LEARNING AND TRANSFER OF UNDERSTANDING IN DYNAMIC DECISION ENVIRONMENTS

by

Bent Erik Bakken

Lic.sci.éc., University of Lausanne (1981)

SUBMITTED TO THE DEPARTMENT OF MANAGEMENT
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

January, 1993 [FEB 1994]

Signature of Author
Sloan School of Management, January, 1993

Certified by
John D. Sterman, Associate Professor of Management
Thesis Supervisor

Accepted by
James B. Orlin, Chairman, Ph.D. Committee
Department of Management

LEARNING AND TRANSFER OF UNDERSTANDING IN DYNAMIC DECISION ENVIRONMENTS

by

Bent Erik Bakken

Submitted to the Department of Management in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Abstract

The thesis is based on findings in two distinct literatures, those of dynamic decision making and transfer of problem solving. The first line of research has hitherto considered dynamic complexity with little emphasis on the context in which the decision making takes place. The second body of literature has investigated how contexts shape subjects' transfer of understanding, but for the most part in dynamically trivial tasks. The experiments here integrate the two literatures by varying both semantic context and dynamic behavior.

The conditions for learning in dynamically complex real estate and oil tanker markets are described. Graduate management student took part in a sequence of two experiments; simulated oil tanker markets, unfamiliar to most subjects, and commercial real estate environments, more familiar to subjects. Performance in and transfer between simulated market conditions were investigated.

Findings indicate that subjects with some prior semantic notions about a market perform and transfer better than subjects in less familiar environments. Drawing additionally on a prior study that showed that highly familiar environments lead to poor learning and transfer in experiments it is suggested that the relationship between context familiarity and learning can be described by an inverted U-shape: Performance is helped by some familiarity, but since the compressed experimental dynamics allow a longer time horizon than that known by most experts, their expertise becomes a burden.

The experiment also manipulated the compression of the experiment. This was done by varying the period of the cyclical markets, and the results showed higher performance in the less compressed environment.

While performance was helped by transparent task conditions, transfer was hindered by the same environments: Context familiarity gave rise to high current performance at the cost of poor transfer to the subsequent setting. In general, however, transfer effects, were weak, indicating that exposure to simulated markets in a few-hour session is not likely to produce learning that will last. It is, however, suggested that reflective exercises may lead to better transfer to environments that also encourage reflectiveness.

In the experiments, decision as well as information acquisition behavior are monitored and related to subject background and performance. The questionnaire data corroborated performance and transfer findings. Subject background influenced first trial performance but could not explain performance in the second trial. The hypothesis that exposure to abstract frameworks found in system dynamics and economics should help transfer of understanding was not supported. Implications for experiential learning in dynamically complex real markets and the use of simulated decision environments to further such learning are discussed.

The findings suggest that designers of computer-supported learning environments must make a trade-off when deciding whether to use familiar or unfamiliar contexts. While highly familiar contexts often interfere with the learning desired by the designer, familiar environments also help transfer back to the workplace.

Thesis Supervisor: John D. Sterman
Title: Associate Professor of Management

Acknowledgements

This thesis is a result of my exposure to my feelings of difficulty in internalizing complex dynamic situations and the effects such cognitive shortcomings have on societal resource allocation. I thank Jay Forrester and Ragnvald Solstrand for pointing out that management research needs to be close to real world decision making. They both have argued that the quality of research as well as society benefits from such a close relationship; the present work is an attempt to follow those guidelines.

John Carroll has been a great help through his wisdom and knowledge about the vast field of decision making research. Paul Andreassen has been of invaluable support in shaping the experimental design into one that could be administered and analyzed.

I also thank Peter Senge for his many ideas of how to get organizations to be more effective in institutionalizing learning processes. He has broadened my professional scope. Hiring me as a research assistant in the program that today encompasses MIT's Organization Learning Center, he gave me an opportunity to meet and discuss with thoughtful real estate executives. Their insights are hopefully reflected in the thesis. John Morecroft and Peter Jebsen organized meetings with oil tanker professionals, whom a learned much from.

Many colleagues at MIT and NDRE have been of assistance. Anjaly Sastry, Per Lindell, and Ernst Diehl in particular provided support and commiseration. My office mate, Christian Kampmann, often took time to help me understand quantitative issues and helped me organize the thesis. Karen Hernandez' MS thesis was of great help, too.

Above all, John Sterman has been of invaluable importance. Not only did he allow me to use his simulation model as a foundation for the experiments, but he has also been most supportive of the questions addressed in the thesis. Providing feedback about substantial as well as form issues, successfully using both stick and carrot to tease out better and more insightful conclusions, giving insightful comments on the many drafts; he has been instrumental in the successful completion of the thesis.

1. Introduction

1.1 Motivation

There exist many situations where feedback from decisions is delayed or otherwise difficult to interpret. Business settings provide common examples. Not only are corporate profits hurting from this, governments often have to remedy corporate mistakes. As an example, losses in the dynamically complex US and Scandinavian real estate industries have forced American, Swedish and Norwegian governments to take over substantial negative bank assets. The total cost of these bail-outs has been over ten billion dollars in the small Norwegian economy of only four million people (Munthe, 1992). Interviews with real estate developers indicate that some of them realize that their judgments have been inconsistent and biased. Yet, reflections and assessment of past decisions appear not be widely shared among developers, investors and others. Decision makers consequently fail to learn from the experience of others and often even repeat their own errors.

The same delays that hinder learning in dynamically complex markets also create conditions for market instabilities and cyclical patterns (Kampmann, 1992; Blanchard and Fischer, 1989, chapter 4; Wheaton, 1990). Though empirical investigations in these markets can address some causes of fluctuations, the interactions of cognitive and market factors are more conveniently addressed in compressed experiments (Smith, 1982; Smith et al., 1987). Such experiments have confirmed that long time constants and supply lines create environments where effects appear unrelated to causes. Decision makers' incomplete mental models are a consequence of such environments; at the same time these incomplete mental models contribute to poor decision making.

The problem of making inferences in environments without reliable feedback has been widely discussed in the judgment and decision making literature (Kahneman and Tversky, 1986; Einhorn and Hogarth, 1978; Hogarth, 1981; Kleinmuntz, 1985; Sterman, e.g. 1989b; Sterman and Paich 1992; Diehl, 1992; Sterman and Kampmann, 1992). Yet, the degree to which simulated decisions may remedy poor real world learning has been given scant attention: In addition to serving as laboratories for researchers to investigate important macro-economic phenomena, compressed decision environments inevitably bring causes and effects closer in real time. Hence, laboratory environments may also help decision makers understand causes that underlie unstable markets, e.g. commodity markets.

It is the latter use of laboratory environments that is investigated in the present study. Drawing on findings in well defined experiments of algebra word problems and similar environments, it is suggested that the context familiarity with a decision environment will significantly influence understanding, performance and transfer processes. In addition, dynamic decision research suggests that short time constants and higher feedback immediacy should help performance in compressed environments.

In this study, subjects interact with a sequence of two trials so that residual transfer effects from the first to the second trial can be investigated in addition to performance in each trial. Subjects' background and understanding will also be monitored and used to explain subjects' performance.

As an example of poor learning, consider recommendations made by real estate specialists who analyzed the causes for the 1974-1976 real estate bust in the US. They suggested that a major cause was a prior over-investment due to a failure of investors to incorporate ongoing and planned construction activity. For instance, Conway and McKinley (1981)[1] stated:

> Hindsight is 20/20, and looking back, it is easy to see what some of the major mistakes of the early seventies were. In regard to feasibility analysis, it now appears that a common mistake was to analyze each project *independently of others*. The repercussions of 1974-76 have been so far-reaching that it is evident that the entire industry *has learned* a lesson. Among alert developers, the approach to every aspect of project planning will be more cautious.

Somehow, the hindsight could not have been 20/20. The losses faced by US and Scandinavian financial institutions (Kindleberger, 1988; Aftenposten, 1992) in the late 1980's and early 1990's were much worse than those of the early seventies. Yet, the Norwegian and US real estate collapses appeared just after new banking regulations were put in effect in both countries. Much public debate focused around the role of changes in banking regulations, leaving out the nature of the real estate market itself (Hernandez, 1991): Such collapses occur in markets where plentiful supply of credit coincide with rosy market predictions even in situations where credit regulations have stayed unchanged for several decades. This suggests that deregulation must have played a triggering more than a causative role. Hoyt (1933), studying land prices in Chicago in the nineteenth and early twentieth century, found recurring cycles of frantic new building, followed by market crashes even in the absence of regulatory changes.

Real estate markets are characterized by long construction lags so that commitments must be made long before the consequences of actions can be evaluated appropriately: Viewed as information feedback systems (Forrester, 1961; Sterman, 1985), real estate markets lack the direct and unambiguous cause to effect links that help people learn (Skinner, 1974; Brehmer, 1980; Einhorn and Hogarth, 1982). Including the time to obtain permits, office buildings may take more than 4 years to complete, at which time a borrower converts construction loans into mortgages and starts paying back his loans. Only at that

[1] Emphasis added

time will a real estate developer and lender have a first "reality check" about the soundness of the decision.

Unfortunately for the prospects for unguided learning, an increasing number of businesses has to deal with problematic action-consequence lags. This stems partly from new products that are built on technologies that require understanding of many related domains at the same time. Moreover, though many efforts are made to reduce development times for new products, their increased interconnectedness often lead to longer product development times. Yet, commitments to investments must be made long before the success of a product is known.

In addition, the organizational environment becomes larger: Market interdependencies increase decision complexity. Increased competition, and the continual breaking up of trade barriers, makes the term "home market" less and less relevant. The world market, with its many cultural and product differentiation facets, is increasingly the only relevant market.

The conditions for natural learning are threatened by other developments, too. The movement towards flatter organizations with fewer layers of management implies that there will be fewer training steps for a typical manager on the corporate ladder and less time for on-the-job management training (Schein, 1992). A flatter organization also implies that managers will have more subordinates and face a more interconnected work environment. Thus, the learning problems in real estate, important as they are, may dwarf those encountered in other industries.

The need to consciously shorten the time between causes and effects in organization is one of the main reasons behind the success of Total Quality Management (Kim, 1989). Many tasks, however, are not easily changed so as to make errors appear immediately after faulty actions have been made. The real estate market is but one example.

The lengthening of action to feedback delays may be a fairly recent phenomenon. For instance, our hunting and gathering forefathers lived in a more transparent world, at least dynamically. The hunter aimed at an animal and shot with his bow and arrow. Deficient decision making would immediately result in the lack of deer for dinner. Compare a hunter to a banker. Instead of instantaneous feedback, a banker receives ambiguous information about the appropriateness of actions taken. When making a loan, he or she gets an up-front fee, and immediate feedback is favorable regardless of the long term risk involved. More meaningful feedback, i.e. information about loan performance when the borrower must start repayment, becomes available several years after the loan has been signed.

Imagine that the loan is not repaid. Then the banker has to make sense of the conflicting evidence. Information about the fact that the loan appeared so beneficial to the bank must now be integrated with the hindsight that decisions produced poor outcomes. The banker has to make attributions about the causative process that produced the ambiguous signals. Such attributions are hindered by covariation in the interest rate, the general economy, the opportunity cost of capital, etc. Combined with long time lags this contributes to an ambiguous, messy causal structure.

On the other hand, training and preparation for complexity have changed since the days of the hunting tribe. Instead of learning hunting by practice in environments where feedback is immediate and unambiguous, we go to school for years in order to be prepared to interpret information in an increasingly complex environment. Similarly, Henry Ford's two-day training scheme to turn farm hands into factory workers has been replaced by advanced training programs for workers who already have 10 years' schooling before they enter a factory. Management commitment to such programs has been said to differentiate industry winners and losers (Womack et al., 1990).

