTM fails to take account of any pro-social resource users at all.

Again, I am less interested here in what might specifically cause preference change in a real world setting, and more in the implications of seeing this kind of change play out in the context of an institutional intervention. Nevertheless, in exploring this question and hypothesis, much hinges on the precise way in which I simulate attitude or preference change. As discussed, there is plenty of general evidence for the potential for attitude or preference change. However, the detail of what particular exogenous factors will do to particular preferences is more complex, and there is no hegemonic theory of how preference sets are influenced by institutional incentives (Pyrko & Noren, 1998). There is consensus only on that different individuals have different capacities for change. Two compelling theories of cognitive response to incentives, cognitive dissonance and reinforcement, are adopted as useful models for application in the present study.

Earlier discussion (section 2.3.10 in Chapter 2) discussed clear empirical evidence that attitudes and behavior are heterogeneous among individuals. The literature is also clear that different individuals have different capacities for change in their attitudes/preferences (Bazerman, 1997). Building on evidence that simple informational campaigns are often ineffective at promoting sustainable behavior (Abrahamse et al., 2005; Schultz, 2002), more recent studies have started to show evidence that the combination of information delivered in a targeted manner, as post-hoc feedback in the wake of a particular behavioral choice, can lead to more significant and sustained changes to attitude and behavior (Lutzenhiser, 1993). While it is not always clear how post-hoc informational feedback works to encourage behavioral change (DiClemente et al., 2000; Smither et al., 2005), several studies indicate a complex between cognitive dissonance (Festinger, 1957; Dickerson et al., 2003; Stone & Cooper, 2001) and reinforcement (Ferster & Skinner, 1957; Bandura, 1971). This is a view of feedback response which accepts

that as much as there is a chance of behavioral change, there is also a chance of no change at all. The theory of cognitive dissonance, in its simplest form, articulates that receiving information describing one's behavior as different from one's beliefs (or vice versa) can lead to either rationalization (altering attitudes to justify the behavior) or behavioral adjustment. Both minimize the dissonance between belief and behavior (Benabou & Tirole, 2006). Reinforcement theory comprises the idea that behavioral choices depend at least in part on the consequences of those choices (Bandura, 1971), whether those consequences are material or non-material (Kurz, 2002). Empirical studies have demonstrated the capability of both individual and social reinforcement to increase sustainability-oriented behavior (Dziegielewski et al., 1983). For both cognitive dissonance and reinforcement theories, there are considerable uncertainties and complexities as to how the mechanisms operate (DiClemente et al., 2000; Smither et al., 2005; Harmon-Jones & Mills, 1999).

Considering fundamental uncertainty as to what underlies preference change, as well as the heterodox model's assumption of variable degrees of consistency in agent decision making, I equip my simulated agents with a factor that determines not how they respond, but whether they respond at all. This factor varies in the population, and is independent of their cognition and behavior. Second, I adopt reinforcement theory to drive attitudinal change in response to positive or negative feedback relative to the agent's current attitudinal orientation. Third, I operationalize the theory of cognitive dissonance as a "stickiness" in attitudinal change when the received feedback departs considerably from the agent's prior orientation. This allows the simulation of both increased and decreased change in response to incentives.

I acknowledge that Question and Hypothesis 3 side-steps the important empirical issue as to what exactly can foster pro-sustainability preferences and behaviors. Consequently, I do not claim that my model will generate any generalizable answers to the question of what will precipitate the preference sets leading to more sustainable behavior. This will likely be contingent in any given context upon the unique nature of the sustainability problem and its social and environmental conditions, as well as pre-existing attitudes, values, emotions and material conditions. Instead, the theoretical benefit of Question and Hypothesis 3 lies in showing what long term dynamics might result from assuming that preference (and so behavioral) change

through institutional design is possible. Since the preference change component in ArkAgent will be parameterized with empirically-derived values from the LAB, Hypothesis 3 may also have practical benefits for the LAB: affirming the hypothesis will add weight to the argument for the more sophisticated use of incentives in any market institution implemented in the basin, as well as showing the potential consequences for LAB sustainability of doing so.

## 3.4   Question and Hypothesis 4

The only difference between Hypothesis 2 and 3 is a newly modified assumption on how static resource user behaviors vary with time and in response to incentives. While the implied experiment will shed light on the impact of behavioral dynamics, the issue of non-static behavioral strategies has implications for institutional design as well. This leads me to formulate the fourth question:

**Question 4** *What are the implications for institutional performance when we modify an institutional design to include mechanisms to positively reinforce attitudinal (and so behavioral) change?*

This is not the same as Question 3: in that instance, we assumed that some agents would be responsive to a package of incentives, with the nature of the response varying according to the model of behavioral change implemented. That approach only explores the impact of a behavioral assumption of the capacity for attitudinal/preference change, and does not alter the nature of institutional incentives to specifically exploit that capacity. The behavioral models also included assumptions that agent change occurs independently of the behaviors and orientations of other agents in any individual's social and/or geographic context. To test the implications of using specific institutional incentives to exploit the capacity for change in attitudes and behavioral strategies, I formulate Hypothesis 4:

**Hypothesis 4** *Exploiting social dynamics in institutional design. An institutional incentive which uses social mimicry and competitiveness dynamics to reinforce behavioral change towards sustainability, will be more effective at improving institutional performance than an institution without this reinforcement component.*

84

Studies of energy conservation behavior have long recognized the potential for reducing energy consumption through personalized, local feedback on energy use (Darby, 2006; Chetty et al., 2008; Rasanen et al., 2008). Only more recently, however, has the influence of other people's energy use habits been explored for introducing competitive dynamics. The social environment influences behavior through the communication and sharing of behavioral norms (Cialdini & Goldstein, 2004), often resulting in pressure to conform (Coleman, 2004). Conformity theory underlies evidence that resource users striving for social status will be more likely to modify their own resource use behavior if they are aware of other resource users doing the same or better. The greater the extent to which other behaviors are visible, the greater the pressures to conform. This is particularly true for behaviors framed institutionally as "desirable" (Ela, 2009). Basic conformism is superseded, however, by competitive conformism. The combination of realtime feedback on resource use, along with tangible goals for the whole community and individuals within the community, is likely to lead to stronger pro-sustainability behavior (Vandenbergh, 2004). This has been most notable in community recycling initiatives (Carlson, 2005).

A variation on this theme, competition within a community fostered through comparative individualized resource consumption reports, has not been explored in-depth by any published research. However, a number of municipal districts and nonprofits have experimented by using targeted resource consumption reports to generate competitive dynamics, and have had considerable success in achieving reductions in consumption (Kaufman, 2009). There is considerable scope to explore the implications of this dynamic for a market institution like the WQTM, particularly since this mechanism introduces the possibility of self-reinforcing feedbacks and more efficient application of institutional incentives.

In Hypothesis 4, I test the effects of implementing individualized comparative post-hoc informational feedbacks, with the assumption that this will generate competitive dynamics. I equip agents with social awareness, to the extent that they will modify their behavior in attempts to competitively conform. Recognizing the uncertainties in these mechanisms, and my fundamental assumption of occasional behavioral inconsistency in the heterodox model, I include a variable chance that agents will in fact be aware of the comparative component of the

informational feedback. The same basic behavioral assumptions included in Hypothesis 3 are also used here: two different models of behavioral change, but with the added complication of reinforcement of different strategies based on what neighbors are doing. For example, if a pro-social agent is surrounded by pro-self agents, the pro-social agent will tend to lose some of its enthusiasm for its pro-social behavior. The reverse is true for a pro-social agent surrounded by other pro-social agents, where the element of competitiveness enters. For more detail, please refer to chapter 4.

The theoretical benefit of affirming Hypothesis 4 would be to show how a general empirical result (the potential for exploiting human tendencies towards both competitiveness and conformism in social settings) can potentially enhance market institutional design. The practical benefit to the LAB would be to show how the overall sustainability of the community can benefit from the combination of a market mechanism and sophisticated incentive packages designed to account for real world variation in preferences and behavioral strategies.

## 3.5   Question and Hypothesis 5

Questions and hypotheses 2-4 have dealt explicitly with the positive effects of sustainability incentives, when mixed in with the basic financial incentives that a WQTM theoretically provides. However, in a real world system, any sustainability incentives will compete for cognitive space with the strong messages of financial self-interest that typically pervade social-hydrologic complexes. The financial incentive traditionally associated with financial markets is assumed to be implicit only. In reality, not only is self-interest a powerful and explicit message, but that it may be self-fulfilling (Ferraro et al., 2005). Laboratory studies have indicated the potential for this to occur (Frank et al., 1993), while survey research suggests a pervasive belief that others are motivated by self-interest (Miller et al., 1996). The market economy and associated institutions normalize and explicitly incentivize selfish, rational behavior (Ghoshal & Moran, 1996; Guagnano, 2001; Kohn, 1990; Ostrom, 1998; Wuthnow, 1991); according to some schools of economics, this is even an objective (Parisi, 2004) (see also section 2.3.9). This leads me to formulate the following question:

**Question 5** *What is the effect on institutional performance of assuming that behavioral change is also susceptible to informational messages communicating financial self-interest, as well as informational messages communicating pro-social benefits?*

Once again recall that "pro-social" is used interchangeable with "pro-sustainability", following precedent in social systems modeling (Casey & Lynne, 1999; Chouinard et al., 2008; Cory, 1999; Cutforth et al., 2001; Lynne et al., 1998).

A central argument in this thesis is that the mainstream water management institutions are not currently designed with heterogeneous behavioral models in mind. Consequently, the novelty of dispensing with the assumptions of rational choice theory is in the design challenge of finding a set of incentives that exploit the pro-social resource users in a population. But the other side of the heterogeneity assumption is that there are still some self-interested resource users, acting in opposition to more sustainable outcomes. Actors in the LAB are likely to be under very explicit pressures from the otherwise implicit self-interest and profit-maximization principles of the market economy (Hofstede, 1991; Kagitcibasi, 2005; Kilbourne et al., 2002; Korhonen, 2002). Family farms of the LAB, as elsewhere in the United States, are increasingly less able to obtain credit through local banks with roots in the community (Leyshon & Thrift, 1995). They are consequently forced to apply the same dominant social paradigm-framed financial calculus as agribusiness corporations (Bartlett, 1993; Dudley, 2000). Maximizing on-farm income is often not a primary motivation for farming (Ohlmer et al., 1998), yet most farmers have to justify their activities solely on this basis. Similarly, interacting with global agricultural markets imparts an explicit price pressure disconnected from the fabric of rural communities or the benefits of local land and water conservation (Lang, 1999; Gardner, 2000; Ikerd, 2002; Zimmerer, 2007). The common pool resource problem is often viewed as a matter of designing institutions to mitigate boundedly rational self-interested behavior. The struggle may instead be to avoid undermining a predisposition to civic-mindedness with institutions that foster and reward rational self-interest (Bowles & Gintis, 2002; Bowles, 2008; Gowdy, 2008).

Self-interest aside, individuals are clearly influenced by multiple factors in their decision making. Environmental psychology research points towards individuals having multiple goals and frames influencing their environmental behavior at any one time (Lindenberg & Steg,

2007). Any single goal will be impacted by the other goals the individual maintains, and actions are conducted within a frame (defined by the goal) determining what are appropriate decisions in that context. Extrinsic incentives can sometimes compete with and even undermine intrinsic incentives (Benabou & Tirole, 2006). This is supported by wider evidence that individuals process increasing amounts of information with decreasing efficiency (Robertson, 1980).

Given all this empirical evidence that the case tested in Hypothesis 4 is unlikely to be the case in reality, Hypothesis 5 is formulated to explore what implications other kinds of informational incentives might have for institutional performance.

**Hypothesis 5** *Non-economic incentives in the face of competing pressures. (i) Including competing informational incentives in a pro-social WQTM will alter the capacity of a pro-social WQTM to achieve sustainable outcomes. (ii) Actual outcomes will be dependent on the initial distribution of behavioral strategies within a population.*

In testing this hypothesis, I use the same behavioral model adopted for Hypothesis 4, except that all resource users are now susceptible to information they receive regarding their own self-interest. This "information" is conceptualized as a proxy for the multitude of different information real resource users will receive, both explicit and implicit, arguing for the pursuit of their own self-interest.

I argue in this hypothesis that, if I assume self-interest and sustainability oriented informational types are equally effective at fomenting behavioral change towards their end of the behavioral spectrum, then the final outcomes for institutional performance and sustainability will depend heavily on the initial distribution of behavioral strategies. If a population is slightly more predisposed initially to pro-social behavior, the sustainability information will tend to move the population more towards that kind of behavior by the end of the simulation, and vice versa for pro-self behavior. But because of the retarding effect of the contrary form of information, any simulation will leave some users distributed in the median zone, more so than would be expected if I had assumed only one form of information.

# Chapter 4

# Methodology

## 4.1 Chapter Summary

I describe the ArkAgent model, which is an agent-based model of the principal actors in the lower Arkansas Basin (LAB). Agents are simulated with a basic behavioral model, and additional behavioral mechanisms which can be switched on and off depending on the experiment. Model development and validation were conducted through the use of collaborative modeling workshops and modified Delphi surveys. Model calibration and verification were conducted using tools provided by the AnyLogic/Eclipse integrated development environment. Five experiments were conducted, which differed principally in how they varied the agent behavioral mechanisms and forms of water quality trading market. Outcomes were judged on the basis of four response factors: institutional participation (success of the water quality trading market in attracting trading activity); sustainability contribution (success in reducing salinity); economic contribution (success in maintaining or enhancing basin incomes); and social contribution (level of conflict and distributional equity).

## 4.2 The ArkAgent model

ArkAgent is a detailed agent-based simulation model parameterized to the social and hydrologic system in the lower Arkansas Basin. ArkAgent is an "artificial society" (Epstein & Ax-

tell, 1996), an agent-based model applied to the bottom-up study of human social phenomena. Departing from the traditional artificial society approach, however, ArkAgent also includes macroscale institutional rules and an agent behavioral model going beyond the traditional concept of bounded rationality. The model incorporates scientifically valid physical process modeling, in common with more traditional decision support models. ArkAgent is a non-equilibrium model, and the short time period of simulation means that it cannot shed light on system dynamics beyond the ten year timescale. The full model documentation is provided in the Documentation folder on the thesis CD.

### 4.2.1 Agents

The agent-based model includes four stakeholder types: *farmer*, *municipal water manager*, *federal water manager*, and *ditch company manager*. Farmers were selected because they hold the majority of water rights in the LAB, are the most responsible for the salinity problem, and would likely have the greatest role in any solution. Ditch company managers were selected for their close relation to farmers and the valley's irrigation system. Municipal water managers and the Federal water manager were selected for the potential role they could have in resolving the salinity issue. One important real world actor, the Colorado state engineer's office, is actively simulated only in basic administrative functions related to the water quality trading market. Many of the real world functions of the state engineer are included implicitly in the function of the GeoDSS: the water allocation model functions much like the division engineer might, making sure that allocations conform as far as possible with legal decrees, and making adjustments in deliveries where necessary to account for priority. The Colorado water court — another important real world institution — is omitted because it was deemed too technically challenging to implement adjudication, and less directly relevant to the hypothesis testing. To facilitate this, ArkAgent makes the assumption that agents never seriously depart from their decreed diversion amounts.

644 farmers were simulated, distributed across 15 ditch companies (each with a ditch manager). The farmer population was developed on the basis of USDA and Colorado Department of Agriculture data on farm numbers in the basin. 3 municipalities are simulated: Pueblo, Col-

orado Springs and Aurora. The Bureau of Reclamation at the Pueblo Dam Office is simulated with 1 agent. Population dynamics are not simulated, due to the short simulation period.

### 4.2.2 Base Behavioral Model

A basic model forms the foundation of agent decision making, across all agent types. Agents are modeled as purposive entities: they pursue goals through the implementation of plans, which consist of linked sets of actions (Rao & Georgeff, 1998). Agents are equipped with resources to evaluate and conduct actions. "Base" goals are established for each agent type, which qualitatively describe the general orientation of the agent's plans. For example, the base goal of the Farmer agent type is "generate income", while the base goal of the Municipal Water Manager agent type is "deliver safe and reliable water". At runtime, each individual agent will differ in the actual goal they pursue, due to their intrinsic Sustainability factor (S-factor; see discussion in next paragraph).

"Plans" are undertaken in order to achieve a particular goal, and consist of linked "Actions". Actions are functions by which an agent can modify its external or internal environment. The structure of each action is pre-defined, but outcomes may vary at runtime due to changes in the agent environment, preferences and resources. Actions for each stakeholder type are based on the real world water use and management actions most likely to impact the quality of water in the Arkansas River, and that are most relevant to participation in a water quality trading market (WQTM). A "Resource" is a quantity of some material (substantive or non-substantive) useful to agent action. An agent's evaluation of decision options is fundamentally controlled by its **S-factor**. This captures the extent to which the agent privileges its own concerns (pro-self) versus the sustainability concerns of the wider community (pro-sustainability): The S-factor is a measure of agent attitude which has a direct influence on behavioral choices, and scales continuously from 0.0 to 1.0. A higher S-factor indicates the agent will tend to take actions that it perceives have sustainability benefits to the wider community. A lower S-factor indicates the agent will tend to take actions it perceives has greater financial benefits to itself. Note that I also refer to a high S-factor as "other-regarding", since it implies concern for wider community benefits.

91

#### 4.2.2.1 Standard (Decision Making) Process

The S-factor is applied to every agent decision process, but operationalized in different ways. The same three step decision process (the "Standard Process" or SP) is applied to any action that any agent considers:

1. **Establish appropriate physical or cognitive limits for decision making**: ArkAgent agents are bounded in their cognition. They are assumed to have limited access to information, and limited resources with which to sort through decision options. Any decision process begins with the establishment of limits on decision options.

2. **Evaluate decision choices based on consistency with the agent-specific goal**: the evaluation process consists of three steps: (1) obtain a decision score for the options available; (2) modify these decision scores by the agent's preferences, principally the agent's S-factor; and (3) weight the final options by the agent's S-factor to make a final decision. In Step 1, each decision option is rated on "pro-self" and "pro-social"/"pro-sustainability" metrics (each ranking from 0.0 to 1.0). The pro-self metric is calculated by the agent through simple financial cost-benefit analysis, applied by the model to the agent's decision options. Note that the municipal and federal water managers have a slightly modified version of this metric: these managers do not necessarily have a profit motive in their job description, and so in these cases the pro-self metric is translated into benefits for mission of the municipality utility or federal agency. The pro-sustainability metric is defined *a priori* for a given decision by a model database external to the agent. In Step 2, both scores are adjusted by the agent's own S-factor before being used in decision making (see discussion below). The separate metrics means that real world complexity in benefits can be incorporated. For example, some options have benefits both to self and to wider sustainability. In Step 3, the agent's S-factor determines to what extent each kind of benefit is weighted in the final decision making.

3. **Consistency modifier**: Human behavior embodies inconsistencies relating to a complex of factors, including emotions, pressures on cognitive resources, and other unknown factors. These complexities are not fully represented in this simple model, but their role is

acknowledged by including a low probability of inconsistency in every repeated behavioral choice, around 1%.

Note that "pro-social" and "pro-sustainability" are deliberately conflated: I argue in this thesis that a pro-social position in the LAB implies a concern for basin-wide sustainability. "Sustainability" in this context is a unified concept incorporating the social, economic and environmental health of the basin community. This is consistent with recent work interpreting social and environmental awareness as a "social" end to a self versus social spectrum (Casey & Lynne, 1999; Chouinard et al., 2008; Cory, 1999; Cutforth et al., 2001; Lynne et al., 1998).

To simplify real decision processes while not sacrificing conceptual richness, I determine base "pro-self" and "pro-sustainability" scores for each decision option. For the pro-self score, a simple algorithm runs to compute an economic cost-benefit analysis based on the information available to the agent at the time (the algorithm satisfices within the physical and cognitive limits of the agent). For the pro-sustainability score, during development I make a qualitative evaluation of the likely community-wide sustainability benefits of each decision option, storing these options (scores) in a model dataset, accessible to agents (but note the discussion in the next paragraph). I define "sustainability" for the purposes of the scoring as a measure of ecological benefit, explicitly excluding salinity considerations. The ecological benefit is divided into riparian (land-river interface) and aquatic (environments of the river itself). The evaluation is necessarily qualitative, since the score is being provided before the action is taken, and is intended as a general measure. This sustainability measure differs from the post-hoc informational incentive I discuss later, since this latter form of information is highly specific to a particular agent action, constructed using an algorithm, and delivered after the action has been completed.

The use of scores is a simplified proxy for what is clearly a much more complex process: the conjunction of preference, skills, experience and other unknown personal and environmental factors in taking a given decision. The S-factor is, therefore, an agent preference which plays a strong role in decision making, but is not the only modifier of this sort. In select decisions, risk tolerance also affects outcomes (see discussion in section 4.3 below). These are necessary simplifications, to ease development complexity and make analysis more tractable. However, one

outcome of this simplification, reduced individual diversity, contrasts with the overall focus of the research on heterogeneity. The model addresses this by conducting a runtime adjustment of the scores for the agent's particular S-factor. The score adjustment tends in the direction of the agent's pre-existing S-factor (see Figure 4.1 below: $SF$, the agent's S-factor; $score_0$ the score before adjustment; $score_1$ the score after adjustment). This is based on the assumption that real world actors will not have the resources to do full analyses of benefits, and will tend to emphasize benefits that align with their pre-existing orientation. The more extreme the S-factor, the heavier the adjustment, consonant with research indicating that individuals seek out and pay most attention to information which aligns closest with their pre-existing orientation, and that the extent to which they do this depends on how extreme their pre-existing views are. The adjustment is limited to 10%.

$$score_1 = score_0 + \frac{(1.0 - SF)}{10.0} \tag{4.1}$$

### 4.2.2.2   Example of the Standard Process - Farmer Agent

Once a year the farmer agent faces the choice of which crops to plant. The SP begins by establishing the physical and cognitive limits for the decision making. In this case, the farmer agent only has so much water, land and capital to apply to growing a crop. These limits automatically narrow the choice of crops down from the full set. Similarly, the farmer's base goal ("generate income") limits the agent from choosing options which would result in bankruptcy. Note that the base goal does not preclude options which do not generate income sufficient to make a profit.

The next step in the SP is to conduct an evaluation of the available choices. Farmer agents can conceive of two kinds of benefits to any given decision choice: personal financial benefit, and sustainability benefit for the wider community ("pro-social" benefits). The agent develops a set of benefit scores for the remaining crop choices. The agent itself calculates a cost-benefit score for each crop, and receives a sustainability score for each choice from the model. The score is the adjusted to account for the pre-existing orientation of the agent.

The third step in the SP is to use the agent's S-factor to weight the benefits for the vari-

ous crop options. For example, if the agent's S-factor is 0.3, this indicates the agent is 70 % concerned with pro-self considerations, and 30 % concerned with pro-sustainability considerations. The crop with the overall highest rating is selected.

The final step is to modify the farmer agent's choice by the low probability of a behavioral inconsistency. In most cases, the highest rated crop will be selected with no change, but there is a 1% chance that another crop will be randomly selected.

### 4.2.3 Experiment-specific cognitive and behavioral details

The Basic Behavioral Model is employed unchanged in Hypothesis 1 experiments, except that in one experiment all agents are limited to S-factors of 0.0. Experiments under Hypotheses 2-5 add several additional behavioral mechanisms: behavioral responses to deliberate post-hoc informational incentives, capacity for attitudinal change, social awareness, and responses to multiple incentive forms. A comprehensive listing of which behavioral and institutional mechanisms apply in which experiment can be found in figure 4.6, and in the associated discussion in section 4.4.

### 4.2.4 Homogeneous versus heterogeneous agents

In Hypothesis 1, I contrast two behavioral models in separate simulation runs. In the first run, all agents are equipped with an S-factor of 0.0, meaning that they exclusively maximize personal profit in all their decisions. In the second, agents are heterogeneously parameterized with S-factors ranging from 0.0 to 1.0. This requires no supplement to the Basic Behavioral Model.

### 4.2.5 Post-hoc informational incentives

Under Hypotheses 2-5, there is an institutional change: I implement a version of the WQTM which supplies post-hoc informational incentives. These incentives quantitatively describe the sustainability benefit (the sustainability rating, SR) of the agent's more recent form of participation in the WQTM. Participation is defined as any decision that results in a market trade; or

**1. SUS and SIS defined exogenously**

| Crop | Sustainability score (SUS) | Self-interest score (SIS) |
|------|---------------------------|---------------------------|
| Alfalfa | 0.7 | 0.2 |
| Corn | 0.2 | 0.4 |
| Wheat | 0.3 | 0.1 |

**2. SUS and SIS adjusted by agent's S-factor (maximum adjustment 10%)**

Agent S-factor = 0.3; SUS adjusted down by 7%, SIS adjusted up by 7%

| Crop | Sustainability score (SUS) | Self-interest score (SIS) |
|------|---------------------------|---------------------------|
| Alfalfa | 0.65 | 0.21 |
| Corn | 0.12 | 0.42 |
| Wheat | 0.27 | 0.107 |

**3. Calculate weighted score for each crop, and select final crop**

Agent S-factor = 0.3

| Crop | SUS*0.3 | SIS*0.7 | Total Score |
|------|---------|---------|-------------|
| Alfalfa | 0.195 | 0.02 | 0.215 |
| Corn | 0.03 | 0.294 | 0.324 |
| Wheat | 0.08 | 0.07 | 0.15 |

Selected: corn

Figure 4.1: Example of scoring, adjustment and weighting processes for a Farmer agent

a change to behavior and/or technology designed to reduce salinity loading. As with the general sustainability score described in the first experiment, the SR is provided as a continuous scale from 0.0 (no benefit to sustainability) to 1.0 (maximum benefit to sustainability). Given the difficulties in calculating the sustainability benefit of a single action by a single agent, a sim-

plifying assumption is used to translate the benefit of this action: the SR is calculated *as if* the agents action were replicated by all agents in the system, and the subsequent change in salinity measured. "Maximum" benefit denotes the case when the change in salinity resulted in reduction of salinity below that years WQTM salinity targets (individual quotas). The behavioral change which accompanies the institutional change is described by Behavioral Mechanism 1, below.

#### 4.2.5.1 Behavioral Mechanism 1: response to post-hoc informational incentives

Behavioral Mechanism 1 defines how any given agent makes use of the post-hoc SR information supplied by the WQTM. Once received for a given past action, the SR is used to modify the reference sustainability score which any given choice always receives, since that score does not otherwise include an estimate of the sustainability benefit for salinity. The existing score ($S_0$) and the Sustainability Rating ($SR$) are averaged to generate a new sustainability score ($S_1$), which then feeds into the Standard Process as normal (as described in Figure 4.2).

$$S_1 = \frac{S_0 + SR}{2.0};$$

(4.2)

### 4.2.6 Hypotheses 3-5: attitudinal change models

Hypothesis 2 explores change that is behavioral only: the receipt of post-hoc informational feedbacks are used to help weight pro-sustainability choices more highly. However, the underlying attitude which has the strongest effect on behavioral choice, the S-factor, is not modified. Hypotheses 3-5, however, involve experiments on agents that have the capacity to change their attitudes as well as their behaviors in response to post-hoc informational feedback. To facilitate these experiments, the ArkAgent attitudinal-behavioral change model incorporates four further mechanisms.

#### 4.2.6.1 Behavioral Mechanism 2: variable likelihood of responding to change incentives

Behavioral Mechanism 2 (BM2) is a change factor (C-factor), a probabilistically distributed propensity to attitudinal change. It is only activated in Hypotheses 2 and later, and deter-

mines the likelihood that an agent's attitudes will undergo change, regardless of the eventual direction and magnitude of change. It is included to factor in the uncertainty in what we know in general about behavioral models of preference change, and also because many complexities in preference change are not represented in ArkAgent.

### 4.2.6.2 Behavioral Mechanism 3: joint reinforcement/cognitive dissonance model of change

Under Hypotheses 3 to 5, the sustainability ranking (SR) agents received under Hypothesis 2 now also has impacts on agent attitudinal change. The agent compares the SR with its own S-factor. Under a reinforcement mechanism, the agent's own S-factor will tend to shift towards the received S-factor. But under a cognitive dissonance mechanism, the magnitude of this change is greatest when the SR is relatively close in size to the agent's own S-factor. The direction of reinforcement is determined by the position of the SR relative to the S-factor. The magnitude of reinforcement ($m$) diminishes with increasing difference between the SR ($SR$) and the agent's S-factor ($S$), as described in Figure 4.3. The size of the BMSF is an important control on the magnitude of response.

$$m = \frac{1}{(|(S - SR)| * BMSF)} \tag{4.3}$$

### 4.2.6.3 Behavioral Mechanism 4: social awareness, conformism and "competition"

Hypothesis 4 tests whether awareness of other agent's S-factors and scores can affect an agent's own S-factor and behavioral choices. Under Behavioral Mechanism 4 (BM4), the agent is able to respond to information regarding other agents' sustainability ratings, which are included along with the post-hoc informational feedback from the WQTM. The information is provided as the mean SR for the agent's 10 nearest neighbors. For non-spatial agents, the SR is the mean for all other agents of the same type. BM4 is designed such that conformist behavior drives adjustment of an agent's S-factor relative to the mean SR. When the mean SR ($SR_n$) is lower than the agent's S-factor, the agent's S-factor weakens. When the mean SR is higher than the agent's S-factor ($S$), the agent's S-factor increases. While the mechanism is essentially conformist in both directions, the latter increase in the agent's S-factor is designed to replicate

a competitive dynamic: the agent seeks not to replicate the received SR, but instead increase to a level at or above the received SR. The magnitude ($m$) of increase is determined by the same calculation used in BM3 (Figure 4.4). The size of the BMSF is an important control on the magnitude of response.

$$m = \frac{1}{(||(S - SR_n)| * BMSF)}$$ (4.4)

#### 4.2.6.4 Behavioral Mechanism 5: susceptibility to multiple incentive forms

The WQTM in Hypothesis 5 provides post-hoc informational feedbacks for both sustainability and self-interest. The self-interest feedback takes the form of an opportunity cost ranking (OCR). When an agent takes an action related to the WQTM, the action is rated according to the agent's personal economic benefit or loss relative to not taking the action. The higher the economic benefit to not participating in the WQTM, the higher the OCR. For example, an OCR of 1.0 would indicate that the agents participation in the WQTM had no economic benefit at all, only cost. As described in Figure 4.5, the OCR ($OCR$) acts just like the other form of post-hoc informational incentive, acting to increase or decrease the agent's current S-factor ($S$). The higher the OCR, generally speaking, the more the agent's S-factor will be reduced, and vice versa. The size of the BMSF is an important control on the magnitude of response.

$$m = \frac{1}{(||(S - OCR)| * BMSF)}$$ (4.5)

#### 4.2.6.5 Potential agent responses to the salinity problem

Given the large and diverse array of conceivable yet practical ways real stakeholders can tackle salinity, both in response to and irregardless of the WQTM, some paring down of options was needed. In consultation with irrigation technology experts at Colorado State University, the following options were selected:

1. Irrigation equipment upgrades: irrigation equipment options simulated by ArkAgent are flood, sprinkler, surge pipe and drip. Flood still dominates in many parts of the LAB, so

for most farmers upgrade options will be sprinkler, surge pipe and drip. ArkAgent does not allow agents to downgrade (i.e. move to a less efficient form of equipment), since not only is this technically challenging to model, but takes into account the considerable investment required to change irrigation equipment.

2. Ditch sealing: in consultation with ditch managers and regional irrigation experts, it was determined that providing agents with the option of polyacrylamide (PAM) for ditch sealing was reasonable. Trials of the sealant have already taken place on the ditches maintained by the Rocky Ford Highline, Lamar, Fort Lyon and Catlin canal companies, and proved successful (Susfalk et al., 2008).

3. Ditch lining: a much older technique, involving lining the ditch with materials ranging from geotextile matting to concrete. I assumed the lining option is concrete (geotextiles are less available), and developed estimates of installation rates, costs and other parameters accordingly.

4. Changes to irrigation patterns: ArkAgent farmers are capable of adjusting their irrigation plans to reduce over-irrigation.

5. Changes to cropping choices: ArkAgent farmers are capable of shifting their cropping preferences intra- and inter-annually, if they wish, to favor more water efficient crops.

6. Changes to acreages: ArkAgent farmers are capable of adjusting the acreage they plant each year, which may reduce irrigation requirements.

### 4.2.7 Model Environment

Physical environment is represented in ArkAgent with a geodatabase-supported raster. Each cell is of a certain type (land, river, lake) and possess a number of attributes based on its type. The grid also supports georeferencing of ditch company location, farmer land and other locational attributes of agents and infrastructural components in the model. The geodatabase is maintained inside and updated by the model. The type, size and attributes of cells, as well as

the size of the grid, are based on an ArcGIS geodatabase maintained by Colorado State University for the lower Arkansas Basin. Agents engage with this environment in a limited variety of ways depending on their type and function. Farmer agents, for example, can query land cells for their soil moisture. Ditch Company Manager agents can query river cells for the flow at their headgate. The extent of the interaction is closely modeled on the real limitations actors face in getting information from their environment. Agents are only allowed to interact with each other in certain circumstances.

### 4.2.8   Model Institutions

#### 4.2.8.1   Water Quality Trading Market

ArkAgent simulates a water quality trading market (WQTM) institution, which is present in and of central importance to all experiments. The institution has a core framework that remains the same between experimental runs, based on typical implementations of WQTMs in the United States: a given required salinity reduction target which is reduced over time, annual pollution quotas distributed at an appropriate resolution, financial (dis)incentives for coming in over or under quota, and a trading system to allow the inter-user transfer of part or whole quotas. The state engineer agent manages the trading system, enforces quotas and distributes benefits. The state engineer does not need to explicitly consider whether trading affects decreed rights (as it might have to do in a real system), since the GeoDSS allocation model will not allow allocations that depart significantly from decreed rights and historical diversions. To simplify the behavioral models and subsequent analysis, farmer and ditch manager agents are automatically participants in the institution, and their behavioral models do not include the option of opting out of the WQTM. This is an acceptable simplification because although most real world WQTMs involving irrigation districts do not explicitly forbid participants from opting out, in many cases farmers are enrolled de facto by virtue of being shareholders in a participating ditch and cannot single handedly refuse to participate. In Colorado, a real institution would have to contend with the possibility of complaints being filed in water court charging injury related to a trade, which could potentially disrupt participation. However, in this in-

101

stance my interest is not on simulating opt in/opt out decision making, but on the dynamics of trade *assuming* participation, and so this real world complexity is omitted.

Following Pharino (Pharino, 2007), I describe the core elements of the ArkAgent WQTM (A-WQTM) in relation to common properties of real world WQTMs (RW-WQTM):

1. **Type of tradable permit**: A-WQTM uses a tradable discharge permit of stock type, referred to as "quota". Quota is defined in quantity units (kg of salt loaded into the system). This follows the most common practice among RW-WQTMs.

2. **Permit lifetime**: A-WQTM permits are finite, and last 1 year from issuance. This is commonly seen in RW-WQTMs, particularly where highly seasonal agricultural systems are concerned.

3. **Allocation strategy**: following Beder (2001), A-WQTM shares permits out according to a fixed proportion of total permits assigned to each polluter. This allows adjustment of actual permit amount (i.e. kg of salt) from year to year, as the total cap is reduced.

4. **Trading system**: A-WQTM has an allowance trading system, where polluters are provided with a quantity of pollution permits (in this case, a kg amount of salt), and will surrender 1 permit for every kg of salt they load into the system. Individual polluters are able to purchase extra quota to cover deficits, and sell any excess quota, through a central clearing house (the second most common structure used in recent RW-WQTMs, Pharino (2007)). This "cap and trade" system is common (Ellerman, 2005). The clearing house has zero transaction costs and instant updating, which is not common (but see note below). All trading is of the type non-point source to non-point source, rare for the United States but necessary in the LAB since technical limitations mean that the major point-source polluters (basin municipalities) cannot be simulated.

5. **Uncertainty**: not all polluters are spatially equal in the LAB. Some may be located above shale beds, others may be located far downstream and so already making use of highly salinized water, where they are clearly not responsible for some portion of the salt load. To address this uncertainty to some degree, trading ratios are used. These are standard

practice in many trading systems (Woodward & Kaiser, 2002), and describe the "number of units of pollution reduction a source must purchase to receive credit for one unit load reduction" (Pharino, 2007, 39). While standard practice, trading ratios are typically not applied to rectify potential spatial inequities in the system: trading ratios are established in the LAB allow downstream users to trade quota at a 1:1 ratio, while upstream users (who have the benefit of clean mountain water) trade at a proportionally less favorable ratio.

6. **Monitoring**: A-WQTM assumes that the Colorado State Engineer has annual access to actual loads into the system, aggregated at the "regional" level. For GeoDSS modeling purposes, the LAB was divided into 20 regions of roughly similar size, and A-WQTM assumes that the State Engineer has access to monitoring data from each region. The GeoDSS model simulates this monitoring network. The State Engineer assesses loads by region annually, checks individual polluter loads against their quota, and determines whether the polluter has come in at, under or over the pre-assigned quota level for the year. This allows trading to be retroactive (i.e. does not rely on predictions of load in the coming year, but forces polluters to trade over pollution they have already emitted), and so addressing some of the significant uncertainties associated with load prediction. This approach is also used in one of the largest RW-WQTMs, the Grassland Area Farmers Tradable Load Program in California (Young & Karkoski, 2000).

7. **Enforcement**: A-WQTM allows the state engineer to levy fines on individual polluters for exceeding their quota. Fines are redistributed to non-exceeding polluters as incentive payments. Fines which are not paid at the end of the following year are canceled, and so by necessity ArkAgent does not simulate defection: if agents are able to pay a fine, they do so. A-WQTM assumes that agents will not go into debt to purchase quota or pay fines; this is a reasonable assumption, since no institution is likely to be effective or last long in a community of family farms if it ends up bankrupting the polluters. The choice of the state engineer is a choice born of convenience, and does not preclude enforcement being in the hands of another actor or actors in any real world system.

8. **Banking**: A-WQTM does not allow inter-temporal banking of quota, and also clears deficits at the end of each year. This avoids temporal clustering, potentially damaging speculation, and problems in ensuring regular inter-annual progress in reductions (Tietenberg, 2000).

9. **Trading boundary**: A-WQTM is limited to the LAB. This follows standard recommendations for establishing trading boundaries to coincide with the watershed (Pharino, 2007).

10. **Setting the cap**: A-WQTM sets the target of reducing salt load at the Kansas state line to a level close to pristine (water entering the system in Pueblo). Each year, caps are decremented slightly from the previous year, in order to reach pristine levels by the end of the WQTM program period, 15 years. In practice, caps are typically set by a combination of political and technical considerations; in the absence of a simulation of this process, the A-WQTM cap is a subjective but reasonable starting point.

11. **Eligibility**: A-WQTM allows only farmers and ditch managers to participate, since the GeoDSS is only capable of quantifying salt loads due to their activities. This is a potential pool of 659 traders, much larger than most trading programs in the US (Breetz et al., 2004).

12. **Additional economic incentives**: many existing WQTM programs have varieties of educational and financial incentives to accompany the basic institution. In A-WQTM, a grant fund exists to support upgrades to canal infrastructure and irrigation equipment, where they are likely to support goals of reduced salt loading. The grant fund is open to contributions by local municipalities, and is administered by the Federal water manager agent.

The one key feature of the A-WQTM not commonly seen is the zero transaction cost clearing house: agents can query this clearing house for realtime information on available quota, current max, min and mean prices, and other information; they can post quota for sale and purchase quota at no charge or delay. While few RW-WQTMs can boast of this, such a structure is increasingly realistic: given the reach of the internet into rural communities (and empirical

evidence from the LAB that many farmers make regular use of the internet for information on crop prices, water availability and weather).

The basic WQTM structure as described above is modified in Experiments/Scenarios 2-5 through the use of post-hoc informational incentives, delivered directly to each agent participating in the market. These were described earlier in section 4.2.5. In Experiments/Scenarios 2 and 3, the information reflects the sustainability performance of the agent in its last act of participation in the market or salinity-reduction related action. In Experiment/Scenario 4, the information reflects the sustainability performance of the agent and the mean performance of neighboring agents, in any salinity reduction efforts. In Experiment/Scenario 5, the information reflects the sustainability performance of the agent, the mean performance of neighboring agents and the financial benefit or loss the agent sustained in any salinity reduction efforts.

### 4.2.8.2  Water Rights

Each water using agent in ArkAgent is equipped with a water right. These rights are based on the real legal decrees established by the Colorado State Engineer. Agents are aware of these rights, their quantity and their priority relative to other rights. ArkAgent assumes that these rights do not change for the duration of the simulation, although it is important to note that agents are aware that rights can change. This is a simplification to avoid having to simulate the complex physical and behavioral dynamics of changing diversion practices to account for shifts in water rights, but still retains the powerful psychological effect that fear of loss of water rights can exert on decision making. The physical model coupled with ArkAgent, MODSIM/GeoDSS (see section 4.2.9 below), simulates the weekly allocation of historic flows over the time period 1999-2007, in accordance with the demands agents make on the system and within the constraints of the water rights framework for Colorado.

### 4.2.9  ArkAgent-GeoDSS Interface

ArkAgent simulates all agent, infrastructure and environmental functions inside the model, with the exception of water allocation among water users and the simulation of salinity dynamics in the surface and ground water. Water allocation is accomplished by coupling with

MODSIM, a hydrologic simulation tool developed by Colorado State University (Fredericks et al., 1998). MODSIM is a river basin network flow model, parameterized to the demand and storage points of the real Arkansas Basin, along with detailed water rights and actual diversion databases, for the period 1999-2007.

Salinity dynamics in the Arkansas Basin are simulated by a MODFLOW extension to MOD-SIM, termed "GeoDSS" by its developers at Colorado State University (Triana & Labadie, 2007). The GeoDSS uses artificial neural networks trained with ground water and surface water measurements, to calculate the water quality implications of a given set of water diversions by agents. GeoDSS, like ArkAgent, runs on a weekly timestep, and is fed information on recharge (through farmland) and seepage (through canals) volumes by ArkAgent. The capabilities of the GeoDSS, and the underlying training of the ANN, limit some of the functionality built into agents. The Federal Water Manager agent is severely limited in its role, because the GeoDSS cannot accurately handle deviations from historic system inflows from the Federal reservoir while the simulation is running. However, realistically, it is unlikely that the Bureau of Reclamation, heavily constrained by the Reclamation Act and contractual agreements, would be able to change much of its deliveries through participation in the WQTM.

ArkAgent simulates crop production internally, by combining historical datasets of crop yield with a number of "growth factors" (soil moisture, air temperature, pest presence) parameterized to the typical needs of the specific crop type being grown. Growth factors are used to update growth progress, and are updated every time step. This growth model is necessarily simple. However, outputs are calibrated for the base (historical) run so that crop growth dynamics are in a reasonable range.

### 4.2.10   Model operation

The core behavioral model (the 3-step decision framework) has no specific parameterization, except that the agent S-factors may be held at zero or varied. The resources an agent has access to are parameterized with historical data sets as far as possible, and simulated with an appropriate distribution where these are absent.

ArkAgent is run on a discrete weekly time step, with time step 0 beginning on October 7,

106

1999, and the last time step (420) ending on October 25, 2007. The simulation time period was selected on the basis of the available set of water quantity and quality data used in MODSIM and the GeoDSS. ArkAgent starts at the beginning of October because October 1 is deemed by the Colorado state engineer to be the beginning of the water year for accounting purposes.

### 4.2.11  Model development

Model development (construction, verification, calibration, validation and parameterization) was accomplished through the use of multiple methods: semi-structured interviewing, participatory modeling, model validation surveying and psychometric surveying. Each of these is described below. Materials used in each methodology, and full results, are included in the Appendices.

#### 4.2.11.1  Construction I: semi-structured interviewing

In January 2007, 23 semi-structured individual interviews were conducted with a group of stakeholders, selected on the basis of the contacts and opinions of an extension officer for the lower Arkansas Valley, and a Colorado State University professor with more than a decade of experience working with stakeholders. The individual interviews sought to explore the following themes:

- The decisions and actions the stakeholder normally took during a typical water year

- The influence of the decisions and actions of other stakeholders

- The central worries of the stakeholder with regard to their use and management of water

- The stakeholder's knowledge of and relationship with major basin institutions

- The stakeholder's knowledge of and relationship to the hydrologic system in the basin

The semi-structured interviewing was intended to lay the groundwork for developing the detail of each agent type, their behavioral models, and other aspects of the overall model structure. Full details are included in the Appendix in section 8.4.2.

### 4.2.11.2 Construction II/Validation I: collaborative modeling workshops

In November 2007 and November 2008, 1 day collaborative modeling workshops were conducted. Each workshop consisted of a simulation game and a brainstorming session. Both workshops were focused both on developing new model structures and validating existing structures. Validation is often defined as a technical process for testing whether the model output matches data acquired from the system being modeled (Troitzsch, 2004; Gilbert & Troitzsch, 2005). For models in subjective domains, however, it is more reasonable to describe validation as a process of testing the suitability of a model for its particular purpose (Qudrat-Ullah, 2005). ArkAgent was developed with the intent to conduct exploratory analyses of a complex and poorly known system. Technical validation under such conditions is effectively impossible (Oreskes et al., 1994). Under the circumstances, I validated ArkAgent with the intent of achieving broad qualitative resemblance with the modeled system, with basin stakeholders as the principal experts in determining this resemblance. The collaborative modeling workshops were the principal medium for engaging basin stakeholders with the model and its development, and followed companion modeling techniques (Barreteau, 2003). See section 8.4.3 in the Appendix for more discussion of the role of collaborative modeling workshops in this study, and other technical details.

### 4.2.11.3 Validation II: model validation surveying

After the second collaborative modeling workshop, it became clear that participation would not be high in subsequent workshops. To conserve resources and time but still accomplish validation tasks, a remotely administered validation process was conducted during July and August of 2009. This involved an iterative survey, designed to seek validation responses on questions relating to various elements of the existing model, and also to probe the opinions of basin stakeholders on final design issues.

The survey process was based on the Delphi technique (Ven & Delbecq, 1974; Stewart, 1987; Linstone & Turoff, 1975). Delphi is a method of seeking and refining the judgment of a group of experts (Jolson & Rossow, 1971). It is typically applied to complex domains where no sin-

gle expert's knowledge is sufficient (Roth, 1990), where alternative non-expert derived datasets are not available (Rixon et al., 2007), where subjective validation criteria dominate (Linstone & Turoff, 1975), and where there are a large number of potential experts (Taylor & Ryder, 2003). A standard Delphi process typically consists of 2 or more rounds. The first round is a straightforward survey exploring the issues of interest, usually distributed to a relatively large number of participants. Anonymous responses to this survey are then statistically processed, and the distributions of responses are returned to the survey participants. Participants then adjust their earlier responses after viewing the collective opinions of the group, and subsequent iterations may be used to further refine the responses. The Delphi process was originally intended to help experts move to consensus over forecasts for some outcome (Linstone & Turoff, 1975), but more recent applications of the technique have applied Delphi successfully without necessarily seeking consensus (Skulmoski et al., 2007), and in many realms apart from forecasting (Taylor & Ryder, 2003).

The form of Delphi applied here shares with standard Delphi only the iterative approach, anonymous response, and the use of statistics to determine how the responses from round 1 were used in round 2. 65 individuals from diverse backgrounds were invited to participate, in anticipation of a high rate of attrition between the invitation and the first round, and between the first and second rounds (Hamilton, 1996). Each round was administered online, using a web-based survey design tool following recommended design guidelines for electronic Delphi (Chou, 2002). The first round sought responses appraising the general accuracy, but not specific details, of the main elements of the base behavioral model and select additional decision settings for different agents. The second round sought to elicit feedback on proposed changes to the model in the wake of the first round. Full details of each round, results and appraisal of results are included in section 8.4.4 of the Appendix.

### 4.2.11.4    Calibration

For a model of such complexity as ArkAgent, traditional calibration is a controversial process. I follow several authors (Gupta et al., 1997; Marks, 2007; Rouchier) in treating formal calibration as of limited value in establishing the "truth" in a model. ArkAgent is not designed as

a predictive tool, and is intended for exploratory analysis of results from changing simple behavioral assumptions in the context of a complex social-environmental system. Nevertheless, ArkAgent was calibrated using four empirical datasets from the historical record (October 1 1999 to October 1 2007):

1. Crop per acre yields: the mean per acre harvest, in appropriate units, for each of the 6 simulated crops

2. Crop fractions of total acreage: the fraction of total irrigable land occupied by each of the 6 crops

3. Diversions: the total acre-feet of water diverted by ditch companies for irrigation purposes

4. Demands: the total acre-feet of water demanded by farmers for irrigation purposes

In the baseline scenario and for each calibration dataset, the annual and inter-annual root mean square error (RM2E) was calculated. The RM2E provides a measure of "fit" between the simulated data and the historic dataset, i.e. how well the particular model parameterization and structure match the historical record for the real world system (Browne & Cudeck, 1993). The RM2E penalizes simulated results with a severity according to how far those results depart from the calibration dataset. A principal limitation of RM2E is that it typically leads to adjustment of existing parameters or addition of new parameters with spurious origins or values. As far as possible, I limited modification of existing parameters to "reasonable" bounds (relying on expert opinion), and limited addition of new parameters to concepts which had ready support in the literature: namely, farming skill level and risk tolerance. The farming skill parameter follows common practice (e.g. Steiglitz et al. (1996); Ziervogel et al. (2005)) in assuming varied capacity for agents to achieve a given task. To limit the potential "over-fitting" of the model (Moore & Doherty, 2006), the skill parameter was used minimally. Addition of the risk tolerance variable draws on precedent in the use of risk tolerance parameters in farming simulations (e.g. Dake et al. (2006); Aull-Hyde & Tadesse (1994); Paudel et al. (2000)). In recognition of the lack of any contextually-constrained dataset for this parameter, risk tolerance

is parameterized using a normal distribution, and the role of risk tolerance is limited to only where absolutely necessary. Aside from the addition of new variables, select other parameters were varied within historical range and/or within bounds of uncertainty. These included: reference water demands for crops; reference insurance premiums for crops; Kc values for the Blaney-Criddle evapotranspiration formula (Allen & Pruitt, 1986); and several other more minor variables. Given the exploratory nature of the model, and the small number of calibration parameters, a 20% error window was allowed, i.e. calibration was deemed complete when all RM2E measures fell within 20% of the historical values over the period of the simulation.

Figure 4.2 indicates that, overall, calibration was roughly at or below threshold during the entire simulation, although the calibration fit clearly improved with time. Figure 4.3 indicates that for most of the years, and overall, each error measure is at or below 20%. After extensive calibration work, however, error measures for wheat yield, sorghum yield and overall crop yields continued to exceed 20% for the first year. This was likely due to model equilibration. Many model outcomes depend on the interaction of one or more other model variables, and the values of these variables will shift from their initial parameterization as a simulation gets underway. Since many of the model processes are cyclical, variables affect by these processes will evolve to a stable range over time; this range may or may not coincide with the original parameterization. In light of this, initially poor calibration performance is to be expected.

Calibration initially included salt load, since there was ample historical data against which the model could tested. However, even after reducing errors in all other calibration datasets to less than 20%, the baseline run of ArkAgent still showed average annual levels of salt around 30% below the real world datasets. I surmised that the model was probably structurally deficient, since the error was systematic year to year. Unfortunately, in the absence of additional time and resources it was not possible to return to stakeholders and address this deficit.

#### 4.2.11.5   Verification

Verification is a test to see "that the software correctly implements the conceptual model the researcher intended" (Midgley et al., 2006, 11). Verification is generally a technical process where a modeler seeks to identify and address internal errors in the model logic. Verification
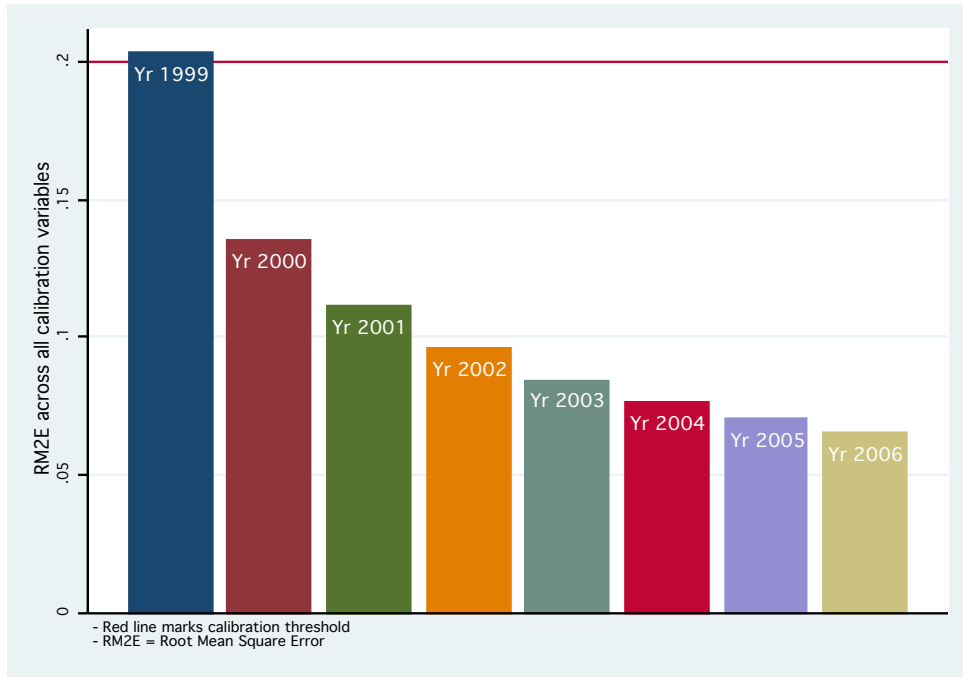
Figure 4.2: Root mean square error across all calibration variables, mean by year, empirical S-factor parameterization. Plot shows that calibration fit increases with time, and apart from the first year, all RM2E variables plot below the 20% calibration threshold.

| | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 |
|---|---|---|---|---|---|---|---|---|
| Vegetable yield | 0.357 | 0.226 | 0.23 | 0.195 | 0.173 | 0.157 | 0.142 | 0.13 |
| Alfalfa yield | 0.202 | 0.196 | 0.132 | 0.117 | 0.106 | 0.099 | 0.096 | 0.086 |
| Corn yield | 0.224 | 0.163 | 0.146 | 0.127 | 0.123 | 0.115 | 0.114 | 0.106 |
| Grass yield | 0.412 | 0.263 | 0.196 | 0.156 | 0.135 | 0.124 | 0.116 | 0.108 |
| Wheat yield | 0.289 | 0.18 | 0.154 | 0.147 | 0.13 | 0.118 | 0.105 | 0.1 |
| Sorghum yield | 0.357 | 0.233 | 0.202 | 0.19 | 0.16 | 0.138 | 0.126 | 0.12 |
| All crops yield | 0.307 | 0.21 | 0.177 | 0.156 | 0.138 | 0.125 | 0.116 | 0.108 |
| Vegetable acres | 0.021 | 0.013 | 0.01 | 0.009 | 0.008 | 0.007 | 0.007 | 0.006 |
| Alfalfa acres | 0.326 | 0.199 | 0.151 | 0.121 | 0.102 | 0.089 | 0.078 | 0.07 |
| Corn acres | 0.12 | 0.085 | 0.063 | 0.047 | 0.038 | 0.032 | 0.029 | 0.026 |
| Grass acres | 0.136 | 0.07 | 0.05 | 0.043 | 0.038 | 0.034 | 0.029 | 0.026 |
| Wheat acres | 0.052 | 0.034 | 0.028 | 0.022 | 0.021 | 0.018 | 0.018 | 0.016 |
| Sorghum acres | 0.004 | 0.01 | 0.009 | 0.007 | 0.006 | 0.005 | 0.004 | 0.004 |
| Diversions | 0.144 | 0.086 | 0.072 | 0.061 | 0.052 | 0.049 | 0.045 | 0.044 |
| Demands | 0.103 | 0.06 | 0.049 | 0.043 | 0.036 | 0.032 | 0.029 | 0.026 |

Figure 4.3: Values for root mean square error across all calibration variables, by year. Indicates that in most cases error measures were at or below 20%, the acceptable margin for the present work.

of ArkAgent was conducted using formal debugging tools (provided by the Eclipse integrated development environment).

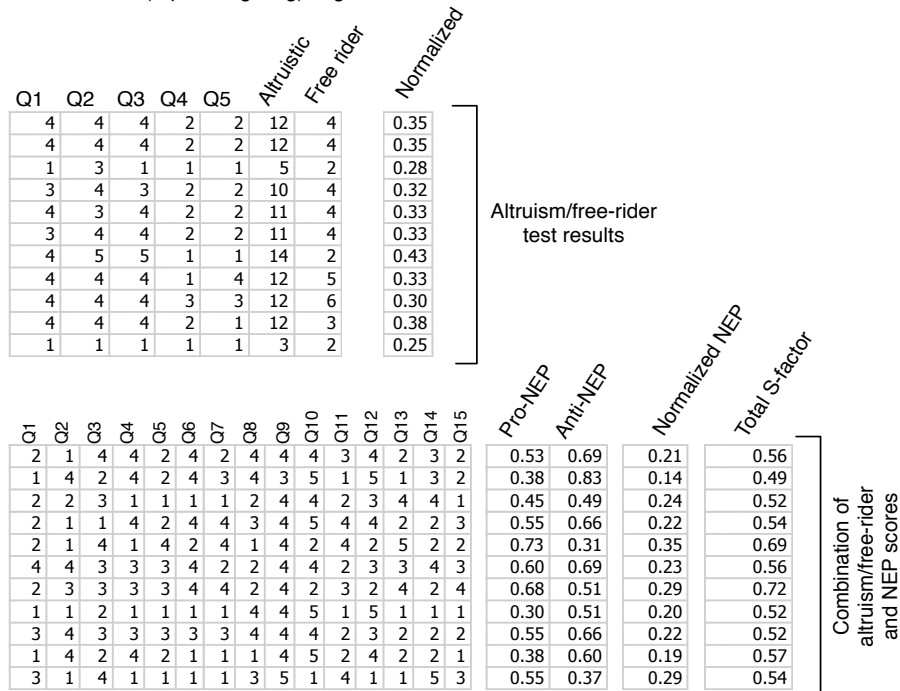### 4.2.11.6 Parameterization: psychometric testing

Three online psychometric tests were conducted in September 2009 to generate an empirical parameterization for agent S-factors (Base Behavioral Mechanism; section 4.2.2) and C-factors (Behavioral Mechanism 2; section 4.2.6.1). Psychometric tests (Likert, 1932; Thurstone, 1927) have the benefit of being accessible to test takers with minimal training and support, and so amenable to large sample surveys. The test was based on test frameworks and questions used in previous studies in environmental psychology, particularly those exploring themes of altruism and pro-environmental behavior (Lusk et al., 2007). The sample set was non-random snowball. The test was sent out to 65 stakeholders. The responses provided information on the actual behavioral heterogeneity of the LAB. Full details of all three separate tests are provided in Appendix section 8.4.5, along with a brief evaluation of the method. Figures 4.4 and 4.5 below summarize the survey responses and the translation of those responses into S-factors and C-factors.

## 4.3 Sensitivity Analysis

Sensitivity analysis is a form of model testing to see how model performance (and modeler conclusions) might change if the model itself is altered. I follow Sterman in arguing that sensitivity analysis should forthrightly tackle structural assumptions in the model wherever possible (Sterman, 1991). The sensitivity analysis I applied to ArkAgent was dictated first by the nature of the simulation. The model is not being applied predictively, and so it is not of primary interest how model outcomes might vary if the empirical datasets varied. Second, the model does not have significant stochastic components, and so replications were not required. Third, sensitivity analysis was limited by constraints on available time and machines (average model runtime was 9-15 hours, and only a small number of machines were simultaneously available for runs). Consequently, sensitivity analyses were conducted in special cases: (1) where no case could be made for initializing the parameter to a particular value, and so a sensitivity analysis could help determine the appropriate range or specific value for the parameter; (2) where a case could be made for a specific initialization, but further insights could be gained by running

**S-factor development process**

1. Based on responses to NEP test, pro- and anti-NEP scores generated, combined to generate NEP score (out of 0.5); based on responses to altruism/free-rider test altruism and free-rider scores generated, combined to generate social interest score (out of 0.5). Scores summed (equal weighting) to generate total S-factor out of 1.0.

| Q1 | Q2 | Q3 | Q4 | Q5 | Altruistic | Free rider | Normalized |
|----|----|----|----|----|-----------|-----------|-----------|
| 4 | 4 | 4 | 2 | 2 | 12 | 4 | 0.35 |
| 4 | 4 | 4 | 2 | 2 | 12 | 4 | 0.35 |
| 1 | 3 | 1 | 1 | 1 | 5 | 2 | 0.28 |
| 3 | 4 | 3 | 2 | 2 | 10 | 4 | 0.32 |
| 4 | 3 | 4 | 2 | 2 | 11 | 4 | 0.33 |
| 3 | 4 | 4 | 2 | 2 | 11 | 4 | 0.33 |
| 4 | 5 | 5 | 1 | 1 | 14 | 2 | 0.43 |
| 4 | 4 | 4 | 1 | 4 | 12 | 5 | 0.33 |
| 4 | 4 | 4 | 3 | 3 | 12 | 6 | 0.30 |
| 4 | 4 | 4 | 2 | 1 | 12 | 3 | 0.38 |
| 1 | 1 | 1 | 1 | 1 | 3 | 2 | 0.25 |

Altruism/free-rider test results

| Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 | Q10 | Q11 | Q12 | Q13 | Q14 | Q15 | Pro-NEP | Anti-NEP | Normalized NEP | Total S-factor |
|----|----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|-----|---------|----------|---------------|---------------|
| 2 | 1 | 4 | 4 | 2 | 4 | 2 | 4 | 4 | 4 | 3 | 4 | 2 | 3 | 2 | 0.53 | 0.69 | 0.21 | 0.56 |
| 1 | 4 | 2 | 4 | 2 | 4 | 3 | 4 | 3 | 5 | 1 | 5 | 1 | 3 | 2 | 0.38 | 0.83 | 0.14 | 0.49 |
| 2 | 2 | 3 | 1 | 1 | 1 | 1 | 2 | 4 | 4 | 2 | 3 | 4 | 4 | 1 | 0.45 | 0.49 | 0.24 | 0.52 |
| 2 | 1 | 1 | 4 | 2 | 4 | 4 | 3 | 4 | 5 | 4 | 4 | 2 | 2 | 3 | 0.55 | 0.66 | 0.22 | 0.54 |
| 2 | 1 | 4 | 1 | 4 | 2 | 4 | 1 | 4 | 2 | 4 | 2 | 5 | 2 | 2 | 0.73 | 0.31 | 0.35 | 0.69 |
| 4 | 4 | 3 | 3 | 3 | 4 | 2 | 2 | 4 | 4 | 2 | 3 | 3 | 4 | 3 | 0.60 | 0.69 | 0.23 | 0.56 |
| 2 | 3 | 3 | 3 | 3 | 4 | 4 | 2 | 4 | 2 | 3 | 2 | 4 | 2 | 4 | 0.68 | 0.51 | 0.29 | 0.72 |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 4 | 4 | 5 | 1 | 5 | 1 | 1 | 1 | 0.30 | 0.51 | 0.20 | 0.52 |
| 3 | 4 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 4 | 2 | 3 | 2 | 2 | 2 | 0.55 | 0.66 | 0.22 | 0.52 |
| 1 | 4 | 2 | 4 | 2 | 1 | 1 | 1 | 4 | 5 | 2 | 4 | 2 | 2 | 1 | 0.38 | 0.60 | 0.19 | 0.57 |
| 3 | 1 | 4 | 1 | 1 | 1 | 1 | 3 | 5 | 1 | 4 | 1 | 1 | 5 | 3 | 0.55 | 0.37 | 0.29 | 0.54 |

Combination of altruism/free-rider and NEP scores

2. Approximately normal distribution fitted to scores, with mean of 0.57 and standard deviation of 0.07. Distribution assumes that extreme values exist but were not sampled, due to small respondent sample. This distribution used to develop agent S-factors in ArkAgent.

Figure 4.4: Summary of results for S-factor survey, with final S-factor values included. Approximately normal distributions (with slight skew) were developed on the basis of these results. Refer to figures 8.1 and 8.2 in the Appendix for the questions in the altruism/free-riding and NEP surveys respectively.

baseline scenarios with alternate initializations; and (3) where further exploration was merited to shed light on analyses and conclusions. The following variables fell under (1):

1. Behavioral Modification Scaling Factor (BMSF): in the Behavioral Mechanisms discussed earlier, the BMSF variable is an important control on how dramatically behavioral characteristics change in response to any given input. Sensitivity analysis suggested that a value of 100.0 was appropriate; variations significantly above or below this value mainly have the effect of speeding up the effects of a given feedback on agent behavior.

2. Stochastic Modifier (SM): in the Base Behavioral Mechanism discussed earlier, the SM is

1. On basis of responses to seven change survey questions, C-factors generated as sum of weighted means (higher value = higher propensity to change).

| Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Weightings | | | | | | C-factors |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4 | 7 | 2 | 2 | 2 | 6 | 0.57 | 0.00 | 0.29 | 0.29 | 0.29 | 0.86 | 0.38 |
| 4 | 3 | 4 | 5 | 5 | 5 | 0.57 | 0.57 | 0.57 | 0.71 | 0.71 | 0.71 | 0.64 |
| 1 | 7 | 4 | 7 | 4 | 4 | 0.14 | 0.00 | 0.57 | 1.00 | 0.57 | 0.57 | 0.48 |
| 3 | 5 | 5 | 3 | 4 | 5 | 0.43 | 0.29 | 0.71 | 0.43 | 0.57 | 0.71 | 0.52 |
| 5 | 5 | 7 | 6 | 5 | 7 | 0.71 | 0.29 | 1.00 | 0.86 | 0.71 | 1.00 | 0.76 |
| 2 | 4 | 2 | 5 | 5 | 5 | 0.29 | 0.43 | 0.29 | 0.71 | 0.71 | 0.71 | 0.52 |
| 4 | 4 | 4 | 5 | 4 | 4 | 0.57 | 0.43 | 0.57 | 0.71 | 0.57 | 0.57 | 0.57 |
| 4 | 4 | 3 | 5 | 4 | 7 | 0.57 | 0.43 | 0.43 | 0.71 | 0.57 | 1.00 | 0.62 |
| 4 | 6 | 2 | 5 | 4 | 4 | 0.57 | 0.14 | 0.29 | 0.71 | 0.57 | 0.57 | 0.48 |
| 2 | 6 | 2 | 7 | 6 | 7 | 0.29 | 0.14 | 0.29 | 1.00 | 0.86 | 1.00 | 0.60 |
| 6 | 4 | 7 | 7 | 7 | 7 | 0.86 | 0.43 | 1.00 | 1.00 | 1.00 | 1.00 | 0.88 |

2. Approximately normal distribution fitted to scores, with mean of 0.59 and standard deviation of 0.14. Distribution assumes that extreme values exist but were not sampled, due to small respondent population. This distribution used to develop agent C-factors in ArkAgent.

Figure 4.5: Summary of results for C-factor survey, with final C-factor values included. Approximately normal distributions (with slight skew) were developed on the basis of these results. Refer to figure 8.3 in the Appendix for the questions in the survey.

a final inconsistency modifier on decisions made by agents. The SM value is a probability that the highest scoring decision option may not actually be chosen after all. The psychometric testing was not designed to gather information for empirically parameterizing this measure, and there is no consensus in the literature on how to characterize classically "irrational" behavior let alone parameterize simulation models. Consequently, some sensitivity analysis was conducted to determine a sensible initial value for the SM. This analysis determined that model outcomes are typically sensitive to changes in the SM on a 1:(0-1.0) ratio. This means that for every 1% change in the SM, model outcomes vary by 0-1%. The actual magnitude of the response depends on the kind of decision involved. Binary decision options ("do A" versus "do not do A") tend to be most strongly affected, since a higher SM increases the chance that a very different decision will be taken. Multiple decision options ("do A", "do B", "do C", ...) tend to be less strongly affected, since a higher SM may not necessarily result in a very different decision relative to the highest scoring option. Given the roughly linear, yet significant effect of the SM on model out-

comes, a nominal value of 0.01 was selected. This translates into a 1% probability that a given choice logic will lead to a sub-rational (i.e. less than highest scoring) decision, or that 1 in every 100 agent decisions will be sub-rational.

The following variables fell under (2):

1. Initial Quota Unit Price (IQP): in any WQTM, an initial unit price for pollution quota has to be set. In some air quality markets, this is done through an auction system, allowing the market to set the price itself. However, auctions of quota are not common in water quality markets, and so the quota price is typically set on the basis of the minimum cost per extra unit of pollution abatement. This information was not available for the lower Arkansas basin, and so based on knowledge of other systems, a value of $0.05 was set as the initial quota. Sensitivity analysis determined that increasing this value by factors of 10 and 100 led to broad increases in average responses across many of the most relevant response factors (institutional performance and economic metrics, in particular). ArkAgent agents are designed to be responsive to quota price, and so this was not unexpected. As will be discussed in Chapters 5 and 6, this sensitivity does not undermine any of the assertions in the hypotheses. However, it was not the case that a simple increase in initial quota price was linearly correlated with the relevant response factor; some unexpected results were seen, and this suggests that the function relating quota price to model outcomes is complex and mediated by other factors.

2. WQTM Management Options (WL): I implemented a context-specific version of a WQTM in ArkAgent, making choices on key features on the basis of a review of past WQTM examples and what seemed appropriate for the LAB conditions. Ultimately, however, choice of one structure over another could be only justified to a limited extent, since I could not know which structures would/not be implemented were a WQTM to be installed in the LAB. One of the most important alternate approaches to WTQM management is the treatment of old quota balances: the ArkAgent WQTM conservatively assumes that at the end of each year a surplus or deficit of quota in each polluter's account is simply zeroed out (after accounting for fines for deficits and incentive payments for

surpluses). A sensitivity analysis conducted to see the effects of allowing both of these features showed that, in broad terms, model outcomes were not greatly altered. The exception, as will be discussed in Chapters 5 and 6, was the effect on market efficiency. It appears these features accentuate pre-existing inefficiencies in the way agents interact with the market.

3. Risk tolerance (RT): as discussed earlier, the risk tolerance parameter was set to a normal distribution in the absence of an empirical dataset. A number of competing risk tolerance models could be advanced for this basin (e.g. risk-seeking; risk-averse), but none can be justified above the normal distribution. Consequently, I conducted a sensitivity analysis to show the effects of picking more and less risk-seeking models. This analysis indicated that while risk strategy sometimes altered the magnitude of response, it rarely altered the overall pattern of outcome or subverted the hypotheses.

For exploratory purposes under (3), the following variables were also varied.

1. Farm skill (FS): the level of skill farmers applied to their activities was varied between monolithically low and high across the population.

2. Length of the WQTM program (PL): the WQTM simulated in ArkAgent was nominally 15 years in length, but I also tested the effects of shifting this down to 8 years and up to 20.

Parameters FS, PL, WL, IQP and RT were varied independently (with two possible values for each parameter) for scenarios 1a through 4. For example, scenario 1a was run separately with low and high farming skill; two more runs were conducted for a WQTM program of 8 years and 20 years, and so on. Scenarios 5a, 5a-LSF, 5a-HSF, 5b, 5b-LSF and 5b-HSF were not varied due to technical problems and subsequent logistical limitations.

## 4.4   Experimental Design

ArkAgent was applied in 11 main simulation experiments. All experiments were run for 420 weekly time steps from October 7 1999 to October 25 2007. Figure 4.6 describes each experi-

ment in terms of the additional behavioral mechanisms and institutional interventions applied. Terminology used in Figure 4.6 is outlined below.

Terminology:

- WQTM-N: a version of a standard water quality trading market in which no post-hoc informational feedbacks are provided to agents.

- WQTM-S1: a version of a standard water quality trading market, in which post-hoc informational feedbacks are provided to agents on the sustainability contribution they made through their most recent act of participation in the market.

- WQTM-S2: identical to the WQTM-S1, but with additional post-hoc informational feedback for the mean sustainability contributions of 5 nearest neighbors (or agents of the same type, for non-geographic populations).

- WQTM-S3: identical to the WQTM-S1 but with the addition of a post-hoc informational feedback describing the personal financial benefit or loss they received from their most recent act of participation in the market.

- WQTM-S4: identical to WQTM-S3, but with nearest neighbor information added in.

- BBM: Base Behavioral Model. Basic functionality for agent cognition included in all experiments.

- BM1-BM5: Behavioral Mechanisms 1 to 5, supplemental cognitive/behavioral mechanisms.

- Empirically varied S-factor: S-factors for all agents parameterized according to a distribution derived from psychometric testing.

- Low-SF skew: S-factors for all agents parameterized according to a beta distribution with a strong kew to lower quartile values.

- High-SF skew: S-factors for all agents parameterized according to a beta distribution with a strong kew to upper quartile values.

Some features to note: Experiment 5a explored the effect of supplying feedback information on the opportunity costs to the individual of taking a particular action versus no action. Experiment 5b did the same, except also allowed awareness of this kind of information among neighboring agents. The stepwise addition of these features was intended to make it easier to identify effects of one behavioral mechanism versus another. Experiments 5a(i) and 5a(ii) were based on 5a and 5b and only set apart by differing distributions of S-factor. Recall Hypothesis 5 asserted that the final outcomes for Experiments 5a and 5b would depend on the initial S-factor parameterization (i.e. our assumptions or understanding of what kind of heterogeneity there was within a population). The two additional runs each for 5a and 5b were intended to explore this.

### 4.4.1 Experimental Response Factors

All five simulated experiments were ranked relative to each other on a set of response factors, which are listed below.

1. Institutional Participation: **number of unique traders** (how many times each agent in the model participated in the WQTM annually, but without incrementing the counter for repeat participation); **clearing success rate** (ratio of quota posted for sale to quota sold); and **volume of quota traded**. Unique participant numbers indicate overall penetration of the WQTM as an attractive decision option within the agent population; clearing success rate indicates how much of the quota posted to the market is clearing to buyers; and the volume of quota traded indicates how much liquidity was present in the market. Overall rankings were decided using the assumptions that a WQTM with the best institutional participation would see unique trader numbers equal to the total population of the basin; an offer success rate of 1.0 (i.e. every unit of quota posted for sale was completely purchased); and the highest total trade volumes.

2. Sustainability Response: achieving sufficient water quality and quantity are important goals of sustainability efforts. Quality contributions were measured by the **salt loaded** during the simulation run. Salt loaded was measured by tracking tons of salt added
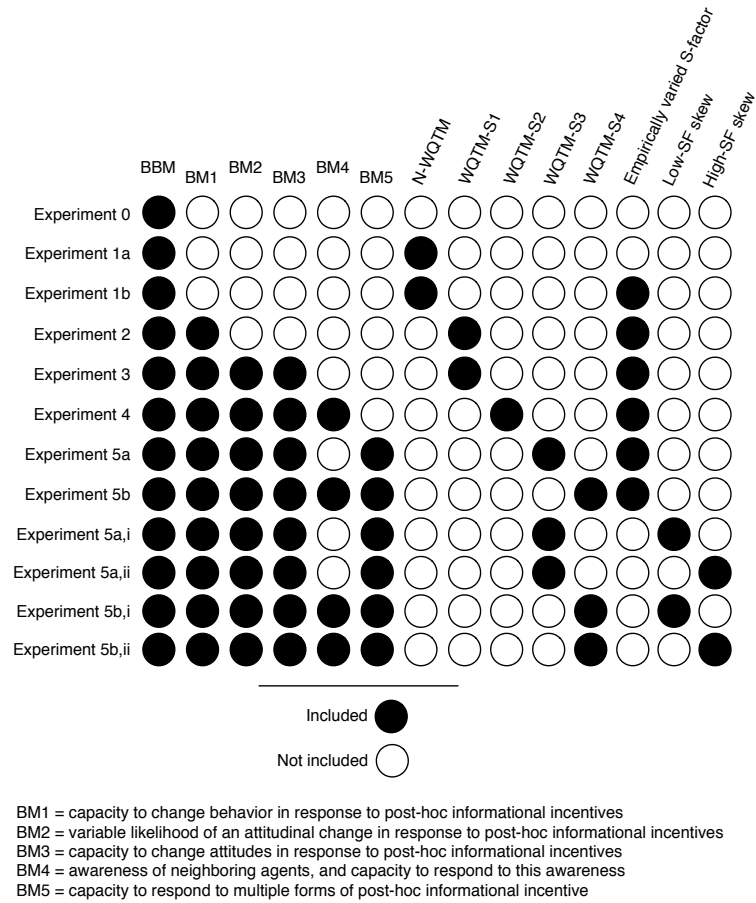
| | BBM | BM1 | BM2 | BM3 | BM4 | BM5 | N-WQTM | WQTM-S1 | WQTM-S2 | WQTM-S3 | WQTM-S4 | Empirically varied S-factor | Low-SF skew | High-SF skew |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Experiment 0 | ● | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Experiment 1a | ● | ○ | ○ | ○ | ○ | ○ | ● | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Experiment 1b | ● | ○ | ○ | ○ | ○ | ○ | ● | ○ | ○ | ○ | ○ | ● | ○ | ○ |
| Experiment 2 | ● | ● | ○ | ○ | ○ | ○ | ○ | ● | ○ | ○ | ○ | ● | ○ | ○ |
| Experiment 3 | ● | ● | ● | ● | ○ | ○ | ○ | ○ | ● | ○ | ○ | ● | ○ | ○ |
| Experiment 4 | ● | ● | ● | ● | ○ | ○ | ○ | ○ | ● | ○ | ○ | ● | ○ | ○ |
| Experiment 5a | ● | ● | ● | ● | ● | ○ | ○ | ○ | ○ | ● | ○ | ● | ○ | ○ |
| Experiment 5b | ● | ● | ● | ● | ● | ● | ○ | ○ | ○ | ○ | ● | ● | ○ | ○ |
| Experiment 5a,i | ● | ● | ● | ● | ○ | ● | ○ | ○ | ○ | ● | ○ | ● | ● | ○ |
| Experiment 5a,ii | ● | ● | ● | ● | ○ | ● | ○ | ○ | ○ | ● | ○ | ● | ○ | ● |
| Experiment 5b,i | ● | ● | ● | ● | ● | ● | ○ | ○ | ○ | ○ | ● | ● | ● | ○ |
| Experiment 5b,ii | ● | ● | ● | ● | ● | ● | ○ | ○ | ○ | ○ | ● | ● | ○ | ● |

Included ●

Not included ○

BM1 = capacity to change behavior in response to post-hoc informational incentives
BM2 = variable likelihood of an attitudinal change in response to post-hoc informational incentives
BM3 = capacity to change attitudes in response to post-hoc informational incentives
BM4 = awareness of neighboring agents, and capacity to respond to this awareness
BM5 = capacity to respond to multiple forms of post-hoc informational incentive

Figure 4.6: Table describing variation in behavioral and institutional models employed in experiments 0-5

to the system over time. Quantity contributions were measured by recording the **acre-feet per week flow of water** in the Arkansas River at the Colorado/Kansas state line, and parsing the resulting time series into simulation mean, inter-annual mean and other statistics. Overall rankings were decided using the assumptions that a scenario with the best sustainability response would have the highest tonnage of salt reductions relative to the baseline scenario; and the highest flow increases relative to the baseline scenario.

3. Economic Response: agent average income levels, and real asset status (cash assets less liabilities) were used to measure any economic contribution of the WQTM variants. Only incomes and assets of farmers and ditch managers were included: in terms of social sus-

tainability, farmers and ditch managers are the most vulnerable to effects of the WQTM. The WQTM is unlikely to have any dramatic effect on the multi-million and even billion dollar budgets of the Colorado Front Range municipalities. Overall rankings were decided using the assumptions that a WQTM with the best economic response would have the highest income increases relative to the baseline scenario; and the highest real asset increases relative to the baseline scenario.

4. Social Response: conflict and equity. Water conflict in the LAB tends to occur when actual river and ditch flows differ from the legally allocated flows. While conflict can arguably force the resolution of problems (Humphreys, 2005; Wick & Bulte, 2006), in this context I assume that conflict is a negative force since it consumes scarce agent resources, and the simulation is not sophisticated enough to adequately represent legal processes. Conflict was operationalized as a subjective perception on the part of farming agents: the level of conflict rises as agents perceive that they are not getting the water they need, their economic prospects worsen, and they believe that these conditions are widespread, and decreases when they feel the opposite in each case. The metric recorded the mean conflict level across all farming agents. Water quality improvements should also be equitable, such that most water users benefit from the WQTM's effects. Equity is important at the very least for helping foster a positive community attitude towards an institution like the WQTM, supporting its longer term viability. This response factor was operationalized as "deficit equity" among ditch companies: if all ditches are experiencing deficit water availability (i.e. they all have less than they need), equity is very high. if only 1 ditch is experiencing deficit, equity is very low. Deficit equity is not a measure of overall quantity sustainability, but indicates how well any negative impacts are being distributed across the population of ditches. Overall rankings were decided using the assumptions that a WQTM with the best social response would have the lowest conflict relative to the baseline scenario; and the highest deficit equity relative to the baseline scenario.

#### 4.4.1.1 Generating scores from response factors

Combined rankings were decided by assigning descending weights for increasing rank positions under each scenario for each response factor. For example, if a scenario achieved 1st, 2nd, 2nd and 4th rankings for each factor respectively, this would translate into total points of 4+2+2+1 = 9 points. A second scenario achieving 1st, 1st, 2nd and 3rd would score at 4+4+2+1 = 11 points, and so place higher than the first scenario. Many of these response factors were generated as non-stationary time series data; in these cases, where simple summary statistics are not appropriate, rankings were accomplished by comparing trends. Trends were calculated using the most appropriately fitting model for the particular response factor.

# Chapter 5

# Results

## 5.1 Summary of results

### 5.1.1 Overall sustainability scores

Figure 5.1, showing the overall sustainability scoring of each scenario, indicates that scenario 2 scored as the most sustainable, while both baseline scenarios scored as the least sustainable. The low scoring peak (around 0.45) indicates that no scenario was straightforwardly more sustainable than any other. These results provide weak support for Hypothesis 1, which argues that a standard neoclassical WQTM (referred to as a "traditional" WQTM) will provide more sustainability benefits under conditions of homogeneous behavioral assumptions than under heterogeneous behavioral assumptions. These results provide much stronger support for Hypothesis 2, which argues that a non-traditional WQTM incorporating post-hoc informational feedbacks will perform better with a behaviorally heterogeneous population than the traditional WQTM: scenario 2 scores higher than scenario 1b. Hypothesis 3 argues that a non-traditional WQTM incorporating post-hoc informational feedbacks will still perform better than the traditional WQTM when we assume agents do not have static preferences and can shift their behavioral strategies with time. These results show this hypothesis to be supported, since scenario 3 (agents with the capacity for change) still outperforms scenario 1b (agents without the capacity for change). Scenario 3 does not perform as well as scenario 2, however, indicating that assuming dynamic preferences can mitigate some of the benefits of
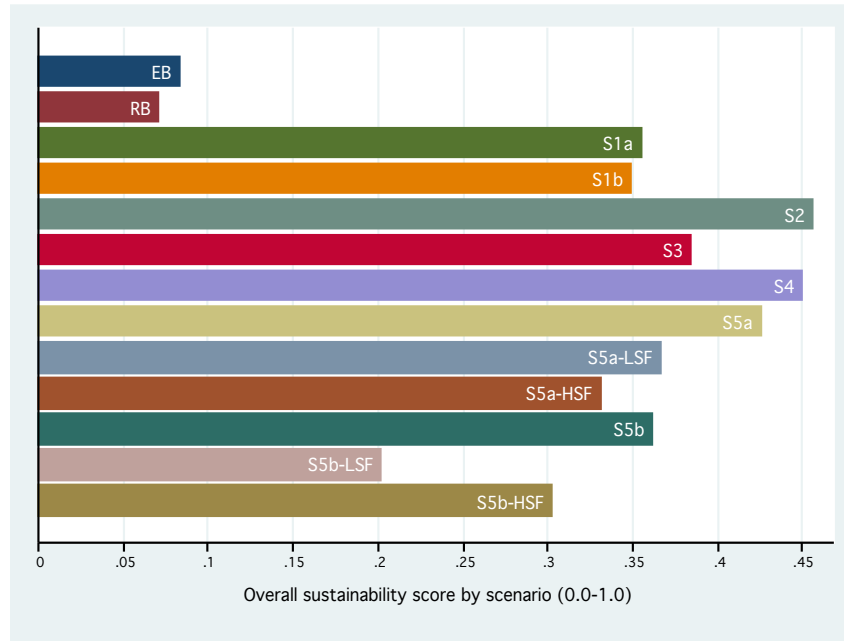
Figure 5.1: Overall sustainability scores across all scenarios. The scores show: weak support for hypothesis 1, since the standard neoclassical WQTM scores higher with attitudinally homogeneous populations than with attitudinally variable populations; stronger support for hypothesis 2, since scenario 2 scores higher than scenario 1b; support for hypothesis 3, since scenario 3 (agents with the capacity for change) still outperforms scenario 1b (agents without the capacity for change); support for hypothesis 4, since scenario 4 scores higher than scenario 3; and finally, mixed support for hypotheses 5. Note that the scores are relative rankings, and do not provide a measure of absolute sustainability. See section 4.4.1.1 in Chapter 4 for detail on how the scores were calculated. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

the non-traditional WQTM. Hypothesis 4 argues that a non-traditional WQTM incorporating mechanisms to exploit neighborhood social dynamics like competitiveness and mimicry will perform better than a non-traditional WQTM without these mechanisms. Figure 5.1 shows support for this hypothesis, since scenario 4 (non-traditional WQTM exploiting social dynamics) scores nearly a point higher than scenario 3 (non-traditional WQTM not exploiting social dynamics). Finally, Hypothesis 5(i) argues that allowing competing incentives in the form of financial opportunity cost information will alter the capacity of a non-traditional WQTM to incentivize more sustainable behavior. I tested this assertion by running "LSF" and "HSF" variants of S5a and S5b: the "LSF" stands for "low S-factor", since all agents in these variants were parameterized with S-factors from a beta distribution skewed to the 0.0-0.2 end of the S-factor range; the "HSF" stands for "high S-factor", since all agents in these variants were

parameterized from a beta distribution skewed to the 0.8-1.0 end of the S-factor range. The bar chart shows that this is indeed the case, since scenario 5a (non-traditional WQTM with opportunity cost information) shows a reduced score relative to scenario 4 and the highest scoring case scenario 2. Adding further weight to the hypothesis is Scenario 5b (non-traditional WQTM with opportunity cost information and social dynamics mechanisms) which scores below several other scenarios including 5a. Hypothesis 5(ii) refines the basic assertion in 5(i), arguing that sustainability outcomes will be dependent on the initial preferences of the population, such that an initially pro-self population will end up with lower sustainability benefits than an initially pro-social population. Figure 5.1 shows that this is the case when social dynamics are present (scenarios 5b-HSF scores higher than 5b-LSF) but not the case when social dynamics are absent (scenarios 5a-LSF scores higher than 5a-HSF).
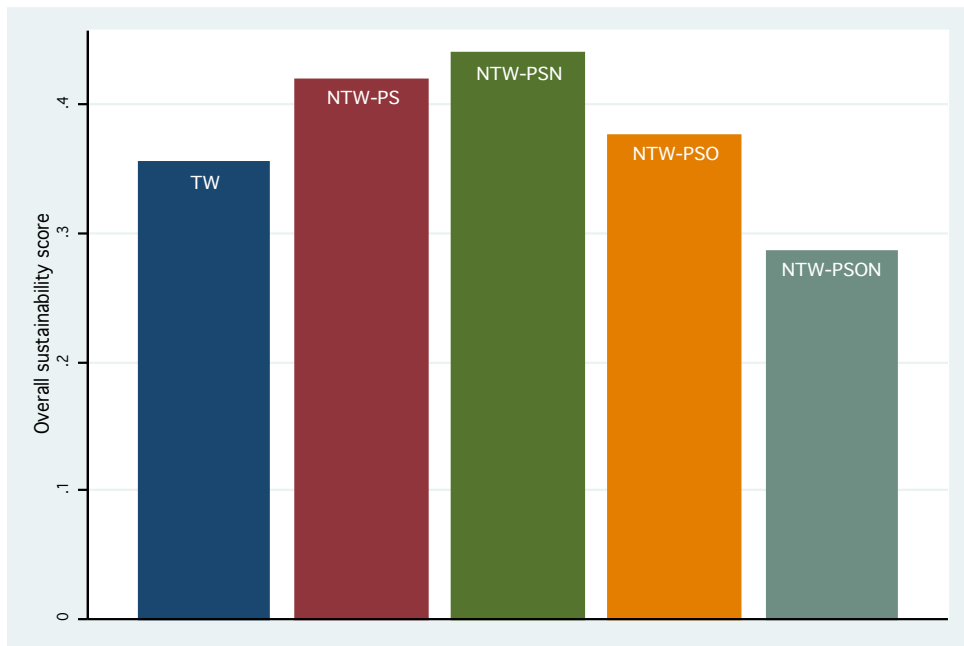


Figure 5.2: Overall sustainability scores grouped by WQTM version (see text for explanation of versions). The plot shows that the most sustainable institutional form was the non-traditional institution, and that this institutional form results in more sustainability benefit except when we allow social dynamics and provide explicit information on opportunity cost. See text for explanation of detail.

An alternate way to organize the data is by WQTM version, figure 5.2. This indicates that the most sustainable institutional form was the non-traditional WQTM with both post-hoc sustainability information and neighborhood information (NTW-PSN). This is supportive of the

overall theory advanced in the thesis, which is that better matching of incentives and exploitation of various dimensions of individual and social behavior can lead to more sustainable outcomes. The plot also indicates that the non-traditional WQTMs out-perform the traditional WQTM (TW), except when we have opportunity cost in a social context (NTW-PSON). This is supportive of the notion that messy reality can intervene in efforts to apply innovative incentives in pursuit of more sustainable individual behavior.