One set of management training programs designed to deal explicitly with dynamic complexity is called the learning lab (Kim, 1989). This approach attempts to recreate real world complexity in terms of lagged relationships between causes and effects. This is done by using interactive simulation models that allow participants to make difficult decisions. Through time compression, dynamics can be "rehearsed" (Senge, 1990), much like a piano player rehearses. Labs may improve attributions regarding delayed relationships, unanticipated side-effects and other dynamic aspects of decision making by allowing participants to experiment with cause-and-effect hypotheses.

Though there is an emergence of research related to how one can embed labs in the organization (Senge 1990), the learning processes that take place during these learning labs are poorly understood (Paich and Sterman, 1992). What kind of learning environments should be chosen? Should they be contextually familiar or remote from participants' daily chores? Should the dynamic compression be high or low? These factors influence how participants will learn in the lab as well as how they transfer learning to the decision environment they return to.

1.2 Method

Several hypotheses will be investigated through simulation experiments that resemble learning labs. Subjects' performance data will be compared across various conditions. Experimental real estate and oil tanker markets were therefore formulated in high frequency (highly compressed) and low frequency (less compressed) dynamic conditions. The two industry contexts were made salient by similar introductory newspaper articles describing the current state of the respective industries. Graduate management students with 2-10 years' work experience participated. They were paid according to performance.

A real learning lab often extends for several days. The research design, though, had to be feasible with paid student subjects. Consequently the 41 students played for half a day, which enabled them to go through a total of 80 decision periods (years). The computer implementation, where subjects had access to various kinds of supporting data, enabled recording of subjects' information acquisition. Questionnaires monitored subjects' causal understanding.

1.3 Outline

Chapter 2 surveys the literature. The first part investigates the mechanisms behind learning and transfer in problem solving tasks. The second part surveys dynamic decision making, with special emphasis on how people learn to make better decisions when they interact with a task environment. Chapter 3 describes the experimental hypotheses of how context and frequency will affect performance, learning and transfer.

The decision contexts are explained in chapter 4. This chapter makes a strong case that conditions for real life learning in oil tanker and real estate markets are poor. The chapter further presents the experimental markets and highlights the different context and frequency conditions. Benchmark performance is also explained. Chapter 5 describes the results of the main treatments and discusses the findings. In chapter 6 these findings are related to additional data about subject background, information acquisition and understanding. Chapter 7 concludes and makes suggestions for the design and implementation of learning labs. It also comments on the implications of the findings for organizational practices, especially with respect to learning. Unanswered questions are addressed and future research laid out in this last chapter.

2 Learning and transfer in dynamic environments: A literature review

2.1 Overview

The first chapter explained that there is an increasing need for individuals and organizations to learn especially about dynamic complexity. The increased need does not, of course, by itself produce time and opportunity for reflection and learning. Future organizations must design and implement tools and processes aimed at developing learning (De Geus, 1988; Senge, 1990).

Learning efforts must be prioritized in order to succeed. This requires that one must classify those decision environments where learning will happen by itself and those environments where the conditions for learning are absent. Where the conditions for learning are absent, one may modify the decision and work environment so that people become more aware of errors and learn. In other instances, however, it will not be easy to change decision environments and therefore learning tools and processes must be designed to improve decision making.

Knowledge about the nature of cognition in static problem solving tasks can be used to help create a typology of decision environments based on how conducive they are to learning. This knowledge is mainly based on experimental studies of transfer. The main findings of these studies are explained early in the chapter. Likewise, the related issue of how people's framings of choice situations create judgmental biases is well documented and is also discussed. The role of information and action feedback in shaping decisions in dynamic environments has also been investigated and is presented later in the chapter. Work that deals with educational issues and attempts at improving performance is presented next. Last, the chapter addresses how learning and transfer in dynamically complex environments will be influenced by the context issues addressed in the problem solving paradigm. Similarly, the learning effects of dynamic patterns of behavior and how these patterns are related to underlying system (i.e. market) structure is also discussed.

2.2 Learning and transfer

In order to design effective learning environments it is necessary to understand the process by which people transfer. Transfer research, concerned with how people use knowledge from one domain in another. A domain can be a scientific field such as statistics. It may also be a more concrete context, such as the ones schoolchildren know from their algebra word problems. "Once upon a time, in a foreign land, there were three missionaries and three cannibals who should cross a river..." is an example of a more concrete context. The transfer research paradigm discusses difficulties people encounter when confronted with tasks that lack familiarity. It has investigated the mechanisms by which problem solvers chooses to start from scratch, or start with a previously created solution, when encountering new problems.

For the purposes of elucidating learning processes in and transfer from learning labs, it is of particular interest to investigate the degree to which prior knowledge interferes with learning in the lab. Furthermore, if learning takes place in the lab, it is important to understand the mechanisms by which lessons are transferred back to a work setting that is necessarily different from the learning laboratory. Hence, the transfer paradigm is of great interest.

2.2.1 Problem solving

When solving problems, people face significant processing limitations and do not go through extensive searches of their memory. Instead people use rules, heuristics, that are evoked by the initial conceptualization of the problem (Kahneman and Tversky, 1974). Heuristic search is the key concept in the Newell and Simon (1972) notion that people use production rules when solving problems.

Problem solving research has used transfer studies to elucidate these cognitive search mechanisms. Simon (1976) used an approach that has become paradigmatic in the problem solving field. Students solve problems based on a written description, such as school children's algebra word problems. In such instances, pupils first conceptualize the problem and later formulate an equation. Then they proceed to solve that formulation. Usually problems are structurally simple and defined in such a way that, once formulated, solving the equation is trivial.

The paradigm uses a sequence of two or more structurally identical, and contextually dissimilar, problems that can be framed with the same mathematical equation. Since one equation can solve both problems, the problems are called isomorphic.

The main question has been to understand how subjects use known solutions to solve isomorphic problems they encounter later. One would hope that the first solved problem could be a "cognitive springboard" for the next one so that a solution to a "Tower of Hanoi" problem is used when subjects encounter the isomorphic "Cannibals and Missionaries".

Early studies (Hayes and Simon, 1974; Simon, 1976; Kotovsky, et al., 1985) showed little evidence of transfer of learning. Subjects do not map the "Tower of Hanoi" equation onto the second problem. Instead, subjects tend to start from scratch again. They use fundamental mathematical building blocks instead just modifying a previous solution.

Lack of transfer can be explained by the nature of mental search processes. People operate in a mental space specific for each problem. Inside this space, rules govern operations on available information. These operations can transform words into equations. For a previous solution to enter the subspace where a new problem is solved, the relevance of the first solution to the second problem must appear to the problem solver. In other words, there must be pointers from the new subspace to a previously solved problem. If problem spaces are contextually rich, the cognitive search will go according to semantic, not mathematical, similarity. Thus, if solutions are classified in terms of the context in which they were used, the likelihood of finding mathematical solutions will be slim; for instance, mathematical equations for a first problem are categorized as solutions to how pegs should be organized in "Tower of Hanoi" and will not be brought to bear in an isomorphic "Missionaries and Cannibals" problem with a boat and people that need to be transported across a river.

Gentner and Toupin (1986) have extended these findings of poor transfer. Instead of investigating how the solution of one problem was applied to solving another, they let subjects solve a sequence of three contextually different isomorphic problems. They found, as did Simon, that little transfer occurred from task 1 to task 2. However, the third problem made a significant difference: A kind of "aha" feeling was activated and subjects used the previously solved problems to attack the new one and Gentner and Toupin suggest that the two earlier cases establish an exemplar that serves to guide the last solution.[2]

Experience is brought to bear when people develop links from general, yet concrete exemplars to the problem at hand. Exemplars are only developed after several exposures to isomorphic problems. Repeated exposures to isomorphic problems in a variety of contextual "disguises" can lead to the development of such exemplars. Once built, exemplars are readily used.

---

[2]A rival hypothesis for their finding is that after the second problem, subjects become conditioned to the experiment and might apply the solution to *any* new problem. However, no transfer took place if the third problem was structurally different from the first two. The study also tried to prompt transfer from the first to the second problem. Such prompting was in general unsuccessful. Indeed, the few students who were helped by prompting, voiced strong dismay; after the prompt, transfer became so obvious that it seemed impossible to the subjects *not* to detect the link between the first and the second problem.

Gentner and Toupin's work provides a sketch of what is required for transfer to take place:

1. An exemplar must exist or be built.

2. There must exist procedures that search for relevant exemplars as well as selection and testing criteria for relevant candidates.

The subsequent chapters will draw upon this idea of exemplars in dynamic decision environments.

2.2.2 Higher level frameworks

The teaching of general scientific principles should help students develop exemplars. Mathematics, economics and statistics are examples of frameworks that should be transferred across domains and ensure better problem solving (Papert, 1980).

Extending findings of problem solving transfer research, Bassok and her co-workers (1989 and 1990) compared how people first solved an "algebra (or physics)" problem and then attempted to solve an isomorphic "physics (or algebra)" problem. They found that people transferred better from algebra to physics than the other way around. Consonant with the argument presented here, Bassok et al. suggested that people expect algebra to be of general use in science-like problems. There are therefore links in subjects' problem spaces from a "science problem" to "equations". However, subjects expect physics knowledge to be domain specific and no search is generated in the "mathematical frame" for "physics solutions".

Transfer is dependent on subjects' possessing skills in the use of a generalizable mapping framework, such as mathematics. But to activate a framework, subjects need to look for transfer opportunities. Such search may be facilitated by application experience. Gilden and Proffitt (1989) and Proffitt and Gilden (1989) have investigated the role of application experience by investigating how physics professors and students solve complex freshman level physics problems.

Freshmen and professors initially performed similarly when confronted with a problem of collision dynamics that involved both rotational and translational forces by working through the problem analytically. While professors were able to quickly solve a second problem without resorting to pen and paper, freshmen did not possess the same level of math application experience, and could not transfer prior solutions to new problems. Professors had easier access to the original equations. These equations were brought into their problem space where they served to solve subsequent problems. Freshmen had to construct the second problem from scratch again.

For simpler problems that involved only one force dimension, no difference between students and professors was found, which indicates that while structural complexity may hinder the application of a prior solution, transparent problems cause easier access to established solutions.

Nisbett and his colleagues (Cheng and Nisbett, 1985; Kunda and Nisbett, 1986; Nisbett, et al., 1987; Larrick, et al., 1989) have investigated under which circumstances subjects apply statistical reasoning in everyday tasks. In a 1987 study, Nisbett, et al. compared how graduate students in law, psychology and chemistry used principles of statistics when solving problems.

The chemistry and psychology students had similar, substantial prior exposure to statistics; the law students had little knowledge of statistics. When confronted with an everyday problem that required use of statistical principles, however, chemistry students performed as poorly as law students. Psychology students were significantly more likely to use appropriate principles from statistics.

Nisbett et al. argue that the difference between statistics taught in chemistry and psychology classes lies in the domain area. Chemistry students get their training in a "science" domain. However, none of the everyday reasoning problems were related to "science." In addition, psychology students had also solved many everyday problems as a part of their statistical training.

These results can be explained in the following way: Subjects who learn statistics as a mathematical discipline and apply it to science will not apply the knowledge for everyday problem solving. Statistics courses must be augmented with application skills in everyday tasks, so that people establish links in their problem space from such tasks to the appropriate statistical framework. Again, transfer may fail because subjects fail to make links from the initial problem space to a meta-rule that tells them "look for an already solved statistics problem".

2.2.3 Shortcomings of the transfer literature

In an algebra word problem, subjects know that the problem has a mathematical solution. Therefore they tend to search for an appropriate equation. Experimental subjects are less likely to look for analogies to known solutions in such an experiment than in real life. In real-world dynamic decision environments, few people have any hope of finding the correct mathematical solution, and may use cognitive processes other than those captured by the problem solving research paradigm.

Few dynamic decision problems can be solved analytically, and those that can require analytical sophistication beyond high school mathematics. However, non-linearities make many dynamic tasks analytically intractable. When implemented as computer simulations, only heuristic methods can guide the search for good solutions. Likewise, decision makers in real tasks *must* resort to heuristic methods.

Even in environments slightly more complex than those studied by Proffitt and Gilden (1989) there is no evidence that people actually try to solve equations; though such analytic solutions exist (Mackinnon and Wearing, 1985).

Despite their failure to capture the full range of problem solving processes used in the real world, transfer studies have resulted in good documentation of the cognitive preconditions for transfer. It remains an empirical question whether the same strict conditions apply for transfer across contextually different dynamic tasks.

2.3 Learning and decision making

2.3.1 Judgement and decision making

While the above studies consider a single problem solved over a period of many minutes or hours, judgement researchers usually put subjects in simple choice situations. Such choices take place in seconds rather than hours. Since monitoring cognitive processes is almost impossible (Payne and Johnson, 1976), judgement research instead describes those environments that induce biased judgment and attempts to map out a typology of such environments.

Judgement and behavioral decision research has shown (Kahneman and Tversky, 1974; Slovic, et al., 1977) that people use simple decision rules that are systematically biased. A number of inconsistencies, biases and fallacies have been defined, such as the "base-rate fallacy" (Bar-Hillel, 1980) and the "hindsight bias" (Fischoff, 1975; Hoch and Loewenstein, 1989). Research into simple choices shows that subjects use different cognitive frames for structurally identical tasks. People treat losses differently than gains, and surviving is different from not dying (Kahnemann and Tversky, 1979).

Criticisms about the validity and robustness of findings have nevertheless been voiced. The fact that people pay attention to spurious cues might indeed be more indicative of the fact that scientists themselves use a biased repertoire of tasks (Cohen, 1982).

One argument states that if biases are dysfunctional, then individuals, organizations and markets will learn to do better. This may happen because people choose good behaviors over bad (Skinner, 1974), or because organizations that stumble upon good business strategies will have better chance of surviving in the marketplace. The latter argument, used by what one could call rational actor economists (e.g. Friedman, 1976), says that organizations that survive will have to mimic aspects of their competitors' behavior that generate the best outcomes, and so over time decision errors will disappear.

Though it is easy to dismiss findings of systematic decision errors by claiming that people learn, the claim is not very convincing without any evidence of how such learning will happen. Since judgmental errors are found with various degrees and types of expertise, there is little general support for the claim that experience makes errors disappear. Likewise, the fact that the entire banking industry used the same risky loan strategies in the 1980's, also indicates that questionable decision practices may survive for years or even decades (Hoyt, 1933), and creates problems for the organizational selection argument.

Another argument states that the environment provides quasi-continuous feedback about what constitutes good and poor outcomes (Hogarth, 1981). Frequent corrections in unambiguous feedback environments will cause good outcomes even though single decisions may be poor. The added effort required to make a good, non-biased, decision may not be worth the time and cost required (Kleinmuntz and Thomas, 1987). The corrective feedback argument underlines that many decision environments are dynamic. A banker making loans to many customers over a number of years provides an example of a decision maker operating in a dynamic decision environment. As stated before, however, this example also shows that occurrences of unambiguous environmental feedback may be less frequent than suggested by Kleinmuntz and Thomas' effort/accuracy argument.

Tversky and Kahneman (1987) recognize the shortcomings of the static, one-shot judgement approach. Still, they are hesitant to propose the study of dynamic environments. Such environments are harder to design (Slovic et al., 1976) and may just confirm or even increase decision biases. Kahneman and Tversky are not upbeat when discussing the prospects for learning in dynamic environments:

> "Effective learning takes place only under certain conditions: it requires accurate and immediate feedback about the relation between the situational conditions and the appropriate response. The necessary feedback is often lacking for the decisions faced by managers, entrepreneurs, and politicians because (i) outcomes are commonly delayed and not attributable to a particular action; (ii) variability in the environment degrades the reliability of the feedback, especially where outcomes of low probability are involved; (iii) there is often no information about what the outcome would have been if another decision had been taken; (iv) most important decisions are unique and therefore provide little opportunity for learning (see Einhorn and Hogarth, 1978). The conditions for organizational learning are hardly better. Learning surely occurs, for individuals and organizations, but any claim that a particular error will be eliminated by experience must be supported by demonstrations that the conditions for effective learning are satisfied."

If feedback from decision environments should keep a decision maker on track in the face of biased decision making, this feedback must be easy to detect and interpret. The quote above suggests many instances in which feedback transparency is low. Certainly, transparent environments exist, and an ecology of feedback environments must be mapped out before one can argue whether they tend to be transparent or not.

The issue of feedback in decision making also raises theoretical questions of what feedback is. One can distinguish between outcome and action feedback. The common use of the term sees feedback as the use of information about the accuracy of a past action or prediction. Outcome feedback is the term used in these situations (Brehmer, 1980).

Another view understands feedback to be information that becomes available as people act and the system responds. This is called action feedback. It is generated by an action or intervention, and consists of information about the system and about the effect of the interaction, which by definition changes the environment (Diehl, 1992).

Outcome feedback is the theoretical feedback construct used in Multiple Cue Probability Learning (MCPL). MCPL has focused on the learning of multiple correlations as a function of experience, noise and the direction of relationships. MCPL builds on the lens model of the psychology of inference (Hammond, et al., 1973) and gives an observer several accesses to the state of the system (Brehmer, 1980), though these are time neutral. A finding has been that subjects are poor at setting up good learning strategies. They have strong prior beliefs regarding relationships and therefore detect positive relationships more readily and in much noisier environments than negative relationships. Curvilinear relationships are only detected in virtually noise-free environments. The fact that they use an inefficient, confirmatory, decision strategy, also hinders learning.

Action feedback deals with tasks that evolve over time: Dynamic decision making tasks.

2.3.2 Dynamic decision making

Only recently have action feedback tasks received systematic attention (Diehl, 1992). The research has asked the question of what cause tasks to be perceived as "opaque" or "transparent" to decision makers (Brehmer, 1988; Kampmann and Sterman, 1992). Results show that human information processing and task characteristics interact to induce decision rules with systematic misperceptions. While subjects do well when the task and/or information feedback is transparent (Mackinnon and Wearing, 1985), internal task feedback and feedback to decision makers often lack such transparency.

Systematic deviations from reasonable decision behavior have been revealed. Poor decisions are not only a problem in experiments; convincing arguments can be made that faulty decision making contributes to business and societal problems. Managers' lack of attention to ordered, yet undelivered goods and assets (Sterman, 1989a and 1989b) may contribute to overall economic instability. As will be discussed further in chapter 4, it appears that environmental cues lacking salience are seriously underweighted when subjects make decisions. Supply lines of commitments to the creation of physical assets are therefore prime importance in dynamic decision making. Supply lines are by their nature related to the time-lag between commitment to action and the consequences of, and thereby feedback about, that action. They also play a crucial role in determining system stability.

The very slow convergence to equilibrium of a system containing time-lags (Kampmann, 1992) may well create oscillatory modes of various kinds in the economy. The lack of subjects' understanding of side-effects (Diehl, 1992) may create dysfunctional outcomes in public policy. Moreover, decision makers appear not to persevere in their hypothesis testing strategies. Yet perseverance is particularly necessary in the real world, where feedback immediacy is rare (Dörner, 1980). This may be why politicians and business managers "vagabond" from one strategy to another without obtaining significant understanding of causal mechanisms.

In fact, there are so many misperceptions in dynamically complex scenarios, that one is hard-pressed to understand that people can set satellites into orbits. Yet precisely the circumstances where people do well must be described, not only those where subjects do poorly (Toda, 1962).

Though decision making errors may disappear with guided experience and reflection, little research has focused on the improvement in decision strategies. Since the aim of decision research has been to document consistencies in decision making errors, changes in decision making with experience have not been revealed.

For instance, Sterman has found consistent decision making in an experiment where subjects made capital acquisition decisions in a simulated economy with substantial time-lags and non-linear fulfillment dynamics. He recreated subjects' decisions with a model of a plausible rule (1989a, 1989b). The same rule can be fitted to most subjects' decisions. Rule parameters were consistent for each subject over the 36 period trial. However, some subjects do not fit the rule very well initially, but Sterman finds that rule consistency is high in the latter part of the sequence. This indicates that rule consistency may increase with experience.

In an attempt to clarify what causes high performance, Bakken (1989a) investigated written reports in the same task. He found that subjects who performed well also mentioned structural and equilibrium features. Though his study was a between-subjects design, it suggests that improved decision performance may result from increased understanding of the feedback structure underlying the problematic system behavior. This again suggests that decisions evolve as subjects' gain experience with a decision environment, something that was corroborated in Bakken (1989b) using subjects' decisions in same simulated economy. He found that performance improved over trials. Fitting a rule with decision weights as variables, he found that these weights evolved towards less biased supply line control, but subjects did not come close to benchmark decision weights used by Ozveren and Sterman (1987).

Kampmann (1992), investigating market stability as a consequence of pricing regimes and production lags, found evidence that subjects use a simple anchoring and adjustment rule. A market that starts out of equilibrium will converge at different speeds, depending on the market clearing and feedback regime used. He found no indication that decision rules change with experience, however.

Dynamic decision making and learning are closely linked. First, a dynamic task requires repeated decisions and so enables learning to be investigated. Secondly, causes of poor decision making may be the same as those that hinder learning: The opacity of a task at one single instance may act so as to induce poor decisions and may make the task even less transparent and perpetuate misperceptions.

Paich and Sterman (1992) investigated human performance in a product lifecycle task by allowing subjects to play a sequence of 5 simulation trials of varying dynamic complexity. Modeling subjects' decision rules, they found evidence that performance suffered as dynamic complexity increased. While experience improved performance on average, the negative effect of feedback complexity on performance was not mitigated by experience.

Brehmer (1988) found that action lags decreased performance in a simulated forest fire, at the same time conditions for learning got worse. While subjects in low lag conditions started out well and improved on each of six subsequent trials, subjects in long lag conditions only improved for 3 trials even though their initial performance was so poor that they had a larger potential for improvement.

In summing up the findings in what is has evolved into a Dynamic Decision Theory (DDT) with relevance to learning, one can argue that

• People seek confirmation for their theories (Einhorn and Hogarth, 1978), and as a consequence they are often stuck in severely suboptimal decision strategies (Dörner, 1980; Sterman, 1989b). Decision makers do not seek out alternative strategies when they are satisfied with outcomes, especially when it would take a long time to test newly generated hypotheses.

• More often than not, decision makers in dynamic environments underestimate or ignore dynamic processes. As a consequence, they leave out concerns for side-effects and self-reinforcing dynamics (Dörner, 1980; Brehmer, 1987; Sterman, 1989b; Fuglseth 1989).

• People fail to adjust their decision strategies to account for delays in the system (Bakken, et al., 1992) and expect feedback to arrive before the system can provide such information.

• Decision makers go into dynamic scenarios with inappropriate scripts based on apparent task characteristics, and they make little or no attempt to challenge the appropriateness of these scripts (Kleinmuntz and Thomas, 1987).