### 5.1.2 Unpacking the overall sustainability scores

The overall scoring information, while providing a useful overview, does not communicate the real complexity underlying many of the results. The overall scores are a composite of social, economic, institutional and environmental scores, and examining these separate scores provides further insights into how the overall scores are developed. Figure 5.3 plots the by-category sustainability scores, and the most important result to note is that a high overall score does not always come from uniformly high ratings across all sustainability categories. Scenario 1a, for example, comes in first on institutional performance, but scores far worse on economic and environmental grounds, particularly relative to Scenario 1b. Scenario 4, alternatively, shows moderate scoring but very consistent performance across sustainability categories.

Support for hypothesis 1 comes largely from the social, institutional and economic dimensions of sustainability, since scenario 1b does better on the environmental front. However, we shall see later how salt load was in fact much higher in scenario 1b than for scenario 1a (figure 5.7), which suggests we must be cautious in using composite measures like the overall sustainability score. Figure 5.3 adds extra support to hypothesis 2, since it shows that not only is scenario 2 a better performer overall, but it shows more consistent improvements across all categories. Similarly for hypothesis 3, not only does scenario 3 outperform scenario 2 overall, but shows more consistent benefits across categories. Much the same can be said for hypothesis 4. In the 5x scenarios, however, we frequently see far greater differences between categories: 5a, for example, has the best economic performance, but much poorer institutional performance. This provides much clearer support for hypothesis 5(i) than the overall scores provide, since
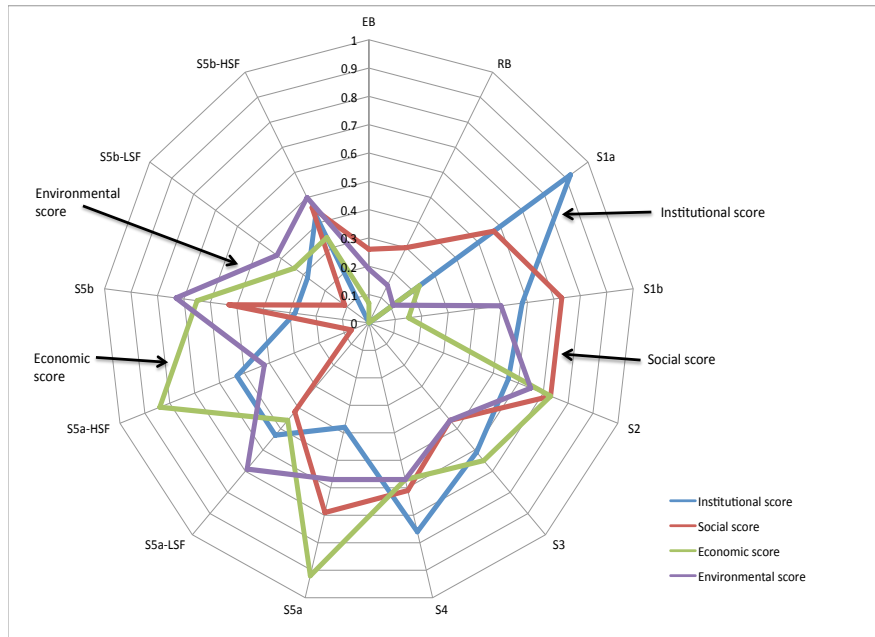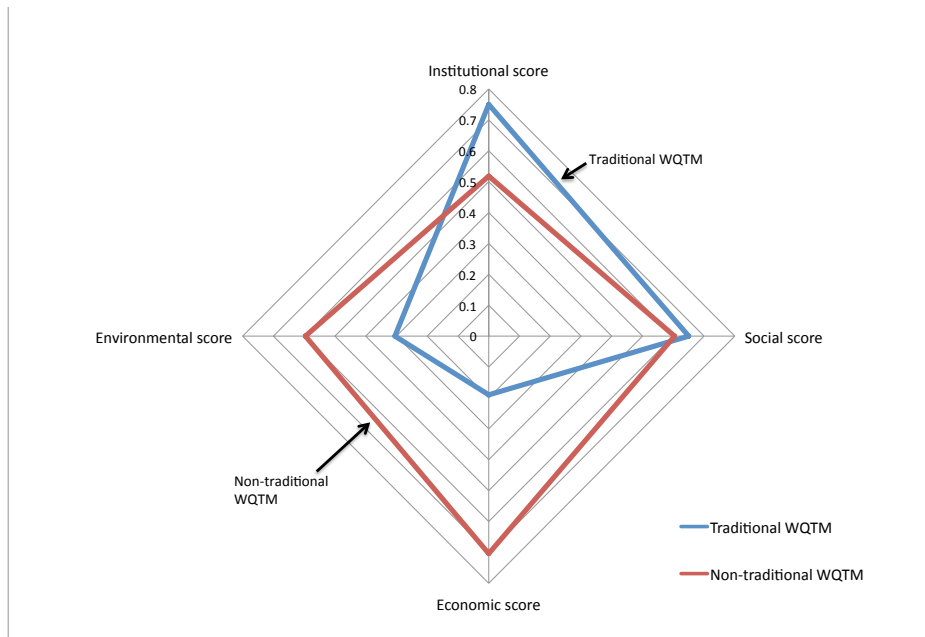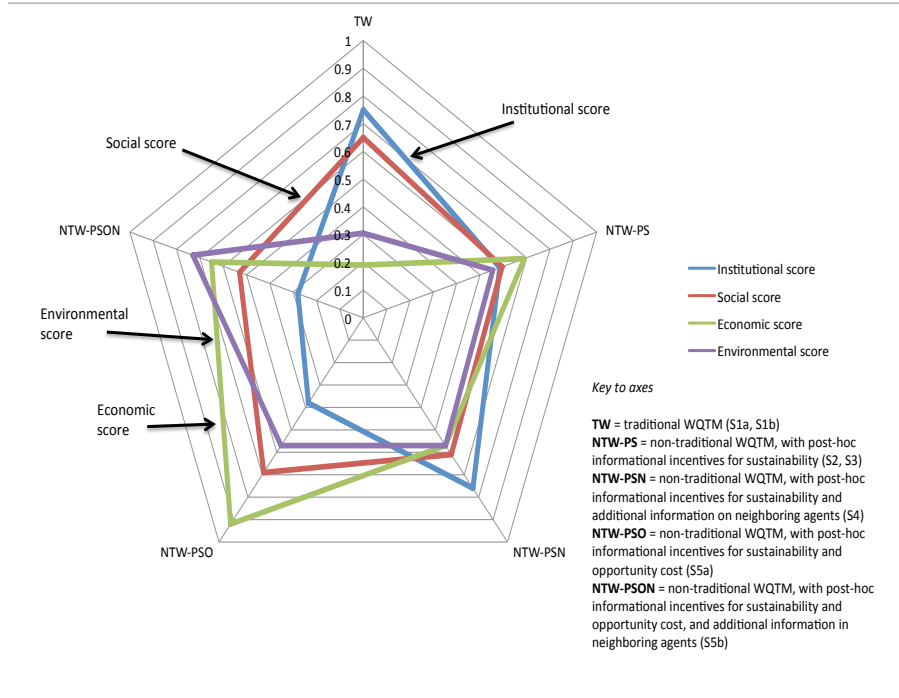
Figure 5.3: Radar plot of sustainability scores by category; the radial scale is the sustainability score, with each radial axis corresponding to a specific scenario. The plot shows that a high overall score does not always come from uniformly high ratings across all sustainability categories. See text for explanation of detail.

it indicates that the addition of opportunity cost information causes considerable disturbance in sustainability benefits relative to earlier scenarios, boosting some while reducing others. Finally, this plot underlines the overall finding that most scenarios do better in most categories than the baseline scenarios.

With regards to patterns not directly related to specific hypotheses, figure 5.4 summarizes scoring data by broad institutional category, traditional and non-traditional WQTMs. The "traditional" category includes scenarios 1a and 1b (this is the the type of WQTM typically seen in most real world air and water emissions trading), while the "non-traditional" category lumps together scenarios 2 through 5b-HSF. The plot shows that the traditional institutional form has benefits skewed towards institutional performance, while the non-traditional form has benefits skewed towards economic and environmental. This is an important result: it shows clearly how

127

making non-traditional assumptions (both institutional and behavioral) causes broad shifts to outcomes both overall and in terms of how sustainability benefits are distributed, relative to the sets of assumptions typically used in traditional social and economic studies.



Figure 5.4: Radar plot of sustainability scores by WQTM version; the radial scale is the sustainability score, with each radial axis corresponding to a specific scenario. The plot shows that the traditional institutional form has its benefits skewed towards institutional performance, while the non-traditional form has benefits skewed towards economic and environmental. See text for more detail.

Lumping together the scenarios in this way is still excessively coarse, particularly since scenarios 2 through 5b-HSF all incorporate rather different institutional assumptions. Figure 5.5 drills further into the institutional distinctions. The most important detail to note from this plot is that the most consistently beneficial institutional form is the NTW-PS, or non-traditional WQTM with sustainability information provided post-hoc. This is the simplest of the non-traditional scenarios, suggesting the important result that adding more institutional and behavioral complexity tends to make sustainability benefits less evenly distributed across social,

economic, environmental and social categories.



Figure 5.5: Radar plot of sustainability scores by WQTM specification; the radial scale is the sustainability score, with each radial axis corresponding to a specific scenario. The plot shows that the most consistently beneficial institutional form is the NTW-PS, or non-traditional WQTM with sustainability information provided post-hoc

### 5.1.3 Reviewing the response factors

Breaking the sustainability scoring out by category does not fully communicate real complexity in the underlying data. The following sections examine individual response factors to determine whether they continue to support the initial findings from the overall and unpacked scoring. I summarize the findings in figure 5.6.

| Response factor | Support for hypotheses | | | | | |
|---|---|---|---|---|---|---|
| | H1 | H2 | H3 | H4 | H5(i) | H5(ii) |
| Water quality | + | + | + | + | ? | + |
| Water quantity | - | - | - | - | +/- | +/- |
| Unique traders | ? | ? | ? | ? | ? | - |
| Clearing success ratio | + | + | + | ? | + | + |
| Volume of trades | + | ? | ? | ? | ? | +/- |
| Income | + | + | + | ? | +/- | +/- |
| Real assets | + | + | + | ? | ? | ? |
| Conflict | ? | ? | ? | ? | ? | ? |
| Deficit equity | ? | ? | ? | ? | ? | ? |

Figure 5.6: Tabulated results showing whether the results for a response factor support ("+"), refute ("-"), variable support or refute ("+/-") or do not provide a clear signal ("?") on each of the 5 hypotheses. More detail can be found in sections 5.1.4 through 5.1.12.

### 5.1.4 Water quality

Water quality is primarily described in ArkAgent using tons of salt contributed by all the land areas of the lower basin through surface and ground water flow to the Arkansas River. This metric will reflect the effects of canal diversions and irrigation practices in these areas, and is displayed in figure 5.7. While the standard deviations indicate some temporal range in the results, this is consistent across the scenarios. The range of results (around 100,000 tons of salt annually, approximately 8% of the mean value across all scenarios) is large in real terms, particularly in view of the fact that many real WQTMs have achieved little or no reductions in salt load. The best performing scenario, 5a-LSF, has a nearly 10% reduction in salt load over the empirical baseline. All differences between scenarios are "significant", since in the absence of stochastic variation these differences necessarily come from changes in parameterization. Recall from Chapter 4 that "more sustainable" outcomes were defined as lower salt loads, since reduced salt loads means higher crop yields, reduced legal conflict with Kansas, as well as

improved soil, riparian and aquatic health.
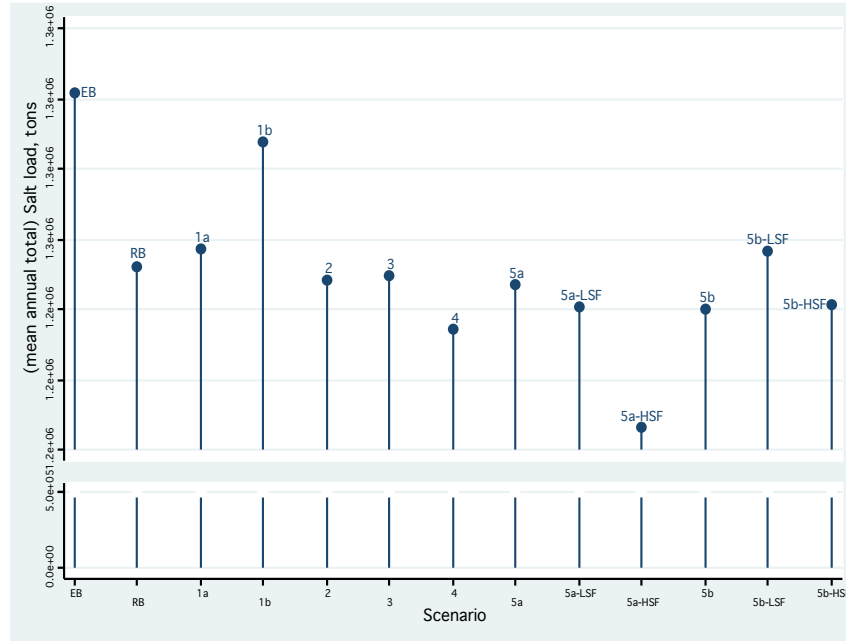


Figure 5.7: Mean sum annual salt load (tons) by scenario. The plot shows that hypothesis 1 is supported, with scenario 1b showing higher mean salt loads than scenario 1a. Note that y-axis has a scale break; focus here is on inter-scenario differences as opposed to range in real world context. Refer to Appendix section 8.1.2 for detailed scenario descriptions, and see text for more detail.

Figure 5.7 shows that hypothesis 1 is supported, since scenario 1b shows higher mean salt loads than scenario 1a. In other words, the traditional WQTM reduces salinity more when paired with an attitudinally homogeneous agent population (scenario 1a) than with an attitudinally heterogeneous agent population (scenario 1b). Hypothesis 2 is well supported, since scenario 2 shows reduced salt load relative to scenario 1b. This means that the use of post-hoc informational feedbacks incentivizes more environmentally sustainable agent behavior than can be achieved by the traditional set of WQTM incentives. Hypothesis 3 is similarly supported, since salt loads under scenario 3 are once again below scenario 1b. This means that even assuming the capacity for attitudinal change among agents, a non-traditional WQTM still works better in incentivizing environmentally sustainable behavior than the traditional institutional form. Figure 5.7 shows support for Hypothesis 4, since scenario 4 reduces salt load even more than the earlier scenarios. This suggests that it was effective for the WQTM in scenario 4 to exploit social dynamics for environmental sustainability. The data show weak support for

the primary assertion of hypothesis 5, since while there is change in the salt loads for 5a and 5b relative to scenario 4, this change is relatively subtle and mixed. This suggests that while the agent population does appear to be sensitive to the addition of opportunity cost information, it is not greatly perturbed. The secondary assertion of hypothesis 5 — contrary to the indications of the overall scoring — is supported by the salt load results: the HSF variants show reduced salt loads relative to the LSF variants. This indicates that the effect of adding opportunity cost information is sensitive to our assumptions on initial preferences among the agent population.

### 5.1.5 Water quantity

Water quantity is described in ArkAgent using the weekly volume of water flowing across the state line (acre-feet per week). This metric will reflect both the overall meteorological conditions (contributions to the river by rainfall and snowmelt) but also the consumptive use by farmers. I define consumptive use here as that portion of water diverted from the river that does not return to the river by surface or ground water flow. Recall from Chapter 4 that "more sustainable" outcomes were defined as higher flows, since more water in the river means improved riparian environmental conditions and more water for Kansas farmers, who have long been in conflict with Colorado agriculture over water deliveries.

Figure 5.8 indicates that the flow data does not support hypothesis 1, since flows are reduced in scenario 1a relative to scenario 1b. This suggests that homogeneously rational agents have higher consumptive use than more attitudinally and behaviorally heterogeneous agents, by some 50,000 acre-feet (significant in real terms). Similarly, hypothesis 2 is not supported, since the addition of post-hoc informational feedbacks in scenario 2 — while increasing flows relative to the rational homogeneous case in scenario 1a, and relative to the empirical baseline — results in lower flows than scenario 1b. This indicates that adding an informational feedback mechanism to a WQTM operating with a heterogeneous population can actually degrade WQTM performance. Hypothesis 3 is unsupported, since S3 does not show improved flows relative to S1, although little change relative to S2. This suggests that allowing for attitudinal change within a population does not greatly impact (slightly worsens) WQTM performance, and the traditional WQTM still performs better. The data continue the pattern with hypoth-
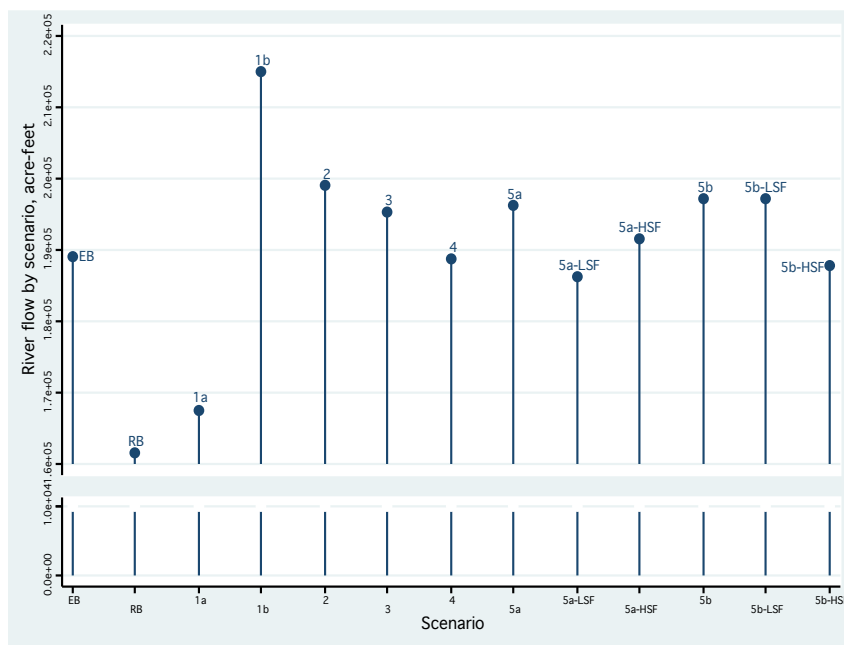
Figure 5.8: Mean sum annual flow (acre-feet) by scenario. Indicates that flow data is not supportive of hypothesis 1. Note that y-axis has a scale break; focus here is on inter-scenario differences as opposed to range in real world context. See text for more detail and refer to Appendix section 8.1.2 for detailed scenario descriptions.

esis 4, also unsupported: S4 shows lower flows than scenario 3. Hypothesis 5(i) is weakly supported, as with the quality data, since both scenarios 5a and 5b show changes — albeit relatively small — in flow relative to S4. The LSF/HSF variants in S5a show support for hypothesis 5(ii), however, since parameterizing agents as more pro-social does result in higher flows relative to the pro-self agents. This is contradicted by the LSF/HSF variants for S5b, however, since the pro-social agents result in lower flows relative to the pro-self agents. This means that if we just assume that agents have access to opportunity cost information but no interaction with their neighbors, overall outcomes are dependent on whether we assumed the agents to be initially pro-self or pro-social, such that pro-social assumptions lead to improved river flows. But if we also assume social dynamics, the relationship is reversed: initially pro-social agents result in lower flows.

133
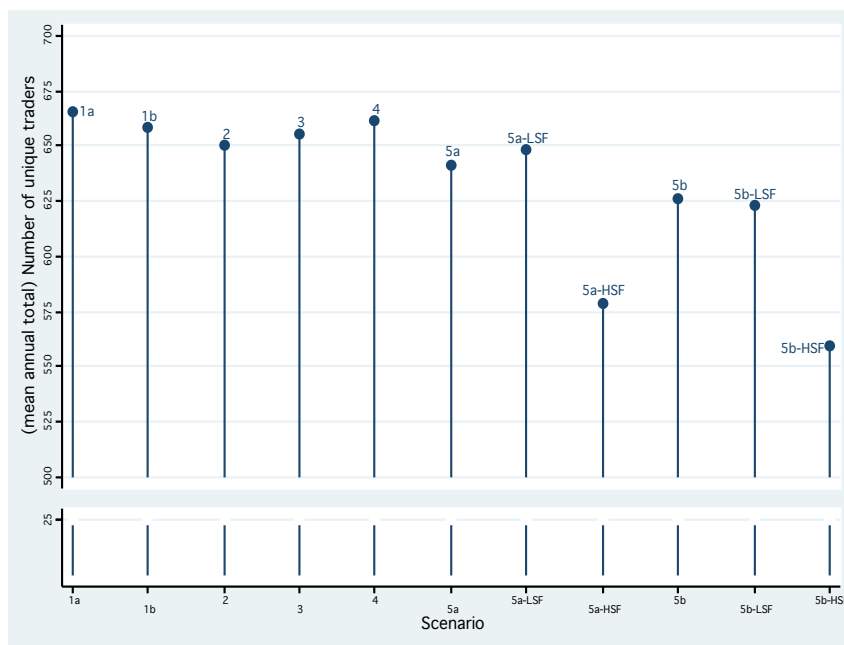
Figure 5.9: Mean sum annual unique trader numbers by scenario. The only finding of note is that the results go against hypothesis 5(ii), since parameterizing agents with a higher S-factor results in degraded institutional performance. Note that y-axis has a scale break; focus here is on inter-scenario differences as opposed to range in real world context. See text for more detail and refer to Appendix section 8.1.2 for detailed scenario descriptions.

### 5.1.6  Unique traders

The "unique traders" response factor was defined as the number of different agents participating one or more times in the WQTM in any given year. I defined a better institutional performance as a higher number of unique traders, since real world WQTMs have traditionally suffered from low participation problems. If more participants can be attracted to the market, in theory this improves the chances that there will be sufficient trading activity to better incentivize pollution reduction. Figure 5.9 indicates that there is mostly very limited variation among scenarios, with typical differences varying between less than 1% to 5% of the total agent population. The exceptions are the S5a and S5b LSF/HSF variants, with differences around 15%. These exceptions contradict the assertion in hypothesis 5(ii), which suggested that agents parameterized with a higher S-factor would result in improved outcomes. In the results, however, both HSF variants show lower trader numbers than the LSF runs. This means that parameterizing agents as more pro-social degrades institutional performance. The general

134

lack of difference between scenarios, particularly among scenarios 1a through 4, suggests that this basic measure of WQTM participation is not greatly influenced by behavioral assumptions or institutional mechanisms.

### 5.1.7 Clearing success ratio

The clearing success ratio (CSR) is the ratio between the number of trade offers introduced to the market, and the number of trades cleared from the market. All WQTMs in ArkAgent utilize a clearing house mechanism, where new offers are transparent to all potential buyers, so the CSR is one measure of market efficiency. A perfect ratio of 1 indicates that the market is clearing efficiently; excessively high ratios indicate that the market is flooded with surplus quota; and very low ratios indicate that not enough willing sellers can be found, and so the market will not effectively incentivize redistribution of pollution reduction costs. Figure 5.10 indicates that hypothesis 1 is well supported: the homogeneously rational agent population coupled with the traditional WQTM has a far better clearing success ratio (CSR) than the heterogeneously rational population with the same institution.
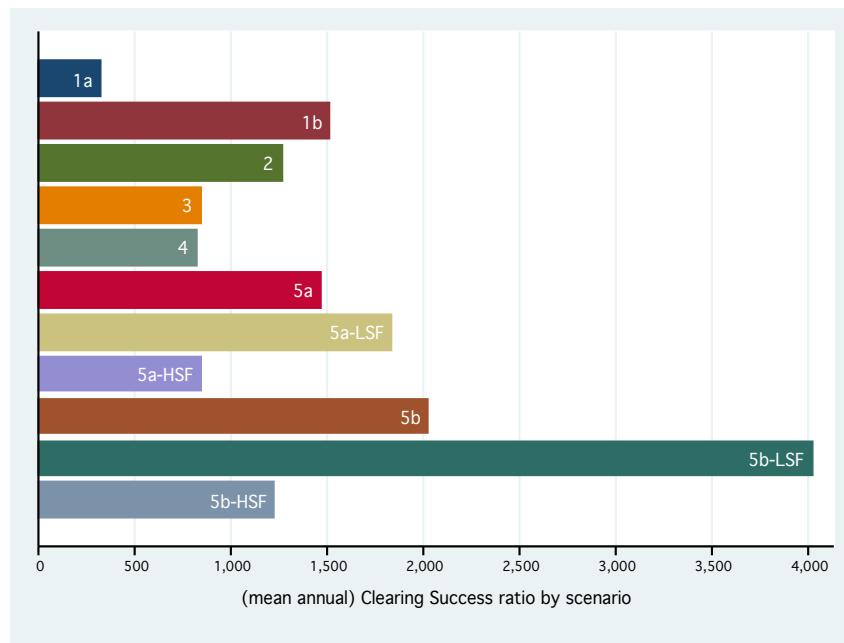


Figure 5.10: Mean clearing success ratio by scenario. The data are supportive of hypothesis 1, since the clearing success ratios drop in scenario 1a relative to S2. See text for more detail and refer to Appendix section 8.1.2 for detailed scenario descriptions.

These results mean that agents in S1a are more responsive to scarcity and pricing signals with the traditional WQTM than the agents in S1b, and the market is more efficient as a result. The results show weak support for hypothesis 2, since the CSR improves in scenario 2 relative to 1b. However, the difference is not large and the CSR is still very high in both scenarios, indicating that the addition of post-hoc informational feedbacks does improve market efficiency but not to any great degree. Hypothesis 3 is also weakly supported, with the mean annual CSR improving still further over both S1b and S2. This indicates that the assumption of a capacity for attitudinal change among agents can result in better price and scarcity signaling and better overall market performance. Hypothesis 4 shows minimal difference in CSR from scenario 3, indicating that the additional of an institutional mechanism to exploit social dynamics does not impact institutional performance. Finally, hypothesis 5(i) is strongly supported, since the CSR worsens in S5a and S5b relative to to S4. This indicates that opportunity cost information has the effect of worsening institutional performance, an effect particularly marked when we also allow for social dynamics (scenario 5b). Hypothesis 5(ii) is strongly supported by the data, since outcomes are sensitive to whether or not the agents were initially parameterized with low (LSF) or high (HSF) S-factors: the higher the S-factor, the better the institutional performance.

### 5.1.8 Volume of trades

Figure 5.11 indicates that hypothesis 1 is strongly supported: S1a shows higher trade volumes than any other scenario, indicating that the pairing of homogeneously rational actors and a traditional WQTM resulted in what we might expect from neoclassical economic theory. Recall that the "volume of trades" response factor, in kg of quota traded, communicates the level of liquidity in the market. Better institutional performance (and more sustainability) is associated with higher levels of liquidity.

The drop off from S1a to S1b (nearly 100,000 tons less in mean trading activity) is significant relative to the range of activity seen among all scenarios (slightly more than 100,000 tons), and indicates that heterogeneous agent populations paired with a traditional WQTM show a much lower level of liquidity. Importantly, however, even in S1b there is still significant liquidity, as there is in all scenarios. This suggests that regardless of behavioral assumptions

136
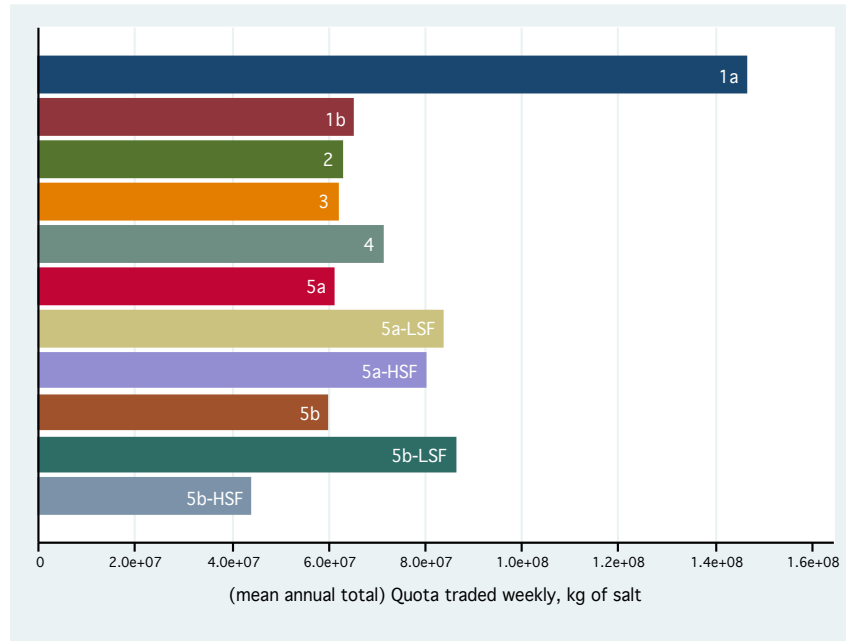
Figure 5.11: Mean sum annual quota trade volumes (kg) by scenario. The results support hypothesis 1, but are less clear on hypotheses 2–5(i). See text for more detail and refer to Appendix section 8.1.2 for detailed scenario descriptions.

and institutional mechanisms, the WQTM institution in the LAB can achieve levels of liquidity exceeding most real world WQTMs. While the results for hypothesis 1 are clear, hypotheses 2, 3, 4 and 5(i) are not as clearly supported (or refuted). There is very little difference in quota volume between scenarios 1b, 2, 3, 4, 5a and 5b. This is not to say that the various behavioral and institutional differences in these scenarios cause no change at all. For example, hypothesis 4 is weakly supported, since S4 shows a few thousand extra tons of trading annually than S3. This indicates that an institution exploiting neighborhood dynamics can increase liquidity by a few percent. But this is still only a few percent, and the changes are less substantial for other hypotheses. Hypothesis 5(ii) is not clearly supported or refuted by the S5a variants, but the S5b LSF and HSF variants show that changing initial attitudinal parameterization can affect liquidity outcomes by nearly 50,000 tons. The results for S5b HSF show that parameterizing agents with higher S-factors can drive down liquidity, but whether or not the parameterization causes this affect depends on whether social dynamics are included. S5a HSF, without social dynamics, shows only a slight drop in liquidity.

### 5.1.9   Income

Figure 5.12 indicates weak support for hypothesis 1, since incomes are higher in S1a relative to S1b. Agent income is a measure of weekly farmer and ditch manager incomes (revenues less expenditures) within the basin. Higher incomes indicate more economically sustainable outcomes, but there is some complexity in this measure since higher mean income does not necessarily indicate equity in the way additional income is distributed, and increased borrowing can also drive up mean incomes.
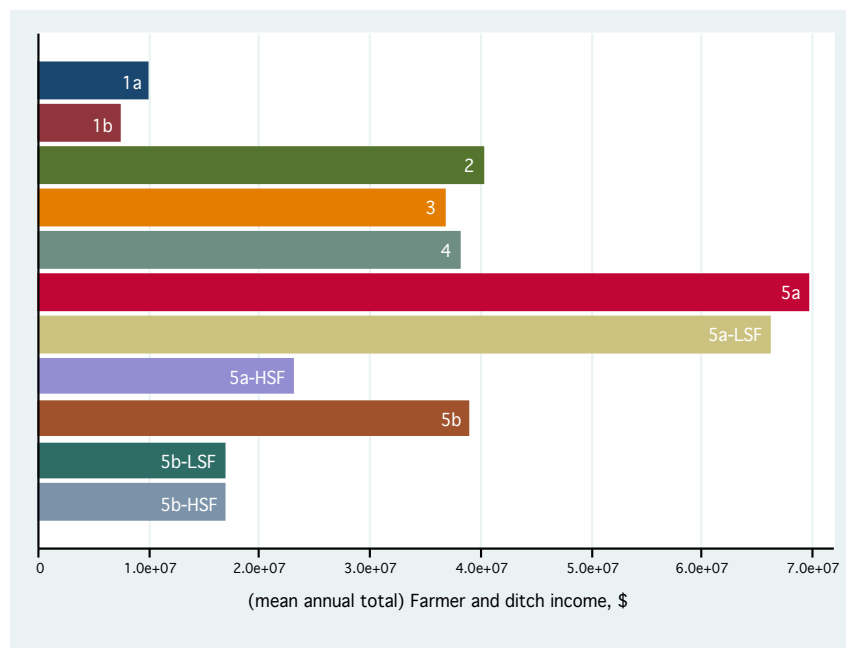


Figure 5.12: Mean sum annual agent incomes by scenario. The data show weak support for hypothesis 1, since incomes are higher in S1a relative to S1b. There is stronger support for hypothesis 2, since incomes are higher in S2 relative to 1b; the same is true of hypothesis 3, while hypotheses 4-5(ii) are less clearly supported. See text for more detail and refer to Appendix section 8.1.2 for detailed scenario descriptions.

Results indicate that the combination of homogeneously rational agents and a traditional WQTM increases incomes relative to the same institution with heterogeneous agents. But we should note that the difference in income is not large relative to the other scenarios and the overall magnitudes. The data provide much stronger support for hypothesis 2, since we see at least a $20 million boost to mean annual incomes in S2 relative to S1b. This indicates that the addition of post-hoc informational incentives is very effective at increasing agent incomes

138

relative to the traditional WQTM. At this point it is not clear what the source of this additional income might be. The data similarly clear in supporting hypothesis 3, since scenario 3 shows clear differences to scenario 1b: the positive effect of post-hoc informational incentives look to be robust even to assumptions of agent capacity for attitudinal change. The data do not, however, provide clear support or refutation for hypothesis 4: scenario 4 shows very little difference from scenario 3, suggesting that the exploitation of social dynamics does not greatly impact agent incomes. Scenario 5a, with the highest mean annual income out of all scenarios, shows very strong support for hypothesis 5(i). S5b does not, however, with very little change relative to scenario 2, 3 or 4. This indicates that while providing opportunity cost information alone appears to substantially increase agent incomes, allowing social dynamics to interfere with this information can wipe out the increase. Similarly, the reduced income in S5a-HSF relative to S5a-LSF indicates that hypothesis 5(ii) is well supported: changing our initial assumptions on agent preferences can have effects on overall outcomes, and that higher S-factors lead to lower agent incomes. But the scenario 5b variants are insensitive to the initial parameterization, indicating that social dynamics may trump assumptions on initial preferences in determining economic sustainability.

### 5.1.10   Real assets

Real assets results (figure 5.13) show basically the same relationships between scenarios (and the same level of support for the various hypotheses) that the results for income in section 5.1.9. An agent's "real assets" are the current cash assets of the agent minus any liabilities that agent was carrying. Higher real asset levels are indicative of more economic sustainability, since they imply more cash resources and less debt.

Real assets are a more reliable indicator of overall economic health since income does not factor in liabilities; it is reassuring that the assets data provide the same overall picture. The data also indicate that agent income has a very strong influence on overall assets, and that agents are not incurring significant liabilities. This is an important result, since it indicates that most of the non-traditional WQTMs (scenarios 2 and above) provide more economic benefit in general terms than the traditional scenarios (scenarios 1a and 1b). The data also show that even
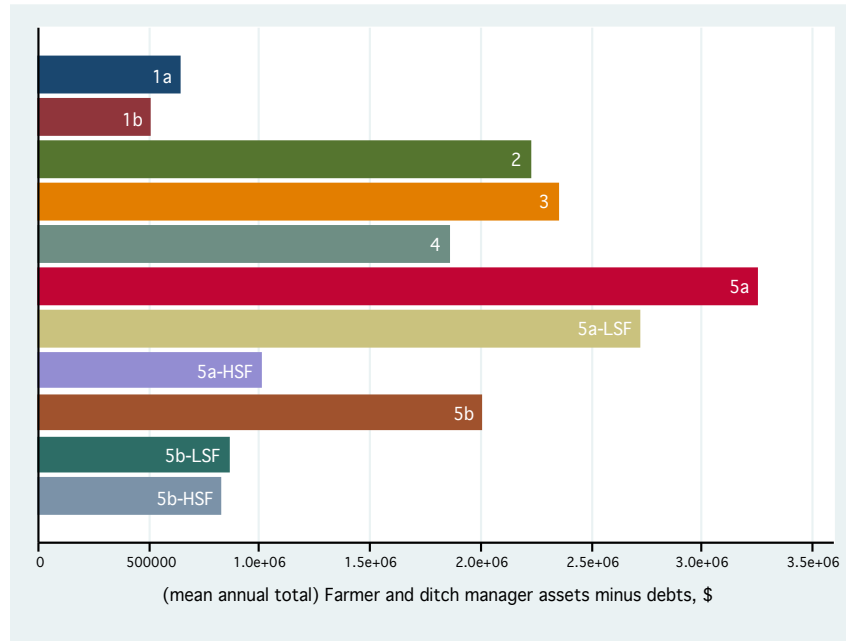
Figure 5.13: Real assets by scenario (mean sum annual agent assets less liabilities), $. Note that the bars are 1 standard deviation, and so communicate variance with time and not error. The plot shows that assets data follows income data closely, implying that debt was not a significant factor in affecting agent asset totals. See text for more detail and refer to Appendix section 8.1.2 for detailed scenario descriptions.

the traditional WQTMs provide more economic benefit than the baseline case.

### 5.1.11 Conflict

The principal result from the conflict data in figure 5.14 is that the social system is relatively conflicted. "Conflict" is a composite of various measures tracking individual agent satisfaction, designed to act as a proxy for the likelihood that agents would precipitate legal conflict over issues related to water use and management in the basin. Higher levels of conflict (lower levels of personal satisfaction) are not constructive in the context of ArkAgent, since legal action would be likely to damage the short term sustainability of any efforts to address salinity through a WQTM institution.

All scenarios show at least a 0.7 out of 1.0 in conflict level, indicating substantial dissatisfaction with the current way water is used and managed in the system. The second main result is that the variation between scenarios is minimal: the range is a little over 2%. The lack of variation is in itself important, since it indicates that while different behavioral and institutional
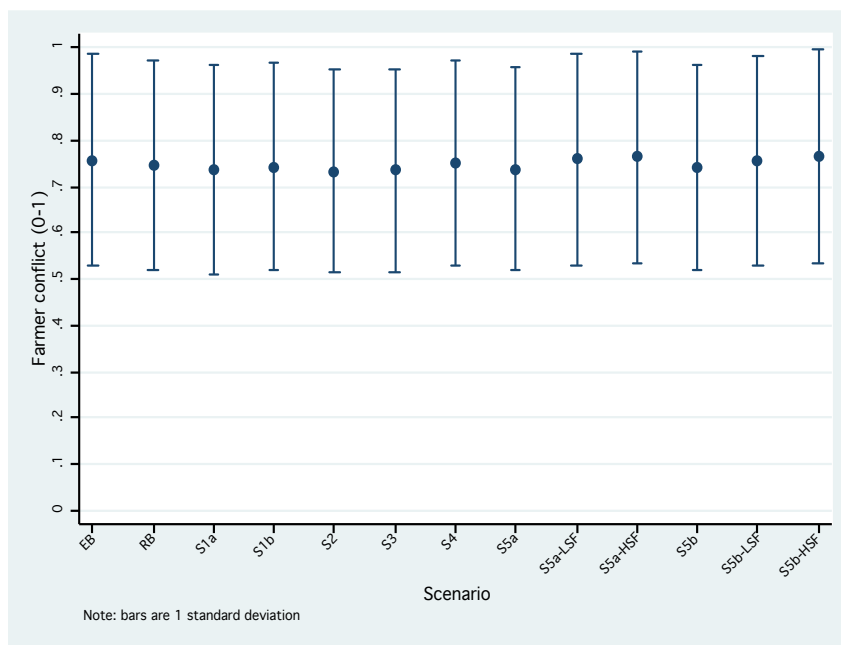
Figure 5.14: Mean conflict level (0.0-1.0) by scenario. Note that the bars are 1 standard deviation, and so communicate variance with time and not error. The plot shows that the system is overall relatively conflicted, but that conflict in the system is minimally impacted by changes in behavioral and institutional assumptions. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

assumptions do cause some shift in conflict within the system, with no set of assumptions does the system reach dramatically higher (or lower) levels of conflict.

### 5.1.12 Deficit equity

The results for deficit equity (figure 5.15) indicate that, as with conflict, levels do not change much between scenarios or between any scenario and the baseline level. (Water) Deficit equity is a measure of fairness in the distribution of costs in the basin: the most important potential cost in the system is a reduction in water availability, since without water very little is possible in the LAB's semi-arid environment. If this cost is evenly distributed across the system, then while the outcome is not ideal it may be more socially palatable. The lack of variation in this result is again important to note, since it indicates that while none of the institutional mechanisms improve on the baseline distribution of water, none of the mechanisms make this any worse.
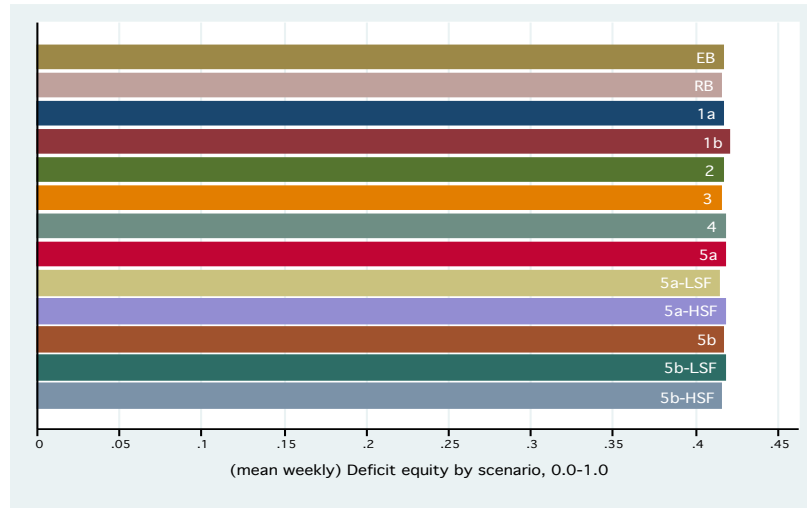
Figure 5.15: Mean deficit equity by scenario. Deficit equity barely shifts between scenarios. Indicates that none of the institutional mechanisms greatly increases or decreases the level of equity in the way water quantity is distributed in the basin. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

## 5.2   Synthesis of hypotheses

The analysis of sustainability scores and response factors showed support for the 5 main hypotheses, but also plenty of unclear signals and mixed responses. In the following section, I synthesize all the various response factors into overall findings relating to each hypothesis; and go into more detail exploring the raw data where necessary. I find that — despite some mixed and unclear results — all hypotheses are weakly to strongly supported by the sustainability scoring and the response factors. Social data (conflict and deficit equity) show very little variation among scenarios, and omitting these from the overall scoring calculations strengthens initial findings. Results from sensitivity testing experiments (where key parameters were varied to extremes) show that all hypotheses are broadly robust to such variation, with notable exceptions being that model results tend to be sensitive to the risk strategy pursued by agents, the initial market price of WQTM quota, and to a lesser extent, the length of the WQTM program. This robustness finding increases the credibility of the model in supporting the hypothesis testing, since it supports my assertions that the model is structurally sound.

### 5.2.1 Synthesis of Hypothesis 1

#### 5.2.1.1 Summary of Hypothesis 1 synthesis

In the following section I analyze the data relevant to Hypothesis 1 and find the following: first, the scoring and raw data support hypothesis 1, in that a heterogeneous behavioral model shows worse sustainability outcomes than a homogeneous behavioral model, when both models are run against a traditional WQTM. The exception to this support is the river flow, which is reduced in scenario 1a relative to 1b. Second, support for hypothesis 1 is stronger when the social metrics (which display little variation among scenarios in any case) are omitted. Third, the homogeneous behavioral model in combination with the traditional WQTM (run in S1a) is not a perfect match, since S1a shows levels of institutional performance far from the pareto-optimal predicated by neoclassical economic theory. Fourth, while parameter variation experiments (setting select parameters to extreme values) show that magnitudes of results change, the overall pattern supportive of hypothesis 1 is broadly robust to such variation. This is clearest for economic and environmental data, and less so for social and institutional data.

#### 5.2.1.2 Detailed synthesis of Hypothesis 1 results

Recalling the hypothesis: *The assumption of universal self-interested, maximizing behavior among resource users is incorrect. A more realistic assumption of behavioral heterogeneity among resource users can help explain why the theoretical benefits of water quality trading markets have not, in general, materialized.*

In hypothesis 1 I assume that "theoretical benefits" be the objective of the market institution in question, namely, a reduction in salt load over time. Testing the hypothesis requires asking the questions (1) "does a heterogeneous behavioral model show worse sustainability performance than a homogeneous behavioral model for the same institution?", (2) "does the homogeneous case show perfect performance, as theoretically predicted?" and finally (3) "is any difference in performance robust to varying other parameters in the model?".

With regard to (1), "heterogeneous behavioral model" refers to the behavioral model used in scenario 1b, and "homogeneous behavioral model" refers to the behavioral model used in

scenario 1a. The overall scoring (figure 5.1) indicates that a heterogeneous model does indeed show worse sustainability performance than a homogeneous behavioral model. But the scoring difference is only slight, so we need to look more deeply into the results. Scoring by individual sustainability categories (see Appendix figures 8.4, 8.7, 8.6 and 8.5 for detail) indicates that under institutional and economic categories, scenario 1a comes out ahead, while under economic and social categories, scenario 1b scores higher. Noting the lack of variation among scenarios for the social response factors (conflict and deficit equity), figure 5.16 shows the overall scoring with the social data omitted. This figure shows that the overall support for hypothesis 1 is strengthened. This result highlights that how we measure the performance of different scenar-
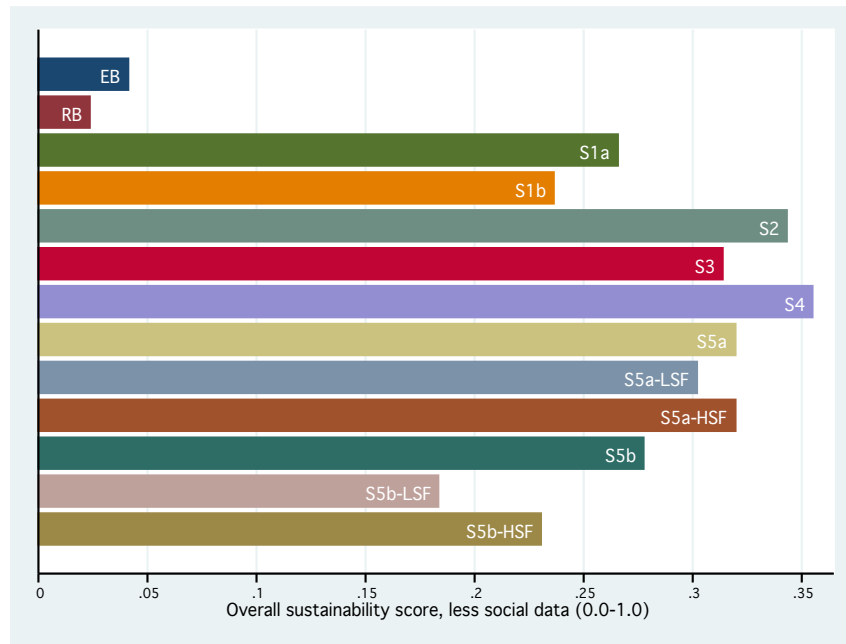


Figure 5.16: Overall sustainability scores computed without social information. The plot indicates that hypothesis 1 is still broadly supported if we omit the social data due to their lack of variation among the scenarios. See text for more detail and refer to Appendix section 8.1.2 for detailed scenario descriptions.

ios — in this model or in any scenario planning context — can be an important determinant of the performance that results. I discuss this issue in more depth in Chapter 6, but suffice to say there is no perfect set of performance metrics. Given the unknown importance of such social metrics (and, indeed, the broad category of social sustainability) to basin stakeholders, I try to avoid removing this component in this and other analyses. Where necessary, however, I omit

the social data to clarify otherwise complex results.

Examining the raw data, figure 5.7 indicates that a heterogeneous behavioral model shows worse performance in terms of salt loading than the homogeneous model. Reducing overall salt load with time is, after all, the primary objective of a water quality trading market. Conversely, however, the flow data (figure 5.8) shows the opposite relation: S1b has a higher flow than S1a. On the institutional side, scenario 1a shows better market performance on all main metrics: a clearing success rate closer to 1, the highest number of unique traders, and the highest amount of quota traded (see Figures 5.10, 5.9 and 5.11). In economic terms, agents in scenario 1a have higher incomes than in scenario 1b (figure 5.12) as well as higher levels of assets (figure 5.13). While there is some debt incursion (figure 5.17), the levels in S1a are still less than the empirical baseline. With regard to question 2 posed above, scenario 1a does not display perfect institu-
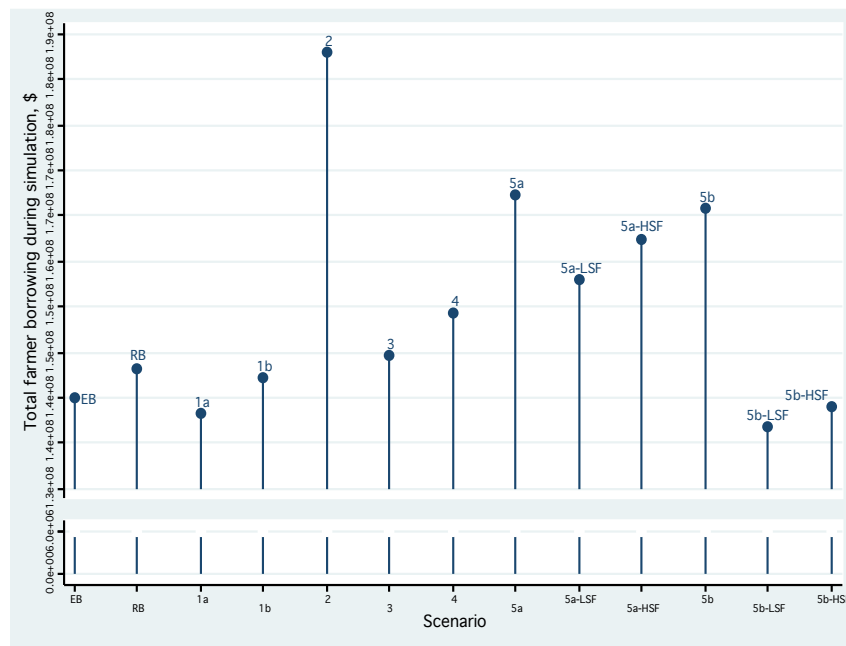


Figure 5.17: Sum of farmer financial liabilities accumulated during the simulation period, ($) across all scenarios. The plot shows that despite returning considerable sustainability benefits overall, scenario 1a does not display perfect institutional properties. Nevertheless, we still see less debt incursion than the baseline, supporting hypothesis 1. Note that y axis includes scale break; this is to make variations between scenarios clearer, since we are less interested in absolute values. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

tional properties, as demonstrated in Figure 5.10. Scenario 1a is clearly the best performing among all scenarios with only around 25 times as much quota being posted to the market as

145

actually sold. But the ideal ratio is 1.0, where all quota is being cleared from the market by willing buyers. The market in S1a is clearly not pareto-optimal.

Answering the third question requires varying model parameters beyond the scope of the experimental setup for hypothesis testing. Recall from section 4.3 in Chapter 4 how this was done: in the form of experiments varying each of a number of parameters across a set of binary values (I refer to individual parameter variation experiments as "variants"). A sampling of relevant parameters for the present question includes: clearing success ratio; salt load; real assets; and conflict.
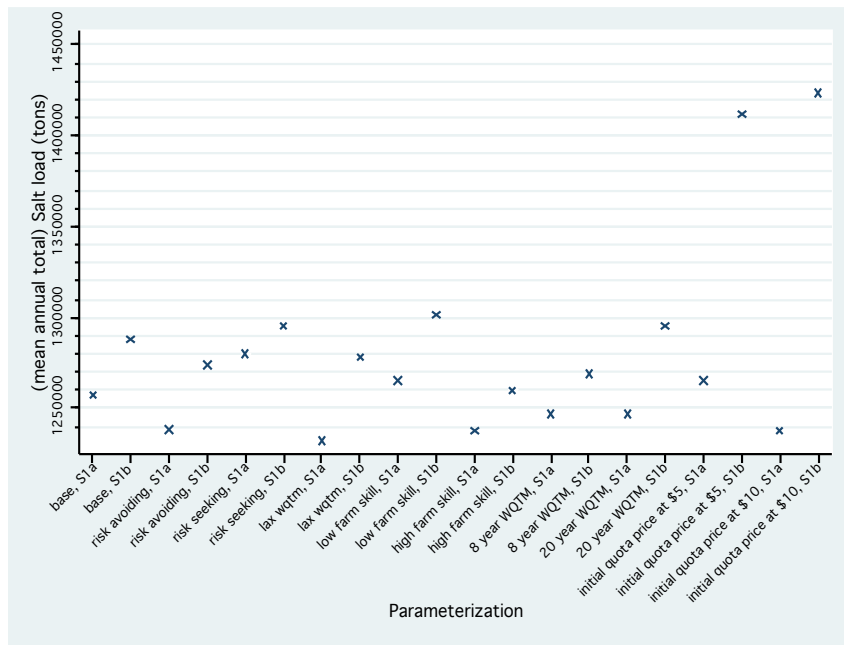


Figure 5.18: Salt load contribution by basin irrigation regions (tons) in scenarios 1a and 1b across all parameter variation experiments. Note that "lax wqtm" refers to an alternate set of institutional rules allowing both quota banking (carrying excess quota over from one year to another) and deficit cancelation (canceling any deficit remaining after a year of trading). Neither are allowed in the WQTM variation simulated in the normal scenarios. This and other parameter variation experiments are described in chapter 4 or in section 8.1.3 in the Appendix. The plot shows that hypothesis 1 is broadly robust to parameter variation, although absolute magnitudes of salt load do change with parameterization. Standard deviations are omitted, since these are large and reflect macroscale controls on salt loading rather than inter-scenario variation, which is the focus of the present analysis. See text for more detail and refer to Appendix section 8.1.2 for detailed scenario descriptions.

Figure 5.18 shows that the results are not robust to parameter variation, if by robust we mean "does not vary outside of original range". However, the basic comparison between 1a and 1b - lower salt in 1a than 1b - is seen under *all* of the parameter variation experiments.

This indicates that even with considerable variation in some of the more subjective behavioral and WQTM parameters, the model still supports the general hypothesis: a neoclassical WQTM performs far worse when we assume heterogeneous behavioral models in the participating agents. Figure 5.19 shows that varying the initial quota price (IQP) has the effect of increasing
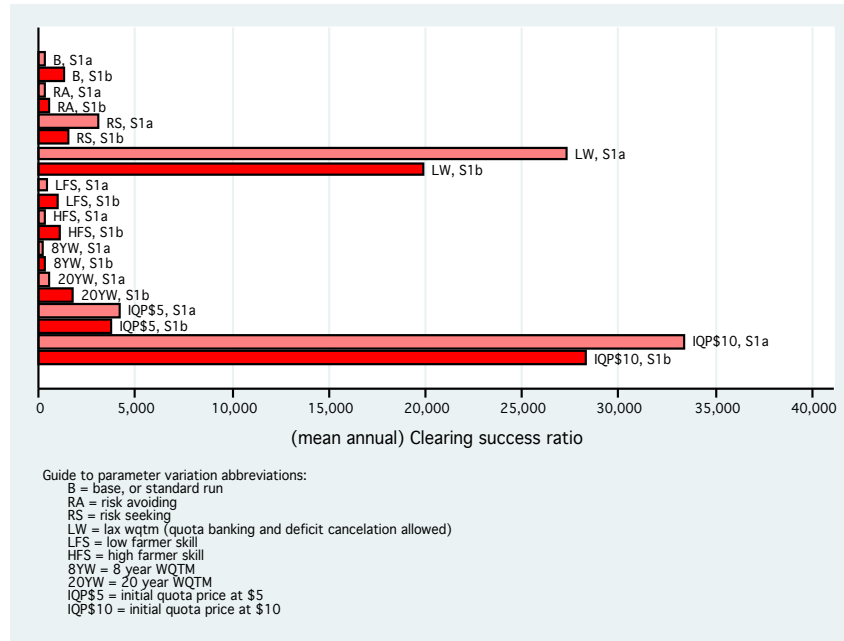


Figure 5.19: WQTM clearing success ratio (CSR) in scenarios 1a and 1b across all parameter variation experiments. Plot shows that varying the initial quota price affects the absolute value of the CSR far more than other variants. See text for more detail and refer to Appendix section 8.1.2 for detailed scenario descriptions.

the ratio many orders of magnitude beyond the effects of most other parameters. "Lax WQTM" is a modification of WQTM institutional controls that allows agents to bank quota from year to year, and cancels any deficits not cleared after one year. The lax WQTM variant has a similar effect to IQP. With both variants, however, there are still improvements from scenario 1a to 1b. Re-plotting without the initial quota price or lax WQTM variants (figure 5.20), there is still broad support for hypothesis 1, except in the case where agents are forced to be risk seeking. In the variant, the ratio improves from S1a to S1b, contra hypothesis 1.

These data indicate that when agents are forced to be risk seeking in their decision making, hypothesis 1 does not hold in relation to institutional performance. In all other cases, however, hypothesis 1 is robust (again, in relation to institutional performance). It is notable that increas-
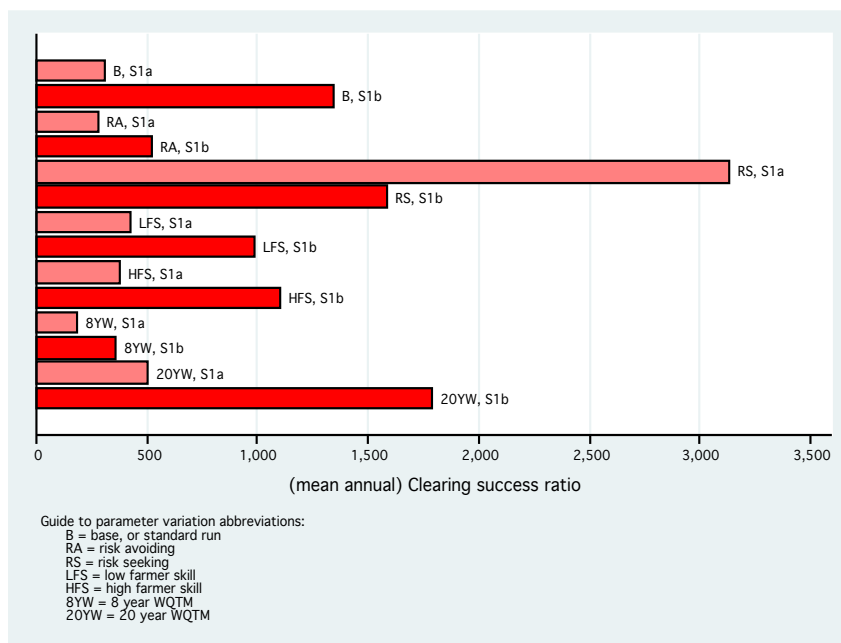
Figure 5.20: WQTM clearing success ratio (CSR) in scenarios 1a and 1b across all parameter variation experiments, omitting variants related to initial quota price and WQTM management options. Plot indicates that, when we omit initial quota price and WQTM management option variants, most parameter variation experiments show support for hypothesis 1, with the notable exception of agent risk strategy. See text for more detail and refer to Appendix section 8.1.2 for detailed scenario descriptions.

ing the initial quota price has the effect of greatly increasing the magnitude of benefits, while not changing the overall pattern between scenarios.

Looking at all the assets data together (figure 5.21), we see that the IQP variants show assets far greater than all other scenarios, and mixed support for Hypothesis 1. Removing the IQP variant data (figure 5.22), and we see strong support for Hypothesis 1 across all variants, in the form of reduced assets from S1a to S1b.

Finally, conflict data (figure 5.23) show that there is very little change in results between variants. The most important result is that raising the initial quota price has the effect of greatly lowering conflict in the system.
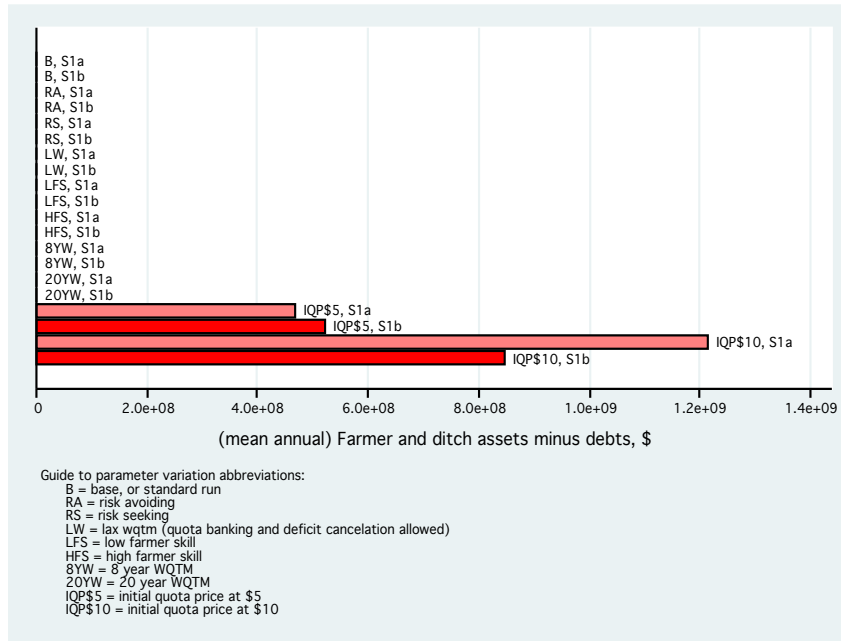
Figure 5.21: Mean annual farmer and ditch manager financial assets ($) in scenarios 1a and 1b across all parameter variation experiments. Note that the bars are 1 standard deviation, and so communicate variance with time and not error. See text for more detail and refer to Appendix section 8.1.2 for detailed scenario descriptions.
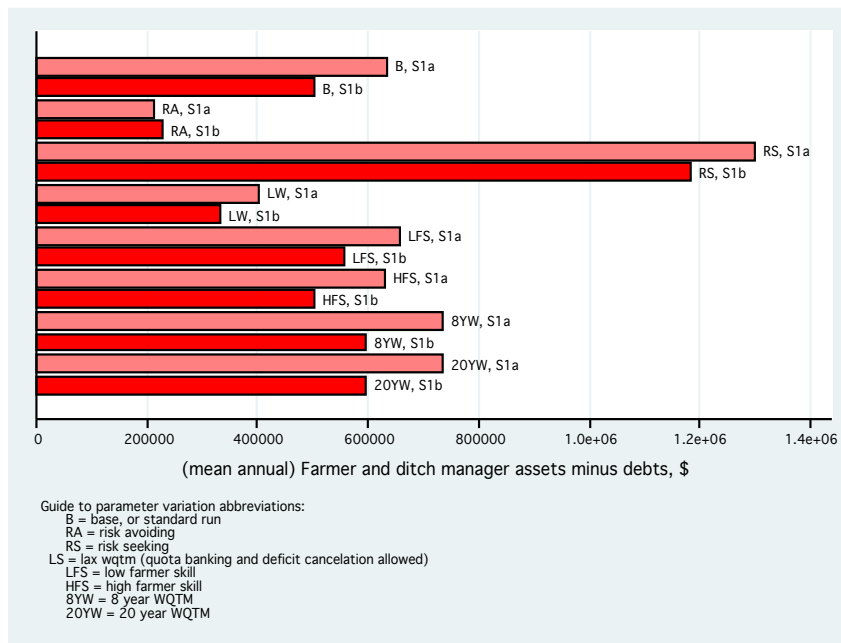


Figure 5.22: Mean annual farmer and ditch manager financial assets in scenarios 1a and 1b across all parameter variation experiments, omitting initial quota price variant data. See text for more detail and refer to Appendix section 8.1.2 for detailed scenario descriptions.
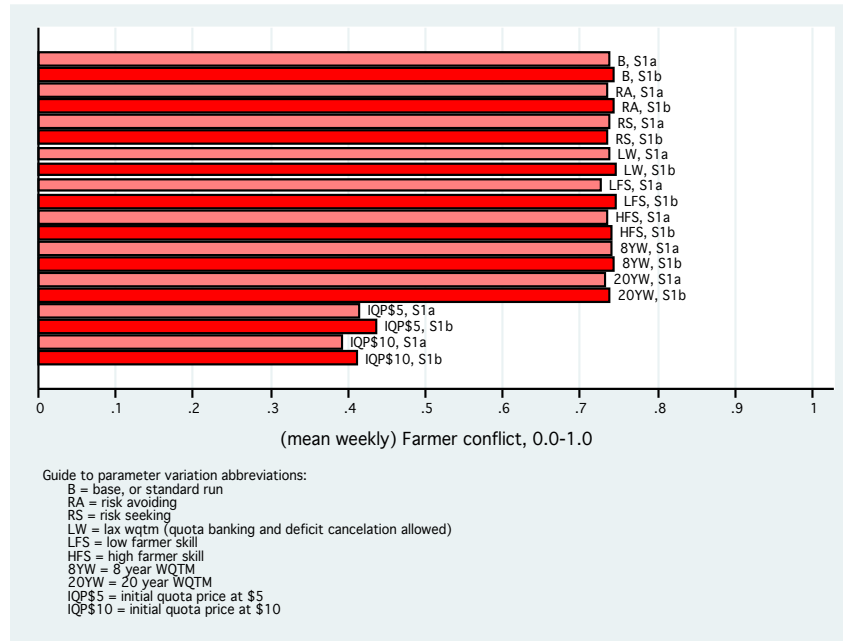
Figure 5.23: Farmer conflict levels in scenarios 1a and 1b across all parameter variation experiments. See text for more detail and refer to Appendix section 8.1.2 for detailed scenario descriptions.

### 5.2.2 Synthesis of Hypothesis 2

#### 5.2.2.1 Summary of Hypothesis 2 synthesis

The principal findings of the synthesis are as follows: first, that the strongest support for hypothesis 2 comes from salt loads, market clearing ratios and agent economic data. There are more mixed results for the other response factors; notably, hypothesis 2 is refuted by the data for river flow. Second, that this finding is robust to parameter variation, except for certain institutional metrics like the length of the WQTM program, and for the risk strategy of agents.

#### 5.2.2.2 Detailed synthesis of Hypothesis 2 results

Recalling the hypothesis: *A pro-social WQTM is a trading institution that incorporates informational incentives, providing feedback to WQTM participants on the sustainability benefits of their WQTM actions. This package of incentives is better matched to a broader spectrum of self- and socially interested resource users, and so will be more robust to behavioral heterogeneity than a neoclassical, "pro-self" WQTM.*

The particular area of interest here is the comparison between scenarios 1b (heterogeneous

150

agents, neoclassical WQTM), and scenario 2 (heterogeneous agents, WQTM providing informational feedback on sustainability benefits). Reviewing the overall sustainability scoring shown in figure 5.1, the data show scenario 2 to be substantially higher scoring than scenario 1b. Looking at the data for institutional scores by category (see Appendix figures 8.4, 8.7, 8.6 and 8.5 for detail), most of the support for hypothesis 1 comes from the economic and environmental scores; on institutional grounds alone, scenario 1b scores slightly higher than scenario 2; on social grounds, scenario 1b is at parity with scenario 2. Noting from hypothesis 1 the dampening effect of the low-variation social data, reviewing the scoring with social data omitted (figure 5.16) indicates that overall support for hypothesis 2 is strengthened.

In terms of environmental sustainability, Figure 5.7 indicates that scenario 2 reduced salt loads better than scenario 1b. However, as we saw with hypothesis 1, figure 5.8 shows that river flows were higher in scenario 1b than 2. On the social side, figure 5.14 indicates that scenario 2 does see slightly less conflict, and (Appendix) figure 5.15 indicates that scenario 2 sees slightly lower deficit equity. But the overall lack of variation among scenario responses for conflict and deficit equity means it is difficult to identify any particular findings in this area. The best we can say is that the scenarios performed at equal levels in terms of social sustainability. Regarding institutional sustainability, figures 5.11 and 5.9 show that S2 had slightly worse institutional performance than S1b. However, figure 5.10 shows that S2 performed substantially better in some years than S1b, and that this may be responsible for the fact that S2 gets an institutional score higher than S1b (see figure 8.5 in the Appendix). In economic terms, S2 performs substantially better than S1b (figure 5.12).

Reviewing figures 5.24, 5.25 and 5.27 show the effect of parameter variations on salt load, assets and clearing success ratio respectively, for scenarios 1b and 2. The results for conflict are omitted, since they once again show very little sensitivity to parameterization.

The salt load plot (figure 5.24) indicates that, broadly speaking, these results are robust to parameter variation. While absolute magnitudes vary, the general trend of reduced salt load in 2 over 1b is shown across all sets of runs.

The raw assets data (figure 5.25) shows that agent financial health is most sensitive to variation in the initial quota price. The higher the initial quota price (IQP), the lower the overall level
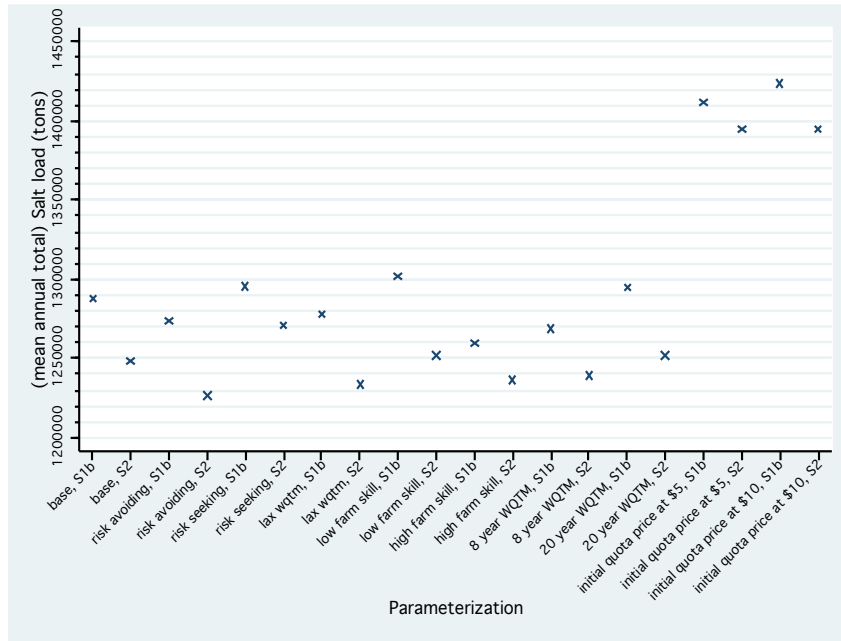
151

Figure 5.24: Salt load (tons) by basin irrigation regions, across parameter variation experiments, comparing results from scenarios 1b and 2. Note that "lax wqtm" refers to an alternate set of institutional rules allowing both quota banking (carrying excess quota over from one year to another) and deficit cancelation (canceling any deficit remaining after a year of trading). Neither are allowed in the WQTM variation simulated in the normal scenarios. This and other parameter variation experiments are described in chapter 4 or in section 8.1.3 in the Appendix. Standard deviations are omitted, since these are large and reflect macroscale controls on salt loading rather than inter-scenario variation, which is the focus of the present analysis. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

of income and assets. The magnitude of the difference between the effects of the IQP variants and all other variants mean that it is easier see other effects when the IQP variants are omitted (figure 5.26). There we see that in all cases the parameter variation data support Hypothesis 2. This indicates that economic performance is not particularly sensitive to varying these model parameters to extremes, although magnitudes of incomes generated and assets accumulated do get affected by initial quota pricing.

In figure 5.27, we see once again that both WQTM management and IQP variants show substantially different magnitudes of response relative to other variants. For the IQP variants, performance does improve from 1b to 2; for the WQTM management variants, the reverse is true. Omitting the IQP and WQTM management data in figure 5.28, we see a mix of sensitivities: more extreme risk strategies lead to worsened institutional performance in 2 over 1b,

152
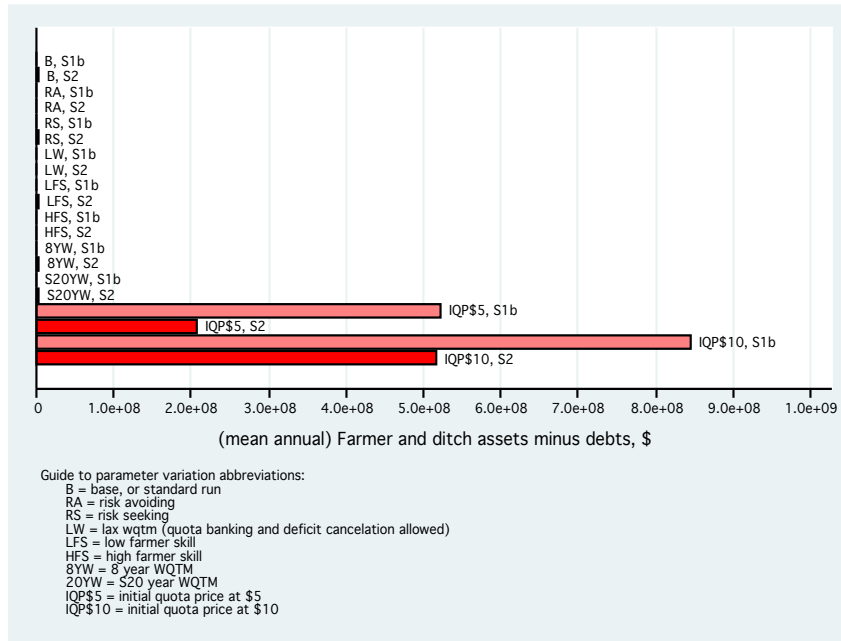
Figure 5.25: Real financial assets ($) for farmers and ditch managers, across parameter variation experiments, comparing scenarios 1b and 2. Plot demonstrates sensitivity of agent financial health to variations in the initial quota price. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.
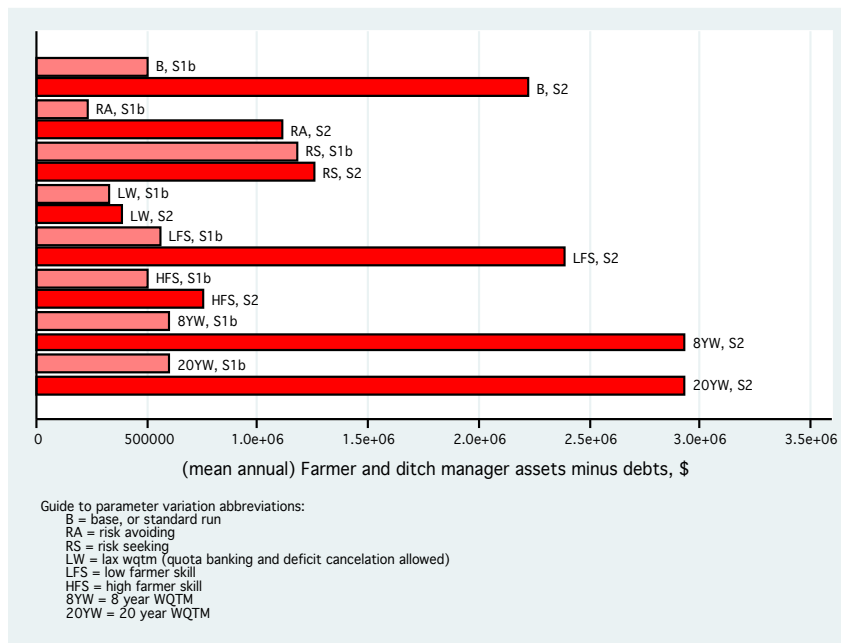


Figure 5.26: Real financial assets ($) for farmers and ditch managers, across parameter variation experiments, comparing scenarios 1b and 2; omitting initial quota price variants. Refer to Appendix section 8.1.2 for detailed scenario descriptions.
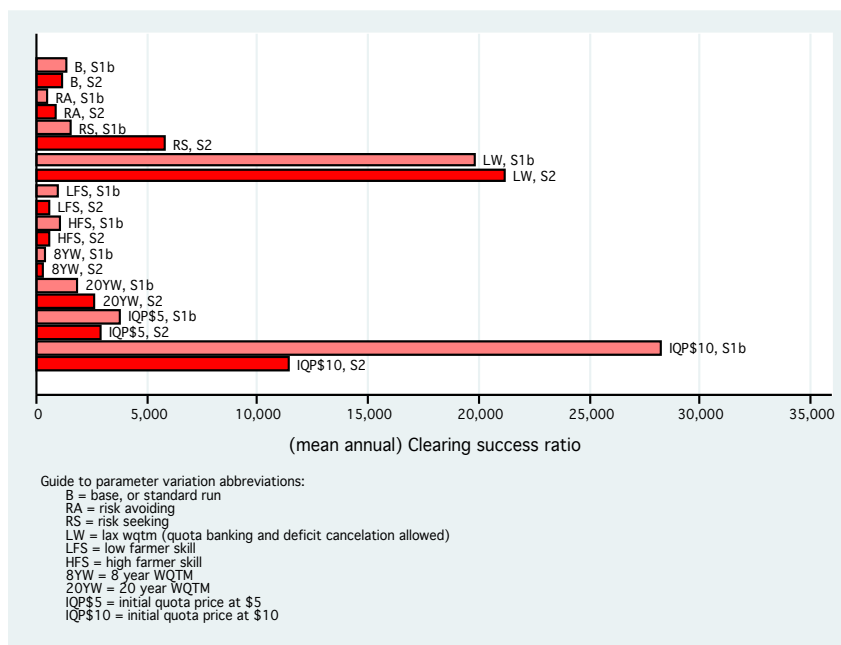
Figure 5.27: WQTM clearing success ratios, across parameter variation experiments, comparing scenarios 1b and 2. Data continue to support the finding that the model is broadly robust to parameter variation, except in the case of initial quota price and quota management options. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

as does increasing the length of the WQTM. Variation in other parameters, on the other hand, does not appear to greatly vary the base trend. The overall picture is that clearing success ratio is a complex measure with minor sensitivity to a broad swath of parameters, but is only greatly impacted by key WQTM decision decisions and initial quota price.

### 5.2.3 Synthesis of Hypothesis 3

#### 5.2.3.1 Summary of Hypothesis 3 synthesis

A synthesis yields the following findings: first, the main thrust of hypothesis 3 — that a non-traditional WQTM with post-hoc informational feedback will perform better in cases where agents are capable of attitudinal change, relative to a traditional WQTM — is supported by all major response factors except river flow. The reduction in flow seen in scenario 3 is not substantial in a real world context, but is nevertheless present and would be a problem for smaller ditches and less senior rights holders. Second, the effect of attitudinal change is to reduce the overall effectiveness of the post-hoc informational feedback; it is only on institutional
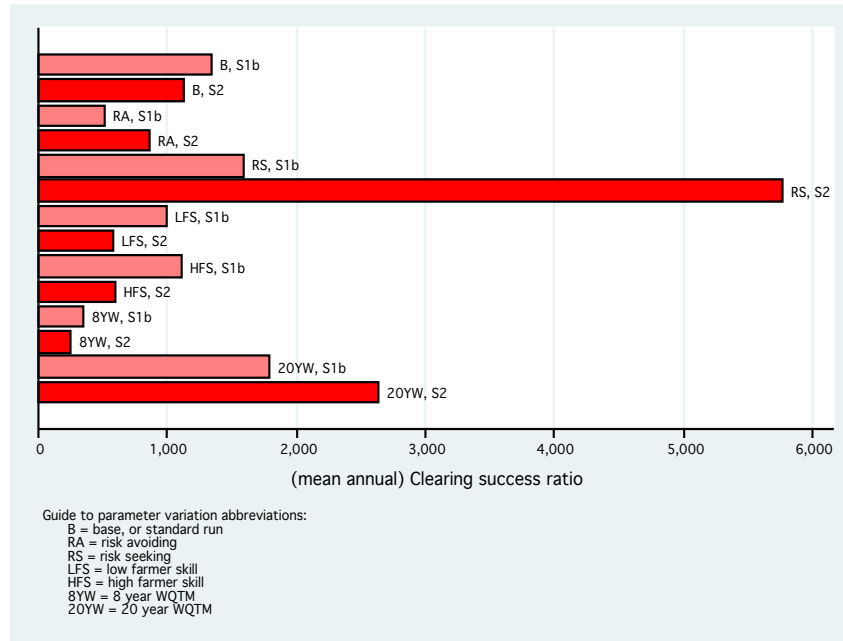
Figure 5.28: WQTM clearing success ratios, across parameter variation experiments, comparing scenarios 1b and 2; omitting initial quota price and WQTM management options variants. Plot demonstrates variable results relative to the baseline: some parameters reverse the base result (which supported for hypothesis 2). Refer to Appendix section 8.1.2 for detailed scenario descriptions.

grounds that the change-enabled scenario 3 shows improved performance relative to scenario 2. Third, parameter variation experiments show the support for hypothesis 3 to be broadly robust. Exceptions were the results for risk strategy and initial quota price, further confirming the sensitivity of the model to these parameters.

### 5.2.3.2    Detailed synthesis of Hypothesis 3 results

Recalling the hypothesis: *Assuming heterogeneous capacity for behavioral change within a resource user population, a WQTM variant that incorporates non-economic incentives intended to promote behavioral change will perform better than a neoclassical WQTM missing such incentives. The S-WQTM will also perform better under the behavioral assumptions of Hypothesis 3 than under the behavioral assumptions of Hypothesis 2.*

There are two separate elements to consider. First, the idea that a WQTM providing post-hoc informational feedback will promote sustainability better than a neoclassical WQTM without such incentives, particularly when we assume a heterogeneous capacity for behavioral

*change.* "Capacity for change" describes the ability of agents to modify their sustainability preferences over time in response to received information. Under earlier behavioral models, agents had a fixed set of sustainability preferences (albeit distributed heterogeneously across the population). In essence, we do not need to see data for a change-enabled population against the neoclassical WQTM, since no incentives would be present to change agent preferences. This means we can make direct comparisons between scenarios 1b and 3.

In terms of overall scoring, figure 5.1 indicates that the hypothesis is supported: scenario 3 has a higher sustainability ranking than scenario 1b. But the difference is ranking is not large, so we need to look a little deeper into the data. In terms of environmental sustainability, scenario 3 scores only slightly worse (see figure 8.6 in the Appendix for detail). Reviewing the summary statistics for salt load and flow (figures 5.7 and 5.8), we see that while scenario 1b performs better than scenario 3 in terms of flow, scenario 3 does much better in terms of salt load. This was a contradiction also seen in hypothesis 2, and raises a problem with the use of unweighted metrics in determining the overall scores. If we narrowly define the objectives of a WQTM to be reducing salt load, hypothesis 3 is well supported: scenario 3 does, in fact, reduce salt load, indicating that making assumptions of behavioral change allows more sustainable behavior over time. However, the environment needs adequate flows, and scenario 3 is on average 24,000 acre-feet down relative to scenario 1b each year. A review of the physical context is appropriate: 24,000 acre-feet represents around 4% of wet year flow measured at Pueblo Dam (Taylor & Luckey, 1974). It is not a substantial difference in the context of natural variation, but 24,000 acre-feet could translate into a significant reduction for smaller ditches. This means that scenario 3 only shows serious flow issues relative to S1b if we restrict our consideration to just the smaller ditches and less senior rights holders.

In terms of economic sustainability, scenario 3 scores much higher than 1b (see figure 8.7 in the Appendix for detail). Reviewing the raw data for income and real assets (figures 5.12 and 5.13) we can better quantify the significance of the difference: an increase in basin income of around $30 million, and an increase in mean real assets of around $1.9 million (note that the real assets metric includes both ditches and farmers, which is why the magnitude exceeds what one might expect for family farms). The picture is one of overall support for hypothesis 3,

156

although given that increased incomes can include increased borrowing, we must be cautious.

Finally, in terms of institutional sustainability, scenario 3 ranks just slightly better than scenario 1b (see figure 8.5 in the Appendix for detail). Reviewing the raw data in figures 5.10, 5.11 and 5.9, we see that the scenarios are barely distinguishable on the basis of trade volume and number of unique traders. However, on the basis of clearing success ratio, scenario 3 shows many times better performance than scenario 1b. Once again, the use of equal weightings for response factors mean that this translates into a minimal ranking difference. But the clearing success ratio is important, since it indicates that while scenarios 1b and 3 may have roughly the same number of traders participating, and see roughly the same amount of quota traded, individuals in scenario 3 have much better luck at selling everything that they put on the market.

Regarding the second element of the hypothesis, which requires a comparison between scenarios 2 and 3, a different picture emerges. Scenario 2 ranks higher than scenario 3 in terms of overall sustainability. Drilling into the score data by category, we see this same pattern replicated in all areas apart from institutional performance. Most of the difference in institutional performance is down to an improved clearing success ratio in scenario 3 versus 2. The overall higher ranking for 2 appears to be largely from the higher economic benefit accruing to agents in that scenario. So, restricting the second element of the hypothesis just to institutional performance, WQTM performance does improve when we assume the capacity for behavioral change but do not change the information feedback instrument. However, on other metrics, performance is little changed or actually worsens.

What do the parameter variation experiments have to say about how the robustness of Hypothesis 3? Robustness comparisons were only developed for scenarios 1b and 3, which limits me to testing the primary assertion of the Hypothesis. Figures 5.29, 5.30, 8.8 and 5.32 show the effect of parameter variations on salt load, assets, conflict and clearing success ratio respectively, for scenarios 1b and 3.

The salt load plot (figure 5.29) indicates that, broadly speaking, these results are robust to parameter variation. While absolute magnitudes vary, the general result of reduced salt load in 3 over 1b is shown across all sets of runs.

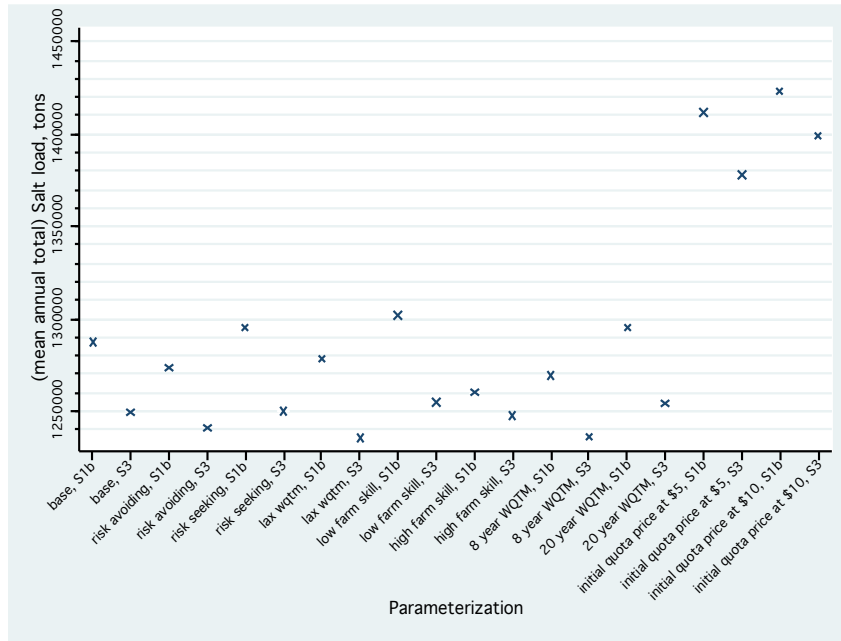The raw data for real assets indicate that, as before, initial quota price (IQP) variants have a

157

Figure 5.29: Salt load (tons) from basin irrigation regions, across parameter variation experiments, comparing scenarios 1b and 3. Note that "lax wqtm" refers to an alternate set of institutional rules allowing both quota banking (carrying excess quota over from one year to another) and deficit cancelation (canceling any deficit remaining after a year of trading). Neither are allowed in the WQTM variation simulated in the normal scenarios. Standard deviations are omitted, since these are large and reflect macroscale controls on salt loading rather than inter-scenario variation, which is the focus of the present analysis. This and other parameter variation experiments are described in chapter 4 or in section 8.1.3 in the Appendix. See text for more detail and refer to Appendix section 8.1.2 for detailed scenario descriptions.

large effect on overall magnitudes (figure 5.30). For both IQP variants, we see assets drop from 1b to 3. Omitting the IQP variants in figure 5.31, assets data indicate that agent financial health is better in scenario 3 than in scenario 1b, regardless of the parameters varied.

As per the normal runs and previous hypotheses, the parameter variation results show that conflict is very robust and shows little variation among scenarios. Once again, conflict decreases with increases in quota price (see figure 8.8 in the Appendix for detail). The full plot for clearing success ratios (figure 5.32) indicates, as before, that response magnitudes are particularly sensitive to choices on WQTM management and initial quota pricing.