As a result, learning may not take place if assumptions and strategies are not challenged from inside or outside the decision environment (Schön, 1983; Salomon, et al., 1991b). To challenge improper beliefs people have about causal relations have been a major focus in improvement research. This is treated in the following section

2.4 Education and improvement research

This section deals with approaches designed to improve decision making. Some are focused on classroom teaching, whereas others are designed with the professional in mind. Especially the latter uses training sessions that also take into account organizational context. Understanding transfer of insight across contexts is critical to the educational community. Similarly, the role of transfer must be understood by designers of learning laboratories. This section surveys and critiques relevant education research and efforts aimed at improving decision making.

Salomon and colleagues (Salomon, 1987; Salomon and Globerson, 1987; Salomon and Perkins, 1989; Salomon, et al., 1989; Salomon, et al., 1991a; Salomon, et al., 1991b; Salomon, 1992) have shown that the ability of high school students to transfer requires use of a higher order skill of searching their own problem spaces for possible solutions. This skill, which Salomon et al. call "mindfulness", can be taught.

Another part of the literature deals with the value of pedagogically oriented simulations, but often lacks the theoretical orientation of the transfer studies (Graham and al., 1989). Accordingly, with a few notable exceptions (see Vennix, 1990 and Teach, 1990), simulation research has mainly documented single instances of use of simulation models, where one speculates about causes of success and failure as in Kreutzer et al. (1992).

Studies are often reported in the business school and gaming literature (see Raia (1966) for an early, but still very useful, conceptual guide), and tend to focus on how business simulations fare as class exercises against written case studies and lectures (Wolfe, 1976 and 1985). A finding has been that students report more enthusiasm about simulations than about case reading. Learning improvements resulting from simulation approaches may be attributed to motivational side-effects of the interactive pedagogy.

By extending the simulation paradigm with tools and processes based on ideas from other fields, training sessions have been carried out designed to help participants access a deeper level of understanding, i.e. understanding that is applicable across environments. Such sessions aim at helping participants learn to improve decision making. Cognitive feedback, learning labs and double loop learning are but three systematic approaches that use simulations as vehicles for learning and insights. The reminder of this section will investigate each approach in more detail.

The cognitive feedback approach (Hammond, 1978; Cooksey, 1986; Steinmann, 1976) starts with a normative model of how people should make decisions. The approach first helps participants realize and later adjust their own inappropriate weighting of decision cues. People in a typical learning session are confronted with their own weights as evidenced by their decision making. The contrast between the actual and normative model indicates how these weights should be adjusted.

Questions exist about the robustness of the learning taking place. In particular, when the graduate of a cognitive feedback seminar returns to the work environment, then the normative decision model is relevant only to the extent that decision problems encountered in the job are similar enough to those encountered in the seminar.

If one assumes that the decision problems facing a typical politician or administrator change all the time, the cognitive feedback approach requires that someone continually create and update normative models. Another criticism of the approach is that it assumes that learning consists in achieving appropriate weighting of cues. However, as was pointed out in the section on problem solving, a main problem tends to be that initial framing is inadequate; in other words key cues may be omitted. It does not help very much to "fiddle" with decision weight parameters if the model is wrong. (But see also Dawes, 1979 and Kleinmuntz, 1990 for the view that simple models are robust.)

The learning lab approach (Kim, 1989; Moissis, 1989; Bakken, et al., 1992; Senge and Sterman, 1992) uses a different point of departure. Both this and the cognitive feedback approaches are base on models of decision problems. In a learning lab, however, the model is used to highlight differences between good and undesirable decisions and consequent system behavior.

The key goal in the learning lab approach is to challenge decision makers assumptions about an underlying phenomenon. Thus, the model serves as a vehicle for "expanding thinking" more than as a cognitive feedback representation of a string of decisions and weights that should be used. In addition to challenging existing assumptions, the learning lab lends itself to exploring dynamics that tend to make real organizations poor learning environments.

Models can be rich in representations of side-effects that are difficult to perceive in the real setting. Both model richness and the fact that learning sessions typically run for many simulated years enable a focus on delayed consequences of actions. Moreover, the lab is designed to encourage experiments that allow new hypotheses to be generated and tested, decision making takes place in a laboratory setting. The compression of time and space as well as tan emphasis put on generation of new ideas may well help overcome some of the learning deficiencies of real decision environments.

Just like the learning lab approach, double loop learning (Argyris and Schön, 1978; Schön, 1983) starts from the premise that organizational processes, such as defensive routines, reduce actors' abilities to generate insights into problematic issues. While the learning labs tend to focus on people's inability to grasp causal relationships and on improving inappropriate understanding about the relationship between structure and behavior in dynamically complex environments, the double loop learning approach focuses on the implicitness of assumptions. By making explicit hidden assumptions, communication and feedback may improve. Also, openness can generate more hypotheses about causal relations.

In absence of diagnostic feedback, decision makers have no way of knowing that they are making erroneous inferences. One may say that people and even organizations fall into decision traps. In order to climb out of such decision traps, they must generate hypotheses that elicit meaningful feedback (Weick, 1977). Moreover, in the "Argyris-type" workshop, process feedback is typically given interactively. Numerous studies in other areas of inquiry have shown that performance, i.e. outcome, feedback can improve behavior, especially if it is unambiguous and provided without delay (Greller, 1980; Nadler, 1979; Tierney, et al., 1986).

Double loop learning sessions encourage feedback exchange and inquiry into causal mechanisms, and aim at improving both quantity and quality of information exchange. In addition, the approach stresses that mental frameworks are open to questioning, and so emphasizes that views can be challenged and changed. The importance of generating hypotheses that suggest solutions different from those already in use, is often pointed out. The exact mechanism by which people use new assumptions is unclear, and little empirical research exists to elucidate whether participants in workshops actually transfer and use the insights in their daily decision environments. A notable exception is Putnam (1989). Using an anthropological approach, he finds that people actually use inquiry skills in their daily decision environments after they have participated in such workshops.

2.5 Discussion

The nature of cognition during transfer of problem solving as well as the influence of feedback structures and parameters in dynamic decision making have been described.

While the transfer and judgment literatures explicitly focus on how problem context hinders the application of correct analogies in static environments, the role of context in experimental dynamic settings has not been investigated. Contextual cues in dynamic decision making could produce worse or improved decision making: The familiarity of a dynamic context should relieve a decision maker from the stress of both remembering names of variables and uncovering a difficult causal structure. This cognitive relief should produce better decision making.

At the same time, a familiar context will also evoke a dense problem space that may contain scripts that act as filters to prevent insights into the structural information. The lack of emphasis on context is especially troublesome in learning laboratory research. In such settings, anecdotal evidence suggests that learning in the laboratory is hindered when have detailed knowledge about the simulation context. In such cases, they may find it difficult to map their own experiences onto the simulation model, which is necessarily a different, and usually a less detailed representation of reality. Yet transfer back to the workplace may well be hindered by too great a leap between laboratory and real world contexts.

Simulation characteristics, i.e. system structures, influence decision making in dynamic tasks. As mentioned, the role of feedback structure in the improvement of performance changes has not been investigated. One may argue, though, that those factors that cause systems to lose transparency, such as delays, side-effects and non-linearities, also may influence learning in a negative way.

The role of system structure in creating learning opportunities is complicated, however. A subject whose decisions show lack of attention to such feedback cues as physical assets under construction will produce a different system behavior than a subject who incorporates the supply lines' tricky structure. In the case of the multiplier-accelerator task (Sterman, 1987), for instance, good decisions produce lower amplitude and faster returns to equilibrium than a poor decision rule. Thus, the sequence of good decisions produces a faster unfolding of the dynamic behavior that again may influence learning.

The lack of context focus in Dynamic Decision Theory, the lack of dynamic problems in the transfer research, and the insufficient focus on learning in dynamic tasks, together suggest a research agenda that can combine the literature and our insight in several ways. The hypotheses and experimental design proposed in the next chapter constitute a contribution and will link previous research on decision making with the new agenda called for here.

3 Experimental hypotheses, approach, and method

3.1 Introduction

The research presented in chapter 2 has elucidated cognitive mechanisms used during problem solving, showing evidence that subjects are influenced by contexts when they search mental problem spaces. Investigations about structural features in dynamic environments, such as delays, side-effects and non-linearities, have indicated poor performance resulting from dynamic complexity. Yet, as mentioned, the research has several shortcomings that can be overcome by investigating two task dimensions: Context familiarity and number of recurrences of a problematic phenomenon, i.e. compression.

Tasks that at the same time are dynamically complex and appear in well-defined contextual disguises make it possible to address theoretical questions of learning and cognitive transfer processes in dynamic environments. Answers to these questions also have value in learning lab design.

This chapter is divided as follows: The first section defines learning and transfer in dynamic tasks. The following section suggests main hypotheses and their operationalization. Thereafter follows an account of how supplementary measures are collected. The experimental implementation is sketched out next. The last section recapitulates the design and the hypotheses.

3.2 Approach

The research approach attempts to clarify what it is that helps and hinders performance and application of understanding in dynamic decision environments. One main explanatory variable for performance and changes in such performance will be transfer from one domain to another. The concept of transfer thus needs further refinement along several dimensions.

In experimental settings performance is dependent on subjects' familiarity with testing instruments. In the approach used here, the decision environment is at the same time a testing instrument. Though much effort has been put into making the user interface as intuitive as possible, a subject needs to become familiar with this instrument, too. Since the sequence of two trials uses user interfaces that differ only in contextual aspects the added familiarity with this instrument will of course help subjects perform better in the second sequence of decisions. Learning of this kind of surface structure is and not of interest because it relates only to superficial aspects of the user interface.

Yet, subjects can make analogies to the deeper structure of other, similar real world and simulated tasks. If there are candidates for such analogies, subjects must sort out their appropriateness to the task at hand and false candidates must be rejected. As we saw in the previous chapter, subjects may already possess appropriate schemas, which are candidates for reasoning by analogy. Subjects must establish their appropriateness to task at hand and reject false candidates. Yet potential exemplars often fail to appear as such, and this is what makes transfer between isomorphic problems so complicated.

Analogies may be classified along several dimensions, but the research reported in the previous chapter indicated that subjects tend to use context as a main source of classification. Analogies could also be classified along the behavioral or structural dimension. In fact, Bassok's research (1990) suggests that the structural dimension works as a prime classifier for physics professors.

Note that the distinction between structure, i.e. the causal mechanisms, and behavior, which is the manifestations of those mechanisms, is very important in dynamic decision environments. In fact, one of the main inference problems in dynamic decision tasks is precisely the counterintuitive relationship between structure and behavior (Forrester, 1970) so that even when the structure is laid out to subjects, subjects do not make the connection between causal feedback structure and corresponding system behavior. Yet in transfer of problem solving tasks there is no such distinction. Once a problem's causal structure is formulated in equation terms, its solution does not present any difficulty.

The research approach must focus on a limited set of transfer issues. It was mentioned that familiarity of the testing instrument is of little interest, and its role must be kept to a minimum. On the other hand, analogies to prior experiences are likely to interfere with learning in the lab and so should be made explicit. This is done by creating familiar and unfamiliar contexts. Likewise, the influence of a structure's behavioral manifestations should also investigated.

Clearly, subjects' understanding of the behavioral manifestations of a structure, their understanding of the structure as well as the relationship between the two is of interest here. Do subjects perform well or poorly because of a lack of attention to the one, the other or because they do not understand how the two interrelate?

The approach used here, a sequence of two similar trials, enables an assessment of performance as well as the impact of prior treatments on later performance. The latter is called transfer. These definitions of performance and transfer do not by themselves distinguish between the various levels of transfer, so there must exist manipulations outside the performance and transfer metrics that address whether analogies are purely related to measurement instruments and thus uninteresting, or whether they also address the interesting issue of the role of contextual analogs and behavioral manifestations.

There are several ways to look at performance and transfer. The choice between various transfer metrics is explained in chapter 5, the results chapter.

Questionnaires and information acquisition data enable investigation of cognitive processes. Comparing performance and process data is important, since they often provide different views of learning and performance processes. Though performance and structural understanding may co-vary in complex dynamic environments (Bakken, 1989a), often they do not (Broadbent, 1978; Broadbent and Ashton, 1986).

A laboratory approach to the study of human decision making has shortcomings. In particular, one must be careful about generalizing from the contrived learning that takes place during an afternoon in a learning lab, with the processes that lead to learning over decades in the real world. As in any experimental laboratory study, the extrapolation of findings to help improve organizational and social policies is problematic, but the high task complexity as well as the use of meaningful market contexts should make this study more amenable to recommendations than most lab experiments. In addition, validity of the results for how people learn in simulated decision environments is high; the experimental decision environments are very similar to the simulations used in learning labs.

3.3 Main hypotheses

A total of four independent and two derived hypotheses are tested. They all relate to the context and compression of the decision environment factors that are manipulated in the experiment. They are stated first and explained in more detail below.

Hypothesis 1

Subjects perform better in a familiar context.

Hypothesis 2

Subjects transfer better from a familiar context.

Hypothesis 2.1

Subjects transfer better to an unchanged context.

Hypothesis 2.2

Subjects transfer better to an unchanged environmental compression.

Hypothesis 3

Subjects perform better in a compressed environment.

Hypothesis 4

Subjects transfer better from a compressed environment.

The hypotheses are derived from the findings reported in chapter 2, yet they have been defined so as to address questions of interest to learning lab designers, as well.

The first hypothesis is stated as "subjects perform better in a familiar context." This hypothesis is not directly tested in any of the studies cited in chapter 2, yet many findings in static contexts can be interpreted to support it. In Nisbett et al's work (Fong and Nisbett, 1986) the approach of investigating performance in a static, everyday problem solving task assumes differences in context familiarity. Indeed the finding that psychology students do better in the experiment is partly explained by the fact that psychology students are more familiar than chemistry students with everyday tasks.

Yet, in dynamically complex and contextually rich tasks there is evidence (Bakken et al. 1992) that domain experience actually hinders performance. Anecdotal evidence suggests that when learning lab participants possess a rich knowledge base about the domain used in the lab, this knowledge may interfere with the new task information to prevent laboratory learning from taking place. In addition to this cognitive interference, subjects with long experience in a domain may feel threatened and become defensive about a simple game portraying a few aspects of their own business, especially if it focuses on other aspects than those decision makers are attuned to. Decision makers' lack of intuition in these aspects may create defensiveness, decrease openness to new ideas, and prevent learning from taking place (Isaacs and Senge, 1992).

Hypothesis 2 stated that "subjects transfer better from a familiar context". This hypothesis deals with the learning that takes place in the lab, and the interaction with the mental models subjects have developed outside the lab. Transfer from a familiar environment should be more efficient than from an unfamiliar one, since familiarity will ensure that problematic structural issues receive more attention than in an unfamiliar environment, where cognitive resources will have to be shared between contextual novelty and structural issues.

Yet, the argument that contextual familiarity should help transfer faces a counter-argument. Though learning in an unfamiliar context is hard, such learning may make it more probable that the little learning that does take place will be less contextual and more structural in nature. If this is true, one may find that poor performance in the familiar environment will coexist with higher transfer performance.

The second hypothesis is elaborated so as to include two related conjectures. The first sub-hypothesis states that "subjects transfer better to the same context". Regardless of whether the first context was familiar or not, this first context will become more familiar when the subject interacts with the decision environment. Thus, if the second context is the same as the first one, the second environment will be familiar. As a consequence, the same arguments that supported the conjecture that context familiarity helps performance, i.e. more available cognitive resources for structural issues, and less likelihood that pointers point to unhelpful areas in the problem space, can be brought to bear for the transfer case.

There is, however, a rival hypothesis that indicates the opposite effect that changes in context should help transfer: Subjects exposed to one context may have an idea about the causes underlying dynamic behavior. In a changed context they will be more likely to replace irrelevant contextual knowledge with informative structural understanding. Subjects transferring to the same context have no similar reason to question their beliefs, and the likelihood of a mental clean-up is smaller.

The second implication of hypothesis 2 concerns the compression in the environment. While there is no contention that subjects have more familiarity with one of the two compression conditions, the first trial will establish familiarity with the structure and hence compression inherent in the task. Few studies directly assess the role of structural, i.e. frequency, change on transfer performance.

The third hypothesis states that "subjects perform better in a compressed environment." The decision environment in a learning laboratory is a simplification of reality where a few select issues may be addressed. As a consequence, learning environments are contextually less dense than the real world; fewer variables are included. Similarly time is compressed so as to enable many decision strategies to be carried in a typical two day session. Game designers and learning laboratory facilitators have the option of choosing between a high compression, and having the interesting behavior unfold more slowly. In a compressed environment, the decision environment changes more from one decision period to the next. More instances of a problematic behavior can be experienced for a given number of decisions. Such "fast" unfolding may lead to more effective learning. But if the dynamic unfolding is too fast, subjects will not be able to realize the effects of these changes and hence causal inferencing may suffer.

For a given system and a given simulated time reference, there is often a relationship between the system's time constants and its natural period, so that a system with long delivery delays and little depreciation will induce a longer period of oscillation than a system with shorter time constants. Thus, the highly compressed scenario can be arrived at by choosing short delivery delays and depreciation time constants. Though the dynamic decision literature has little direct reference to system frequency or dynamic compression, information and delivery delays influence decision behavior. Subject control and hence performance decreases with longer time lags (Diehl, 1992; Kampmann and Sterman 1992).

However, structure may have a dual effect: Time lags decrease decision control and performance, yet the same long time lags may enable subjects a "slow motion" view of the environment. This slower unfolding of dynamics may again further insight and learning.

The dynamic decision making literature shows higher performance in short time lag than in long time lag settings. In such environments markets will be more compressed, i.e. exhibit higher system frequency. Similarly, anecdotal evidence from learning labs also suggest that learning labs where subjects go through many instances of a problematic behavior are effective. Performance will improve with higher compression as long as the time compression is plausible.

The fourth hypothesis states "subjects transfer better from a compressed environment". This hypothesis is supported by the common observation that repetition is conducive to performance. The necessary condition for transfer must be followed by the sufficient condition that the environmental feedback must provide cues about the underlying phenomenon to be understood. The question of whether feedback in the complex, dynamic environment is decipherable has not been addressed in similar environments before. The reverse hypothesis of poor learning from compression can be supported by the argument discussed under the third hypothesis, namely that feedback transparency increases with slower unfolding of dynamics as the decision environment changes less from period to period and changes become easier to understand.

3.4 Experimental design and procedure

3.4.1 Overall design and implementation

The hypotheses were investigated using a factorial design with 4 factors. These were context, frequency, change in context and change in frequency. Each factor was administered in two levels:

Context: Familiar, unfamiliar.

Frequency: High compression, low compression.

Context change between trials: No change, change.

Frequency change between trials: No change, change.

The compressed environment contained four peaks of high prices, while the less compressed environment contained two such peaks. The three factors of frequency, frequency change and context change were straightforward experimental manipulations, but the context familiarity represented an important experimental choice. The commercial real estate industry served as the familiar context and the international oil tanker market as the unfamiliar context. The definition of real estate as more familiar than oil tankers was made on the following grounds:

First, the importance of the two industries to the US economy is very different: While real estate fluctuations in the 1985-1992 period caused one major bank to fail after another, oil tankers were hardly represented on any major bank's balance sheet.

Consequently, the popular business press, which was followed closely in the 1987-1992 period, had more than 20 articles about the real estate industry for every article about the oil tanker business. In addition to articles about the industries, there were also articles about other aspects of the assets. In particular, the oil spill from the supertanker Exxon Valdez in Alaska the spring of 1989 received much attention. However, this attention largely vanished a year before the experiments took place and should not have influenced subjects much. It is reasonable to conclude that management students would be much more exposed to real estate than to oil tankers. In addition, most students were personally involved in the real estate industry through their own apartments and condominiums and those of friends and family. No similar personal experience can be expected about oil tankers.

In addition to experimental treatments affecting performance and transfer, previous research indicates that there exist measurable factors that contribute to subject understanding, performance and transfer. In particular, Nisbett et al (1986) as well as Bassok (1990) and Proffitt and Gilden (1989) all indicate that academic training influences subjects' framing of situations.

The experimental markets have been developed using building blocks from System Dynamics and Control Engineering (Forrester, 1961; Friedland, 1986). System Dynamics and Control Engineering background should help performance because these educational programs stress the dynamic nature of systems and the factors contributing to instabilities. Transfer should be high, because the conceptual framework may act so as to filter out contextual information and help subjects draw attention to systemic features. Similarly, background in Economics should help subjects understand the underlying forces behind the task and increase both performance and transfer.

Background in fields conducive to understanding the task should improve performance. Thus, courses in Finance and Accounting will make understanding of interest payments and depreciation rules easier. Corporate strategy courses that focus on long range issues should help prevent a short term focus. Similarly, Computer Science background should make subjects more comfortable with computers and thus alleviate the fear of making technical blunders and increase performance for two reasons: Familiarity with the testing instrument will be high and this familiarity may alleviate stress that tends to limit decision repertoire and transfer.

Though graduate management students were used as subjects, there was some variance in prior work experiences. Thus performance and transfer could be related to context and length of prior work experience.

In the next chapter market compression is operationalized as frequency of the market. Since frequency = 1/period, the compressed market showed a shorter period of instability than the less compressed market.

The context conditions are called F (familiar), UF (unfamiliar), and the frequency treatments are called H (high) and L (low). Each market will be of one of the four types: FH, FL, UFH, UFL. *Performance* is the notion used to measure how well subjects did in each of the four markets. Performance will, according to the hypotheses, be high when context is familiar and/or when compression is high.

The sequence of treatments will be one of the 16 types FH_1 FH_2, FH_1 FL_2, FH_1 UFH_2, FH_1 UFL_2,,,,, UFL_1 FL_2, UFL_1 UFL_2 where subscripts denote trial number. The metric that denotes residual effects from the first trial on performance in trial 2 is denoted *transfer* in the following.

3.4.2 Measurements

The performance metric for each subject was chosen so as to assess how well each subject did relative to a benchmark rule. Benchmark rules, when psychologically plausible and requiring little effort and yielding high performance can serve as an objective standard against which to measure impacts of differential treatments (Paich and Sterman, 1992; Kampmann, and Sterman 1992). The benchmark rule is explained in chapter 4.

Instead of comparing the benchmark to the subject "en bloc" over 40 simulated years, the benchmark rule was updated by the decision environment produced by the subject after each decision period. This was done because the treatment conditions vary in terms of profit opportunities. The metric measured the difference between the subject and decision rule performance for each decision period, given identical conditions. The performance metric was a normalized measure of the sum of annual differences as will be discussed in chapter 4.

In order to investigate the mental processes that produce outcome differences, a questionnaire was designed. This questionnaire was distributed before and after each game and contained 24 questions relating to the game the subject was about to play, or had just finished playing. Each question asked subjects to describe a relationship between two variables in the game. Subjects had a choice of 5 classifications of each relationship: Increase (immediate), Increase (delayed), No relationship, Decrease (immediate), Decrease (delayed). The answers were compared to the correct ones and one point was given for each fully correct answer and an answer that showed the right direction but the wrong delay value was given half a point.