Omitting these from the plot in figure 5.33, we see that across all variants, apart from risk strategy, institutional performance improves from 1b to 3. In both risk seeking and risk avoidance strategies, institutional performance worsens from 1b to 3.
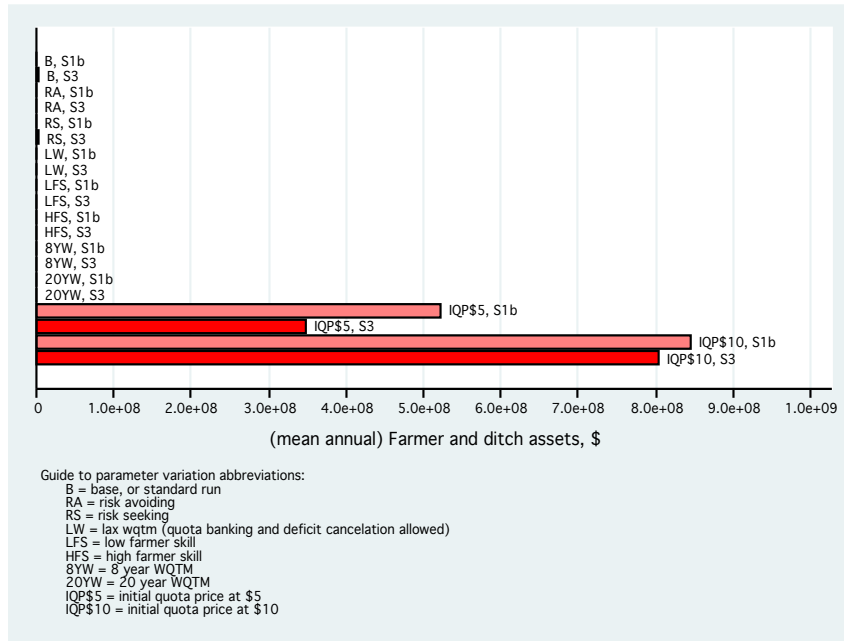
Figure 5.30: Farmer and ditch manager real financial assets ($), across parameter variation experiments, comparing scenarios 1b and 3. Initial quota price has a strong effect on overall magnitudes of outcomes. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.
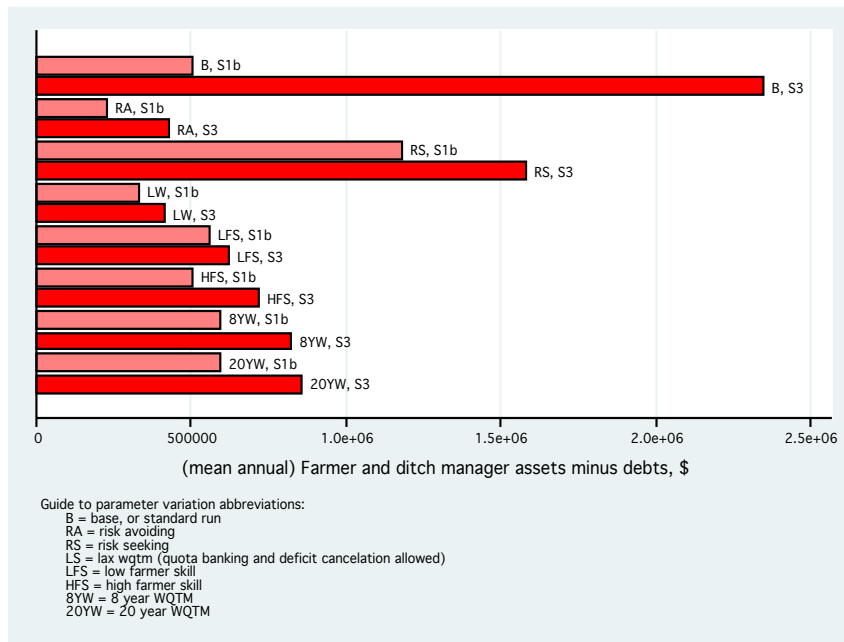


Figure 5.31: Farmer and ditch manager real financial assets ($), across parameter variation experiments, comparing scenarios 1b and 3; omitting initial quota price variants. Now that IQP data is dropped, it is clearer that the data support hypothesis 3. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.
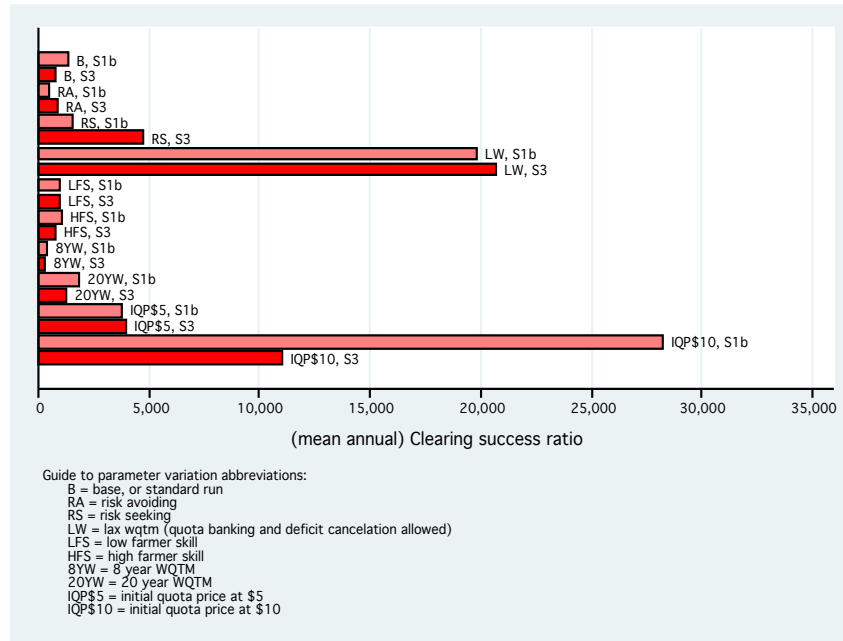
Figure 5.32: WQTM clearing success ratios, across parameter variation experiments, comparing scenarios 1b and 3. Response magnitudes are particularly sensitive to choices on WQTM management (deficit cancelation and quota banking) and initial quota pricing. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

### 5.2.4  Synthesis of Hypothesis 4

#### 5.2.4.1  Summary of Hypothesis 4 synthesis

The principal findings of the synthesis are as follows: first, that hypothesis 4 is generally supported by the data at aggregate and individual levels. One exception is the economic performance data, which indicates that while incomes and assets are relatively high, agents incur a great deal of debt in scenario 4. Second, that these results are broadly robust to variation in parameters, with the exception once again of risk strategy and initial quota price. It is notable that, in some cases, varying parameters to extremes provides even stronger support for the hypothesis.

#### 5.2.4.2  Detailed synthesis of Hypothesis 4 results

Recalling the hypothesis: *An institutional incentive which uses social mimicry and competitiveness dynamics to reinforce behavioral change towards sustainability, will be more effective at improving institutional performance than an institution without this reinforcement component.* Also recall
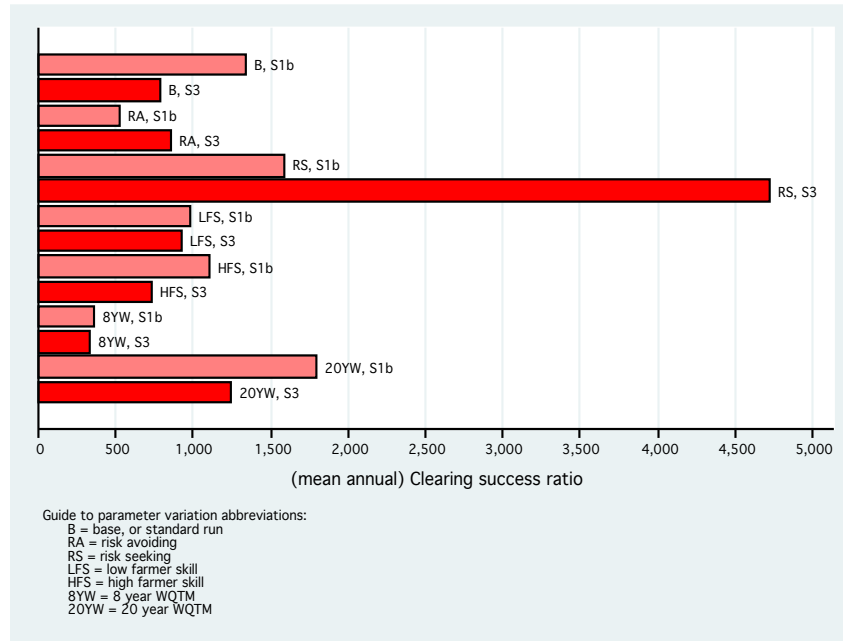
160

Figure 5.33: WQTM clearing success ratios, across parameter variation experiments, comparing scenarios 1b and 3; omitting WQTM management options and initial quota price variants. With these omitted, it becomes easier to see that institutional performance improves from 1b to 3 across all variants, apart from risk strategy. This indicates the broad robustness of hypothesis 3. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

from section 4.2.6.3 in Chapter 4 that the "competitiveness" component is driven by a deliberately inaccurate conformism logic. When the received information on neighboring agents is sustainability-rated above their own S-factor, they will seek to modify their own attitudes closer to the received rating. The existence of a magnitude modifier in this logic (which increases as the agent gets closer to the received rating) means that the agent will usually be driven to increase its S-factor above the received rating, resulting in an effectively "competitive" dynamic.

For this hypothesis, the most directly comparative scenarios are 3 and 4. In scenario 4, additional information is added to the basic informational incentive instrument. In addition to information on the sustainability performance of the agent's recent WQTM-related action, information on the sustainability performance of up to 5 neighboring agents is also supplied. To strongly support this hypothesis, the results must show that scenario 4 had better overall sustainability than scenario 3.

Examining figure 5.1, this is clearly true. Scenario 4 ranks as the second most sustainable

scenario, while scenario 3 ranks as the 4th most sustainable. A similar pattern is found separately in the environmental, social and institutional scoring (see Appendix figures 8.6, 8.4, 8.5 for detail), but interestingly, not in the economic scoring (8.7). Scenario 4 shows roughly the same annual income as scenario 3, but displays around $600,000 less in mean assets. This indicates that agents are incurring higher debt levels (which we can see in the liabilities plot, figure 5.17). Focusing on institutional performance, the data show that the WQTM under scenario 4 performs the best among all scenarios with attitudinally-heterogeneous agents; the only scenario with better institutional scoring is the homogeneous case, scenario 1a. This is to be expected, since the WQTM in that scenario enjoys an agent population with a near-perfect behavioral match to the incentives it provides.

How much better in real terms is scenario 4? The most important metric is salt load, since this indicates the overall effectiveness of the WQTM in managing water quality issues. Figure 5.7 shows that while scenario 4 does not have the lowest loads, it does have the lowest loads among empirically parameterized runs (i.e. agent attitude and behavioral models parameterized using LAB stakeholder data). The range from the highest loading empirical scenario (scenario 1b) to scenario 4 is around 30,000 tons of salts annually. Relative to the baseline run, scenario 4 is a year-on-year reduction of around 5%.

What do the parameter variation experiments have to say about how the robustness of hypothesis 4? Figures 5.34, 5.35, 8.9 and 5.37 show the effect of parameter variations on salt load, assets, conflict and clearing success ratio respectively, for scenarios 3 and 4.

The salt load plot indicates that, broadly speaking, these results are robust to parameter variation. Most variants show a decrease in salt load from scenario 3 to 4. The exceptions are risk seeking behavior and setting initial quota price to $5; in both cases salt load increases from scenario 3 to 4.

As we have seen repeatedly, figure 5.35 indicates that agent financial health is most sensitive to variation in the initial quota price. The higher the initial quota price, the lower the overall level of income and assets. In the quota price variants, we see a decrease in assets from 1b to 3. Omitting these data to make analyzing the other data points easier, figure 5.36 indicates that assets actually improve across most variants, with the exception of the 20 year WQTM run.
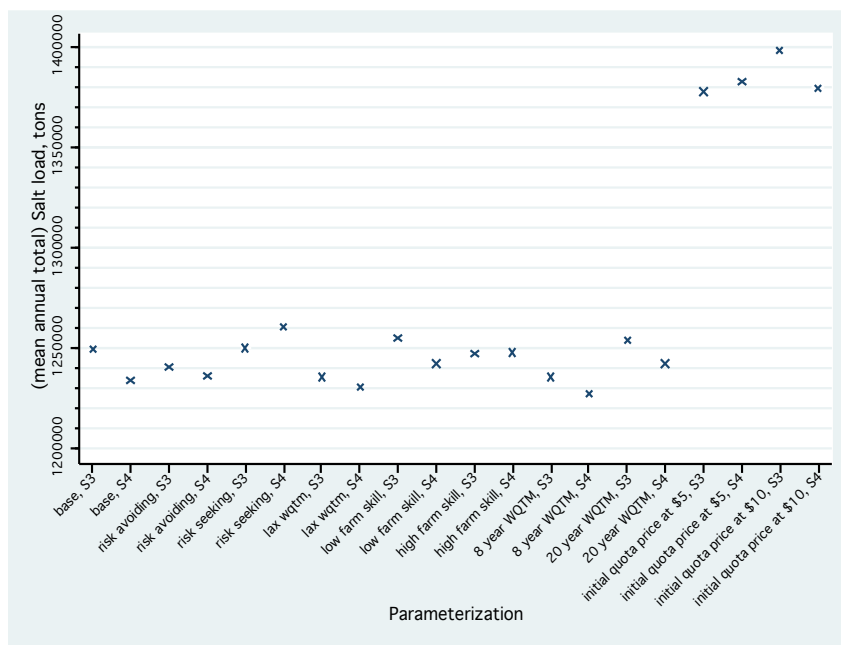
Figure 5.34: Salt load (tons) into river by basin irrigation regions, across parameter variation experiments, comparing scenarios 3 and 4. Note that "lax wqtm" refers to an alternate set of institutional rules allowing both quota banking (carrying excess quota over from one year to another) and deficit cancelation (canceling any deficit remaining after a year of trading). Neither are allowed in the WQTM variation simulated in the normal scenarios. Standard deviations are omitted, since these are large and reflect macroscale controls on salt loading rather than inter-scenario variation, which is the focus of the present analysis. This and other parameter variation experiments are described in chapter 4 or in section 8.1.3 in the Appendix. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

This actually shows stronger support for hypothesis 4 than the base runs.

The parameter variation results for conflict once again show that this response factor is robust and shows little variance among scenarios. We also again see that conflict decreases with increases in quota price (see Appendix figure 8.9 for detail).

Finally, clearing success ratios (figure 5.37) worsen when the initial quota price is raised. Omitting these results (figure 5.38), clearing success rate only shows substantial improvements in scenario 4 over 3 in the cases of relaxed WQTM management, and for risk seeking behavior. The sensitivities here suggest that both risk seeking behavior and allowing quota banking and deficit cancelation increase the economic rationality of agent behavior in the system. All other variants either show poorer institutional performance, or little discernible change, indicating the outcomes are broadly robust to these parameters.
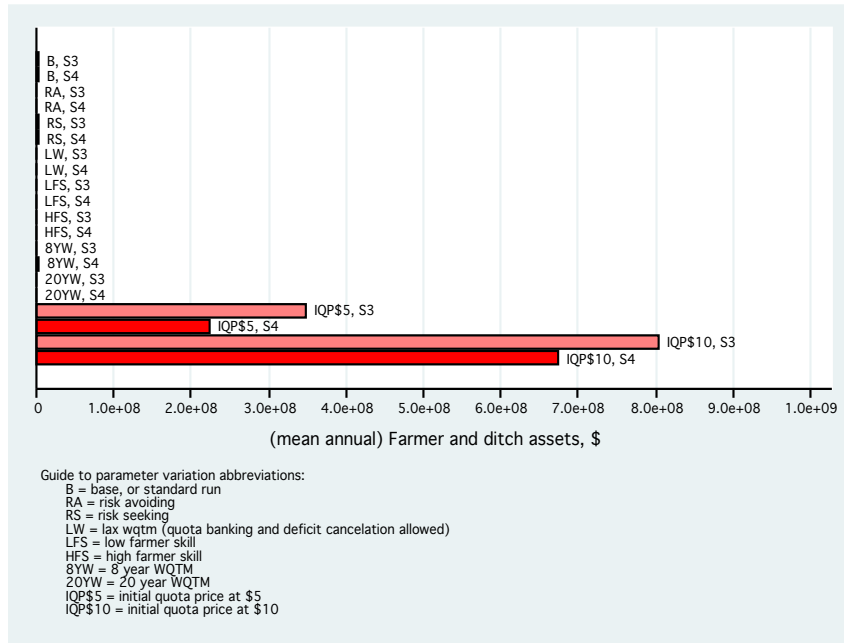
Figure 5.35: Farmer and ditch manager real financial assets ($), across parameter variation experiments, comparing scenarios 3 and 4. Agent financial health is most sensitive to variation in the initial quota price (inverse relation). Refer to Appendix section 8.1.2 for detailed scenario descriptions.
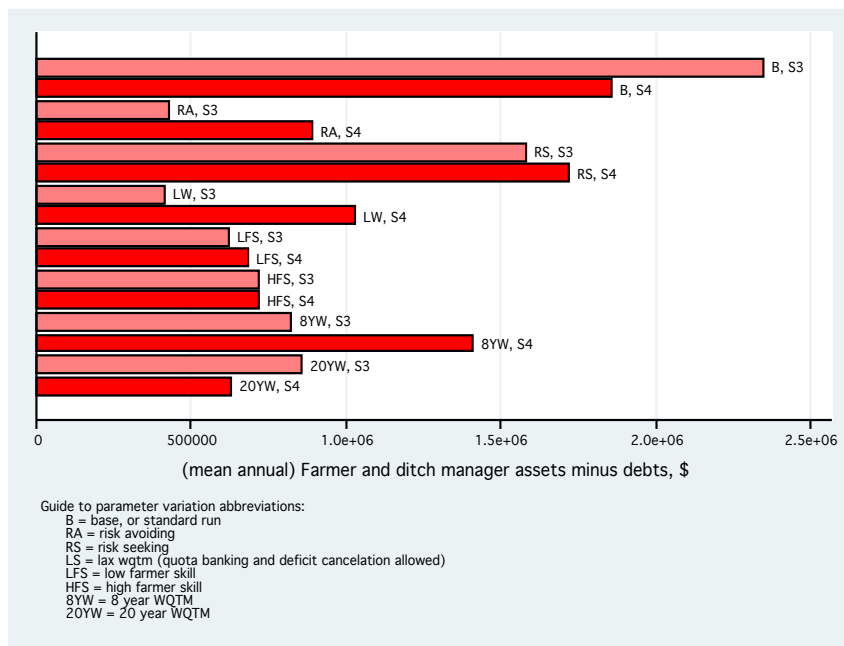


Figure 5.36: Farmer and ditch manager real financial assets ($), across parameter variation experiments, comparing scenarios 3 and 4; omitting initial quota price variants. Indicates stronger support for hypothesis 4 than many of the base (non-parameter variation) runs. Refer to Appendix section 8.1.2 for detailed scenario descriptions.
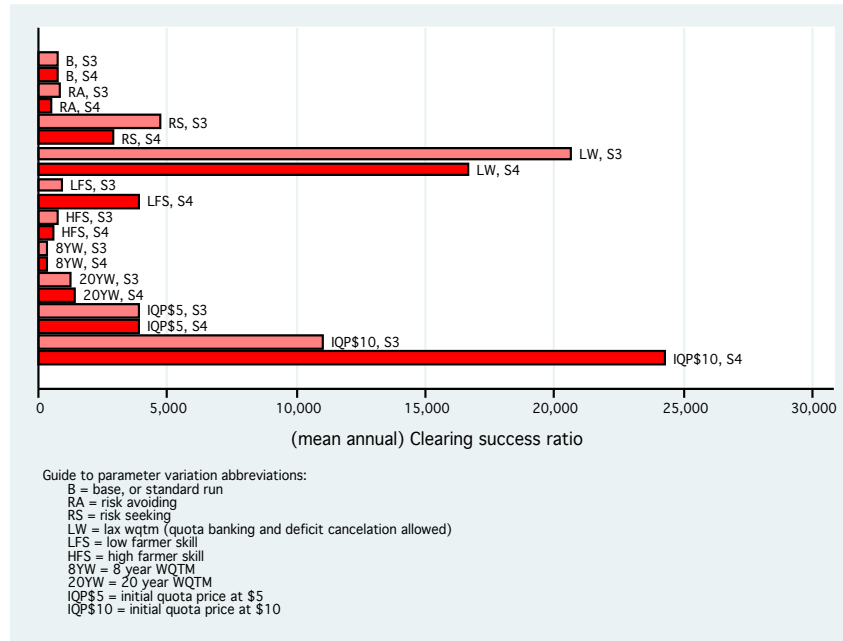
164

Figure 5.37: WQTM clearing success ratios, across parameter variation experiments, comparing scenarios 3 and 4. Ratios worsen when we raise the initial quota price, suggesting once more sensitivity to this parameter. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

### 5.2.5  Synthesis of Hypothesis 5

#### 5.2.5.1  Summary of Hypothesis 5 synthesis

The main findings of the synthesis are as follows: first, that hypothesis 5(i) and (ii) are only weakly supported by the overall sustainability scoring data, with stronger support among certain response factors including river flow and agent incomes. Second, that the social scoring controls whether or not the overall scoring support for hypothesis 5(i) and (ii) is strong, i.e. omitting the social scores strengthens support. Third, that support for both elements of hypothesis 5 is influenced by whether or not we include social dynamics. Allowing for social dynamics has inconsistent results, in some cases supporting and in others refuting the hypothesis.

#### 5.2.5.2  Detailed synthesis of Hypothesis 5 results

Recalling the hypothesis: *(i) Including competing informational incentives in a pro-social WQTM will alter the capacity of a pro-social WQTM to achieve sustainable outcomes. (ii) Actual outcomes will*
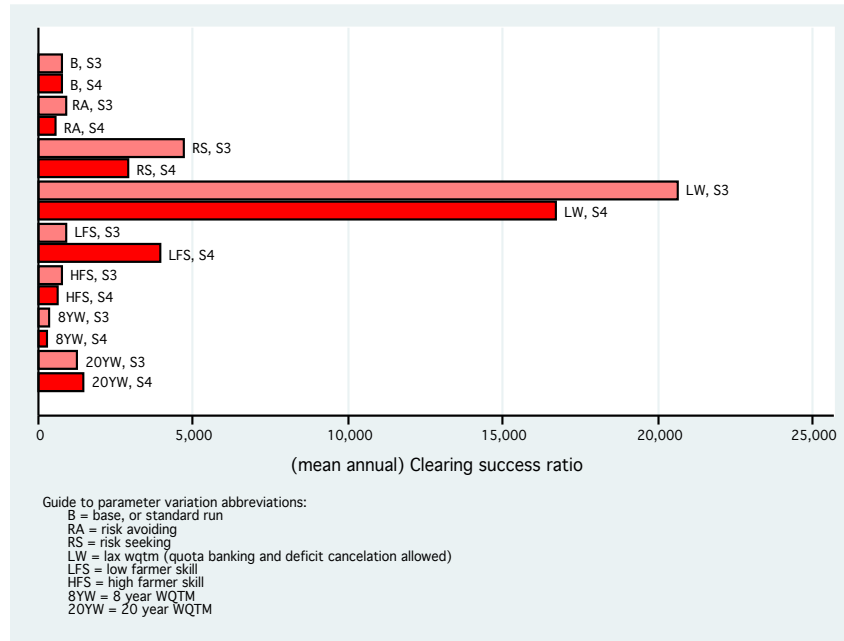
165

Figure 5.38: WQTM clearing success ratios, across parameter variation experiments, comparing scenarios 3 and 4; omitting initial quota price variants. Indicates that we only see substantial improvements in scenario 4 over 3 in the case of relaxed WQTM management and risk seeking behavior. This supports that notion that it is more economically rational in this system to be risk seeking, and that changing WQTM management to allow both deficit cancelation and quota banking allows more economically rational behavior. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

*be dependent on the initial distribution of behavioral strategies within a population.*

Two sets of experiments were conducted to test each assertion in the hypothesis. Beginning with (i): in scenario 5a, the WQTM included both pro-social and pro-self informational feedback (NCW-PSO); in scenario 5b, the NCW-PSO was supplemented by providing both forms of information on neighboring agents (NCW-PSON). Scenario 5a and 5b were then run with the S-factor empirically parameterized, providing results useful in answering (i). Then a second set of runs were conducted: one using the S5a institutional setup (the NCW-PSO), but alternately forcing the S-factor parameterization towards the pro-self end of the scale (5a-LSF) and the pro-social end of the scale (5a-HSF); the other using the S5b institutional setup (the NCW-PSON), and with the same alternate high/low forcing for the S-factor parameterization.

The idea in conducting runs with both S5a and S5b was not only to test the hypothesis as stated, but also to see how robust the hypothesis was to changing other elements of the WQTM incentive structure (specifically, the addition of social dynamics). Results supportive

166

of hypothesis 5(i) would show that scenarios 5a and 5b were different from the most similar and previously tested scenario, S4. Results supportive of hypothesis 5(ii) would show that the 5a-LSF/5b-LSF runs scored worse on sustainability metrics than 5a-HSF/5b-HSF runs.

Turning to hypothesis 5(i) first: examining figure 5.1, scenario 5a does show a different ranking from scenario 4, though the difference is not substantial (scenario 4 ranks second, while scenario 5a comes in third). Scenario 5b shows a larger difference in ranking, coming in below not only scenario 5a but also 2, 3 and 4. The results show weak support for hypothesis 5(i).

Turning to hypothesis 5(ii): the results are mixed. For the 5a set, the LSF scenario ranks slightly higher than the HSF scenario. For the 5b set, the LSF scenario ranks substantially lower than the HSF scenario. The hypothesis is neither strongly supported nor refuted for the case where opportunity cost is the only supplement to sustainability information. Conversely, the hypothesis is more strongly supported for the case where opportunity cost information is provided in the context of social dynamics.

If we more closely examine the by-category sustainability scoring, there is a larger difference in ranking between 5a/5b and other scenarios in the economic, environmental and institutional categories (see Appendix figures 8.7, 8.6 and 8.5 for detail). While the social category also shows difference, recall that the raw data underlying the social scores shows only subtle variation between scenarios. If we re-plot the overall scoring without the social scores (figure 5.16, we see an overall shift: scenarios 5a and 5b now show more difference from other scenarios. Furthermore, we now see that the LSF runs rank lower than the HSF runs, as the hypothesis predicts. This provides additional support for the point discussed earlier regarding the problems in using complex ranking metrics, but also strengthens support for hypothesis 5 since the only major factor complicating the overall scores for hypothesis 5 is a set of data which shows minimal variation among scenarios.

Hypothesis 5(i) states that competing informational incentives will alter the capacity of a pro-social WQTM to achieve sustainable outcomes. How well do the raw data support this assertion? Comparing statistics for salt load in scenario 4, 5a and 5b (figure 5.7), both 5a and 5b do show a small difference in loading relative to 4: around 10,000 tons per year higher loads. Examining flow results (figure 5.8), scenarios 5a and 5b achieve around 8,000 acre-feet higher

volumes of flow than 4. On the institutional side, results indicate that both scenarios 5a and 5b have considerably worse clearing success ratios, lower trade volumes, and barely distinguishable differences in unique trader numbers. On the economic side, 5a shows dramatically larger incomes, while 5b shows very little difference (results mirrored in the assets data). Finally, on the social side, we see slightly less conflict in 5a and 5b, and very little difference in deficit equity figures. Noting the limited real world significance of the differences in environmental results, yet much more significant results economically, the best we can say is that the raw data weakly supports the first assertion of Hypothesis 5, but the level of support is affected by whether or not we include social dynamics.

The second assertion states that the initial parameterization will control the overall outcome. Once again setting aside the scoring, how well do the raw data support this assertion? The LSF to HSF transition causes a reduction in salt load. This is true for both 5a and 5b, although the reduction is larger for the 5a set. River flows increase from LSF to HSF in 5a, but decrease in 5b. In the institutional category, the LSF to HSF transition appears to cause substantial improvement in clearing success rate, most markedly so for the 5b results. Trading volumes appear to be little different in LSF or HSF for 5a, but drop substantially from LSF to HSF for 5b. Interestingly, the LSF to HSF transitions for unique trader numbers are the only scenarios to show noticeable differences relative to other runs: in both 5a and 5b, the HSF scenario results in lower trader numbers. In economic terms, the 5a LSF to HSF transition sees a very substantial decrease in income (from around $60 million total annual income to around $20 million), while the same for 5b shows no change. As we have come to expect, real assets reflect the same differences. Finally, for the social response factors, the LSF-HSF transition shows minimal difference across both 5a and 5b. The second assertion is well supported in that the initial parameterization choice for sustainability factor clearly is causing a difference in outcome. In Chapter 3 I argued that an initially more pro-social parameterization (i.e. HSF) would result in more sustainable outcomes, and vice versa for anti-social parameterizations (i.e. LSF). On the basis of the scoring, this assertion is broadly supported. However, caution is advised: we need to discount the weakly variant social data to get to this finding, and the underlying raw data are highly variable.

## 5.3 Trend analysis

The preceding analysis has been largely conducted using summary statistics, mostly annual or weekly means. This is the most straightforward approach for a ranking and scoring exercise, but can miss important time-dependent trends.

Time series data for environmental and social variables show evidence for macro-scale controls over and above any effects of the WQTM institutions. It appears that river flow may exert macro-scale control over salt loads. It also appears to be influencing levels of social conflict over and above the effects of individual scenarios. A comparison of figures 5.39 and 5.40 support the idea of macro-scale control of water quantity on water quality, which suggests that whatever support the data provides for the main hypotheses, the effect of institutional interventions in reality will be dependent on macroscale physical dynamics.
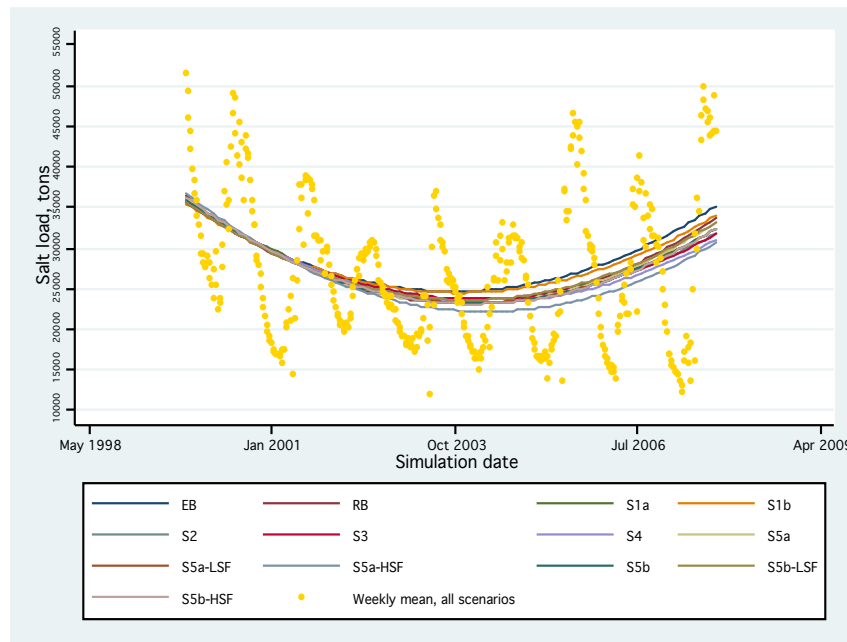


Figure 5.39: Salt load (tons) contributed to river by basin irrigation regions; quadratic trend based on weekly time series. The plot shows that differences in sustainability benefits among scenarios become more pronounced with time. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

Aside from underlining the strong seasonality in basin water quantity and quality dynamics, figures 5.41 and 5.42 suggest that over an increased simulation period we might see increase differences in system response among different scenarios. Since ArkAgent is a non-equilibrium
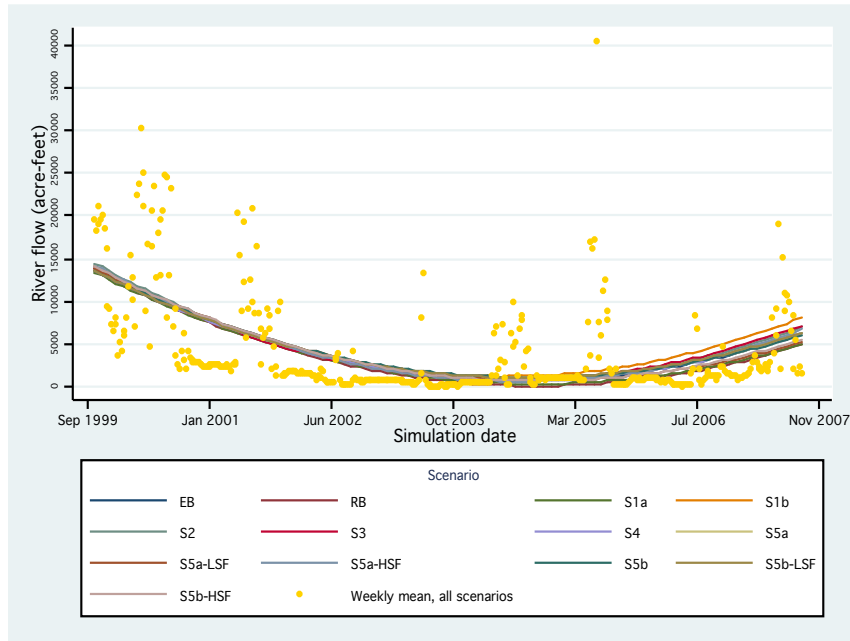
Figure 5.40: River flow (acre-feet) at the Colorado/Kansas state line, weekly time series, fitted with a quadratic model. Quadratic trend supports the assertion that river flow acts as a macro-scale control on salt load. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

model, there is no empirical basis for extrapolating trends. However, given that dramatic shifts in trend are rarely seen across large and complex systems like the LAB, it is reasonable to make some qualitative suggestions for where these scenarios might evolve over a longer simulation period. Most of the environmental and economic data show increasing divergence among scenarios with time, suggesting that in the decade following this scenario we might expect to see increased differences in system response, and that overall results seen in the averaged data would become more distinct. For example, projecting forward on a decadal timescale we might expect S2 and S4 — the most sustainable scenarios in overall terms — to show increasing levels of sustainability with time relative to the baseline.

Plotting salt load and flow data as weekly means (figures 5.43 and 5.44) yields some clear intra-annual trends, and shows how the different effects of the different behavioral and institutional assumptions are much clearer at this scale, particularly in the summer. Salt loads decline in the off-season when ditches are not irrigating, and increase again in the summer, which is to be expected since recharge and seepage rates will be highest when more water is being moved through the system. An implication for discussion is how an institution should take account of
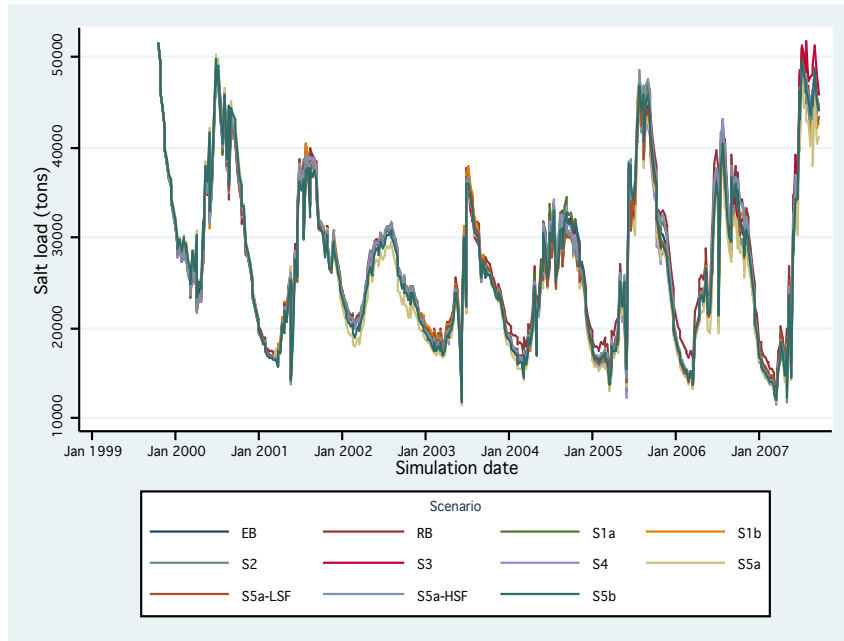
Figure 5.41: Salt load (tons) contributed to river by basin irrigation regions, weekly time series. Plot shows that salt loading is strongly seasonal intra-annually, and that salt load decreases in the middle years of the simulation. The inter-annual trend suggests macro-scale control on overall salt trend. Refer to Appendix section 8.1.2 for detailed scenario descriptions.
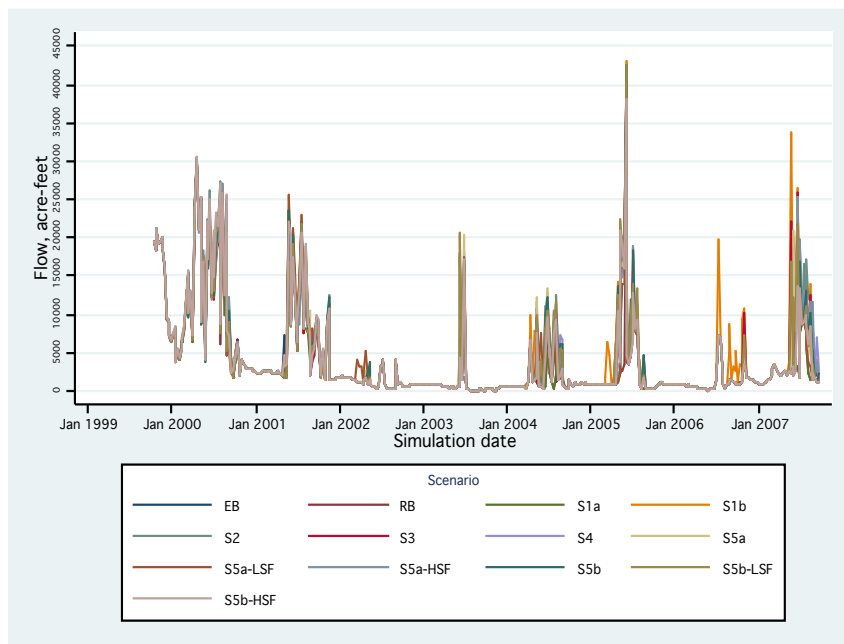


Figure 5.42: River flow (acre-feet) at the Colorado/Kansas state line, weekly time series. Plot shows expected strong annual seasonality in flow and strong macro-scale similarities with salt load trend. This indicates flow may act as macro-scale control on salt load. Refer to Appendix section 8.1.2 for detailed scenario descriptions.
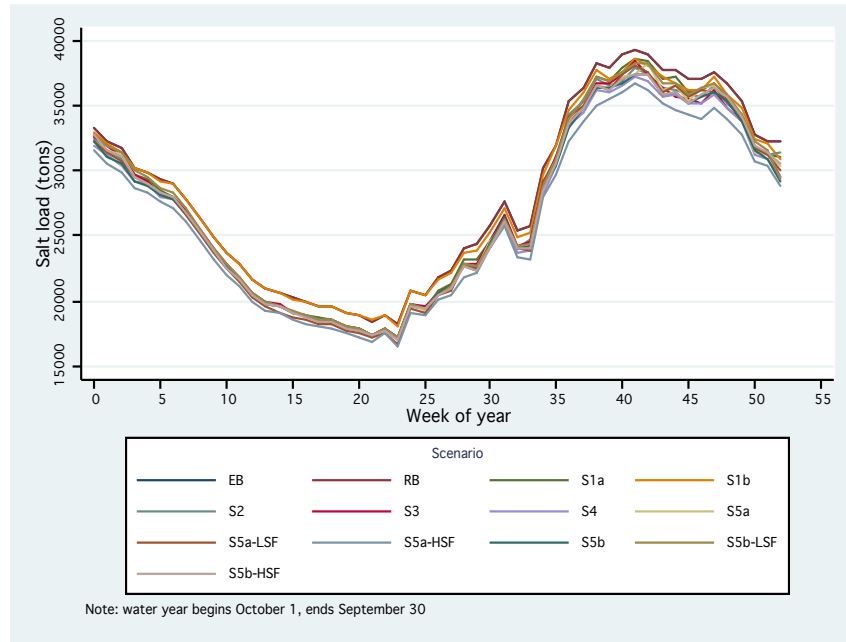
171

these strong intra-annual patterns.



Figure 5.43: Salt load (tons) contributed to river by basin irrigation regions, time series based on intra-annual weekly mean. Plot shows that salt loads correlate with irrigation activity, and that differences between the sustainability benefits of different scenarios are much clearer at this time scale. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

The timing of initial quota allocation (conducted by the state of Colorado at the beginning of each water year) exerts strong controls on the temporal distribution of a number of institutional variables, including the number of traders in the market (figures 5.45 and 5.46), clearing success ratio (figure 5.47) and quota trade volumes. Strategic behavior by agents appears to be a secondary control on the timing of trading activity. These results demonstrate how decisions on the temporal organization of an institution can have knock-on effects on institutional dynamics.

The clearing success ratio data demonstrate how institutional efficiency declines with time. This trend is echoed with the trend of reduced market liquidity with time (figure 5.48). Both data sources suggest that quota allocation policy may need to be more adaptive to market conditions.

The WQTM provides a clear economic boost in the form of incentive payments. Results for income with time (figure 5.49) show a substantial annual boost in agent income coinciding
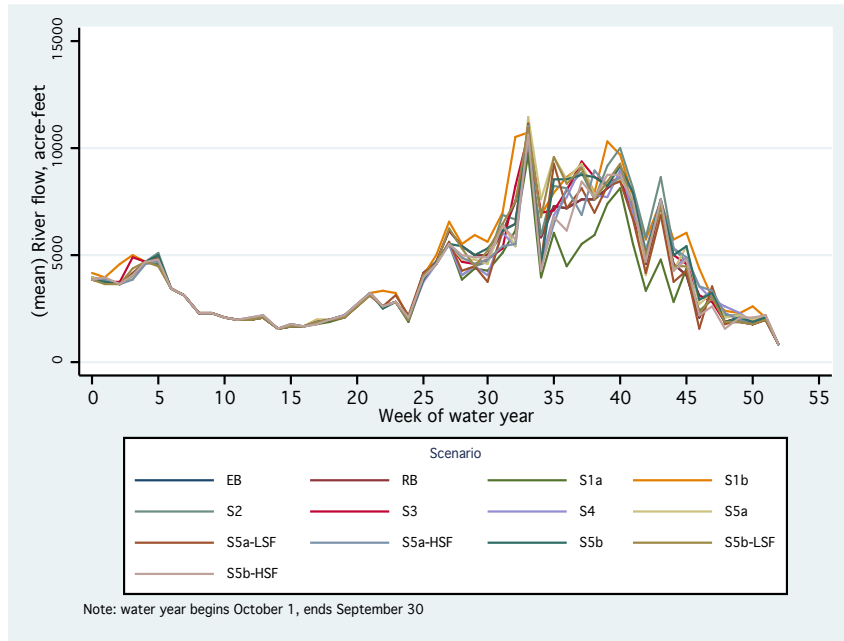
Figure 5.44: River flow (acre-feet) at the Colorado/Kansas state line, intra-annual weekly mean time series. Plot shows that flow has expected strong seasonality related to climatic patterns, and also that differences between scenarios become more pronounced in the summer irrigation period. Refer to Appendix section 8.1.2 for detailed scenario descriptions.
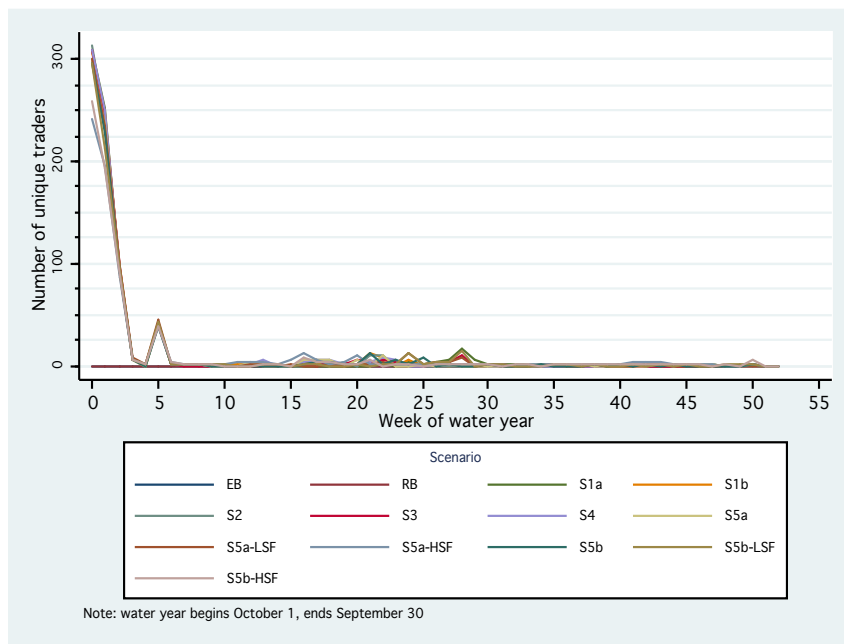


Figure 5.45: Number of unique traders in the WQTM, intra-annual weekly mean time series, by scenario. May provide an indication of both the periodicity imposed on market dynamics by the decision to undertake quota allocation in October each year. Refer to Appendix section 8.1.2 for detailed scenario descriptions.
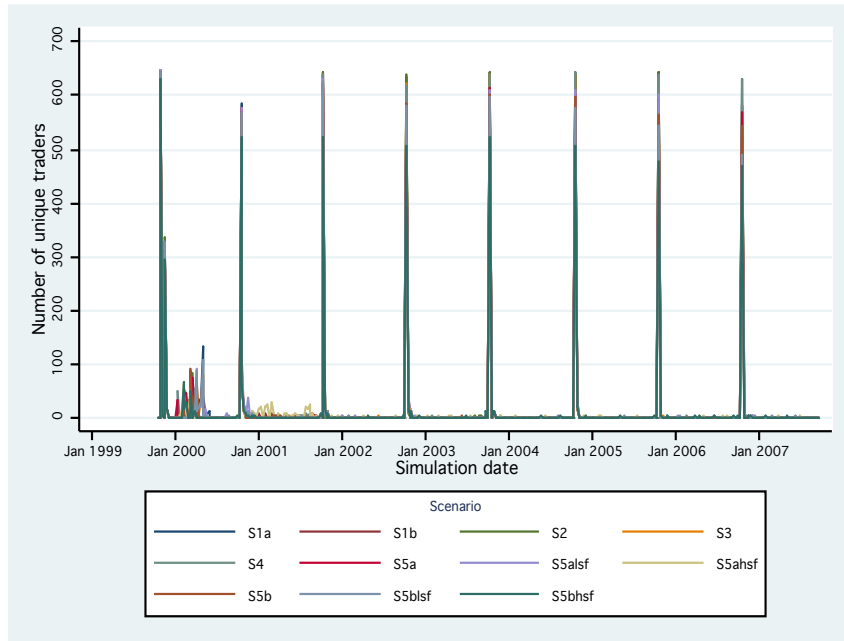
Figure 5.46: Number of unique traders in the WQTM, time series by scenario. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.
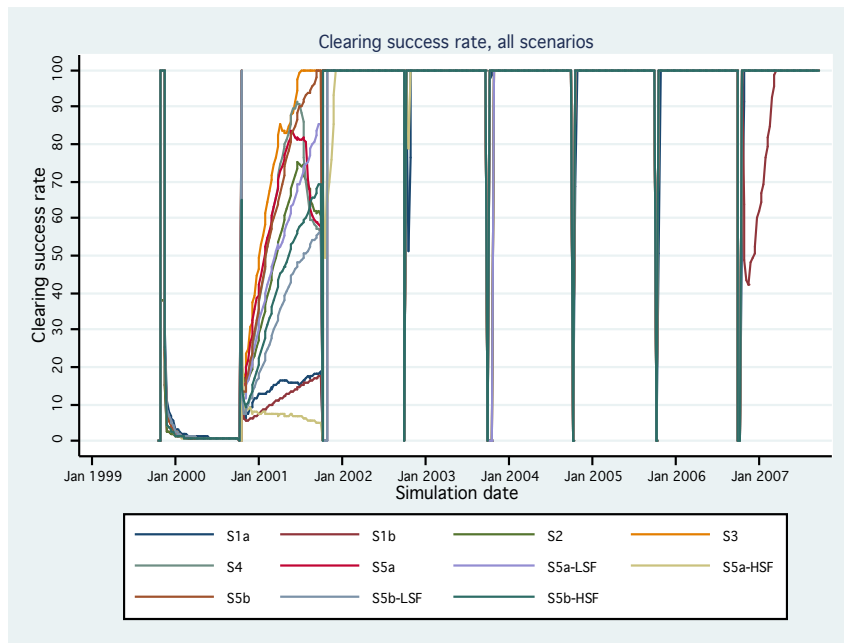


Figure 5.47: WQTM clearing success ratio (CSR), time series, by scenario. Plot shows that market efficiency is highest in the first few years, and thereafter the market shows consistent over-capacity; note that the figure is a trimmed plot which normalizes ratios to the range 0-100, with the effect of removing outlier points in several scenarios. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

Figure 5.48: Volume of quota (kg) traded in the WQTM, time series by scenario. Note how market liquidity declines over time across all scenarios. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

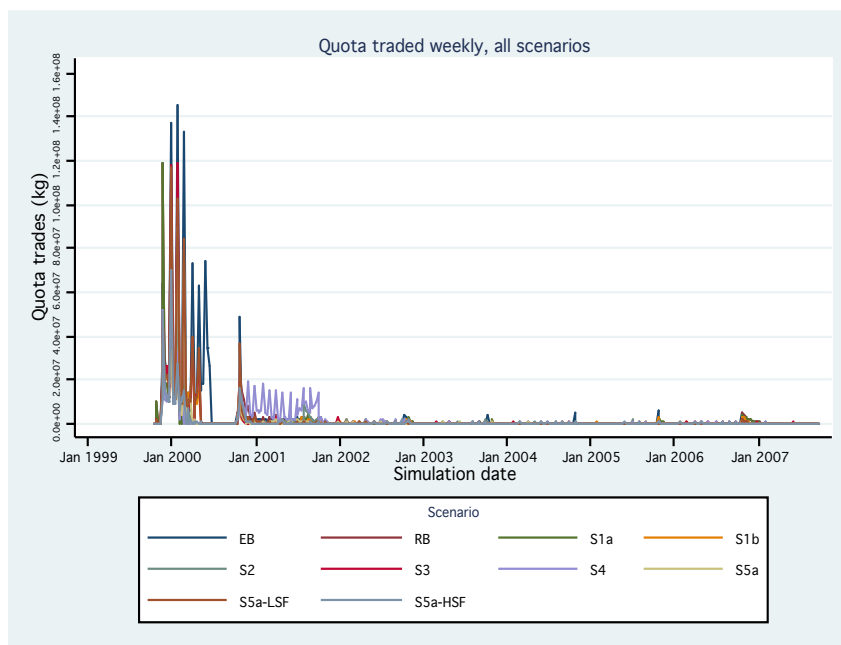with the point when the State Engineer allocates fine monies as incentive payments to those agents who did not exceed their quota. Income at other periods of the year is derived from sale of crops and off-farm income, and the much smaller nature of these contributions indicates the potential economic boost of the WQTM institution.

Finally, results for conflict with time (figure 5.50) appear to be strongly influenced by model equilibration issues in the first two years of the simulation. Over the first year or so, conflict rises sharply before stabilizing in late 2000. This may be an model "warm-up" issue. Conflict is treated as purely an output variable, and so not subject to any parameterization at the start of the model run: all farmers across all scenarios start with a conflict perception of 0. While under even the baseline conditions the system does not have zero conflict levels, it still takes the conflict perception in each agent some time to react to the agent's condition.

If we trim the plot to around a year into the simulation (figure 5.51), conflict shows a decreasing trend over time from the initial stabilization point across all scenarios. This suggests that the model maybe be evolving towards a more stable point. Different scenarios show 1-10% variation around this trend, however, suggesting that these scenarios do have a real effect on

175
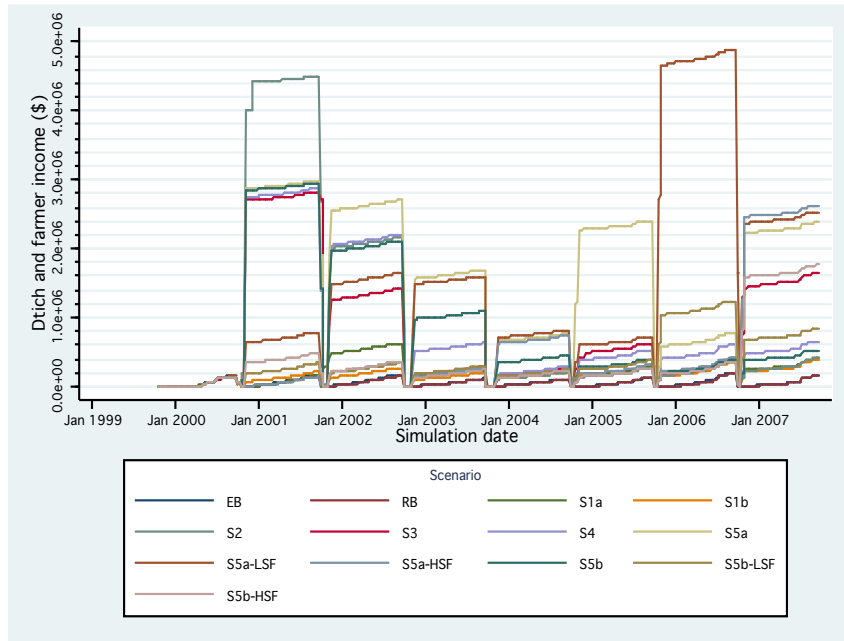
Figure 5.49: Ditch manager and farmer income ($), time series, by scenario. Income shows strong annual cycles related to agent accounting practices, but also indicating the strong contribution of WQTM incentive payments once a year, and lesser contributions from on- and off-farm income. Refer to Appendix section 8.1.2 for detailed scenario descriptions.
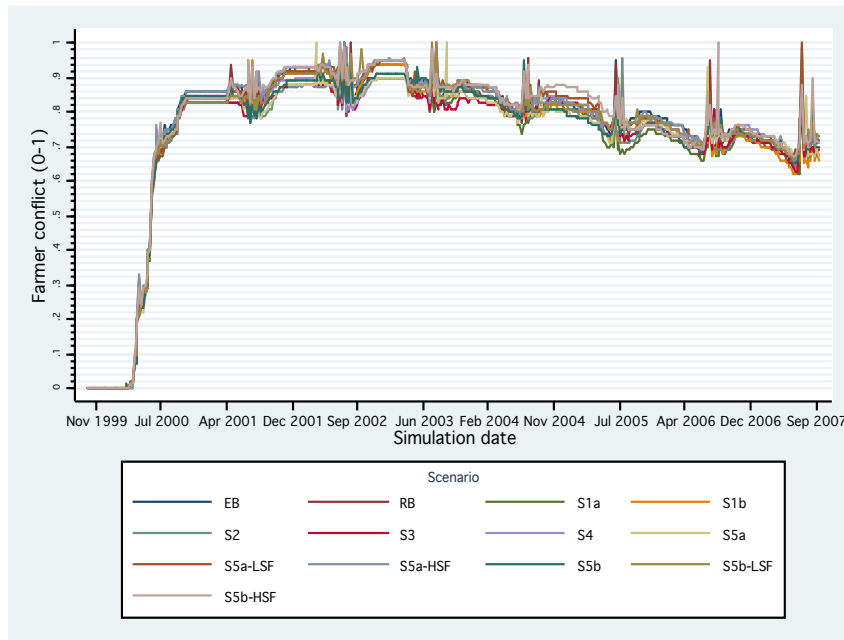


Figure 5.50: Farmer conflict (0.0-1.0), time series, by scenario. Sharp rise initially may be due to model equilibration issues. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

conflict even if the overall means are similar.
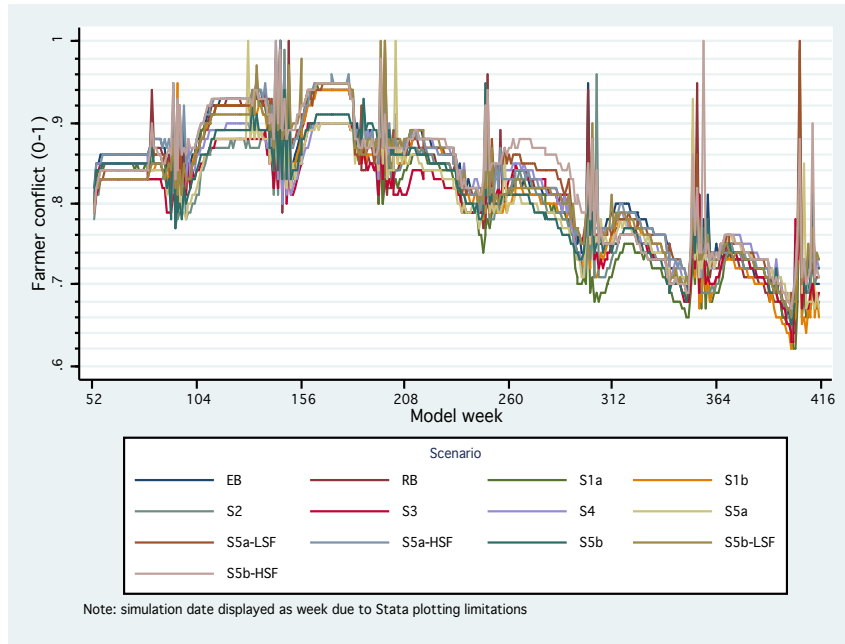


Figure 5.51: Farmer conflict (0.0-1.0), time series, by scenario; trimmed to simulation period 2000-2007. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.



Figure 5.52: Farmer conflict (0.0-1.0), intra-annual weekly mean. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

Reassuringly, intra-annual results (figure 5.52) continue to support the notion that adding
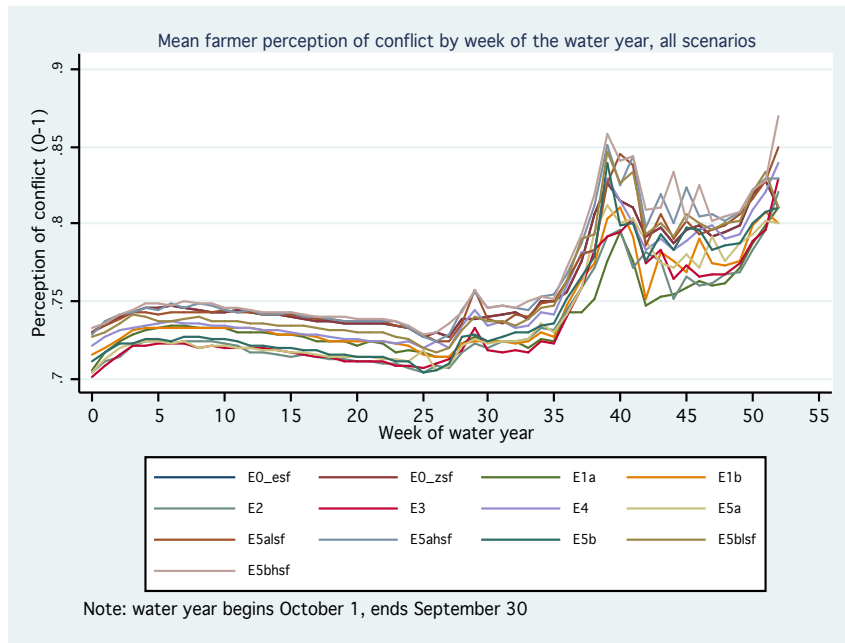
non-traditional incentives can improve the social sustainability of an institution: S2 consistently shows some of the lowest levels of conflict. But note that since the intra-annual means differ by less than 5%, the impact of institutional variation is limited. Spikes in conflict levels during the summer are likely due to the fact that perceptions of water scarcity rise at this time, when temperatures and demand are highest and water availability can become constrained. This continues to support the implication that institutional design should play close attention to temporal adaptation in institutional mechanisms.

## 5.4 Model development methods

In chapter 4, I characterized the model development methods (individual interviewing, collaborative modeling workshops, Delphi iterations and psychometric surveying) as useful for improving the quality of the model. But these techniques also yielded results not directly related to the model development or the five hypotheses: the general suitability of each method's application to agent-based modeling; and insights into the broader challenge of enhancing sustainability in the Lower Arkansas Basin. For reasons of brevity, a detailed discussion of these results is included in the Appendix in section 8.4. I summarize some of the most important results in the following paragrapgs.

In Appendix section 8.4.2, I highlight how the individual interviewing results support my overall framing of the social system in the LAB as consisting of stakeholders with diverse attitudes and strategies towards water use. The results emphasize how non-monetary ties to the community and the land keep farmers in farming through profitable and unprofitable times. The results also indicate how many stakeholders view the hydrologic system more as an engineered structure for water delivery than as a complex environmental feature. The use of individual interviewing in model development was helpful in setting the scene and seeding various areas of model design, but getting stakeholders to talk about specifics was very challenging: most interviewees wanted to talk instead at a more general level on issues relating to the basin as a whole.

In section 8.4.3, I describe how the use of a non-computational role-playing game had lim-

ited direct benefit to the model, but also shed light on stakeholder dynamics and proved a good means by which the modeling process could be introduced. The computational role-playing game was an extremely effective — although logistically challenging — means of obtaining detailed stakeholder input to the model. The workshops suffered from low attendance, which is a problem frequently seen with participatory events in the LAB. It is also a broader and rarely acknowledged problem in collaborative modeling.

Finally, in the latter parts of sections 8.4.4 and 8.4.5, I describe the benefits of applying Delphi in model development and psychometric testing in model parameterization. I show that Delphi filled important gaps in model development not met by the workshops. Similarly, I show that psychometric testing was a useful tool in tackling a parameterization issue not easy to handle in a participatory setting.

# Chapter 6

# Discussion

## 6.1   Summary of Chapter

The results from Chapter 5 were broadly supportive of the 5 hypotheses. Some of the key findings I discuss in the present chapter include repeated evidence that traditional WQTMs are simply inappropriate for real world complexity in stakeholder attitudes and behavioral strategies (section 6.2). I discuss the evidence that incorporating non-traditional (explicit non-economic) incentives in a WQTM institution can achieve better sustainability outcomes than traditional WQTM incentives. Annual reductions in salt load with such incentives range from 4% to 8% relative to the empirical baseline. This is robust even to assumptions of non-static actor preferences with time (section 6.4), and that these findings can even be enhanced if non-economic incentives also exploit social dynamics (section 6.5). I also discuss the finding that even explicit economic informational feedback encouraging selfish behavior can also be effective in supporting sustainable behavior, if paired with sustainability information within the context of the WQTM institution (section 6.6). I argue that all this calls for institutional designers — planners, politicians and other advocates in the public and private sectors — to consider additional, much more innovative incentive measures in their market institutions. This will allow them to address the empirical fact that real basins incorporate a broader spectrum of stakeholder attitudes and behavioral strategies than can be accommodated by boundedly rational choice theory and institutional designs based on that theory (section 6.3). I also dig into

the data to address the LAB context, and find that the institutional interventions could lead to more sustainable basin salinity outcomes; that ditch managers have particular power to affect salinity through their management decisions; that none of the scenarios would greatly impact the social stability in the basin, and none would increase the amount of out-of-basin water transfer relative to the baseline case.

There are, however, fundamental limitations to the model that generated these results. These are, principally, inaccuracies and omissions arising from necessary simplification; systematic underestimation of salt loads relative to reality; logistical limitations on sensitivity testing; and overarching limitations on the simulation time period, due to constraints imposed by ArkAgent's integration with a separate water quality and quantity simulation tool. I review these in detail in section 6.9, but in considering the credibility of my assertions and implications, remember that ArkAgent's purpose is in testing the implications of theoretical assertions on social and environmental systems within a more complex context than a purely abstract environment. It is not a tried and tested decision support tool, although it could well become that with further work. Recall that the thesis treats ArkAgent as a tool for theoretical testing and exploratory analysis, which means the present chapter is not a discussion of what *will* happen in the LAB should a WQTM be implemented in that context. It could well be, however, a discussion of what *may* happen as long as scenario projections are understood in the context of their institutional and behavioral assumptions.

Section 6.8 draws on results from the test of each hypothesis to suggest model improvements (focusing on those which were suggested by the results, and not just those that would be a natural next step for any modeling project of this nature). I highlight the omission of the water courts system in the present model, along with unsophisticated simulation of the state engineer, as a concern for a model with aspirations to be policy-relevant. I suggest several improvements to agent behavioral logic, particularly in relation to the economic behavior of agents, how they respond to received information, and their behavior in social networks. I suggest various institutional modifications, including expanding treatment of more traditional economic topics like transaction costs; simulating municipal policies and technologies in relation to their salt contribution; and the testing of actor-specific incentive packages. An essential

next step would be to better gather data on and understand finer resolutions of social and environmental model dynamics (regional and individual levels). Better data collection on economic decisions would be helpful analytically. Critically important among all possible changes would be modifying both ArkAgent and GeoDSS to allow more than eight years of simulation. The use of better municipal growth simulation is warranted, since this impacts municipal water leasing behavior. Unfortunately, I estimate the scale of many of these changes to be considerable, particularly if further empirical work is required and if changes to the GeoDSS are necessary.

## 6.2   Hypothesis 1: self-interested maximizers versus varied and complex resource users

### 6.2.1   Summary of section

In the following section 6.2.2, I explain that the improved sustainability outcomes in scenario 1a over 1b arose from the fact that the market institution in 1a provided incentives with a better match to agent behavioral models. Much of the reduction in salt loads in 1a came from improved levels of seepage reduction by ditch managers; much of the seepage reduction came from the extensive use of polyacrylamide sealant. This comes in spite of increased diversions in S1a driven by more economically rational decision making. The contradictory results for flow and salt load (a lower salt in S1a but a higher flow, relative to S1b) indicate that a single package of incentives is limited in how many areas of sustainability it can affect. In general, scenario 1a did not conform entirely to the predictions of rational choice theory, since while S1a incorporates key elements of the rational actor model, the scenario also incorporates concessions to the real world context it simulates. This does not, however, affect the most important finding of the S1a/S1b comparison: when we introduce even a small measure of heterogeneity and complexity into a simulation, the institutional performance and sustainability benefits of a neoclassical market institution immediately starts to degrade. This suggests that incorrect behavioral assumptions *can* explain worsened performance in a water quality trading market, which is a significant finding in light of the real world lack of success with WQTMs across the

western United States.

In section 6.2.3 I discuss the fact that higher incomes in S1a are derived partly from improved application efficiencies, but mainly from increased levels of incentive payments through the WQTM. I note that this increase in payments comes at the expense of increased debt loads for some agents. There is also evidence that the total amount of quota allocated to the system may have been insufficient to meet demand. Despite this, I argue that the goals of the WQTM program as simulated may have been too lenient, and the agents may have been capable of even more salt load reduction with a shorter term program. Assuming more attitudinally variable populations (as in S1b) has the same effect as lengthening the WQTM program — it reduces the rate of sustainable change in the basin — suggesting that behavioral characteristics in a population can affect not only the overall level of change but how quickly it is achieved. A third finding is that the ArkAgent WQTM generated considerable sustainability benefits even without full participation from the basin population. A fourth point is that the WQTM was effective in communicating to agents the negative effects of salinity on the value of water. Fifth, the data on regional salt load contributions suggest that the improvements in environmental sustainability (for S1a over S1b) were distributed relatively evenly among all 20 regions of the basin. Anomalously high salt loads in the westernmost region are due to a combination of high flows, municipal discharges and upstream contributions. Overall, the data indicate an extremely complex relation between geography and salt load.

Finally, in section 6.2.6 I outline some important real world implications. Foremost among these is the implication that institutional designers — planners, politicians and other advocates in the public and private sectors — should consider additional measures to address the incentive mismatch uncovered by the comparison between scenarios 1a and 1b: institutional designers should not assume that financial incentives are enough, if they believe that the pool of potential participants for an institution departs in any way from the rational actor ideal. Of more practical relevance to the LAB system is the finding that ditch managers have especially important roles to play in addressing the salinity problem, and that the use of polyacrylamide sealing is one of their most important tools. On the basis of the finding that only the economically rational agents in S1a were any good at judging appropriate levels of participation

184

with the WQTM, I argue that institutional designers should be ready for unexpected levels and modes of participation by real world actors. On the basis of sensitivity testing which shows initial quota price choices to have significant effects on magnitudes of certain economic outcomes, I suggest that institutional designers should focus more on mitigating price volatility than finding the "right" way to set an initial quota price. I argue that institutional designers should consider more than just one dimension of sustainability when making institutional design decisions: the model results show that some design decisions impact some areas of sustainability more than others. Finally and perhaps most importantly, I discuss implications that institutional design needs to be smart in time and space: attention to geographic distribution of salt load, actor types and other basin features can reduce market inefficiencies and improve sustainability outcomes; attention to subtle variations in societal and physical conditions during any given year can lead to opportunities to enhance the effectiveness of behavioral change interventions, and may even increase the acceptability of a new institution in a conservative community.

### 6.2.2   Hypothesis 1: reviewing the arguments

Results shown in Chapter 5 that Hypothesis 1 (firstly, that the assumption of universally self-interested, maximizing resource users is incorrect, and secondly, that dispensing with this assumption can help explain poor real world performance of WQTMs) was well supported by the data. In this discussion, I compare the "orthodox" model (a single and economically-focused rationality distributed homogeneously across the agent population; scenario 1a) against the "heterodox" model (variable and not necessarily economically-focused rationality distributed heterogeneously across the agent population; scenario 1b). The comparison of scenarios 1a and 1b does indeed show more sustainability benefits for scenario 1a. Why did we see this improvement?

The answer from ArkAgent logic is that performance worsened from the rational actor case because incentives provided by the WQTM were not as well aligned with agent behavioral models. As discussed in earlier chapters, the ArkAgent behavioral model is strongly controlled by the Sustainability Factor (S-factor). Values of the S-factor are used to weight different kinds

of information and make final decisions. In the homogeneous model, all agents have their S-factors set 0. This means that they pay no attention to anything other than monetary gain and loss, acting as selfish economic maximizers in everything that they do. Of course, participation in the WQTM frequently makes good economic sense: upgrading irrigation equipment leads to less wasted water, allowing a farmer to plant more crops; upgrades also reduce salt load, which reduces the likelihood of being fined for exceeding a pollution quota. Consequently, when we assume agents are single-mindedly focused on the economic case for action, in most cases (and even assuming some component of non-rational decision making, as the model does) the agents participate eagerly and fully in the WQTM.

Conversely, when we start to vary the S-factors across agents, some farmers and ditch managers are no longer absolutely driven by economics. This is not say that economic considerations are discounted. They still play a role, since nowhere in ArkAgent does a behavioral model ignore economic considerations. But with higher S-factor values, economic considerations are weighted as less important than other goals. In the empirical case, agents have their S-factors parameterized using psychometric testing data from the real system (i.e. real stakeholder perspectives distilled into a behavioral model). This behavioral model departs considerably from the self-interested actor model: many agents now add more weight to considerations of community and sustainability. This means that when they approach a WQTM decision, it is not always in their best interest to participate.

This behavioral logic drives practical actions for reducing salt load. In ArkAgent, there are three potential ways in which salt load might be reduced: by reducing seepage through canals by lining or sealing the bottom of the canal; by reducing recharge through irrigated fields by increasing the efficiency of irrigation; and by simply diverting less water, since this means a lower volume of water can seep or recharge. Consequently, for scenario 1a we might expect diversions to be lower, and with more seepage and recharge reductions relative to scenario 1b. Figure 6.1 indicates that diversions under the homogeneous scenario were in fact considerably higher — by around 4,000 acre-feet annually per ditch — than the heterogeneous scenario. Considering the behavioral motivations of agents in scenario 1a — make money at all costs — this is not surprising. Just as diverting less water can reduce the risk of fines in the WQTM,

it can also cost farmers money by reducing the acreage they can plant and the yields they achieve. Furthermore, hyper-rational agents will be more likely to divert up to their legal decree, since they will be more concerned about personal loss if any un-diverted portion is lost due to dereliction (as per Colorado water law). See also the more extensive discussion of this in section 6.4.2.
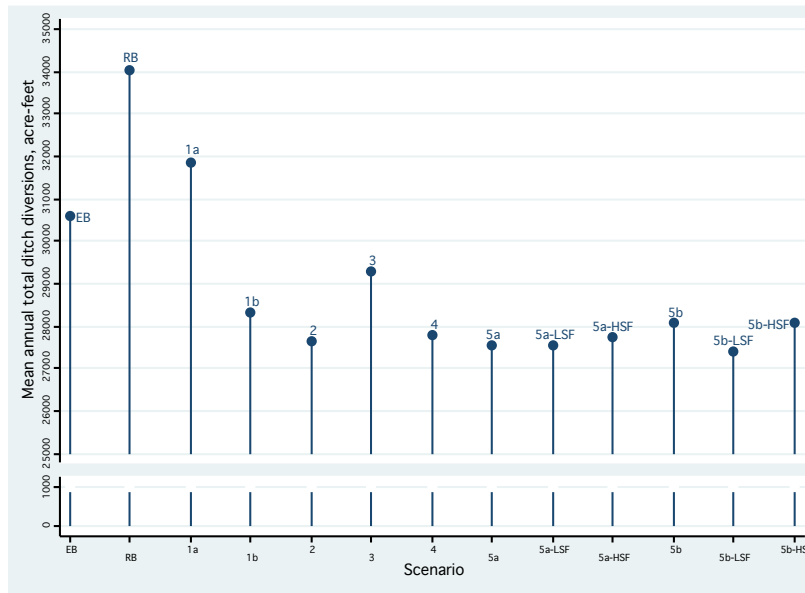


Figure 6.1: Mean annual ditch diversions (acre-feet), across all scenarios. Shows how diversions under scenario 1a were higher than scenario 1b, contra hypothesis predictions. More economically rational agents appear to divert more water than they absolutely need. Note the scale break: this is included to emphasize inter-scenario variation, since absolute ditch diversion values are less important in this context. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

Where the differences between scenarios really begin to emerge is in the recharge and seepage. In figure 6.2, we see that scenario 1a had average seepage reductions of around 40% compared with historic levels, whereas scenario 1b achieved only around 25%. A different pattern is seen in figure 6.3, where scenarios 1a and 1b only differ in mean weekly recharge reduction by around 2%, with scenario 1b coming out on top. Note that this plot omits standard deviations for more clear viewing of inter-scenario differences; figure 6.4 indicates that there was also considerable temporal variation.

A look at the time series (figure 6.5) clarifies the relationship. While scenario 1b has high recharge reduction on average, we see that scenario 1a eventually reaches parity by the end of
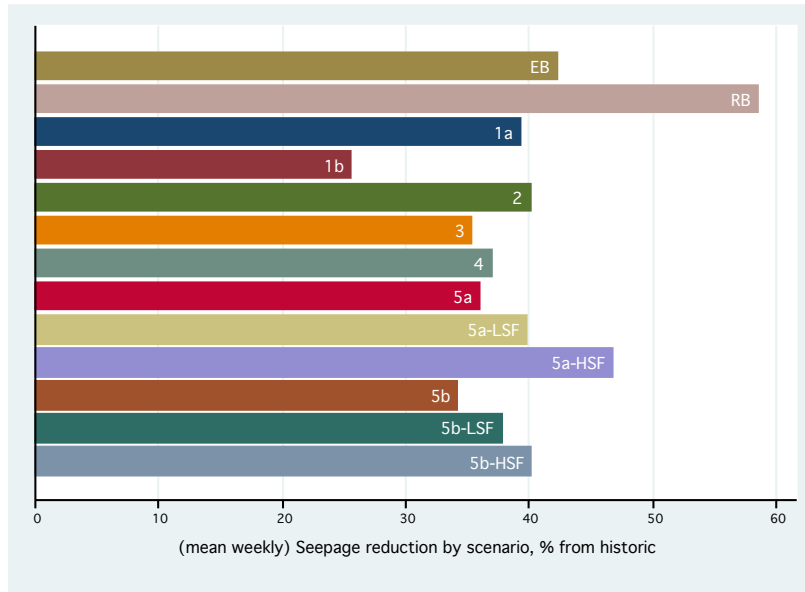
Figure 6.2: Seepage reduction, % from historic seepage, across all scenarios. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.
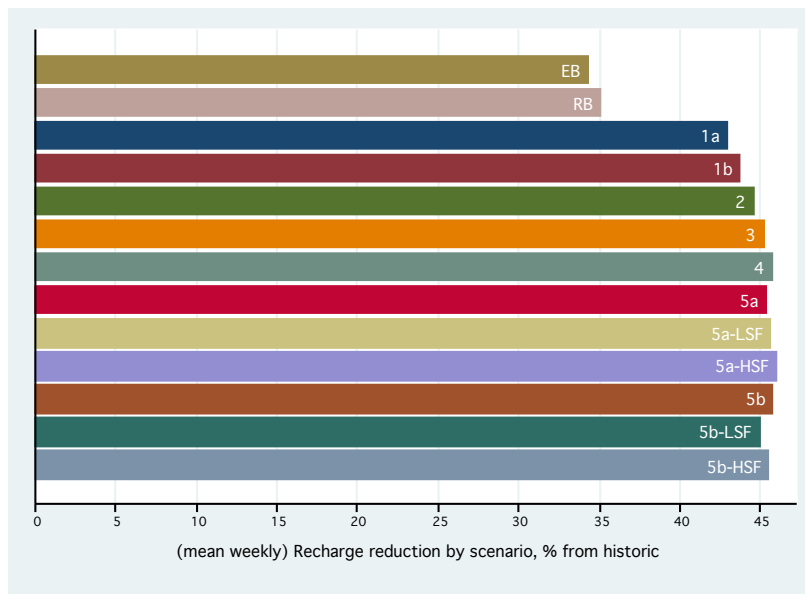


Figure 6.3: Recharge reduction, % from historic recharge, across all scenarios. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

the simulation. The salt load metric is a cumulative one, but in truth what matters most is that a longer term trend towards salt load reduction is set in place.

The detail behind the differences in the environmental sustainability of scenarios 1a and 1b, then, is that while both scenarios did not differ greatly in terms of how much recharge over
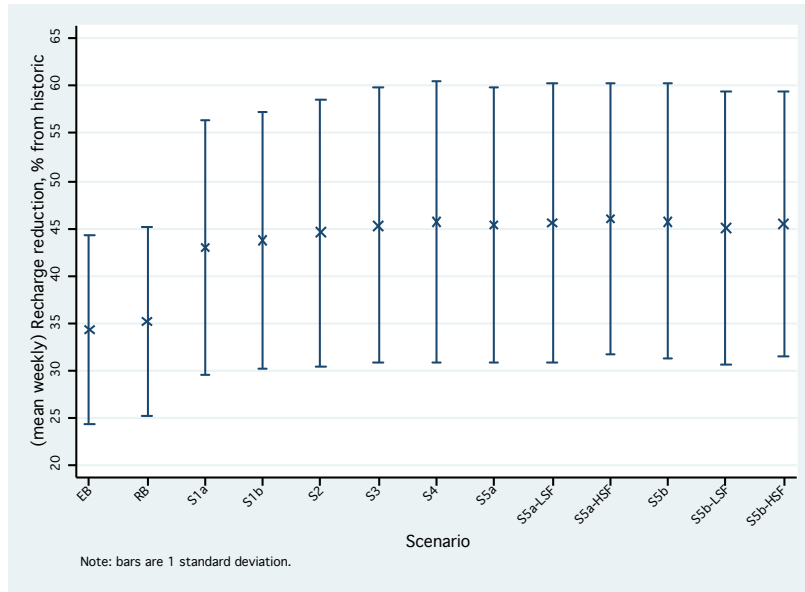
Figure 6.4: Recharge reduction, % from historic recharge, across all scenarios, including bars at 1 standard deviation. While the mean weekly values indicate relatively small average shifts in efficiency, this disguises quite considerable variation over the simulation period. This is even easier to see in figure 6.5 below. Refer to Appendix section 8.1.2 for detailed scenario descriptions.
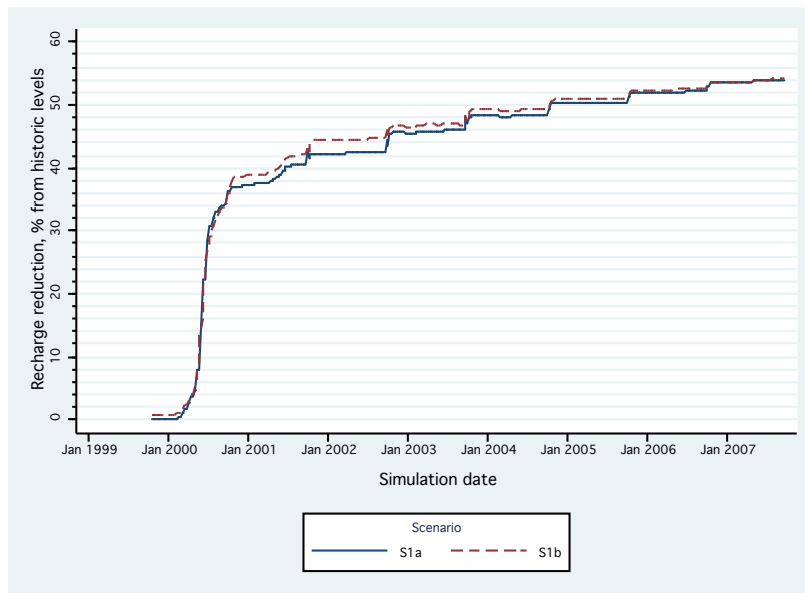


Figure 6.5: Recharge reduction, time series, comparing scenarios 1a and 1b. Note in particular how the levels of recharge reduction initially, scenarios 1a and 1b evolve together with time. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

the fields was reduced, scenario 1a was far more effective at reducing seepage and so resulted in overall lower salt loading. We see this pattern despite higher diversions in scenario 1a. An

implication of this is that it may be harder to see serious differences between scenarios when we consider only the farmers and the incentives they have to change their behavior and equipment. But when we also consider ditch managers, who have different sets of options and interests (particularly the use of polyacrylamide sealant, which is a cheap and highly effective way to reduce seepage), this can make the difference in overall salt loading. This is not to discount the importance of targeting incentives at farmers. Without their efforts to reduce recharge, overall salt loads would still be higher. But it does suggest that paying attention to ditch managers and their decision making may have significant payoffs. On this subject, recall from Chapter 4 that pollution quota is allocated initially to ditch managers who then re-allocate portions of the quota to individual farmers on the ditch based on a mean seepage/recharge ratio for the basin and individual farm acreages. The re-allocation is conducted automatically, without allowing the ditch managers to make custom allocations based on any prior knowledge they might have about individual farmers and farming practices, suggesting that the re-allocation is unlikely to be as efficient as it could be. Assuming that ditch managers would, in reality, be far more capable of tailoring allocations to farmer salt contributions (with their detailed knowledge of pre-existing salinization, for example), I suggest that even though ArkAgent shows the potential role of ditch managers, this role may be *underestimated*.

Other metrics explored in Chapter 5 included economic, institutional and social measures. The social measures showed little response to the change in behavioral model from 1a to 1b, and this is something seen across all scenarios. I discuss this separately at the end of the chapter. On the economic side, what might be driving increased incomes in S1a relative to S1b? Agent incomes are composed of (1) returns from off-farm income, (2) income from selling crops and (3) income from taking out additional debts. With regard to (1), ArkAgent does not connect off-farm income to any activities related to the WQTM, so we can discount this dimension. With regard to (2), crop incomes will go up when yields go up and/or prices per unit increase. The pricing databases are kept the same among scenarios, so any increases in income from selling crops will come from increased yields, and those are dependent on water supply and the efficiency with which it is applied. Finally, in regard to (3), the ArkAgent WQTM includes a mechanism by which fine money paid by agents polluting more than their quota, is redis-

tributed to agents not exceeding their quota. So an increase in income from 1a to 1b could conceivably come from increased efficiencies in water application, redistributed benefits from the WQTM, or incurring additional debt.

Diversions data (figure 6.1) show that, as expected, diversions are higher in 1a over 1b. However, since figure 6.7 shows that peak irrigation application efficiencies increased by around 10% from scenario 1a to 1b, it seems unlikely that there were higher returns due to improved crop yields. This leaves us with option (3). The data show that roughly $40 million more in incentives (re-distributed fines from polluters exceeding their pollution cap) were paid out in 1a than 1b. If no additional money is added to the system, redistribution of resources among agents will not change the mean: to see higher mean incomes due to fines and incentive payments, the fines must drive an increase in debt, as in fact is the case (figure 6.6).
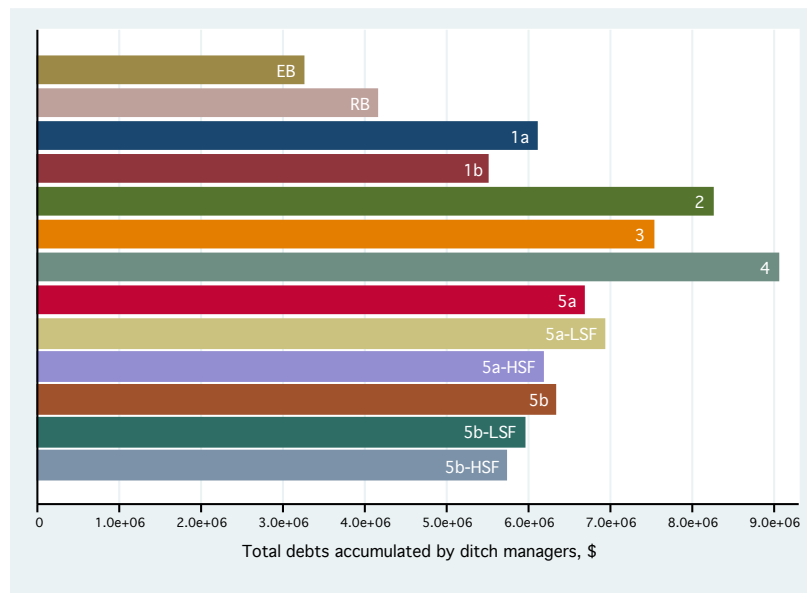


Figure 6.6: Debts accumulated by ditches, $, across all scenarios. S1a has higher debt incursion, probably due to higher fine rates for agents exceeding quotas. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

This suggests that additional income in S1a may come at the expense of increased outgoings for deficit-carrying agents, and incur higher debts for some agents. Interestingly, quota status information (mean state of agent quota accounts) indicates that higher economic rationality (in S1a) led to more delayed responses to market incentives, relative to S1b. This is explicable when we consider that market participation and salinity reduction are two of a large range

191

of economically-relevant decisions that agents make, and the more rational agents become, the more they will weigh off particular choices on economic grounds. It appears that, in this instance, it was more economically rational to take the fine for exceeding quota than addressing excess pollution directly by reducing salt load or purchasing quota. We should note further, however, that because fine amounts are indexed to quota prices, it cannot actually have been more economically rational to pay a fine versus buying more quota: this implies that the initial quota allocation to the community was too low, while reducing pollution directly was still the least economically rational option overall.
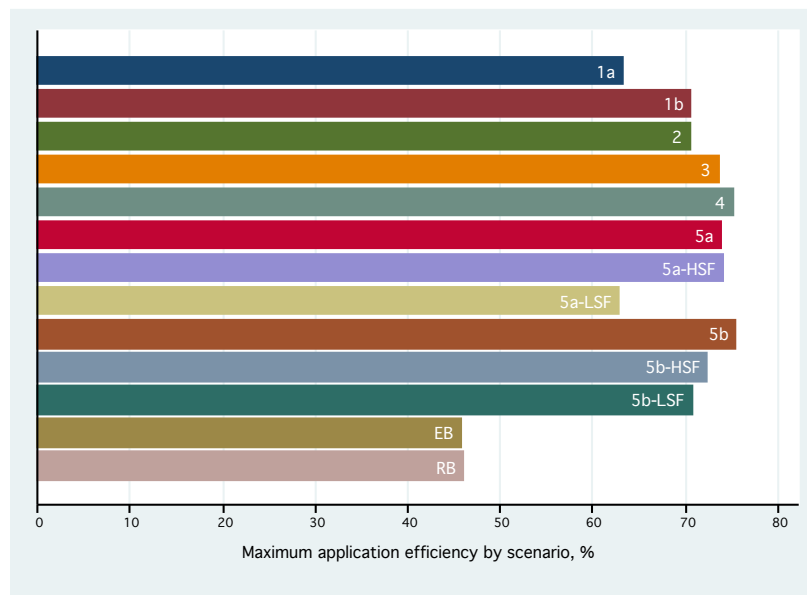


Figure 6.7: Maximum on-farm irrigation application efficiency, 0-100%, across all scenarios. Irrigation application efficiency increases close to monotonically with time, and so plotting the maximum value provides an indication of overall performance. Note how application efficiencies increase from scenario 1a to 1b. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

The Chapter 5 analysis identified that performance of the 1a WQTM was far from perfect, particularly in terms of its institutional performance. I pointed out that the market in scenario 1a was far from pareto-optimal. We should note, however, that ArkAgent is not a rational actor model as classically defined ("orthodox" models, in Chapter 2). It is full of messy reality, both environmental and social. The agents in scenario 1a are hyper-rational insofar as they single-mindedly pursue self-interest, and do not value anything but monetary cost and benefit information in decision making. However, the restrictions of modeling a real world system dictate

that they not have perfect foresight, nor have access to all the information they would need to make a perfectly rational decision. So while the rational agent market was not pareto-optimal, given the nature of ArkAgent the results are even more supportive of hypothesis 1. They suggest that even when we assume a less-than-perfect model of rationality distributed homogeneously across an agent population (S1a), this model still performs better than the population with variable rationality from individual to individual (S1b). It is reasonable to assume that, were we to simplify ArkAgent to allow perfect foresight, information access and other concessions to neoclassical theory, we would see even better performance. Conversely, the ArkAgent behavioral model is still nowhere near as complex as reality. Plenty of simplifications were necessarily made to translate stakeholder-provided material into a technically and logistically-feasible model. The ArkAgent farmers are undoubtedly far more rational and better resourced computationally than the average LAB irrigator, simply by the way that many exceedingly complex decision processes had to be boiled down into functional algorithms. Because of this, improving the model realism would be likely to generate even poorer performance for the heterogeneous population relative to the homogeneous group. In other words, it is highly unlikely that reality is simpler and more rational than the model: the model demonstrates that when we introduce just some heterogeneity and complexity into the picture, performance of a neoclassical institution immediately starts to degrade.

This raises an important question: does the degradation in institutional performance in the varied behavioral model case (scenario 1b) provide sufficient support for the assertion that more realistic assumptions would explain the widespread real world failure of WQTMs? It is not possible (and it is not my intent) to dismiss all other possible explanations for WQTM failure. Many real world cases suffer from problems due to genuine issues with transaction costs, complexities in the natural system and other issues not addressed in ArkAgent. Further, without a much more thorough and extensive exploration of all possible dimensions of WQTM design and implementation, it is not possible to conclude that poor behavioral assumptions are exclusively responsible for the failure of WQTMs in the United States. However, what ArkAgent does show is that poor behavioral assumptions can be responsible for worsened performance in a WQTM, and this is something that had not been clearly demonstrated before.

My case is strengthened by some of the results from the sensitivity testing. I used these in Chapter 5 to support the robustness of Hypothesis 1, but these results also support my present argument. The results (assuming linear behavior) show that reductions in WQTM performance under heterogeneous behavioral assumptions are robust to the length of the WQTM program when varied over 8 to 20 years; robust to the choice of initial quota price when varied between $0.1 and $10; and robust to whether or not we allow agents to bank quota and forgive their deficits after a year. Given that these are all important design considerations, we must seriously consider the idea that WQTM designs not directly related to incentive structures may have less effect on overall performance than has been assumed in the past. At the very least, the ArkAgent results certainly suggest that typical treatment of the problems underlying WQTM failure is too narrow.