The questions are shown below in figure 3 for the tanker case. Only one line was shown at a time to subjects. The lines were distributed in random order. The real estate questionnaire was structurally identical, with the wording used in the real estate context.

..Relative to Tankers

An increase in...

Mkt Tonnage	leads to...		in... **Newbuilding Starts**
Average Life Time of Tankers	leads to...		in... **Depreciation**
Mkt Tonnage u/Constr	leads to...		in... **Secondhand Price**
Spot Rate	leads to...		in... **Mkt Tonnage**
Mkt Tonnage u/Constr	leads to...		in... **Mkt Tonnage**
Secondhand Price	leads to...	○ **immediate increase**	in... **Spot Rate**
Operating Profits	leads to...	○ **delayed increase**	in... **Operating Costs**
Mkt Tonnage u/Constr	leads to...	○ **no change**	in... **Spot Rate**
Interest on Bank Balance	leads to...	○ **immediate decrease**	in... **Bank Balance**
Newbuilding Starts	leads to...	○ **delayed decrease**	in... **Capacity Utilization**
Transaction fees	leads to...		in... **Loans**
Average Life Time of Tankers	leads to...		in... **Secondhand Price**
Newbuilding Starts	leads to...		in... **Secondhand Price**
Price Elasticity of Newbuilding	leads to...		in... **Mkt Tonnage**
Mkt Tonnage	leads to...		in... **Mkt Tonnage u/Constr**
Appreciation realized	leads to...		in... **Transaction fees**
Loans	leads to...		in... **Interest paid on Loans**
Operating Costs	leads to...		in... **Operating Profits**

Figure 3: Questionnaire to assess quality of causal understanding

The questionnaire was designed to assess the process by which people learn and transfer, and focuses on factors that dynamic decision researchers claim to be problematic to decision makers (Brehmer 1988; Sterman 1989a and b; Diehl, 1992; Kampmann and Sterman, 1992; Paich and Sterman, 1992). The findings in prior research indicate that the mental models subjects have for solving decision tasks tend to underestimate the impact of delays. Supply line information (e.g market tonnage under construction) seems to be particularly poorly integrated.

A subject's information acquisition can tell about which factors enter his or her problem space. One must expect that subjects who do well access the information used by the benchmark decision rule, since the rule is a condensed and efficient way of combining information and making decisions. In particular, information about the market supply line gives the decision maker a glimpse into what the future supply will be and prepares for decision making much more effectively than current financial results. As Fuglseth (1989) has shown, decision makers in unstable markets use a shorter time horizon than researchers (and normative rules). Thoughtful decision makers use longer decision horizons (Bakken, 1990b), and tend to focus on inherent market dynamics more than current operating income.

Another question that needs to be addressed is whether subjects can be predisposed to see similarities between real estate and oil tanker markets. One must expect that subjects who rate market dynamics of the oil tanker and commercial real estate industries as very similar will transfer better from the one market to the other than subjects who do not have such a prior notion of similarity. It is likely that subjects who indicate similarities have an understanding of the shared underlying structure. An understanding, or at least attention to such a structure, should help both performance and transfer.

Consequently, subjects were also asked to rate the similarities between different industries in terms of similarity. Similarity was operationalized in three dimensions: Physical, market dynamics and industry structure. The physical dimension reflects whether subjects think of two industries as contextually alike. Such similarity would help if it goes together with other proximity aspects, like market dynamics. Unwarranted similarity between objects in subjects' mental models may cause conceptual interference or cognitive dissonance (Heider, 1956; Akerlof and Dickens, 1982) and hinder transfer.

Market dynamics refer to time constants, speed of change in the market and so forth. Industrial structure reflects how an industry is organized in terms of size of firms, their interrelationships, etc.

The actual definition of each dimension shown to subjects is provided below in exhibit 3.4.2.

Dictionary; Dimension of Similarity

Physical Appearance : How the production facilities look; How the products (services) provided by these facilities look and feel.

Industry Structure : How the organizations that provide these products (services) are structured internally. How the firms that make up the industry typically interact between them and with their customers. Industry concentration. Barriers to entry etc.

Market Dynamics : The speed with which facilities and products are discarded from the market due to obsolscence and wearing out. Elasticity of demand and supply. The speed with which the market reaches equilibrium after external shocks.

Exhibit 3.4.2: Definition of similarity dimensions

The questionnaires are found in chapter 6.

3.4.3 Treatment administration

A sequence consisted in two trials that altogether required about 4 hours. Most subjects took a short break of less than 1/2 hour between the two trials. A trial consisted of reading a newspaper article, and a brief explanation of the game, completing the above mentioned questionnaire, making two decisions every period for 40 decision rounds, and finished with the same questionnaire. The next trial was run the same way. The questionnaires had a time limit of 25 seconds per question, but no other activity was limited by time.

All information, including newspaper articles, decision making games and questionnaires, was displayed on a 9" standard black-and-white Apple Macintosh screen. Readings, questionnaires and games were implemented in a sequence of information spreadsheets, using Wingz software from Informix. This enabled a recording of active mouse clicks and keystrokes, including their sequence and timing. Subjects scrolling of introductory texts, their decisions, access to graphs and tables, as well as the value of all variables in time were monitored. Each subject generated about 1/2 megabyte of information, most of which was related to the unfolding of the dynamic decision environment.

The introductory newspaper and explanatory text are shown in the appendices. The game design and decisions are explained in chapter 4, the questionnaire layouts in chapter 5. The complete experimental sequence of activities is shown in figure 3.4.3 for reference.

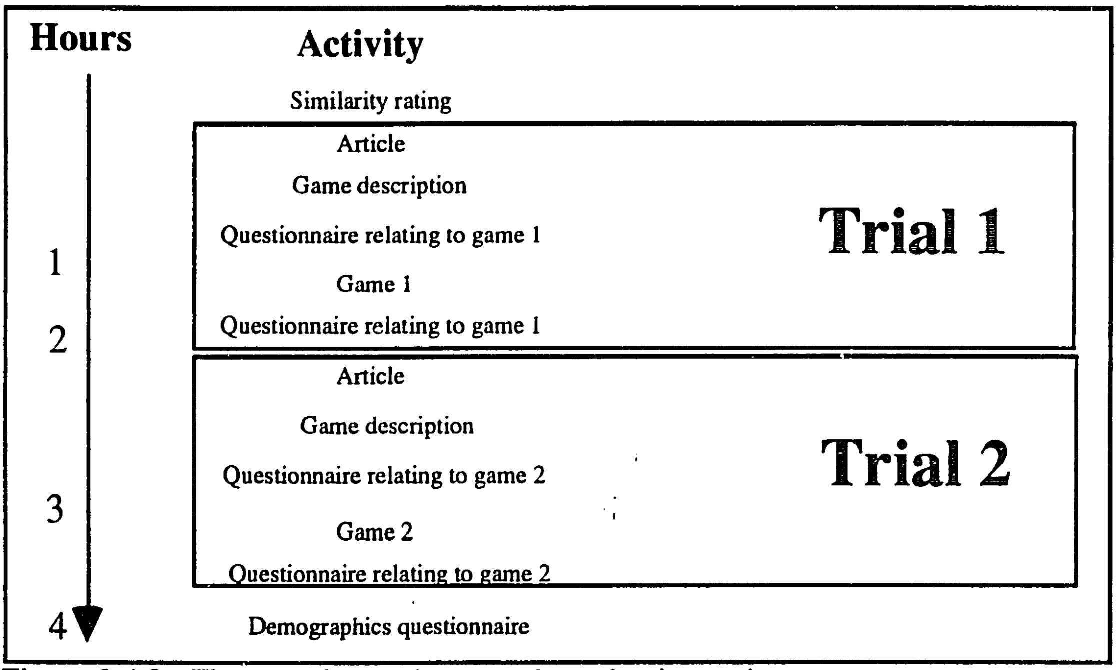


Figure 3.4.3: The experimental set-up along the time axis

3.4.4 Subjects

Forty-one MBA student subjects volunteered for the experiment by signing up after announcements in classes. These announcements included information about the reward system, but subjects were also informed that the decision games had been used by managers in large multinational companies. They were told that the expected pay was $8 per hour, that "previous masters have earned over $100 for the four hours" and that the experiment would run for at least 4 hours. The average pay was about $31.60 for the entire four hour session.

Considering that the expected pay is not much in an academic program where tuition was over $16000 per year, one must assume that volunteering was based more on inquisitiveness and perhaps an interest in management flight simulators than the financial reward. Yet, once recruited, subjects appeared to be very focused on their own results.

The experiments took place between December 1990 and March 1991 in a period were real estate problems figured prominently in the news. Of course, no communication between students was allowed during the experiment. Participants also promised not to talk to other students about their experiences until they got feedback (including pay) in the mail from the experimenter.

The students were told to use as much time as they wanted. Subjects reported that they enjoyed the experimental markets, but some found the timed questionnaire boring. Since the questionnaire was repeated four times, answers to later questionnaires might be less reliable than earlier ones. One subject left before even completing the first trial; this subject's results were discarded.

3.5 Summary

Six hypotheses will be tested; two of which relate to performance, and four to transfer. The hypotheses state that context familiarity and dynamic compression should improve performance. Similarly, initial context familiarity and high compression should improve transfer, but transfer will be degraded by changes in either context or compression. Since subjects' academic and work experience differ, it will also be possible to investigate the impact of background on performance and transfer.

The operationalization of performance and transfer is shown in chapter 5, while questionnaires and information acquisition metrics are shown in chapter 6.

4. Operationalization: Real and experimental decision environments

4.1 Introduction

The previous chapters have indicated the need to study learning and transfer in decision making environments that are both dynamic and contextually well defined. Dynamic environments enable a sequence of decisions where performance may improve, yet inherently poor feedback transparency creates problematic conditions for causal attribution and subjects may not learn much. Poor outcome feedback transparency acts so as to degrade inferences (Brehmer, 1980). In dynamic tasks with action feedback, however, decision making is further complicated by the fact that initial decisions shape future decision options. For instance, a real estate developer who is waiting cautiously during a market up-swing will not benefit from a profitable market as much as an aggressive developer. Thus the cautious developer will have less financial leverage in the late stages of a boom than more aggressive competitors.

This chapter presents a class of feedback environments: Stock management tasks. These tasks have several desirable properties. First, they are well defined (Diehl, 1992). Second, the real world abounds in phenomena that may be conceptualized as stock management terms. The chapter uses this stock management to first describe and later analyze Oil Tanker and Real Estate markets and how investors perceive them. In this description, an emphasis is put on their systemic nature, and on the interaction between the markets' structural properties and decision makers (in)ability to make causal inferences about how the markets operate. Most of this chapter deals with decision makers and the real markets. At the end of the chapter, however, the experimental markets are described. High performance characteristics and the issue of poor conditions for learning in the described real markets are discussed.

4.2 Stock Management

While the control of first order stock management tasks often yields good decision performance (Mackinnon and Wearing, 1985), higher order stock control tasks are notoriously hard to manage (Meadows, 1969; Sterman 1989a). Inherent physical delays and price rigidities caused by institutional factors make systems lose transparency and feedback about these processes may create instabilities, i.e. slow and erratic returns to equilibrium after a shock to the system (Kampmann and Sterman, 1992). Asset price fluctuations create transaction and profit opportunities for the savvy contrarian in unstable stock management systems, yet decision makers may not perceive systematic instabilities (Randers, 1984a; Drummond and Maidment, 1989).

Figure 4.2.1 shows the generic stock management description in stock-and-flow terms. (For a more detailed account of the descriptive methodology, see Forrester, 1961 and Morecroft, 1985.) A decision maker uses a goal function and available information to determine a desired stock. The decisions then lead to the acquisition of assets that serve to achieve the stock goal.

Actions, or decisions, are based on information about system states. Decisions are again constrained through cost and other limitations. The accumulation into stocks takes time and so system states may have changed from the time of a decision until the stock reaches its originally desired level. In the mean time, the desired stock itself may have changed considerably. One reason that goals change is that the initial profit opportunity was perceived by other investors as well. Delays in information about system states create similar control problems as do physical delays (Diehl, 1992).

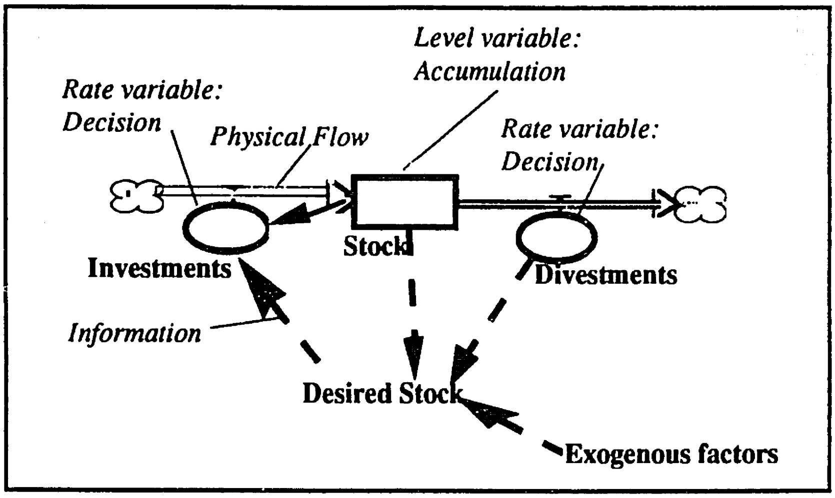


Figure 4.2.1: The decision maker regulates investments so as to achieve the desired stock.

The first order stock management task in figure 4.2.1 has been shown to be easy to control (Mackinnon and Wearing, 1985); the stock is immediately changed by investments. Systems of higher order, i.e. with more accumulators, can be unstable. Such a case is shown in figure 4.2.2. While the relationship between demand and supply determine prices and profit opportunities, construction initiation leads only to an increase in assets under construction. If agents do not act according to a rational expectations model as is the case in dynamically complex markets, (Kampmann, 1992) prices evolve as a function of available assets. These must be <u>completed</u> before they add to supply and affect market conditions. Instabilities arise because at the time the stock is in equilibrium, the supply line is either above or under the equilibrium level, pushing the stock up or down from its equilibrium level.

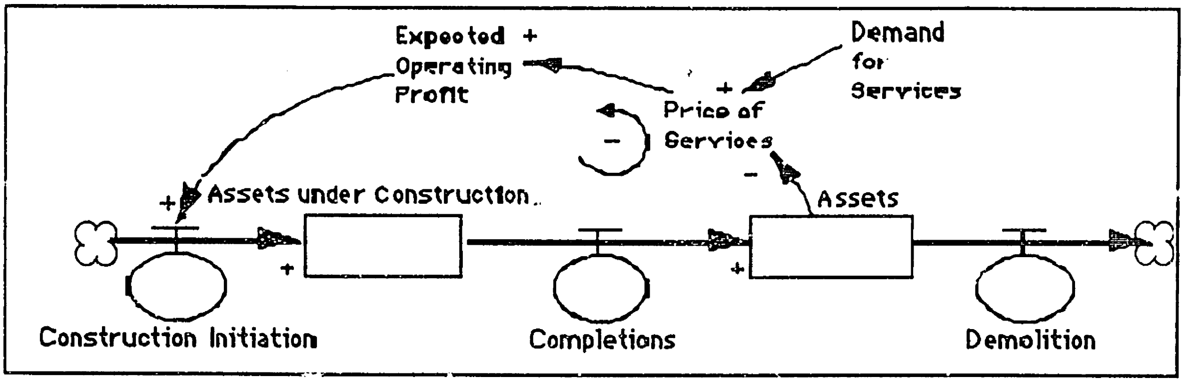


Figure 4.2.2: The feedback structure of a second order stock management task.

In figure 4.2.2, the desired stock is determined by the disequilibrium signal inherent in the price. The relationship between price and demand is determined by demand elasticity and its magnitude also determines system stability. If demand is inelastic, i.e. largely unresponsive to prices, then prices must fluctuate significantly to clear the market when supply and demand are out of equilibrium. Such price fluctuations are characteristic of inexpensive complementary products that are a small part of some end-product. Oil transportation is an inexpensive complementary product that counts little in the final gasoline price (Zannetos, 1960). Similarly, the rental cost for a typical office user constitutes such a small fraction of total personnel and other expenses; space requirements are also largely unaffected by prices.

Unavailability of substitute products also contributes to low demand elasticity. When gasoline prices fluctuate, there is little short run availability of synthetic fuel or electric engines that would reduce oil transportation demand by a shift in demand for the primary product when oil transportation prices increase. This further contributes to price volatility and accentuates instability in higher order stock management tasks.

Higher-order stock management tasks can induce decisions that cause stocks to overshoot their desired levels (Meadows, 1969). Such decisions are at the same time largely responsible for economic instability. As shown in figure 4.2.3, the outcome of such decisions in terms of fluctuations should appear transparent to decision makers even if the causal structure of the underlying system remains unclear. It appears that subjects, if anything, have a bias towards inferring systematic patterns in high variance time series data.

As suggested in figure 4.2.4, however, the tendencies of a stock management task to produce instabilities may not be perceived by looking at shorter term indices alone. In figure 4.2.4, subjects will have to use knowledge about the causative process in addition to outcome indicators to be able to infer that figure 4.2.4 is indeed a fluctuating system of the same type as the transparent type of 4.2.3, again caused by the feedback structure in figure 4.2.2. Indeed, figure 4.2.4 is similar to the pattern seen by decision makers in real, low frequency markets. Consequently, learning about the inherent and long term behavior of a market system can be enhanced by elongating the historical sampling (Bakken, 1990b) or by inferring the inherent stability properties in the causal nature of the stock management system (Bakken, 1989b).

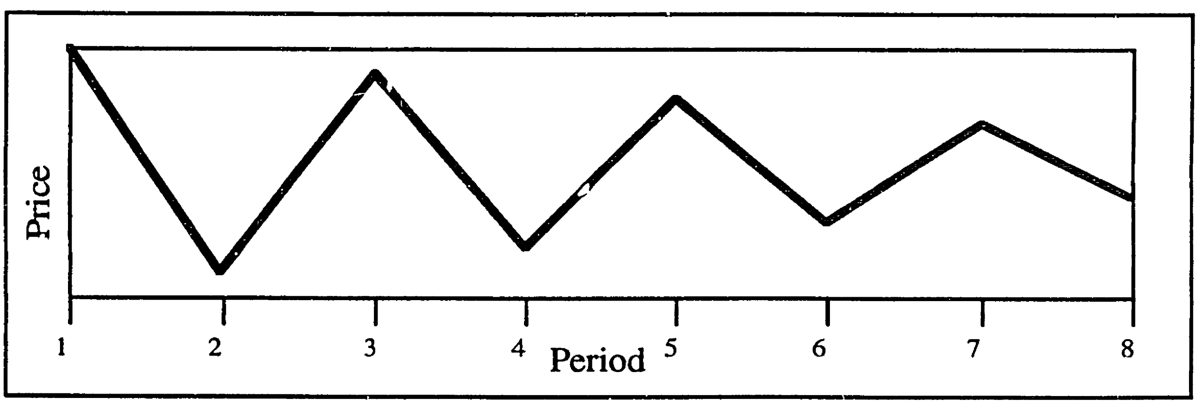


Figure 4.2.3: High frequency oscillations.

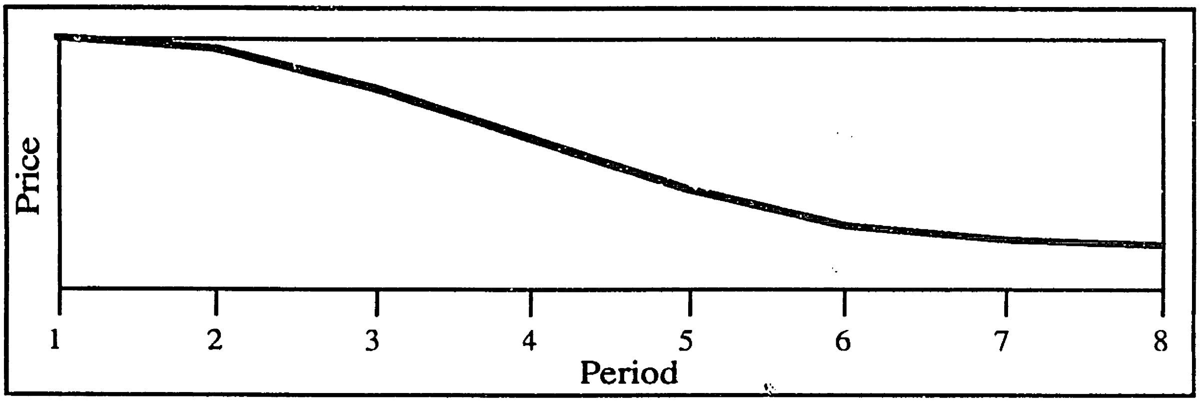


Figure 4.2.4: Low frequency cycle with a 16 year period.

4.3 Oil tankers and real estate as instances of unstable markets.

Oil tanker and real estate markets can be seen as stock management systems. As such, they are characterized by low demand elasticity, long delivery delays and long asset life times. There exist many dynamic models of the oil tanker industry (Raff, 1960; Randers, 1984b) built to gain insight into these instabilities. There also exist models of real estate (Laurent, 1970). Models are commonly built around the structure shown in 4.2.2, but with details distinguishing assets along the age, size and type dimensions (Maidment and Drummond, 1989).

In this section the two markets are described in more detail, with a special emphasis on how decision makers perceive the markets, their own role in them and the poor conditions for learning. The section provides evidence that several aspects of the causal nature of the system are not apparent to decision makers, since the precondition for effective learning, timely and transparent feedback availability , is not met.

4.3.1 The oil tanker market.

Figure 4.3.1.1 below shows monthly oil tanker transportation spot rates for the last 40 years. Though much industry wisdom says that tremendous peaks are caused by wars and other unpredictable events, the capacity utilization figures in figure 4.3.1.2 show that wars have little impact when utilization rates are low. As an example, the Iran-Iraq war, arguably the one that had most impact on oil shipments, had far less consequence for transportation rates than did the Yom Kippur war. This is due to the fact that in 1973 there was a shortage of tankers, while there was a surplus in 1979-81.