### 6.2.3    Hypothesis 1: discussion of additional metrics

Output from ArkAgent came in far more detail than the 9 response factors, along with diversions, recharge and seepage reduction. Under the same categories of sustainability (institutional, economic, environmental and social), there are additional measures we can use to refine our understanding of how scenarios differed and why, and identify any important implications that result.

In analyzing these select additional metrics, I find that overall salt reduction goals set by the WQTM were realistic, and lengthening the WQTM program period made these goals easier to attain. I also find that the ArkAgent WQTM generates considerable sustainability benefits even with more limited participation in the WQTM. I note from data returned on the dollar value of water per acre-foot that the WQTM successfully embeds the cost of salt load in the value of water to individual agents. Finally, I find that breaking the salt load down by regions continues to support the overall hypothesis, while highlighting that the largest contributions came from the largest ditches in the westernmost part of the basin.

In the institutional category, one interesting metric is in the evolving status of all the trading accounts in the WQTM. A trading account was assigned to every participating agent in the WQTM, and this recorded market accounting measures like current quota availability, fines

and incentive payments. Recording overall quota availability with time allows us to see how quickly or slowly the agents in the system reacted to WQTM incentives, and also sheds light on whether the overall goals for salt load reduction set by the WQTM were realistic. Figure 6.8 indicates that agents in both scenarios 1a and 1b initially struggled to reduce their salt load, as evidenced by the fact that the average account took around 4 years to finally reach a surplus. After mid-2003, the average account status showed a slight decline each year through the end of the simulation.
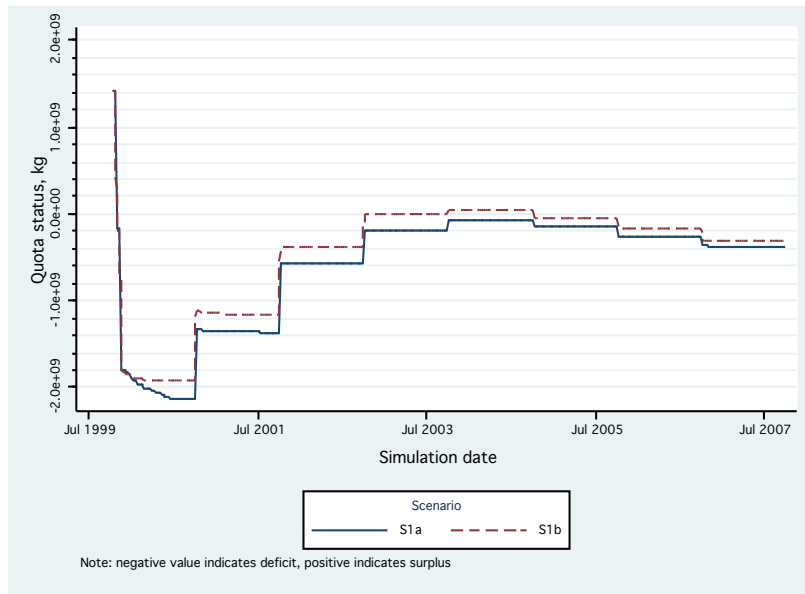


Figure 6.8: Status of quota allocations to trading agents, comparison between scenario 1a and 1b. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

The WQTM goal — reducing salt loads from Pueblo to the state line to be equal to or less than salt loads west of Pueblo — were clearly realistic and may in fact have been too lenient. Agents were able to meet salt reduction goals within four years, stay ahead of those goals for a year, and then were slightly behind salt reduction goals for the remaining three years. In the sensitivity testing, I experimented with both shorter and longer time periods. The range of response this resulted in is depicted in figure 6.9.

In this figure we see the basic trend reproduced (scenario 1a with more deficits than 1b), but shifted: in the shorter term market, the deficits are more extreme; in the longer term market, the deficits are less extreme. If we review the time series data for the same two scenarios, varying

Figure 6.9: Status of quota allocations to trading agents, parameter variation by program length, comparing scenarios 1a and 1b. Indicates how increasing the length of the WQTM program can make it harder for agents to meet their annual pollution reduction targets. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

the program length (figure 6.10), we can identify the overall effect of lengthening the program: agents adjust to salt reduction goals quicker, and are able to keep up with those goals for longer.

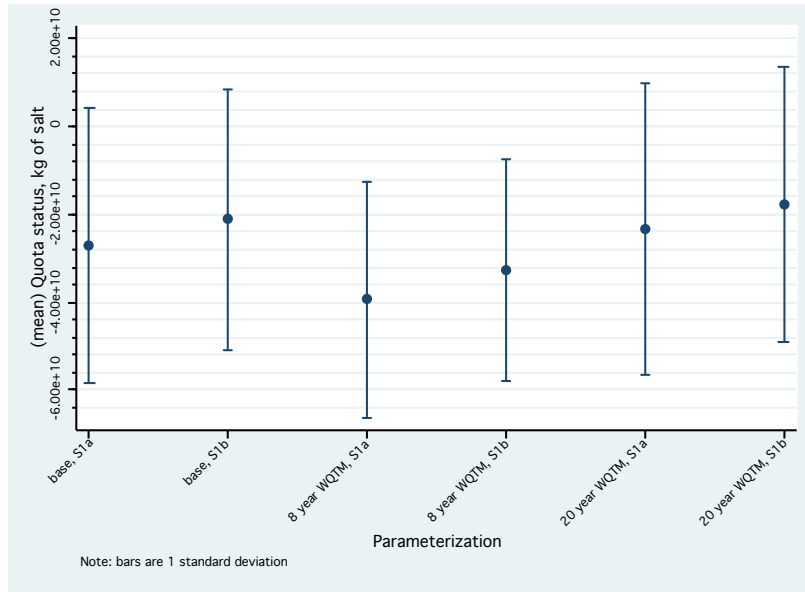We see these effects because while the overall goal of salt reduction does not change, the length of the program determines by how much the assigned quotas will be reduced each year. The shortest programs have a much larger inter-annual decrease in allotted quota, putting much more pressure on agents. The reverse is true for the longest programs.

A final element to consider on the subject of quota status is the systematic differences between scenarios 1a and 1b. While the trends are similar, scenario 1b typically shows lower deficits than 1a overall. The effect of using a heterogeneous behavioral model is akin to lengthening the program: deficits are less extreme, and surpluses last for longer. The reason for this may lie, again, in the fact that decisions relating to the WQTM take several forms, and incentives can combine in unusual ways to achieve a goal. Highly self-interested agents (in scenario 1a) clearly value reducing salt load by upgrading irrigation equipment and lining or sealing canals. However, it appears they are also minimizing their expenditure on quota. In other words, self-interested agents need not necessarily be fully bought into the idea of the WQTM
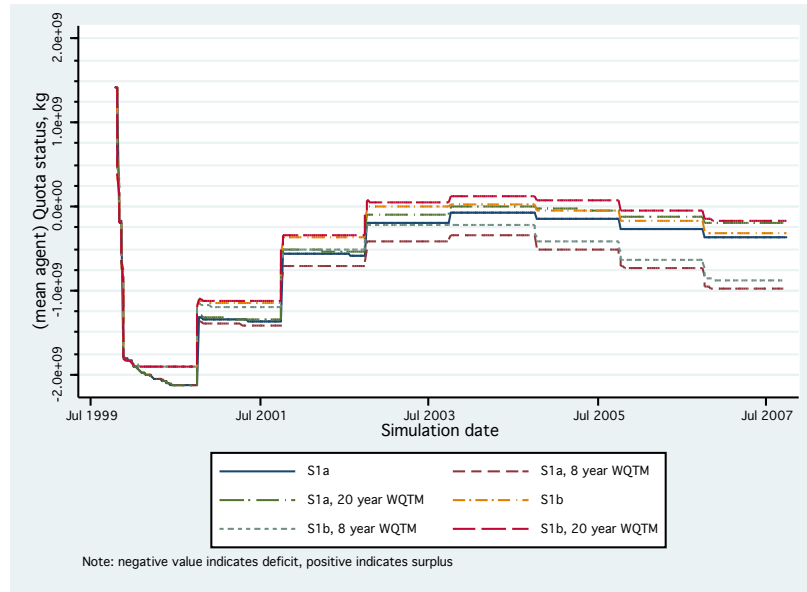
Figure 6.10: Status of quota allocations to trading agents, time series, parameter variation by program
length, comparing scenarios 1a and 1b. Shows that in the scenarios with longer WQTM pro-
grams, agents adjust to salt reduction goals quicker and keep up the salt reduction achieve-
ments for longer. See text for discussion and refer to Appendix section 8.1.2 for detailed
scenario descriptions.

to make salt reductions: they just need a push in that direction to provide added incentive to
actions which are already good business sense. On the other hand, socially and sustainably
oriented agents are aware of and value other benefits to participating in the WQTM (and so are
more enthusiastic about reducing their deficits), but may not be as single-mindedly focused on
the economic benefits when considering investments in equipment upgrades and ditch sealing.

In the economic category, ArkAgent also tracked the mean value of water across all agents.
This measure embodies the value returned to the farming agents from their use of an acre-foot
of water. In the non-WQTM (baseline) scenarios, this is the use of water for irrigation and/or
sale to municipalities. In the WQTM scenarios, the value of water will also reflect any cost
incurred by the salt load inherent in using that acre-foot of water. Figure 6.11 describes the
evolution of water value for farmers over the time series.

In the time series, we see that scenario 1a mostly enjoys higher peaks of water value. The
raw data shows that scenario 1a has on average a higher mean $ per acre-foot value ($6067.29)
relative to 1b ($5650.55). Water value in both scenarios appears to follow a seasonal trend, with
spikes during the harvest (when the most value is returned, since irrigation is the primary use),
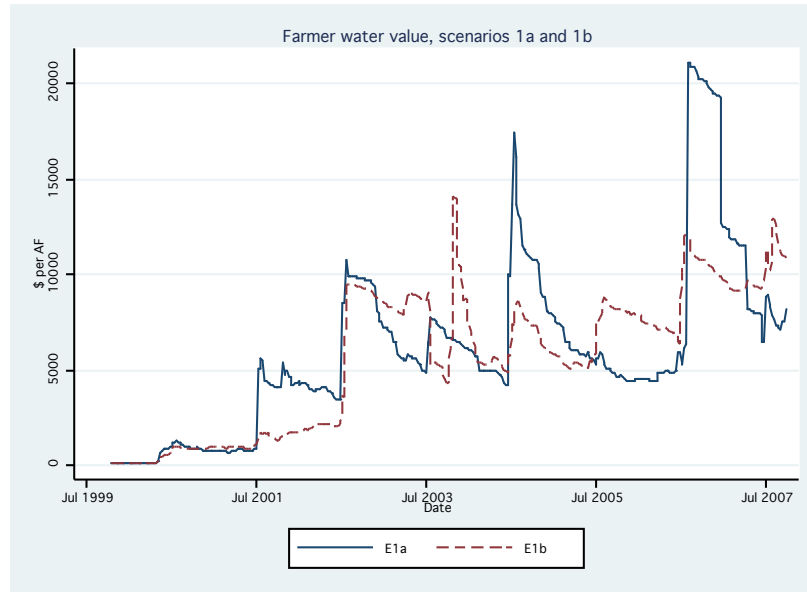
Figure 6.11: Agricultural value of water ($ per acre-foot), comparing scenarios 1a and 1b. Scenario 1a sees higher means and higher peaks of water value, with water value following a seasonal trend related to agent activities and WQTM dynamics. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

and possible secondary increases during the year associated with WQTM dynamics. The water value results clearly support the assertion that the better performing scenario in economic terms is 1a, and it is likely that the reduced salt load allows an increase in water value. This implies that the WQTM mechanisms do a reasonable job of making salt load costs more tangible to stakeholders. As discussed in Chapter 2, salinity build up is often treated as an intangible externality, difficult to quantify and consequently difficult to incorporate into the market system. Evidence shows here shows that the WQTM can help address this problem. The seasonality in the value of water with time suggests that there may exist unexploited potential for improving sustainability outcomes. It indicates that water is not always equally valuable to all agents: at some times of the year, it becomes much more valuable — like crucial irrigation months — suggesting that at these times institutional mechanisms (market or otherwise) would have greater leverage in encouraging farmer/ditch manager improvements in irrigation efficiency and reductions in seepage/recharge. I discuss these further, along with the broader issue of temporal dynamics in institutional mechanisms, in section 6.2.6 below.

The discussion of environmental metrics so far has not touched on geographic variation.

ArkAgent was limited in the extent to which geographic variation could be captured, but salt load contributions by region was one of the few metrics disaggregated in this way. As discussed in Chapter 4, the GeoDSS model simulates salt load in return flows from 20 discrete regions. Each region was defined according to a combination of empirical data and technical need. Region 0 is the westernmost region, while region 19 is the easternmost. We can break down the regional salt load data to show overall contributions by region, organized by scenario (figure 6.12).
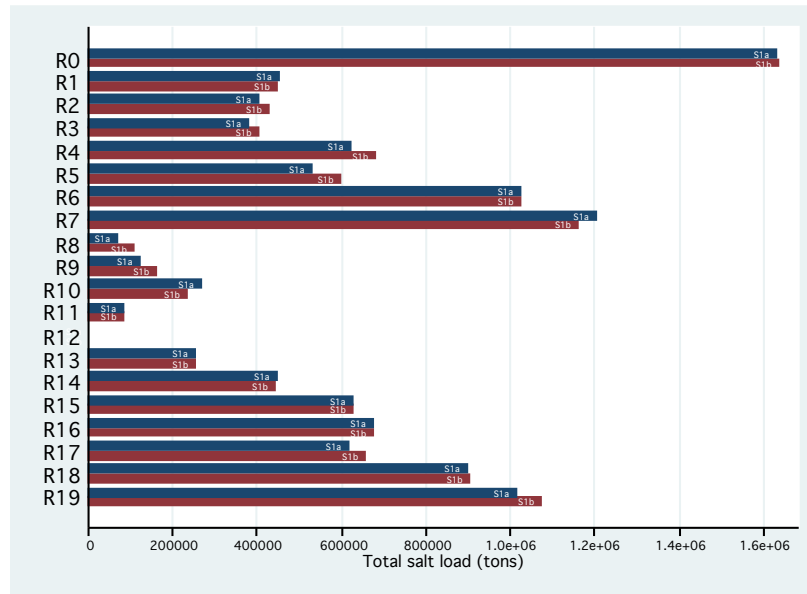


Figure 6.12: Salt loads (tons) contributed to river by basin regions, broken down by basin region, comparing scenarios 1a and 1b. Across most regions, scenario 1b shows an equal or higher salt loading than scenario 1a (exception regions 7 and 10), providing support for the hypothesis. High loads in region 0 probably relate to higher flows from Fountain Creek, and additional salt contributions from municipal outflows. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

As far as Hypothesis 1 is concerned, the most useful information in this plot is that across most regions scenario 1b shows an equal or higher loading than scenario 1a, with notable exceptions in regions 7 and 10. It is unknown what causes the exceptions, but the overwhelming trend is clear. However, there is more interesting material here, raising the issue of how to handle geographic variation in these complex systems.

The regional contributions vary considerably, with the highest loading from region 0 and the lowest from region 13. Region 0 has a particularly high load for several reasons: first,

this region is associated with high flows, since there are few major diversions upstream of this region, and there are large contributions from Fountain Creek (pers. comm. Temeepattanapongsa 2010). High flow rates can drive up absolute salt load, since higher quantities of water can carry more salt. Second, there are several major sources of salt that will contribute to higher loads. Historically, around 35% (and occasionally up to 50%) of salt in the river at Avondale (east of Pueblo and in the center of Region 0) was contributed by Fountain Creek. Around the same contribution comes from the mainstem river (contributions from upstream ditches and the Pueblo Board of Water Works treatment plant) (Ortiz, 2004). The Fountain Creek salt is derived from various municipal treatment outfalls, and from natural loading in return flows. Finally, the location within this region of the Bessemer Ditch (a senior ditch with high diversions; see figure 6.13, noting that the figure region 1 = the plot region 0) is also likely to increase the load. Contributions from Pueblo Board of Water Works, ditches in the Middle and Upper Arkansas basins, and municipalities along Fountain Creek are not addressed in ArkAgent. The GeoDSS is not capable at present of simulating either volumetric changes to Fountain Creek inflow, nor changes to its salt load due to changes to municipal treatment technologies or practices; the same goes for changes at the Pueblo municipal discharge. However, the plot and the map indicate that a real world institution should either provide allowances to ditches (the Bessemer is a good candidate) close to the inflow point, so that they are not unfairly penalized for the salt load contributions of Colorado Springs and other municipalities. Alternatively (and this is more ideal), the municipalities of Fountain Creek and Colorado Springs should be included in the water quality trading market, and the trading market area should be extended to the Middle basin (Canon City to Pueblo).

Apart from region 0, contributions vary so significantly among regions because there are extremely complex relations between region size, location, ditch size, location of diversion points, diversion amounts and timing. Some general trends can be explained, however. We likely see a reduction in salt load contribution from the middle basin regions (R8-12 in the plot; R9-13 on the map) because there are fewer ditches passing through these regions, and because the size of these regions is smaller in relation to the western regions. The monotonic increase in load as we move east is probably due to the steady accumulation of salt moving east, since all the
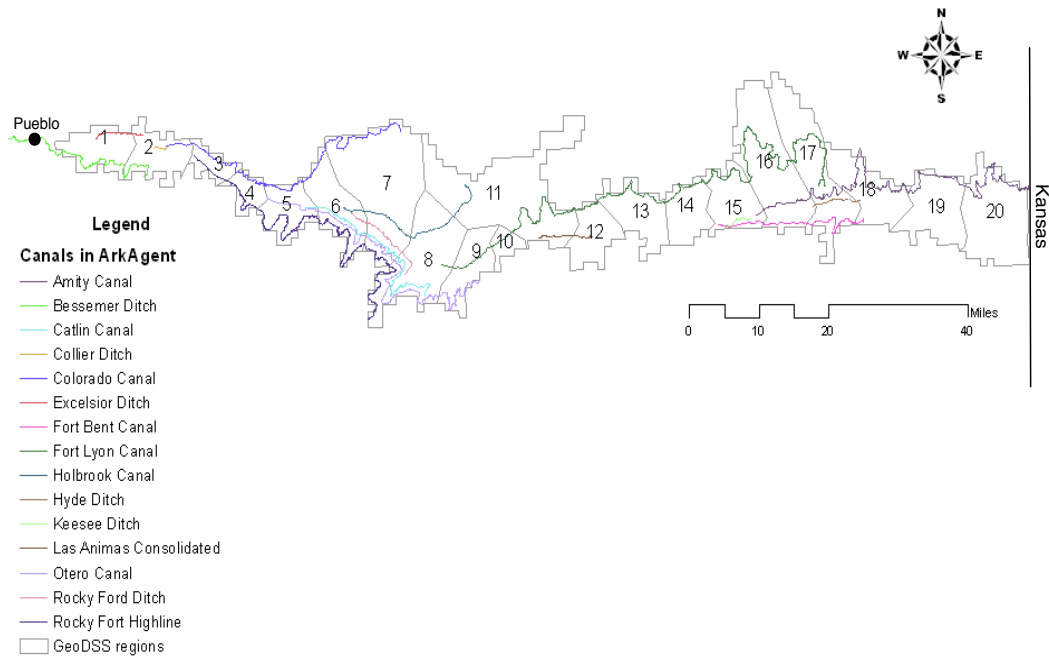
Figure 6.13: Regions of the lower Arkansas Basin used in tracking regional loads. Ditches also labeled. Note that region 1 in this map equates to region 0 in figure 6.12. Note how more ditches are clustered in the western portion of the basin, and how basin regions vary in size. Salt loads from individual regions are affected by the complex interrelations between ditch diversion volumes, length, size of the region and number of ditches within a region. The ArkAgent WQTM tries to take account of the ditch miles in each region, to ensure balanced quota allocation.

regions are roughly the same size. Ditches further east have smaller and less senior diversion rights, but the salt load is cumulatively higher because these are downstream in the system. In other words, more pollution is returned to the river in more eastern regions per unit of seepage and recharge. The plot and map provide support for the use of trading ratios (which were included in ArkAgent), so that the eastern ditches are not unfairly penalized for the loading of other agents. Once again, both figure 6.12 and the map indicate the importance of building appropriate geographic adjustments into the WQTM. ArkAgent's WQTM included trading ratios and adjusted quota allocations for each region according to the total mileage of ditches and

total diversion volumes in each region. This is still a relatively coarse approach to handling the obvious geographic complexities, and as I discuss in section 6.2.6, real world institutional designers should probably go further.

### 6.2.4 Hypothesis 1: anomalies

One anomaly of particular interest is the relatively poor performance of the WQTM in scenario 1a, in terms of clearing success ratio (CSR). Scenario 1a had a CSR of around 25, indicating that, on average, 25 times as much quota was posted for sale versus what was actually cleared from sale. This does not seem ideal. Why should the most economically rational agents have such poor performance relative to the predictions of neoclassical theory? The first thing to remember is that these are not orthodox *homo economicus*. There are plenty of intentional and unintentional elements of the agent logic which result in less economic rationality rather than more. Even with high levels of economic rationality, then, we still see flawed market performance. Second, note that the ArkAgent WQTM is designed such that posting quota to the market for sale incurs minimal costs, financial or otherwise. No doubt the incentives (and CSR results) would be altered if posting quota cost some small fee each time: this would encourage traders to consolidate their sale offers, and also perhaps limit over-posting. Agent logic is also limited in the extent to which it provides negative feedback on unsold quota. Agents are equipped with a variable tolerance for failure to sell quota over time, but clearly this tolerance is either set too high in ArkAgent, or is simply insufficiently instrumental in affecting outcomes.

A second anomalous result is the dominance of lining over sealing in scenario 1b. As figure 6.14 indicates, scenario 1a displays an expected result: the cheaper and quicker method of sealing using polyacrylamide (PAM) dominates (note that the sealant decays over time due to natural degradation, and must be renewed each year). Ditch managers in scenario 1b, however, never seal more than 10% of the ditch miles, while 40% of ditch miles are lined with concrete by the end of the simulation. One explanation is that the more diverse attitudinal orientations of scenario 1b agents led to less economically rational decision making: on purely economic grounds, applying PAM is the natural choice. But ditch lining also has its benefits, particularly a longer lifespan and incidental benefits in reducing ditch maintenance costs. It is also the

option ditch managers will be more familiar with, and the more heterogeneous set of interests

among ditch managers in scenario 1b will allow such considerations to increase in importance.
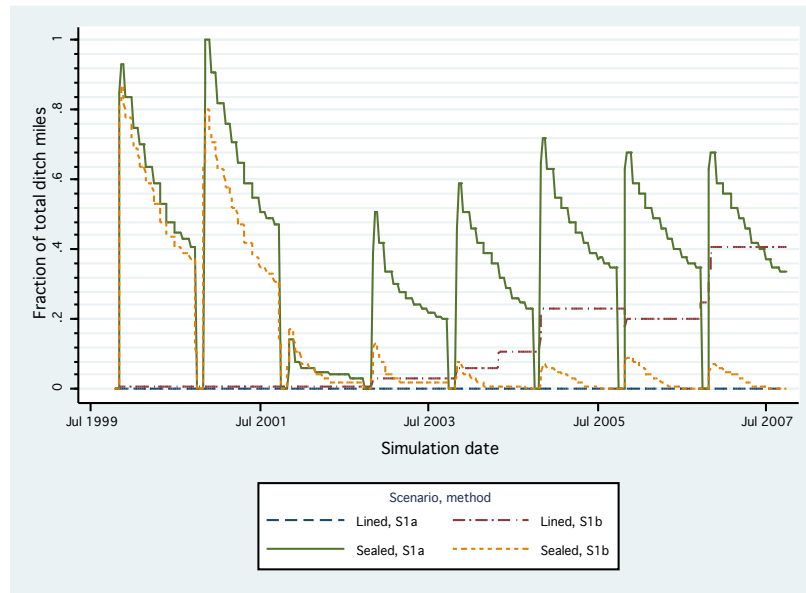


Figure 6.14: Fraction of total ditch miles lined and/or sealed, comparing scenarios 1a and 1b. Plot shows that the cheaper and quicker method of polyacrylamide sealing dominates. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

### 6.2.5   Hypothesis 1: discussion of parameter variation

Results of sensitivity testing from Chapter 5 highlighted some important findings.

1. *Specific effects of parameter variation on salt load*: varying the parameters does not greatly shift the range of responses seen with time, but does tend to alter the mean load. Variation in levels of risk tolerance towards either risk-seeking or risk-avoiding ends of the spectrum appear to either drive up or drive down salt loads. This is likely due to the fact that more cautious agents tend to consider longer term planning horizons, and so will amortize benefits over longer periods. Variation in farm skill has straightforward effects: low skill drives up loads, high skill reduces them. This is because many of the potential salt reduction activities farmers can engage in depend on a general level of skill: the higher the farming skill, the better the agent will perform a particular activity. Salt load actually appears to be broadly robust to changes in the length of the WQTM program,

although varying the program length away from the 15 year baseline appears to allow greater potential reductions (in the 1a scenario) and peaks (in the 1b scenario). This is an interesting result without a clear cause. We might expect a longer program to show lower reductions with time, since longer programs have smaller year to year reduction targets, but this is clearly not the case. Finally, varying the initial quota price (IQP) up to $5 appears to increase salt loads with time, while upping the price to $10 results in lower salt loads with time. The $10 result is to be expected, since higher quota prices provide more incentives for agents to make changes to their irrigation and ditch management practices. But is not clear why $5 prices result in higher salt loads. It does suggest that the salt load reduction response is a more complex function of the IQP than first thought.

2. *Effect of initial quota price (IQP) on institutional performance*: increasing the IQP appears to dramatically worsen the problem of a quota glut on the market, although the data show that regardless of IQP this problem diminishes from 1a to 1b. The cause of this lies in the agent logic for posting quota, which relies heavily on prospective returns as a measure of potential benefits from participating in the market. Higher initial prices simply increase the volume of quota posted to the market, and reduce its overall efficiency.

3. *Effects of IQP on agent financial health*: it is little surprise that higher IQPs result in lower agent assets. Since fine levels are tied to quota pricing, initial fines for agents will be much higher for higher IQPs. While we might expect the quota price to shift over time from the initial State Engineer-set value, this may be enough to cause significant declines in financial health across the agent population.

4. *Broad support for Hypothesis 1*: the parameter variation appears to indicate that the assertions of Hypothesis 1 are robust to variation in the various social and economic parameters included in the experiments.

### 6.2.6   Hypothesis 1: real world implications

1. Traditional WQTMs do not operate well in simulations where actors have diverse attitudinal and behavioral characteristics. Institutional designers, in addition to considering

the many things that can go wrong with the standard WQTM, should consider additional measures to address the mismatched incentives problem. It is simply not practically effective to foist neoclassical economic theory on complex basin settings. Since most basin settings in the American West are considerably more complex than the LAB, this implication has broad relevance.

2. Even in the absence of a WQTM, a targeted program to encourage PAM usage in the LAB — particularly in the larger canals — may go a long way toward reducing salt loads. Scenarios 1a and 1b show that ditches have a significant role to play in reducing salt loads in the LAB, and that they have a relatively straightforward option: spreading polyacrylamide (PAM) to seal the base of the ditch and reduce seepage. Although the sealant typically has to be renewed each year, it is still far cheaper than lining the ditch with concrete (Susfalk et al., 2008). The model actually shows that concrete lining is preferred by managers in scenario 1b. Some part of this may be model overestimation, but the strong trend suggests it cannot entirely be due to structural error. The results suggest that while PAM is clearly the most rational economic choice, targeted interventions may have to work to either support or overcome pre-existing preferences for other options like concrete lining.

3. Institutional designers may expect a lack of liquidity not because a WQTM is incorrectly implemented, but because real actors do not participate in markets in the same way a population of homogeneous economically rationally agents do. Results from scenarios 1a and 1b highlight a problem with excessive market liquidity: too much quota put up for sale by too many agents. The situation found more typically in real WQTMs is a severe lack of liquidity. Part of this problem is a modeling issue, since the logic could be improved to improve the sophistication with which agents handle posting decisions (see the discussion of further work in section 6.8). But much of this problem arises from a rational actor issue: it is entirely rational for self-interested agents to keep posting quota if there is little cost and they can potentially gain from doing so. Such "rationality" is not seen in the real world, but is often implicitly assumed by policy makers. Since most, if

not all, real world settings will have less economically rational actors than the neoclassical ideal, the implication is that in most cases WQTMs can expect reduced liquidity relative to theoretical predictions.

4. Institutional designers may need to be concerned with the length of the WQTM program, since shorter programs can put unnecessary economic pressure on actors. Results from sensitivity testing showed that increasing the length of the WQTM program reduced the overall deficits in quota suffered by agents, which in turn reduced the fines levied on these agents and increased the sustainability of the institution. Real world WQTMs will have their program goals and periods set by physical context, as well as the social and economic context. But if at all possible, institutional designers should take heed of the implication from these results that shorter programs, while reaching targets sooner, may put difficult economic pressures on participants. This implication may be relevant to any context where the basin agricultural community has limited economic means and is dominated by smaller family farms; it may be less relevant to large agribusiness systems with significant financial resources.

5. ArkAgent results show initial quota price (IQP) to have significant magnitude effects on economic and institutional dimensions of the WQTM. However, the fact that changes in IQP affect magnitudes more than directions of outcomes, coupled with the lack of a consensus among economists on the right way to arrive at initial prices for pollution permits (e.g. Cramton & Kerr (2002); Stavins (1995)), suggests that institutional designers may want to pay more attention to mitigating price volatility than setting the "right" initial price. The IQP sensitivity testing does not qualitatively change outcomes — Hypothesis 1 is still broadly supported — but does affect magnitudes. This indicates that the ArkAgent WQTM is effectively using price signals to alter behavior (Newell et al., 2005), but also that the initial choice of price can cause significant shifts in the size of financial costs and benefits. In theory, the initial price should be at the marginal abatement cost of pollution (Sartzetakis, 1997; Tietenberg, 2006), but in the LAB setting the marginal abatement cost is not easily discoverable. One alternative is to auction off quota and allow

the market to set the price, but auctions are fraught with complexity (Cramton & Kerr, 2002), and in the LAB case may not even be effective (since the participants in the market will not be clear themselves on how to estimate abatement costs). In such a case, there may be no "right" way to price quota initially. Since pricing does appear to be volatile during all simulation runs, designers should considering including institutional mechanisms for use if pricing proves to be volatile (Jacoby & Ellerman, 2004). Highly volatile prices drive up fine amounts — which may have more of a damaging economic effect in the real world than ArkAgent is able to simulate — and so should be avoided. ArkAgent simulation results suggest it is less likely that prices will be too low to drive sufficient pollution reduction. Several real world pollution trading institutions have suffered from price volatility (Convery & Redmond, 2007; Ellerman & Joskow, 2008), as evidenced by the debate over appropriate ways to address volatility. However, I cannot argue that the reasons behind price volatility in ArkAgent extend to real world settings dissimilar to the LAB, particularly since the experience in real world WQTMs is so limited as to provide little evidence of any pricing dynamics at all. This underlines the fact that ArkAgent was not designed to test alternate theories of initial quota pricing. There is a much more extensive and more technically accomplished debate extant in resource economics, and so I encourage the interested reader to seek out these sources.

6. The seasonally and even sub-seasonally variable dollar value of water with time indicates that institutional designers need to play closer attention to the potential for using temporal dynamics to further sustainability goals. Since all irrigated systems in the American West have strong seasonality in water deliveries, this implication is relevant to a wide variety of settings. The fact that water is more valuable to farmers than at other times could be used to "temporally target" institutional interventions. For example, it is likely that farmers will be more receptive during the irrigation season to efficiency-related interventions, since the value of saving water will be most clearly demonstrable at that time. Conversely, farmers may be more supportive of the overall goals of the institution if they know that institutional managers are able and willing to adjust quota allocations during the year.

7. Results for the geographic distribution of salt load underline the importance of using institutional mechanisms which handle regional variation either implicitly or explicitly. Since the kind of geographic variation seen in ArkAgent is a feature of most if not all basin settings, this is an implication with relevance in a wide variety of contexts. ArkAgent's WQTM scales quota by ditch size, which accounts for some of the regional bias, but also uses trading ratios to avoid penalizing farmers and ditch managers in the east for the salt load contributions of water users in the west. But ArkAgent does not handle geographic variation arising from heterogeneous geology, economic conditions, actor locations, actor contributions and other elements. The ArkAgent results highlight that institutional designers should be creative in assessing and targeting geographic variation. Results for geographic variation also suggest that any WQTM should include Fountain Creek and Colorado Springs as responsible polluters and traders.

## 6.3 Hypothesis 2: adopting non-economic incentives to exploit behavioral variation

### 6.3.1 Summary of section

In the following discussion (section 6.3.2), I explain that scenario 2 showed better sustainability outcomes than scenario 1b because the more diverse incentives provided by the S2 WQTM allowed the targeting of a broader spectrum of behavioral models than the S1b WQTM. This is particularly important when we consider that the ArkAgent logic did not guarantee an agent response to feedback information provided by the WQTM. Ditch managers were once again critical actors in salt load reduction, and that the informational feedback provided in S2 did a better job of incentivizing the cheapest and quickest means at their disposal — application of polyacrylamide sealant — for salt load reduction. I discuss the economic benefits of the WQTM in S2, showing that higher effectiveness of the S2 WQTM meant higher fines for polluters and income redistribution to non-exceeding agents. However, I show that this increased overall debt accumulation in the population, but since the high rate of debt accumulation was limited to the first few years of the simulation, I argue that the overall economic picture was positive. I

also highlight that S2 increased revenues from water leased to municipalities. Exploring results on this theme show that S2 makes more water available to agents, so they are more willing to lease surplus rights to the cities. This result compounds the benefits of other salt load reduction efforts: more water leased to cities in effect means water left in the river (at least in the short term), further reducing the salt loads from irrigation. Finally, providing non-economic informational feedback can sometimes work against the overall sustainability goals of a market institution, since ArkAgent results show that — on occasion — these generated behavioral choices contributing to institutional inefficiency.

Following findings from the discussion of hypothesis 1, section 6.3.4 highlights the finding that decision decisions on initial quota pricing have, more than any other institutional parameter, serious effects on the magnitude of change (both good and bad) that the market institution precipitates. In the same institutional design vein, I show how the ArkAgent results highlight the need to keep institutional controls tight: simulating a relaxation of trading and deficit controls resulted in reduced institutional efficiency.

Finally, I offer some real world implications in section 6.3.5. Principally, ArkAgent results show the considerable potential for including explicitly non-economic incentives along with the more traditional implicit economic incentives that a WQTM provides. I argue that we can go even further, and develop non-economic incentives targeting specific stakeholder groups which may have particular capacity for fostering more sustainable change in a basin. Another implication of particular relevance to any basin with salinity problems where agriculture-to-urban water transfers are common (several Californian basins fit this description), is that leasing activity can actually enhance sustainability outcomes, if delays in municipal use result in less irrigation diversions. More relevant to the LAB context in particular is the finding that water leases do not necessarily increase under the alternative WQTM trialled in S2, which might allay political fears in the community about water moving out of basin. Finally, the discussion in section 6.3.5 highlights the phenomenon of path dependency in finding solutions to a natural resources management problem: each decision made closes off (or makes less likely) other future decision options, so that a single initial intervention can force later system evolution down a unique and not necessarily reversible path. The implication is that institutional design-

ers should be aware of path dependency, while ArkAgent shows it is possible to simulate and explain the phenomenon.

### 6.3.2 Hypothesis 2: reviewing the arguments

In Chapter 5, I showed that the data supported the overall hypothesis 2 (adding appropriate behavioral incentives to a market institution can improve its performance when our assumptions on population behavior depart from neoclassical assumptions). In particular, the data showed that scenario 2 (the WQTM with incentives added) performed considerably better in terms of salt load and river flow than scenario 1b (the WQTM without incentives). How and why did this occur?

Recall from Chapter 3 that the post-hoc informational incentive messaging (PHIIM) system in ArkAgent works by returning informational feedback to agents after every particular action they take relating to the WQTM. Consider the example of when an agent purchases some quota. The following time step the agent will receive an estimate of the sustainability benefit accruing to the system had every other participant in the system taken a similar level of action in that previous time step (the State Engineer generates these estimates using canned simulation scenarios at various levels of seepage and recharge). The sustainability rating contained in the feedback message acts on the agent's perception of the sustainability benefit for whatever action is being rated, generating a new sustainability benefit based on the mean of the old benefit and the received rating. The results from scenarios 1b and 2 indicate that this can be a powerful mechanism, improving the ratings of WQTM-related actions, and increasing the likelihood that the agent will in fact take that action. If we examine the evolution of the mean internal sustainability rating for selling quota (figure 6.15), for example, we see a steady increase over time, no doubt contributing to both the strong interest in posting quota to the market and the overall sustainability of the system. As the trace shows, sustainability ratings are liable to some fluctuation during the simulation, as the relative benefit of a given action is computed using historical data appropriate for the time of year when the action was taken. At different times of year, in other words, the single action of selling quota, or upgrading irrigation equipment, or some other WQTM-related action, will vary in how effective it is in reducing salt load.
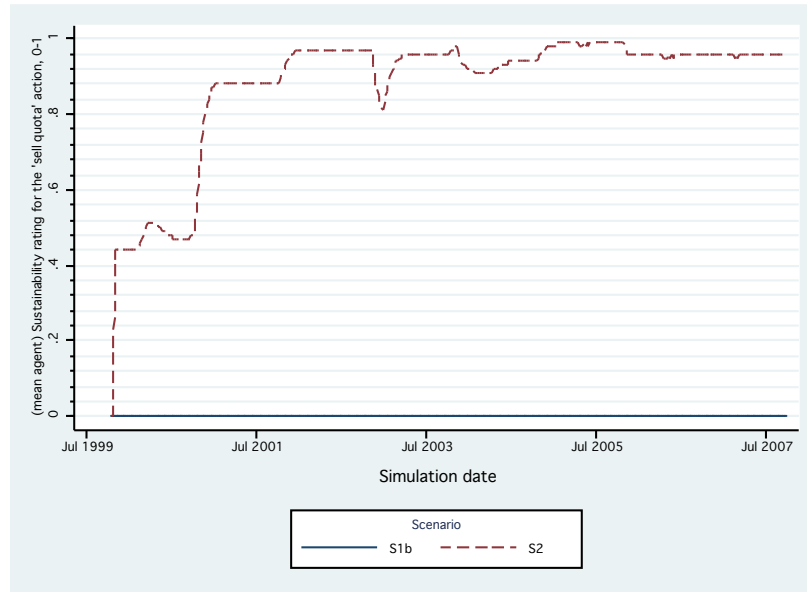
Figure 6.15: Mean sustainability rating among agents for the "sell quota" action (0.0-1.0), comparing scenarios 1b and 2. Indicates how the rating for the action increased with time, as more and more sustainable feedback was returned. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

Note that the PHIIM response mechanism in agents included a chance of non-response. As real world data has shown, all agents are not equally susceptible to influence through the provision of feedback, and ArkAgent reflects this. In addition, even if a particular action gets a higher sustainability rating through averaging with the feedback rating, this does not necessarily improve the chances that this action will be taken. The final decision weighting always depends on the agent's own S-factor. If the agent has a low S-factor, it makes little difference whether the pending action has a sustainability rating of 0 or 1, since it will receive little weight in the decision. Consequently, Hypothesis 2 is robust in the face of both a non-response percentage, and the varied predispositions of individual agents.

While Chapter 5 showed that institutional performance did not greatly vary from scenario 1b to 2, there is evidence the PHIIM system worked to enhance the attractiveness of certain actions. Which actions in fact saw greater attention? Recall that the only physical ways salt load can be reduced is by reductions in seepage, recharge and the amount of water pulled out of the river (which also relates to seepage and recharge by controlling the amount of water flowing in the ditch and spread over the fields). Figures 6.16 and 6.17 indicate that, as we might

expect, reductions in seepage and recharge are higher in scenario 2 than scenario 1b.
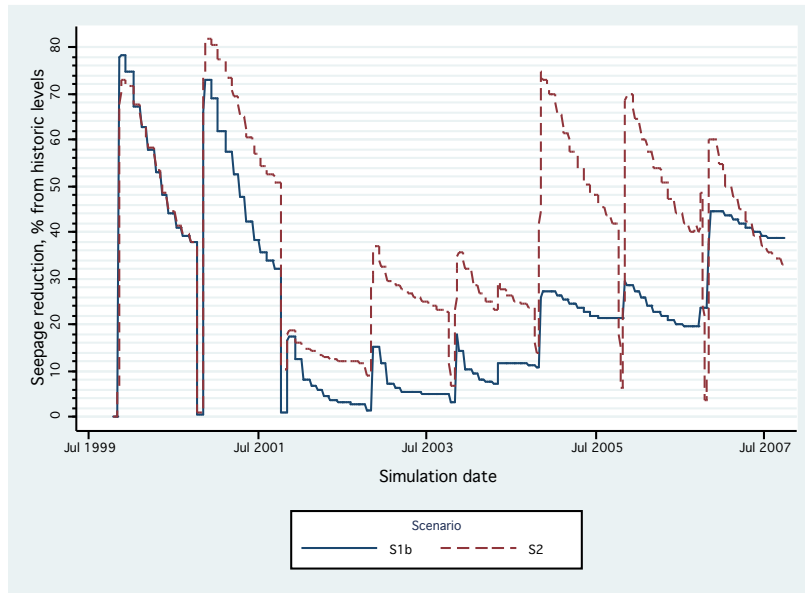


Figure 6.16: Seepage reduction, % from historic seepage levels, time series, comparing scenarios 1b and 2. Reductions in seepage and recharge are higher in scenario 2 than scenario 1b, supporting the assertion in hypothesis 2 regarding the effectiveness of non-economic incentives with behaviorally heterogeneous populations. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

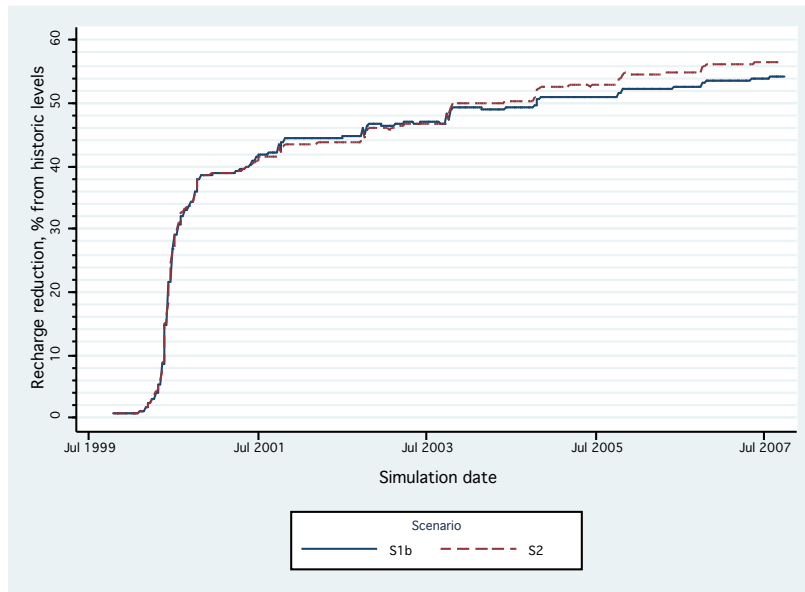

Figure 6.17: Recharge reduction, % from historic recharge levels, time series, comparing scenarios 1b and 2. We see less a less dramatic difference between scenarios relative to seepage, perhaps suggesting incentive packages should more closely target ditch managers. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

Once again, it is ditch manager actions which appear to be most sensitive to institutional changes. If we examine the sealing and lining data in figure 6.18, we can identify two interesting results: first, an increase in sealing in scenario 2 over 1b; second, a reduction in lining over scenario 2 over 1b.
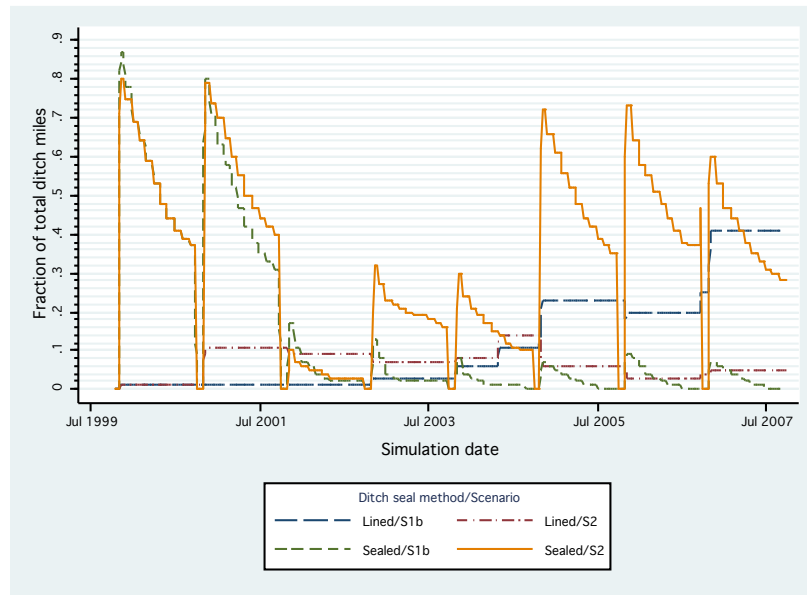


Figure 6.18: Fraction of total ditch miles lined, time series, comparing scenarios 1b and 2. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

Adding informational feedback favors the cheaper and quicker seepage reduction options, while penalizing the use of concrete lining. This is most likely to be due to the development of inertia in the sustainability rating of the sealing action. Since sealing is a quick, cheap and easy action to take, it dominates the early years of the simulation in both scenarios. In scenario 2, however, this results in the sustainability rating of sealing changing faster than concrete lining (ratings are only changed when a feedback is received), and a higher rating over time (since the feedback ratings of sealing are likely to be high). This inertia would matter less if it were not also the case that agents are only allowed to expend their resources on a few activities each time step, and any emphasis on one action can reduce attention on other actions. As sealing grabs more and more of the limelight, ditch managers have less resources to devote to lining. In other words, it appears that the addition of informational feedback causes a shift in trend away from lining and toward sealing, in a form of path dependency: as early decisions favor sealing,

213

the informational feedbacks help lock in this particular behavioral choice over time. This is not a problem per se, since sealing is an effective and appropriate way to reduce seepage. But it raises some practical implications for institutional design which I discuss below in section 6.8.

On the economic front, we saw in Chapter 5 that adding informational incentives appears to improve the income and asset status of agents. But we also see that trading volumes in the WQTM do not greatly vary between the two runs, so what else might explain this economic variance? As discussed earlier, there are only a few ways that agents can make money: through traditional means like planting and harvesting crops, through slightly less traditional means like selling water rights to cities, through selling quota on the market or receiving incentive payments, and through taking out additional debts. Beginning with quota selling, incentive payments are the redistributed fines charged to agents remaining with a quota deficit even after a year of trading. In figure 6.19, we see clearly how the incentives paid in scenario 2 greatly exceed those paid in scenario 1b. This is no doubt a significant contribution to the improved economic outlook across many agents. Recall that the income figures are averages, which indicates that this cannot be a simple redistribution of existing resources. In fact, what we are seeing is that higher levels of fines through the WQTM are driving higher levels of debt accumulation by both farmers and ditches (see figures 6.6 and 5.17). While – in principle – higher incomes indicate improved sustainability, if much of this gain is coming from increased levels of debt incursion we can hardly call the higher incomes "sustainable".

However, these are figures for *accumulated* debt. If we review time series plots for debt incursion across all scenarios (figures 6.20 and 6.21), while there is an initial spike in debt incursion in all WQTM scenarios, this debt ends up being worked back down with time. Overall, Scenario 2 saddles farmers with around 30% more debt relative to the baseline, and ditches with around 100% more than they would be used to, which are manageable figures for actors and institutions well accustomed to carrying large amounts of debt. Considering the considerable sustainability benefits that are associated with these additional debts, and the fact that the debts disappear with time, we can continue to support the general thesis that S2 is more sustainable than S1a.

We have explored how high levels of fines drove high levels of incentive payments and this
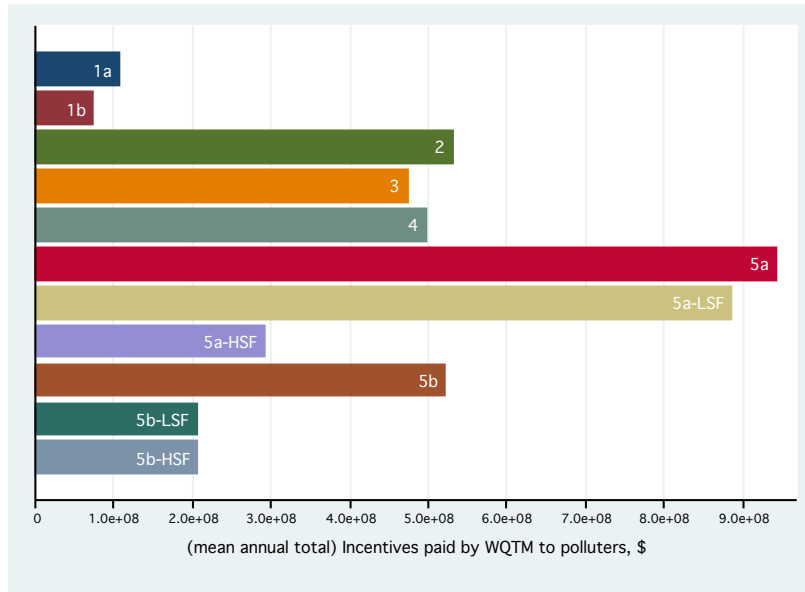
Figure 6.19: Incentives paid (total $) by the WQTM to agents not exceeding their pollution quota, across all scenarios. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.
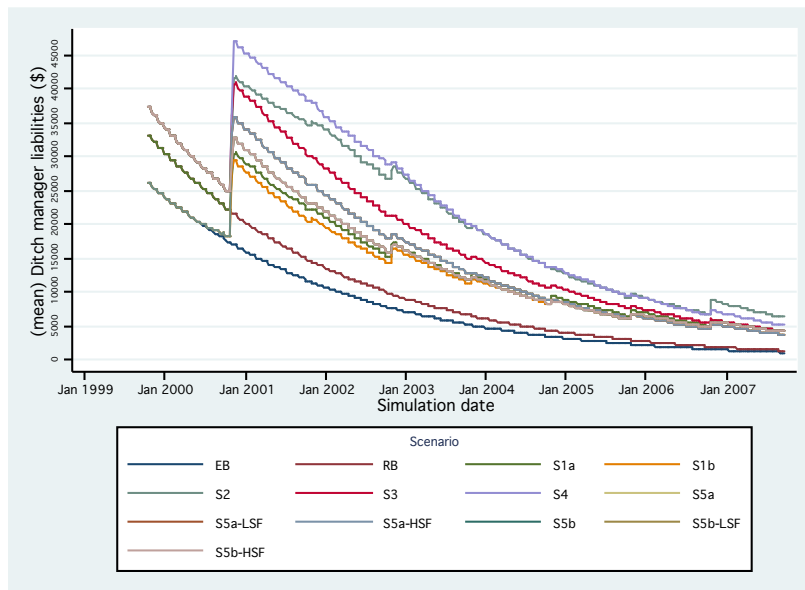


Figure 6.20: Debts accumulated over time by ditches ($). Highlights initial debt incursion in scenarios with WQTMs, followed by a steady decrease in debt load. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

in turn drove high rates of debt accumulation., But nothing that across several metrics (quota traded and quota status), there appears to be little difference in the overall market activity, why did incentive payments increase so much in scenario 2 over 1b? The answer lies in the pegging

215
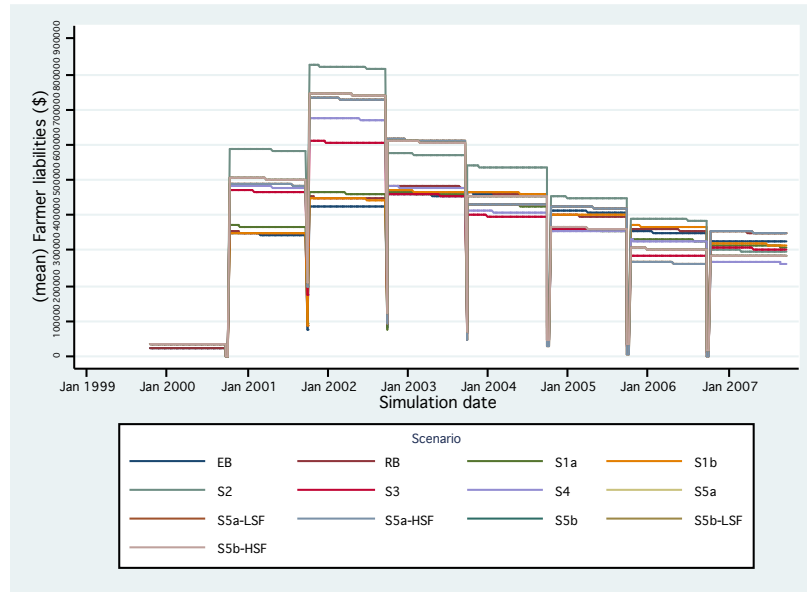
Figure 6.21: Debts accumulated over time by farmers ($). Highlights initial debt incursion in scenarios with WQTMs, followed by a steady decrease in debt load. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

of fine levels to the unit price of quota. This is standard practice in WQTM design, since it is advantageous for salinity outcomes to ensure that the per unit fine for exceeding quota stays above the price of buying quota. That way, agents will theoretically be more inclined to trade in the market than simply take the fine. As we see in figure 6.22, the mean per unit price in scenario 2 exceeds that of scenario 1b by one to two orders of magnitude.

It is not clear why this is the case. Quota pricing is strongly linked to perceptions of scarcity, but as shown in figure 6.23, there is very little difference over time in the scarcity rating (ratio of demand to supply) between the two scenarios. This is an anomalous result, which I put aside for now but revisit later in section 6.3.3.

Another option for increasing economic performance is to lease more water for municipal use. Major leasing arrangements have been in place in the LAB for several years, taking the place of less politically palatable water sales. Figure 6.24 shows that around 200,000 more acre-feet were leased by ditches to cities in scenario 2 than scenario 1b. Note that this is a decline from the historical case (EB on the plot), but represents a considerable increase over the rational actor scenario of 1a, and the baseline run for rational actors (RB on the plot). This increase would certainly contribute to improved incomes in scenario 2 over scenario 1b. But

216
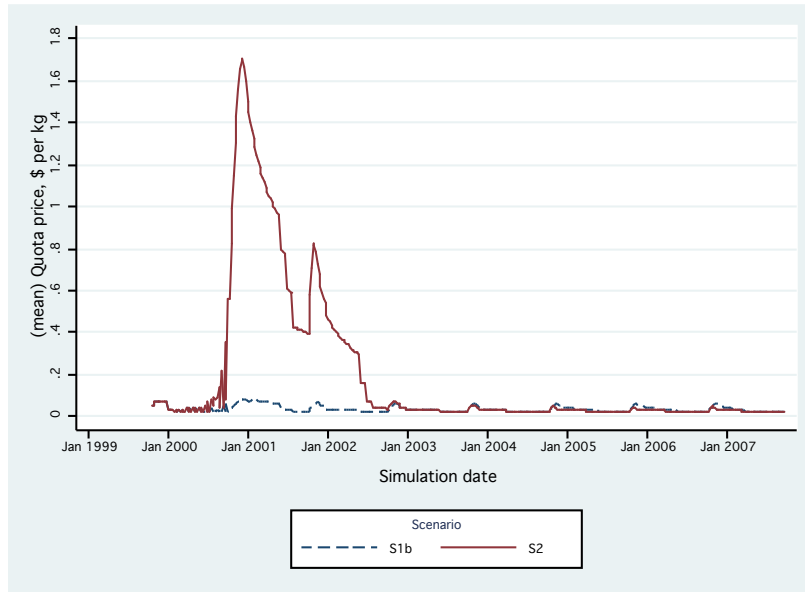
Figure 6.22: Market price in $ of 1 kg of quota, time series, comparing scenarios 1b and 2. Mean per unit price in scenario 2 exceeds that of scenario 1b by one to two orders of magnitude. This is an anomalous result, explored in detail in the text. Refer to Appendix section 8.1.2 for detailed scenario descriptions.
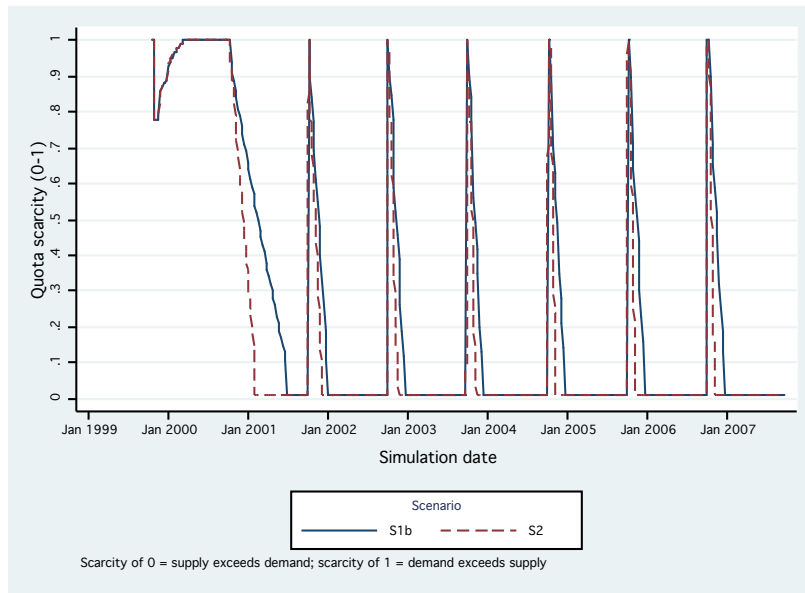


Figure 6.23: Scarcity rating for quota in the market (0.0-1.0). Value of 0.0 indicates no scarcity, value of 1.0 indicates maximum scarcity. Scarcity ratings show very little difference with time, an anomalous result considering much more dynamic quota price trends and given that agent logic is designed such that scarcity impacts agent pricing decisions. See later section on anomalies for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

217

what might be driving this increase in leasing?
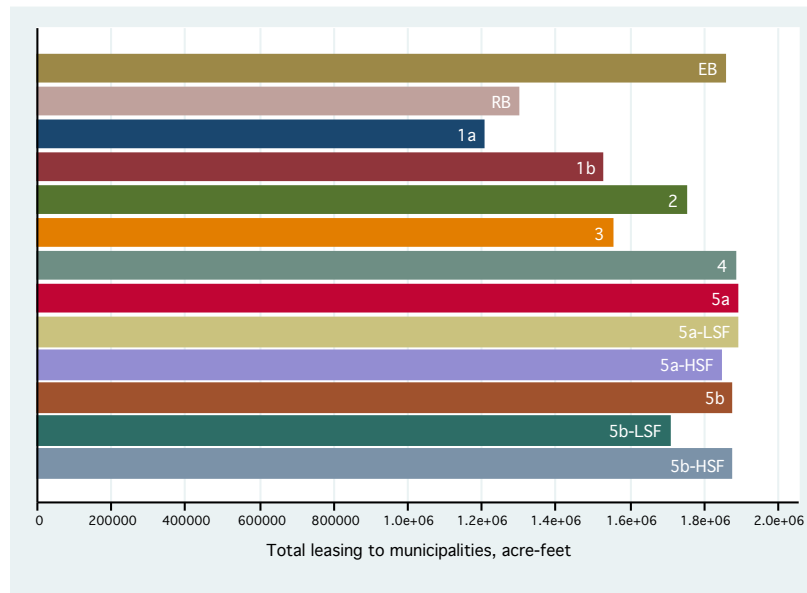


Figure 6.24: Volume of water (acre-feet) leased to Front Range municipalities, across all scenarios. Note how more water is leased in scenario 2 relative to 1b, indicating that providing post-hoc sustainability information feedbacks to agents may act to increase leasing volumes. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

It appears that providing post-hoc informational feedbacks drives up the likelihood of water being leased to cities, and also that more economically rational agents do not favor leasing. The economic benefits of leasing will vary year to year, depending on fluctuations in crop prices, overall water levels, summer temperatures, availability of government subsidies and so on. ArkAgent does include historical fluctuations in all these variables, but the fluctuations are constant from scenario to scenario. What is not kept constant is the economic value of water, which depends on a number of factors including crop yields. The decision logic for ditches to lease water to cities is not activated until a leasing request is received from a municipality; such requests were expected during the simulation, since municipalities over this historical period were actively seeking leases to meet future shortfalls, and the same population and water need calculations were included in ArkAgent. Once a lease is received, a ditch manager agent will consult with shareholders (farmers) as well as conduct its own decision logic regarding the value of water in terms of the proposed lease rate (per acre-foot) relative to the agricultural value of water (i.e. keeping water in irrigation), and the sustainability benefits of the action.

This gives one reason why the more economically rational agents of scenario 1a showed less interest in leasing than the more variable populations in other scenarios: they apply no weight to the environmental or any other sustainability benefits of reduced diversions. The fact that agents in scenario 2 still leased less water than in the baseline scenario (EB on the plots) suggests that participation in the non-traditional WQTM increases the value to agents of keeping water within the basin. However, data for water value shows a broader, contradictory trend: from scenarios 1a through 5a, the agricultural value of water increases, yet the leasing volumes also increase.

An explanation for this apparent contradiction lies in the fact that agents will in all likelihood lease more water if they feel they have spare water to lease (i.e. unused out of their right). Because participation in the WQTM tends to drive down diversions (it increases the cost of diverting over and above what is needed, and application efficiencies improve so that less water is needed), it is likely we are seeing the result of more water being available than would otherwise be. This appears to override the improved agricultural value of water in scenario 2 relative to 1b, which is not surprising since a water lease represents a guaranteed source of income for a given volume of water, something farmers are not accustomed to receiving. We see this trend across scenarios 1a through 5a, with agricultural water values increasing but also leasing volumes going up. This suggests that leasing behavior is particularly robust to normal economic fluctuations in crop pricing and other determinants of return on farming investments, not least because the price cities are willing to pay for water does not fluctuate in the same way.

One final note on the impact of leasing. Leasing within ArkAgent (and indeed, in the real LAB) does not necessarily imply the physical removal of water from the system. Particularly over the historic time period simulated by ArkAgent, cities in and outside of the LAB were not particularly concerned with actually using any leased water rights, but instead appeared to use such rights purchases as hedges against potential drought and to avoid much larger capital expenditures. So, in effect, the only difference that leasing made to the ArkAgent hydrology was to keep that leased water in the river and reduce diversions (see later figure 6.1). It is important to note that this will also have contributed to reduced salt loads.

Returning one last time to the issue of increased income in S2 over S1b: note that there are

demonstrable improvements in crop yields (figure 6.53) in scenario 2 over 1b, and so this no doubt also contributes to increased incomes in S2. These improvements in yields are likely to come from increased application efficiencies, and a slight shift to higher returning crops like vegetables (see Appendix figure 8.11).

Turning at last to institutional performance, Chapter 5 suggested that while scenario 2 did perform slightly worse than scenario 1b in terms of market efficiency, much of this difference came from the clearing success ratio. This identifies a potential problem arising from the use of information feedbacks: just as informational feedbacks can foster desired improvements in participation, this can backfire. Selling quota, in particular, appears to have greatly improved its sustainability rating over time, but this may have well contributed to the glut of quota added to the market for sale. Such a glut can theoretically drive down quota prices and lead to worsened participation rates in the long term, although this was not seen within the model time period.

### 6.3.3   Hypothesis 2: anomalies

The main anomaly in comparing data from scenarios 1b and 2 is that the quota pricing for scenario 2 in the first couple of years greatly exceeds anything seen in scenario 1b. In fact, the pricing in scenario 2 over this period exceeds that of any other scenario (figure 6.25).

This anomaly is likely due to a combination of the informational incentives driving less rational trading behavior, and quirks of the simple trading logic. Since the only difference between scenarios 1b and 2 is the addition of the informational incentive, the informational incentive is actually driving less rational trading behavior, even though behavioral orientations do not change between scenarios. This is because the informational incentives reinforce existing behavioral logic. Even if the underlying logic is less economically rational to begin with, adding the informational incentives will reinforce this behavior (recall that these incentives encourage more *sustainable* behavior, not economically rational behavior, although these may coincide on occasion).

According to the underling logic, when an agent decides to post quota for sale, it decides upon an asking price on the basis of a number of factors, including its perception of scarcity

220
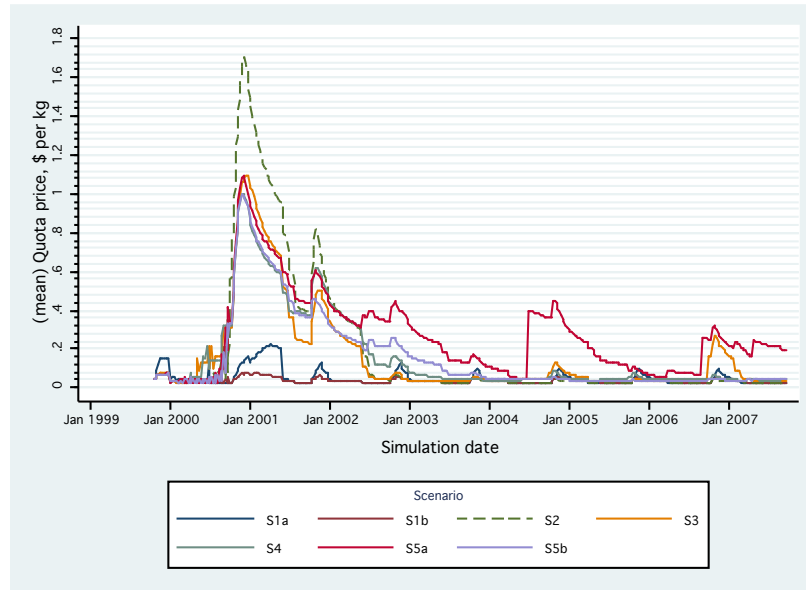
Figure 6.25: Price of 1 kg of quota in $, time series, across all scenarios. Pricing of quota in scenario 2 exceeds all other scenarios in the initial years of the simulation. This is likely due to a combination of less rational trading behavior driven by non-monetary incentives, and by quirks of the ArkAgent trading logic. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

in the market, the current mean price of quota already on the market, and how long quota posted for sale has been stuck unsold. Crucially, the agent re-evaluates its pricing strategy *every time* it considers posting quota. One potential emergent result of running this logic with less economically rational agents is runaway pricing: each time the agent reconsiders posting surplus quota for sale, it sees continued scarcity, continued demand, and previous quota has not been on the market long enough to cause any depression in pricing. As expected, the beginning of simulation year 2000 is marked by a dramatic increase in the number of quota sale offers made, from less than 50 a week to more than 500 (figure 6.26).

The difference between scenario 1b and 2 is largely down to the "de-rationalizing" effect that informational incentives have on this posting logic: the informational incentives do not penalize for excessive posting and consequent price volatility, but they do return positive feedback for participation in general.

While the combination of simplistic pricing logic and informational incentives is likely driving much of this posting and price volatility, could the minimal transaction costs built into the clearing house also be having an effect? Transaction costs have long preoccupied resource
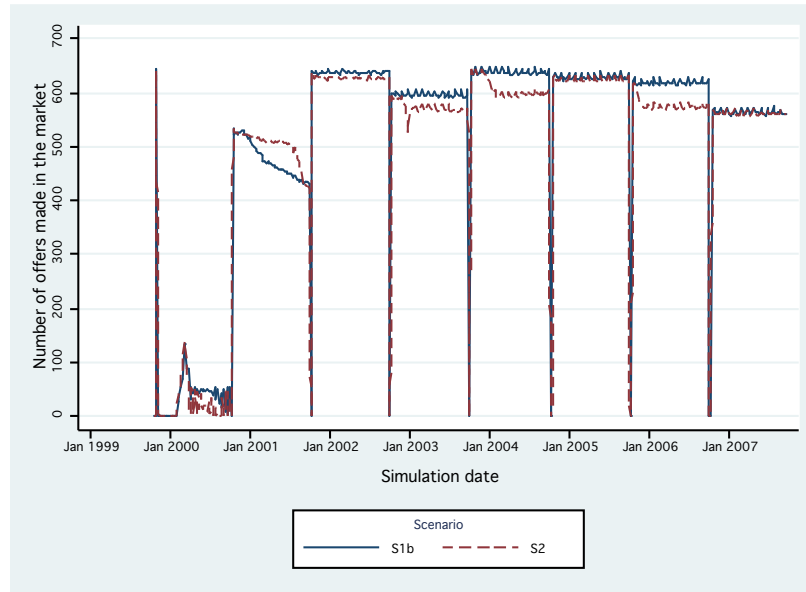
Figure 6.26: Trade offers made, time series. A "trade offer" is a single instance where an agent places quota for sale in the marketplace (the WQTM quota clearing house). See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

economists (Fang et al., 2005; Hoag & Hughes-Popp, 1997; Stavins, 1995), and recognizing the depth and breadth of study in this area, I do not intend to make a lengthy analysis of this issue. Indeed, any real transaction costs were deliberately left out of ArkAgent since the model was not intended to be a study of neoclassical explanations for WQTM failure. Nevertheless, transaction costs are real, and have measurable effects in real markets. With the introduction of transaction costs to ArkAgent, the exuberant posting behavior exhibited by the simulated actors in ArkAgent would likely be tempered considerably. But assuming that institutional designers usually do not set out to maximize transaction costs, and the internet-based clearing house adopted in ArkAgent is not an outlandish concept, it is hard to believe that transaction costs alone would negate the "runaway posting" behavior seen in ArkAgent.

A key lesson from this discussion is that institutional designers may need to be careful to avoid generating motivations for perverse behavior, when applying non-economic incentives. Price volatility is something to be avoided; in a real system, it will be more disruptive than the simulation here implies. Another key lesson from the transaction cost discussion is not that costs were too low, or that higher costs would damage market performance; instead, ArkA-gent's excessive trading frequencies, and associated price volatility, suggest we need to build

richer simulations of agent trading and pricing behavior. It is perfectly reasonable to say, in general, that high transaction costs deter market participation. But consider the case of the ill-fated water quantity trading market attempted in the LAB during the 1980s. Completing a trade at that time required overcoming several onerous bureaucratic hurdles which no doubt prevented quick and easy trades. But perhaps more importantly, farmers faced the general opprobrium of their colleagues if they participated in an institution which moved water out of the basin. Indeed, in water quantity markets among several western states, Colorado stands first in percentage of trade applications — 70% — formally protested or opposed (Howe, 1998). In this case, do we count social resistance to trading water out of the basin as a "transaction cost", or is it, in fact, a feature of more complex and varied behavioral models than neoclassical studies account for? ArkAgent's simple trading and pricing models probably generate unrealistic trading frequencies and price volatility, but I propose that the solution is not to do a better job of simulating transaction costs. Instead, we should be focused on gathering empirical data that helps determine the various dimensions of trading and pricing decision making in this particular basin context.

### 6.3.4 Hypothesis 2: parameter variation

1. *Specific effects on salt load*: the only parameter variation experiments which show a substantial departure from the baseline in terms of mean salt load reduction, are the initial quota price (IQP) runs. In this case, setting the IQP higher appears to drive up the mean level of salt load reduction, although we do still see a reduction in salt load from 1b to 2. Contrasting this with the IQP results from Hypothesis 1, where we saw that changing the IQP shifted salt load means up or down by a small amount, we get further insight into the complexity of the function relating IQP and salt load response.

2. *Effects of WQTM management on magnitudes of response*: the Chapter 5 analysis showed that relaxing controls on quota banking and deficit carryover meant that institutional performance greatly worsened. This is predictable, since allowing agents to store surplus quota will simply result in more quota being re-added to the market each year.

3. *Broad support for Hypothesis 2*: environmental, institutional and economic data show that Hypothesis 2 is broadly robust to variation in social and economic parameters.

### 6.3.5   Hypothesis 2: real world implications

From the comparison of data from scenarios 1b and 2, a number of real world implications can be drawn.

1. Non-economic incentives have considerable potential. Institutional designers should be considering alternative incentive forms in terms of how explicit these incentives are presented to stakeholders (i.e. going beyond the traditionally implicit financial incentives of the WQTM), and in what form these incentives take (i.e. adding a non-economic twist to the traditional economic emphasis). This is an implication relevant to any basin setting where stakeholders have variable attitudes and behavioral strategies. This is likely to be particularly true for basins dominated by family farms that have deep ties to the community and the land (Filson, 1996). The extent to which this might hold in settings dominated by major agribusiness concerns — a setting more common in the U.S. midwest — is a matter of speculation, but we should also be wary of ascribing monolithic beliefs and strategies to corporations: some large agribusiness concerns have shown particular concern for sustainability in their operations (Jansen & Vellema, 2004).

2. ArkAgent results show the potential for actor-specific incentive packages in the LAB. Discussion under Hypothesis 1 highlighted the importance of ditch managers in ensuring good environmental outcomes. The results for Hypothesis 2 underline this conclusion. Since ditch managers can have important roles in salt load reduction, the use of incentives to exploit this potential is particularly well-advised. While this is most relevant to the LAB, it is common in many basin settings for actors to have varied capacity for sustainability change. Where such disparities can be identified, it may be appropriate to target incentives at specific actor groups with especial effort made to target those actors with the greatest potential to alter sustainability outcomes (like ditch managers, in this case).

3. ArkAgent results highlight the sustainability benefits of municipal leasing, in particular the reduced salt loads and increased incomes for farmers. Given the politically sensitive nature of water leasing, this should not be taken as evidence that water leasing is universally "good" or "bad" for any particular system. The results do indicate, however, that leasing may have a partial role in improving the economic and environmental outcomes in the LAB. Results related to Hypothesis 2 indicate that the non-traditional WQTM actually depresses leasing levels relative to the baseline, suggesting that a real world WQTM might be seen as beneficial in the eyes of those who would prefer to keep Arkansas Basin water within the basin. On that subject, it is important to note that while other scenarios (4, 5a an 5b) show slightly increased levels of leasing relative to baseline, no WQTM scenario indicates that considerable more leasing would take place with such an institution. Since many people in the LAB do regard leasing as potentially disruptive, this indicates the WQTM would at the very least not exacerbate any existing problems. Inter-basin leasing is common in the American West, but the range of buyers and sellers, coupled with the variety of hydrologic settings for the leases (Chong & Sunding, 2006), makes it difficult to determine the generalizability of this implication.

4. Incentives have side-effects. The discussion of Hypothesis 2 raises the potential problem of unexpected side-effects due to incentive packages. The ArkAgent market was flooded with unsold quota to an extreme which is in some measure to do with simple agent logic, but mainly to do with the fact that the informational feedbacks encouraged a perverse tendency to keep posting quota regardless of actual benefit to doing so. The potential for such side-effects to occur in real world systems, immeasurably more complex than ArkAgent, is a serious possibility. Since we have limited understanding of what motivates water users and managers in most basin settings, this has relevance in any complex water system beyond the LAB. One other important side-effect was the increased debt accumulation driven by higher levels of fines in Scenario 2. While this appears to be a negative side-effect, the debt accumulation was not crippling, and may in fact be a necessary instrument forcing agents to react appropriately to market incentives.

5. Institutional designers need to develop a more sophisticated, empirically-based understanding of trading behavior. The ArkAgent experience suggests that institutional designers should consider as much empirical evidence as possible when deciding to what extent something like transaction costs will affect trading behavior. The transaction cost of trading quota is something ArkAgent deliberately avoids simulating. This undoubtedly results in higher levels of trading than might otherwise be expected. This is in line with the argument that mitigating transaction costs is critically important in improving WQTM performance (Hoag & Hughes-Popp, 1997). However, much of the excessive trading frequencies and volatile pricing in ArkAgent are likely to come not from any particular absence of transaction costs, but from insufficiently sophisticated treatment of agent logic related to trading and pricing. The failure of real world quantity markets in this basin have demonstrated that anything involving the non-traditional use of water is fraught with social complexity. We do real world actors a disservice by asserting that anything as simple as transaction cost would single handedly drive trading behavior. Once again, this kind of implication will be relevant wherever we do not have a deep understanding of what drives water users and managers, which is most cases.

6. ArkAgent results show path dependence, implying institutional designers involved in complex basin settings must use forecasting techniques before they implement interventions. The Hypothesis 2 discussion uncovered path dependency in the choices of ditch managers. An initial preference for sealing over lining led to irreversible reinforcement of this preference over time, and an overall shift towards ditch sealing as the main means of seepage reduction. There are two real world implications relevant to most real basin settings that embody considerable social and biophysical complexity: the first is that institutional designers should remember that agents frequently face competing options in responding to an institutional incentive: should the manager seal or line? Should the farmer upgrade to drip or sprinkler? Once a given option is selected, this reduces the potential for other options to be selected in future, for any number of reasons including simple behavioral inertia (what has been done before is typically favored in future), real limitations on resources, and potential positive feedbacks which lock in behavioral

choices. The second implication is that institutions can cause path dependence. Institutional designers should take care to develop a good understanding of what potential shifts in system behavior may result from an institutional intervention, because it may be very difficult to undo these shifts once they have occurred. Forecasting tools (computational or otherwise) are essential to developing this understanding.

## 6.4 Hypothesis 3: capacity for attitudinal change and effects on institutional performance

### 6.4.1 Summary of section

In the following discussion (section 6.4.2), I find that scenario 3 follows much the same trend as scenario 2 in results, and since scenario 3 barely differs in parameterization and structure from scenario 3, the same reasoning applies: the explicit non-economic incentives of the non-traditional WQTM drove more sustainable behavior over time, relative to the results seen with the traditional WQTM. Noting, however, that scenario 2 shows better sustainability outcomes than scenario 3, I explain this different outcome by showing that the addition of S-factor dynamics in S3 results in an initial adjustment period of lower S-factor values (for farmers) and for overall lower S-factor values (for ditch managers). Since lower S-factors are associated with more economic rationality, this has the effect of driving behavior away from more sustainable decisions. I dig deeper into the data to show that the reasons for the reduced S-factors vary by agent type, with farmer S-factors affected by an adjustment period reflecting the relative "un-sustainability" of their initial decisions, and ditch manager S-factors affected by the reduced exposure these agents get to informational feedbacks. I also discuss show the much more dynamic attitudes and behaviors seen in scenario 3 generate some interesting results, including a shift by ditch managers away from ditch sealing and towards lining (subverting the sustainability objectives of the institution). Finally, I emphasize that while S3 shows reduced performance relative to S2, it is also more complex (and so most similar to the real system), and it is encouraging to see the sustainability benefits still exceeding those gained through the traditional WQTM.

In the discussion of real world implications in section 6.4.4, I suggest that the results for scenario 3 continue to add weight to the idea that policy makers applying real world WQTMs should consider alternatives to the traditional set of implicit financial incentives. I also point out, however, that incorporating non-traditional incentives into a traditional WQTM structure can have initially counterintuitive results in cases where actors are vulnerable to attitudinal change, and that as a result positive outcomes may take several years to emerge. This counsels patience for institutional designers when trialling new institutional forms. Finally, I re-emphasize that different actors may respond very differently to different incentives and to how those incentives are delivered, and that this suggests institutional designers should be careful to implement non-economic incentives in blanket fashion.

### 6.4.2 Hypothesis 3: reviewing the arguments

Hypothesis 3 argues that assuming variable capacity for behavioral change within a population, a WQTM variant that incorporates incentives intended to promote behavioral change (i.e. non-traditional WQTM) will perform better than a traditional WQTM missing such incentives. It also argues that the non-traditional WQTM will also perform better under the behavioral assumptions of Hypothesis 3 than under the behavioral assumptions of Hypothesis 2.

The analysis in Chapter 5 argued that on all sustainability metrics, scenario 3 performed better than scenario 1b, which is broadly supportive of at least the first assertion in hypothesis 3. This means that non-traditional WQTMs are better at fostering attitudinal change among a population of heterogeneous agents, even when we assume that agents have a capacity for attitudinal change in response to informational feedback. In essence, the non-traditional market institution plus diverse incentives is more robust to attitudinal dynamics than the traditional market institution. Note, however, that on most metrics scenario 3 results in less sustainability benefit than scenario 2. This indicates that the effect of assuming attitudinal dynamics is to diminish the effect of diverse incentives. It should be no surprise by now that S3 performed better than S1b, since S3 has almost exactly the same setup as S2 with the exception of added attitudinal dynamics. In other words, the post-hoc informational feedback mechanism coupled with implicit financial incentives worked to incentivize more sustainable decisions on the part

of agents. Explaining the performance decrement in S3 relative to S2 takes a little more work, and I address this in the following discussion.

The analysis of hypothesis 3 results in Chapter 5 showed that scenario 3 experienced more salt load reduction than 1b, but lower levels of river flow. As before, I begin by exploring how and why salt load was reduced. Examining a plot for seepage and recharge for 1b and 3 (figures 6.27 and 6.28), we see that scenario 3 exceeded scenario 1b for both. Since differences were more significant for seepage than recharge (yet again), examining a plot for sealing versus lining (figure 6.29) indicates a similar trend to past analyses: ditch sealing dominating seepage reduction efforts. Since scenario 3 has the same core behavioral model as scenario 2, the same explanations apply and I will not dwell further on this.
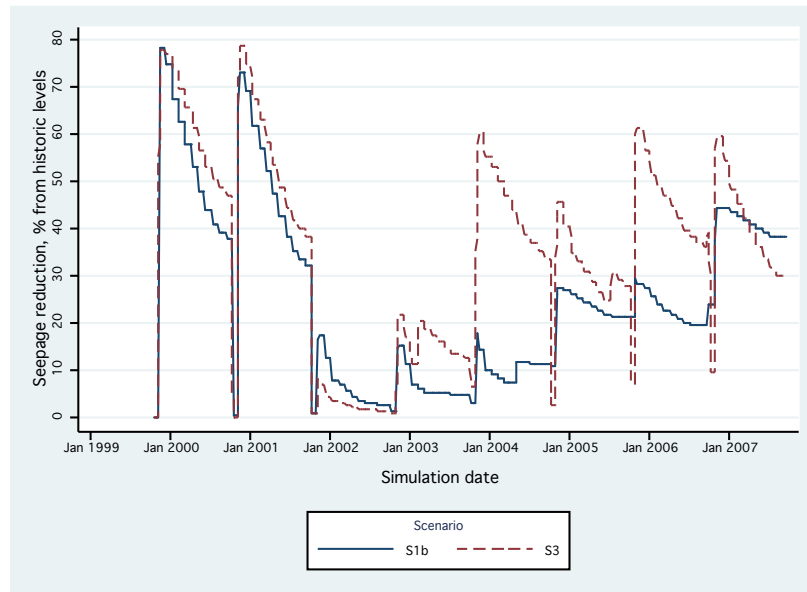


Figure 6.27: Seepage reduction, % from historic seepage levels, time series, comparing scenarios 1b and 3. Seepage reduction is higher in scenario 3 relative to scenario 1b. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

Another interesting feature of the S1b/S3 comparison identified in Chapter 5 was the lower volume of river flow in S3. Examining diversions data (figure 6.1), we see that scenario 1b diversions were an annual average of nearly 2,000 acre-feet higher. Three possible factors could be at work: higher diversions because of higher perceived or actual irrigation needs; higher diversions because of reduced leasing of water to municipalities; and higher diversions due to abandonment clauses in water rights. Regarding the first possibility, we see in figure 6.30 that,

229
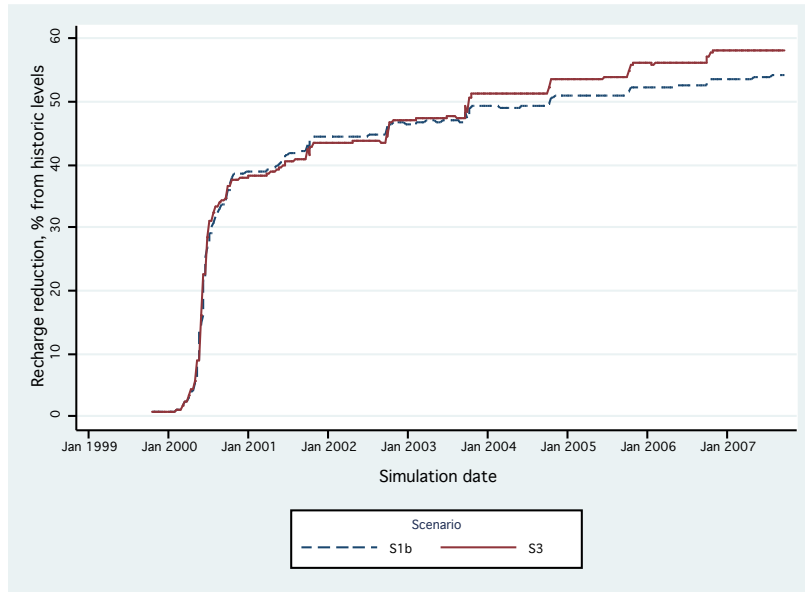
Figure 6.28: Recharge reduction, % from historic recharge levels, time series, comparing scenarios 1b and 3. Recharge reduction is higher in scenario 3 relative to scenario 1b. Refer to Appendix section 8.1.2 for detailed scenario descriptions.
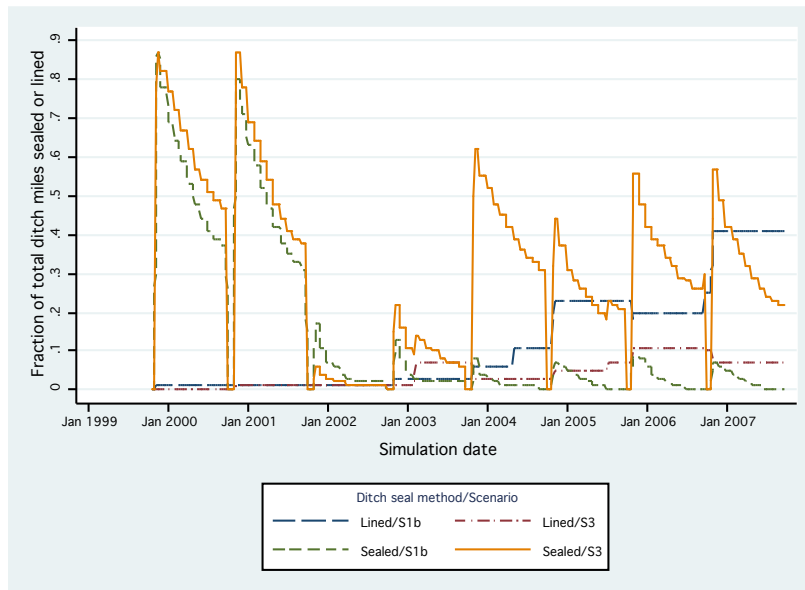


Figure 6.29: Fraction of total ditch miles lined and/or sealed, time series, comparing scenarios 1b and 3. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

in fact, farmers plant less land in scenario 3 than they do in 1b. Furthermore, we see in figure 6.31 that mean basin soil moisture in scenario 3 was little changed over the period of simulation relative to S1b, suggesting that farmers were not suffering from any particularly severe deficit

230

in moisture (or at least, no more so than S1b).
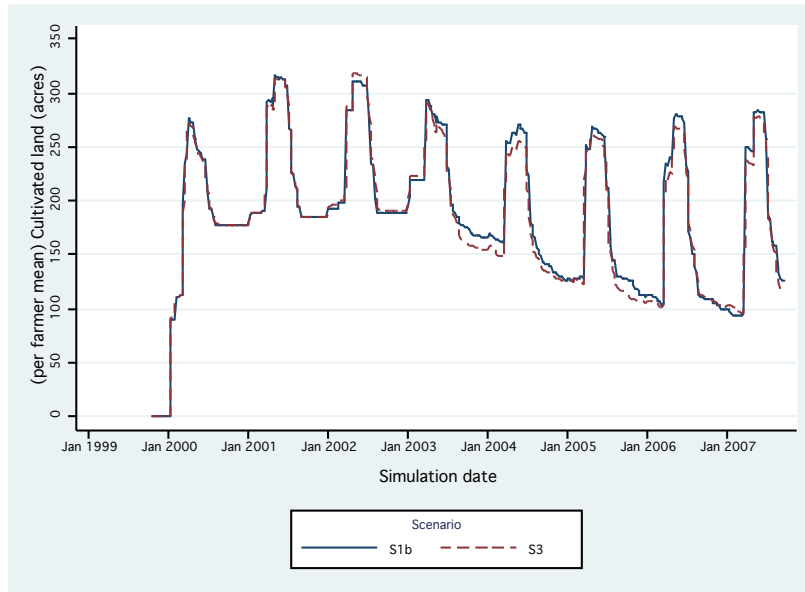


Figure 6.30: Acres of land cultivated in the basin, time series, comparing scenarios 1b and 3. Farmers plant slightly less land over time in scenario 3 relative to scenario 1b, suggesting lower acreages in S3 may be partly responsible for reduced diversions in that scenario. Refer to Appendix section 8.1.2 for detailed scenario descriptions.
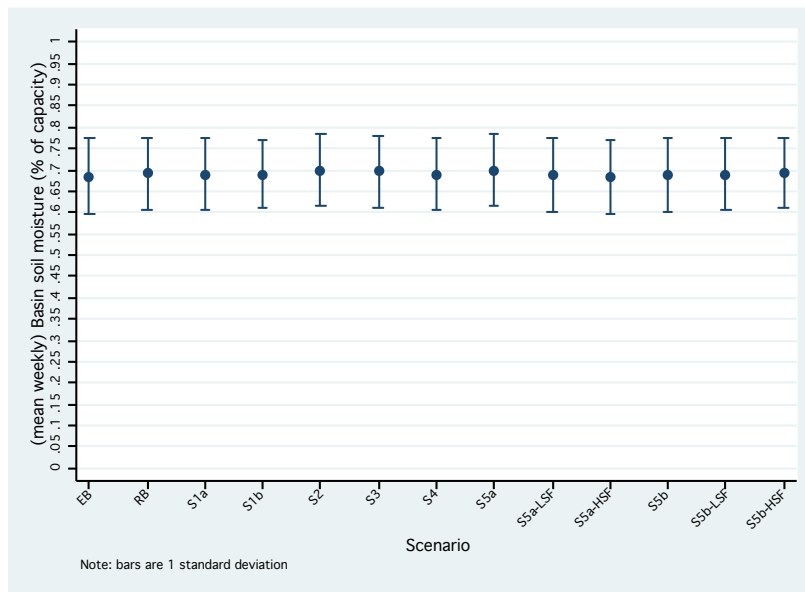


Figure 6.31: Mean basin soil moisture across all simulated grid cells, % of maximum capacity. Note the minimal variation in soil moisture among the scenarios relative to the baseline. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

There is another layer of complexity here, however. Farmers can not only choose to plant

more crops, but they can choose to plant more water intensive crops. ArkAgent simulates the planting of vegetables, alfalfa, corn, grass, wheat and sorghum. Of these, vegetables are by far the most water intensive crop, followed by grasses, alfalfa, sorghum, wheat and corn. Figure 6.32 has acreages for alfalfa and grass omitted, since these together make up more than 80% of all plantings and if plotted would make identifying trends in other crops difficult. In any case, acreages for both these crops decrease from scenario 1b to 3. For the remaining crops, we see that per farmer mean total acreages across all crops barring sorghum increase from scenario 1b to 3. The increases across these crops, particularly vegetables, could certainly have contributed to the increased diversions. However, the acreage increases are still slight relative to the magnitude of increase in diversions.
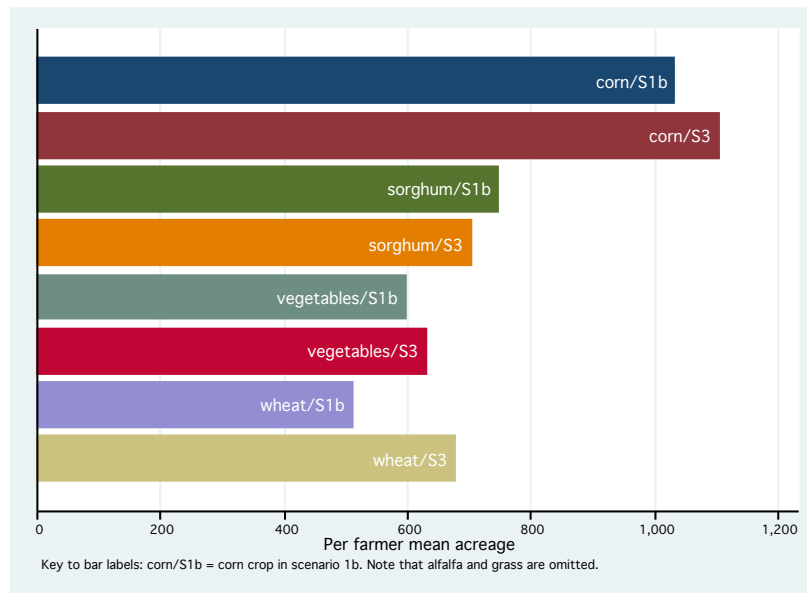


Figure 6.32: Acres cultivated by crop, comparing scenarios 1b to 3. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

Regarding the second possibility (reduced leasing), figure 6.24 indicates that leases to cities in scenario 3 were roughly the same as in scenario 1b. This brings us to the third possible explanation: the effect of abandonment clauses. Briefly mentioned in Chapter 2, these legal specifications in water right decrees hold that a water right not fully appropriated for a period of years becomes derelict and may be reassigned to competing rights holders or back to the state (i.e. the river). While in practice these clauses are rarely activated in Colorado, the threat

of doing so acts as a real impetus to farmers to divert their full appropriation. On purely economic grounds, this is rational: if there is a risk valuable water will be lost if it is not diverted, and there is minimal immediate cost to diverting the full amount, then a rational actor will always do so. In the ArkAgent results, since farmers are neither planting irrigating more acres of crops nor increasing their irrigation rates, this suggests that after they have used up a portion of their water right for irrigation, they are continuing to divert water to maintain that portion of the right as "active". In ArkAgent logic, the portion of this right that they divert depends in part on their risk strategy (more risk averse agents will divert more water), but more importantly on their S-factor. Farmers with a lower S-factor will divert more of their spare water depending on how "rational" they consider themselves. Diversion of this additional water is not tracked in ArkAgent, but we have an indirect measure in the farmer sustainability factor (figure 6.33). As expected, the mean S-factor in scenario S1b is around 0.57 (as per the empirical parameterization from the psychometric testing), and the value does not change because the change mechanism is not activated.



Figure 6.33: Farmer S-factor (0.0-1.0), weekly time series, comparing scenarios 1b through 3. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

The mean S-factor in S3, however, shows considerable variation downwards in the first three years of the simulation, and lies below the mean for 1b in most of the remaining years.

Comparing diversions in these first three years (figure 6.34), and we see that, indeed, the lower SF in scenario 3 drives higher mean diversions.
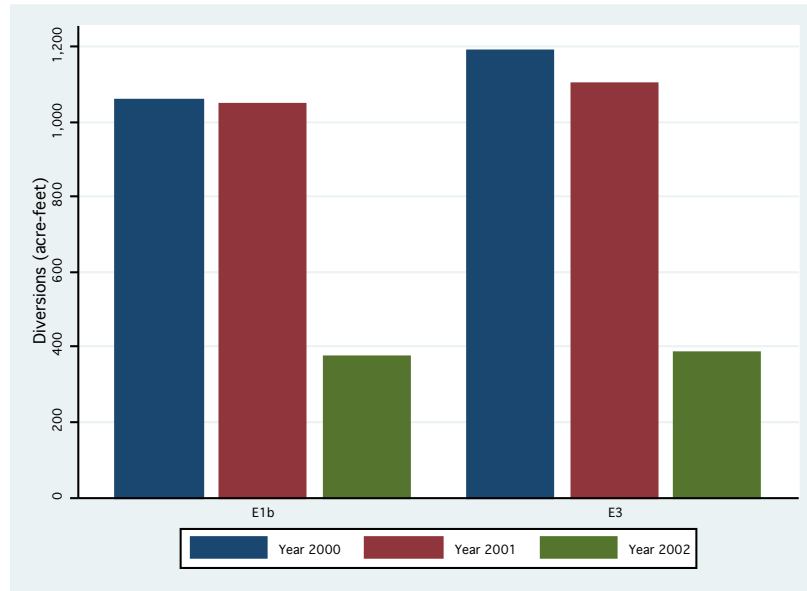


Figure 6.34: Mean annual ditch diversions (acre-feet), comparing scenarios 1b through 3 in 2000 and 2001. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

While this in all likelihood solves the problem of what caused higher diversions in scenario 3, a new question arises: what drove the agent S-factors downwards in those first three years, and what kept them lower than the mean S-factor for scenario 1b during most of the simulation? Further, what does this mean for the overall mechanism behind the reduced performance in S3 relative to S2? Recall from Chapter 3 the S-factor modification equation (figure 6.1), where $S$ = the agent's original S-factor, $SR$ = the received sustainability rating from the informational feedback, and $BMSF$ = a constant helping determine the magnitude of change.

$$m = \frac{1}{(|(S - SR)| * BMSF)} \tag{6.1}$$

For the mean S-factor to decrease the agents must be receiving sustainability ratings lower than their own S-factors. The decrease with each time step appears to be small, suggesting that the received ratings were considerably lower than the agent S-factor values. After a year of decrease, values start to increase again before stabilizing back up at around 0.5-0.6 from 2003

onwards. There is, in fact, a relatively straightforward explanation. Agents start the simulation by making decisions which are not strongly informed by sustainability, since they have little information on sustainability for each choice. Consequently, as these decisions get made, the feedback returned is negative from a sustainability perspective. This drives the agent attitudinal orientation downwards, as the joint reinforcement component of the change model kicks in (agents tend to shift their orientation towards the rating they receive). The inflection point is reached in 2000, when the poor sustainability ratings received are now higher than the agent S-factor values. This starts to reverse the trend, which in turn shifts the decisions that get made towards the sustainability spectrum. Barring a few serious but short lived decrements in the trend (around September 2002), the trend is then upwards and more stable. It is not known if the decreasing trend in the last year of the simulation would continue, or if the mean S-factor values would show such cycles over time.

The farmers are not only the agents in the system with importance roles to play in salt load reduction, of course. If we review the same data for ditch managers (figure 6.35), we see a different trend.
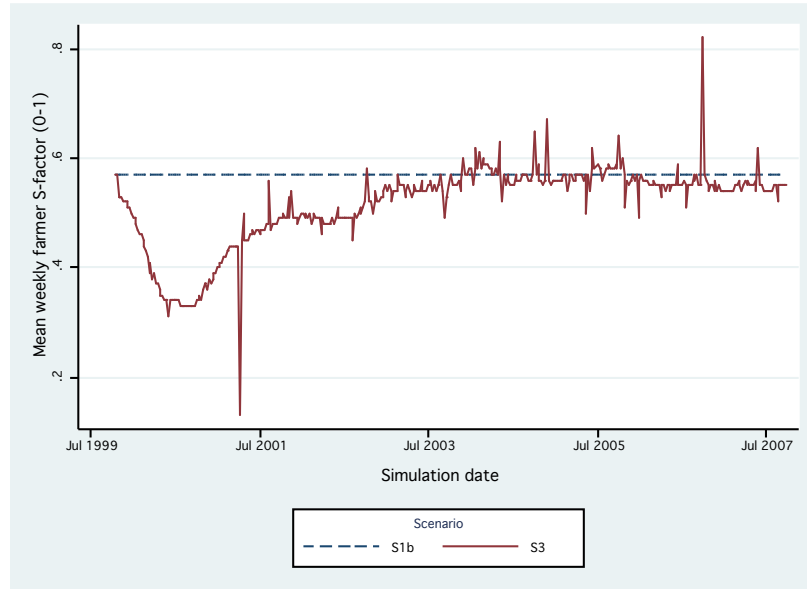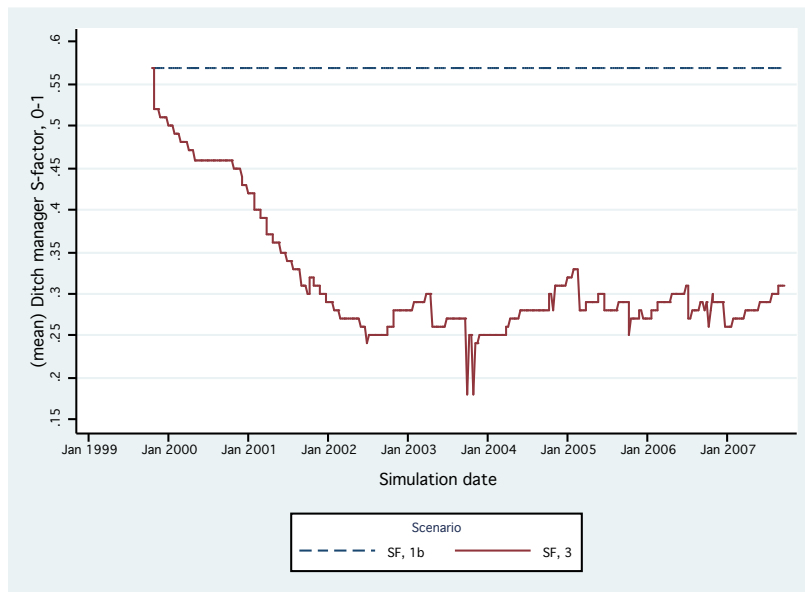


Figure 6.35: Ditch manager S-factor (0.0-1.0), weekly time series, comparing scenarios 1b and 3. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

Instead of a trough followed by a recovery, we see a steady decrease in sustainability atti-

tudes, then fluctuation around a mean of approximately 0.28 (a relatively pro-self orientation). The explanation advanced for the farmer trends is not applicable in this case, but we should not expect it to be. Ditch managers and farmers have very different decision options and frequencies. Farmers make a lot of decisions every day. Ditch managers make far fewer decisions, spaced farther apart. The result is that they get far fewer individual feedback messages from the State Engineer, and far less opportunity for attitudinal change. This explains the more muted variance and lack of a sustained climbing trend. There also appears to be weak seasonality in the ditch manager S-factor values after summer 2002. We see peaks broadly associated with the winter period (October through April), and particular lows at the end of the irrigation season in September. It may be that the increases in sustainability factor during the winter are related to the dominance of maintenance and lining activities in this period, both of which are likely to get better sustainability ratings. At other times of year, it will be harder for ditch managers to maintain good sustainability ratings. These data are also useful in explaining an overall trend towards reduced levels of sealing in the latter two thirds of the simulation (see figure 6.29). Ditch manager S-factors had evolved towards the less sustainable end of the spectrum by January 2002, and it is at this point that we see some of the lowest levels of sealing and lining. The S-factors recover somewhat after that, and we see corresponding recoveries in the amount of sealing and lining ditch managers undertake.

Turning to economic performance, Scenario 3 showed significantly improved incomes and assets over 1b. We cannot look to increased water leasing income (in the case of scenario 2), since this shows little change in 3 relative to 1b. But we do see higher fines in the WQTM and so higher incentive payments. This is a very similar situation to that seen with scenario 2, and the explanation given in the discussion of Hypothesis 2 above is likely to still hold (including the higher levels of debt accumulation). Also, in the economic category, we revisit the water value metric first examined in the discussion of Hypothesis 1. In figure 6.36, we see that scenario 3 follows a similar trend to scenario 2 in the first few years: a significant peak, followed by decline and fluctuation through 2005.

To understand these curves, and what it says about scenario 3 relative to 2 in particular, it is important to note how water value is calculated: whenever they harvest a crop, farmers
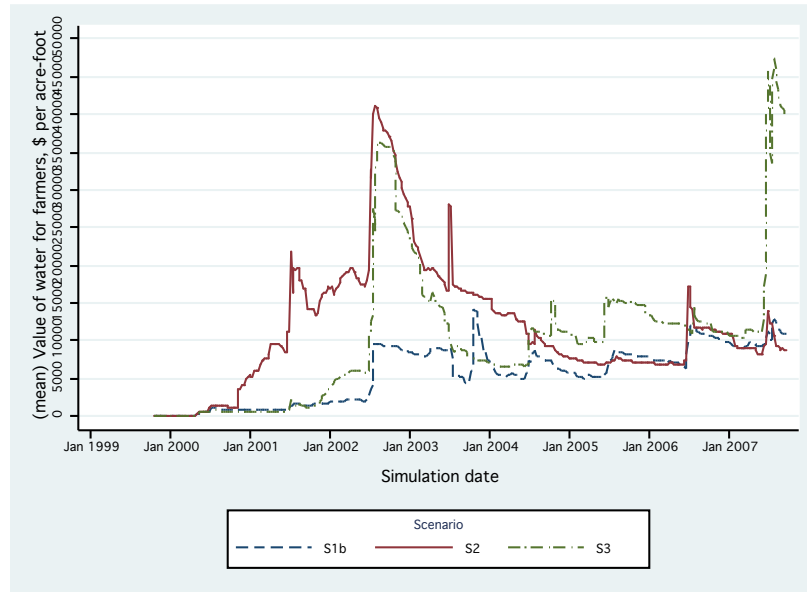
Figure 6.36: Agricultural value of water, $ per AF, time series, comparing scenarios 1b, 2 and 3. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

determine the operating income they brought in since the beginning of the water year. They use this value to determine the per acre-foot value of water they used during the simulation. Note that the farmers do not distinguish between income from on-farm, off-farm or other sources. Consequently, since neither crop yields nor crop prices were dramatically up in this period, the significant increases in the water value during 2002-2003 are most likely due to income returns from the WQTM. As discussed earlier, monetary incentives paid out by the WQTM during this period were significant (see figure 6.19), and so this can explain the overall increase in the value of water. Furthermore, the monetary incentives paid out were slightly less in scenario 3 than 2, and we see reduced water value.

With regard to institutional performance, Chapter 5 identifies scenario 3 as showing a much improved clearing success ratio over scenario 1b, yet still not as good as 1a. This suggests that the effect of assuming the potential for personal attitudinal change and linking that change to decision making, is to improve the "rationality" with which agents engage in quota trading. The earlier plot showing farmer S-factor over time is important: it shows on average that the personal attitudinal orientations of agents in scenario 3 were far more economically rational than those in scenario 1b. Given this, it is entirely explicable that the agents behaved in a

more economically rational way when it came to participating in the WQTM. These effects are also evidence when we review the quota status metric. Figure 6.37 shows that the principal difference that behavioral change has on the evolution of quota deficits is to slow the rate of adjustment to pollution reduction requirements. Scenario 3 sees slightly larger deficits than either scenarios 1b or 2. In fact, it shows a trend similar to that seen in scenario 1a. Since agent S-factor values in scenario 3 at one point approach those of 1a, this is to be expected.
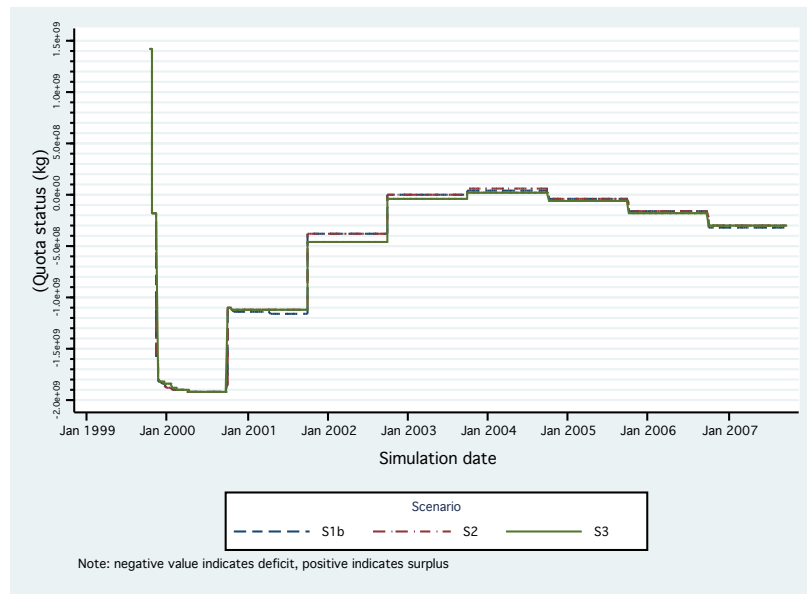


Figure 6.37: Status of quota allocated to trading agents, weekly time series, comparing scenarios 1b, 2 and 3. Plot shows how behavioral change in scenario 3 slows the rate of adjustment to pollution reduction requirements. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

The second component of Hypothesis 3 addressed the performance of scenario 3 relative to 2, i.e. the effects of adding the capacity for personal attitudinal orientations to change during the simulation. A comparison of 2 and 3 on the basis of salt load (figure 5.7) indicates that this behavioral mechanism is to blame for slightly lower reductions in salt load in S3 versus S2. Examining reasons for this reduction (figures 6.38 and 6.39) turns up an interesting result alluded to in Chapter 5: higher levels of sealing in scenario 2 than 3. Not only does the behavioral change mechanism reduce overall sustainability gains, but does so by acting primarily on seepage reduction through ditch sealing. Scenario 3 actually shows improved rates of concrete lining of ditches.
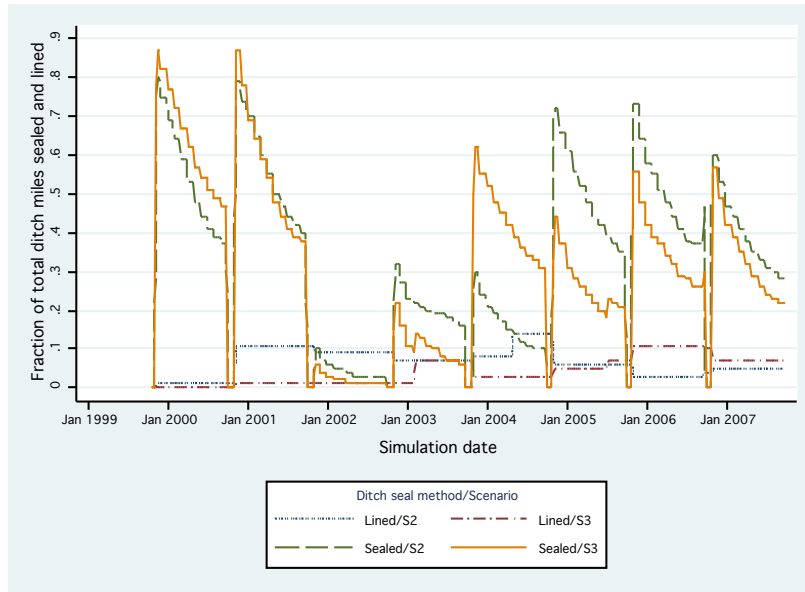
Figure 6.38: Fraction of total ditch miles sealed and/or lined, weekly time series, comparing scenarios 2 and 3. Scenario 2 enjoys higher levels of sealing relative to S3, indicating that preferences being non-static acts to reduce seepage reduction activities. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.
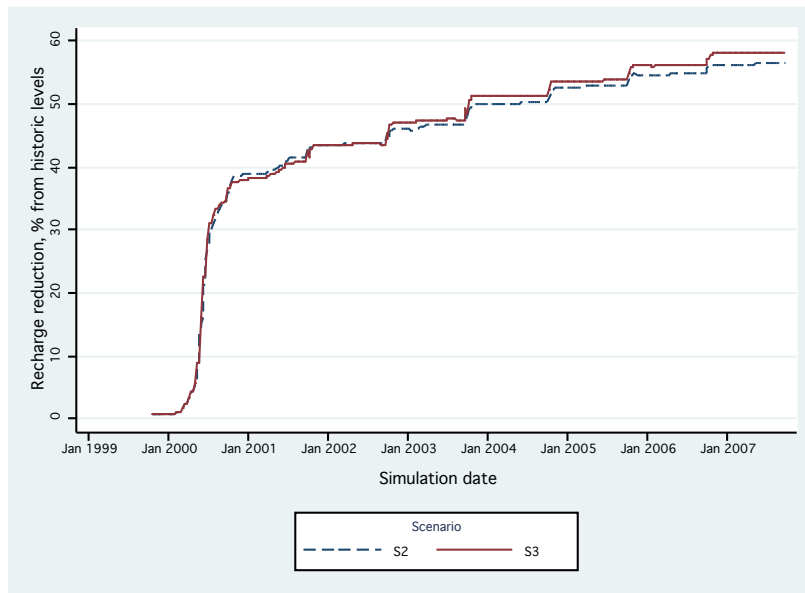


Figure 6.39: Recharge reduction. % reduction from historic recharge levels, weekly time series, comparing scenarios 2 and 3. Scenario 3 enjoys higher levels of lining relative to S2, indicating that preferences being non-static acts to increase the value of ditch lining to ditch managers. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

In the Hypothesis 2 discussion, I pointed out how an inertial effect could be generated by positive feedback on ditch sealing. The effect this had was to promote sealing as a preferred

option for the remainder of the simulation, but this was only possible because the agent S-factors were not able to shift. Now, in scenario 3, since the values were in fact able to change with time, we see a different result: instead of the dominance of sealing, we see a much more complex interplay between sealing and lining with time. The self-reinforcement we saw in scenario 2 could not take place because even though the sustainability rating of the sealing action no doubt continued to be high, agent sustainability factors were dropping for other reasons, and the greater ratings for sealing were weighted as less important. There is greater complexity in these traces than there is space here to analyze, but the key lesson from the sealing/lining question (as well as the earlier discussion of the clearing success ratio) is that allowing for attitudinal change sometimes subverts the objectives of the original incentives. And, of course, it underlines the dangers in leaving out dynamics of behavioral change from models of complex human-environmental systems.

The two key points to draw from this discussion so far are that (1) adding informational incentives to a neoclassical WQTM still results in better sustainability outcomes than the plain neoclassical WQTM, but that (2) agents capable of attitudinal change can reduce the overall performance of the WQTM. Most of this performance decrement is due to relative declines in the sustainability factor across the farming agent population (see figure 6.40; compare with figure 6.41).

The decline appears to be due to an initial adjustment process whereby agents initially become more pro-self in their attitudinal orientation, but then move back towards more sustainable behavior with time. It is important to emphasize that the WQTM in scenario 3 still shows better sustainability benefits with a heterogeneous population than the neoclassical version. This provides robust support for the core theory at the heart of all these hypotheses, since the assumptions in 3 are becoming more and more realistic.

### 6.4.3 Hypothesis 3: parameter variation

The parameter variation experiments show that agent risk strategy has the potential to impact institutional performance under the specific conditions seen in S3. While the trend across most metrics is to see institutional performance improve in scenario 3 over 1b, the effect of varying
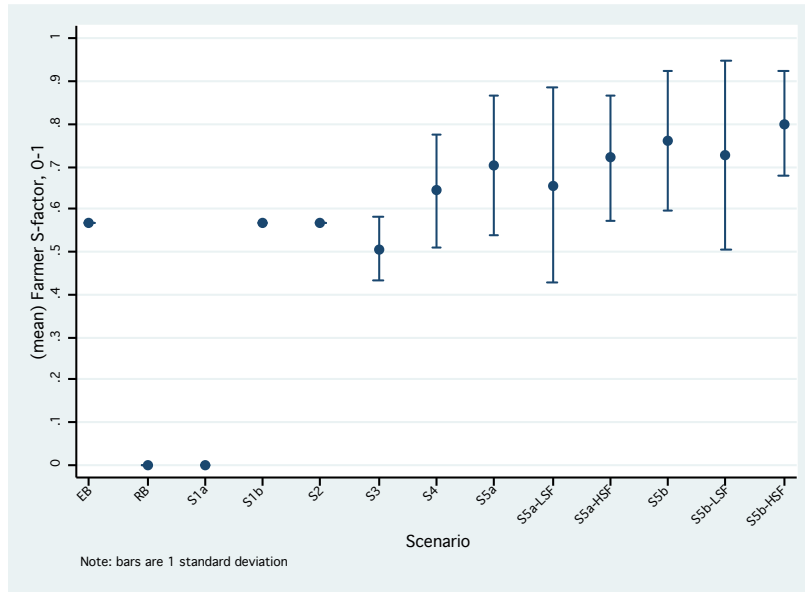
Figure 6.40: Mean farmer S-factor (0.0-1.0), across all scenarios. The lower (more economically rational) S-factor drives less sustainable agent behavior in S3 relative to S2. Note that bars are 1 standard deviation, and indicate variation with time rather than error. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.
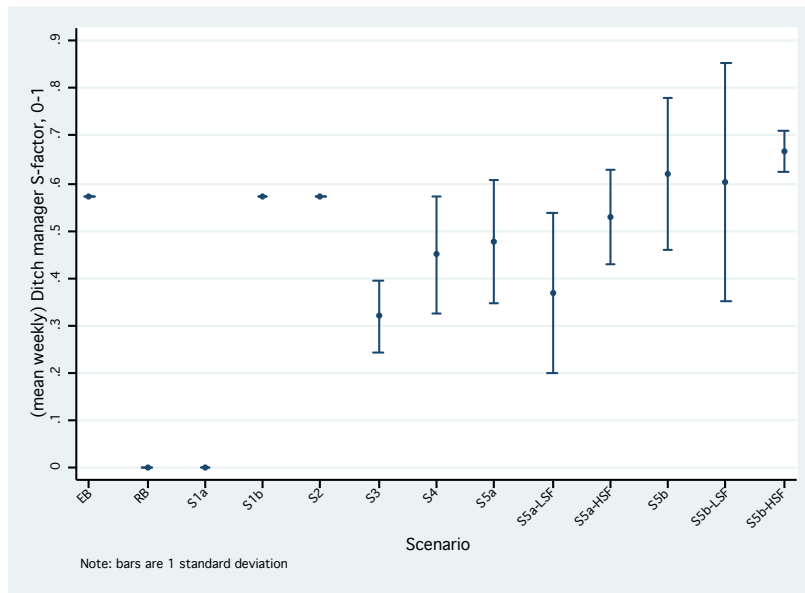


Figure 6.41: Mean ditch manager S-factor, across all scenarios. The lower (more economically rational) S-factor drives less sustainable agent behavior in S3 relative to S2. Note that bars are 1 standard deviation, and indicate variation with time rather than error. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

risk strategy to the extreme (either monolithically risk-seeking or risk-avoiding) is to reverse the trend. Additionally, running the WQTM with risk seeking agents results in much poorer

institutional performance in either 1b or 3 relative to the baseline, while running it with risk avoiding agents actually improves institutional performance relative to the baseline. These are expected results: risk strategy is one component of the agent logic driving decisions to sell or purchase quota, so that more cautious agents will post less quota and also purchase less quota. It also appears that the combination of risk seeking behavior, the capacity for attitudinal change, and a WQTM providing informational feedback, is to significantly reduce market efficiency.

Despite the sensitivity of results to risk strategy, we should note that even under extreme variation of environmental and economic response factors, results continue to indicate support for hypothesis 3. Weaker support from social and institutional metrics is related to specific aspects of the logic (in the case of clearing success ratio) and broader lack of sensitivity to model dynamics (in the case of conflict and other social metrics).

### 6.4.4   Hypothesis 3: real world implications

The extent to which we draw real world implications from these results depends on how much faith we have in the joint reinforcement/cognitive dissonance model of attitudinal change. As I discussed in Chapter 2, there is considerable empirical support for the existence of these dynamics in real human psychology. I also acknowledged, however, that these undoubtedly do not capture all the potential mechanisms of attitudinal change. To address some of the resultant uncertainty, I included a probabilistic parameter that controlled whether agents were in fact susceptible to attitudinal change, and parameterized this using the results from psychometric testing.

Assuming some empirical relevance, these results continue to add weight to one of the core assertions of the thesis: that non-traditional WQTMs provide more sustainability benefits than traditional WQTMs in simulated cases where populations of actors are variable in their attitudes and behavioral strategies. Since empirical evidence argues that most real world cases do involve attitudinally and behaviorally variable stakeholders, the implication for institutional designers is, once again, that their institutions may be more effective if they include alternatives to the traditional set of implicit financial incentives.

These results show that incorporating non-traditional incentives in a traditional market institution can have counterintuitive results. Instead of fostering sustainable attitudes among the agent population, the WQTM incentives initially encouraged unsustainable attitudes among farmers. But as we saw, once these attitudes became particularly strongly expressed in farmers, the pendulum began to swing the other way, and we saw individual farming agents start to shift back towards more sustainable attitudes and behavior. Institutional designers should expect some initial dynamism in how real world institutional incentives foster attitudinal and behavioral change, and that these dynamism may not always work in favor of the intended goals. They may also have to deal with a lag of some period between introducing incentives and seeing concrete attitudinal and behavioral change. Such lags can cause political problems, particularly for initially unpopular interventions, and institutional designers might have to work to lower initial expectations and emphasize the potentially slow response of large and complex systems. Both issues — unexpected outcomes and time lags — are products of the interaction of simple interventions with complex actors, and so this implication is relevant to most basin settings.

The data also indicate the importance of understanding the different effects of the same incentives on different agents. We saw this in the comparison of ditch manager and farmer S-factor values over time, which displayed very different trends arising from very different sets of decisions. While the incentives for farmers led to a strong sustainability orientation in the population, among ditch managers the opposite was true. If institutions are to rely on feedback incentives after individual actions, then institutional designers must be aware that the direction and magnitude of effects may be dependent at least partly on how frequently stakeholders get feedback, what decisions stakeholders are typically taking and when they take them seasonally. Since most basin settings are composed of a variety of actors, who vary in the kinds of decisions they take and when, this may be a broadly relevant implication.

## 6.5 Hypothesis 4: exploiting social dynamics in institutional design

In summary, a deeper analysis of the data reaffirms the initial assertion in Chapter 5:

### 6.5.1 Summary of section

I find that the addition of neighborhood awareness, with its new forces of competitiveness and conformism, acts to increase the overall level of environmental sustainability attained by the system, and the overall institutional performance of the WQTM. This chimes with empirical findings that providing neighborhood information on energy use can lead to overall reduced levels of consumption. Behind these improvements in ArkAgent lie increased volatility in agent attitudes and behavioral choices, and a risk that the environmental improvements are coming at an economic cost which will register as unsustainable on the longer term.

In the discussion in section 6.5.2, I first show that the addition of incentives to exploit social dynamics (scenario 4) has different effects depending on the agent type: faster rates of change but no overall shift in attitudinal dynamics for farmers, but genuinely new attitudinal dynamics for ditch managers. I argue that these outcomes are due to the differences in population size for farmers and ditch managers: the smaller population for ditch managers, for example, means changes in attitude are echoed around the network faster and farther, providing for more significant changes in attitude with time. I also show that increased consumptive use due to increased irrigation efficiency is at least partly to blame for reductions in diversions in scenario 4 without corresponding increases in river flow. This echoes earlier findings that decisions taken to reduce salt load may not necessarily increase the levels of water flowing in the system. An important net result of these dynamics is that ditch managers end up with much more pro-social attitudinal orientations, and this drives all sorts of improved sustainability benefits via their decision making. Another interesting finding discussed in this section is that the effect of exploiting social dynamics is to drive up leasing relative to scenarios without these dynamics. I argue that this is because of the higher S-factors among agents, and the availability of surplus water due to reduced diversions for crop irrigation. I also discuss in this section the unfortunate result that the more sustainable behavior by ditch managers also causes them to incur larger debts, because their diminishing concern for economic considerations reduces their ability to appraise project plans with appropriate economic rationality.

In the discussion of real world implications (section 6.5.4), I argue that the results indicate that social dynamics are well worth exploiting through appropriate non-economic WQTM in-

centives. While the representation of such dynamics in ArkAgent is simplistic, the results are still encouraging for institutional designers seeking to improve institutional efficiency. However, I also caution that while exploiting social dynamics appears to be useful, it can also lead to more volatility in responses, and less certainty for a policy maker with regard to what magnitude of change is seen. On the basis of the results for ditch lining, I suggest that the use of incentives to exploit social dynamics can sometimes lead to less economically rational decisions. Institutional designers should be broadly cognizant of the risk that their incentives, particularly applied in a social context, can end up having unexpected results. Finally, I discuss an implication from the ArkAgent results that it may be worth considering the limits on effectiveness for a given incentive. No institutional innovation, however innovative, is guaranteed to move a social system all the way towards a given outcome. Given this implication, the use of complex systems simulation may be useful for exploring how far a given innovation might be able to shift behavior.

### 6.5.2 Hypothesis 4: reviewing the arguments

In Chapter 5, I argued that the overall sustainability scoring information strongly supported Hypothesis 4 (that an institutional incentive which uses social mimicry and competitiveness dynamics to reinforce behavioral change towards sustainability, will see more sustainability benefits than an institution without this mechanism). Refer to section 4.2.6.3 in Chapter 4 for clarification on the nature of the "competitiveness" mechanism. On the environmental, social and institutional fronts, scenario 4 (including mimicry and competitiveness dynamics) outperformed scenario 3 (which omitted these dynamics). Why might we expect this kind of improvement? Recall from Chapter 4 the fourth behavioral mechanism (figure 6.2), where $S$ is the agent's S-factor, $SR_n$ is the mean rating received for a given action by the agent's neighbors, and $BMSF$ is a constant helping determine the magnitude of response.

$$m = \frac{1}{(|(S - SR_n)| * BMSF)} \tag{6.2}$$

As described in that chapter, this equation encapsulates both competitive dynamics (when

the neighborhood rating exceeds the agent's rating, this drives the agent's attitudes towards exceeding that rating) and conformism (when the neighborhood rating is below the agent's rating, this drives agent's attitudes down towards that rating). The effect for farming agents, as we see in the comparison between scenario 3 and scenario 4 data in figure 6.42, is to accelerate the effects of negative or positive attitudinal change. Scenario 4 shows a faster increase in pro-self attitudes, a deeper initial trough, faster recovery, and then a higher overall level than scenario 3. Figure 6.43 plots the same data to emphasize the difference in mean and variance: in both cases higher for scenario 4. In other words, neighborhood dynamics do not generate fundamentally new trends in attitudinal change, but both enlarge the range of possible change and weight average attitudes towards the more sustainable, pro-social end of the scale.
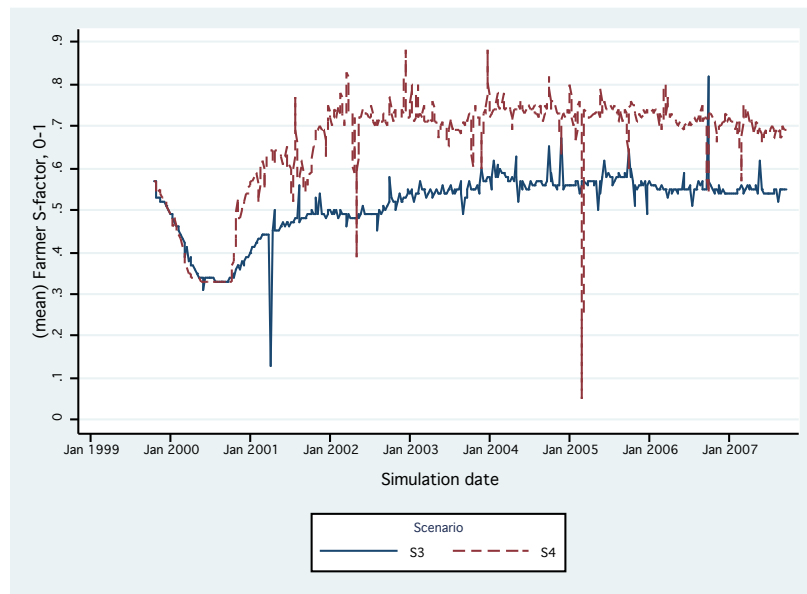


Figure 6.42: Mean farmer S-factor, weekly time series, comparing scenarios 3 and 4. The addition of social network effects accelerates the effect of negative or positive attitudinal change. Note that S4 is the top trace, and S3 is the bottom trace. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

The addition of competitive and conformist dynamics has a dramatically different effect on ditch managers. As we see in figure 6.44, the scenario 4 trend for ditch managers is now much closer to that of the farming agents. The plot in figure 6.45 illustrates how the shape of the curve and overall trend has shifted, with mean and variance increasing significantly.

For ditch managers, competitive and conformist dynamics have indeed generated new
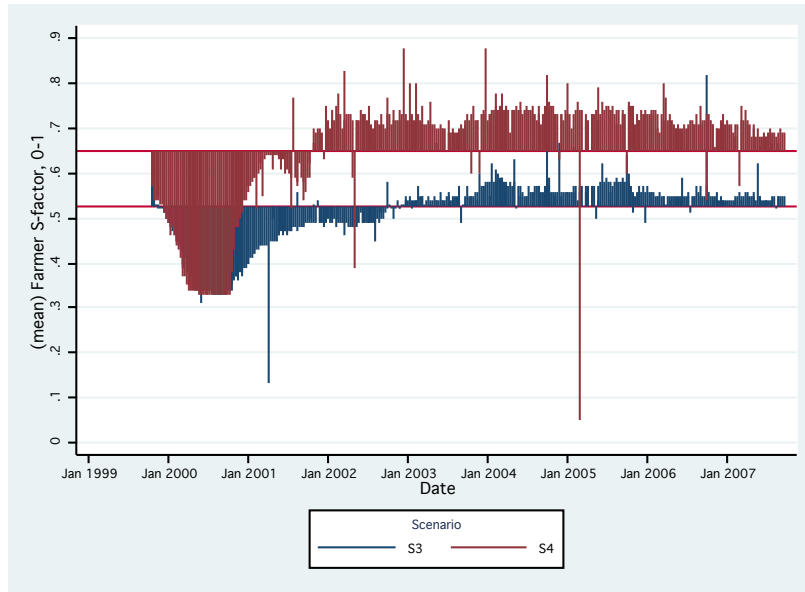
Figure 6.43: Mean farmer S-factor, weekly time series, spikeplot, comparing scenarios 3 and 4. Plotted to emphasize higher mean and variance in scenario 4 relative to scenario 3. Refer to Appendix section 8.1.2 for detailed scenario descriptions.
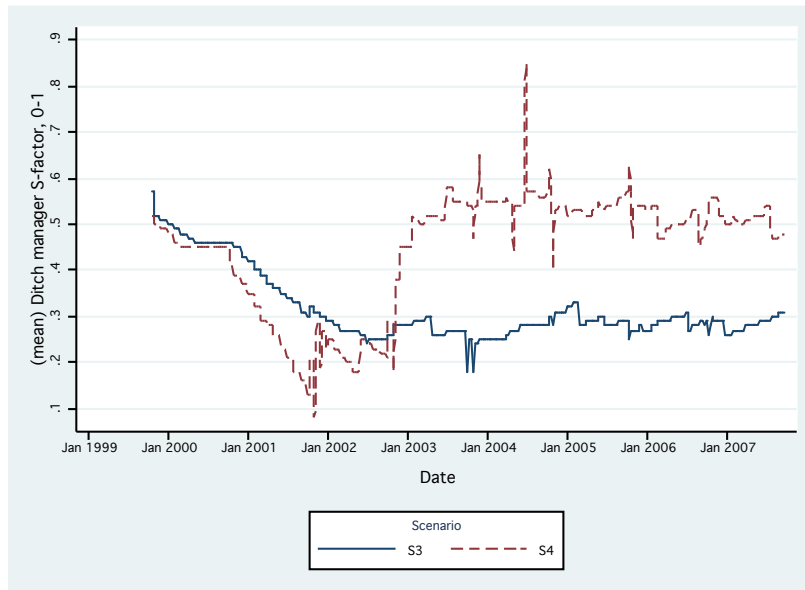


Figure 6.44: Mean ditch manager S-factor, weekly time series, comparing scenarios 3 and 4. Plot shows that ditch manager preferences, while qualitatively similar in overall trend, show considerable differences in mean, variance and timing of trend shifts with time relative to farmers. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

trends in attitudinal orientation over time. Some of this change is due to the effect these dynamics have in accelerating the rate of attitudinal change: we see this in the initial decline in SF
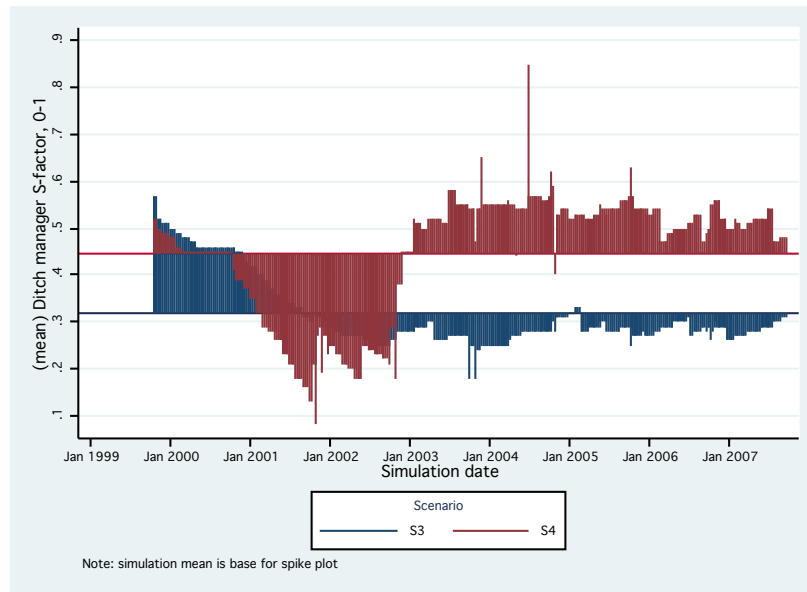
Figure 6.45: Mean ditch manager S-factor, weekly time series spikeplot, comparing scenarios 3 and 4. Compare with figure 6.43 to emphasize the differences in preference evolution with time for farmers and ditch managers. Note that S4 is the upper trace, and S3 is the lower trace. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

values up to around July 2001, and again in the rapid rate of increase around in the late summer of 2002. However, the dramatic shift in SF values upwards after late summer 2002 likely has its cause in the fact that the pool of potential "neighbors" for ditch managers is far smaller than for farmers, and "neighbors" may not necessarily be geographically proximate. The first issue is a reality of numbers: there are only 15 ditches simulated in ArkAgent, so only 14 potential neighbors for any ditch manager. The second issue relates to how ArkAgent simulates neighbors for the farming agent population. Interviewing data suggested farmers interact most closely with farmers within geographic reach, through chance encounters on county roads and in local towns. A farmer in the eastern LAB is unlikely to meet with a farmer from the western LAB as frequently as he would a farmer from his own area. The basin is simply too large for that, and for obvious reasons the lives of farmers are not peripatetic enough to balance out this effect. According to the same interview data, however, ditch managers have regular contacts with other managers that do not necessarily relate to geographic proximity. The apparently simple task of delivering water to shareholders may in reality involve very complex informal and virtual "trades" of water among ditches with widely separated diversion points. So a ditch

manager in the far eastern part of the system may be communicating with managers far to the west. Consequently, ArkAgent simulates the "neighborhood" of ditch managers to be unbiased by geography.

The effects of the smaller pool and the relative unimportance of geographic bias means that shifts in attitudinal orientation propagate even faster and more dramatically among the ditch manager population. In other words, the sphere of influence of any ditch manager can be multiplied by the fact that the connectivity of the overall social network is far higher than that of farmers. Once a trend towards more sustainable behavior starts, we can expect the trend to accelerate as ditch managers receive information on their "neighboring" ditch managers, who in turn share that information with their neighbors. Interestingly, the small size of the ditch manager pool means that ditch managers may end up being influenced by their own orientations, reflected back through the small but highly connected social network in the community.

These behavioral trends will naturally have a strong effect on the salt loads and diversions seen in scenario 4. Reviewing the environmental data for salt load (figure 5.7), we see how scenario 4 has an annual average sum around 15,000 tons less than scenario 3. Reviewing the data for flow (figure 5.8), however, and we see that scenario 4 has one of the lowest annual average flow levels. The lowest, in fact, among all empirically parameterized runs. Reviewing seepage and recharge plots (figures 6.46 and 6.47) indicates that scenario 4 mostly shows higher levels of seepage and recharge reduction than scenario 3.

The differences may appear slight, but over time such differences add up to a considerable reduction in total tonnage. Following the trend seen in all other scenarios, most of the reduction appears to come from ditch manager decisions on sealing and lining. Reviewing the time series for these data (figure 6.48), we see that not only does scenario 4 have — on average — higher levels of sealing than scenario 3, but also a clear lead on ditch lining. Recalling the S-factor data, we can explain the much increased levels of lining in scenario 4 on the significantly different evolution of ditch manager sustainability factors over time. Lining is a much more expensive activity and so additional weighting must be given to its sustainability benefits for it to be favored in ditch manager decision making. This additional weighting takes place when the ditch manager S-factor values start to shift upwards after 2002.
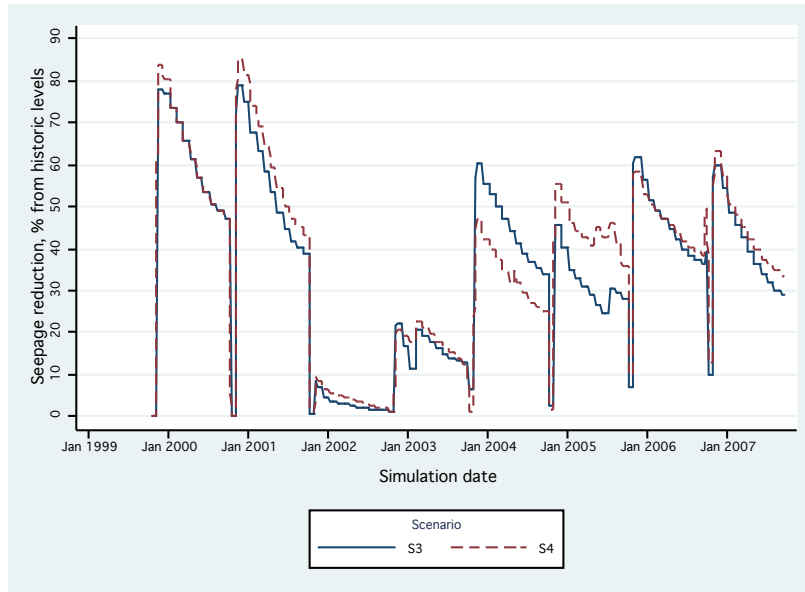
Figure 6.46: Seepage reduction, % from historic, weekly time series, comparing scenarios 3 and 4. Reduced seepage (increased seepage reduction) in scenario 4 may help explain reduced salt loads in that scenario. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.
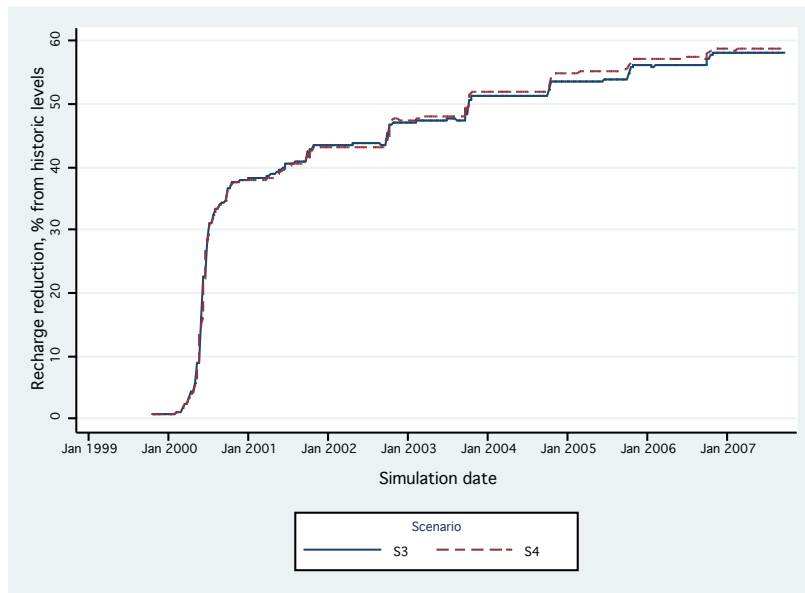


Figure 6.47: Recharge reduction, % from historic, weekly time series, comparing scenarios 3 and 4. Reduced recharge (increased recharge reduction) in scenario 4 may help explain reduced salt loads in that scenario. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

Returning to the flow issue, and reviewing an earlier plot (figure 6.1), we see that scenario 4 actually shows considerably lower diversions than scenario 3, yet river flows do not increase
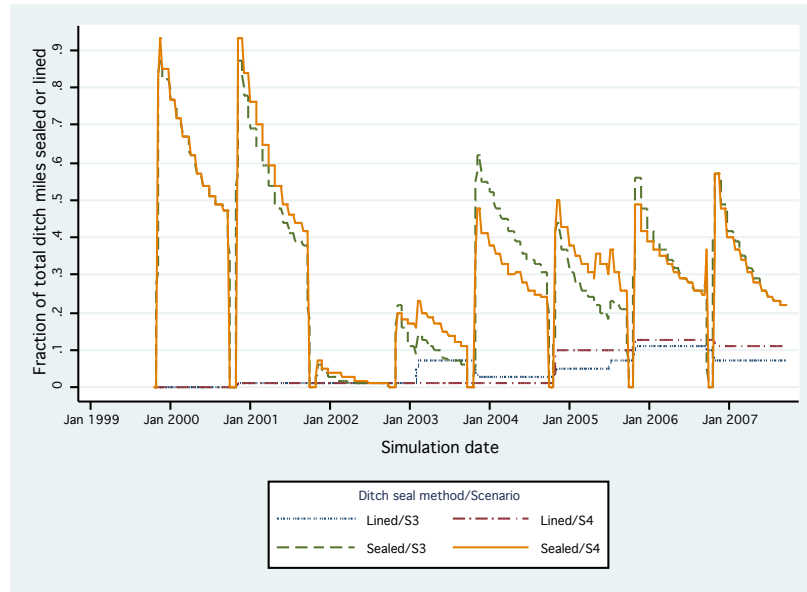
Figure 6.48: Fraction of total ditch miles sealed and/or lined, weekly time series, comparing scenarios 3 and 4. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

proportionately. This raises an issue we have not touched on so far: the potential for increasing irrigation efficiency (and reducing salinity levels) without actually increasing river flows. This is a widespread fear among farmers in many irrigated systems: by shifting towards more efficient application systems, consumptive use will actually increase. Consumptive use is the portion of water applications to fields which are lost through incorporation in plant tissue and evapotranspiration. Consumptive use is therefore not returned to the river through surface runoff or ground water recharge. Sprinkler systems, for example, apply water more efficiently such that less runs off the field and relatively more water is taken up by the plant and lost through evapotranspiration. The net result is that while less water is being diverted, more water is being "lost" from the system so that downstream users see lower flows. As we see in figure 6.49, application efficiencies in scenario 4 increase over scenario 3 beginning in late 2002, and consumptive use (figure 6.50) responds accordingly.

The data echo earlier findings that increased sustainability for salt load does not necessarily imply increased sustainability for river flow. While some salt load reduction techniques also make more water available in the system (like reduced diversions), other techniques can in fact reduce the water available.

251
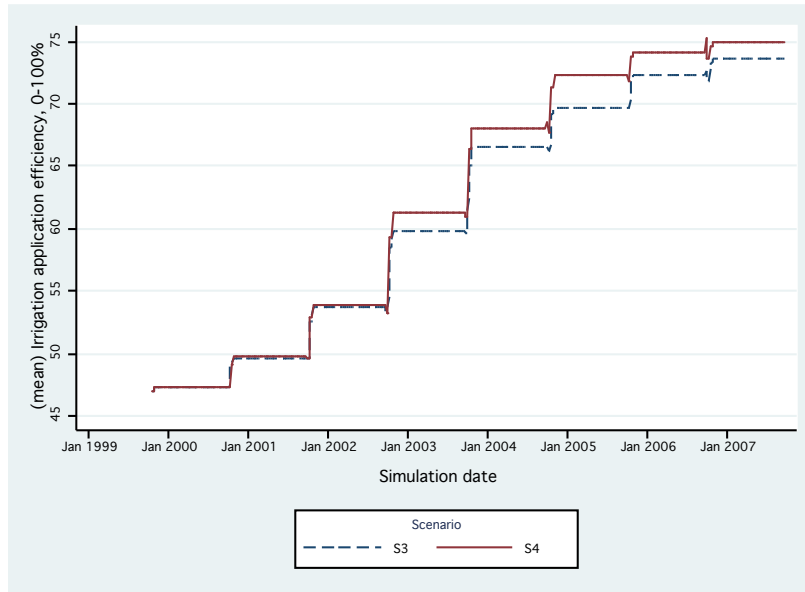
Figure 6.49: Farm irrigation efficiency (0-100%), weekly time series, comparing scenarios 3 and 4. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.
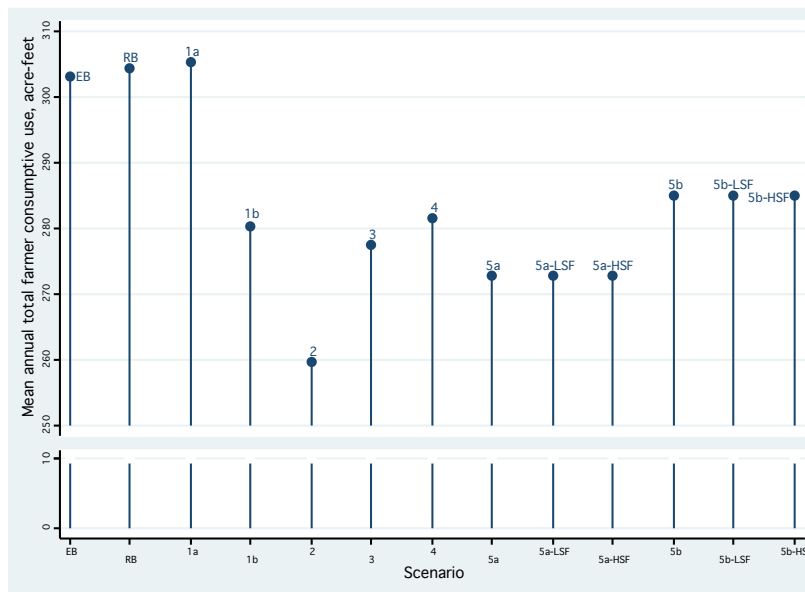


Figure 6.50: Mean of cumulative annual on-farm consumptive use (acre-feet), across all scenarios. Increased application efficiencies in scenario 4 relative to scenario 3 drive up consumptive use in the same way. Note scale break: this is included to place emphasis not on absolute values but on inter-scenario variation. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

Turning now to the institutional data, we see that the WQTM in scenario 4 enjoys the lowest clearing success ratio (figure 5.10) and the highest volume of quota traded (figure 5.11) among

252

all empirically-parameterized scenarios. The number of unique traders does not vary greatly from the previous scenarios, a result I address later in the General Discussion section. Underlying the improved clearing success ratios and quota trade volumes in scenario 4 may be the lower trough that ditch manager and farmer sustainability factors reach in the first few years of the simulation. As discussed earlier, a higher level of economic rationality is linked to better clearing success ratio, since how responsive agents are to market signals depend strongly on their economic rationality. Reviewing another dimension of institutional performance, quota status, figure 6.51 illustrates that there is really very little distinction between scenarios 3 and 4. If anything, scenario 4 shows slightly less severe quota deficits in the latter part of the simulation period.
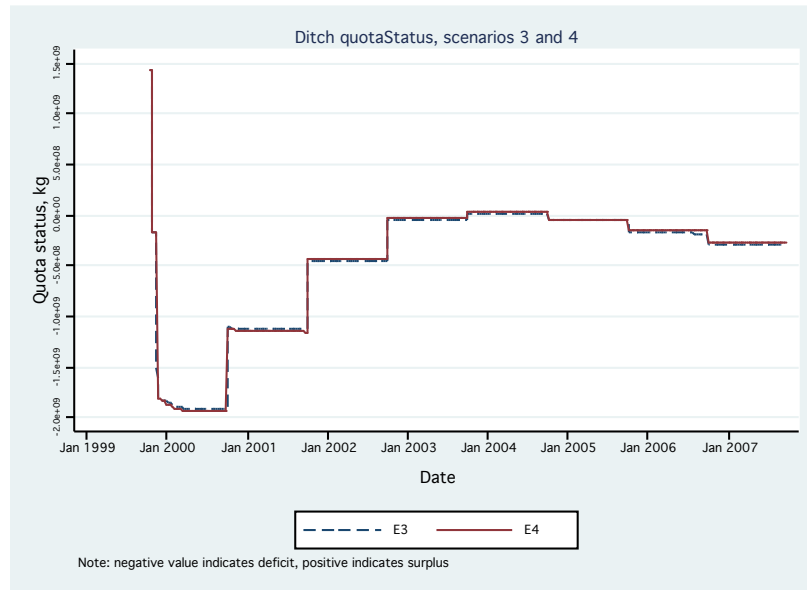


Figure 6.51: Status of allocated quota among trading agents, weekly time series, comparing scenarios 3 and 4. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

Reduced deficits means that agents are exceeding their pollution quotas by fewer kg of salt each year. It also means, in theory, that fewer fines will be incurred by agents as a result. Yet we see higher fine levels in scenario 4 relative to scenario 3. As in earlier discussions, it is likely that while fines are being levied on less quota exceedance in volumetric terms, fines are being levied at a higher per unit value. Recall that fines are tied to quota prices to ensure that the cost of exceeding quota is higher than the cost of purchasing quota to avoid exceedance. Indeed, in

figure 6.52 we see slightly increased quota prices in scenario 4.
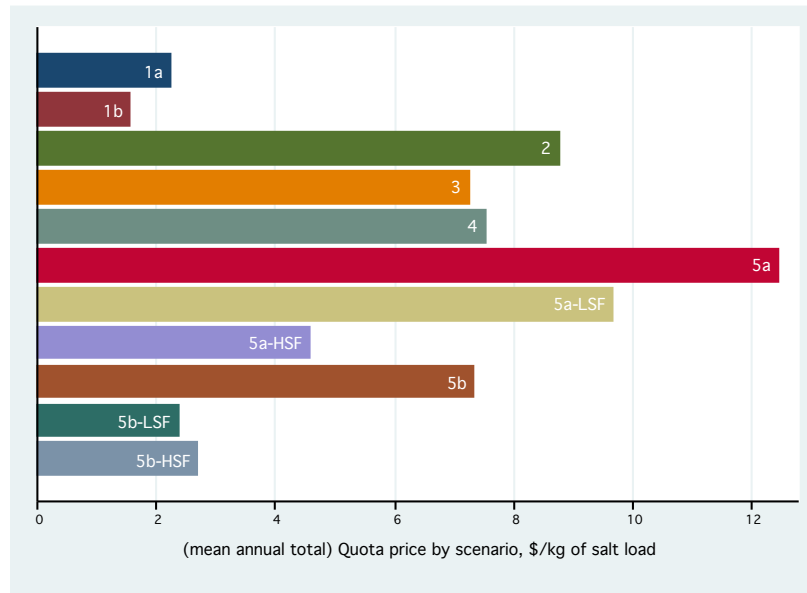


Figure 6.52: Price of 1 kg of quota ($), across all scenarios. Quota price is slightly increased in scenario 4, relative to scenario 3, which drives up fine amounts and accounts for additional levels in this scenario. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

Quota price setting is a relatively complex logic. It factors in agent perceptions of scarcity, the agent's sustainability factor, available funds (on the part of purchasers), risk tolerance (on the part of sellers), and other factors still. With such complex logic, it is not always easy to tease out particular causes for trends. However, in general, higher perception of scarcity should drive up prices, while the reverse should happen for lower scarcity (assuming that risk strategy does not shift during the simulation, which in ArkAgent is the case). Furthermore, agents with higher sustainability factors will be less sensitive to scarcity, while agents with very low sustainability factors will be highly sensitive. This may be a potential contributor to the differences between scenarios 3 and 4: while the mean S-factor value for both farmers and ditch managers in scenario 4 is higher than in scenario 3, the S-factor time series has a lower bottom point in scenario 4. The likely effect of this is to heighten the sensitivity of agents to scarcity, and drive up quota prices. The difference is not significant, though, and so we do not expect to see much of a shift in quota price.

Finally, turning to the economic data, we see from earlier figures 5.12 and 5.13 that agents in

scenario 4 enjoy slightly higher incomes but lower asset totals over time. The higher income is likely due to a combination of higher incentive payments (figure 6.19), increased debt accumulation (as before) and increased municipal water leasing in scenario 4 versus 3. It is certainly not due to higher crop yields, as we see in figure 6.53.
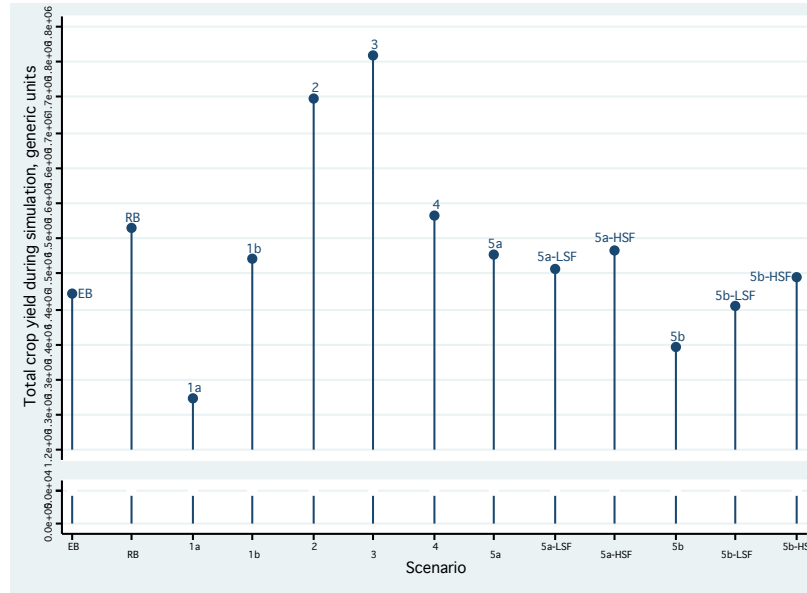


Figure 6.53: Total crop yields (in generic units) from harvests during the simulation, across all scenarios. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

Like borrowing, gains from municipal leasing represent new injections of wealth into the system. It appears that the effect of incorporating neighborhood ratings is to drive up the amount of leasing that takes place, to some of the highest levels among all scenarios. The increase is certainly significant relative to scenario 3: around 35,000 more acre-feet were leased annually in scenario 4 compared with scenario 3. Drivers behind this increase include the much higher S-factor values in scenario 4, and the additional availability of spare water with increased application efficiencies and reduced diversions. Once again we see that these factors are enough to overcome some of the highest agricultural water values seen in any scenario (figure 6.54).

Much as with scenarios 2 and 3, the combination of non-traditional incentives, a WQTM and behaviorally heterogeneous agents seems to be driving increased indebtedness (figures 5.17 and 6.6). In this case, debts for for scenario 4 even exceed liabilities for scenario 3, particular for
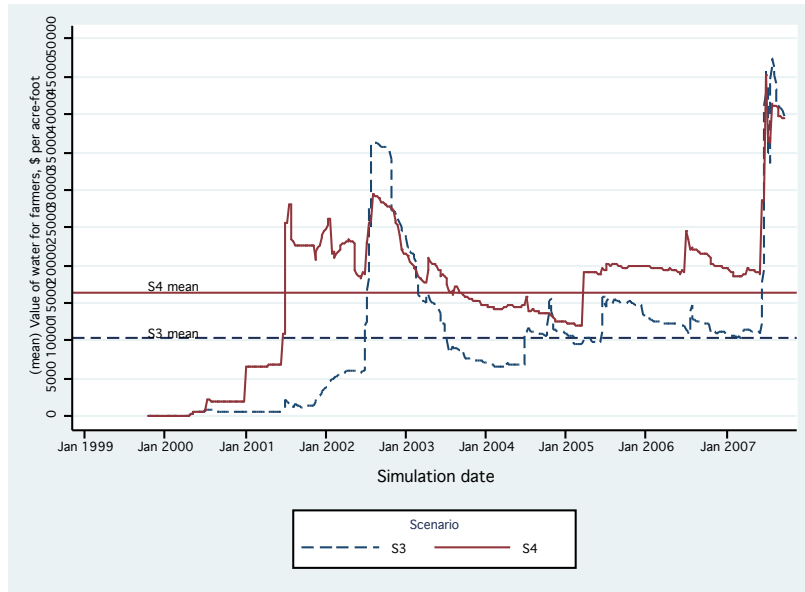
Figure 6.54: Agricultural value of water ($ per acre-foot), weekly time series, comparing scenarios 3 and 4. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

ditch managers. While there is a slight increase in incentive payment in S4 relative to S3 (figure 6.19) it does not appear to be significant enough to account for the increased borrowing in S4. An additional explanation may be that the increased level of capital projects in S4 (more ditch miles sealed and lined) is driving additional indebtedness.

### 6.5.3  Hypothesis 4: parameter variation

The dynamics seen when comparing the results of parameter variation on scenarios 3 and 4 do not generate any new findings relative to previous experiments, and the plots are not reproduced here. In general, the results of these variant experiments lend weight to Hypothesis 4, by showing it to be broadly robust to variation in key parameters.

### 6.5.4  Hypothesis 4: real world implications

The results for hypothesis 4 shows how effective it can be to exploit social awareness through institutional incentives. How much potential benefit we assume for social awareness incentives depends on how much we trust the behavioral model in ArkAgent, and in turn how much we trust the empirical results on which this model is based. Since most irrigation systems embody

complex social networks both formal and informal, there are likely to be many settings where institutional designers can exploit the power of social networks to magnify the effectiveness of incentives. Indeed, many realms of business already work hard (with considerable success) to exploit these networks.

The ArkAgent model suggests that institutional designers may not be able to easily gauge the magnitude of response to exploiting social networks. They may, however, be more confident about the direction of change, if they have a reasonable grasp on the right social interaction model for the setting. ArkAgent results show that, for a competitiveness/conformism social interaction model, a socially networked response to incentives shifts agent attitudes towards the desired end of the spectrum, while increasing the temporal variance of attitudes. As agents receive conflicting and confirming impressions from their neighbors, their own attitudes become more volatile. Sensitivity testing of ArkAgent shows how we can force variations in the size of the incremental attitudinal change with every neighborhood feedback received, yet still reproduce the overall trend (figure 6.55).

Finally, the results for Hypothesis 4 echo implications from earlier hypotheses: different agents differ in their response to specific kinds of incentives, and this may indicate potential for modifying incentives to exploit these differences. For the LAB case, the ArkAgent results argue for targeting ditch managers with social networking incentives, since these actors have a particularly small and close-knit professional community, which may magnify the effectiveness of incentives. Since real world actors in modern irrigated basins in the American West show at least as many differences as those represented in ArkAgent (it being, after all, a gross simplification of reality), this is an implication likely to have broad geographic relevance. Of course, ArkAgent results very likely underestimate the true complexity of social interactions in the LAB and elsewhere. This means that the present implication will be most relevant in cases where interaction among actors is highly circumscribed by actor type. For example, the wisdom of targeting specific stakeholders is less clear if all stakeholders communicate freely regardless of their "type".

The results relevant to Hypothesis 4 show that while the WQTM can drive positive behavioral change (more ditch lining, for example), institutional designers in the LAB should
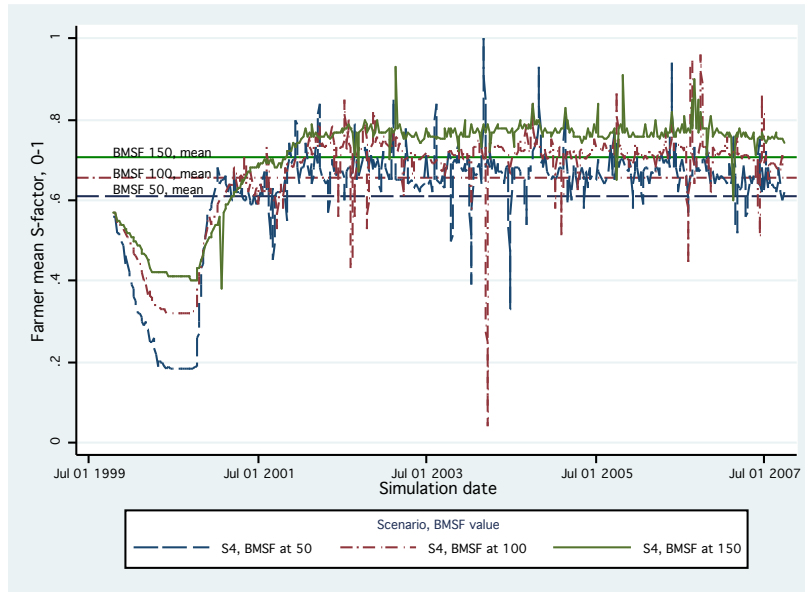
Figure 6.55: Farmer S-factor under parameter variation ("BMSF" constant), weekly time series. The BMSF constant contributes to the magnitude of behavioral change in response to informational stimuli, but does not affect the overall shape of the curves. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

not necessarily assume that this response will always be economically rational. Scenario 4 assumes varied attitudinal orientations, some more economically rational than others. As a result, ditches implement capital projects like ditch lining to varied levels of economic sustainability: the more economically rational ditches avoid assuming too much debt; the less economically rational ditches assume more debt, and this can be particularly burdensome on the decadal scale. Even if we have a healthy dose of skepticism regarding how economically irrational real stakeholders will be, we clearly cannot assume absolute infallibility in things like long term capital projects planning. This is just as likely to hold in other basins as in the LAB. The ArkAgent results highlight the importance of playing out the second order effects of institutional interventions: incentives can promote good behavior, but particular enthusiasm for certain good behaviors can lead to bad outcomes on the longer term. Institutional designers could make provisions — extra grant money, perhaps — to handle any such eventualities.

Scenario 3 and 4 evidence demonstrates how the effectiveness of a given incentive type (i.e. in fostering attitudinal change) may have an upper limit. In the ArkAgent model a given incentive has a finite capacity to shift attitudes in a given direction. This is not an unexpected

result, since incentives which do not create fundamentally new decision options for agents will have natural limits on how far they can perturb attitudes from the baseline. This has corollaries in many interventions in the real world where what is sought is a shift in existing behavior, not a shift to entirely new behaviors. One example is fostering reduced energy consumption: an electricity utility does not need energy consumers to start engaging in entirely new forms of energy consumption; it will be happy with alterations in existing consumption rates. The results for Hypothesis 4 suggest institutional designers may want to consider whether their proposed incentives will drive entirely new behaviors, or whether they will simply shift the parameters of existing choices and if so what the potential ceiling of the shift might be. Since forecasting such outcomes is challenging with standard policy analysis tools, this is another instance where the use of complex systems simulation model may be particularly beneficial. While it is unlikely to accurately predict outcomes, such a model — like the example provided by ArkAgent — can shed light on envelopes of possibility and foster discussion over potential alternative outcomes given different assumptions.

Because sustainability is such a complex metric, a given incentive will not necessarily align gains in all areas of sustainability. The ArkAgent model shows that while salt load is reduced by institutional interventions, flows in the river may be negatively impacted. Institutional designers in the LAB could put some care into determining alternative incentives that might address the problems arising when farmers simply make more use of newly surplus water (from improved application efficiencies). The fear of water right abandonment with increased irrigation efficiency, while not operationalized as such in ArkAgent, is a real feature of the LAB. It may be necessary to shift the legal basis of abandonment in Colorado to improve the chance of gaining on both flow and quality bases at the same time. This may also be true of other basin settings, since many irrigation systems with prior appropriations-driven allocations suffer from the same perverse incentives.

Shifting behavioral attitudes can have cascading effects across different decision sets. In ArkAgent we saw how a shift in a single attitudinal variable — the S-factor — caused improvements in application efficiency after irrigation equipment was upgraded; this in turn was a partial contributor to reductions in diversion volume; which in turn made more water avail-

able for leasing, which made leasing to municipalities a more viable option. Real world institutions in the LAB that try to foster changes in attitude and behavior may find that cascading effects can have multiple potential benefits. As I suggested in the preceding paragraph, perhaps equally possible is that cascading effects may not be so positive. The cascade arises because of the strongly interconnected nature of water management decisions in the LAB: there are rarely decisions made upstream which do not influence people downstream, and the infrastructural interdependencies in the LAB can even spread dependencies out of the basin. Since most hydraulic societies share these characteristics, this implication may have broad relevance.

## 6.6   Hypothesis 5: non-economic incentives in the face of competing pressures

### 6.6.1   Summary of section

Even when agents are assumed to be aware of and susceptible to influence by information favoring more self-regarding and less sustainable behavior, this may not reduce the effectiveness of using a non-traditional incentive to foster more sustainable behavior. This is due to the fact that (a) even if an agent is aware of such information, this does guarantee the information is included in decision making, and (b) the fact that the various incentives forms are competing against one another, and so tendencies towards self-regarding behavior may be well balanced. In fact, a non-traditional market institution using diverse incentives may be even more effective at encouraging sustainable behavior when we allow competing incentives. It is not fully clear why this may be the case, but it is likely that the particular context of the LAB provides for more than a 50-50 chance that a given decision will have both economic and sustainability benefits (e.g. application efficiency improvements). Further, since the empirical parameterization starts the population slightly favoring more sustainable outcomes, this may generate an inertia in favor of more sustainable outcomes over. However, examining time series data suggests that initial orientation (S-factor parameterization) matters less and less as time goes by: institutional structures have more influence on agent attitudes over time.

In the discussion in section 6.6.2, I offer a number of explanations for some complex and

contradictory results. I treat hypothesis 5(i) first. Digging into the raw data, I suggest that the clear differences between S5a/S5b and S4 are likely due to different trends in attitudinal change among the agent population: opportunity cost rating (OCR) information appears to drive agents to *more* sustainable attitudinal orientations. In other words, even when agents are assumed to be aware of and susceptible to influence by information favoring more selfish behavior, this does not necessarily reduce the effectiveness of the non-traditional WQTM in incentivizing more sustainable attitudes and outcomes, and may even increase its effectiveness. However, detailed results show that sustainability benefits tend to decrease slightly in some cases (e.g. salt load) and increase in others (e.g. river flow) in scenarios 5a and 5b relative to S4. I explain these complex results by suggesting that the OCR information sometimes supports and sometimes acts against sustainable decision making: this is the result of having multiple routes to sustainable outcomes, and multiple decisions with different combinations of economic and sustainability benefits. With regard to hypothesis 5(ii), I discuss the result that the initial parameterization does appear to influence the overall sustainability benefits in a scenario, when opportunity cost information is factored in. But by reviewing time series results I show that this is unlikely to hold over the longer term: the outputs of the very differently parameterized LSF and HSF runs evolve toward one another during the simulation period, and seem likely to get even closer if we were able to run the simulation further into the future. I suggest that this indicates the institutional structure is a more powerful control on agent attitudinal change than the initial parameterization of agent attitudes. This does not preclude the reality that starting with a more pro-self or pro-social attitudinal orientation across a population imparts inertia to that orientation, which can take a while to overcome.

In the discussion of real world implications (section 6.6.4), I suggest that the hypothesis 5 results imply that non-traditional and explicit incentive packages in market institutions may have the potential to dramatically improve sustainability outcomes across a wide range of social conditions. On the basis of how the data shows overall outcomes to be less dependent on the initial attitudinal orientations of actors, and more on the package of incentives used, I argue that institutional designers should be less concerned with the fact that a population shows a tendency towards one view versus another. Instead, they should focus on how effective their

proposed institutional form will be in incentivizing appropriate behavioral choices in the long term; in the short term they can probably expect some unexpected and undesirable results, but the chances are that over the long term, the typical stakeholder response to the incentive package may evolve towards a more productive equilibrium.

### 6.6.2 Hypothesis 5: reviewing the arguments

The agent behavioral models simulated in scenarios 5a and 5b are the most sophisticated, and likely the most realistic, of all the models. Each scenario was effectively split into three: the base case, with agent sustainability factors empirically parameterized (S5a, S5b); the low S-factor case, where agent S-factors were distributed at the low end of the scale (LSF); and the high S-factor case, where agent S-factor values were distributed at the high end (HSF). This enabled the analysis first of how the system responded with the introduction of competing incentives with and without the exploitation of social dynamics (S5b versus S5a); and second of how outcomes in each case depended on our assumptions as to the pre-existing set of agent attitudes.

Hypothesis 5(i) argued that the inclusion of competing informational incentives (i.e. pro-economic information) would alter the capacity of a non-traditional WQTM to achieve sustainable outcomes. Hypothesis 5(ii) enlarged on this by suggesting that the kind of outcome we could expect (i.e. more or less sustainable) would depend on the initial distribution of S-factors within the agent population. The initial finding in Chapter 5 was that the first and second parts of Hypothesis 5 were only weakly supported.

Beginning with hypothesis 5(i), and starting with the most important response factor from a sustainability perspective — salt load — the effect of providing opportunity cost rating (OCR) information is to increase loads in both S5a and S5b relative to S4. On the basis of the agent behavioral mechanisms, this is expected, since the OCR information acts to drive agents to make more selfish decisions and show less concern for basin-wide sustainability benefits. But what really underlies this result? As per previous scenarios, we can begin by reviewing the recharge and seepage reduction data for S4, S5a and S5b (figures 6.2 and 6.3). Figure 6.2 shows that seepage reduction is actually lower for both S5a and S5b than S4. The difference is only a

few percent in both cases, but may well be enough to contribute the few extra thousand tons of salt per year we see in the salt load data. The recharge reduction plot supports this conclusion, since we see either parity or reduction in S5a and S5b from S4. Based on the seepage data and the plots comparing sealing and lining fractions among scenarios, most of this reduction comes from a decrease in the amount of sealing taking place (figure 6.56). Interestingly, the lining data shows no such reduction (figure 6.57), with higher rates of lining in S5a than S4, and values holding steady in S5b.
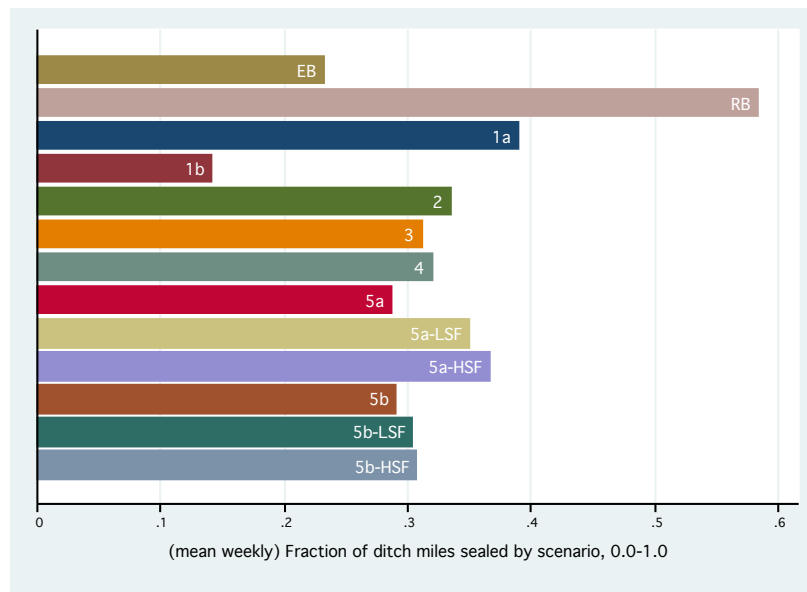


Figure 6.56: Fraction of ditch miles sealed, across all scenarios. Plot included here to show how sealing activities reduce in S5a and S5b relative to S4. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

What might be making a ditch manager agent less likely to seal but more likely to line its ditch when opportunity cost information is presented? Given the design of the behavioral mechanisms, the principal explanation is that sealing received a higher opportunity cost rating, while lining received a lower rating. The ArkAgent mechanism for incorporating OCR information is such that higher OCR ratings drive down the sustainability rating of a given action, and the S-factor of a given agent receiving the rating. This is because OCR indicates the potential gain or loss to doing a particular activity relative to doing nothing. Sustainable activities which have a higher OCR will tend to push agents towards away from those activities (both in decisions and attitudinally), since they are being rated as more costly than doing nothing. In
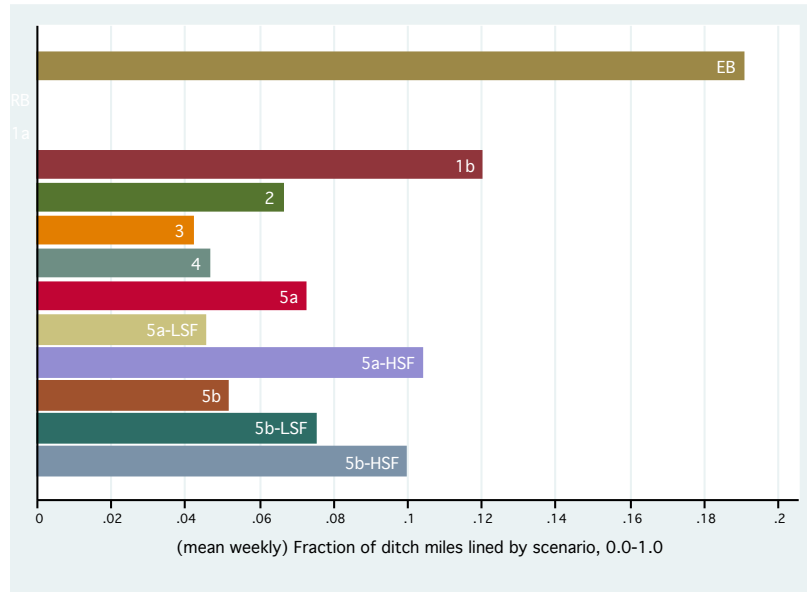
Figure 6.57: Fraction of ditch miles lined, across all scenarios. Plot included here to highlight that while we see reductions in sealing activity from S4 to S5a/S5b, no such change is seen with lining. As a side note, it is interesting to see that the most economically rational agents engage in no ditch lining at all. This speaks to the expense and effort required in lining relative to sealing extensive ditch systems. Refer to Appendix section 8.1.2 for detailed scenario descriptions.

the case of sealing, the shift in ratings had the effect of reducing the overall decision score. In the case of lining, the lower OCR increased its decision score. This result is somewhat counterintuitive, since one might expect sealing to have a relatively low OCR (more benefit than cost): it costs very little, with potentially large returns. Conversely, lining typically has a higher OCR: adding a concrete bed to canals is very expensive, yet the potential gains through reduced seepage are not any greater relative to sealing. What lies behind this unexpected result?

Unfortunately, the resolution of data collection in ArkAgent does not allow the querying of the individual OCR values received by agents. We can, however, examine the mean sustainability factor data for 5a and 5b, since this will provide some idea of the kind of feedback agents were receiving. Figure 6.58 shows that the S-factor ratings for both S5a and S5b were consistently higher than those for S4. It is also clear, however, that fluctuations in attitude are far more dramatic in S5a and S5b: the S5a S-factor time series shows, at one point, a shift from 0.8 to 0.2 and back again in a matter of weeks. I address this higher variance in the anomalies section below.

It appears that the effect of allowing OCR information to reach agents is to promote more
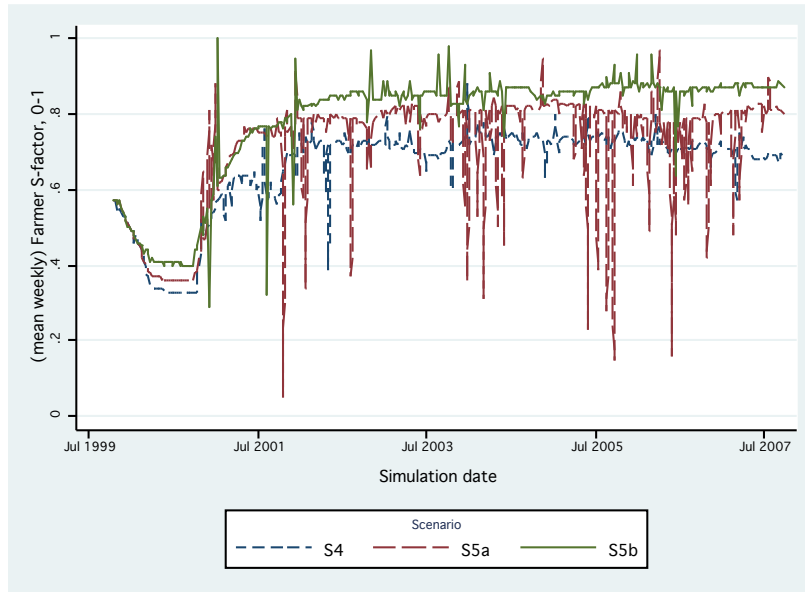
Figure 6.58: Weekly mean farmer S-factor (0.0-1.0), weekly time series, comparing scenarios 4, 5a and 5b. S-factor ratings for S5a and S5b are consistently higher than those for S4. It appears that the effect of allowing opportunity cost information to reach agents is to promote more sustainable attitudes in the population. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

sustainable attitudes in the population. This fits well with the formulation of Hypothesis 5: the empirical parameterization used in S5a and S5b starts off slightly more predisposed to sustainability (mean of around 0.57), and the result is that the population moves towards a more sustainable end of the spectrum over time. The ditch manager results show a similar, if slightly less clear trend. Figure 6.59 shows how 5a shows a slight increase in SF over time, while 5b shows a much larger rise in SF over time, relative to scenario 4. The dramatic temporal variance seen with the farmer SF values is not seen with the ditch managers.

The S-factor data for both farmers and ditch managers suggest that the OCR feedback agents received was more of a complement than a counterweight to the sustainability rating information. It appears in many cases — ditch lining, for example — that the addition of opportunity cost information rendered sustainability-benefitting actions more attractive than they would have been without that information. This is an entirely unexpected result, given the original assumption that the OCR information provided to agents was going to be driving agents towards self-interested profit-maximization. It is possible that the opportunity cost for many sustainable actions was lower than expected, relative to more self-interested options, and
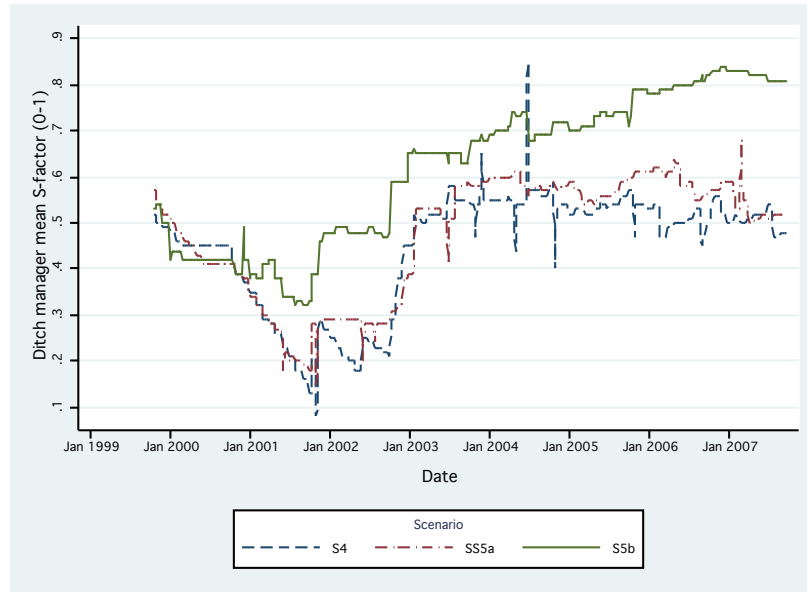
Figure 6.59: Weekly mean ditch manager S-factor (0.0-1.0), weekly time series, comparing scenarios 4, 5a and 5b. Note reduced variance in the ditch manager attitudinal dynamics, relative to the farmer case. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

so simply served to accentuate the benefits of taking that action.

There is a certain contradiction here: S5a and S5b show that the attitudes of agents become more sustainable with time, suggesting that they are getting good sustainability feedback from their actions, yet we still see higher salt loads. To explain this, we may have to look to results for river flow. This alternate measure of sustainability does not always follow the same trend as salt load. Indeed, figure 5.8 shows that flow in both S5a and S5b is several thousand acre-feet higher on an annual basis. This may help explain the contradiction of agents improving their sustainability attitudes while more salt is loaded into the system. The salt load measure is a volumetric one: higher flows will also generally see higher overall loads of salt, because every extra acre-foot of water results in that much extra salt being added to the system. To test this idea, if S5a and S5b are more sustainable than S4 then we should see a lower instantaneous salt concentration (as opposed to cumulative load) at the Colorado/Kansas state line. This would allow us to say that less salt was loaded into the system per acre foot, but the higher flows resulted in higher overall volumetric loads. Figure 6.60 indicates this is not the case: salt concentrations in S5a and S5b were roughly the same as S4. On the hand, the data do

show that absolute salt loading from the irrigation system did not increase significantly. So the contradiction may be resolved: while data apparently show more sustainable attitudes among agents yet higher salt loads, we see that much of the higher salt load could be due to increased flow volumes.
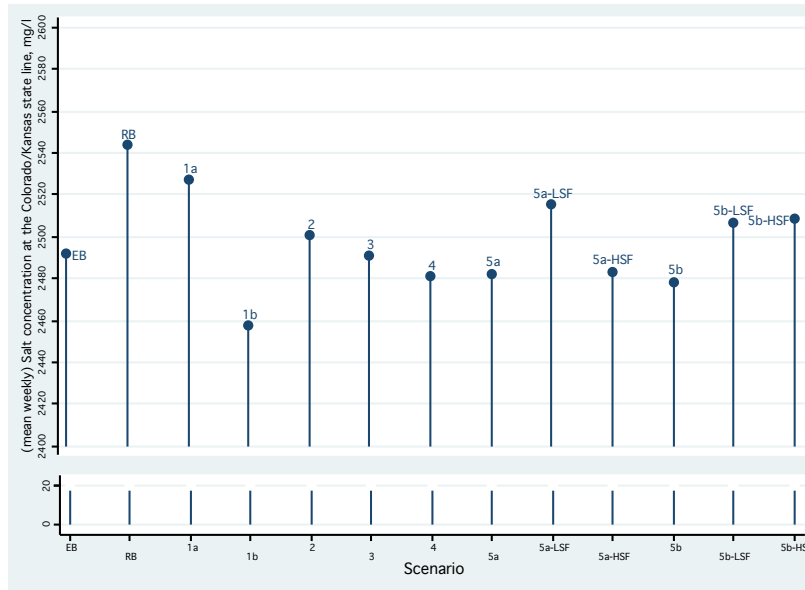


Figure 6.60: Mean weekly salt concentration (mg/L) at the Colorado/Kansas state line, across all scenarios. Salt concentrations in S5a and S5b are roughly similar to S4, leading to the conclusion that higher overall salt loads in S5a and S5b were driven by higher flows. Note scale break: this is used to emphasize inter-scenario variation, since the absolute salt concentration is not of concern here. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

This explanation must remain tentative. Salt concentration at a given point in time is the product of complex physical processes in surface and ground water occurring the length of the basin, and may not always reflect the levels of salt being loaded into the system across some or even all regions (pers. comm. Temeepattanapongsa 2010). I have chosen to make more use of salt load since this reflects a cumulative region across the regions. Salt concentration can reflect salt loads pre-existing in the river without the contribution of basin regions, and the effect of flows unrelated to diversions or return flows from individual regions. The main point to take away from this is that there is no simple relationship between human actions and physical response, particularly when we start to explore subtle differences between actions.

What is clear, however, is that river flow increases in S5a and S5b. What may be behind this?

Recall from previous discussions that the amount of water diverted has strong effects on river flow, and there are several reasons why farmers might change their diversions. Reviewing the diversions data (figure 6.1), there is little to distinguish diversions in 5a and 5b from scenario 4. There are slight reductions in diversions for scenario 5a, which will go some way to explaining the higher flows in that scenario. But in scenario 5b, diversions actually increase. If diversions are not changing substantially, or they increase, what can drive higher flows? One possible explanation is increased levels of runoff due to reduced application efficiencies and lower levels of seepage and recharge reduction. While ArkAgent does not measure runoff directly, figure 6.61 shows that, indeed, we do see reductions in application efficiency across both scenarios 5a and 5b relative to 4.
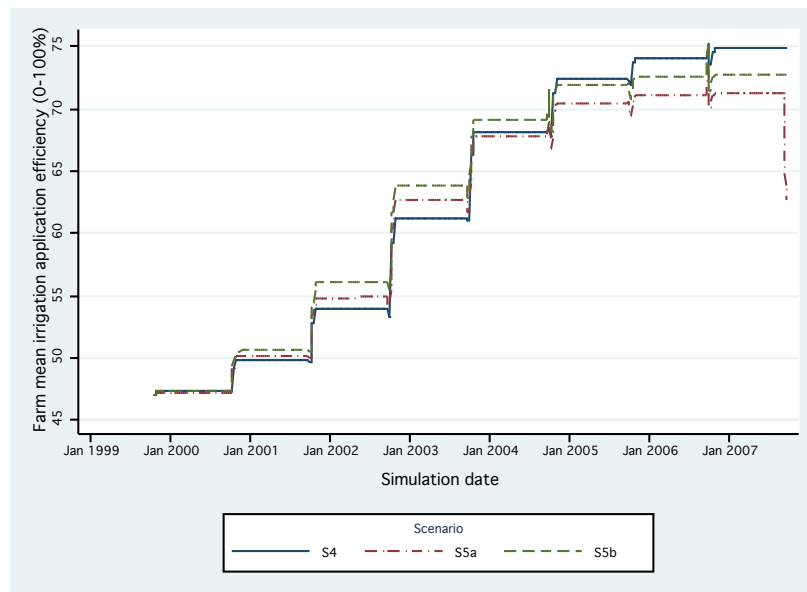


Figure 6.61: Farm irrigation efficiency (0-100%), weekly time series. Application efficiency decreases in scenarios 5a and 5b relative to S4, indicating that opportunity cost information acts to reduce the likelihood of farmers upgrading their irrigation equipment. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

Since we have already seen how 5a and 5b have higher levels of seepage and recharge, and figure 6.24 indicates the amount of water leased to municipalities does not change, we can reasonably infer that the higher flows in the Arkansas River are due to reduced efficiencies in application, coupled with higher recharge and seepage.

What does all this have to say about the overall effect of OCR information on environmental

sustainability? The strongest conclusion so far is that the addition of OCR does not wreck the environmental benefits of the non-classical WQTM and its incentive packages. On the contrary: with regard to most metrics, scenarios 5a and 5b perform well. They still show lower salt loads than most other scenarios tested. But the effects of providing OCR to agents do appear to weaken preferences for certain actions which have proven key to sustainability benefits in previous scenarios, namely irrigation equipment upgrades and ditch sealing. Ironically, OCR information appears to shift agent attitudes towards the more sustainability-oriented end of the spectrum, and this probably comes from occasional alignment between what is economically and environmentally rational for a given set of decision options.

Environmental sustainability is not the only measure of success in these scenarios. We saw from Chapter 5 that incomes in S5a are dramatically increased over S4, or indeed any other scenario barring S5a-LSF. What might be behind this increase? As per previous scenario analysis, the first place to look is the WQTM fine and incentive payment system, and any associated borrowing necessary to facilitate payment of fines. And, as per previous scenarios, figures 6.19, 6.6 and 5.17 indicate that a primary driver of increaed incomes in S5a is higher fine redistribution and associated increases in debts. This, in turn, suggests that some agents especially struggled in S5a to reduce their deficits, but as we see in figure 6.62, this does not appear to be the case. The mean deficit in the system was not much worse in S5a than any other scenario. The next possibility is that quota pricing was significantly higher in S5a, and figure 6.52 confirms this. Figure 6.63 illustrates the variations in quota price with time, highlighting that while scenario 5a follows a similar trend to S4 and S5b for most of the simulation, the scenario hits two peaks in January 2005 and January 2007 which are not replicated in other scenarios.

These two extra peaks are more than enough to drive up overall incomes. As discussed earlier, the mean quota price develops through the price setting decisions of individual agents, and these decisions depend on a subtle balance between the agent's attitudinal orientation, their perceptions of scarcity in the market, and their perceptions of the existing mean price for quota. We already know that perception of scarcity did not greatly vary between scenarios, but sustainability factors did. Because of the way the price setting logic is designed, higher S-factors will typically result in agents setting lower initial asking prices relative to the mean.
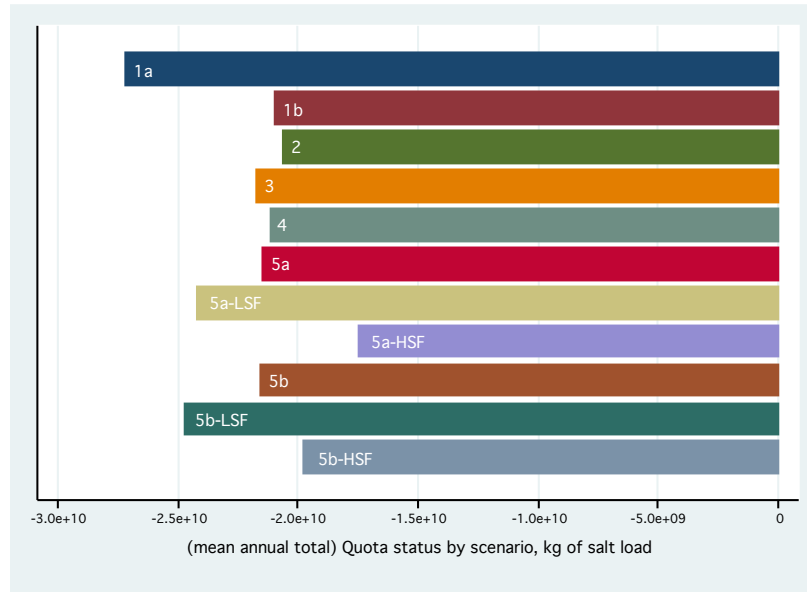
Figure 6.62: Mean annual total quota status by scenario, kg of salt load. "Quota status" is compiled from the various quota trading accounts of individual agents: a negative value indicates that, on average, individual agents were polluting beyond their individual allowances. It may also indicate there was not enough quota in the system to meet demand, but this is not necessarily the case (agents may simply choose not to purchase quota and remain in deficit). Note that overall scarcity of quota in S5a is little different from S4 and S5b, suggesting that agents in S5a were not struggling to meet pollution reduction targets/quota purchasing any more than they were in S4 and S5b. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

Yet what we see is contradictory: the higher S-factor values seen in S5a led to some of the highest quota prices of any scenario. I will revisit this apparent contradiction in the anomalies discussion (section 6.6.3) below.

Finally, note from Chapter 5 that the institutional performance worsens in S5a and S5b relative to S4. Clearing success ratio (CSR) and trade volumes are down in both cases, although numbers of unique traders are apparently unaffected. What might be causing reduced performance in both cases? From previous scenarios, we know that clearing success ratios and trade volumes are adversely affected by lower levels of economic rationality (or higher S-factors). On this basis alone, we would expect to see worse CSRs and trade volumes in scenarios 5a and 5b, since these maintain higher mean S-factors. But undoubtedly the same set of factors that have affected other CSR and trade volume figures in previous scenarios are also at play here, including the quota posting and pricing logic and agent perceptions of scarcity. The poor insti-
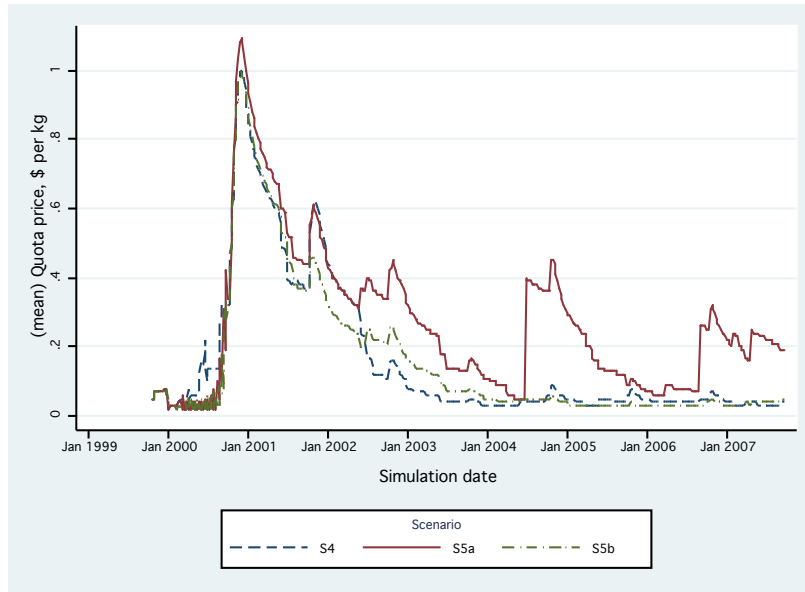
Figure 6.63: Price of 1 kg of quota ($), weekly time series, comparing scenarios 4, 5a and 5b. Plot included to show how S5a is distinguished from S4 and S5b by peaks in January 2005 and January 2007. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

tutional performance in S5a and S5b act to support hypothesis 5 in that they show incorporating OCR information in feedback can help worsen outcomes from a sustainability perspective. But, paradoxically, the poorer institutional performance comes not from any drive towards more economic rationality, but from the opposite: more orientation towards sustainability and less consideration for economic drivers.

Hypothesis 5(ii) puts forward the theory that overall outcomes would be sensitive to initial parameterization. The initial analysis in Chapter 5 showed that at least on the basis of salt load, this was true. The low S-factor (LSF) parameterizations came out with higher overall loads than the high S-factor (HSF) parameterizations. Looking at the seepage and recharge data (figures 6.2 and 6.3), we see that both recharge and seepage are reduced more in the HSF runs than the LSF runs, and that more of the reduction in overall salt load comes from seepage reduction. As to where the seepage reduction comes from, we can see on figures 6.56 and 6.57 that both lining and sealing contribute roughly equally (and both increase by roughly equal amounts from LSF to HSF runs).

On the basis of river flow, however, the results are mixed: higher flows in S5a-HSF relative

to S5a-LSF, but lower flows in S5b-HSF relative to S5b-LSF. The diversions data (figure 6.1) indicate that the amount of water pulled off at the canal headgate does change in 5b-HSF relative to 5b-LSF, explaining the reduced flows. We expect agents with lower S-factors to draw more water from the canal, so why are the more sustainability-disposed agents in 5b-HSF diverting more water? To answer this, we need to first see whether, in fact, the S-factor values in 5b-HSF remained at high levels throughout the simulation. Reviewing the time series for sustainability factors under 5b-LSF and 5b-HSF (figures 6.64 and 6.65), we see some interesting features.
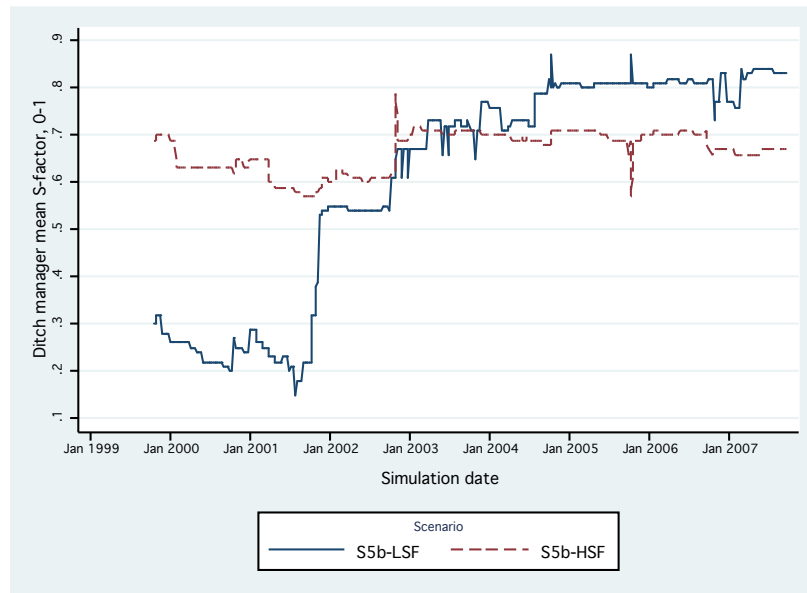


Figure 6.64: Mean weekly ditch manager S-factor (0.0-1.0), weekly time series, comparing scenarios 5b-LSF and 5b-HSF. Contra hypothesis 5(ii) predictions, the LSF run ends with agents having a more sustainability-oriented set of attitudes than than the HSF run. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

In the ditch manager plot, the LSF run ends with agents having a more sustainability-oriented set of attitudes than the HSF run, even though the LSF agents started off with mostly pro-self orientations. Conversely, the farmer response is much closer to the prediction of hypothesis 5: the HSF agents start more oriented towards sustainability and end in much the same way, barring occasional fluctuations. We do see, though, that the LSF agents narrow the gap considerably, and end up with S-factors very close to the HSF agents. This data helps us explain the counterintuitive results for flow: the LSF agents in S5b-LSF actually end more attitudinally oriented to sustainable behavior by the end of the simulation than the HSF agents.
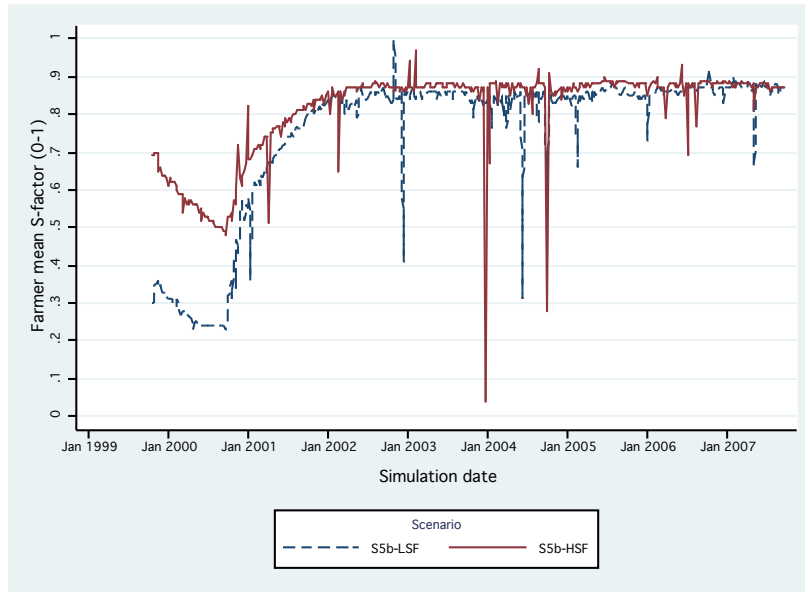
Figure 6.65: Mean weekly farmer S-factor (0.0-1.0), weekly time series, comparing scenarios 5b-LSF and 5b-HSF. In support of hypothesis 5(ii) predictions, the LSF and HSF runs end much as they start, barring occasional fluctuations. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

If we review the time series for S5a-LSF and S5a-HSF, we see broadly similar trends displayed. While the farmer data shows significantly more short term fluctuation, the overall trend is the same: an initial decrease, followed by a steady climb to a peak of around 0.8. The ditch manager data also shows broadly the same trend as the S5b runs, although importantly we do not see the LSF agents exceeding HSF agents in sustainability attitudes over time. This helps cement the explanation offered for flow in the previous paragraph, since we do not see any puzzling reduction in flow from LSF to HSF for the S5a runs.

Looking beyond measures of environmental outcome and attitudinal change, we see much more straightforward improvement in performance from LSF to HSF for institutional measures, supporting the second assertion in hypothesis 5. We can also now explain the otherwise anomalous result that clearing success ratios improve from S5b-LSF to S5b-HSF. Previous results have indicated a lower mean S-factor is a prerequisite for better (more economically rational) WQTM performance. As the time series for S5b-LSF and S5b-HSF show, the S5b-HSF ends with a more economically rational set of agents, running counter to what we might expect for this scenario. On the economic front, it is worth looking harder at the very significant income

drop from S5a-LSF to S5a-HSF, since the magnitude of this shift alone (around $40 million) would merit attention. In the analysis of scenario S5a, I suggested that it was higher quota pricing that drove higher incentive payouts to traders and pushed both borrowing and overall incomes up accordingly. Comparing quota price with time in S5a and S5a-LSF (figure 6.66), S5a-LSF follows a similar pattern — albeit with lower magnitude — until January 2006, when the quota price spikes dramatically. These two features appear to contribute most of the overall income boost we see in 5a-LSF, since a reduced quota price in 5a-HSF reduces fines individual agents are paying, and overall levels of debt accumulation (e.g. see figure 6.6).
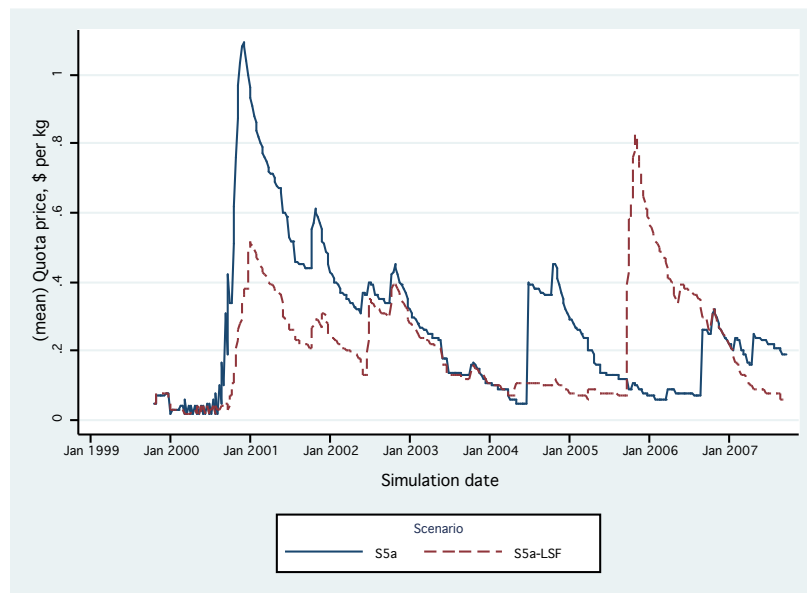


Figure 6.66: Price of 1 kg of quota ($), weekly time series, comparing scenarios 5a and 5a-LSF. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

Comparing all three traces (figure 6.67), we begin to see the effect that the LSF/HSF parameterizations have on the system. The HSF parameterization appears to depress WQTM price setting to the extent that the order of magnitude spike in September-January 2001 is barely seen. At the other extreme, the LSF parameterization shows some evidence of the spike. The overall behavior of quota pricing appears to be decidedly skewed: instead of a monotonic increase in quota price as the mean S-factor value in the population decreases, we see a peak around the mean value. The reason for this behavior is unknown, but will be revisited in the anomalies section below.
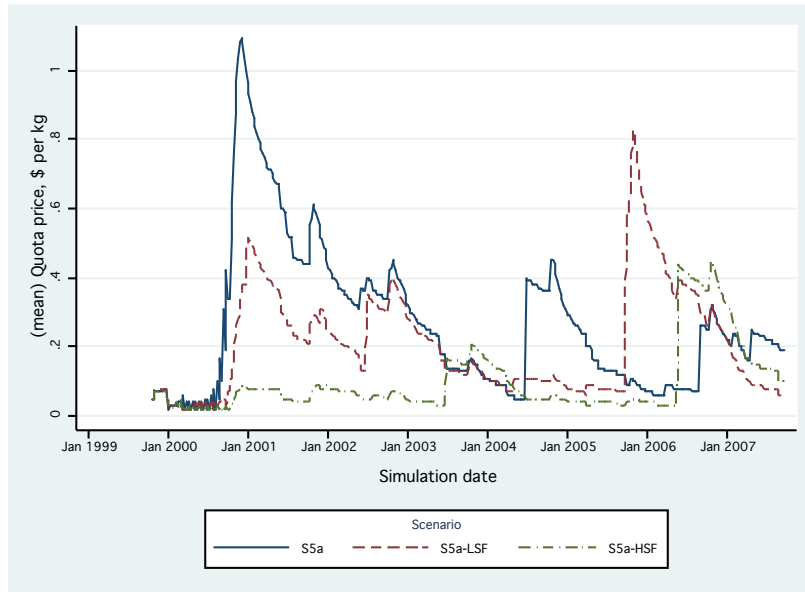
Figure 6.67: Price of 1 kg of quota ($), weekly time series. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

The most intriguing aspect of this analysis has been the S-factor results, displayed again in figures 6.68, 6.69, 6.70 and 6.71.
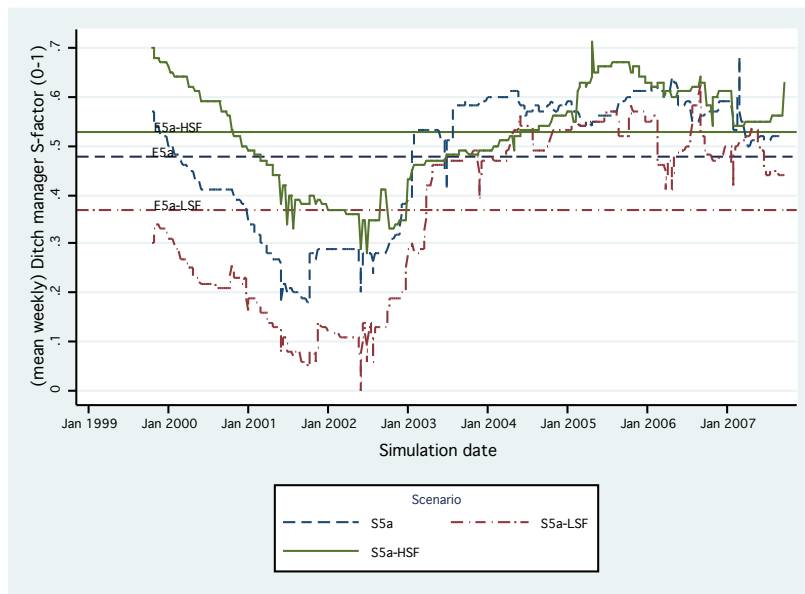


Figure 6.68: Mean weekly ditch manager S-factor (0.0-1.0), weekly time series, comparing scenarios 5a, 5a-LSF and 5a-HSF. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

Beginning with the ditch manager results (figures 6.68 for S5a and 6.70 for S5b), there are
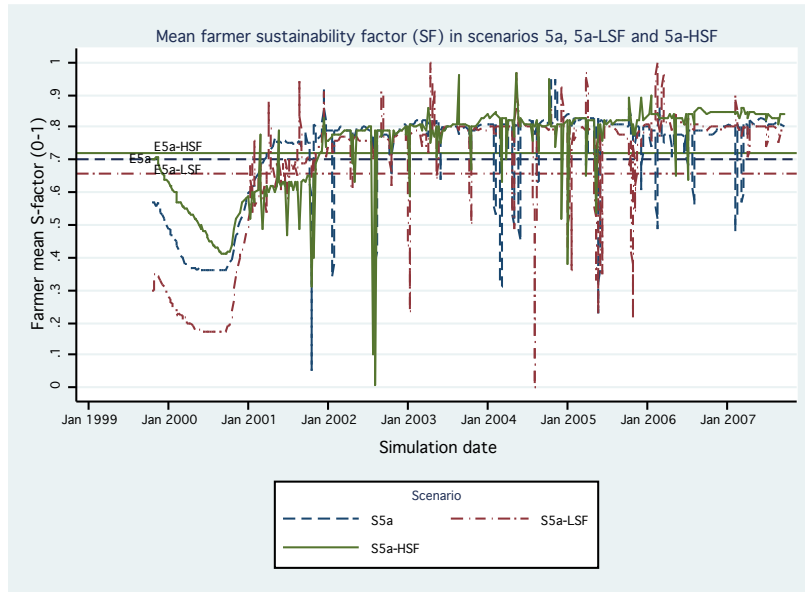
Figure 6.69: Mean weekly farmer S-factor (0.0-1.0), weekly time series, comparing scenarios 5a, 5a-LSF and 5a-HSF. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.



Figure 6.70: Mean weekly ditch manager S-factor (0.0-1.0), weekly time series, comparing scenarios 5b, 5b-LSF and 5b-HSF. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

two particularly interesting features to highlight. First, the fact that the evolution of ditch manager S-factor values over time is broadly supportive of hypothesis 5, since the initial parameterization appears to control the "mean" attitude evolving in the system. LSF parameterized

276
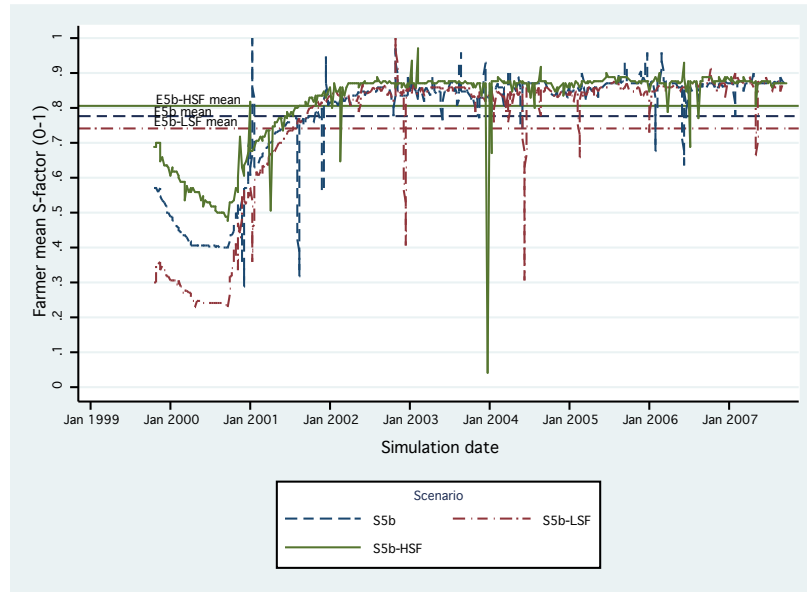
Figure 6.71: Mean weekly farmer S-factor (0.0-1.0), weekly time series, comparing scenarios 5b, 5b-LSF and 5b-HSF. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

agents start with the most pro-self attitudes, and maintain that mean over time; conversely, HSF parameterized agents start with the most pro-sustainability attitudes, and also maintain that mean over time. Second, the fact that using just the "mean" attitude to characterize the results misses out on some unexpected results. Specifically, while the different parameterizations may start out significantly different (means of 0.7, 0.57 and 0.3 for empirical, LSF and HSF parameterizations respectively), they do not end that way (means of 0.6, 0.5 and 0.45 respectively). Turning to the farmer results, we can highlight the exact same two features: "mean" attitudes controlled by parameterization, but very little daylight between parameterizations at the end of the run. In both DCM and farmer data, it appears that the main reason the mean attitudes are distinguishable is because they start so far apart and the simulation takes a while to bring all the means together. From time series analysis, it looks like the second assertion holds very weakly, and is likely not to hold at all if the simulation period were to be extended for a few additional years. The broader and more important point is that the effects of a given incentive *appear to be robust to assumptions on the skew in my distributions of attitudinal orientation* among the agent population. The likely reason for this is that the particular non-economic and economic incentives provided in S5a and S5b, combined with the set of decisions available to each agent

277

type, act as more powerful controls on agent decision making and attitudinal change than the initial attitudinal parameterization. Recall how in figure 6.58, the S-factor curves were distinctly different based on the different scenario parameterization, and then note in figure 6.71 how the curves end up in a similar place: clearly, the scenario parameterization supercedes the effect of the parameterization.

Something that has received only peripheral treatment so far is the effect of social competitiveness and conformism dynamics on addition of opportunity cost information to the feedback package agents receive. The reason why S5a and S5b were both simulated with LSF and HSF parameterizations was to see to what extent findings were impacted by the addition of these social dynamics (scenario 5b). While it is not always to tease out explicable differences, the clearest result appears to be that the addition of social competitiveness and conformism dynamics dampens the system response. For example, salt load reductions from LSF to HSF are less in S5b than in 5a, and the increase in flow seen between S5a-LSF and S5a-HSF is much more muted between s5b-LSF and S5b-HSF. It seems that the social networks farmers and ditch managers are embedded within acts as a brake on the rate of change in whichever direction that is occurring. This makes intuitive sense, since the influence of neighbors is really nothing more than a process of repeated averaging: in receiving information on neighbors, an agent allows its own S-factor to be shifted towards the collective mean. This process will penalize any dramatic shifts away from the mean by individual agents, but also allows change to propagate to more individuals.

### 6.6.3 Hypothesis 5: anomalies

The uneven causal relationship between agent S-factor and quota pricing is an important anomaly that needs attention. The ArkAgent logic assigns a relatively significant role to the agent's S-factor when the agent sets its initial asking price for quota it is posting to the market. If the agent perceives market scarcity, the agent will increase its asking price above the mean market price by a certain fraction of the existing price. The fractional increase is determined by the inverse of the agent's S-factor. The theoretical result is that, when scarcity is perceived, a higher agent S-factor will result in a lower increase to the asking price above the market mean. Despite

this, we typically see the *opposite* effect. Take the peak in quota price around September-January 2003 as an example. In order of increasing mean sustainability factor, the scenarios are ranked 3, 4 and 5a. Yet in the pricing curves at the same time, the scenarios ranked in order of increasing mean price are 3, 4 and 5a. The anomaly, then, is that the relationship should be reversed. As discussed earlier, scarcity appears to play little in the complex temporal dynamics of quota price: mean levels of scarcity across all scenarios vary very little.

One potential explanation for this anomaly is in the way the prices on quota offers already on the market are changed with time. In ArkAgent, if a quota offer sits on the market for more than a few weeks, agents are allowed to drop their asking price. The design of this mechanism is such that the higher an agent's sustainability factor, the slower the agent drops the price. The reasoning behind this design is that lower S-factor agents are more economically savvy, and so according to economic theory will react faster to market signals (e.g. their quota is overpriced). Consequently, one contributory reason to the fact that we see the higher pricing peaks in the scenarios with higher mean agent S-factor is that agents in the higher S-factor scenarios are slower to drop their prices. For example, following the data for scenario 3 (a low S-factor scenario) in figure 6.72, the mean price drops at a much faster rate in the post-peak phase than other scenarios with higher SF values. Unfortunately this is not always the case, so at least part of the pricing dynamic remains unexplained.

### 6.6.4  Hypothesis 5: real world implications

These results continue to suggest that non-traditional incentive packages in market institutions may have the potential to dramatically improve sustainability outcomes across a wide range of social conditions in the LAB, and imply further that this theory is robust to more realistic conditions like the existence of competing informational influences on agents. This implication is likely to have broad relevance in any basin setting where attitudes and strategies are variable within a stakeholder population, and where stakeholders have access to or are bombarded with competing forms of information. Indeed, there are few basin settings in the American West with monolithic stakeholder groups, and the variety of lobbies in water management settings (e.g. recreational, aesthetic, ecological, agricultural) suggest that most real world stakeholders

279
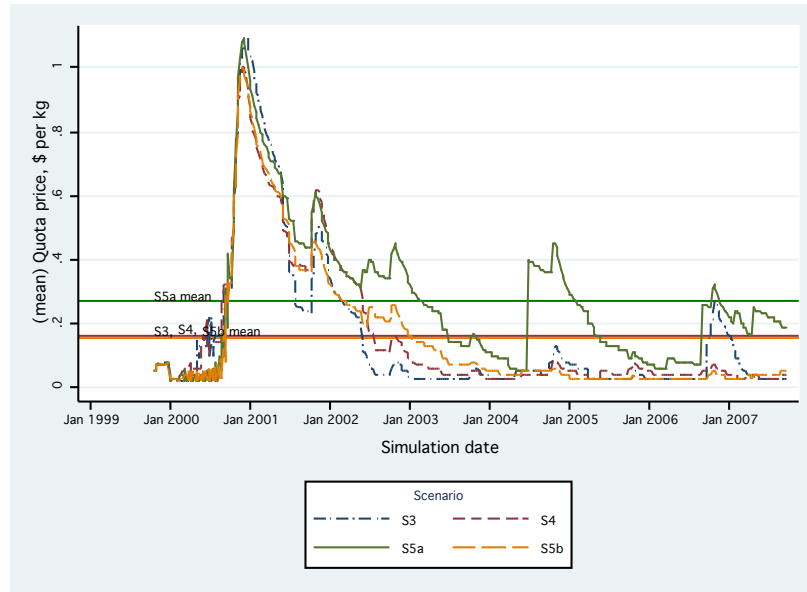
Figure 6.72: Price of 1 kg of quota ($), weekly time series. Plot shows some evidence that the scenarios with more economically rational agents (lower S-factors) show quota price declines faster than scenarios with less economically rational agents. Note that this is not reliably the case, however. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.

experience competing influences.

Hypothesis 5 also asserted that the effect of competing informational influences on WQTM outcomes would depend on the initial set of attitudes in the stakeholder population. The data shows that this is mostly the case over the simulation time period. This suggests that the use of post-hoc informational feedback may be particularly effective in an LAB population already tending towards more sustainable attitudes and more sustainable behavior, and that this tendency need not be significant. However, it is likely that extending the simulation period would show that initial attitudes do not ultimately control sustainability outcomes. The central real world implication of this is that it may not matter so much in the long term if the LAB stakeholder population starts out with attitudes resolutely not supporting more sustainable behavior. The Hypothesis 5 results show that even if this is the case, the use of post-hoc informational feedback may still be warranted if there is a reasonable likelihood of some sustainable behaviors also having economic benefits. Stakeholders in many irrigated agricultural systems in the American West have historically been politically conservative and shown resistance to concepts associated with liberal politics like "sustainability" (Knoke, 1977). However, if we

assume that (a) stakeholder attitudes and strategies are not immutable and can shift with the right mix of incentives and disincentives, and (b) that since sustainability includes economic considerations, there will always be elements of a sustainable water policy that improves economic conditions, the implication that initial stakeholder attitudes do not matter as much as how they evolve over time may hold far more widely than the LAB.

## 6.7 General Discussion

ArkAgent was a large and complex model with very rich output. There is not room to explore in depth many components of the output that are both interesting and relevant to the hypotheses, but in the following section I summarize some of the most interesting results.

### 6.7.1 Hypothesis-relevant findings: sustainability scoring, crop acreages, soil moisture, farming practices and macroscale controls on salt load

*Sustainability scoring*: general results on sustainability scoring support the overall picture that changing **our behavioral assumptions can cause shifts in expected outcomes for market institutions**, and also that **adding non-traditional incentives can increase environmental and economic outcomes without necessarily greatly improving the level of economic performance of the market institution**. Non-traditional WQTMs showed better economic and environmental sustainability benefits, but worse institutional benefits, relative to traditional WQTMs: the non-traditional WQTMs with non-economic incentive packages tended to drive agents to less economically rational behavior, which has generally positive results for more sustainable choices, but sometimes undermines the efficient operation of the market institution. This shows we do not necessarily need a particularly efficient market to achieve sustainability goals, if we couple the implicit financial incentives of the market with explicit sustainability incentives. The best performing institutional form was the basic non-traditional WQTM, and adding further behavioral and institutional complexity degraded performance. This emphasizes how **sustainability benefits across and even within categories are not always acting in parallel**. Benefits to salt load can reduce benefits to river flow; improvements to institutional efficiency may come

281
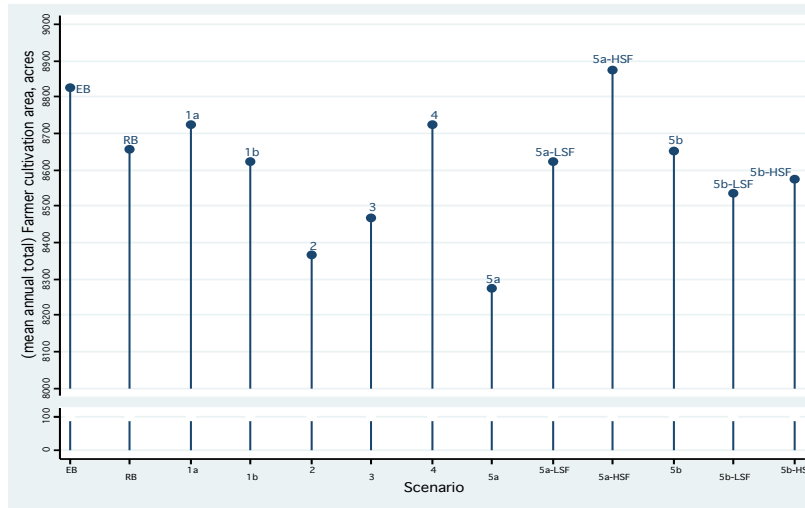
Figure 6.73: Mean annual totals for farmer cultivation of land, comparison across all scenarios. Overall levels of variation are low (range of approximately 500 acres), supporting the idea that sustainability benefits can be achieved without significant shifts in acreage. Note the scale break: this is included to allow emphasis of the differences between scenarios.

at the cost of environmental gain. The implication is that it is inadvisable to dismiss efforts to improve environmental sustainability in the LAB as an automatic loss for the economy, and equally so to equate high levels of market efficiency with better economic outcomes.

*Crop acreages*: crop acreage results indicate that the addition of any kind of WQTM institution drives down crop acreages relative to the baseline. With the addition of more sophisticated non-economic incentives, there are subtle shifts into more water intensive crops on smaller acreages. Results show that significant sustainability benefits can be obtained by relatively small reductions in overall acreage (see figure 6.73) and small shifts in crop mix: alfalfa continues to dominate all scenarios, as expected (see Appendix figure 8.10), and most institutionally-driven gains happen through overall acreage reductions and a shift to more water intensive crops (see Appendix figure 8.11).

These shifts are likely driven by improvements in application efficiency as farmers upgrade their irrigation equipment in response to WQTM incentives (see figure 6.7), and also pressure to achieve higher per acre incomes from a smaller land base. A focus on water intensive crops might sound counterintuitive for sustainability, but if these water intensive crops provide higher returns on far smaller acreages, there is net benefit to both environmental and economic sustainability. The key is to ensure that total acreages of crops with moderate water needs and
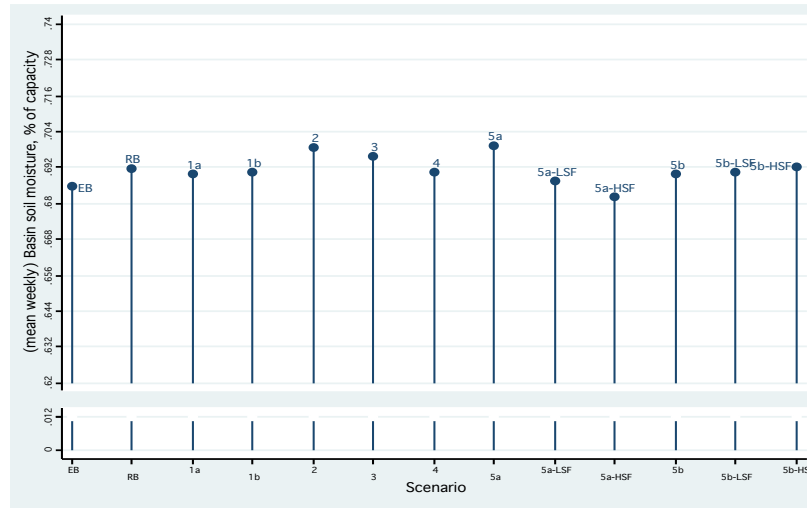
Figure 6.74: Mean weekly basin soil moisture, comparison across all scenarios. No scenario greatly impacted soil moisture levels relative to the baseline. Note the scale break: while mean levels of moisture show minimal variation in absolute terms across scenarios, there are subtle variations (1-2%) which are easier to see with a focused scale.

moderate returns are reduced.

*Soil moisture*: results for mean soil moisture levels across the basin (figure 6.31) show how all empirically-parameterized scenarios showed very slightly higher levels of soil moisture than the baseline, but the differences are not substantial (the differences are clearer to see in figure 6.74). At best, the soil moisture data show that none of the interventions greatly impacts soil moisture levels.

*Farming practices*: results for irrigation equipment choices show a **clear trend towards sprinkler irrigation and away from traditional flood techniques across all scenarios and forms of institutional intervention** (see Appendix figures 8.12 and 8.13). The greatest differences are seen comparing the baseline cases to *any* form of institutional intervention: in the baseline cases, flood irrigation dominates; in any of the other institutional intervention scenarios, non-flood irrigation dominates. This indicates that even the neoclassical institutional interventions may be capable of driving broad changes in irrigation technology. Tweaking the specific incentives associated with an institution can achieve more subtle but nevertheless important shifts in technology use: the WQTM with basic post-hoc informational feedbacks drive shifts to sprinkler irrigation; exploiting social networks in incentive delivery drives further upgrades to drip irrigation. It is reassuring that the irrigation upgrade curves approximate theoretical curves

associated with innovation diffusion theory, specifically the innovator/early adopter/majority phases (Meade & Islam, 2006), although achievement of the majority phase is more rapid than previous studies suggest is realistic (Colman, 1968; Diederen et al., 2003). This raises two important caveats: first, ArkAgent does not simulate many of real world obstacles to equipment upgrades, like variability in capital cost, availability and repair needs for various kinds of irrigation technology (Specht, 2008; Upendram & Peterson, 2007). Because of this, the ArkAgent results should not be taken as any indicator of what rates of uptake we might see in the real LAB. Second, ArkAgent's limited simulation period, non-equilibrium nature and simple behavioral assumptions means that we cannot make any assertions on whether these efficiency gains will be maintained, exceeded or lost over time.

Irrigation technology upgrades were supported by a federal grant program included in the WQTM simulation, and results generally indicate that **the grant program was most effective in supporting efficiency upgrades when the WQTM did not provide sufficient financial support** (positive correlations of 0.4 between higher agent incomes and lower grant uptake; see Appendix figure 8.16, comparing with figure 5.12). The relation is complex, however, suggesting fertile areas for further work.

Results for farmer BMP ("Best Management Practices"; a broad family of changes to traditional farming practice designed to reduce the salt load from irrigation Qadir & Oster (2004)) show — as expected — how the **most economically rational farmers are better at taking up on opportunities for learning more about salinity BMPs** (see Appendix figure 8.14), and that **all institutional interventions result in higher BMP skill levels relative to the baseline** (see Appendix figure 8.15) but that informational incentives are more effective than plain economic incentives at driving up skill levels. Results also indicate that skill level improvements are less consistent for more behaviorally heterogeneous populations, but that non-economic incentives can mitigate this. Given that ArkAgent makes gross simplifications in representing BMP, I hesitate to draw any truly novel implications. However, results do imply that LAB institutional designers should explore the potential for incentivizing change to basic farming practices as well as improvements to technology. The LAB infrastructure for communicating salinity reduction practices is already in place, with an extensive network of NRCS offices and specialists.

Encouraging more irrigators to make use of these resources could lead to significant benefits at minimal cost.

*Macroscale controls on salt load*: general similarities in the shape of salt load trends with time across all scenarios are likely derived from macroscale control of absolute salt load volumes by river flows (compare figures 5.39 and 5.42). Macro-control of salt load by inflow volume does not invalidate any assertions on the effects of various WQTM scenarios, because all the scenarios receive the same base flow into the system, and so any inter-scenario variation will be due to changes in the diversion and salt management practices of agents. However, this macro-control does impose limits on what a given institution can achieve in terms of salt load reduction. Large hydrologic systems are complex, with lags and macro-cycles of geochemistry and climate that WQTM institutions may be able to do very little about. The WQTMs simulated in ArkAgent took place during the tail end of a drought period (see Appendix figure 8.17), and were fortunate in the sense that this made it "easier" to reach salt reduction goals. Wetter periods, of course, would make it more difficult to see the effects of stakeholder behavioral and technological change; extremely wet periods would drown out anything an institution of this kind could hope to achieve.

### 6.7.2   Explaining anomalies: social measures, unique traders and quota gluts

*Social Measures*: while the lack of variability in the two social measures of sustainability — conflict and deficit equity — meant they were very little use in distinguishing scenarios, **both conflict and deficit equity do have benefits in indicating a low overall level of social disruption caused by the WQTM**. Both conflict and deficit equity may be most useful in telling us when within any given year the highest risk of social problems might be seen, since both measures show strong seasonality (figures 5.52). The lack of inter-scenario variation in conflict levels is in some measure a mathematically-predictable result given the composite nature of the conflict variable, but also an indication that none of the institutional interventions caused any serious damage to the social system. With regard to deficit equity, the lack of variation among scenarios supports the assertion that none of the institutional interventions worsened distributional equity within the system. Both conflict and deficit equity data show strong seasonal

trends. Conflict peaks in the early summer and again in late summer, while deficit inequity is highest in the winter months. This implies that tracking social outcomes on a seasonal basis may be even more important than an annual or inter-annual scale, and institutional design may need to play closer attention to temporal sensitivities of the social system.

The issue of lack of inter-scenario variation in social measures raises the **broader challenge of finding the right performance metrics for the institution**. While there is no "right" set of metrics, further work with stakeholders squared with mainstream ideas from the literature may improve the salience and credibility of the ArkAgent metrics. This would be challenging, since a diverse range of indicators for sustainability are in use depending on the field (e.g. ecological economics versus planning: Rees (1996) and Briassoulis (2001)). Different sustainability measures would be unlikely to alter my fundamental conclusions, because in drawing my broad conclusions I have paid closest attention to the underlying data.

*Unique traders*: the lack of variation in model results for number of unique traders each year, and relatively high levels of participation overall (see figure 5.9) both suggest this model area may need technical improvement. However, it also provides some useful information on annual patterns of participation: virtually all traders jump into the market as soon as the State Engineer finishes determining quota allocations in September each year, imposing strong periodicity on WQTM operations (see figure 5.46). Departures from this pattern (e.g. the first year of trading) reflect changing balances between supply and demand.

*Quota gluts*: the gluts in quota seen on the market after the first year of trading occur across all scenarios, and since these are rarely seen in real institutions, the cause likely lies in an **insufficiently sophisticated depiction of the decision logic behind posting quota for sale**. Higher transaction costs and less frequent attention to market dynamics by real farmers would likely drive down real quota volumes. This overestimation of quota introduced to the market identifies an area for further improvement to the model, but does not necessarily overestimate overall trading activity. This is because the actual level of quota traded is controlled by buyer behavior, and there is no indication from the results that this logic was particularly off the mark.

### 6.7.3 Discussion of modeling methods

Modeling data are not the only sources of material for discussion. I developed ArkAgent by applying a mix of methods, including individual interviewing, collaborative modeling, Delphi iterations and psychometric surveying. While I did not intend the application of these methods to be a study of their utility or validity, the experience of their application may have import for collaborative modeling in general. In the following section, I first review the overall suitability of the methods to the core objective of the study (development of a complex systems model and its application in testing of theoretical solutions to a practical problem). I find that the methods were broadly effective in facilitating the construction of a complex systems model, although none of the methods were implemented without problems. I suggest some specific fixes for each method in isolation. Many of the problems with the methods stemmed from poor participation rates. I argue that this stems largely from the fact that the ArkAgent process was not sufficiently well resourced or indeed intended to allow engagement with large numbers of stakeholders, and because ArkAgent was not embedded within an overarching political framework addressing the salinity issue. I suggest the potential that ArkAgent might have if such a framework were to be implemented in future: results from interviewing and anecdotal responses to the workshops suggest that the basin may be ready for a collaborative modeling initiative more closely tied to a decision making process. In Appendix section 8.4.1, I take this line of reasoning further to argue that complex systems modeling in the absence of a decision making framework is common and underserved by available theory and practice. Such modeling can still have significant benefits, and so we need to focus more attention on how to improve modeling practice in such contexts. I propose one possible set of steps by which collaborative modeling can be pursued in contexts that require a mix of exploratory decision support and close stakeholder engagement. Supplemental to this discussion is Appendix section 8.2, which explores in more depth the results from the individual interviewing and workshops with regard to how sustainability might be achieved in the LAB. In this section I recount how the basin community shares a narrative of decline in their way of life, and recognizes the environmental and economic dimensions of this decline without necessarily describing these as part of "sustainability". The concern of many stakeholders with finding solutions to an existential crisis

suggests that a broadly defined sustainability initiative may gain traction in the basin. Legal obstacles are serious and potentially insurmountable, however, and so immediate prospects for change may rely on innovative institutional design and on harnessing the considerable ingenuity of the stakeholder community.

### 6.7.4   Benefits of and problems with the modeling methods

The combination of individual interviewing, collaborative modeling workshop, Delphi validation surveying, and psychometric testing was broadly effective in facilitating the construction of a complex systems model.

- The individual interviewing allowed a scene-setting process, guiding the structure of the initial model, the design of the modeling workshops, and informed later analyses. This "scene setting", a part of modeler education, is critical to allow the modeler to become broadly conversant in the social and physical characteristics of a basin setting. Modeling is a subjective and iterative process, whereby our own theories and hypotheses are continually confronted by reality. It is rarely possible to arrive in a real world context as complex as the LAB and immediately begin to test a pre-existing hypothesis (Barreteau, 2003).

- The collaborative modeling workshops provided the core inputs to model structure and content and exposed stakeholders to the model and modeling in a novel manner. It was during these workshops that the most information was gathered on model structure and content, and was the main period during which stakeholders were confronted with a computational portrait of their entire basin. I did not attempt to measure "systems learning" through these workshops, but at least anecdotally there was evidence that stakeholders were engaged with elements of the system they had paid scant attention to prior, and were brought into contact with other stakeholders they would not normally meet. For example, at the second workshop, an eastern basin ditch manager discussed agricultural issues with western basin municipal engineers; there is no other basin forum at which such a discussion could take place.

- The Delphi validation surveying allowed more elicitation of stakeholder feedback in a much more controlled way than possible through workshops, and allowed focus on more detail than is possible in a workshop. The predominant use of diagrams in the survey was an effective way to communicate model complexities without resorting to technical language.

- The psychometric testing allowed treatment of elements of the model difficult to address with other methods, specifically, parameterization of agent behavioral models. The test was designed to elicit the stakeholders' own attitudes, rather than their opinions on appropriate model content.

However, none of the techniques applied was without problems. I review some of these problems below, and suggest potential fixes for future collaborative modeling efforts.

- The interviewing was difficult to structure, and yielded a large volume of information not related to the focus of the model or thesis (not in itself wrong, just inefficient). Relative to the finished product, the initial interviewing ranged very widely in subject, mostly because it is virtually impossible to design a semi-structured interviewing guide which can anticipate the design needs of a complex model within a multi-year project. This suggests that the use of such interviews should be restricted to general scene setting as opposed to generating ideas for specific model structures. Alternatively, it may suggest that the modeling process should be compressed in time as much as possible, with minimal delays between beginning collaborative modeling work after the initial interviewing. This should make it easier to link the interviewing to later model design.

- The workshops had poor attendance. Some of this was due to deliberate design decisions due to logistical limitations and the political context for the work, and did not necessarily lead to a poorer quality model. Nevertheless, if the project were re-run with a more significant decision making role, additional publicity and community engagement efforts would be helpful. A clear general lesson for future collaborative modeling efforts in the American West is to work hard to find a home for the project within the state apparatus. ArkAgent was conducted within the auspices of a U.S. Federal government agency.

While the Bureau of Reclamation is well regarded in the LAB, the Federal government does not have the same role in day-to-day basin dealings that the state agencies — like the Colorado Water Conservation Commission, the State Engineer's office, the Governor's office, and so on — maintain. More intensive work to engage these agencies would improve participatory outcomes.

- The workshops involved disproportionately large effort in preparation relative to their benefits. In particular, constructing the gaming interface emerged to be far more challenging and time consuming than was originally anticipated. A more staged approach could have been useful here. Instead of the three main workshop events planned, perhaps five to ten smaller workshops could have been used to slowly introduce the model development process. This would also have allowed more time to develop components of the model, its gaming interface and other associated workshop activities.

- The Delphi validation was challenging to implement for such a complex model and with largely non-technical audiences (Delphi panels are more traditionally implemented with expert groups). This suggests that this technique should be used sparingly, with attention to good panel and survey design. But it also suggests that more frequent, smaller surveys addressing single model design questions may be more effective than single, infrequent surveys addressing multiple areas of model design.

- The psychometric testing had limited relevance to stakeholders. The questions on environmental and social attitudes, while contextualized as far as possible, were a departure from previous ways in which stakeholders were engaged with the model, and not clearly relevant to stated model objectives. Completion of the survey at that late stage in model development probably relied mostly on the goodwill of participants towards the modeling process. This is difficult to overcome, since psychometric testing is by definition different from other collaborative modeling methods, but remains useful in addressing parameterization needs. This all suggests that more ground work in collaborative fora would have been helpful in preparing the way for the use of psychometric tests at a later date.

One important reality of this study was that ArkAgent was not implemented within a formal decision process, nor was it designed to focus on helping stakeholders to come to specific decisions on the salinity issue. The theoretical objectives of the study along with logistical limitations and political realities dictated this necessary modification to more traditional collaborative modeling practice. As I argue in section 6.9 below, however, this *does not* undermine the basis of the arguments I make in this thesis, nor the execution of my hypothesis testing. The model is well validated for such a complex instrument, particularly given the limited development timetable and resources available. The model proved an effective tool in testing the five hypotheses as laid out. However, the absence of a clear political niche for the project did pose challenges to applying the chosen collaborative modeling methods. Without buy in from major basin actors, other stakeholders throughout the basin found it more difficult to pledge the time to get involved. It was also harder to defend the potential benefits of the study when stakeholders felt that nothing was moving at a higher political level on the salinity issue. Perhaps most importantly, the lack of an overarching political process meant the lack of a "home" for model development, which increased the barrier to easy acceptance of the effort by stakeholders. There is little in the collaborative modeling literature that directly addresses the case of modeling in an institutional vacuum. The ArkAgent experience, while not directly intended to comment on this context, does offer some insights into how modeling can and should be conducted in such settings. There is not space in the present section to discuss these, but I tackle the question in more detail in Appendix section 8.4.1.

Assuming the model development process were re-implemented within a formal decision process, what kind of different outcomes might we expect? All evidence points towards the likelihood that conditions would become much more amenable to a more classical collaborative modeling effort. Evidence from the individual interviewing suggested the persistence of multiple conflicts in the basin, indicating the community could well use a formal process to guide and mediate the development of solutions to problems of water scarcity and water quality degradation. Most stakeholders expressed the need for some basin-wide means of enhancing communication, understanding and fostering collaboration. A commonly heard suggestion was "involve lawyers less, whenever possible". Indeed, when asked directly for feedback on

291

the idea of collaborative modeling, responses were overwhelmingly positive: several stakeholders suggested that the present time was a watershed moment, with people becoming more and more interested in collaborating over the future of the basin. Given the clear appetite for collaboration, the diminished interest in litigation, and the positive reception of the modeling techniques in the interviewing and the workshops, I can only surmise that a collaborative modeling effort linked to a formal decision making process would be viewed favorably and be useful to advancing decision making in the basin.

## 6.8 Model improvements

### 6.8.1 Changes to how the legal and political system is modeled

ArkAgent deliberately excluded sophisticated simulation of the Colorado state engineer, and omitted the Colorado water courts and other water quality- and quantity-related state agencies, for reasons of parsimony, technical ease and theoretical relevance. While sophisticated simulation of these actors was unnecessary to explore the present hypotheses, for ArkAgent to have real relevance to policy makers in the basin it must do a better job. This is particularly true of the operation of the WQTM, which raises all manner of options for who should manage the institution and how. Including the water courts would greatly enlarge the scope of the model, since it would imply increased sophistication in all agent logic; the inclusion of additional agents (water court judges, potentially also lawyers, various agency staff); and a number of additional model dynamics. Further work in this area should be systematic and incremental, ideally conducted along lines of hypothesis testing as per the current study.

### 6.8.2 Changes to agent behavioral logic

**ArkAgent should be enhanced to improve agent trading and pricing logic**. Results from all hypotheses and scenarios, particularly S1 and S2, indicated often volatile market dynamics. While such dynamics are certainly possible in real systems, it is unlikely real trading behavior would reach such extremes, given additional realities of transaction costs and behavioral complexities unexplored in ArkAgent. The simulation of enhanced pricing and trading logics

would also open up another avenue for sensitivity testing, and demonstrates the considerable potential for this kind of model in exploring ranges of potential societal responses to institutional interventions. For example, the model could be adapted to explore transaction costs related to: (1) search and information; (2) bargaining and decision; and (3) monitoring and enforcement (Stavins, 1995). With regard to (1), this could involve building in lags and costs in agent access to the clearing house. With regard to (2), this could involve adding road blocks and delays to the quota trading process, so that agents find harder to negotiate over price and quantity. With regard to (3), this could involve building in defection potential to agent behavior, so that market behavior is no longer guaranteed to be honest and open. This could extend to actually simulating decisions to opt in/opt out of the market, which would make the market simulation far more realistic.

Results from scenario 3 suggest that **ArkAgent should be modified to improve the joint reinforcement/cognitive dissonance model of attitudinal change**. Cognitive neuroscience and computational psychology are continually improving on our understanding of what drives attitudinal change (Sun, 2008). A future version of ArkAgent could address some of these more sophisticated models, since it already has the basic cognitive framework necessary to support them.

Results from scenario 4 suggest that **ArkAgent should be modified to test variations on the core logic used to control how agents react to the sustainability ratings of their neighbors' actions**. The competitiveness/conformism model is a reasonable start but by no means captures all possible nuance of influence. For example, the competitiveness/conformism model does not account for changes in a social network over time, existence of different kinds of social networks where different kinds of information are exchanged, or differential weighting by actors of information received from different nodes in the network (individual interviewing suggests the latter is particularly important). Furthermore, the "competitive" component of Behavioral Mechanism 4 (see 4.2.6.3 in Chapter 4) is at best a weak corollary of real competitive social behavior.

### 6.8.3   Changes to institutional design

**Scenario 1 results indicate that there is much to be gained from modifying ArkAgent to allow tests of alternate explanations for WQTM failure**. While the results do strongly support Hypothesis 1, the true real world relevance of H1 can only be tested if ArkAgent were to allow the systematic variation of the particular dimensions of WQTM design which other researchers have argued are to be blamed for WQTM failure (for example, transaction costs and monitoring/enforcement effectiveness). ArkAgent was not designed to include these, since Hypothesis 1 was not formulated to discount other possible sources of failure. Furthermore, to make the assertions on the value of post-hoc informational incentives even more robust, ArkAgent would need to run in settings where the set of actions varied, and where the potential economic and environmental benefits also varied.

**The GeoDSS should be improved to allow simulation of changes in municipal policy and technology towards water treatment and discharge**. Because of existing limitations in the GeoDSS, ArkAgent did not include the municipalities of Fountain Creek and Colorado Springs in the trading market. This may have passed an excessive burden for pollution reduction onto farmers and ditch managers. With modifications to the GeoDSS so that these municipalities can be included, the WQTM could be run again to see how outcomes would change with the addition of several important polluters.

**ArkAgent should be modified to allow the use of non-economic incentive packages organized around and tailored by stakeholder group**. Scenario 2 showed that ditch managers were particularly capable of causing reductions in salinity, and since there are fewer managers, targeting these actors might be more efficient. ArkAgent currently treats all agents the same, supplying the same kind of informational feedback to any given agent completing a specific action. The real world is far more heterogeneous than even the generous assumptions of ArkAgent, and tailoring incentives even closer to this heterogeneity may be effective. A modified ArkAgent could easily test this hypothesis.

### 6.8.4   Changes to model operation

Results from all scenarios suggest that **ArkAgent should be modified to aggregate data at finer resolutions, particularly the regional level**. Due to technical limitations, ArkAgent did not gather all of its data at a regional level, but results show that there is considerable complexity at this level that would be illuminated by more social and physical data. A better understanding of regional complexity would also guide institutional design decisions for exploiting these variations. This could be done in two ways: first, the model could be substantially reorganized to gather data at a regional level, so as to explore the various ways in which sustainability outcomes are distributed. This will allow the second step, which is to re-design the model to trial institutional mechanisms which explicitly both account for and exploit regional variation. Understanding and exploiting geographic variation stretches may address market inefficiencies, and may also allow targeting of specific actors by geography. Results so far have indicated that one actor type — ditch managers — have more capacity for sustainability change than the average farmer. But the managers of the larger ditches will have greater capacity than managers of the smaller ditches, which argues for spatial targeting of institutional mechanisms at those larger ditches.

Results from scenario 3 suggest that **ArkAgent should be improved to allow more sophisticated treatment of the economic value of water**. ArkAgent was designed to reflect the limited ability of farmers to engage in detailed financial analyses of their various activities. This need not change. However, to ease future analyses, ArkAgent could include improved logic to more accurately parse out the various influences on the value of water for individual agents.

Following on from the above suggestion, results from testing Hypothesis 4 **underline the need to improve the transparency with which ArkAgent exposes exposes monetary decisions**. The current structure of the model aggregates the results of different economic decisions, presenting only the overall budgetary information for a given agent. It also does not allow the aggregation of financial data by different social and geographic groupings, or distinguishing what decisions were taken at what time. As I have touched on before, this is not the result of a haphazard design decision: there are real computational limitations on what data can be gathered. Nevertheless, this reality limits the sophistication of my explanations, particularly

with regard to the complex temporal dynamics of agent income, assets, liabilities, water value and quota pricing.

A critically important future improvement would be **modifying both ArkAgent and the GeoDSS to allow a 10–20 year simulation period** (and perhaps even longer). All scenarios show sensitivity to variation in WQTM program lengths, suggesting there is room for more study in this area. The non-equilibrium nature of the model means that we are limited to the short eight year period of simulation when discussing implications. Allowing a longer simulation period would expand the analytical reach of ArkAgent, and make it more useful to policy makers in the LAB and beyond.

### 6.8.5 Changes to scenario design

Results from all scenarios indicate that **future model runs should test alternate urban growth scenarios**, since this has important impacts on water leasing behavior by municipalities, and this can affect a variety of outcomes in the LAB. ArkAgent uses the historical baseline to drive simulated urban growth, but in the longer term the scenario of continued growth is not a given, particularly at the more local scale. The effects of reduced regional growth on basin sustainability are not easy to predict, and so ArkAgent would clearly have a role in helping develop ideas about what some outcomes for the LAB might be.

### 6.8.6 Feasibility of further work

The proposals listed in the sections above make suggestions in several general categories. While none of the suggestions are impossible, it is important to be realistic about the level of effort required. Changes to agent behavioral logic (section 6.8.2) would require considerable additional empirical work with stakeholders in addition to the normal development time to design, implement and test new features. The existing agent behavioral models are based on empirical work, and it be would inconsistent with the ArkAgent development approach to start to add on purely theoretical behavioral modifications. Modifications that affect social interactions would require more development than just changes to the core agent behavioral logic, and more challenging verification. Changes to institutional design (section 6.8.3) would

require considerable effort in re-architecting the WQTM institution within ArkAgent. These changes would also imply adding experimental control structures that would manipulate the new institutional features. This is particularly true of incentive mechanisms, since the existing incentive structures are complex. Changes to model operation (section 6.8.4) are mostly straightforward to implement, but would incur considerable processing and memory penalties. It is not clear whether the model platform currently used for ArkAgent could currently handle any such improvements. Any modifications required of the GeoDSS would have to be conducted with the full cooperation of Colorado State University, but their exceptional support in making changes necessary for the existing ArkAgent setup suggest that this would not cause any problems. Finally, changes to scenario design (section 6.8.5) would be relatively straightforward to implement, unless empirical work turned up additional complexity that proved necessary to include.

## 6.9 Model limitations

ArkAgent has a number of basic limitations that should be underlined so that any assertions I make, and real world implications I suggest, are received with appropriate caution. For reasons of brevity, details of specific limitations are included in the Appendix. Categories of limitations addressed there include: limitations arising from the unavoidable need to simplify the real world into a computationally tractable model (Appendix section 8.3.1); limitations arising from non-equilibrium modeling (Appendix section 8.3.2); limitations arising from deliberate choices to simulate some things and not others, on theoretical or other bases (Appendix section 8.3.3); limitations arising from logistical constraints, particularly the computational demands of ArkAgent relative to the available machines for simulation (Appendix section 8.3.4); and finally, limitations arising from the coupling with the GeoDSS water quality/quantity simulation tool (Appendix section 8.3.5). In the following section, I discuss the implications of the model limitations.

Does the fact that ArkAgent had myriad technical, logistical, conceptual and theoretical limitations mean the many real world implications of the model are not to be trusted? Lim-

itations imposed by logistical constraints and design considerations are clearly unavoidable, present in all modeling efforts, and do not greatly subvert the implications of Arkagent. The problems ArkAgent addresses fall into the category of "squishy" or "wicked" (Lachapelle et al., 2003): the problems themselves are not clearly defined, and in many respects the LAB system is not particularly well understood. Most issues of natural resources management (NRM) fall into this category, and it is precisely because of this that models are often appropriate tools for testing and formulating NRM policy. In simulation-based studies of wicked problems, whether we trust the model depends on its overall credibility. The meaning of "credibility" is very much dependent on the objectives of the study and the intended use of the model (Balci, 1998). The credibility of complex system models is often established by assessing "usefulness with respect to some purpose" (Barlas, 1996, 184). Since this assessment process will necessarily not be fully objective, it is meaningless to hold every systems model to a perfect standard of objective validity. Indeed, as long as there is no single model for theory confirmation in the natural sciences (ibid), absolute validation is an impossible endeavor. Is ArkAgent useful with respect to its purpose? I defined ArkAgent's purpose at the beginning of the discussion chapter as "testing the implications of theoretical assertions on social and environmental systems within a more complex context than a purely abstract environment". ArkAgent includes complexity that abstract models cannot handle, and it is in fact this complexity which is particularly crucial to the proper testing of the hypotheses. This more limited purpose — ArkAgent was never intended to be a finished decision support tool, at least in this iteration — implies that we must hold ArkAgent to a different standard of validity. But it also means we should not overreach. For example, we cannot say anything about what *will* happen in the LAB should a WQTM be implemented; only what *might* happen if we accept the behavioral assumptions embedded in the model, and even then we cannot discount the potentially confounding future behavior of a non-equilibrium model run outside the current eight year period.

This should not be taken as an excuse for making no effort to improve model validity. In ArkAgent's three year development period, I sought to ensure that the model was as well validated as possible given the technical and logistical limitations. ArkAgent's initial design was based upon intensive stakeholder interviews and design discussions with local and regional

experts. During development, the model was iteratively reviewed by groups of stakeholders and again by local and regional experts, in accordance with accepted practices for face validity testing (Gass, 1983). Preliminary outcomes were reviewed with researchers who have spent over a decade working on the problem of salinity in the LAB. Finally, sensitivity analysis was performed on areas of concern, and the model structure adjusted accordingly. Nevertheless, mindful of the many potential problems with such a complex model, I have always been careful to apply ArkAgent in a retrospective manner.

It is unreasonably cautious to absolutely distrust ArkAgent's output, even if it has not undergone a traditional calibration, verification and validation track. Outcomes from this model do have some bearing on reality, and I believe this is sufficient to support — at the very least — general discussion of potential real world implications.

# Chapter 7

# Conclusions

## 7.1   Answering the thesis questions

This study began with posing five questions addressing different aspects of the relation between institutional design, behavioral variation, and sustainability outcomes in large water systems. I offer concluding answers to each question in the following section.

**Question 1**: *What is/are the potential impact(s) of assuming more heterogeneous behavioral models on the success or failure of water quality trading markets?*

ArkAgent results show that assuming more varied behavioral models can greatly shift the kind of performance we might expect for a given institutional form. If we limit this question to just neoclassical water quality trading markets (WQTMs), the ArkAgent results demonstrate that many of the putative benefits of a market institution fall away. The ideal market institution and the ideal set of agents leads to lower levels of salt, better economic returns and improved overall institutional performance, relative to the ideal market institution with a less than perfect set of agents. Populations of agents with variable behavioral models at their core show that even the ideal market institution cannot get even close to the predictions of the 1st and 2nd Welfare Theorems. This goes beyond simply showing that neoclassical economic theory is not perfect. It shows that such theory is inadequate to describe and predict the behavioral choices of agents with assemblages of different motivations. This is even the case when those assemblages include the rational self-interest that neoclassical economic theory maintains as a

foundational assumption.

Conversely, what the ArkAgent results also show is that simulated market institutions can function relatively well in populations of variable behavioral models. Nowhere near as well as the perfect theoretical case, but reasonable functionality nonetheless. It has never been the intent of this thesis to dogmatically undermine the application of market institutions to natural resources management problems, nor has it been to compare such institutions to non-market regulatory alternatives. Market institutions have many benefits in managing NRM problems, and there is room for both regulatory and market-based solutions. Instead, one major intent has been to show that making erroneous assumptions about stakeholder behavior can lead to considerably poorer institutional performance than neoclassical theories might lead one to expect. The ArkAgent results show clearly that this can be the case. They do not allow us to blame incorrect behavioral assumptions entirely for poor performance of WQTM institutions, since there are many institutional design dimensions that can be incorrectly or inappropriately formulated, even assuming a perfectly economically rational set of stakeholders. But the results do suggest we should be opening up debate beyond simple rational actor and boundedly rational actor theories of stakeholder behavior. Of course, ArkAgent does incorporate important elements of the boundedly rational actor model. It has been my position throughout this study that this remains *one possible* model among a much broader behavioral spectrum than traditional studies of water resources management institutions have allowed. So ArkAgent also incorporates agents with particular preferences for sustainability — defined along multiple axes — in ways that may or may not coincide with their own self-interest. An important theoretical implication of ArkAgent is that we both should and *can* study alternate behavioral models; we need not stick exclusively to the tried and tested territory of the boundedly rational actor.

Since ArkAgent shows clearly we can study more complex assemblages of behavioral models, the ArkAgent development exercise also tells us something about how this should be done. For one, intensive engagement with an extensive array of real people should be at the core of model development. This model shows that relying on neoclassical abstractions sheds little light on actual performance of real systems; conversely, if we are to get models relying on a

variety of new behavioral assumptions to improve our understanding, we must root model development in empirical data gathering. Such work should incorporate a diversity of techniques from the stable of participatory modeling. Models like ArkAgent are extremely sensitive to how the core of agent behavior is constructed, and the relevance and utility of such models requires that this core is built through close work with the stakeholders being simulated.

**Question 2**: *Given a population of resource users with different behavioral models distributed in space, what happens to a WQTM institution designed with incentives that work best with a behaviorally-homogeneous population? In other words, how does the performance of this kind of institution compare with a similar institutional form incorporating more diverse incentives?*

The ArkAgent results show that a "pro-social" market institution — a WQTM incorporating incentives designed to attract agents with other than purely economically self-interested behavioral models — can be more effective at achieving sustainability goals than the basic neoclassical version of the same institution. ArkAgent does not show that the inclusion of more inventive incentives will *always* lead to better market institutional performance. Instead, it shows that if we make the reasonable (and empirically founded) assumption that a population of actors will vary in how they view the world and how they take actions within it, a market institution exploiting this diversity will *always* do better than the basic neoclassical alternative.

It is critically important to remember that this is not a result derived from an highly abstracted conceptual system. ArkAgent incorporates significant structural complexity in both society and environment; variation in behavioral models; fluctuations in physical parameters and responses; and all manner of potential path dependencies. What ArkAgent shows is that this core idea — that incentives should be better matched to the real complexity in stakeholder attitudes and behavioral strategies — is robust even under a myriad of potentially confounding influences. Moreover, sensitivity testing conducted on ArkAgent indicates that this conclusion is robust to varying many key parameters in the design of the WQTM and the nature of the agent behavioral models.

ArkAgent results show that a pro-social market institution is more effective at fostering sustainability because it harnesses the attention and potential salt reducing activities of a greater proportion of the stakeholder population. By providing an alternate form of incentive beyond

the financial, the pro-social market institution reaches farther into individual decision making than a neoclassical institution can hope to achieve. This is, after all, the idea at the heart of market institutions in natural resources management: provide incentives for more sustainable behavior in a way that allows participants to choose their own best solution. Market institutions with more diverse incentives simply enlarge the potential range of decisions that can be affected by the operation of the market.

ArkAgent results suggest that the use of diverse incentives may be even more effective when they are focused to target particular stakeholder groups. The LAB highlights a reality in many such systems, where different stakeholders have very different capacity to affect a given problem. A neoclassical institution assumes that these differences will be overcome by the mechanisms of the market, such that the stakeholder with the lowest cost in salt reduction will simply do so and pass the surplus quota onto other polluters less able to afford reductions. But the LAB case indicates that stakeholders are not usually aware of which is the most effective solution when they start addressing a problem, and many other complicating factors (including past experience, what is traditionally done, the opinions and recommendations of friends, and so on) will intervene to stop a stakeholder from making the most efficient decision. The use of more diverse incentives in ArkAgent suggests that there is plenty of room within the WQTM structure to help the least-cost-pollution-reduction process along by focusing certain incentives on certain agents who might need a gentle push in the right direction. The consideration of such apparently sensible provisions is only possible because ArkAgent dispenses with the restricting assumption that the best and only way to incentivize behavior is through the exploitation of self-interest.

Finally, in a dose of realism, ArkAgent results suggest that the use of diverse incentives does not always lead to predictable results. ArkAgent results showed that while one outcome was to reduce salt load, another was to drive down flows in the river and another was to reduce market efficiency by a considerable margin. Additionally, the combination of unsophisticated trading behavior by less economically rational agents, and the fixing of fines to quota prices, led to considerable increases in agent borrowing over the empirical baseline. All these outcomes show that even an institution with more sophisticated incentive structures cannot

304

hope to counter all the potential obstacles generated by a real world setting, and may generate undesirable side effects. Conversely, the fact that the non-traditional WQTMs still showed sustainability gains despite such obstacles indicates that we do not necessarily need highly efficient markets to achieve more sustainable behavior, and that sometimes unfortunate side effects (like higher borrowing) are necessary to facilitate sustainable change.

**Question 3**: *What are the implications for institutional performance of assuming a population of behaviorally heterogeneous agents, where behavioral strategies are dynamic with time and responsive to institutional incentives?*

ArkAgent results in this area imply that, even in the case where we assume agents are capable of attitudinal change in response to external influences, the market institution with diverse incentives still performs robustly. It does not perform quite as well as the previous case where we did not assume that agents were capable of attitudinal change, but we still see significant sustainability benefits relative to the baseline case (no institutional intervention) or the case where a neoclassical institution was implemented.

The principal reason why attitudinal change appears to worsen sustainability outcomes overall is that the initial effect of informational feedback (the incentive with which the non-traditional WQTMs are equipped) is to drive agents away from the ideal salt reducing attitudes and behaviors. Change-equipped agents actually end up with more sustainable attitudes and taking more sustainable actions, but the dynamics of change mean a period of adjustment where actions might actually be worse with the use of informational incentives. In other words, the implications for institutional performance may depend on the timescale of interest. If a behavioral incentive is needed to drive change in a short period of time, outcomes may be initially counterproductive. But with patience and time, incentives may drive the right kind of attitudinal shifts across a population.

Finally, ArkAgent results show that institutional performance is strongly affected by the the extent to which the path of attitudinal change varies according to the particular domain of an agent: the decisions it takes, how frequently it encounters feedback and how capable the agent is of translating that feedback into positive change. The ideal case is where agents are frequently encountering feedbacks and can easily take actions to remediate any poor feedback

they receive. But as ArkAgent demonstrates, the reality is more often the reverse: agents receiving frequent feedback often do so on actions which are relatively routine with limited room for change; agents receiving feedback more rarely can often theoretically achieve a more salt reduction through fewer actions. While this and the previous point are outcomes entirely derived from specific model assumptions, their broader lesson is that institutional performance is dependent not only on how much change we assume agents are capable of, but how this change is distributed in time and within the population. An apparently straightforward result, but it indicates that when behavioral dynamics are included, a whole host of new institutional design considerations must also be added.

**Question 4**: *What are the implications for institutional performance when we modify an institutional design to include mechanisms to positively reinforce behavioral change?*

ArkAgent results show that the inclusion of incentives in a market institution to exploit social competitiveness and conformism leads to better sustainability outcomes than a comparable institution omitting these incentives. This is despite the fact the competitiveness mechanism used in ArkAgent is relatively weak, and tends to dampen the competitive drive if an agent becomes too far ahead of its neighbors. Sustainability outcomes are better across multiple metrics, most critically in the area of salt load reduction. These improved dynamics result from exploiting the potential for social networks to echo and magnify behavioral trends. The effect appears to be accentuated by the relatively small agent population simulated in ArkAgent (less than 700), and the fact that some critical agents (in terms of potential salt load reduction) exist within even smaller networks of agents similar to themselves.

ArkAgent results indicate that incentives exploiting social competitiveness and conformism will vary in their absolute effectiveness partly depending on which agent types have the highest capacity to affect sustainability outcomes. In the LAB, the ditch managers have the highest individual capacities to affect such outcomes, and with their small social networks the neighborhood rating system appears to work especially well.

Note that sensitivity testing of ArkAgent suggests that the general result is broadly robust to assumptions on how much attention agents pay to their neighbors, and how competitive or conformist an individual agent is. With less competitive and conformist agents, the effects

306

are less magnified, but still present. This particular incentive structure has upper limits to the amount of attitudinal and behavioral change it can foster: beyond a certain level, the mechanism is self-limiting, with agents shifting into a stable set of actions that provide consistent but not improving sustainability benefits.

Finally, model results indicate that exploiting behavioral change to positively reinforce desirable attitudes and behaviors can sometimes backfire: as ArkAgent actors became more oriented towards pro-social and sustainable activities, their concern diminished with being economically rational on every decision, resulting in some undesirable economic outcomes (like increased debt loads). The implication for institutional performance — particularly market institutions — is that if the incentives provided by an institution act to drive agents towards pro-social behaviors, some dimensions of sustainability dependent on pro-self behavior may suffer.

**Question 5**: *What is the effect on institutional performance of assuming that behavioral change is also susceptible to informational messages communicating financial self-interest, as well as informational messages communicating sustainability benefits?*

ArkAgent results show that the existence of competing informational messaging does not significantly reduce institutional performance. This is principally due to the fact that effects of the competing financial information (in the form of opportunity cost feedback) are mitigated by the diverse nature of the attitudinal orientations of individual agents. In the empirical parameterization there as many agents who are significantly affected by this information as completely ignore it.

In perhaps the most surprising result, ArkAgent shows that influences communicating financial benefits as opposed to sustainability benefits can actually drive agents towards more sustainable behavior. This is primarily because the WQTM acts to increase the opportunity costs to not participating in salt reduction efforts, causing increasing coincidence between actions with high costs if they are not taken, and those actions that have high sustainability benefits.

The initial assumptions on the distribution of behavioral models across a population appear to control overall outcomes on the less than decadal timescale. While ArkAgent did not

simulate beyond an 8 year period, the time series suggest that sustainability performance may be evolving to a common steady state regardless of initial parameterization. This is related to the fact that institutional structures appear to control behaviors in the long term more than the initial attitudinal orientation of agents in the system. The only effect the existence of competing financial information appears to have is to bias the time series towards slightly more or less sustainability benefit during its evolution towards the common steady state.

## 7.2  Real world implications

Much of the preceding discussion has talked about implications of the ArkAgent results in relatively abstract and theoretic terms. In the following section, I summarize potential real world implications of these results, both for the specific context of the Lower Arkansas Basin, and for western water settings more generally.

### 7.2.1  General implications for institutional design

The generalizability of the ArkAgent results depends on the extent to which we accept the empirical findings on which much of the ArkAgent theory is grounded: that in most cases, real resource users do not resemble boundedly rational *homo economicus*. I offer the following general implications with the assumption that in most basin settings, this empirical finding holds. Second, I remind the reader that ArkAgent is a non-equilibrium model, and does not provide any guarantee that trends displayed over the simulated eight year period will continue or change over an extended simulation period. ArkAgent is best understood in this context as a simulation of the introduction of an institution, and not its long term evolution.

#### 7.2.1.1  Incentive structures

**In general, institutional designers should consider the potential benefits from using more diverse and inventive incentive structures, particularly explicit informational feedbacks communicating sustainability benefits**. This does not preclude concern with traditional issues like transaction costs and monitoring/enforcement. But results from ArkAgent imply that we

should not resort to simple neoclassical explanations for poor performing markets, when it is clear that different — and empirically grounded — attitudinal and behavioral assumptions can lead to very different market dynamics. ArkAgent results show that incentives need not (and cannot) be as complex as the attitudes and behaviors of real actors, but can still be effective in exploiting these attitudes and behaviors. ArkAgent results show that importing innovative measures from energy reduction initiatives (like post-hoc informational feedbacks and "smart" metering of consumption/pollution) may hold considerable promise in exploiting variable attitudes towards sustainability among stakeholders. ArkAgent results show that the use of such incentives may be robust even to the presence of countervailing information emphasizing more self-interested and less sustainable behavior. ArkAgent does not imply that post-hoc informational incentives will *always* be successful. The most important implication from the ArkAgent work is that such incentives *can* be successful within a complex social-hydrologic system. This argues for a shift in institutional design away from bases in theoretical argument and political conviction, towards simulation-mediated studies of institutional alternatives using more realistic assumptions than allowed by neoclassical economic theory.

**Institutional designers should not stop their use of innovative incentives with informational feedback**. Arkagent's results do not imply that informational feedbacks are the only or best way to harness behavioral heterogeneity within a stakeholder population. If we accept the premise that real stakeholders have complex attitudes and behavioral strategies, possibilities open up for all sorts of incentive targeting. The literature on the spread of technological innovation has a long history of studying populations of innovators to determine when, how and why individuals might take up a particular technological innovation. Beginning with the classical diffusion model (Rogers, 1995), researchers have developed sophisticated and testable theories on innovation diffusion, and this has been translated into real strategies of business and government (Meade & Islam, 2006). We need the same kind of approach to institutional design for water resources sustainability. ArkAgent cannot offer fundamentally novel insights into what additional forms of incentive might be useful — that is a role for policy makers — but the model is an example of one way we can design and test ideas for new incentive forms. There is no reason that the ArkAgent approach cannot be replicated in a variety of different

basin settings.

**Institutional designers should consider designing incentives to encourage improvements in farming and ditch management skills**. Trading markets are not traditionally thought to have any direct role in improving levels of education among individuals. However, many pollution abatement techniques are not necessarily known by stakeholders, and simply providing a monetary incentive (as per the neoclassical trading market) may not be enough to promote uptake. ArkAgent's simulation of farmer skill, as well knowledge of salinity best management practices, raises potential synergies between more traditional educational outreach efforts (e.g. USDA NRCS; Cooperative Extension services) and incentive delivery by the market institution. For example, post-hoc informational feedbacks could be expanded to include not only estimates of what the sustainability benefit or cost of the actor's previous decision was, but also suggest actor-specific ways to address salinity issues. Note, once again, that ArkAgent does not and cannot demonstrate such an approach to be successful in all cases. ArkAgent merely shows how the presence of such dynamics in a complex system structure can affect sustainability outcomes.

**Institutional designers should also take note of geographic complexities in designing and implementing incentives**. Simulated salt loads from different parts of the LAB likely depended on complex combinations of spatial factors. For ditches, these include ditch clustering, ditch lengths, and ditch position relative to the east and west ends of the basin. For farmers, these include the distribution irrigation equipment across their lands, locations of their lands relative to near-surface shale beds, and locations of their lands relative to the ditch and the river. While some of these complexities are unique to the LAB, many are not. It behoves institutional designers to find as many ways as possible to tailor institutional incentives to where they will potentially be most effective geographically.

**Institutions need to tailor their incentive forms and delivery characteristics to the seasonal and sub-seasonal climatic, economic and other cycles of a given irrigated system, to make those incentives more effective**. This tailoring could take the form of the timing, frequency and form of incentive delivery. ArkAgent results demonstrated how, even in the absence of a WQTM, the LAB is a highly seasonal system: everything from the value of water to

the timing of operational decisions is driven by the cycles imparted by climate and the basin's water distribution infrastructure. This is common across many large irrigated systems in semi-arid climatic zones. Consequently, institutional designers should consider the potential for tailoring incentives on the temporal scale. For the most part, this should take the form of more flexibility. Instead of delivering informational feedbacks the same way the whole year round, an institution may be more effective if it delivers those incentives during planning periods in the year, so that it is more likely they affect planning decisions. Alternatively, informational feedbacks delivered more frequently during the height of the irrigation season will almost undoubtedly be more effective than those same feedbacks delivered during the end of the irrigation season, when making efficient use of scarce water may be less of a prime concern for farmers.

**Institutional designers should consider tailoring incentive structures to specific actor types**. In any given basin setting, it is very unlikely that all actors will have the same capacity for reducing salt load. While traditional economic theory holds that it is important that marginal abatement costs vary within a population of polluters, so that the market can function efficiently to concentrate pollution abatement efforts where they cost the least, it may be even more effective to encourage the concentration of pollution reduction effort with the actors who can most effectively reduce pollution. This may or may not coincide with least cost abatement, and requires the conceptualization of the core trading market objective as pollution reduction at all costs rather than pollution reduction at least cost. A market that focuses solely on achieving perfect least cost pollution abatement may miss opportunities to exploit the willingness of individual stakeholders to go beyond what is economically rational. In the LAB, ditch managers have considerable capacity to reduce salt load, and may have a variety of interests in doing so above and beyond just minimizing cost. Institutional designers in this instance could engineer a species of incentive specifically for ditch managers, and achieve greater pollution reduction as a result. The broader implication of this is that it may be inappropriate to apply resource management policies "equitably", in the sense that not all actors should necessary face the same set of institutional mechanisms. When an institution is designed with such broad mechanisms that they can apply to all actors, there is likely to be a loss in effectiveness at an

individual level. ArkAgent results suggest that an institution could be designed with a richer array of actor-specific mechanisms, reducing inefficiencies and perhaps enhancing effectiveness in general.

**Institutional designers should consider the power of social networks in propagating behavioral change**. ArkAgent results demonstrate that incentives designed to exploit even simple social dynamics, with very crude assumptions, can increase sustainability benefits. This is in line with empirical evidence, particularly from the energy consumption literature, that social networks can and should be exploited for this purpose (Mankoff et al., 2007; Seyfang, 2006; Spaargaren, 2003). From chance coffee shop interactions with fellow farmers to sophisticated informal contracts between ditch managers, modern irrigation systems exhibit a great variety of social networks with capacity to propagate (and perhaps resist) institutional interventions. ArkAgent provides the theoretical implications of exploiting an indisputably real feature of hydraulic societies; institutional designers should take the cue to develop and test real instruments for exploiting social networks. This also supports the previous implication on tailoring incentives to specific actor types: social networks can often be highly stratified, and so different actor types may have social networks with variable capacity to enhance incentives. Institutional designers should take note of differences in social networks, if any, and potential implications for incentive design.

### 7.2.1.2  Explaining WQTM failure

**Researchers and analysts exploring reasons for market institutional failure in a given real world context should pay more attention to behavioral dynamics**. For example, while lack of liquidity is commonly seen as a structural problem in many real world WQTMs, ArkAgent simulations show how behavioral diversity, more than any particular structural issue, can be to blame. While ArkAgent does not (and was not intended to) offer guidance on where in particular behavioral diversity more than institutional structure will affect market liquidity, ArkAgent does demonstrate how it is both possible and reasonable for behavioral dynamics to affect market liquidity. Common sense reasoning suggests that institutional designers should begin by trying to understand if this will be the case in their particular setting; then they can, if necessary,

move proactively to avoid liquidity issues with appropriate stakeholder engagement.

High transaction costs are often seen as driving poor market performance (Hoag & Hughes-Popp, 1997; Pharino, 2007). ArkAgent does not explicitly treat transaction costs, and so has a limited contribution to make to the extensive literature on this subject. Still, ArkAgent results suggest that behavioral dynamics in the absence of transaction costs can drive large swings in trading frequency and quota pricing with time. While transaction cost effects are real and can have significant effects on market performance, ArkAgent results indicate that institutional designers should not neglect the need to develop a deeper understanding of what might drive trading and pricing decisions among potential market participants. Past experience with water quantity trading markets in the LAB suggests that social factors other than low tolerance for transaction costs may be reducing participation rates, and ArkAgent demonstrates how this can occur.

Some authors argue that the considerable uncertainties present in trading systems dominated by non-point source trading can mitigate market performance. ArkAgent results show that sustainability benefits can be achieved through a market instrument even in the face of uncertainty. The theoretical implication is that if traders are familiar with uncertainty, and are able to call on more than just economic rationality to make decisions, uncertainty may become less of a problem. Reasoning on this basis, I can go further: the reality is that farmers in the LAB have sophisticated strategies for handling the many unknowns in their operational decisions. Such rich behavioral strategies and acceptance of omnipresent uncertainties may allow traders in the LAB to interact with markets that have uncertain dynamics. The LAB is slightly unusual in having particularly low resilience in its water supply due to overallocation and a tough climate, and so irrigators in more comfortable settings may not have the same adaptive capacity. The ArkAgent results do not shed light on likely responses in such settings, but highlight the importance of understanding behavioral responses to uncertainty before making institutional design choices.

### 7.2.1.3 Designing institutional parameters

**Institutional designers should be more prepared to take the long view**. ArkAgent results show that the choice of WQTM program length can make a difference to how much economic pressure is put on stakeholders. While the ultimate benefits to environmental sustainability appear unlikely to change with longer programs, the time scale over which those benefits are sought can dramatically alter the pricing signals for stakeholders, the time they have to adjust their behavior, and the overall financial health of the system. If a longer salinity reduction program leads to a more stable institution and less burden on economically vulnerable participants while still achieving salinity reduction benefits, the slower rate of improvement may be worthwhile. This is of most relevance to communities with similar economic vulnerabilities to the LAB: mostly family farms, marginal agricultural conditions and distance from major markets. Larger and more economically stable communities may be able to bear more economic pressure to achieve sustainability benefits quicker.

**Institutional designers should pay closer attention to managing price dynamics during the operation of an institution, and be less concerned about how the initial price is set**. ArkAgent sensitivity testing suggests that while the choice of initial quota price does affect magnitudes of outcomes, particularly if quota exceedance fines are tied to quota prices, qualitatively similar outcomes are seen regardless of initial price. More of a concern is the considerable volatility seen in pricing, particularly initially. Such volatility can damage the economic stability of the basin, and should be mitigated with appropriate tools. ArkAgent results suggest that institutional designers should set an initial price relatively low, and be prepared to use additional tools during the program to address excessively high or low pricing. While ArkAgent's behavioral dynamics and institutional structures are based on and tailored to the LAB, the results mostly resonate with existing findings in the resource economics literature, suggesting wider generalizability.

**Institutional designers should take care to address geographic and temporal biases**. Sources of geographic and temporal variation within a basin can lead to systematic biases that will reduce market efficiency and reduce potential sustainability benefits. One well known systematic bias is the upstream/downstream problem, such that downstream users will always suffer

higher salt loads regardless of what they themselves do. ArkAgent results show how these biases can work against market efficiency, and also demonstrate how the use of appropriate institutional tools — like trading ratios — can mitigate their effects. Less well understood and explored are the effects of temporal biases. In many hydraulic societies like the LAB, some actors have strong temporal controls on their decision making, while others do not. If an institutional designer assumes that all actors have the same temporal restrictions (i.e. they all face similar scheduling in their main operational decisions), this will tend to favor some actors above others: those who can make trading decisions at any point will likely come off better than those who can only occasionally make trading decisions, reducing market efficiency. Since all social-hydrologic systems have considerable biases on geographic and temporal grounds, this is an implication with broad relevance.

#### 7.2.1.4 Awareness of complex system dynamics

**Institutional designers should be aware that some paths to or away from sustainability objectives may not be recoverable**. The simulated LAB, like many complex social-hydrologic systems, shows strong path dependency in response to subtle shifts in actor decision making. Aside from being an academic curiosity, the lesson to institutional designers is that some paths to or away from sustainability are irreversible. All real world social-hydrological systems are complex, and there is plenty of evidence from other fields of study that path dependence is a feature of these kinds of systems. It implies that implementing agencies should take especial care in exploring beforehand what potential paths to sustainability may be preferred and how to promote those preferred paths. While ArkAgent does not shed light on the issue of communicating the concept of path dependency to stakeholders, experience and common sense suggests that institutional designers may want to communicate the potentially irreversible implications of particular decisions, since this can emphasize the importance of making the right decisions as often as possible. The use of models like ArkAgent may be helpful in this kind of effort.

**Institutional designers should be aware that a given complex system may show a considerable time lag before responding noticeably to a given incentive**. ArkAgent showed that

the behavioral response to sustainability incentives was slow to occur, and initially trended in the opposite way to what was intended. Given that the real world is considerably more variable than the limited level of heterogeneity included in the ArkAgent scenarios, institutional designers can expect even longer lags and perhaps more temporal complexity in reality. This counsels patience in institutional design and caution in setting expectations too high. Political cycles are far, far shorter than the physical and social response time of large and complex basins, so institutional designers may have to be careful to reconcile the disparity in timescales between political expectations and system response.

**Institutional designers should be wary of categorizing a given population as "pro" or "anti" a particular sustainability outcome**. ArkAgent results show how it is theoretically possible to achieve sustainability outcomes regardless of whether or not a given community is initially receptive to an institutional intervention and a specific incentive package. ArkAgent results demonstrate how a population attitudinally oriented against sustainability outcomes may take longer to move towards more sustainable behavior, and that initial outcomes may even worsen initially. Institutional designers may simply need to allow more time in these cases, and as mentioned above, take care not to set expectations too high. Few populations of water users can be monolithically characterized as for or against a particular sustainability outcome, suggesting that this implication may have broad relevance across basin settings.

### 7.2.2 General implications for salinity reduction in the Lower Arkansas Basin

#### 7.2.2.1 Possibilities for salinity reduction in the LAB

**Above all else, ArkAgent demonstrates that salinity reduction is possible in the Lower Arkansas Basin**. The model demonstrates that annual salinity reductions of up to 10% can be achieved using a modified water quality trading market institution without causing dramatic societal and physical disruption. ArkAgent's results suggest that salinity reduction may be possible through incentivizing changes in farmer and ditch manager strategies for water use and management. The model does not indicate that a salinity reduction effort would necessarily improve the equity in water distribution and income within the basin, but it would

316

certainly not make things worse. It is important to note that while the ArkAgent WQTM was designed to be theoretically possible within the existing legal structure, ArkAgent does not simulate creation of the institution itself.

**However, the WQTM in the LAB setting does not have infinite capacity to improve sustainability outcomes**. Results from ArkAgent indicate that even the most well incentivized scenario under the most favorable institutional and behavioral assumptions cannot achieve more than around 80% seepage and recharge reductions relative to historic, probably regardless of how long the institution was tried. In fact, we should assume that the 80% ceiling overestimates capacities for behavioral and technological change within the basin, due to the simplicity with which many aspects of salinity reduction are simulated. Any Colorado agencies seeking to roll out novel institutions should be aware that if the stakeholder group is not necessarily going to be greatly changed by the intervention, there will be upper limits to what can be achieved through the intervention. Any given mix of stakeholders has a finite capacity for behavioral change, both given the natural limits of human cognition, and the fundamental physical limits imposed by the difficult climatic and hydrographic conditions in the LAB. Evidence from time series data suggests that river flows exert overwhelming macroscale control on what levels of salt reduction can be achieved by a given institutional form. These data also follow general findings in the WQTM literature that institutional design (in the LAB, and elsewhere) must be ready to adjust to climatic shifts. Finally, we should note again that ArkAgent is a non-equilibrium model that provides no guarantee of future performance, and may exaggerate the pace of change since it omits many behavioral, political, economic and other institutional brakes that are seen in the real world. While trends point towards continued salinity reduction beyond the simulated eight year period, there is absolutely no guarantee that this would be the case. Consequently, one of the most important areas for further work is an extension of the simulation time period.

**Perhaps most importantly, ArkAgent results show that an institutional intervention to address salinity need not necessarily clash with the prior appropriations framework**. ArkAgent was run with the assumption that no solution could take precedence over the allocation of water according to legal decrees. The results show that considerable benefits were gener-

317

ated without dramatic shifts in water allocation patterns. It also shows that improvements in irrigation efficiency did not necessarily result in increases in consumptive use. This is a particularly important implication in Colorado, where efforts to address salinity are limited by the fact that the prior appropriations framework has considerable legal privilege. It is also important in relation to the Colorado-Kansas interstate compact governing Arkansas River flows, since this compact is sensitive to the timing and volume of water delivery. A caveat to recall is that since the water courts were not simulated in ArkAgent, these results have nothing to say about whether irrigators and ditches would perceive injury and file formal complaints.

### 7.2.2.2   Costs and benefits of salinity reduction through a WQTM

**Salinity reduction would cost farmers and ditch managers in the investment required to improve equipment and line ditches, and for some actors, in initially higher debt loads to meet financial obligations to the WQTM**. ArkAgent results show that the economic costs of the institutional intervention are front-loaded in the first few years of the simulation: higher investment costs in technology upgrades, and higher levels of borrowing to pay off fines incurred by exceeding quota. Clearly some stakeholders will lose out in the early years of a WQTM, since it is unlikely that the fines for exceeding quota are being distributed equitably (i.e. not all polluters are exceeding their quota by the same amunt). But this is an unavoidable feature both of the market and of the problem: someone, somewhere, must pay something to reduce salinity loads. Furthermore, the associated borrowing does decrease with time, dropping to to levels below the empirical baseline by the end of the simulation. In addition, ArkAgent underestimates the financial benefits from salinity reduction, since the model does not simulate improved crop yields from reduced soil salinity, or the potential economic gains to longer term shifts in land use.

A market institution for salinity reduction in the LAB would not necessarily drive up levels of water leasing to municipalities relative to the historical condition. However, more sophisticated incentive structures targeting sustainable change do tend to increase leasing relative to more traditional market mechanisms. Many of the scenarios indicate that a WQTM institution would in fact lower the volume of these transfers by providing crucial additional

income to those agents otherwise prone to raising more cash through leasing their water rights. Water transfers have proven to be a controversial topic in the LAB (Smith, 2008), so it is helpful to demonstrate that a WQTM need not necessarily increase water exports from the basin.

**Sustainability gains in terms of salinity may involve trade-offs in other categories of sustainability**. Policy makers are accustomed to unexpected and undesirable side-effects to environmentally desirable policies, and the LAB is no exception. ArkAgent results show that reduced salt load was typically accompanied by lower flow volumes in the Arkansas River (and vice versa for higher salt loads). This goes beyond the simple fact that higher river flows often increase absolute tonnages of salt: ArkAgent shows shifts towards higher consumptive use among farmers as part of behavioral changes to reduce salinity, in turn reducing return flow volumes. This suggests that real world institutions to address salinity in the LAB may also need to address both legal and political concerns within the basin and with Kansas regarding effects on river flow.

### 7.2.2.3   Stakeholder roles and available tools in salinity reduction

**Ditch managers have the most potential in the LAB for helping achieving reductions in salt load**. ArkAgent shows that the ditch managers have capacity to reduce salinity disproportionate with their numbers. In fact, ArkAgent probably underestimates this capacity, since it does not equip simulated ditch managers with the ability to tailor quota allocations among shareholders with their likely considerable experiential knowledge of real basin conditions. ArkAgent results suggest that the use of polyacrylamide (PAM) sealants to reduce ditch seepage may be a particularly appropriate strategy for these actors. Again, ArkAgent's crude simulation of PAM application logic suggests the results may underestimate potential seepage reduction: real world ditch managers have knowledge of ditch conditions that can improve the efficiency of PAM application.

**Farmers have a lesser but still significant potential among actors in the LAB to achieve reductions in salt load**. While ArkAgent shows the potential for improve efficiency through irrigation equipment upgrades, much of the salt load reduction comes from shifting away from extensive grass and corn to more intensive vegetable crops. It appears that the dominance of

alfalfa in this system is little altered by the WQTM institution, but that the remaining 10-20% of cropland is where farmer agents feel most comfortable shifting their crop mix to more water intensive crops over smaller acreages. ArkAgent results provide limited support for the use of federal and state grant making capacity to facilitate irrigation equipment upgrades, but the relationship is complex and further work is needed to fully determine the potential role of outside financial support. Finally, ArkAgent results show that farmers have an important role to play with regard to any water surplus to their needs but which they are still legally entitled to divert. Long tradition and water law conspire to encourage farmers to divert as much as possible, which is both inefficient and contributes to increased salt load. ArkAgent results for some of the scenarios show the potential improvements both to river flow and salt load if farmers divert less of this surplus.

**Municipal actors, while not directly implicated in the simulated WQTM, can still impact WQTM outcomes**. The ArkAgent scenarios demonstrates ways in which municipalities can get involved in the salinity reduction effort: they can lease from the basin as a hedge against long term shortage; and they can contribute to loan and grant instruments that support infrastructure improvements for both ditches and farmers. Both roles were important in supporting salinity reduction in the ArkAgent runs. Leasing reduces diversions, so reducing exposure of river flows to salt-loading shales, and if the leasing does not physically remove water from the basin (most leases so far have not done so), the extra water can dilute existing salt loads to reduce concentrations. Municipalities are the main sources of out-of-basin funding for actors inside the basin, and have considerable financial means; ArkAgent simulates limited contributions by individual municipalities, on the assumption that municipalities have political and material interests in ensuring that the water in the basin is less polluted and that the communities of the basin bear goodwill towards municipal involvement. Finally, regional load data suggest that the in-basin municipal actors (Colorado Springs, Fountain Creek and Pueblo) can and should participate in any WQTM implemented in the basin. Contributions to river salt load through treatment plant return flows in Fountain Creek have substantive effects on Arkansas River flows, and so these municipalities should not be excluded. ArkAgent did not, however, simulate this participation due to technical limitations.

**Federal actors could have an indirect role in providing financial support for irrigation and ditch infrastructure improvements**. ArkAgent simulates a Federal grant program, modeled on the "Water 2025" program run for some years by the Bureau of Reclamation. While no detailed study was made of the effects of this program, ArkAgent results clearly show that grants were utilized in both ditch lining and irrigation equipment upgrades. ArkAgent does not tell us anything about the political viability of such a program, or the effects of transaction costs on grant disbursement, but does indicate how useful a well-run Federal grant-making effort can be in supporting basin sustainability outcomes.

Since the state actors were not simulated in ArkAgent in any detail, these results shed no light on their potential role. I reiterate the importance of addressing this as a priority area for further work. This will help define more clearly what the role of the state agencies can and should be, and with appropriate modification ArkAgent would be an excellent tool for the job.

## 7.3 Next steps for the lower Arkansas Basin

ArkAgent has shown that salinity reduction is possible in the LAB, and that a water quality trading market may be an effective means to achieve this reduction. Like all simulations, Ark-Agent can only tell us what the outcomes are for a given set of assumptions, and cannot offer empirically grounded insights on anything not included in the model. However, the individual interviewing suggested that there was an appetite for innovation and change in the management of salinity in the basin, and so in the following section I offer a vision for next steps in the basin, steps towards the ultimate goal of returning the Arkansas River to a healthy aquatic and riparian state, and a positive resource for basin society and economy. Where possible, I base this vision on empirical model outcomes; where this is not possible, I base the vision on experience and evidence from other settings.

**Salinity Reduction Step 1** *The Colorado Water Conservation Board should convene a formal lower-basin-wide collaborative modeling process, run by both state agency professionals and in-state university researchers. The process should focus on communicating that it is possible to address water quality while not impinging on water availability, disrupting existing agricultural operations, or adding onerous reg-*

*ulation. A collaborative modeling process is more preferable to alternative collaborative fora, since communicating the links between individual behavioral change and basin-wide benefit will be greatly aided by a computational tool that makes these links explicit. Model development should focus on communicating potential benefits to individuals and the basin community at large. It is important that such a process access as wide a range of stakeholders as possible, both in terms of geography and interest, and make it clear that the WQTM design is not fixed and is open to stakeholder contributions. The basin-wide process could easily build on the ArkAgent experience, as well as existing basin-wide fora for engagement (such as the Arkansas Basin Roundtable). This step would likely take 12-24 months, depending on the size of the team working on the process.*

**Salinity Reduction Step 2** *The Colorado Water Quality Control Commission should conduct a review of the science to determine the appropriate annual targets for salinity reduction at different locations within the basin; and the means by which salinity can be tracked (either through direct monitoring or indirect simulation or a combination of the two). The science review should be followed by a legal review to determine what tools the Commission has available under existing law to monitor and enforce participation in the market. The key is to design an institution which can fit within existing regulatory frameworks: changing water law in Colorado to better accommodate the institution is simply not an option. Given the wealth of existing scientific data on water quality, including the long term efforts of Colorado State University in the lower basin, it is likely that a review would take around 12 months to complete, and could be conducted in parallel with the collaborative modeling exercise.*

**Salinity Reduction Step 3** *The Colorado Water Quality Control Commission should establish the market framework, consisting of two main components: first, an online clearing house facilitating the fast and inexpensive trading of quota; second, a management, monitoring and enforcement division within the appropriate state agency but incorporating as many basin stakeholders as possible. As far as possible, complaints or grievances arising from the operation of the market should be handled through the specialized management team, or failing that through the existing water courts. Transparent management would set apart this framework from other water management approaches in Colorado. Management of the market should at the very least include the use of diverse economic and non-economic incentives, such as equipment upgrade grants to ditch managers and farmers, best management practice workshops, and*

*informational feedbacks. Incentive delivery should exploit social dynamics within the basin to enhance incentive effectiveness. The market should provide for events out of the control of stakeholders, such as exceptional rainfall and particularly adverse economic conditions. Depending on funding availability, the setup stage may take 6-12 months.*

**Salinity Reduction Step 4** *In accordance with adaptive management principals (Pahl-Wostl, 2007b), the Water Quality Control Commission should then initiate a five year trial period of the market in one year increments. At the end of each year, independent reviews should be conducted to ascertain the over-all effectiveness of the market and the effectiveness of individual incentives used by the market. Market structures should be flexible enough to allow the use of new mechanisms when existing mechanisms do not appear to be working or when new mechanisms may be supplementally beneficial. Managers would need to be mindful, however, of the potential time lags in system response to institutional and behavioral changes. Additional collaborative fora should be routinely scheduled throughout the lower basin, to address emerging problems and highlight benefits.*

Assuming typical institutional lags, the lower Arkansas Basin could be seeing the benefits of salinity reduction within seven years, including a two year initial review and setup period. Assuming the WQTM could survive beyond an initial five year time frame, the basin could expect to see mounting benefits — improved crop yields, improved water quality, reduced legal tension with Kansas, a more resilient agricultural community — on the scale of ten to twenty years.

Obstacles to this ideal process are numerous. There is likely to be political resistance both in the legislature and in state agencies, neither of which have shown much enthusiasm for institutional innovation when it comes to water management. There will be resistance within the basin community from stakeholders who refuse to acknowledge either the problem or its severity. There will also be the normal delays associated with complex institutional action in a water system not accustomed to rapid change. Effective use of collaborative process tools (modeling or otherwise) may help address many of these obstacles, but the creation of this institution would not be guaranteed. In Appendix section 8.5, I discuss in more detail some of the potential legal and political obstacles to these proposed changes.

## 7.4   Practical implications for modeling methods

### 7.4.1   Agent-based modeling as tool for institutional analysis and design

The ArkAgent model has demonstrated the potential for applying agent-based modeling to "higher resolution" institutional analysis and design. Many alternate modeling approaches – not just overtly economic methods — treat populations as monolithically lumped parameters, but ArkAgent shows that this can miss out on the analytical and design implications of variations in behavior within groups of similar actors, as well as variations between groups of different actors. The ArkAgent results focused on tailoring incentives to general behavioral heterogeneity within a population, but also showed that some groups of factors have considerably more leverage for change than others, suggesting that incentive design could also be tailored to these groups. While much policy analysis treats the evolution of institutions in time and space very simplistically, the novel approach demonstrated by ArkAgent allows for much closer inspection of temporal and spatial trends, and may encourage more focus on institutional adjustment at these scales.

ArkAgent is not novel in its sophisticated treatment of individual behavioral models, but it does break new ground in coupling a peer-reviewed hydrologic model with a sophisticated behavioral simulation. Coupling social and biophysical models has long been recognized as of critical importance to achieving a better understanding of natural resources management challenges (Evans & Moran, 2002; Matthews, 2006). Due to the various technical and logistical challenges involved, this has rarely been done in the field of water resources management (although other fields have been more successful; Parker et al. (2008)), and it has certainly never been done in the context of salinity management. ArkAgent advances the field insofar as it demonstrates the feasibility of the integration; overcomes a typical challenge in model coupling (communication between programming languages and platforms) using a relatively simple approach; maintains a high resolution without sacrificing overall spatial scale; links the biophysical and social systems in multiple direct and indirect ways (Parker et al., 2008); and argues for building a social systems model as a supplement to an existing hydrologic tool. Since hydrologic models are usually pre-existing, given the long pedigree of hydrologic modeling, this

may allow faster progress towards the ideal of much wider use of integrated social-biophysical models.

Finally, the ArkAgent example suggests that large-scale integrated modeling of hydraulic societies is both possible and potentially insightful. ArkAgent incorporates both rural and urban actors interacting with a spatially large water system with considerable legal and infrastructural complexity. Results from the simulations show how interactions between rural and urban actors can drive changes in basin conditions, and how individually small modifications to infrastructure implemented across the entire basin can affect sustainability outcomes. Charting paths to hydraulic sustainability will at least partially depend on an understanding of interactions between diverse actors at scale, and ArkAgent demonstrates the value of fine-grained but large-scale simulation modeling in achieving this understanding.

### 7.4.2   Implications for collaborative modeling methodology

It is important to emphasize that the ArkAgent model development exercise was not designed to evaluate one collaborative modeling methodology versus another. Conducting a proper evaluation of alternate collaborative modeling methodologies requires a replication setup and a focus on the process itself, while ArkAgent's development was solely designed to maximize benefits to model quality, credibility and salience. The chosen mix of methodologies was suitable for purpose, but this does not suggest that the chosen approach is any better or worse than the large number of potential alternatives.

With that in mind, the combination of individual interviewing, collaborative modeling workshop, Delphi validation surveying, and psychometric testing was broadly effective in facilitating the timely construction of a reasonably comprehensive complex systems model. The resulting tool was capable of simulating scenarios in line with the study hypotheses, as well as contributing to stakeholder engagement benefits like systems learning and new communication opportunities among stakeholders. Problems with individual methods can be addressed by specific fixes (see section 6.7.3).

There is ample evidence that were the ArkAgent process to be carried forward into a decision making framework oriented towards finding and implementing salinity solutions, the

325

modeling would be an effective support of more efficient decision making involving far more stakeholder involvement than is currently the norm in the LAB. ArkAgent could play a role in addressing specific decision support concerns, or in furthering the political process by providing a collaborative forum for stakeholder communication and shared systems learning.

The ArkAgent modeling process resembles an intermediate application of the collaborative modeling canon: while not directly focused on a specific decision making outcome, and not expressly designed to attract broad stakeholder participation, the model has still yielded considerable practical insights for the basin. The participatory elements were less broad than is traditionally seen with companion or mediated modeling, but still contributed both to the model and — in all likelihood — to stakeholder understanding of the salinity issue. In Appendix section 8.4.1, I reflect in more detail on this intermediate role for such models, suggesting that the ArkAgent context may be more common than we think. I argue that this kind of modeling calls for adaptive and eclectic use of a wide variety of modeling methods, and suggest one potential set of steps based on the ArkAgent experience.

## 7.5 Theoretical contributions

I initially argued that the Common Pool Resources (CPR) literature is excessively straitened in the behavioral models it assumes, to the extent that it tends to filter the nature of resource management (NRM) problems through the specific behavioral lens of self-interest. **ArkAgent shows that it is possible to conduct NRM analyses using different analytical foundations**. ArkAgent builds on the empirical findings that people are not always single-mindedly pursuing their self-interest, and takes this insight a step further. It shows concretely that a real system can be usefully modeled using much more diverse behavioral assumptions, and that it is reasonable and practicable to make behavioral heterogeneity the core assumption of the model (rather than a single specific behavioral model, like the B-RCT). It matters little if my choice of a particular set of alternative behavioral models turns out to be drastically off target. I will have still demonstrated that the outcomes we expect can be critically affected by the behavioral assumptions we make to begin with.

**This study opens up a new area of institutional design for sustainability**. ArkAgent shows that in testing our institutional ideas we can not only vary the institution but also our fundamental assumptions on behavioral responses to that institution. ArkAgent shows how new institutional mechanisms can be tested with alternative sets of behavioral models, and that this testing can occur using a real world context as a basis. ArkAgent results show that broadening the incentives provided by a market institution can improve its performance, and this may help improve both institutional efficiency and potential sustainability gains in real systems. ArkAgent results also show how the social impact of an institution can be gauged, and how institutional design could be adjusted to minimize the negative sides of particular social impacts. ArkAgent is, of course, limited in what variations of institutional design it can test, and what set of behavioral assumptions can be explored, but it represents a first step towards the wider application of decision support models in the interests of better institutional design: parameterized with rich empirical data, structured to include new institutional ideas, and run over both historical and future periods to analyze past performance and potentially shed light on future outcomes.

The ArkAgent model is not novel in general terms of using more sophisticated behavioral models to explore policy questions, but it does go further than most models in dispensing with the typically uncritical adoption of bounded rationality as a core behavioral assumption. ArkAgent also provides a concrete demonstration that a widely-used behavioral framework (the generic "belief-desire-intention" structure; Rao & Georgeff (1998)) can be used to develop non-traditional behavioral assumptions. ArkAgent does break new ground in **linking the societal model to a sophisticated and scientifically-validated surface and ground water model** (also incorporating geochemical simulation), and doing so in a way that provides for complex and dynamic two-way interactions between agents and their hydrologic environment.

## 7.6   Addressing the problem frame

In Chapter 1 I defined the core problem of the thesis as one of institutional failure within the context of hydraulic sustainability. I defined the challenge of "hydraulic sustainability" as the

task of ensuring that large populations dependent on extensive infrastructure for their water supply needs develop healthy societal and environmental foundations for the future. I identified institutional failure in attempts to address salinity in the Arkansas Basin and around the western United States, and used a water quality trading market simulation to test assertions on potential reasons for the broad failure of this kind of institution. What light, if any, did ArkAgent shed on the broad question of institutional failure in water quality trading market institutions, and on the wider issue of hydraulic sustainability?

One of the new contributions of ArkAgent is to **treat this issue explicitly using a complex systems simulation**. There has been no previous modeling effort that has targeted the issue of salinity within the context of hydraulic sustainability, and certainly none that has focused on a specific setting in the American West. At the very least, ArkAgent shows that exploratory analysis of institutional failure can take place using complex systems simulation, and without making empirically weak assumptions on stakeholder behavior. While ArkAgent was never intended to analyze all the potential flaws of the complex institution that is a water quality trading market, it does open the door to alternative analyses that might shed more light on this problematic issue. While this is not a new contribution per se, ArkAgent continues to support the use of non-equilibrium modeling in the analysis of complex systems (Epstein, 1999). More novel is the use of non-equilibrium modeling in the analysis of the introductory period of a complex institutional intervention: ArkAgent shows that even if logistical constraints dictate a limited simulation period, there can still be considerable insights gained.

ArkAgent results provide logically rigorous evidence that **our behavioral assumptions, not just basic features of institutional design, can dramatically alter the success of an institution and its outcomes**. While ArkAgent's simulated WQTM does not undergo "failure" per se, it does show poorer performance under empirically-based assumptions of behavioral heterogeneity than under assumptions of homogeneously and boundedly rational actors. Consequently, ArkAgent results suggest that one additional reason why real world WQTMs have consistently failed to engage participants may lie in the mismatch between the incentives a WQTM provides — typically purely financial, since this is what economic theory predicts will be most effective — and the complex, diverse and varied attitudes and behavioral strategies

that real stakeholders possess.

Perhaps most importantly, ArkAgent shows that **it is possible to conceive of a sustainability challenge involving a common pool resource as potentially something other than a "social dilemma"**. With this new framing, the LAB salinity challenge becomes a problem not of mitigating individual self-interest, but of identifying and targeting the diverse personalities and strategies inherent in such a large and complex system. As the ArkAgent results demonstrate, while this adds uncertainty it also allows us to be much more inventive in the way we design our institutions. With careful study of a given context, and attention to the real attitudes and behavioral strategies of a basin community, institutional interventions can be potentially much more successful. ArkAgent results also show that we may be able to start to move away from the artificial division between "market" and "government" solutions. It matters less whether we employ the free market or regulatory mandates, as long as we make sure that individuals are empowered to make the changes needed.

ArkAgent illustrates that **we do not necessarily need new infrastructure solutions for our existing environmental problems in large social-hydrologic systems**. The sustainability gains realized in the ArkAgent simulations come purely from shifts in water use and management by basin stakeholders, in response to an institutional intervention. No new technologies were introduced to the system, although existing technology (in the form of simulation tools, irrigation equipment and new canal lining techniques) did play a part in improving sustainability outcomes.

Furthermore, ArkAgent shows that **an institutional intervention need not mandate high-level institutional changes to be successful**. The ArkAgent WQTM relies on leveraging small changes at an individual stakeholder level towards basin-wide shifts in outcome, potentially minimizing expenditure for government agencies while ensuring that sustainable change by individuals receives appropriate reward. The institution in ArkAgent has a relatively small footprint in terms of economic and social disruption. The principal benefit of this is that it is relatively reversible, particularly useful if real world outcomes turn out to be greatly and unexpectedly different from simulated results.

It is important to note that the ArkAgent model skirts some significant issues: the problem

of getting an institution implemented in the first place; the potential issues with what can happen when political interests become aligned against an existing institution; and the extent to which the legal system can be used in support or in opposition to the institution. ArkAgent does not, in fact, allow any changes to the governing frameworks (legal and otherwise) already in place in the basin. As such, it is clear that much work remains to be done. However, one of the intents with ArkAgent was to demonstrate what is possible if stakeholders decide to collaborate over a potential solution. ArkAgent focused on a relatively cheap and minimally disruptive option, partially in the hope that by showing positive outcomes this may encourage the real stakeholders to be more proactive in their own efforts to find solutions. ArkAgent is clearly not a thorough exploration of all the potential options for reducing salinity in the LAB, but it certainly provides an example of one potential pathway to increased social and environmental sustainability in this basin and others like it.

## 7.7   Further work

I discussed specific areas of model improvement in section 6.8 in Chapter 6. In this final section I focus on general ideas for further work raised by the ArkAgent modeling effort.

Subsequent iterations of ArkAgent should focus at least some attention on **the simulation of the legal and political dynamics associated with water law in Colorado**. This implies a larger and more sophisticated role for the state engineer; the deliberate simulation of various state agencies; and some version of the water courts system. Without this, the model cannot be truly relevant to policy makers. Such relevance will also be enhanced by other carefully targeted improvements in sophistication. Indeed, many of the specific modeling improvements suggested in section 6.8 of Chapter 6 revolve around enhancing the model's sophistication in the way agents are simulated, the components of the WQTM institution, and in the supporting framework of the model and the GeoDSS. This model enhancement would be best accomplished within a longer term (two to five year) modeling project.

In keeping with the theme of improving model sophistication in general would be a **new focus on the drivers of behavioral heterogeneity within a resource user population**. ArkA-

gent assumed that behavioral heterogeneity existed a priori, and apart from the incorporation of an attitudinal change mechanism, did not make any attempt to simulate the development of behavioral variation. There is evidence that behavioral heterogeneity has roots in environmental conditions, the cultural milieu, legal setting and other factors, all of which are fair game for simulation with the kind of high resolution behavioral modeling that ArkAgent makes available.

Specific elements of ArkAgent could be packaged as **discrete modules available for modeling work in other basin settings**. Modifying the entire model for another basin setting would be a mammoth endeavor, and probably not worth the effort. The development process requires engagement with stakeholders regardless of how much model is already built, and if too much model is provided readymade, this short-circuits the opportunity for stakeholders to contribute. However, select components of ArkAgent may have considerable utility in other settings. The core behavioral model ("WaterAgent"), and associated modified-BDI structure, would be useful either as a blueprint for development in other programming languages, or as a Java library ready for use in other Java-based models (Kock, 2008). To get to that point, considerable work would be necessary both technically and logistically to isolate and package the selected model code, and then to make it available through an appropriate online channel.

The ArkAgent framework provides a rare example of linked social-hydrologic simulation, and **a concomitant opportunity to trial sophisticated agent behavioral models in conjunction with a physical simulation model**. Few water resources models have successfully integrated social and physical simulation, and too many sophisticated behavioral models from computational psychology and other fields languish, disconnected from the much larger range of available physical models. Unusually, while ArkAgent builds on relatively simple ideas from various behavioral modeling fields, it also incorporates sophisticated physical process modeling. The lessons from this development effort could be codified into recommendations for future linked models of a similar nature. Further, trialling more sophisticated and empirically-grounded behavioral models would not only shed further light on many of the hypotheses tested in this study, but also greatly expand methodologies for linking social and physical models in the realm of water resources management.

**ArkAgent could be re-focused to gather comprehensive data at regional and sub-regional scales, and test novel institutional mechanisms for exploiting geographic and temporal variation**. Few institutions in water resources management successfully account for spatial and temporal variation in complex basin settings, and few decision support tools are available to help improve institutional design in this way. ArkAgent's results identified geographic variation as an important element affecting model outputs, but also suggested the possibility of institutional mechanisms that could exploit such variation. An important caveat here is that to enable such novel data collection and scenario testing, the model would either need a great improvement in efficiency, or improvements in the efficiency of the modeling platform it runs on.

Finally, ArkAgent should be taken back to stakeholders. Logistical limitations meant that the existing model was not returned to stakeholders either in workshop form or through appropriate online means. At the very least, exposing the finished model to stakeholders would offer stakeholders the chance to learn about their basin in a novel way, as well as exploring their own particular hypotheses on the salinity issue. At best, returning the model to the basin could drive political interest in dealing with the salinity problem.

# Chapter 8

# Appendix

The Appendices include material that is either non-essential to the main arguments but of interest nonetheless; or essential but at a higher level of detail than is appropriate in the main text; or is of use as a general reference to the main text. Note that additional material is included on the dissertation CD; this material was too lengthy to include in the main Appendix.

Section 8.1 describes important terms used in chapters 5 through 7. Section 8.2 provides some tangential but nevertheless interesting insights into stakeholder visions for basin sustainability (mostly based on individual interviewing results). Section 8.3 provides a detailed description of all the model limitations referenced in section 6.9 of chapter 6. Section 8.4 includes a variety of material relating to the model development methods, including a detailed exposition of results from the interviewing and workshops, and descriptions of additional materials available on the CD attached to the dissertation. Section 8.5 offers some educated speculation on potential obstacles to implementing a real WQTM in the LAB. Section 8.6 describes how the model can be installed and run on any computer running Windows XP. Finally, section 8.7 includes a number of plots referenced in chapters 5 and 6 but omitted from those chapters for space reasons.

## 8.1   Glossary of terms

### 8.1.1   Response factors

Response factors are ArkAgent output variables used to score all scenarios on overall sustainability.

1. Water quality: tons of salt loaded into the Arkansas River; environmental metric

2. Water quantity: acre-feet of flow at the Colorado/Kansas state line; environmental metric

3. Unique traders: sum of unique buyers and sellers in the water quality trading market (WQTM); institutional metric

4. Clearing success rate: ratio of quota volume (kg) offered for sale to quota volume (kg) sold; institutional metric

5. Volume of trades: total volume of quota traded (bought and sold); institutional metric

6. Agent income: mean income for ditch managers and farmers; economic metric

7. Real assets: cash assets less liabilities for ditch managers and farmers; economic metric

8. Conflict level: level of subjective conflict felt by the farming population; social metric

9. Deficit equity: extent to which any water deficits are felt by all or some of the farming population; social metric

### 8.1.2  Scenarios

The scenarios are usually referred to in full (e.g. "scenario 1a"), but to economize on space, are abbreviated in charts. The following list briefly describes each scenario, and lists the plot abbreviation.

- Empirical baseline - attitudinally and behaviorally heterogeneous agents; behavioral models were empirically parameterized, and this run was calibrated to the historic data; this scenario did not include a WQTM. EB on graphs.

- Rational baseline - attitudinally and behaviorally homogeneous agents = "rational baseline" or RB on graphs; behavioral models were forced to a low S-factor value; this scenario did not include a WQTM. RB on graphs.

- Scenario 1a - classical WQTM, homogeneous actors = S1a on graphs

- Scenario 1b - classical WQTM, heterogeneous actors = S1b on graphs

- Scenario 2 - WQTM utilizing post-hoc informational incentives = S2 on graphs

- Scenario 3 - WQTM utilizing post-hoc informational incentives, agents capable of attitudinal change = S3 on graphs

- Scenario 4 - WQTM utilizing post-hoc information incentives and information on neighbors = S4 on graphs

- Scenario 5a - WQTM utilizing post-hoc informational incentives with opportunity cost = S5a on graphs

- Scenario 5b - WQTM utilizing post-hoc information incentives with opportunity cost and information on neighbors = S5b on graphs

In addition, as part of testing Hypothesis 5(ii), two extra runs each for 5a and 5b were conducted:

- Scenario 5a-LSF - same WQTM as 5a, but with Sustainability Factors parameterized using a non-empirical, beta distribution skewed to the low end of the scale = S5a-LSF on graphs

- Scenario 5a-HSF - same WQTM as 5a, but with Sustainability Factors parameterized using a non-empirical, beta distribution skewed to the high end of the scale = S5a-LSF on graphs

- Scenario 5b-LSF - same WQTM as 5b, but with Sustainability Factors parameterized using a non-empirical, beta distribution skewed to the low end of the scale = S5b-LSF on graphs

- Scenario 5b-HSF - same WQTM as 5b, but with Sustainability Factors parameterized using a non-empirical, beta distribution skewed to the high end of the scale = S5b-LSF on graphs

### 8.1.3 Parameter variation experiments

A set of model parameters were selected for sensitivity testing on the basis of their importance in model logic. These are described, along with their abbreviations on the parameter variation plots, below. The values of the varied parameters represented extremes relative to the baseline run, which is plotted on all parameter variation charts as "base".

1. Risk strategy: in the "risk avoiding" variant, a given scenario was run with all agents parameterized to avoid risk in their decision making wherever possible. In the "risk seeking" variant, a given scenario was run with all agents parameterized to be risky in their decision making wherever possible.

2. WQTM management options: in the "lax wqtm" variant, the WQTM in a given scenario was modified to allow quota banking and deficit cancelation. This meant that at the beginning of each new water year, the State Engineer would allow traders to keep any surplus quota ("bank" it) and use it in future years. It also meant that the State Engineer would cancel any quota deficit (i.e. the result of the agent polluting over quota) incurred more than one year ago.

3. Farm skill: in the "low farm skill" variant, a given scenario was run with all farmers parameterized to have low levels of skill in farming practices. In the "high farm skill" variant, the agents were parameterized to have very high levels of farming skill.

4. WQTM length: the length of the WQTM determines the annual decrement in total quota allocation, intended to move the basin towards a target level of salt load at the end of the program. In the "8 year WQTM" variant, the length of the WQTM was set at 8 years, while in the "20 year WQTM" variant the length of the program was set at 20 years.

5. Initial quota price: the price of a unit of quota (1 kg of salt) was set in the baseline run to $0.5. In the extreme variants, the price was set at $5 ("initial quota price at $5") and $10 ("initial quota price at $10").

## 8.2 Sustainability insights from qualitative methods

The individual interviewing and workshops yielded some additional results that shed light on overall efforts to enhance sustainability in the basin. The basin community shares a narrative of decline in their way of life, and recognizes the environmental and economic dimensions of this decline without necessarily describing these as part of "sustainability". The concern of many stakeholders with finding solutions to an existential crisis suggests that a broadly defined sustainability initiative may gain traction in the basin. Legal obstacles are serious and

potentially insurmountable, however, and so immediate prospects for change may rely on innovative institutional design and on harnessing the considerable ingenuity of the stakeholder community.

The interviewing identified a shared fear that the agricultural way of life was being lost as water uses shifted and increasing economic and climatic pressures were felt in the basin. Stakeholders are clearly aware that the basin is under stress in a number of social and environmental dimensions, and that something needs to be done. There also appeared to be an appetite for change: interviewees volunteered sustainability-oriented initiatives like an expansion of the USDA-funded Conservation Reserve Program (fallowing salinized, marginal or otherwise unproductive land for conservation purposes), a broad shift from irrigated cropping to dryland ranching or protected prairie, and even an unprompted suggestion for some sort of trading system to address salinization. Many spoke of the need to enhance the resilience of the system, so that the agricultural economy was more stable and the environmental conditions less prone to fluctuation. Basin stakeholders are interested in "sustainability" largely from the perspective of preserving the population and well being of basin communities, but recognize that the health of basin society depends on maintaining good environmental conditions. While salinity is not top of the list of concerns for most stakeholders — water scarcity far exceeds water quality as a worry — most stakeholders recognize it as a problem.

The broad awareness of sustainability challenges by stakeholders, and an apparent interest in making change, suggests the conditions are ripe for an initiative to further sustainability in the basin. But members of the basin community have complex and varied concerns, and similarly diverse perspectives on what solutions are needed. This diversity suggests that any sustainability initiatives should touch on more than just salinity (as important as this issue is, and clearly there is room for education in this regard). Again and again, stakeholders in interviewing and workshops raised the concept of resilience, that their community could survive in the face of overwhelming macro-scale pressures from the outside world, yet still engage with that world to the basin's benefit. It may be that a successful sustainability initiative in the LAB will be one that addresses community concerns with basin resilience.

Many interviewees raised the lack of institutional flexibility in the system. Most stakeholders echoed the core assertion of this thesis, in that water movement in the basin is no longer a naturaldynamic but is governed by the legal system and limited by the available delivery infrastructure. Most stakeholders also suggested that the existing legal framework is deficient. Stakeholders spoke variously of the need for more flexible laws, regulations and operational practices; more innovation in decision making; and updated laws and regulations reflecting the modern reality of the basin. In chapter 2 I discussed the stagnation of water quality control efforts due their conflicts with water allocation law. I mentioned that one reason water quality suffers is that it became a concern much later in the state's history of water lawmaking, long after allocation to beneficial uses in priority had become the cornerstone of water management. Stakeholders are clearly also cognizant of this tension, and other side effects of an archaic legal doctrine. They are aware that the legal system poses obstacles to virtually any effort to be innovative in the way water is managed, and are realistic about how challenging it is to make changes to the legal framework: when asked for realistic predictions on the future state of the basin, not a single stakeholder mentioned "changed or improved law and regulations".

By being an obstacle to change, the legal system is also an obstacle to sustainability. Any efforts to improve basin resilience by tackling areas of sustainability like salinity and water availability will have to contend with the existing legal system, at least until there is more

success in legislating change. But it may be possible to harness stakeholder skills in making things work despite legal obstacles. One long interview with a ditch manager made it clear that the community of ditches engages in constant and complex negotiations to ensure that water gets where it is supposed to go, and that not all of these interactions conform with prior appropriation doctrine. The stakeholders of the basin have shown considerable ingenuity so far in making a modern agricultural system work with a 19th century body of water law.

## 8.3 Detailed descriptions of model limitations

### 8.3.1 Limitations from simplification

The first category of limitation is rooted in a number of large and small simplifications to do with the fact that ArkAgent is still just a model, albeit a very complicated one. It is readily apparent from the discussion that one of the most significant of those limitations was the agent decision logic associated with WQTM trading behavior (specifically, the posting of surplus for sale). The logic clearly was too simple, but this was more from exigency rather than choice. Another important limitation identified in Chapter 4 relates to the calibration performance of salt load: ArkAgent systematically underestimates real world salt load by around 30%. Fortunately, since the effect does appear to be systematic, we can correct upwards any quantitative data we apply to the real world. A third important limitation is that ArkAgent does not simulate soil salinity. This is technically feasible, but would have taken additional and unavailable resources to integrate soil salinity with the GeoDSS. This removes a potential feedback for farmer and ditch manager decision making, since farmers would theoretically notice improved yields due to reduced salt accumulation. However, this limitation may not be critical since, given the short period of time simulated by ArkAgent, it is unlikely we would see significant shifts in soil salinity on a basin scale.

### 8.3.2 Limitations from non-equilibrium modeling

ArkAgent, like many agent-based models, is a non-equilibrium model. Many complex social systems may theoretically attain equilibrium on certain grounds, but this is usually only possible after very long lead times (and in some cases is theoretically not guaranteed) (Axtell, 2000; Epstein, 1999). In such systems, non-equilibrium analysis is both necessary and useful since it is non-equilibrium conditions we will have to design our institutions to handle, for probably all of their functional existence. The limitation of non-equilibrium modeling, particularly for such a complex model as ArkAgent, is that we can say little with confidence about longer term dynamics beyond the end of the simulation period. For example, results for irrigation efficiency improvements suggest a longer term trend of improvement. However, there is no guarantee that such improvements will continue in the model (due to non-equilibrium conditions), and certainly no evidence from reality that irrigation improvements are irreversible or that anything close to 100% is even physical possible. Similarly, while the salt load trends among scenarios appear to be evolving over time to become more different from each other, there is no guarantee that this will continue.

### 8.3.3 Limitations by design decision

The second broad category of limitations arises from deliberate choices in what was not simulated, on the basis of theoretical and practical considerations. Prominent among these was the choice not to simulate transaction costs and WQTM monitoring and enforcement. In both cases I made the deliberate assumption that they would not play a part in model outcomes, since the focus of my hypotheses is elsewhere. Other limitations of this kind include the decision not to simulate any significant federal intervention in the system, on the basis that the U.S. Congress and the EPA have shown little collective interest in this level of interference with state level salinity issues, and partly on the basis that simulating this interference would not be relevant in the context of the hypotheses being tested. I also omitted the possibility of certain behavioral choices by individual agents, such as water sales, usually either on the assumption that such activities are becoming increasingly rare in the basin, or that the implementation of the behavioral choice would add unnecessary complexity not directly impacting the testing of my hypotheses. I address many of these limitations in the further work sections of the discussion.

### 8.3.4 Limitations by logistics

The third broad category of limitations arises from logistical limitations on how many runs of the model were possible. This arose from the fact that ArkAgent takes between 9 and 15 hours to run on a single computer, depending on the specifications of the machine. The problem was compounded by the fact that ArkAgent was developed on the AnyLogic simulation platform, which locks its license keys to the hardware on which the platform runs. University research licenses for AnyLogic cost around $5,000 each, and so it was financially prohibitive to run the model on a cluster (which would otherwise have been a straightforward solution). Consequently, I had to decide how many model variants could be run, and to what extent stochastic processes could be allowed to influence model outcomes. Ultimately, it was not possible to include stochasticity in the model, due to the impossibility of running adequate numbers of model variants. It was also not possible to conduct a full sensitivity testing across all possible parameters.

### 8.3.5 Limitations by model integration

The final category of limitations arises from constraints imposed by the coupling with the GeoDSS model. Due to this model's design, it was not possible to vary input parameters of the hydrologic system (e.g. the flow of water arriving at the westernmost cell of the simulated river). This predicates any true testing of what impact shifts in the underlying physical system might have on the social system. This is one of the most critical tasks for further work, since as I discuss later, the macro-scale trends imposed by the hydrologic cycle can be significant.

Perhaps the most important limitation arising from the coupling with the GeoDSS is that ArkAgent is limited by its historicity. The GeoDSS has only been parameterized for 8 years of data, since this is the only period for which sound empirical data exist. This determined the time period over which the social model could be run. A more traditional modeling effort would see model validation take place over a different historical time period than the calibration dataset. This improves the confidence that the design and implementation of the model do a reasonable job of reproducing real world structures. Consequently, in a strictly technical sense, ArkAgent cannot be understood as a "credible" model until it has undergone a further

cycle of development and improvement in close cooperation with stakeholders, policy makers, researchers and other domain experts (Gass, 1983). Only then would it resemble a true decision support tool.

## 8.4  Modeling Methods

### 8.4.1  Complex systems modeling in the absence of a strong political framework

Complex systems modeling in a political vacuum is an understudied challenge, yet such models can still be useful. ArkAgent provides an example of how an intermediate form of model, bridging the methodological divide between companion modeling (where the model is loosely related to reality and the principal objective is process-oriented; Barreteau et al. (2007)) and decision support (where the model is closely tied to reality and focused on a particular decision objective; Antunes et al. (2006); Keyes & Werick (1995)) can survive and thrive in the absence of a decision process. Instead of seeing the problems with ArkAgent as stemming from the "wrong" environment for collaborative modeling, I argue that ArkAgent is a pioneer in a *new application area*. I term this "exploratory collaborative modeling for sustainability". "Sustainability" implies an overarching concern for sustainable natural resource management, while "exploratory" conveys the way this kind of modeling seeks to balance specific decision support objectives with more general exploration of system dynamics and novel interventions. The modeling remains "collaborative" because of the importance of engaging stakeholders with sustainability issues, finding solutions, and opportunities to learn and interact.

Existing case studies of collaborative modeling, whether companion, mediated, shared vision, participatory or any other form, often describe a process either insulated as a purely research-oriented study, or as a process embedded within a broader and active political effort to address a particular resource management issue. Rarely is there much discussion of how to conduct a collaborative modeling exercise in the absence of strong political support or a pre-existing decision making process. Most studies emphasize the critical benefits of a strong institutional framework, and suggest that this kind of modeling may not be useful in such a setting (Belt, 2004). However, there are clear arguments why collaborative modeling can be useful under such conditions and is still worth doing: systems learning can lead to more sustainable resource management practice by individuals with or without institutional-scale decisions; increased awareness of the utility of systems models through collaborative model development can lead to more frequent and more nuanced use of such tools; agent-based models, in particular, absolutely need the contributions of stakeholders if behavioral models are to be credible both socially and scientifically; and, most hopefully, a model like ArkAgent may open up new debate on issues long out of favor, and even drive new political interest in the problem. Assuming that we do need collaborative modeling regardless of the political climate, what does the ArkAgent experience provide in the way of lessons?

The ArkAgent experience argues for developing an initial model based on early semi-structured interviewing and secondary data. This initial model will be useful in attracting further interest in the modeling effort, in guiding the evolving thoughts of the modeling team, and providing a useful seed for subsequent methods. Importantly, this basic model should incorporate a simple gaming interface: starting early with the interface is helpful, and the gaming approach is one of the most effective ways to mediate interaction with a complex model. This should then be followed by much more highly structured interviewing sessions with individual

stakeholders. These should be tailored to the stakeholder, and make extensive use of the basic model. This kind of interviewing has the primary objective of attracting participants to modeling workshops, and the secondary objective of refining the overall focus of the basic model. Following this second round of interviewing, collaborative workshops should be conducted. The workshop numbers should be tailored to the basin geography and to a maximally diverse group of stakeholders, since this form of collaborative modeling will often lack a clear stakeholder group. The workshops should also be staged as close together as possible, to maintain interest and momentum. These workshops should still follow the broad ideals of companion modeling, such that the model be re-worked during and after each workshop and very little should be left unexposed. However, the requirements of more complex systems models may predicate the full re-working of a model, so the model may should not necessarily be exposed in its entirety in the initial workshops. Finally, remote survey and expert panel methods may be useful later in the process for filling in design gaps, addressing more technical issues, and meeting late-stage budget constraints. There is, no doubt, much further work to be done in improving upon this selection of methods.

### 8.4.2   Individual interviewing

The individual interviewing proved most useful in developing an overall picture of stakeholder concerns and perspectives on the basin. The interviewing provided additional strong support for the thesis assertion that stakeholders in the LAB have diverse perspectives and strategies towards their water use and management. This was particularly evident in the variety of orientations stakeholders maintained towards issues of concern. For example, some farmers were concerned about preserving the agricultural focus of water use, while others were also interested in preserving aesthetic and recreational values. In general, not only did groups of similar stakeholders have different attitudes and strategies from other groups, but attitudes and strategies continued to vary even within groups. Many stakeholders expressed a fear of losing water and so the basin's agricultural way of life. Many of the farmers described deep family connections to the basin; others were concerned that their children would not continue in the profession; and nearly all talked of the precipitous decline of many rural communities in the basin in the latter half of the 20th century. The overwhelming impression is that many members of the basin community have remained in the basin not because of the great profits to be made off LAB agriculture (there are hardly any, particularly relative to other irrigated basins in the West), but because of deep-seated familial and experiential ties to the land.

The individual interviewing results also shed light on important characteristics and dynamics of the basin unrelated to model design or the hypotheses. Interviewing confirmed that stakeholders share the assertion of this thesis that the LAB is a highly engineered water resource system, in which water movement is governed less by hydrology than by the legal framework of prior appropriation. One stakeholder described the Arkansas River as "one big drain", indicating to what extent a complex hydrologic feature has been reduced to a unidimensional component of the agricultural system. Many interviewees noted the significance of the major physical forcers of the system, particularly climatic variation and the impact of ground water geochemistry on salt loading. Farmers in the LAB are well aware that they operate at the mercy of significant physical and economic forces out of their control, and so may be well suited to handling a salinity control institution which is also at the mercy of exogenous forces. The interviewing showed that in addition to the salinity problem, there exist a number of ongoing conflicts among stakeholders. These include concerns over keeping water in the basin versus

leasing or selling water to Front Range municipalities; reconciling recreational/aesthetic values of water with more traditional agricultural values; conflicts between ground water and surface water users (echoing similar conflicts elsewhere in the state); and broad contrasts between the attitudes of urban and rural stakeholders. Interviewing showed stakeholders to be concerned principally about water scarcity; potential competition with urban water users; institutional inefficiencies and inflexibility; and water quality.

The individual interviewing was useful in setting the scene for model development, and helped seed various elements of the design. It also yielded some interesting information unrelated to the model and its core hypotheses, particularly in regard to the nature of water conflict in the basin. But this method also had several limitations as a tool in model development. First, individual interviewing was not particularly effective at testing specific ideas on model development in the absence of the actual model. Repeated attempts to put specific questions to a stakeholder (like "please describe a couple of ways in which you make use of water on a daily basis") were either ignored or were quickly derailed by another issue of concern that the stakeholder was more interested in talking about. Even with the model present, it was difficult to keep the stakeholder critiquing elements of the model as opposed to embarking on a discussion of a separate issue that a model element raised.

The guide used in the semi-structured interviewing is included on the CD in the Interviewing folder. Interviewing transcripts cannot be provided due to confidentiality agreements signed with participants.

### 8.4.3   Collaborative Modeling Workshops

Detailed descriptions of the two implemented collaborative modeling workshops and the one planned (but not implemented) workshop are included in the Workshops folder on the dissertation CD. The workshop descriptions are loosely based on earlier reports made to research supervisors at the Bureau of Reclamation in 2007 and 2008. They are not complete, since many of the conclusions from the original reports have already been discussed in the main thesis text. On the dissertation CD, the Workshop I, II and III folders hold the workshop presentations, stakeholder booklets and other supporting materials. What follows is a summary of the more detailed reports.

Three collaborative modeling workshops were planned, and two implemented, between 2007 and 2009. The first modeling workshop, conducted in October 2007, used a role-playing game (non-computational) to test initial ideas on various foundational model structures (the focal issue, the mix of actors, institutions worth simulating and so on), and to gather information on the ways different stakeholders interacted over the issues. Many of the results were folded directly into model development, but a number of outcomes are of general interest. First, it was clear that all stakeholders were adept at taking on different roles, and showed excellent understanding of the various operational duties and constraints of stakeholders other than themselves. Second, all stakeholders showed a clear interest in and willingness to explore consensus-based approaches to conflict resolution. This was in marked difference from other basin stakeholder gatherings I attended. Finally, many stakeholders expressed strong interest in the use of collaborative modeling to build a common understanding of the water system, and had several ideas of how such a tool might be applied to resolve problems of common interest. The first workshop showed that the use of non-computational role-playing games was of some use in gathering additional information for incorporation in the model, but that the exercise was most useful in providing a gentle introduction of the model and modeling effort

to a captive audience. Asking stakeholders to suspend their disbelief and interact with a computer simulation of a complex model may be easier when they have already had to suspend their disbelief and interact with a non-computational model of the same.

The second workshop, conducted in November 2008, used a role-playing game consisting of the simulation model wrapped in a custom-built interface. The interface allowed stakeholders to take on one of the four main roles in the basin, and make annual decisions relevant to that agent type. The same workshop was scheduled on three separate days, and the first day was productive. The attending stakeholders provided feedback on the simulated options presented to them, the results their decisions appeared to have in the simulation, and on the interface itself. Model criticisms were generally of three types: (1) the model is too simplistic in area..., and what is missing is...; (2) the model is incorrect in area..., and should instead be...; and (3) the model is entirely missing mechanism... Stakeholders provided detailed written comments on both the model and the model interface, and were also verbally debriefed after each round of the game. The use of the computer-based role-playing game proved to be very helpful in model development, and no serious problems presented themselves. However, it was clear that stakeholders differed widely in their expectations for the complexity of the interface: one farmer, unaccustomed to the use of computer models, struggled initially to make sense of the interface; alternatively, municipal water engineers in attendance were immediately at home with the interface and directed most of their criticism at the excessive simplicity of the model. Second, the technical limitations of the model meant that only one screen was available at any one time, which meant logistical challenges in keeping stakeholders occupied when not engaging with the game. This also meant that considerably more time was spent on the gaming portion of the day than was planned, and an additional agent design activity was never initiated (although attendees were briefed on the approach).

The third workshop was not completed due to a combination of poor weather and inadequate replies to initial invites from western basin stakeholders. Other researchers in the basin have reported that attendance is a recurring problem at research events, and this was clear from not only the level of attendance at the first two workshops, but also participation rates in other stakeholder engagement efforts. The attendance issue is a largely unacknowledged problem in the collaborative modeling literature, which typically assumes enough stakeholders to make the chosen method work.

### 8.4.4   Delphi survey

This section provides details on the two Delphi survey rounds administered in the summer of 2009. The aim of the first round (questionnaire) was to obtain information on specific modeling validation and design questions which emerged after the early model development phase. Questions in the first round round focused on several objectives, relating to the various components of the Base Behavioral Model:

1. Validating the design of agent plans: the Delphi sought to gather stakeholder opinion on the realism, correctness and completeness of each set of agent plans.

2. Validating the design of agent actions: the Delphi sought to gather stakeholder opinion on the realism of how agent actions were depicted.

3. Validating the choice and design of agent resources: the Delphi sought expert opinion

on the validity of each resource, and appraise the realism of how each resource was described.

4. Designing agent responses to drought, high salinity and economic crisis: three of the major external perturbations likely to afflict the real world system are drought, accumulating levels of salt in river water and valley soils, and economic crisis. The Delphi sought to gather material for designing the space of possible agent responses to each external perturbation.

5. Designing agent responses to a water quality trading market: there has been no water quality trading market in the LAB, and little experience of how agents might respond to the market. The Delphi sought to gather material for designing agent responses to any new market instrument.

The survey sought responses appraising the general accuracy, but not specific details, of each model element. The first round made extensive use of graphical depictions of agent plans and actions. This is in accordance with accepted Delphi, surveying and other model development practice (Linstone & Turoff, 1975; Corben & Wolstenholme, 1994). Only 8 first round responses were received. Summary statistics were generated. Where only a narrow spread was observed, the mean responses were adopted; where the spread was wider, the responses were grouped according to the distribution. Some questions were only incompletely answered, and so performing statistical operations was not always necessary. Some of the free-form responses were discarded for inflammatory comments and irrelevant material.

These results were used in the second round to propose removals, additions and modifications to the existing model. These modifications were packaged in a similar graphical format to the first round, and returned to the stakeholders in the second round of the process. Respondents were asked to rate the proposed modifications. 12 second round responses were received, with improved completion rates. Results from this second round did not warrant review of additional modifications, so the second round was the last in the series. The Delphi folder of the CD contains PDF versions of the online surveys from round 1 and round 2, and stakeholders responses to both rounds.

The Delphi method applied in ArkAgent's development used iterative surveying to help validate specific elements of the model. It represented a much more controlled validation, with only specific elements of the model exposed diagrammatically to stakeholders. Because of its narrow focus, the Delphi surveying did not generate any results not related to specific aspects of model development. With regard to its methodological utility, the Delphi method proved to be effective: the survey addressed a far broader swath of model structure than was possible in the workshops, and allowed the eliciting of stakeholder opinion at a level of detail that had proved difficult to maintain in the individual interviewing. This was partly because of the much greater degree of control I had as survey designer — forcing stakeholders to focus in on specific elements of specific agent plans and actions, and limiting the response options — and the greater information density that the use of diagrams enabled.

### 8.4.5 Psychometric testing

This section provides full details of the three psychometric tests used to parameterize agent S-factors and C-factors (the Base Behavioral Mechanism and Behavioral Mechanism 2).

The first two tests were intended to gather data on S-factors in the population. The first test was based on work by Lusk and others, who formulated a set of questions testing for altruism and propensity for free riding (Lusk et al., 2007). Previous research (Rushton et al., 1981; Webb et al., 2000) suggests that such tests can be good predictors of "pro-social orientation, completion of organ donor cards, actual contributions to charities, and peer ratings of altruism" (Lusk et al., 2007, 508). Accepted procedure in psychometric test development is to generate initial scales, pre-test these scales, and then refine to a final scale (Churchill, 1979; Nunnaly & Bernstein, 1994). Logistical limitations meant that these procedures could not be followed for this present research. Since the Lusk scale has been pre-tested, and applied with relative success, it was deemed sufficiently valid for present use. Test takers were asked to indicate their agreement (on a Likert scale) with the 5 statements described in Figure 8.1:

Indicate how strongly you agree or disagree with the following statements:

1. I am willing to make sacrifices for the good of those around me

2. I enjoy contributing to charities and other non-profit organizations

3. Paying taxes is important because they fund programs such as schools and roads from which everyone benefits

4. I am comfortable taking benefits even if I don't contribute

5. My personal happiness is more important than the well-being of the average American

Figure 8.1: 5 statements measuring pro-social orientation, based on the Lusk scale

Lusk and others report that the objective of this formulation was to gather information on the extent to which a respondent receives utility from the utility of others (their level of altruism), and the obverse, the extent to which a respondent pursues his or her own self interest (their propensity to free ride). Respondent altruism is determined by summing the Likert values for items 1 to 3, then normalizing the response to a mean of 0. Respondent free-riding is determined by summing the Likert values for items 4 and 5, again normalizing to a mean of 0.

The second test was based on the New Ecological Paradigm (NEP). The New *Environmental* Paradigm was introduced in 1978 by Dunlap and Van Liere (Dunlap & van Liere, 2008), and has since become one of the most widely used measures of proenvironmental orientation (Arcury & Christianson, 1990; Dunlap, 2008; Ji, 2004; Lalonde & Jackson, 2002; Schultz et al., 2000; Wiidegren, 1998). A modified version was proposed in 2000, the "New Ecological Paradigm" (Dunlap et al., 2000). The NEP measures "degree of endorsement (from low to high) of an ecological worldview" (Dunlap, 2008, 9), by asking test takers to record their agreement with 15 statements. 3 statements each are devoted to the underlying concepts, which have all proved resilient in studies of whether they actually all measured the same construct: "the reality of limits to growth... antianthropocentrism... the fragility of nature's balance... rejection of exemptionalism... and the possibility of an ecocrisis" (Dunlap et al., 2000, 431). For present pur-

poses, however, one statement from each of these concepts has been removed, to facilitate a more concise survey. The final set of statements is described in Figure 8.2.

Indicate how strongly you agree or disagree with the following statements:

1. We are approaching the limit of the number of people the earth can support
2. Humans have the right to modify the natural environment to suit their needs
3. When humans interfere with nature it often produces disastrous consequences
4. Human ingenuity will insure that we do NOT make the earth unlivable
5. Humans are severely abusing the environment
6. The earth has plenty of natural resources if we just learn how to develop them
7. Plants and animals have as much right as humans to exist
8. The balance of nature is strong enough to cope with the impacts of modern industrial nations
9. Despite our special abilities humans are still subject to the laws of nature
10. The so-called "ecological crisis" facing humankind has been greatly exaggerated
11. The earth is like a spaceship with very limited room and resources
12. Humans were meant to rule over the rest of nature
13. The balance of nature is very delicate and easily upset
14. Humans will eventually learn enough about how nature works to be able to control it
15. If things continue on their present course, we will soon experience a major ecological catastrophe

Agreement/disagreement options:

*Strongly agree - mildly agree - unsure - mildly disagree - strongly disagree*

Figure 8.2: 15 statements making up the New Ecological Paradigm scale, after Dunlap et al., 2000

Following Dunlap et al. (2000), agreement with the eight odd-numbered statements, and disagreement with the seven even-numbered statements, indicates a pro-NEP response. To make the results compatible with the Lusk altruism test, responses are normalized to a mean of 0.

These two tests were adopted to identify the proportion of stakeholders who were both altruistic and proenvironmental. Without doing specific testing for both altruism and proenvironmental orientation, an assumption is made that all altruistic stakeholders will also behave altruistically on environmental questions, which is not necessarily the case. Specific testing allowed the identification of stakeholders who may have been altruistic but did not have proenvironmental orientation, as well as those stakeholders who were not altruistic but were proenvironmental. The former case translated into a lower S-factor, since altruism in ArkAgent is strongly identified with proenvironmental behavior. The latter case translated into a higher S-factor, on the assumption that a proenvironmental orientation is correlated with altruistic be-

havior in environmental contexts, but such behavior is not necessarily motivated by altruism towards society at large. For the purposes of ArkAgent, the actual motivation for altruism is less important than its display.

To calculate the final S-factor, which is a 0.0-1.0 scale, the two components were summed, and then normalized to a mean of 0.5. The S-factor score is consequently designed to indicate pro-sustainability behavior when the score is greater than 0.5, and anti-sustainability behavior when the score is less than 0.5. A score of 0.5 indicates equal weight between pro-sustainability and anti-sustainability behavior.

The third test was intended to gather data on the propensity towards attitudinal (or fundamental behavioral) change within the population - the "C-factor". The questions were based on a survey structure first used by Waddell and Sohal (Waddell & Sohal, 1999). Following most empirical studies, the survey structure adopted here focuses on resistance to change, inferring propensity to change as the reciprocal of the resistance to change rating. Waddell and Sohal's approach is to use questions addressing specific components of resistance to change. Semantic differential responses to each question (Friborg et al., 2006; Osgood, 1957) not only provide information on the level of resistance to change in the respondent, but also the respondent's attitude to change (positive or negative). This accounts for the cases where someone with negative attitudes towards change may not actually be resistant to change, if they view change as inevitable; for the cases where someone with positive attitudes towards change may actually resist change; and for the cases where the individual is apathetic and so neither supports nor opposes change which is imposed on them. Waddell and Sohal conduct a detailed study of the components of resistance to change, yet not all of these items are relevant to the LAB context. Consequently, a modified survey structure was adopted, shortened and contextualized relative to the original Waddell and Sohal survey. A pre-test was used to shorten the original 12 questions to a final 6, tailored to the LAB context. The survey structure is described in Figure 8.3.

Each question relates to a particular facet of resistance to change:

- 1 = fear of change

- 2 = disruptive effects of change

- 3 = broader culture of change in the community

- 4 = acceptability of change

- 5 = negative associations with change

- 6 = apathy regarding change

Positive or negative weightings for change were assigned to each bipolar scale. For example, for question 2, "easier" equates to a positive C-factor contribution, while "harder" denotes a negative C-factor contribution. Each question contributed a 0.0-1.0 component of change to an equal component of the final C-factor scale (6 components in total).

Response rates for all the tests, administered online, were 11 out of 60 invitees (65 original invitees, 5 failed e-mail addresses). Invitees were reminded twice by e-mail during the 4-8 week period each survey was open. Response rates were therefore approximately 18%. These rank as poor response rates compared with the average for online surveys (Sheehan, 2001), but good compared with past research experience in the lower Arkansas Basin (pers. comms.

> People are always discussing options for changing the way water
> is managed in the basin. The following questions are intended to
> explore how you feel about such change in general.
>
> 1. Describe your feelings about the likelihood of any change happening
> to the way water is managed in the LAB. [*Apprehensive - Confident*]*
>
> 2. What effect do you think such change would have on how easy or
> difficult it is to do your job? [*Easier - Harder*]
>
> 3. How do your colleagues/acquaintances respond to the possibility of
> change? [*Reject - Accept*]
>
> 4. Do you think that change in the valley's water management is
> inevitable? [*Never - Always*]
>
> 5. Generally, what is change? [*Bad - Good*]
>
> 6. Are you indifferent to all types of change? [*Always - Never*]
>
> *Semantic differential scale runs on units 1-7, with labels only on points 1 and 7

Figure 8.3: 6 questions making up the C-factor survey, based on Waddell and Sohal's resistance to change survey

Smith 2009, Gates 2009). The resulting 95% confidence interval for the S-factor and C-factor, assuming a sample size of 60 and a population size of 663, is +/-7.2%. This is an acceptable level of error for generating empirical distributions for parameterizing individual agents, but developing a distribution with such a small number of results means that distribution tails are likely to be underrepresented. Normal distributions (with slight skew) for both C-factor and S-factor values were produced with this assumption.

Psychometric tests are rarely used in conjunction with collaborative modeling, although they are often used in simulation modeling more generally. The S-factor and C-factor values were arcane concepts mostly related to theoretical elements of my study, and not appropriate subjects to focus on in a collaborative session. Psychometric testing consequently proved to be an effective way to address elements of model parameterization which were too complex or too dull to tackle in a collaborative session.

The original online survey and results are included on the CD in the Psychometrics folder.

## 8.5   Obstacles to progress for a real world WQTM

In section 7.3 of chapter 7, I offered an idealized vision for change in the Lower Arkansas Basin, that would address social, economic and environmental problems generated by high levels of basin salinity. In this present section, I offer a more detailed discussion of selected potential

major obstacles that this idealized process might encounter. None of these are derived from model results, and are offered as educated speculation on the basis of experience and a reading of the relevant literature. It is not an exhaustive list.

The continued inferiority of water quality regulation in the eyes of Colorado water law would likely hamper a major water quality initiative like the proposed WQTM. As discussed in the literature review, Colorado water quality law is a latecomer, separate from and subordinate to the prior appropriation framework. Even if a WQTM poses no theoretical threat to existing diversion patterns, there would still likely be stiff opposition from the legal community, and from individual irrigators who suspected that damage would be done to their decreed rights. The current legal framework resists change, penalizes creative thinking and insists on various hydrologically-blind provisions that create complexities and inefficiencies in water allocation. The best legal climate for the WQTM is one in which water quality and quantity are finally linked in a coherent way, but this could be a significant legal battle.

The state of Colorado — in the form of its water courts and administrative agencies — could generate transaction cost overheads that might affect the success of any proposed WQTM. Ark-Agent made the simplifying assumption that transaction costs were very small for the purpose of hypothesis testing, but this is clearly untenable for any real world institution: transaction costs in the Colorado water courts can be exceptionally high. It is not known, of course, whether such transaction costs would translate into a proposed WQTM, or if in fact these costs would prove to be a problem (in light of the foregoing discussion, we should not make assumptions in this regard).

Any real market's success could be greatly affected by farmer and ditch interest in honest participation. I used ArkAgent to show that uncertainty need not necessarily greatly affect simulation outcomes, but I restricted my discussion to uncertainty associated with the determination of actual salt loads. There are additional and considerable uncertainties associated with whether or not farmers would break the rules, whether or not monitoring would catch any such infractions, and whether enforcement would be capable of providing disincentives for such behavior. Assuming that farmers and ditches break the rules (which, again, we should not assume), much of the market's success could hinge on effective monitoring and enforcement.

Climate is an important macro-scale influence on the success of the market, and as was seen in a WQTM in California, a series of exceptionally wet years could offset any salinity reduction from earlier years (Pharino, 2007). Other macro-scale influences include the national health of the agricultural industry, since trends on a national scale can have important effects on basin economic conditions. A period of low agricultural prices and market surplus in key basin commodities would likely slow trade volume and frequency irregardless of any change in quota allowances; it might also make it politically unpalatable to levy fines for exceeding quota.

Finally, continued lack of guidance on the use of TMDLs by Congress and the EPA (Andreen, 2007) could generate political resistance. Since the LAB WQTM as proposed relies on a version of a TMDL (an annual load cap), there is the possibility that state agencies would resist implementation in the absence of a clearer federal stance on this issue.
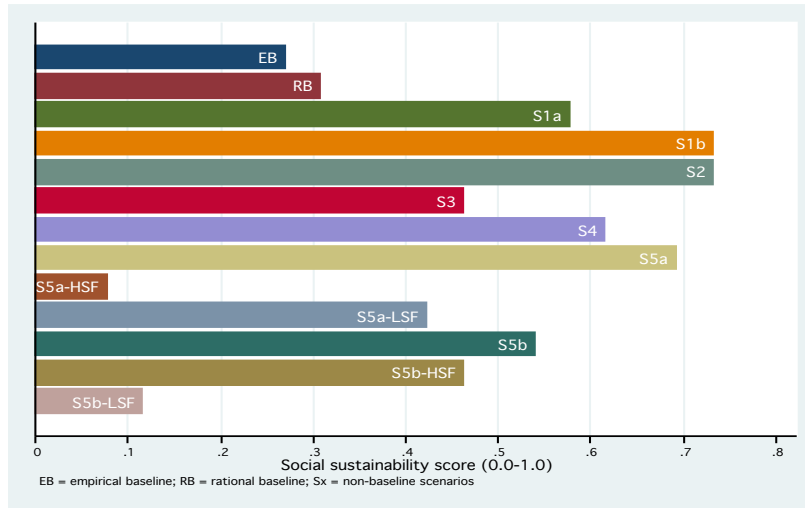
Figure 8.4: Social sustainability scores by scenario.

## 8.6 ArkAgent: installation and operation guide

The ArkAgent installation and operation guide can be found in the Model folder on the CD accompanying this thesis.
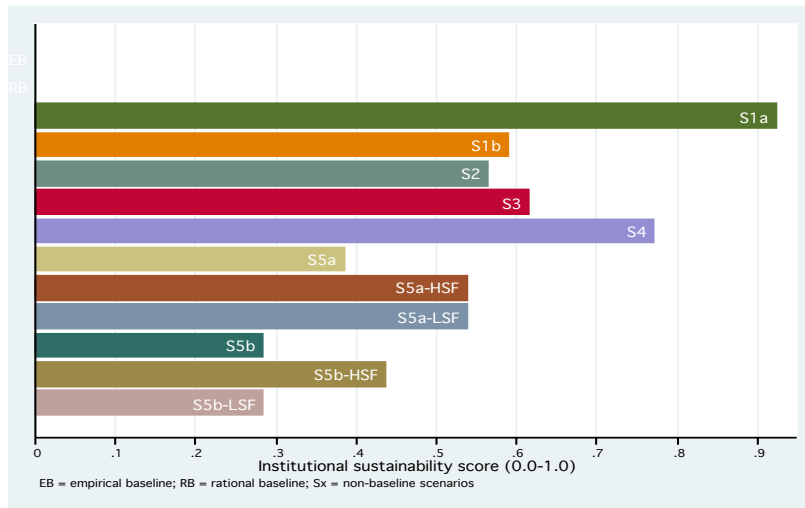
## 8.7 Additional plots

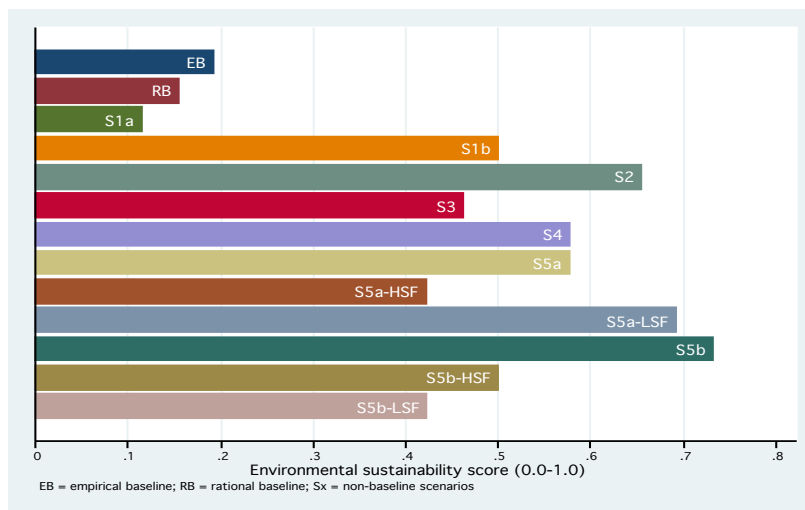Figure 8.5: Institutional sustainability scores by scenario.



Figure 8.6: Environmental sustainability scores by scenario
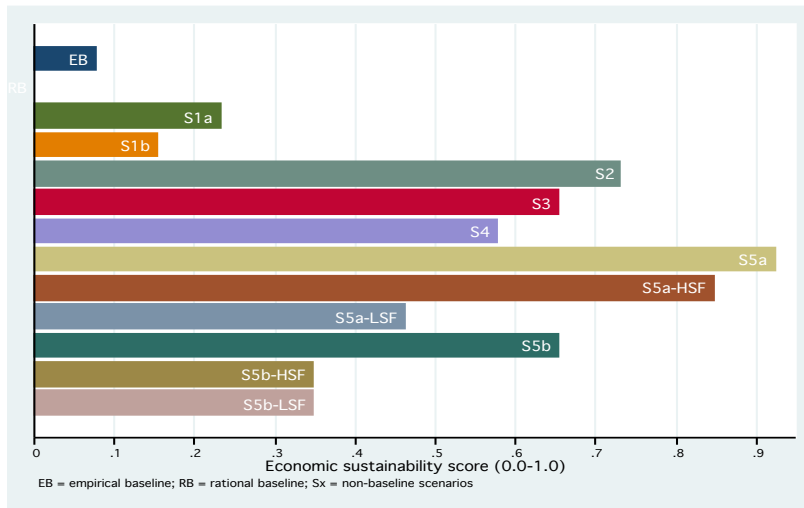
Figure 8.7: Economic sustainability scores by scenario
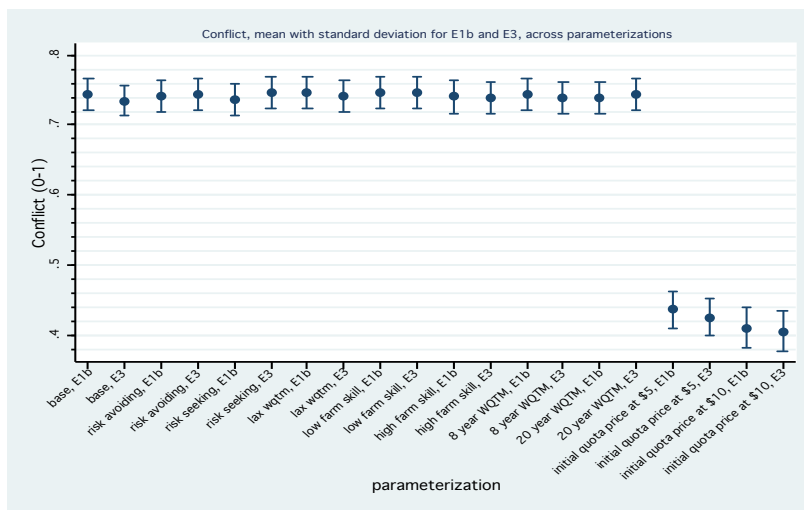


Figure 8.8: Conflict, parameter variation, comparison between scenarios 1b and 3
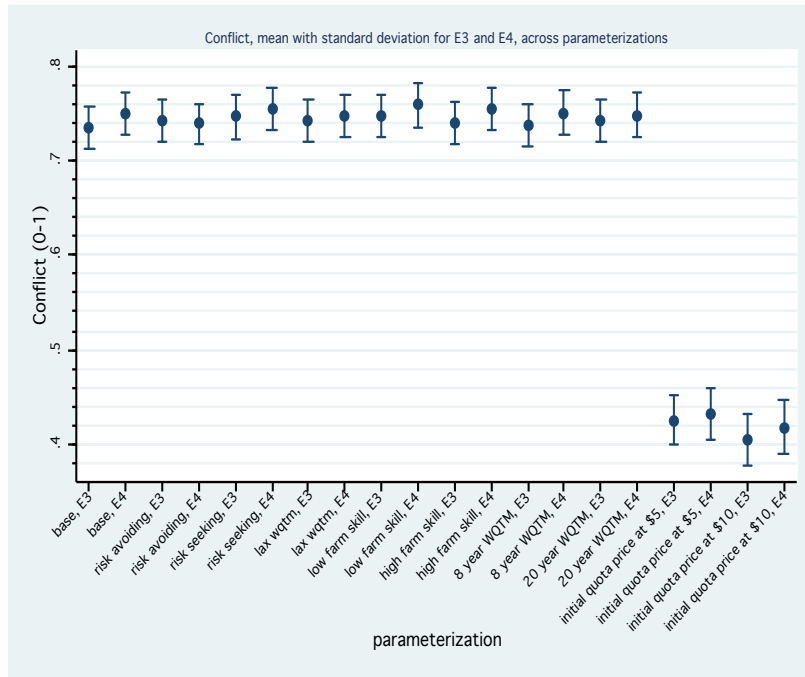
Figure 8.9: Conflict, parameter variation, comparison between scenarios 3 and 4
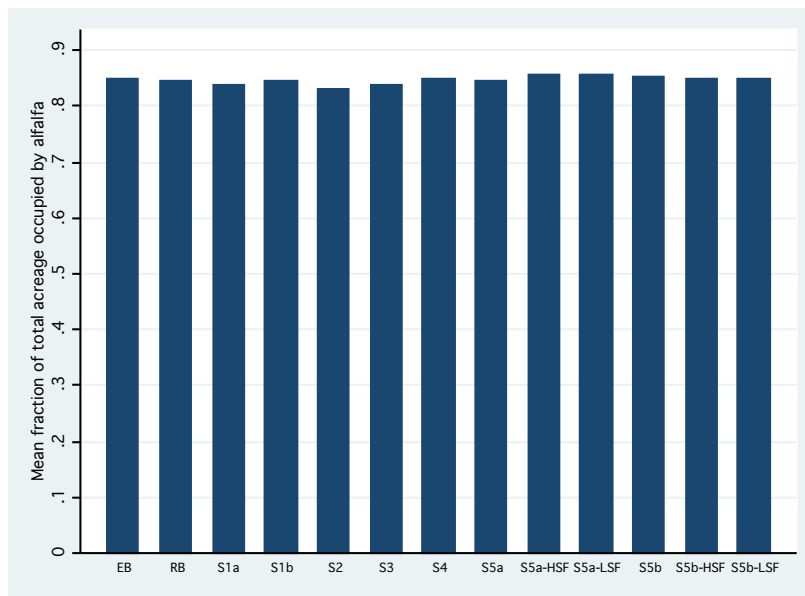


Figure 8.10: Fraction of total acreage planted with alfalfa (0.0-1.0), across all scenarios. Plot shows that none of the various behavioral and institutional assumptions greatly change the dominance of alfalfa as the crop of choice for most farmers in the LAB. This is to be expected, since alfalfa is the crop that has dominated planting in the basin for many years, and remains the most reliable choice for most farmers. It is also a multi-year crop, and so farmers have less flexibility in shifting acres out of alfalfa. Refer to Appendix section 8.1.2 for detailed scenario descriptions.
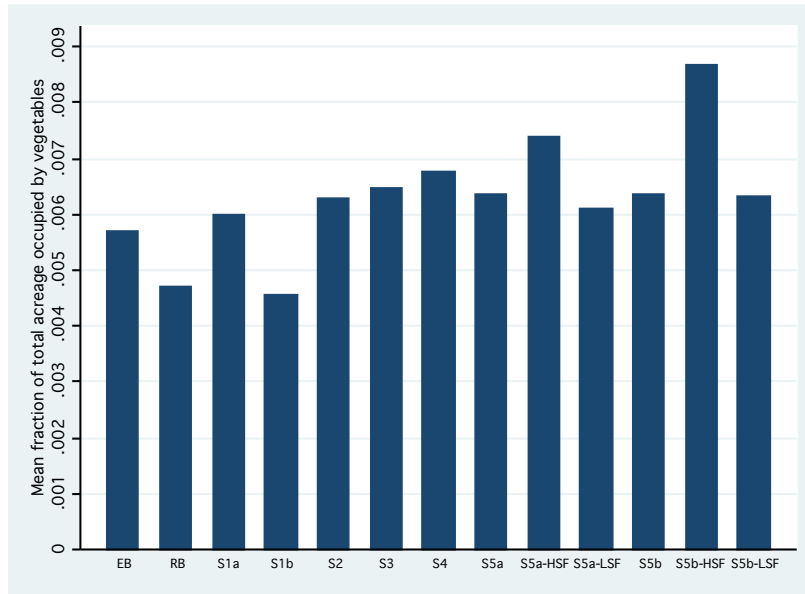
Figure 8.11: Fraction of total acreage planted with vegetables (0.0-1.0). Note the weak trend towards higher vegetable acreages in scenarios 1b through 4; in other words, additional behavioral and institutional complexity appears to drive farmers towards more water intensive, smaller acreage crop types. Refer to Appendix section 8.1.2 for detailed scenario descriptions.
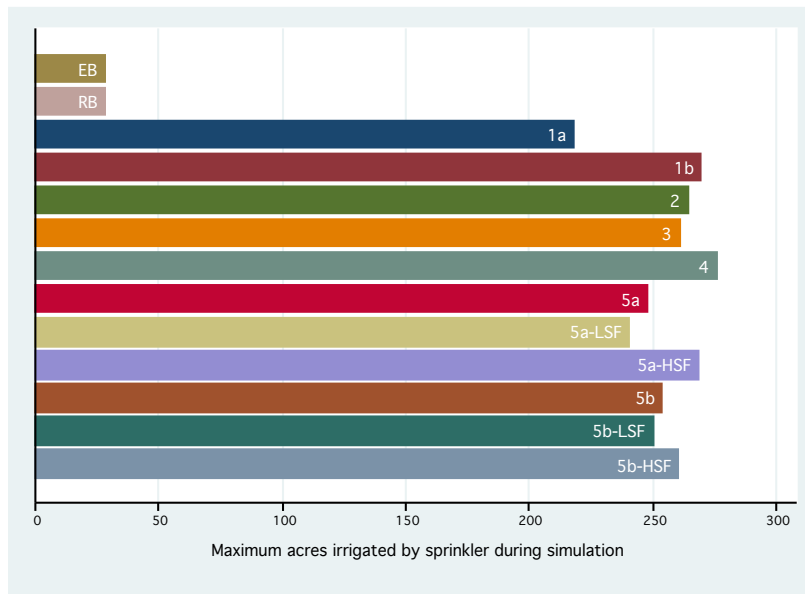


Figure 8.12: Maximum area of land irrigated by center-pivot sprinkler technology during the simulation period. Max value is plotted because ArkAgent does not simulate downgrading of technology, and so this value represents an overall measure of upgrade levels for the simulation. Plot indicates that much of the shift away from the baseline case (which is dominated by flood irrigation) is taken up by center-pivot sprinkler upgrades. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.
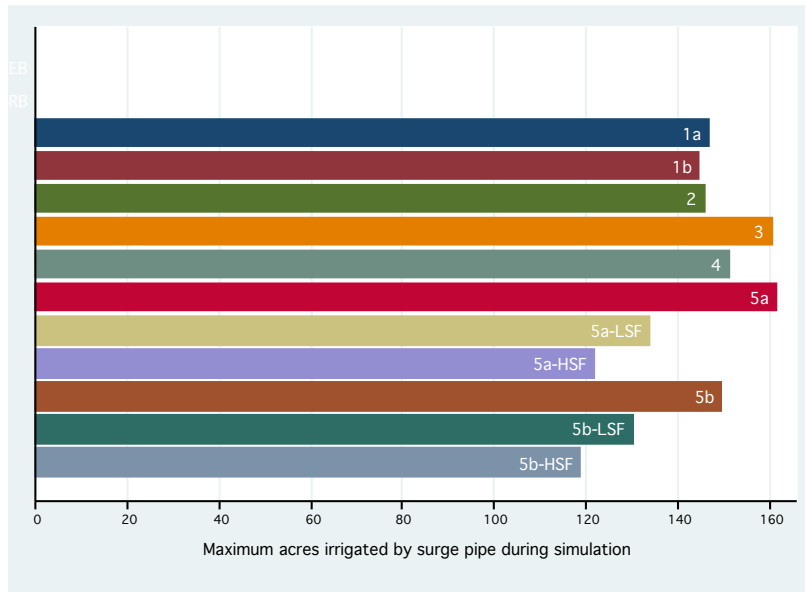
Figure 8.13: Maximum area of land irrigated by surge pipe technology during the simulation period. Plot indicates how, as with sprinkler upgrades, most of the change happens when we move from the baseline to an institutional scenario (i.e. scenarios EB and RB to 1a and above). See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.
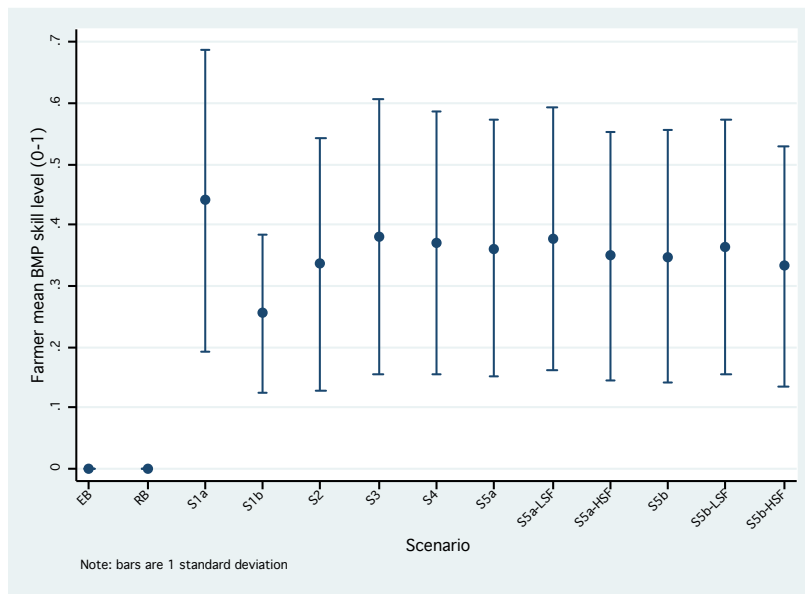


Figure 8.14: Mean weekly farmer Best Management Practice (BMP) skill level (0.0-1.0). Note how the most economically rational farmers end up the best mean level of education on salinity. This is to be expected, since highly economically rational decision makers will plump for the great cost-benefit ratio of salinity management education. Refer to Appendix section 8.1.2 for detailed scenario descriptions.
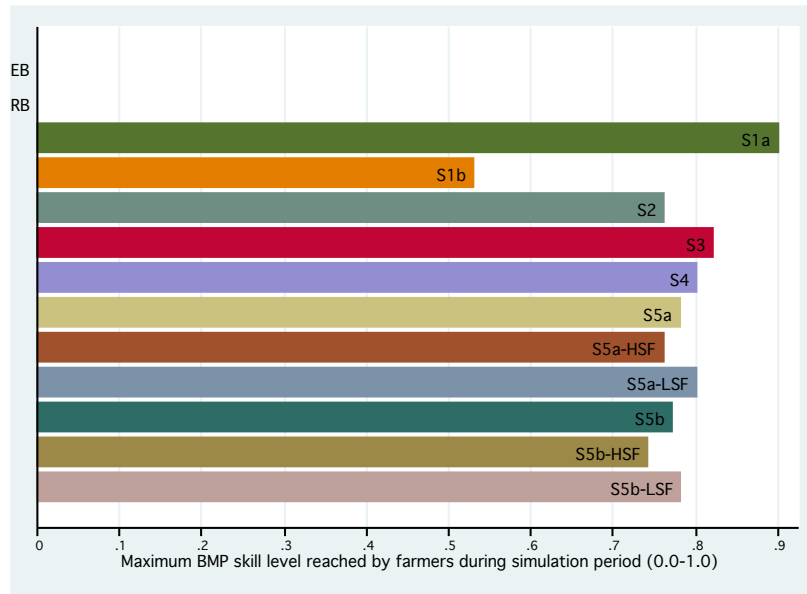
Figure 8.15: Maximum level of Best Management Practice (BMP) skill achieved by farmers (0.0-1.0). All institutional interventions drive up skill levels relative to the baseline. Since ArkAgent does not simulate the degradation of a skill level once it is reached, this figure provides a good indication of the overall relation between BMP and scenario. Results are broadly supportive of the main thrust of the thesis: non-economic incentives can mitigate otherwise poor performance by a market institution with an attitudinally heterogeneous population. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.
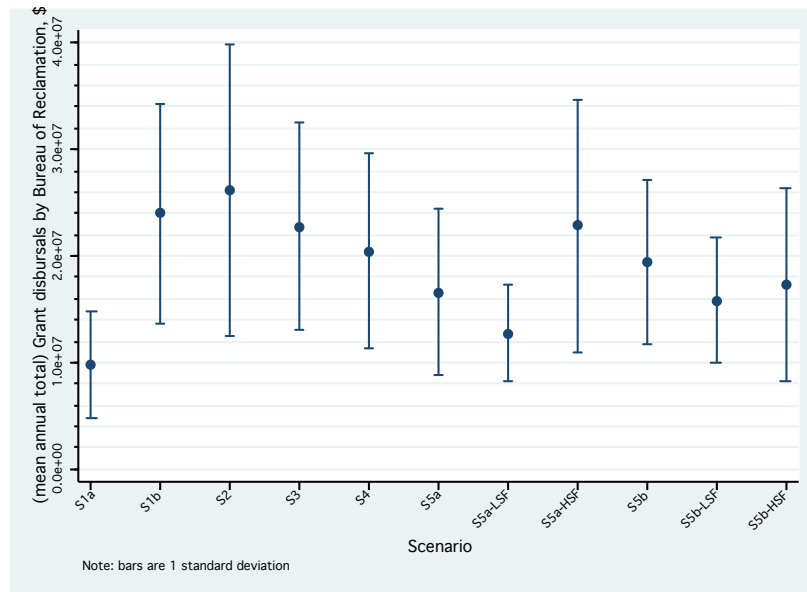


Figure 8.16: Mean annual total grants disbursed by the Bureau of Reclamation ($). Grants support irrigation equipment upgrades and ditch maintenance and improvement activities. See text for discussion and refer to Appendix section 8.1.2 for detailed scenario descriptions.
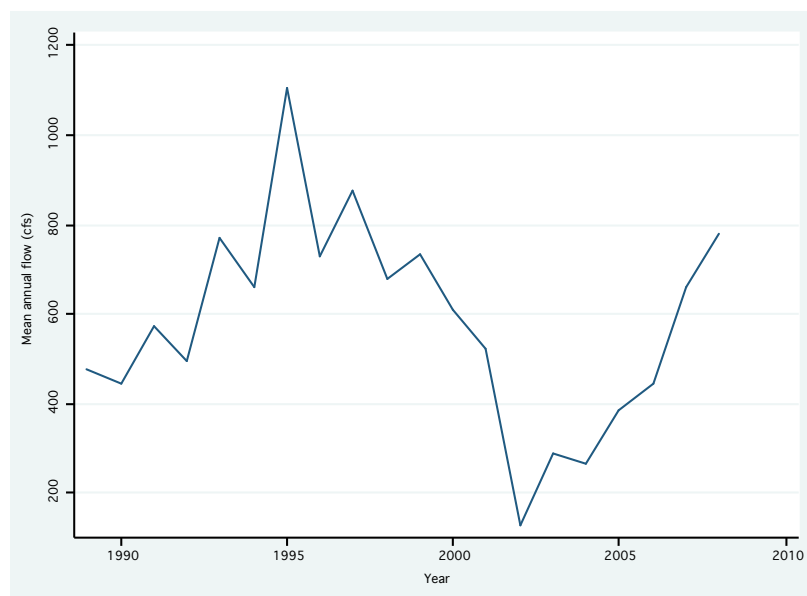
Figure 8.17: Historic flow (acre-feet, recorded quarterly) at the Pueblo Dam gauge. Since this gauge is upstream of all the ditches, and indicates a low flow period during the same span of time as the simulation run, this supports my assertion that flow dynamics unrelated to human activity provide a macroscale control on salt loads.

# Bibliography

W. Abrahamse, et al. (2005). 'A review of intervention studies aimed at household energy consumption'. *Journal of Environmental Psychology* **25**:273–291.

W. Adams, et al. (2003). 'Managing Tragedies: Understanding Conflict over Common Pool Resources'. *Science* **302**(5652):1915–1916.

W. Adger (2003). 'Social Capital, Collective Action, and Adaptation to Climate Change'. *Economic Geography* **79**(4):387–404.

W. Akpalu (2008). 'Fishing regulations, individual discount rate, and fisherman behavior in a developing country fishery'. *Environment and Development Economics* **13**:591–606.

R. Allen & W. Pruitt (1986). 'Rational Use of the FAO Blaney-Criddle Formula'. *Journal of Irrigation and Drainage Engineering* **112**(2):139–156.

M. Altman (2006). *Handbook of Contemporary Behavioral Economics*. ME Sharpe.

H. Amilon (2008). 'Estimation of an adaptive stock market model with heterogeneous agents'. *Journal of Empirical Finance* **15**:342–362.

L. An, et al. (2005). 'Exploring Complexity in a Human-Environmental System: An Agent-Based Spatial Model for Multidisciplinary and Multiscale Integration'. *Annals of the Association of American Geographers* **95**(1):54–79.

J. Anderies, et al. (2006). 'Loss of Resilience, Crisis, and Institutional Change: Lessons from an Intensive Agricultural System in Southeastern Australia'. *Ecosystems* **9**:865–878.

T. Anderson & P. Hill (1975). 'The Evolution of Property Rights: A Study of the American West'. *Journal of Law and Economics* **18**(1):163–179.

W. Andreen (2007). 'Motivating Enforcement: Insitutional Culture and Clean Water Act'. *Pace Environmental Law Reviews* **24**:67.

P. Antunes, et al. (2006). 'Participatory decision making for sustainable development - the use of mediated modelling techniques'. *Land Use Policy* **23**:44–52.

T. Arcury & E. Christianson (1990). 'Environmental worldview in response to environmental problems: Kentucky 1984 and 1998 compared'. *Environment and Behavior* **22**:387–407.

L. Arrington (1984). 'Rural Life among Nineteenth-Century Mormons: The Woman's Experience'. *Agricultural History* **58**(3):239–246.

L. Arrington & D. May (1975). 'A Different Mode of Life: Irrigation and Society in Nineteenth-Century Utah'. *Agricultural History* **49**(1):3–20.

S. Atkinson & D. Lewis (1974). 'A Cost-Effectiveness Analysis of Alternative Air Quality Control Strategies'. *Journal of Environmental Economics and Management* **1**:237–250.

R. Aull-Hyde & S. Tadesse (1994). 'A Strategic Agricultural Production Model with Risk and Return Considerations'. *Agricultural and Resource Economics Review* **4**(8):37–46.

R. Axelrod (1997). *Simulating Social Phenomena*, chap. Advancing the Art of Simulation in the Social Sciences. Springer.

R. Axtell (2000). 'Why Agents? On the varied motivations for agent computing in the social sciences'. Tech. rep., The Brookings Institution.

J. Baden (1998). *Managing the Commons*, chap. A New Primer for the Management of Common-Pool Resources and Public Goods. Indiana University Press.

J. Baland & J. Platteau (1996). *Halting Degradation of Natural Resources: Is There a Role for Rural Communities?* Clarendon Press.

O. Balci (1998). *Handbook of Simulation: Principles, Methodology, Advances, Applications and Practice*. John Wiley and Sons.

A. V. D. Ban (1998). 'Supporting Farmers, Decision Making Processes by Agricultural Extension'. *Journal of Extension Systems* pp. 55–68.

A. Bandura (1971). *The Nature of Reinforcement*, chap. Vicarious and Self-Reinforcement Processes. Academic Press.

B. Barber & T. Odean (2008). 'All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors'. *The Review of Financial Studies* **21**(2).

E. Barbier (2004). 'Water and Economic Growth'. *Economic Record* **80**.

P. Bardhan & J. Dayton-Johnson (2002). *The Drama of the Commons*, chap. Unequal Irrigators: Heterogeneity and Commons Management in Large-Scale Multivariate Research. National Academy Press.

Y. Barlas (1996). 'Formal aspects of model validity and validation in system dynamics'. *System Dynamics Review* **12**(3):183–210.

T. Barnett, et al. (2005). 'Potential impacts of a warming climate on water availability in snow-dominated regions'. *Nature* **438**(17):303–309.

O. Barreteau (2003). 'Our Companion Modeling Approach'. *Journal of Artificial Societies and Social Simulation* **6**(1).

O. Barreteau & F. Bousquet (2000). 'SHADOC: a multi-agent model to tackle viability of irrigated systems'. *Annals of Operations Research* **94**:139–162.

O. Barreteau, et al. (2003). 'Agent-Based Facilitation of Water Allocation: Case Study in the Drome River Valley'. *Group Decision and Negotiation* **12**:441–461.

O. Barreteau, et al. (2007). 'Contribution of simulation and gaming to natural resource management issues: An introduction'. *Simulation and Gaming* **38**(2):185–194.

C. Barrow (1998). 'River basin development planning and management: A critical review'. *World Development* **26**(1):171–186.

R. Barthel, et al. (2005). 'Linking the physical and the socio-economic compartments of an integrated water and land use management model on a river Linking the physical and the socio-economic compartments of an integrated water and land use management model on a river basin scale using an object-oriented water supply model'. *Physics and Chemistry of the Earth* **30**:389–397.

P. Bartlett (1993). *American dreams, rural realities*. UNC Press.

M. Bazerman (1997). *Environment, ethics and behavior: the psychology of environmental valuation*. Lexington Books.

N. Becu, et al. (2002). 'How bad isn't the agent-based model CATCHSCAPE?'. *Proceedings of the 2002 IEMSS Meeting* .

S. Beder (2001). 'Trading the Earth: The Politics behind Tradable Pollution Rights'. *Environmental Liability* **9**(2):152–160.

J. Beedell & T. Rehman (2000). 'Using social-psychology models to understand farmers' conservation behavior'. *Journal of Rural Studies* **16**(1):117–127.

A. Belsky, et al. (1999). 'Survey of Livestock Influences on Stream and Riparian Ecosystems in the Western United States'. *Journal of Soil and Water Conservation* **54**:419–431.

M. V. D. Belt (2004). *Mediated Modeling: a system dynamics approach to environmental consensus building*. Island Press.

R. Benabou & J. Tirole (2006). 'Incentives and prosocial behavior'. *American Economic Review* **96**(5):1652–1679.

I. Benenson & P. Torrens (2004). 'Geosimulation: object-based modeling of urban phenomena'. *Computers, Environment and Urban Systems* **28**:1–8.

J. Bennett (1969). *Northern Plainsmen: Adaptive Strategy and Agrarian Life*. Aldine.

T. Berger (2001). 'Agent-based spatial models applied to agriculture: a simulation tool for technology diffusion, resource use changes and policy analysis'. *Agricultural Economics* **25**(2):245–260.

M. Binford, et al. (1997). 'Climate Variation and the Rise and Fall of an Andean Civilization'. *Quaternary Research* **47**:235–248.

W. Blomquist (2006). 'Crafting Water Constitutions in California'. *Working Paper* .

W. Blomquist, et al. (2004a). 'Building the agenda for institutional research in water resources management'. *Journal of the American Water Resources Association* pp. 925–937.

W. Blomquist, et al. (2004b). *Common Waters, Diverging Streams: Linking Institutions and Water Management in Arizona, California, ....* Washington D.C.: RFF Press.

I. Bohnet (2007). 'Institutions and Trust'. *Rationality and Society* **19**(1):99–135.

S. Boissau & J. Castella (2003). 'Constructing a Common Representation of Local Institutions and Land Use Systems through Simulation- ...'. *Simulation and Gaming* **34**(3):342–357.

B. Bolin, et al. (2008). 'Fate of the Verde: Water, environmental conflict, and the politics of scale in Arizona's Central Highlands'. *Geoforum* **39**(1494-1511).

E. Bonabeau, et al. (1999). *Swarm intelligence: from natural to artifical systems*. Oxford University Press.

J. Booker, et al. (2005). 'Economic impact of alternative policy responses to prolonged and severe drought in the Rio Grande Basin'. *Water Resource Research* **41**:1–15.

S. Born & W. Sonzogni (1995). 'Integrated environmental management: strengthening the conceptualization'. *Environmental Management* **19**(2):167–181.

C. Boronkay & W. Abbott (1997). 'Water Conflicts in the Western United States'. *Studies in Conflict and Terrorism* **20**:137–166.

H. Bouwer (2002). 'Integrated Water Management for the 21st Century: Problems and Solutions'. *Journal of Irrigation and Drainage Engineering* **128**(4):193–203.

S. Bowles (2008). 'Policies Designed for Self-Interested Citizens May Undermine "The Moral Sentiments": Evidence from Economic Experiments'. *Science* **320**:1605–1609.

S. Bowles & H. Gintis (2002). 'Social Capital and Community Governance'. *The Economic Journal* **112**(483):419–436.

S. Bowles & H. Gintis (2004). 'The evolution of strong reciprocity: cooperation in heterogeneous populations'. *Theoretical Population Biology* **65**:17–28.

H. Breetz, et al. (2004). 'Water Quality Trading and Offset Initiatives in the U.S.: A Comprehensive Survey'. *Unpublished Manuscript by Dartmouth College* .

S. Bretsen & P. Hill (2006). 'Irrigation Institutions in the American West'. *Economic History Association* .

H. Briassoulis (2001). 'Sustainable Development and its Indicators: Through a (Planner's) Glass Darkly'. *Journal of Environmental Planning and Management* **44**(3):409–427.

J. Broome (2007). 'Is rationality normative?'. *Disputatio* **2**(3):162–178.

D. Brown & D. Robinson (2006). 'Effects of Heterogeneity in Residential Preferences on an Agent-Based Model of Urban Sprawl'. *Ecology and Society* **11**(1):46.

S. Brown (1979). 'Maximum Utilization Collides with Prior Appropriation in A-B Cattle Co. v. United States'. *Denver Law Journal* **57**:103.

T. Brown (2006). 'Trends in water market activity and price in the western United States'. *Water Resources Research* **42**:1–14.

M. Browne & R. Cudeck (1993). *Testing Structural Equation Models*, chap. Alternative Ways of Assessing Model Fit. Sage.

G. Brundtland (1987). 'Our Common Future'. Tech. rep., World Commission on Environment and Development.

L. Bruni & R. Sugden (2008). 'Fraternity: Why the market need not be a morally free zone'. *Economics and Philosophy* **24**:35–64.

J. Burkhalter & T. Gates (2005). 'Agroecological Impacts from Salinization and Waterlogging in an Irrigated River Valley'. *Journal of Irrigation and Drainage Engineering* pp. 197–210.

R. Burton, et al. (2008). 'Exploring Farmers' Cultural Resistance to Voluntary Agri-environmental schemes'. *Sociologia Ruralis* **48**(1):16–38.

D. Burtraw, et al. (2005). 'Economics of Pollution Trading for SO2 and NOx'. *Annual Review of Environment and Resources* **30**:253–289.

D. Burtraw, et al. (1997). 'The Costs and Benefits of Reducing Acid Rain'. *Resources for the Future Discussion Papers* .

D. Burtraw & E. Mansur (1999). 'Environmental Effects of SO2 Trading and Banking'. *Environmental Science and Technology* **33**:3489–3494.

C. Camerer & E. Fehr (2006). 'When does "Economic Man" Dominate Social Behavior?'. *Science* **311**:47–52.

C. Camerer, et al. (2005). 'Neuroeconomics: how neuroscience can inform economics'. *Journal of Economic Literature* **43**:9–64.

W. V. Campen (1992). *Agriculture, economics, and sustainability in the Sahel*, chap. Incentives for sustainable land use: some general aspects of constraints on design, implementation and financing. With special reference to southern Mali. Royal Tropical Institute.

A. Carlson (2005). 'Social Norms and Individual Environmental Behavior'. *Environmental Law Reports* **35**.

S. Carpenter, et al. (1999). 'Ecological and Social Dynamics in Simple Models of Ecosystem Management'. *Ecology and Society* **3**(2):4.

S. Carpenter, et al. (2001). 'From Metaphor to Measurement: Resilience of What to What?'. *Ecosystems* **4**:765–781.

F. Casey & G. Lynne (1999). *Flexible Incentives for the Adoption of Environmental Technologies in Agriculture*, chap. Adoption of Water Conserving Technologies in Agriculture: The Role of Expected Profits and the Public Interest. Kluwer Academic.

T. Cason & C. Plott (1996). 'EPA's New Emissions Trading Mechanism: A Laboratory Evaluation'. *Journal of Environmental Economics and Management* **30**:133–160.

CDPHE (2002). 'Colorado Water Quality Control Act'. Tech. rep., Colorado Department of Public Health and the Environment.

M. Chetty, et al. (2008). 'Getting to Green: Understanding Resource Consumption in the Home'. *Proceedings of the 10th International Conference on Ubiquitous Computing* .

S. Childs & R. Hanks (1975). 'Model of Soil Salinity Effects on Crop Growth'. *Soil Science Association of America Journal* **39**:617–622.

H. Chong & D. Sunding (2006). 'Water Markets and Trading'. *Annual Review of Environment and Resources* **31**:239–264.

C. Chou (2002). 'Developing the e-Delphi system: a web-based forecasting tool for educational research'. *British Journal of Education Technology* **33**(2):233–236.

H. Chouinard, et al. (2008). 'Will Farmers Trade Profits for Stewardship? Heterogeneous Motivations for Farm Practice Selection'. *Land Economics* **84**(1):66–82.

G. Churchill (1979). 'A Paradigm for Developing Better Measures for Marketing Constructs'. *Journal of Marketing Research* **16**:64–73.

R. Cialdini & N. Goldstein (2004). 'Social Influence: Compliance and Conformity'. *Annual Review of Psychology* **55**:591–621.

G. Clore & J. Huntsinger (2007). 'How emotions inform judgment and regulate thought'. *Trends in Cognitive Sciences* **11**(9):393–400.

R. Coase (1960). 'The Problem of Social Cost'. *Journal of Law and Economics* **3**:1–44.

S. Coleman (2004). 'The Effect of Social Conformity on Collective Voting Behavior'. *Political Analysis* **12**:76–96.

G. Colman (1968). 'Innovation and Diffusion in Agriculture'. *Agricultural History* **42**(3):173–188.

F. Convery & L. Redmond (2007). 'Market and Price Developments in the European Union Emissions Trading Scheme'. *Review of Environmental Economics and Policy* **1**(1):88–111.

D. Corben & E. Wolstenholme (1994). 'A hypermedia-based Delphi tool for knowledge acquisition in model building'. *Journal of the Operational Research Society* **45**(6):659–672.

G. Cory (1999). *The Reciprocal Modular Brain in Economics and Politics: Shaping the Rational and Moral Basis of Organization, Exchange and Choice*. Kluwer Academic.

R. Costanza & L. Wainger (1991). *Ecological Economics: The Science and Management of Sustainability*. Columbia University Press.

R. Costanza, et al. (1993). 'Modeling Complex Ecological Economic Systems'. *Bioscience* pp. 545–557.

J. Coval & T. Shumway (2005). 'Do Behavioral Biases Affect Prices?'. *The Journal of Finance* **60**(1):1–35.

T. Cox, et al. (1991). 'Effects of Ethnic Group Cultural Differences on Cooperative and Competitive Behavior on a Group Task'. *Academy of Management Journal* **34**(4):824–847.

P. Cramton & S. Kerr (2002). 'Tradeable carbon permit auctions: How and why to auction not grandfather'. *Energy Policy* **30**:333–345.

J. Cross & M. Guyer (1980). *Social Traps*. University of Michigan Press.

H. Crumpler & P. Grossman (2008). 'An experimental test of warm glow giving'. *Journal of Public Economics* **92**:1011–1021.

D. Culver, et al. (2007). 'Biological Survey of Bent's Old Fort National Historic Site, Otero Country, Colorado.'. Colorado natural heritage program, Colorado State University.

L. Cutforth, et al. (2001). 'Factors Affecting Farmers' Crop Diversity Decisions: An Integrated Approach'. *American Journal of Alternative Agriculture* **16**(4):168–176.

C. Dake, et al. (2006). 'Specification of Environmental Emission Trading Options in a Spatial Multi-Agent Simulation Model of Pastoral Farming'. In *Proceedings of the New Zealand Agricultural and Resource Economics Society Conference 2006*.

H. Daly & J. Farley (2003). *Ecological Economics*. Island Press.

S. Darby (2006). 'The effectiveness of feedback on energy consumption'. *Environmental Change Institute Reviews* .

W. Dare & O. Barreteau (2003). 'A role-playing game in irrigated system negotiation: between play and reality'. *Journal of Artificial Societies and Social Simulation* **6**(3):6.

B. de Mandeville (1924). *The Fable of the Beeds, or Private Vices, Publick Benefits*. Clarendon Press.

P. Dead & E. Schlager (2002). *Integrating Geographic Information Systems and Agent-Based Modeling Techniques for Understanding Social and Ecological Processes*, chap. Models of Individual Decision Making in Agent-Based Simulation of Common-Pool Resource Management Institutions. Oxford University Press.

P. Deadman (1999). 'Modelling individual behavior and group performance in an intelligent agent-based simulation of the tragedy of the commons'. *Journal of Environmental Management* **56**:159–172.

J. Delleur (2003). 'The Evolution of Urban Hydrology: Past, Present, and Future'. *Journal of Hydraulic Engineering* **32**(2):563–574.

J. Dent, et al. (1995). 'Simulation of ecological, social and economic factors in agricultural systems'. *Agricultural Systems* **49**(4):337–351.

J. DeShazo (2002). 'Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency'. *Journal of Environmental Economics and Management* **44**:123–143.

C. Dickerson, et al. (2003). *Readings about the Social Animal*, chap. Using Cognitive Dissonance to Encourage Water Conservation. WH Freeman and Company.

C. DiClemente, et al. (2000). 'The Role of Feedback in the Process of Health Behavior Change'. In *Proceedings of the Inaugural Meeting, American Academy of Health Behavior*, vol. 25, pp. 217–227.

P. Diederen, et al. (2003). 'Modernisation in agriculture: what makes a farmer adopt an innovation?'. *International Journal of Agricultural Resources, Governance and Ecology* **2**(3-4):323–342.

T. Dietz, et al. (2005). 'Environmental Values'. *Annual Review of Environment and Resources* **30**:335–372.

N. Doidge (2007). *The Brain That Changes Itself*. Viking Books.

W. Dolfsma (2002). 'Mediated Preferences: How Institutions Affect Consumption'. *Journal of Economic Issues* **36**(2):449–457.

J. Doran (1995). *Recent Trends in Data Type Specification*, chap. Agent-Based Modelling of Ecosystems for Sustainable Resource Management. Springer.

B. Dowdle (1984). 'Perspectives on the Sagebrush Rebellion'. *Policy Studies Journal* **12**(3):473–482.

K. Dudley (2000). *Debt and dispossession: farm loss in America's heartland*. University of Chicago Press.

R. Dunlap (2008). 'The New Environmental Paradigm Scale: From Marginality to Worldwide Use'. *The Journal of Environmental Education* **40**(1):3–18.

R. Dunlap, et al. (2000). 'Measuring endorsement of the New Ecological Paradigm: A revised NEP scale'. *The Journal of Social Issues* **56**(3):425–442.

R. Dunlap & K. D. van Liere (2008). 'The New Environmental Paradigm'. *The Journal of Environmental Education* **40**(1):19–28.

B. Durham, et al. (2002). 'Integrated Water Resources Management - through reuse and aquifer recharge'. *Desalination* **152**:333–338.

I. Dyner & C. Franco (2004). 'Consumers' Bounded Rationality: The Case of Competitive Energy Markets'. *Systems Research and Behavioral Science* **21**(373-389).

B. Dziegielewski, et al. (1983). 'Evaluation of drought management measures for municipal and industrial water supply'. Tech. Rep. 83-C-3, Institute for Water Resources.

T. Eggertsson (1990). *Economic behavior and institutions*. Cambridge University Press.

J. Eheart & T. Ng (2004). 'Role of Effluent Permit Trading in Total Maximum Daily Load Programs: Overview and Uncertainty and Reliability Implications'. *Journal of Environmental Engineering* **130**(6):615–622.

I. Ehrlich & R. Posner (1974). 'An Economic Analysis of Legal Rule Making'. *Journal of Legal Studies* **3**:257–186.

P. Ehrlich & J. Holdren (1971). 'Impact of Population Growth'. *Science* **171**(3977):1212–1217.

L. Eisenbud (2007). *The Conceptual Foundations of Quantum Mechanics*. AMS Bookstore.

J. Ela (2009). 'Law and norms in collective action: maximising social influence to minimize carbon emissions'. *UCLA Journal of Environmental Law and Policy* .

A. Ellerman (2005). 'A Note on Tradable Permits'. *Environmental and Resource Economics* **31**:123–131.

A. Ellerman & P. Joskow (2008). 'The European Union's Emissions Trading System in Perspective'. Tech. rep., Pew Center on Global Climate Change.

B. Ellison (1995). 'Autonomy in Action: Bureaucratic Competition Among Functional Rivals in Denver Water Politics'. *Review of Policy Research* **14**(1):25–49.

W. Encinosa, et al. (2007). 'The sociology of groups and the economics of incentives: Theory and evidence on compensation systems'. *Journal of Economic Behavior And Organization* **62**:187–214.

J. Epstein (1999). 'Agent-Based Computational Models and Generative Social Science'. *Complexity* **4**(5):41–61.

J. M. Epstein & R. Axtell (1996). *Growing Artificial Societies: Social Science from the Bottom Up*. MIT Press.

C. Ervin & D. Ervin (1982). 'Factors Affecting the Use of Soil Conservation Practices: Hypotheses, Evidence, and Policy Implications'. *Land Economics* **58**(3):277–292.

R. Estes (1995). *Tyranny of the bottom line: why corporations make good people do bad things*. Berrett-Koehler Publishers.

T. Evans & E. Moran (2002). 'Spatial Integration of Social and Biophysical Factors Related to Landcover Change'. *Population and Development Review* **28**:165–186.

E. Fama (1965). 'The behavior of stock market prices'. *Journal of Business* **38**:34–105.

F. Fang, et al. (2005). 'Point-Nonpoint Source Water Quality Trading: A Case Study in the Minnesota River Basin'. *Journal of the American Water Resources Association* **June**:645–658.

D. Feeny, et al. (1990). 'The Tragedy of the Commons: Twenty-Two Years Later'. *Human Ecology* **18**(1):1–19.

E. Fenichel (2009). 'Heterogeneity among agent types and second-best management for non-market ecological services'. In *2009 AAEA and ACCI Joint Annual Meeting*.

M. Ferge (1999). *Irrigated Eden: the Making of an Agricultural Landscape in the American West*. University of Washington Press.

F. Ferraro, et al. (2005). 'Economics Language and Assumptions: How Theories can Become Self-Fulfilling'. *Academy of Management Review* **30**(1):8–24.

C. Ferster & B. Skinner (1957). *Schedules of Reinforcement*. APA Books.

L. Festinger (1957). *The Theory of Cognitive Dissonance*. Stanford University Press.

S. Feuillette, et al. (2003). 'SINUSE: a multi-agent model to negotiate water demand management on a free access water table'. *Environmental Modelling and Software* **18**:413–427.

P. Feyerabend (1970). *Beyond the Edge of Uncertainty*, chap. Problems of Empiricism. Prentice-Hall.

G. Filson (1996). 'Demographic and Farm Characteristic Differences in Ontario Farmers' Views about Sustainability Policies'. *Journal of Agricultural and Environmental Ethics* **9**(2):165–180.

M. Finke & S. Huston (2003). 'Factors Affecting the Probability of Choosing a Risky Diet'. *Journal of Family and Economic Issues* **24**(3):291–304.

A. Fischer & J. Bliss (2008). 'Behavioral assumptions of conservation policy: conserving oak habitat on family-forest land in the Willamette Valley, Oregon'. *Conservation Biology* **22**(2):275–238.

B. Fischoff (1991). 'Value elicitation: is there anything in there?'. *American Psychologist* **43**:835–847.

I. Fisher (1930). *The Theory of Interest as Determined by Impatience to Spend Income and Opportunity to Invest It*. MacMillan.

C. Folke (2003). 'Freshwater for resilience: a shift in thinking'. *Philosophical Transactions of the Royal Society London* **13**:2027–2036.

M. Fourcade & K. Healy (2007). 'Moral Views of Market Society'. *Annual Review of Sociology* **33**:285–311.

J. Fowler & O. Smirnov (2005). 'Dynamic Parties and Social Turnout: An Agent-Based Model'. *American Journal of Sociology* **110**(4):1070–1094.

I. Fox (1976). 'Institutions for Water Management in a Changing World'. *Natural Resources Journal* **16**:743.

R. Frank, et al. (1993). 'Does studying economics inhibit cooperation?'. *Journal of Economic Perspectives* **7**:159–171.

J. Frankel & K. Froot (1986). 'Understanding the US Dollar in the eighties: the expectations of chartists and fundamentalists'. *Economic Record* **4**(2):24–38.

S. Frederick (2006). 'Valuing future life and future lives: A framework for understanding discounting'. *Journal of Economic Psychology* **27**:667–680.

J. Fredericks, et al. (1998). 'Decision support system for conjunctive stream-aquifer management'. *Journal of Water Resources Planning and Management* **124**(2):69–78.

B. Frey (1999). 'Morality and Rationality in Environmental Policy'. *Journal of Consumer Policy* **22**:395–417.

O. Friborg, et al. (2006). 'Likert-based vs. semantic differential-based scorings of positive psychological constructs: a psychometric comparison of two versions of a scale measuring resilience'. *Personality and Individual Differences* **40**(5):873–884.

A. Friedman (1953). *On the methodology of positive economics*. University of Chicago Press.

J. Friedman, et al. (1998). 'Downstream effects of dams on channel geometry and bottomland vegetation: Regional patterns in the great plains'. *Wetlands* **18**(4):619–633.

M. Friedman (1955). 'Leon Walras and His Economic System'. *The American Economic Review* **45**(5):900–909.

B. Gardner (2000). 'Economic Growth and Low Incomes in Agriculture'. *American Journal of Agricultural Economics* **82**(5):1059–1074.

H. Gardner (1917). 'Cooperation Among the Mormons'. *The Quarterly Journal of Economics* **31**(3):461–499.

R. Gardner & R. Young (1985). 'An Economic Evaluation of the Colorado River Basin Salinity Control Program'. *Western Journal of Agricultural Economics* **10**(1):1–12.

R. Gardner & R. Young (1988). 'Assessing Strategies for Control of Irrigation-Induced Salinity in the Upper Colorado River Basin'. *American Journal of Agricultural Economics* **70**(1):37–49.

S. Gass (1983). 'Decision-Aiding Models: Validation, Assessment, and Related Issues for Policy Analysis'. *Operations Research* **31**(4):603–631.

T. Gates, et al. (2002). 'Monitoring and Modeling Flow and Salt Transport in a Salinity-Threatened Irrigated Valley'. *Journal of Irrigation and Drainage Engineering* **128**(2):87–99.

T. Gates, et al. (2006). 'Toward Optimal Water Management in Colorado's Lower Arkansas River Valley: Monitoring and Modeling to Enhance Agriculture and Environment'. *Colorado State University* .

J. Geanakoplos (2004). 'The Arrow-Debreu Model of General Equilibrium'. *Cowles Foundation Papers* **1090**.

E. Gerber & J. Lewis (2004). 'Beyond the Median: Voter Preferences, District Heterogeneity, and Political Representation'. *Journal of Political Economy* **112**(6):1364–1383.

S. Ghoshal & P. Moran (1996). 'Bad for Practice: A Critique of the Transaction Cost Theory'. *Academy of Management Review* **21**(1):13–47.

G. Gigerenzer (2006). *Contemporary debates in cognitive science*, chap. Bounded and Rational. Blackwell.

N. Gilbert & K. Troitzsch (2005). *Simulation for the social scientist*. Open University Press.

R. Gimblett, et al. (2001). 'The simulation and visualization of complex human-environment interactions'. *Landscape and Urban Planning* **54**(1):63–79.

H. Gintis (2000). 'Beyond Homo Economicus: evidence from experimental economics'. *Ecological Economics* **35**:311–322.

H. Gintis (2004). 'Modeling cooperation among self-interested agents: a critique'. *The Journal of Socio-Economics* **33**:695–714.

K. Glanz & R. Mullis (1988). 'Environmental Interventions to Promote Healthy Eating: A Review of Models, Programs and Evidence'. *Health Education and Behavior* **15**(4):395–415.

P. Gleick (2000). 'The changing water paradigm: a look at twenty-first century water resources development'. *Water International* **25**(1):127–138.

R. Glennon & P. Culp (2002). 'The Last Green Lagoon: How and Why the Bush Administration Should Save the Colorado River Delta'. *Ecology Law Quarterly* **28**(903):1–92.

S. Goetz & D. Debertin (2001). 'Why Farmers Quit: A County-Level Analysis'. *American Journal of Agricultural Economics* **83**(4):1010–1023.

K. Goff, et al. (1998). 'Simulated Effects of Irrigation on Salinity in the Arkansas River Valley of Colorado'. *Ground Water* **36**(1):76–86.

R. Goldstone & M. Janssen (2005). 'Computational models of collective behavior'. *Trends in Cognitive Sciences* **9**(9):424–431.

B. Gomez & J. Wilson (2006). 'Cognitive Heterogeneity and Economic Voting: A Comparative Analysis of Four Democratic Electorates'. *American Journal of Political Science* **50**(1):127–145.

J. Gomez-Ibanez & J. Kalt (1986). 'Controlling Acid Rain'. Tech. rep., John F Kennedy School of Government.

M. Goodrich, et al. (2000). 'Satisficing Revisited'. *Minds and Machines* **10**:79–110.

S. Gorelick & L. Lefkoff (1990). 'Simulating Physical Processes and Economic Behavior in Saline, Irrigated Agriculture: Model Development'. *Water Resources Research* **26**(7):1359–1369.

J. Gowdy (2008). 'Behavioral Economics and Climate Change Policy'. *Journal of Economic Behavior And Organization* **68**:632–644.

GPO (2001). 'Federal Register 41817'. Tech. rep., Government Printing Office (GPO).

M. Grauer, et al. (eds.) (1984). *Plural Rationality and Interactive Decision Processes*, no. 248. Springer.

R. Gregory, et al. (1993). 'Valuing environmental resures: a constructive approach'. *Journal of Risk and Uncertainty* **7**:177–197.

R. Griffin & S.-H. Hsu (1993). 'The Potential for Water Market Efficiency when Instream Flows have Value'. *American Journal of Agricultural Economics* **75**(2):292–303.

N. Grigg (1996). *Water Resources Management: principles, regulations and cases*. McGraw-Hill Professional.

M. Grodzinka-Jurczak, et al. (2006). 'Effects of an educational campaign on public environmental attitudes and behaviour in Poland'. *Resources, Conservation and Recycling* **46**:182–197.

G. Guagnano (2001). 'Altruism and Market-Like Behavior: An Analysis of Willingness to Pay for Recycled Paper Products'. *Population and Environment* **22**(4):425–437.

L. Gunderson (1999). 'Resilience, Flexibility and Adaptive Management - Antidotes for Spurious Certitude?'. *Ecology and Society* **3**(1):7.

L. Gunderson & C. Holling (2002). *Panarchy: Understanding Transformations in Human and Natural Systems*. Island Press.

H. Gupta, et al. (1997). 'Toward improved calibration of hydrologic models: multiple and noncommensurable measures of information'. *Water Resources Research* **34**(4):751.

O. Gurerk, et al. (2009). 'Motivating teammates: The leader's choice between positive and negative incentives'. *Journal of Economic Psychology* **In Press**.

T. Gurung, et al. (2006). 'Companion Modeling, Conflict Resolution, and Institution Building: Sharing Irrigation Water in the Lingmuteychu Watershed, Bhutan'. *Ecology and Society* **11**(2):36.

W. Guth & H. Kliemt (2004). 'Perfect or Bounded Rationality?'. *Analyse and Kritik* **26**:364–381.

GWP (2000). 'Integrated Water Resources Management'. Tech. Rep. 4, Global Water Partnership.

R. Hahn (1989). 'Economic Prescriptions for Environmental Problems: How the Patient Followed the Doctor's Orders'. *The Journal of Economic Perspectives* **3**(2):95–114.

M. Haigh & J. List (2005). 'Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis'. *The Journal of Finance* **60**(1):523–536.

D. Hamilton (1996). 'Knowledge Acquisition for Multiple Site, Related Domain Expert Systems: Delphi Process and Application'. *Expert Systems with Applications* **11**(3):377–389.

E. Hanak (2005). 'Stopping the Drain: Third-party Responses to California's Water Market'. *Contemporary Economic Policy* **23**(1):59–77.

H. Hansen (2007). 'An Adaptive Land-use Simulation Model for Integrated Coastal Zone Planning'. *Lecture Notes in Geoinformation and Cartography* pp. 35–54.

K. Happe, et al. (2006). 'Agent-based Analysis of Agricultural Policies: an Illustration of the Agricultural Policy Simulator AgriPoliS, its Adaptation and Behavior'. *Ecology and Society* **11**(1):49.

G. Hardin (1968). 'The Tragedy of the Commons'. *Science* **162**:1242–1248.

E. Harmon-Jones & J. Mills (1999). *Cognitive dissonance: Progress on a pivotal theory in social psychology*, chap. An introduction to cognitive dissonance theory and an overview of current perspectives on the theory. American Psychological Association.

M. Harris (1979). *Cultural Materialism*. Random House.

S. Harter (2001). *The construction of the self: a developmental perspective*. Guilford Press.

M. Haw, et al. (2000). 'A pinch of salt: landowner perception and adjustment to the salinity hazard in Victoria, Australia'. *Journal of Rural Studies* **16**(2):155–169.

C. Hay (2004). 'Theory, Stylized Heuristic or Self-fulfilling Prophecy? The Status of Rational Choice Theory in Public Administration'. *Public Administration* **82**(1):39–62.

T. Herman (1959). 'One Chinese View of the Hydraulic Society'. *Annals of the Association of American Geographers* **49**(3):324–325.

G. Hesseling (1996). *Improved natural resource management: the role of formal organizations and informal networks and institutions*, chap. Legal and institutional incentives for local environmental management. Leiden University.

D. Hillel (1992). *Out of the Earth: Civilization and the Life of the Soil*. University of California Press.

J. Hird (1991). 'The Political Economy of Pork: Project Selection at the U.S. Army Corps of Engineers'. *American Political Science Review* **85**(2):429–456.

D. Hoag & J. Hughes-Popp (1997). 'Theory and Practice of Pollution Credit Trading in Water Quality Management'. *Review of Agricultural Economics* **19**(2):252–262.

G. Hobbs (1985). 'Federal Environmental Law and State Water Law: Accommodation or Preemption'. *Natural Resources and Environment* **23**.

G. Hobbs (1994). 'Ecological integrity, new Western myth: a critique of the Long's Peak report'. *Environmental Law* **24**.

G. Hobbs (1997). 'Colorado Water Law: An Historical Overview'. *Water Law Review* **1**(1).

G. Hobbs (2002). 'Priority: the most misunderstood stick in the bundle'. *Environmental Law* **32**:37–59.

G. Hobbs (2003). *Citizen's Guide to Colorado Water Law*. Colorado Foundation for Water Education.

G. Hobbs (2007). 'Colorado, Centennial State at the Headwaters'. *Ecology Law Currents* **36**:220–226.

G. Hofstede (1991). *Cultures and organizations: Software of the Mind*. McGraw-Hill.

S. Holland & M. Moore (2003). 'Cadillac Desert revisited: property rights, public policy, and water resource depletion'. *Journal of Environmental Economics and Management* **46**(1):131–155.

C. Hommes (2006). *Handbook of Computational Economics*, chap. Heterogeneous Agent Models in Economics and Finance. North Holland.

J. Hopper & J. Nielsen (1991). 'Recycling as Altruistic Behavior'. *Environment and Behavior* **23**(2):195–220.

R. Horan & M. Ribaudo (1999). 'Policy Objectives and Economic Incentives for Controlling Agricultural Sources of Nonpoint Pollution'. *Journal of the American Water Resources Association* **35**(5):1023–1036.

C. Howe (1998). *Markets for Water: Potential and Performance*, chap. Water Markets in Colorado: Past Performance and Needed Changes. Springer.

C. Howe, et al. (1990). 'The economic impacts of agriculture-to-urban water transfers on the area of origin: a case study of the Arkansas River Valley in Colorado'. *American Journal of Agricultural Economics* **72**(5):1200–1204.

T. Huerta (2008). 'Simulating Institutional Controls on Consumption Patterns in the Commons'. *Systems Research and Behavioral Science* **922**.

M. Humphreys (2005). 'Natural Resources, Conflict, and Conflict Resolution: Uncovering the Mechanisms'. *Journal of Conflict Resolution* **49**(4):508–537.

R. Hunt (1988). 'Size and the Structure of Authority in Canal Irrigation Systems'. *Journal of Anthropological Research* **44**(4):335–355.

J. Ikerd (2002). 'The Globalization of Agriculture: Implication for Sustainability of Small Horticultural Farms'. *Actae Horticulturae* **638**.

H. Ingram (1967). *Water Politics: Continuity and Change*. University of New Mexico Press.

H. Ingram (2006). 'Water as a multi-dimensional value: implications for participation and transparency'. *International Environmental Agreements* **6**:429–433.

J. Irwin, et al. (1993). 'Preference reversals and the measurement of environmental values'. *Journal of Risk and Uncertainty* **6**:5–18.

H. Jacoby & A. Ellerman (2004). 'The safety valve and climate policy'. *Energy Policy* **32**:481–491.

W. Jager, et al. (2000). 'Behavior in commons dilemmas: Homo economicus and Homo psychologicus in an ecological-economic model'. *Ecological Economics* **35**:357–379.

W. Jager & H. Mosler (2007). 'Simulating Human Behavior for Understanding and Managing Environmental Resource Use'. *Journal of Social Issues* **63**(1):97–116.

K. Jansen & S. Vellema (2004). *Agribusiness and Society: Corporate responses to environmentalism, market opportunities and public regulation*. Palgrave Macmillan.

M. Janssen & S. Carpenter (1999). 'Managing the resilience of Lakes: A Multi-agent Modeling Approach'. *Ecology and Society* **3**(2):15.

M. Janssen & E. Ostrom (2006). 'Empirically Based, Agent-based Models'. *Ecology and Society* **11**(2).

M. Janssen & E. Ostrom (2007). *Inequality, cooperation, and environmental sustainability*, chap. Adoption of a New Regulation for the Governance of Common-Pool Resources by a Heterogeneous Population. Princeton University Press.

S. Jeffrey & V. Shaffer (2007). 'The Motivational Properties of Tangible Incentives'. *Compensation Benefits Review* **39**(44-52).

M. Jensen (2002). 'Value Maximization, Stakeholder Theory, and the Corporate Objective Function'. *Business Ethics Quarterly* **12**(2):235–256.

C. Ji (2004). 'Factor structure of the new environmental paradigm scale: evidence from an urban sample in southern California'. *Psychology Reports* **94**(1):125–130.

L. Johnson, et al. (2006). 'Patterns and pathways in the post-establishment spread of non-indigenous aquatic species: the slowing invasion of North American inland lakes by the zebra mussel'. *Biological Invasions* **8**:475–489.

A. Johnston (2002). 'Treating International Institutions as Social Environments'. *International Studies Quarterly* **45**(4):487–515.

M. Jolson & G. Rossow (1971). 'The Delphi Process in Marketing Decision Making'. *Journal of Marketing Research* **8**(4):443–448.

C. Kagitcibasi (2005). 'Automony and Relatedness in Cultural Context: Implications for Self and Family'. *Journal of Cross-Cultural Psychology* **36**:403–424.

D. Kahneman (2003). 'Maps of Bounded Rationality: Psychology for Behavioral Economics'. *American Economic Review* pp. 1449–1476.

V. Kandiah & P. Rao (2004). 'Water Infrastructure Interdependencies: A Complex Adaptive System'. *Journal of Environmental Hydrology* **12**.

S. Kaplan (2000). 'Human Nature and Environmentally Responsible Behavior'. *Journal of Social Issues* **56**(3):491–508.

L. Kaufman (2009). 'Utilities Turn Their Customers Green, With Envy'.

C. Kaysen (1957). 'The social significance of the modern corporation'. *The American Economic Review* **47**(2):311–319.

D. Kenney (2005). *Water Rights Reform: Lessons for Institutional Design*, chap. Prior Appropriation and Water Rights Reform in the Western United States. IFPRI.

D. Kenney, et al. (2004). 'Use and Effectiveness of Municipal Water Restrictions During Drought in Colorado'. *Journal of the American Water Resources Association* pp. 77–88.

A. Keyes & W. Werick (1995). *Virtual Droughts and Shared Visions: Some Innovations from the National Drought Study*. Diane Publishing Co.

S. Khan, et al. (2006). 'Can irrigation be sustainable?'. *Agricultural Water Management* **80**(1):87–99.

W. Kilbourne, et al. (2002). 'The role of the dominant social paradigm in environmental attitudes: A multinational examination'. *Journal of Business Research* **55**(193-204).

D. King (2005). 'Crunch Time for Water Quality Trading'. *Choices: The magazine of food, farm and resource issues* **20**(1):71–76.

A. Kirman (1992). 'Whom or What does the Representative Individual Represent?'. *The Journal of Economic Perspectives* **6**(2):117–136.

M. Klaes & E. Sent (2005). 'A Conceptual History of the Emergence of Bounded Rationality'. *History of Political Economy* **37**(1):27–60.

D. Knoke (1977). 'Political Structure of Rural America'. *The Annals of the American Academy of Political and Social Science* **429**(1):51–62.

B. Kock (2008). 'Going beyond the unitary curve: incorporating richer cognition into agent-based water resources models'. *Eos Trans. AGU* **89**(3).

A. Kohn (1990). *The brighter side of human nature*. Basic Books.

A. Kollmuss & J. Agyeman (2002). 'Mind the Gap: why do people act environmentally and what are the barriers to pro-environmental behavior'. *Environmental Education Research* **8**(3):239–261.

N. Kolodny (2005). 'Why be rational?'. *Mind* **114**(455):510–563.

L. Konikow & E. Kendy (2005). 'Groundwater depletion: A global problem'. *Hydrogeology Journal* **13**:317–320.

S. Kopelman, et al. (2002). *The Drama of the Commons*, chap. Factors Influencing Cooperation in Commons Dilemmas: A Review of Experimental Psychological Research. National Academy Press.

J. Korhonen (2002). 'The Dominant Economics Paradigm and Corporate Social Responsibility'. *Corporate Social Responsibility and Environmental Management* **9**:67–80.

G. Kuehne, et al. (2007). 'There's more than one type of farmer: acknowledging farmers' diversity - an Australian perspective'. *International Journal of Interdisciplinary Social Sciences* **2**(2):179–186.

T. Kuhn (1962). *The Structure of Scientific Revolutions*. University of Chicago Press.

M. Kunc & J. Morecroft (2007). 'Competitive dynamics and gaming simulation: lessons from a fishing industry simulator'. *Journal of the Operational Research Society* **58**:1146–1155.

T. Kurz (2002). 'The Psychology of Environmentally Sustainable Behavior: Fitting Together Pieces of the Puzzle'. *Analyses of Social Issues and Public Policy* pp. 257–279.

P. Lachapelle, et al. (2003). 'Barriers to Effective Natural Resource Planning in a Messy World'. *Society and Natural Resources* **16**:473–490.

I. Lakatos (1968). *The Problem of Inductive Logic*. North Holland.

R. Lalonde & E. Jackson (2002). 'The New Environmental Paradigm Scale: Has It Outlived Its Usefulness?'. *The Journal of Environmental Education* **33**(4):28–36.

T. Lang (1999). 'The complexities of globalization: the UK as a case study of tensions within the food system and the challenge to food policy'. *Agriculture and Human Values* **16**:169–185.

P. V. Lange & D. Kuhlman (1994). 'Social value orientations and impressions of partner's honesty and intelligence: A test of the might versus morality effect'. *Journal of Personality and Social Psychology* **67**(1):126–141.

P. V. Lange & W. Liebrand (1991). 'Social value orientation and intelligence: A test of the Goal Prescribes Rationality Principle'. *European Journal of Social Psychology* **21**(4):273–292.

S. Lea & P. Webley (2006). 'Money as tool, money as drug: The biology psychology of a strong incentive'. *Behavioral and Brain Sciences* **29**:161–209.

E. Leach (1959). 'Hydraulic Society in Ceylon'. *Past and Present* **15**(1):2–26.

R. Lempert (2002). 'Agent-based modeling as organizational and public policy simulators'. *Proceedings of the National Academy of Sciences* **99**(3):7195–7196.

T. Lepper (2008). *Reregulating the flows of the Arkansas River: Comparing forms of common pool resource organizations*. Ph.D. thesis, Colorado State University.

A. Leyshon & N. Thrift (1995). 'Geographies of Financial Exclusion: Financial Abandonment in Britain and the United States'. *Transactions of the Institute of British Geographers* **20**(3):312–341.

G. Libecap (2008). 'Chinatown Revisited: Owens Valley and Los Angeles - Bargaining Costs and Fairness Perceptions of the First Major Water Rights Exchange'. *Journal of Law, Economics, and Organization* (1-28).

J. Lichatowich (1999). *Salmon without Rivers: a History of the Pacific Salmon Crisis*. Island Press.

E. Lichtenberg & R. Zimmerman (1999). 'Information and farmers' attitudes about pesticide, water quality, and related environmental effects'. *Agriculture, Ecosystems and Environment* **73**(3):227–236.

A. Ligtenberg, et al. (2004). 'A design and application of a multi-agent system for simulation of multi-actor spatial planning'. *Journal of Environmental Management* **72**:43–55.

R. Likert (1932). 'Technique for the Measurement of Attitudes'. *Archives of Psychology* **140**.

W. Lin, et al. (1974). 'An Empirical Test of Utility vs. Profit Maximization in Agricultural Production'. *American Journal of Agricultural Economics* **56**(3):497–508.

A. Lindbeck, et al. (1999). 'Social Norms and Economic Incentives in the Welfare State'. *The Quarterly Journal of Economics* **64**(1):1–35.

S. Lindenberg & L. Steg (2007). 'Normative, Gain and Hedonic Goal Frames Guiding Environmental Behavior'. *Journal of Social Issues* **63**(1):117–137.

B. Linner (2004). *The return of Malthus: environmentalism and post-water population-resource crises*. White Horse Press.

H. Linstone & M. Turoff (1975). *The Delphi Method: Techniques and Applications*. Addison-Wesley Publishing Company.

B. Lomborg (2001). 'The truth about the environment'. *The Economist* .

M. Luckert & W. Adamowicz (1993). 'Empirical Measures of Factors Affecting Social Rates of Discount'. *Environmental and Resource Economics* **3**:1–21.

D. Lueck (1995). 'The Rule of First Possession and the Design of the Law'. *Journal of Law and Economics* **38**:393.

J. Lusk, et al. (2007). 'Public Preferences and Private Choices: Effect of Altruism and Free Riding on Demand for Environmentally Certified Pork'. *Environmental and Resource Economics* **36**:499–521.

I. Lustick (2000). 'Agent-based modelling of collective identity: testing constructivist theory'. *Journal of Artificial Societies and Social Simulation* **3**(1):1.

L. Lutzenhiser (1993). 'Social and Behavioral Aspects of Energy Use'. *Annual Review of Energy and Environment* **18**:247–289.

G. Lynne (1995). 'Modifying the Neo-Classical Approach to Technology Adoption With Behavioral Science Models'. *Journal of Agricultural and Applied Economics* **27**(1):67–80.

G. Lynne, et al. (1988). 'Attitudes and Farmer Conservation Behavior'. *American Journal of Agricultural Economics* **70**(1):12–19.

G. Lynne, et al. (1998). 'Attitudes and Producer Conservation Behavior'. *American Journal of Agricultural Economics* **70**(1):12–19.

A. Maass (1951). *Muddy waters: the Army Engineers and the Nation's rivers*. Harvard University Press.

A. Maass, et al. (1962). *Design of Water Resource Systems: New techniques for Relating Economic Objectives, Engineering Analysis, and Governmental Planning*. Harvard University Press.

E. MacKaay (2000). *Encyclopedia of Law and Economics*, vol. 1, chap. History of Law and Economics. Edward Elgar.

W. Magdych (1984). 'Salinity Stresses Along a Complex River Continuum: Effects on Mayfly (Ephemeroptera) distributions'. *Ecology* **65**(5):1662–1672.

U. Maki (2002). *Fact and fiction in economics: models, realism, and social construction*. Cambridge University Press.

J. Mankoff, et al. (2007). 'Leveraging Social Networks to Motivate Individuals to Reduce their Ecological Footprints'. In *Proceedings of the 40th Hawaii International Conference on System Sciences*.

R. Marks (2007). 'Validating Simulation Models: A General Framework and Four Applied Examples'. *Computational Economics* **30**(3):265–290.

G. Marshall (2004). 'Farmers cooperating the commons? A study of collective action in salinity management'. *Ecological Economics* **51**:271–286.

R. Mathevet, et al. (2003). 'Agent-based simulations of interactions between duck population, farming decisions and leasing of hunting rights in the Camargue (Southern France)'. *Ecological Modelling* **165**(2):107–126.

R. Matthews (2006). 'The People and Landscape Model (PALM): Towards full integration of human decision making and biophysical simulation models'. *Ecological Modelling* **194**(4):329–343.

R. Matthews, et al. (2007). 'Agent-based land-use models: a review of applications'. *Landscape Ecology* **22**:1447–1459.

N. Maxwell (1972). 'A Critique of Popper's Views on Scientific Method'. *Philosophy of Science* **39**(2):131–152.

D. Maybery, et al. (2005). 'Categorizing farming values as economic, conservation and lifestyle'. *Journal of Economic Psychology* **26**:59–72.

M. McClure, et al. (2003). 'A Large-Scale, Multispecies Status Assessment: Anadromous Salmonids in the Columbia River Basin'. *Ecological Applications* **13**(4):964–989.

A. McMakin, et al. (2002). 'Motivating Residents to Conserve Energy Without Financial Incentives'. *Environment and Behavior* **34**:848–865.

J. McNeely (1993). 'Economic Incentives for Conserving Biodiversity: lessons for Africa'. *Ambio* **22**(2):144–150.

N. Meade & T. Islam (2006). 'Modelling and forecasting the diffusion of innovation - A 25-year review'. *International Journal of Forecasting* **22**:519–545.

D. Meadows (1972). *The Limits of Growth: A Report for the Club of Rome's Project on the Predicament of Mankind*. Universe Books.

C. Meghir & L. Pistaferri (2004). 'Income Variance Dynamics and Heterogeneity'. *Econometrica* **72**(1):1–32.

M. Meisner (1963). 'The Despotism of Concepts: Wittfogel and Marx on China'. *The China Quarterly* **16**:99–111.

E. Michel-Guillou & G. Moser (2006). 'Commitment of farmers to environmental protection: From social pressure to environmental conscience'. *Journal of Environmental Psychology* **26**:227–235.

D. Midgley, et al. (2006). 'The Building and Assurance of Agent-Based Models: An Example and Challenge to the Field'. In *Proceedings of the Agent-Based Models of Market Dynamics and Consumer Behavior*.

F. Milenski (1990). *F Milenski*. Trails Publishing Co.

D. Miles (1977). 'Salinity in the Arkansas Valley of Colorado'. *EPA* (EPA-1AG-D4-0544).

D. Miller (1999). 'The norm of self-interest'. *American Psychologist* **54**:1053–1060.

K. Miller, et al. (1996). 'Global change in microcosm: US Water Institutions'. *Policy Sciences* **29**:271–290.

T. Moe (1984). 'The New Economics of Organization'. *American Journal of Political Science* **28**:739–777.

T. Moe (1997). *Perspectives on public choice*, chap. The positive theory of public bureaucracy. Cambridge University Press.

C. Moore & J. Doherty (2006). 'The cost of uniqueness in groundwater model calibration'. *Advances in Water Resources* **29**:605–623.

D. Mueller (1994). *The philosophy of economics*, chap. The corporation and the economist. Cambridge University Press.

L. Mumford (1967). *Technics and Human Development*. Harcourt Brace Jovanovich.

J. Murphy & J. Stranlund (2007). 'A laboratory investigation of compliance behavior under tradable emissions rights: Implications for targeted enforcement'. *Journal of Environmental Economics and Management* **53**:196–212.

J. Muys (1977). *Water Needs for the Future*, chap. Allocation and Management of Interstate Water Resources: The Emergence of the Federal-Interstate Compact. Westview Press.

M. Nabli & J. Nugent (1989). 'The New Institutional Economics and Its Applicability to Development'. *World Development* **17**:1333–1347.

J. Nelson (2005). 'Is economics a natural science'. *Cosmos and History: The Journal of Natural and Social Philosophy* **2**:261–269.

N. Nelson & A. Keeler (2005). 'Water Quality Trading in Georgia: Addressing Two Potential Impediments to Success'. In K. Hatcher (ed.), *Proceedings of the 2005 Georgia Water Resources Conference*.

P. Nelson (1970). 'Information and Consumer Behavior'. *The Journal of Political Economy* **78**(2):311–329.

B. Newell (2005). 'Re-visions of rationality?'. *Trends in Cognitive Sciences* **9**(1):11–15.

R. Newell, et al. (2005). 'Fishing quota markets'. *Journal of Environmental Economics and Management* **49**:437–462.

T. Nguyen, et al. (2006). 'A Guide to Market-Based Approaches to Water Quality'. *CSREES Publications* .

P. Nichols, et al. (2001). *Water and Growth in Colorado: A Review of Legal and Policy Issues*. Natural Resources Law Center.

D. Nielsen, et al. (2003). 'Effects of increasing salinity on freshwater ecosystems in Australia'. *Australian Journal of Botany* **51**:655–665.

M. Nishibe & H. Uemura (2005). 'The Socio-economics of Institutions and Evolution'. *Evolutionary Institutional Economics Review* **2**(1):1–5.

P. Norris & S. Batie (1987). 'Virgina Farmers' Soil Conservation Decisions: An Application of Tobit Analysis'. *Southern Journal of Agricultural Economics* pp. 79–91.

B. Norton, et al. (1998). 'The evolution of preferences: why sovereign preferences may not lead to sustainable policies and what to do about it'. *Ecological Economics* **24**:193–211.

B. Norton & A. Steinemann (2001). 'Environmental Values and Adaptive Management'. *Environmental Values* **10**:473–506.

J. Nunnaly & I. Bernstein (1994). *Psychometric Theory*. McGraw-Hill.

E. Nyhus & P. Webley (2006). *Handbook of Contemporary Behavioral Economics*, chap. Discounting, Self-Control, and Saving. ME Sharpe.

T. Odean (1999). 'Do Investors Trade Too Much?'. *The American Economic Review* **89**(5):1279–1298.

B. Ohlmer, et al. (1998). 'Understanding farmers' decision making processes and improving managerial assistance'. *Agricultural Economics* **18**:273–290.

F. Olander & J. Thogersen (1995). 'Understanding of Consumer Behavior as a Prerequisite for Environmental Protection'. *Journal of Consumer Policy* **18**:345–385.

N. Oreskes, et al. (1994). 'Verification, validation and confirmation of numerical models in the earth sciences'. *Science* **263**(5147):641–646.

R. Ortiz (2004). 'Methods to Identify Changes in Background Water-Quality Conditions Using Dissolved-Solids Concentrations and Loads as Indicators, Arkansas River and Fountain Creek, in the Vicinity of Pueblo, Colorado'. Tech. Rep. 5024, United States Geological Survey.

C. Osgood (1957). *The measurement of meaning*. University of Illinois Press.

K. Oshio (1997). 'Who Pays and Who Benefits? Metropolitan Water Politics in Twentieth-Century Southern California'. *The Japanese Journal of American Studies* **8**:63–90.

E. Ostrom (1990). 'Governing the Commons: The Evolution of Institutions for Collective Action'. *Cambridge University Press* .

E. Ostrom (1995). 'New Horizons in Institutional Analysis'. *American Political Science Review* **89**(1):174–178.

E. Ostrom (1998). 'A Behavioral Approach to the Rational Choice Theory of Collective Action'. *The American Political Science Review* **92**(1):1–22.

E. Ostrom (2005). *Understanding Institutional Diversity*. Princeton University Press.

E. Ostrom, et al. (1999). 'Revisiting the Commons: Local Lessons, Global Challenges'. *Science* **284**:278–282.

E. Ostrom, et al. (2002). *The Drama of the Commons*. National Academy Press.

E. Ostrom, et al. (1992). 'Covenants with and without a sword'. *American Political Science Review* **86**(2):404–417.

V. Ostrom, et al. (1988). *Rethinking Institutional Analysis and Development*. Institute for Contemporary Studies.

V. Ostrom & E. Ostrom (1999). *Polycentricity and local public economies*, chap. Public Goods and Public Choices. University of Michigan Press.

D. O'Sullivan & M. Haklay (2000). 'Agent-based models and individualism: is the world agent-based?'. *Environment and Planning A* **32**:1409–1425.

C. Pahl-Wostl (2002). 'Towards sustainability in the water sector - the importance of human actors and processes of social learning'. *Aquatic Sciences* **64**:394–411.

C. Pahl-Wostl (2007a). 'The implications of complexity for integrated resources management'. *Environmental Modelling and Software* **22**(5):561–569.

C. Pahl-Wostl (2007b). 'Transitions towards adaptive management of water facing climate and global change'. *Water Resources Management* **21**:49–62.

D. Pannell, et al. (2006). 'Understanding and promoting adoption of conservation practices by rural landowners'. *Australian Journal of Experimental Agriculture* **46**:1407–1424.

D. Pannell, et al. (2001). 'Rethinking the externality issue for dryland salinity in Western Australia'. *The Australian Journal of Agricultural and Resource Economics* **45**(3):459–475.

D. Parfit (1971). 'Personal Identity'. *The Philosophical Review* **80**(1):3–27.

F. Parisi (2004). 'Positive, Normative and Functional Schools in Law and Economics'. *European Journal of Law and Economics* **18**:259–272.

D. Parker, et al. (2008). 'Complexity, land-use modeling, and the human dimension: fundamental challenges for mapping unknown outcome spaces'. *Geoforum* **39**:789–804.

K. Paudel, et al. (2000). 'Effect of risk perspective on fertilizer choice by sharecroppers'. *Agricultural Systems* **66**(2):115–128.

J. Pennings & R. Leuthold (2000). 'The Role of Farmers' Behavioral Attitudes and Heterogeneity in Futures Contracts Usage'. *American Journal of Agricultural Economics* **82**(4):908–919.

L. Pereira, et al. (2002). 'Irrigation management under water scarcity'. *Agricultural Water Management* **57**(3):175–206.

J. Perry (1996). 'Measuring Public Service Motivation: An assessment of construct reliability and validity'. *Journal of Public Administration Research and Theory* **6**(1):5–22.

W. Pesendorfer (2006). 'Behavioral Economics Comes of Age: a review essay on advances in behavioral economics'. *Journal of Economic Literature* **44**:712–721.

J. Peterson, et al. (2007). 'Choice Experiments to Assess Farmers' Willingness to Participate in a Water Quality Trading Market'. In *Proceedings of the American Agricultural Economics Association Annual Meeting 2007*.

C. Pharino (2007). *Sustainable Water Quality Management Policy*. Springer.

D. Pisani (2002). *Water and American Government*. University of California Press.

C. Plott (1967). 'A notion of equilibrium and its possibility under majority rule'. *American Economic Review* **57**(4):787–807.

A. Poteete & E. Ostrom (2004). 'Heterogeneity, Group Size and Collective Action: The Role of Institutions in Forest Management'. *Development and Change* **35**(3):435–461.

T. Presser & H. Ohlendorf (1987). 'Biogeochemical Cycling of Selenium in the San Joaquin Valley, California, USA'. *Environmental Management* **11**(6):805–821.

H. Purnomo, et al. (2005). 'Developing multi-stakeholder forest management scenarios: a multi-agent systems simulation approach applied in Indonesia'. *Forest Policy and Economics* **7**(4):475–491.

J. Pyrko & C. Noren (1998). 'Can we change residential customers' energy attitudes using information and knowledge?'. Tech. rep., Lund Institute of Technology.

M. Qadir & J. Oster (2004). 'Crop and irrigation management strategies for saline-sodic soils and waters aimed at environmentally sustainable agriculture'. *Science of the Total Environment* **323**(1-3):1–19.

H. Qudrat-Ullah (2005). 'Structural Validation of System Dynamics and Agent-Based Simulation Models'. In *Proceedings of the 19th European Conference on Modelling and Simulation*.

A. Rao & M. Georgeff (1998). *Readings in Agents*. Morgan Kaufmann.

T. Rasanen, et al. (2008). 'Reducing energy consumption by using self-organizing maps to create more personalized electricity use information'. *Applied Energy* **85**(9):830–840.

W. Rauch, et al. (2005). 'Integrated Approaches in Urban Storm Drainage: Where Do We Stand?'. *Environmental Management* **35**(4):396–409.

W. Rees (1996). 'Revisiting carrying capacity: Area-based indicators of sustainability'. *Population and Environment* **17**(3):195–215.

A. Reeson (2008). 'Institutions, Motivations and Public Goods: Theory, Evidence and Implications for Environmental Policy'. *CSIRO Working Paper Series* **1**.

M. Reisner (1993). *Cadillac Desert: The American West and its Disappearing Water*. Penguin.

M. Reisner & S. Bates (1990). *Overtapped Oasis: Reform or Revolution for Western Water*. Island Press.

S. Resnick, et al. (2000). 'One-year Age Changes in MRI Brain Volumes in Older Adults'. *Cerebral Cortex* **10**:464–472.

B. Richter, et al. (2003). 'Ecologically sustainable water management: managing river flows for ecological integrity'. *Ecological Applications* **13**(1):206–224.

M. Richter (1971). *Preference, Utility and Demand Theory*, chap. Rational Choice. Harcourt Brace Jovanovich.

J. Rieskamp, et al. (2006). 'Extending the bounds of rationality: evidence and theories of preferential choice'. *Journal of Economic Literature* **44**:631–661.

W. Riker (1982). *Political Equilibrium*, chap. Implications from the disequilibrium of majority rule for the study of institutions. Kluwer-Nijhoff.

R. Rindfuss, et al. (2003). 'The changing institutional context of low fertility'. *Population Research and Policy Review* **22**:411–438.

H. Ritzema, et al. (2008). 'Subsurface drainage to combat waterlogging and salinity in irrigated lands in India: Lessons learned in farmers' fields'. *Agricultural Water Management* **95**(3):179–189.

A. Rixon, et al. (2007). 'Perspectives on the art of facilitation: a Delphi study of natural resource management facilitators'. *Australasian Journal of Environmental Management* **14**:179–191.

R. Robertson (1980). 'Small group decision making: the uncertain role of information in reducing uncertainty'. *Political Behavior* **2**(2):163–189.

S. Roch & C. Samuelson (1997). 'Effects of environmental uncertainty and social value orientation in resource dilemmas'. *Organizational Behavior and Human Decision Processes* **70**(3):221–235.

C. Rodriguez-Sickert, et al. (2008). 'Institutions influence preferences: Evidence from a common pool resource experiment'. *Journal of Economic Behavior And Organization* **67**:215–227.

E. Rogers (1995). *The Information Systems Research Challenge: Survey Research Methods*, vol. 3, chap. The Critical Mass in the Diffusion of Interactive Technologies in Organizations. The Free Press.

P. Rogers, et al. (2002). 'Water is an economic good: How to use prices to promote equity, efficiency, and sustainability'. *Water Policy* **4**:1–17.

C. Roncoli (2006). 'Ethnographic and participatory approaches to research on farmers' responses to climate predictions'. *Climate Research* **33**:81–99.

R. Roth (1990). 'A Delphi Approach to Acquiring Knowledge from Single and Multiple Experts'. *Proceedings of the 1990 ACM Conference* .

J. Rouchier (????). 'Data gathering to build and validate small-scale social models for simulation. Two ways: strict control and stakeholder involvement'. Unknown conference.

O. Rudenko (1991). 'Conditions for applicability of the Newtonian approximation of the theory of relativity and the correspondence principle'. *Russian Physics Journal* **34**(7):601–603.

J. Rushton, et al. (1981). 'The Altruistic Personality and the Self-Report Altruism Scale'. *Personality and Individual Differences* **2**:293–302.

E. Ryan (2001). 'Public Trust and Distrust: The Theoretical Implications of the Public Trust Doctrine for Natural Resource Management'. *Environmental Law* **31**:477–496.

R. Ryan, et al. (2003). 'Farmers' motivations for adopting conservation practices along riparian zones in a mid-western agricultural watershed'. *Journal of Environmental Planning and Management* **46**(1):19–37.

S. Rysgaard, et al. (1999). 'Effects of salinity on NH4+ adsorption capacity, nitrification, and denitrification in Danish estuarine sediments'. *Estuaries and Coasts* **22**(1):21–30.

D. Saari (1998). 'Connecting and resolving Sen's and Arrow's theorems'. *Social Choice and Welfare* **15**:239–261.

I. Sanchez-Cuenca (2008). 'A Preference for Selfish Preferences: The Problem of Motivations in Rational Choice Political Science'. *Philosophy of the Social Sciences* **38**(3):361–378.

C. Sargent (1994). 'Incentives for the sustainable management of the Tropical High Forest in Ghana'. *Commonwealth Forestry Review* **73**(3):155–163.

T. Sargent (1993). *Bounded rationality in Macroeconomics*. Clarendon Press.

A. Sarker, et al. (2008). 'A common-pool resource approach for water quality management: An Australian case study'. *Ecological Economics* **68**:461–471.

E. S. Sartzetakis (1997). 'Tradeable Emission Permits Regulations in the Presence of Imperfectly Competitive Product Markets: Welfare Implications'. *Environmental and Resource Economics* **9**:65–81.

J. Sax (2000). 'Environmental Law at the Turn of the Century: A Reportorial Fragment of Contemporary History'. *California Law Review* **88**(6):2375–2402.

T. Schelling (1984). *Choice and Consequence*. Harvard University Press.

E. Schlager (2006). 'Challenges of governing groundwater in U.S. western states'. *Hydrogeology Journal* **14**:350–360.

E. Schlager (45). 'Rationality, Cooperation, and Common Pool Resources'. *American Behavioral Scientist* **5**(801-819).

M. Schluter & C. Pahl-Wostl (2007). 'Mechanisms of Resilience in Common-Pool Resource Management Systems: an Agent-Based Model of Water Use in a River Basin'. *Ecology and Society* **12**(2):4.

A. Schotter (1981). *The economic theory of social institutions*. Cambridge University Press.

G. Schoups, et al. (2005). 'Sustainability of irrigated agriculture in the San Joaquin Valley, California'. *Proceedings of the National Academy of Sciences* **102**(43):15352–15356.

P. Schultz (2002). *New tools for environmental protection*, chap. Knowledge, Information, and Household Recycling: Examining the Knowledge-Deficit Model of Behavior Change. National Academies Press.

P. Schultz, et al. (2000). 'Acculturation and ecological worldview among Latino Americans'. *The Journal of Environmental Education* **31**(2):22–27.

S. Schumm (1963). 'Sinuosity of Alluvial Rivers on the Great Plains'. *GSA Bulletin* **74**(9):1089–1100.

C. Seligman, et al. (1981). *Advances in environmental psychology*, chap. Encouraging residential energy conservation through feedback. Erlbaum.

A. Sen (1977). 'Rational Fools: A Critique of the Behavioral Foundations of Economic Theory'. *Philosophy and Public Affairs* **6**(4):317–344.

R. Sethi & R. Franke (1995). 'Behavioural Heterogeneity Under Evolutionary Pressure: Macroeconomic Implications of Costly Optimisation'. *The Economic Journal* **105**(430):583–600.

G. Seyfang (2006). 'Ecological citizenship and sustainable consumption: Examining local good networks'. *Journal of Rural Studies* **22**(4):383–395.

K. Sheehan (2001). 'E-mail survey response rates: a review'. *Journal of Computer-Mediated Communication* **6**(2).

J. Sheppard (1994). 'The Corporate Moral Person: The Organization's Personality and Its Board'. *Journal of Business and Society* **7**(2):151–164.

J. Sherow (1990). *Watering the valley: development along the High Plains Arkansas River, 1870-1950*. University Press of Kansas.

R. Shiller (1989). *Market Volatility*. MIT Press.

B. Simon (2002). 'Devolution of Bureau of Reclamation Constructed Water Facilities'. *Journal of the American Water Resources Association* **38**(5):1187–1196.

H. Simon (1945). *Administrative Behavior*. The Macmillan Company.

H. Simon (1979). 'Rational Decision Making in Business Organizations'. *American Economic Review* **69**:493–513.

H. Simon (1990). *Reason in Human Affairs*. Stanford University Press.

G. Skulmoski, et al. (2007). 'The Delphi Method for Graduate Research'. *Journal of Information Technology Education* **6**.

P. Slovic (1995). 'The construction of preference'. *American Psychologist* **50**:364–371.

A. Smith (1937). *The Wealth of Nations*. Modern Library.

A. Smith (1994). 'Incentives in community forestry projects: a help or a hindrance?'. Tech. Rep. Rural Development Forestry Paper 17c, ODI.

E. Smith & F. Conrey (2007). 'Agent-Based Modeling: A New Approach for Theory Building in Social Psychology'. *Personality and Social Psychology Review* **11**(1):87–104.

M. Smith (2008). 'Considerations for Agricultural to Urban Water Transfers'. Tech. rep., Arkansas Basin Roundtable Water Transfer Guidelines Committee.

M. Smith, et al. (2009). 'The economics of spatial-dynamic processes: Applications to renewable processes'. *Journal of Environmental Economics and Management* **57**:104–121.

V. Smith (1991). 'Rational Choice: The Contrast between Economics and Psychology'. *The Journal of Political Economy* **99**(4):877–897.

V. Smith, et al. (1988). 'Bubbles, crashes and endogenous expectations in experimental spot asset markets'. *Econometrica* **56**:1119–1151.

J. Smither, et al. (2005). 'Does Performance Improve Following Multisource Feedback? A Theoretical Model, Meta-Analysis, And Review of Empirical Findings'. *Personnel Psychology* **58**:33–66.

W. Smythe (1969). *WE Smythe*. The Conquest of Arid America.

C. Sneddon, et al. (2002). 'Contested Waters: Conflict, Scale, and Sustainability in Aquatic Socioecological Systems'. *Society and Natural Resources* **15**(8):663–675.

K. Sontheimer (2006). *Handbook of Contemporary Behavioral Economics*, chap. Behavioral versus Neoclassical Economics: paradigm shift or generalization? ME Sharpe.

G. Spaargaren (2003). 'Sustainable consumption: a theoretical and environmental policy perspective'. *Society and Natural Resources* **16**:687–701.

O. Spate (1959). 'The Hydraulic Society'. *Annals of the Association of American Geographers* **49**(1):90–95.

J. Specht (2008). 'Improving the Efficiency of Water and Energy Use in Nebraska's Irrigation Soybean Production Systems'. *NCESR White Papers and Presentations* .

R. Stavins (1995). 'Transaction Costs and Tradable Permits'. *Journal of Environmental Economics and Management* **29**:133–148.

R. Stavins (1998). 'What Can We Learn from the Grand Policy Experiment? Lessons from SO2 Allowance Trading'. *Journal of Economic Perspectives* .

R. Stavins (2000). *Public policies for environmental protection*, chap. Market-Based Environmental Policies. Resources for the Future.

K. Steiglitz, et al. (1996). *Market-Based Control: A Paradigm for Distributed Resource Allocation*, chap. A computational market model based on individual action. World Scientific.

J. Sterman (1991). *Managing a Nation: the Microcomputer Software Catalog*, chap. A Skeptic's Guide to Computer Models. Westview Press.

P. Stern (1999). 'Information, Incentives, and Proenvironmental Consumer Behavior'. *Journal of Consumer Policy* **22**:461–478.

J. Steward (1953). *Irrigation Civilizations: A Comparative Study*. Pan-American Union.

T. Stewart (1987). 'The Delphi Technique and Judgmental Forecasting'. *Climatic Change* **11**:97–113.

J. Stiglitz (2010). *Freefall: America, Free Markets, and the Sinking of the World Economy*. WW Norton and Company.

J. Stone & J. Cooper (2001). 'A Self-Standards Model of Cognitive Dissonance'. *Journal of Experimental Social Psychology* pp. 1–16.

R. Sugden (2000). 'Credible worlds: the status of theoretical models in economics'. *Journal of Economic Methodology* **7**(1):1–31.

J. Sumpsi, et al. (1996). 'On farmers' objectives: a multi-criteria approach'. *European Journal of Operational Research* **96**:64–71.

R. Sun (2008). *The Cambridge Handbook of Computational Psychology*. Cambridge University Press.

R. Susfalk, et al. (2008). 'Evaluation of Linear Anionic Polyacrylamide (LA-PAM) Application to Water Delivery Canals for Seepage Reduction'. Tech. Rep. 41245, Desert Research Institute.

R. Swart, et al. (2004). 'The problem of the future: sustainability science and scenario analysis'. *Global Environmental Change* **14**:137–146.

A. Tarlock (2001). 'The Future of Prior Appropriation in the New West'. *Natural Resources Journal* **41**:769.

A. Tarlock (2003). 'A First Look at a Modern Legal Regime for a Post-Modern United States Army Corps of Engineers'. *University of Kansas Law Review* **52**:1285–1327.

D. Tarlock & S. V. de Wetering (1999). 'Growth Management and Western Water Law from Urban Oases to Archipelagos'. *Hastings West-Northwest Journal of Environmental Law and Policy* **163**.

J. Taylor & S. Ryder (2003). 'Use of the Delphi Method in Resolving Complex Water Resources Issues'. *Journal of the American Water Resources Association* **39**(1):183–189.

O. Taylor & R. Luckey (1974). 'Water-Management Studies of a Stream-Aquifer System, Arkansas River Valley, Colorado'. *Ground Water* **12**(1):22–38.

R. Thaler & S. Benartzi (2004). 'Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving'. *Journal of Political Economy* **112**(1):164–188.

O. Thebaud & B. Locatelli (2001). 'Modelling the emergence of resource-sharing conventions: an agent-based approach'. *Journal of Artificial Societies and Social Simulation* **4**(2):3.

L. Thurstone (1927). 'Psychological Analysis'. *American Journal of Psychology* **38**:368–389.

T. Tietenberg (2000). *Property Rights, Economics and the Environment*, chap. Tradable Permit Approaches to Pollution Control: Faustian Bargain or Paradise Regained? JAI Press.

T. Tietenberg (2006). *Emissions trading: principles and practice*. RFF Press.

D. Tilman, et al. (2002). 'Agricultural sustainability and intensive production practices'. *Nature* **418**:671–677.

J. Tirole (2002). 'Rational irrationality: some economics of self-management'. *European Economic Review* **46**:633–655.

N. Torrecilla & J. Martinez-Gill (2005). 'The New Culture of Water in Spain: A Philosophy Towards a Sustainable Development'. *E-WAter* pp. 1–20.

N. Traore, et al. (1998). 'On-Farm Adoption of Conservation Practices: The Role of Farm and Farmer Characteristics'. *Land Economics* **74**(1):114–127.

E. Triana & J. Labadie (2007). 'Geo-MODSIM: Spatial Decision Support System for River Basin Management'. *Proceedings of the 2007 ESRI International User Conference* .

K. Troitzsch (2004). 'Validating Simulation Models'. In *Proceedings of the 18th European Simulation Multiconference*.

A. Turner & A. Penn (2002). 'Encoding natural movement as an agent-based system: an investigation into human pedestrian behavior in the built environment'. *Environment and Planning B: Planning and Design* **29**:473–490.

G. Turner (2008). 'A comparison of The Limits to Growth with 30 years of reality'. *Global Environmental Change* .

R. Turner (2005). *Sustainability: Critical Concepts in the Social Sciences*, chap. Sustainability: principles and practice. Taylor and Francis.

A. Tversky & D. Kahneman (1974). 'Judgment under Uncertainty: Heuristics and Biases'. *Science* **185**(4157):1124–1131.

A. Tversky & D. Kahneman (1981). 'The framing of decisions and the psychology of choice'. *Science* **211**:453–458.

A. Tversky, et al. (1988). 'Contingent weighting in judgement and choice'. *Psychology Review* **85**:371–384.

A. Tversky, et al. (1990). 'The Causes of Preference Reversal'. *The American Economic Review* **80**(1):204–217.

S. Upendram & J. Peterson (2007). 'Irrigation Technology and Water Conservation in the High Plains Aquifer Region'. *Journal of Contemporary Water Research and Education* **137**:40–46.

USBR (1973). *Fryingpan-Arkansas Project Colorado*. US Department of the Interior Bureau of Reclamation.

USBR (2010). 'Southern Delivery System Environmental Impact Statement'.

USEPA (1996). 'Wetlands, Oceans and Watersheds'.

USEPA (1997). '1996 Compliance Record: Acid Rain Program'. Tech. rep., U.S. Environmental Protection Agency.

USEPA (2003). 'Final Water Quality Trading Policy'.

J. van den Bergh, et al. (2000). 'Alternative models of individual behavior and implications for environmental policy'. *Ecological Economics* **32**:43–61.

M. Vandenbergh (2004). 'From Smokestack to SUV: The Individual as Regulated Entity in the New Era of Environmental Law'. *Vanderbilt Law Review* **515**:554–556.

M. Vandersande, et al. (2001). 'Tolerance of five riparian plants from the lower Colorado River to salinity, drought and inundation'. *Journal of Arid Environments* **49**(1):147–159.

T. Veeman & J. Politylo (2003). 'The Role of Institutions and Policy in Enhancing Sustainable Development and Conserving Natural Capital'. *Environment, Development and Sustainability* **5**:317–332.

A. V. D. Ven & A. Delbecq (1974). 'The Effectiveness of Nominal, Delphi, and Interacting Group Decision Making Processes'. *The Academy of Management Journal* **17**(4):605–621.

L. Vincent (2003). 'Towards a smallholder hydrology for equitable and sustainable water management'. *Natural Resources Forum* **27**:108–116.

T. Vukina, et al. (2008). 'Do farmers value the environment? Evidence from a conservation reserve program auction'. *International Journal of Industrial Organization* **26**:1323–1332.

D. Waddell & A. Sohal (1999). 'Measuring resistance to change: an instrument and its application'. *International Journal of Business Performance Management* **1**(4):353–367.

R. Wade (1988). *Village Republics: Economic Conditions for Collective Action in South India*. ICS Press.

M. Wallace & J. Moss (2002). 'Farmer Decision-Making with Conflicting Goals: A Recursive Strategic Programming Analysis'. *Journal of Agricultural Economics* **53**(1):82–100.

R. Weaver, et al. (1996). 'Efficacy of Standards vs. Incentives for Managing the Environmental Impacts of Agriculture'. *Journal of Environmental Management* **43**:173–188.

D. Webb, et al. (2000). 'Development and Validation of Scales to Measure Attitudes Influencing Monetary Donations to Charitable Organizations'. *Journal of the Academy of Marketing Science* **28**:299–309.

W. Webb (1981). *The Great Plains*. University of Nebrasks Press.

P. Werhane (1985). *Persons, Rights, and Corporations*. Prentice-Hall.

J. Wescoat (1986). 'Impacts of Federal Salinity Control on Water Rights Allocation Patterns in the Colorado River Basin'. *Annals of the Association of American Geographers* **76**(2):157–174.

K. Wick & E. Bulte (2006). 'Contesting resources - rent seeking, conflict and the natural resource curse'. *Public Choice* **128**:457–476.

O. Wiidegren (1998). 'The New Environmental Paradigm and Personal Norms'. *Environment and Behavior* **30**(75-102).

J. Wiklund, et al. (2003). 'What Do They Think and Feel about Growth? An Expectancy-Value Approach to Small Business Managers' Attitudes towards Growth'. *Entrepreneurship: Theory and Practice* **27**.

J. Wilkins-Wells & W. Epley (2003). *Urbanization of Irrigation Systems*, chap. Urbanization of Irrigation Systems. US Committee on Irrigation and Drainage.

J. Wilkins-Wells, et al. (1999). 'Operational Costs of Canal Companies and Irrigation Districts in the Intermountain Region'. In *Proceedings of the Western Agricultural Economics Association Annual Meeting*.

C. Wilkinson (1989). 'Aldo Leopold and Western Water Law: Thinking Perpendicular to the Prior Appropriations Doctrine'. *Land and Water Law Review* **24**(1).

C. Wilkinson (1991). 'In Memoriam, Prior Appropriation 1848-1991'. *Environmental Law* **3**(1).

J. Willock, et al. (1999a). 'The Role of Attitudes and Objectives in Farmer Decision Making: Business and Environmentally-Oriented Behavior in Scotland'. *Journal of Agricultural Economics* **50**(2):286–303.

J. Willock, et al. (1999b). 'Farmers' Atittudes, Objectives, Behaviors, and Personality Traits: The Edinburgh Study of Decision Making on Farms'. *Journal of Vocational Behavior* **54**:5–36.

A. Wils, et al. (1998). 'Threats to sustainability: simulating conflict within and between nations'. *System Dynamics Review* **14**(2):129–162.

P. Wilson, et al. (1988). 'Risk Perceptions and Management Responses of Arizona Dairy Producers'. *Journal of Dairy Science* **71**(2):545–552.

K. Wittfogel (1957). *Oriental Despotism*. London: Earthscan.

A. Wolf (1998). 'Conflict and cooperation along international waterways'. *Water Policy* **1**:251–265.

C. Woodka (2007). 'Area farmland vanishing at a steady rate'. *The Pueblo Chieftain Online* .

C. Woodka (2009). 'Protecting ag land key to state's health'.

R. Woodward (2003). 'Lessons About Effluent Trading From a Single Trade'. *Review of Agricultural Economics* **25**:235–245.

R. Woodward & R. Kaiser (2002). 'Market Structures for U.S. Water Quality Trading'. *Review of Agricultural Economics* **24**(2):366–383.

D. Worster (1982). 'Hydraulic Society in California: An Ecological Interpretation'. *Agricultural History* **53**(6):503–515.

D. Worster (1985). *Rivers of Empire: Water, Aridity and the Growth of the American West*. Oxford University Press.

R. Wuthnow (1991). *Acts of compassion*. Princeton University Press.

C. Yeh (2008). 'The effects of intelligence on price discovery and market efficiency'. *Journal of Economic Behavior And Organization* **68**:613–625.

R. Young (1986). 'Why Are There So Few Transactions Among Water Users?'. *American Journal of Agricultural Economics* **68**(5):1143–1151.

T. Young & J. Karkoski (2000). 'Green Evolution: Are Economic Incentives the Next Step in Nonpoint Source Pollution Control'. *Water Policy* **2**:151–173.

S. Zekster, et al. (2005). 'Environmental impacts of groundwater overdraft: selected case studies in the southwestern United States'. *Environmental Geology* **47**:396–404.

G. Ziervogel, et al. (2005). 'Agent-based social simulation: a method for assessing the impact of seasonal climate forecast applications among smallholder farmers'. *Agricultural Systems* **83**(1):1–26.

K. Zimmerer (2007). 'Agriculture, livelihoods, and globalization: the analysis of new trajectories (and avoidance of just-so stories) of human-environment change and conservation'. *Agriculture and Human Values* **24**:9–16.

E. Zube (1987). 'Perceived land use patterns and landscape values'. *Landscape Ecology* **1**(1):37–45.

F. Zybach (1960). 'Self-propelled sprinkling irrigating apparatus'. Tech. Rep. 2,941,727, US Patent Office.