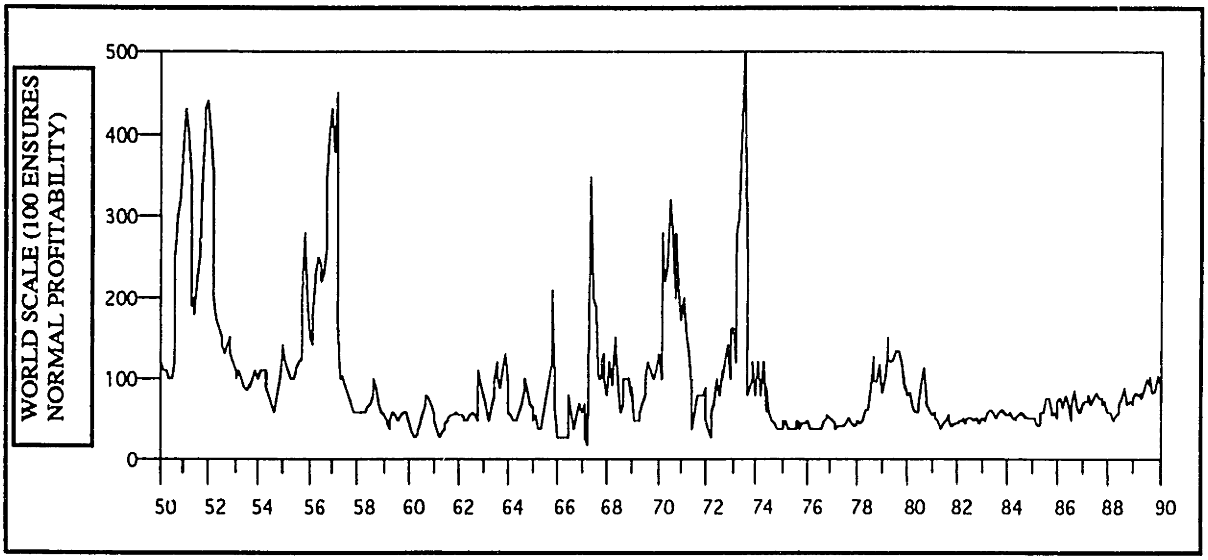


Figure 4.3.1.1: Spot rates for transporting oil on medium tankers from the Arabian Gulf to the US east costs for medium sized oil tankers. Sources: Randers (1984b), Fearnleys (1983-1988) and Drewry Shipping Consultants (1990).

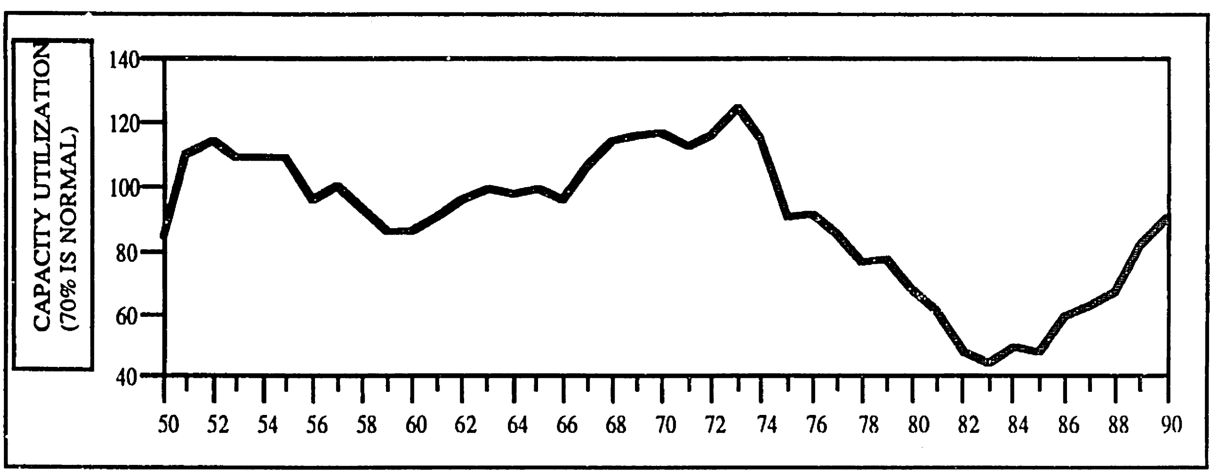


Figure 4.3.1.2: Oil tanker industry capacity utilization. Sources: Randers (1984b), Fearnleys (1983-1988) and Drewry Shipping Consultants (1990).

When petroleum was first put to use, the crude oil was usually produced, refined, and consumed on the same continent. However, rapid consumption growth after World War II, the depletion of US reserves, and the development of the huge Middle East oil fields meant consumption could only be met by transporting crude oil across the oceans.

The average size of oil tankers grew quickly in the nineteen-fifties, sixties and seventies. Construction technology improved so as to enable the design of huge tankers, some of which can carry 3.5 million barrels of oil (over 100 million gallons). By the early eighties, the deck of a typical supertanker was twice the size of a football field. Many yards expanded capacity in the sixties and seventies, but even in these highly productive shipyards, it can take several years to complete a supertanker.

In Scandinavia, shipyards used to have much the same importance for industrial activity as the auto industry has in the US. Fluctuations in prices led to erratic construction activity that again caused problems for economic stability. However, Scandinavian shipyards lost much market share to the Japanese in the sixties. In the seventies and eighties, South Korea emerged as an important shipbuilder. Just as the Japanese enjoyed a labor cost advantage over the Scandinavians in the sixties, the Koreans enjoyed a similar cost advantage in the seventies and eighties.

In the tanker owner and operating industry, often called "the biggest poker game in the world" (Rawlinson and Porter; 1983), enormous fortunes have been won and lost. The players include the major oil companies, who operate tankers in order to secure deliveries of crude oil to their refineries and keep a fleet of smaller tankers that moves refined products. The oil companies are dependent on reliable deliveries to their refineries and distribution systems and place great value on securing smooth deliveries.

Independent ship-owners own and usually also operate tankers on single voyages or long term charter contracts with the oil companies. They compete among themselves. Many of these independents are based in Greece, Hong Kong and Scandinavia, making the industry highly international. Companies register their ships in countries with little or no taxes, sailing under "flags of convenience."

As shown in figure 4.3.1.1, the tanker market is very volatile. Ship-owners have enjoyed periods of extraordinary profits and, also, periods of extended depression. In the early seventies, transportation rates were so high that ship-owners could pay back loans on highly leveraged new supertankers with income produced in a couple of month-long round-trip voyages between the Arabian Gulf and the US. Tankers have otherwise a useful life time of 20 years.

The rapid consumption growth as well as the shift towards non-US crude production created a long period of prosperity among tanker owners in the sixties and early seventies. To many investors, supertankers appeared to be in chronic short supply. The problematic years around 1960 were soon forgotten and trend extrapolations of transportation rates and demand made during the late sixties (Bakken, 1990b) projected into the future. The crescendo came in 1972 when world-wide shipyard order books were about 50 % of the operating tonnage (Fearnleys, 1982).

According to one shipowner, it was frantic (Bakken, 1990b, p 4):

> "The 1956 Suez crisis lead to uncertainty and rate hikes. Subsequently there was a lot of investments before 1960 and thus the early sixties was a hard time. The 1967 closing of the Suez canal also led to good rates. But the markets did not become frantic until '72 when shipyards all over the world were flooded with orders. The sad thing is that governments from Japan to UK contributed to the over-investment through subsidy programs to aid shipyard workers... Watching the mountain of tonnage on order in 1972, we decided to liquidate our tanker holdings. We told the purchaser that we were diversifying into oil-rigs and other oil-related activities; which we indeed did..... Our selling out in 1972-1973 was determined by our technical knowledge. The tankers got bigger, but time to complete tonnage also dropped much... It seems to me that many contractors totally missed the fact that tonnage was coming on stream very fast. We, however, owned the Akers shipyard in Oslo and so could foresee the glut of capacity coming on so soon."

The recession induced by the 1973 oil embargo and tripling of crude oil prices combined with the large number of new tankers coming on stream in the following years produced a long and sobering period for shipowners and lending banks. Seventeen years after the market crash, new entrants saw rosy prospects and older, more experienced owners could only sigh (Salpukas, 1989):

> "That's shipbuilding. It's always been up and down. There are those who hope it will not be as bad as yesterday. Of course it will be as bad as yesterday. The only thing I can hope for is that it will not be as bad again for such a long time. "

Conditions for learning

The erratic nature of earnings combined with the long delivery delays and life times create poor conditions for learning. In addition, investments tend to be highly leveraged so that shipping investors become risk loving (Salpukas, 1990). Learning is also hindered by a common defeatist attitude among investors. These attitudes can be exemplified by the "gut feeling" comment of one Greek ship-owner (Rawlinson and Porter; 1983), indicating that there is no predictability in the markets, and nothing to reflect about:

> "Business -- I know nothing about business, I only know how to make money".

Ship-owners have often looked into short-term rosy predictions with a myopic bias (Fuglseth, 1989). However, though the average life expectancy of shipowning firms is low and bankruptcies are common, there also exist firms that have been in the industry for decades. Some of these have developed a more philosophical attitude to causal inference and data analysis. This enables a historic view long enough to see the cyclicality unfolding and an understanding of the business' causal structure. For instance, one ship-owner compared operating tankers to ocean racing. He stressed the importance of history and of memory (Bakken, 1990b, p 2):

> "After having raced for many years, you realize that the current follows general patterns depending on time of the day, the direction of the wind, as well as the specific conditions of that day. After some time, if you process and think what happened that day, you appreciate that certain patterns recur. To appreciate these patterns memory and thoughtfulness is required...
>
> I think the clue was to build up a memory... Every race was like a business cycle. They were all different, but still, common factors were present. Likewise, in the business field it is mandatory to build up a database with many personal experiences... In addition,

> political and economic history says a lot. In my case, the story of my family and bankruptcies back to the middle of the last century has proven to be a rich source of data."

The fact that the account is atypical is also underlined by a recent study (Fuglseth, 1989) where talk-aloud protocols compared shipping researchers with many years of research background to professionals with similar period in the same field. Fuglseth found that researchers differed from professionals in that researchers distinguish between demand and supply and also between immediate and long term consequences, i.e. they have a mental capacity to keep them separate and to combine them.

Fuglseth's work shows that compared to researchers' cognitions, professionals' mental models are simpler and their causal chains shorter. Of course, the professionals need to have an understanding of contextually rich aspects of the business that the researchers do not have. Thus it appears that in the competition for scarce cognitive resources, professionals tend to simplify into schemas that confound phenomena of different time horizons.

The lack of distinction in time frames can also be explained by the much higher frequency of information about immediate issues in information sources used by professionals. Inherent higher uncertainty in the longer time horizon and the higher ease with which short term predictions can be validated by available data, may also create conditions where short term issues loom bigger in professional minds.

4.3.2 The real estate market[3]

Real estate markets play a pivotal role in western economies, and in the 1985 to 1992 period were foremost in creating billion dollars losses in US, Japanese, Norwegian and

---

[3] Hernandez (1991) has provided a good account of the feedback structure as well as the causes of poor learning in real estate markets. The account has been used here as a main reference, since I co/supervised the thesis, helped her design a semi-structured interview scheme etc., and worked very closely with her in that study .

Swedish banks that have cost taxpayers billions, too (Munthe, 1992; Reve, 1992). Both US and Scandinavian banking systems have been put under enormous stresses due to soured real estate loans. These losses occurred within a short time frame and under a wide range of quite stable (Japan) and changing (Norway and US) regulatory environments and suggests that a common attribution, namely that the decision makers were unprepared for a less regulated banking environment (Tranøy, 1992), cannot be the only cause.

Yet, as shown above, only recently have instabilities in real estate markets received scholarly attention. Older, less analytical, accounts have long stressed and linked actors' decision myopia to cyclical behavior (Hoyt, 1933). Recently, Wheaton and Torto (1988) showed that there are systematic fluctuations in real estate values and vacancy rates and that there exists a 10 to 12 year cycle in average US vacancy data. By definition, local data would show more erratic fluctuations. The Wheaton and Torto data are shown in figure 4.3.2.1.

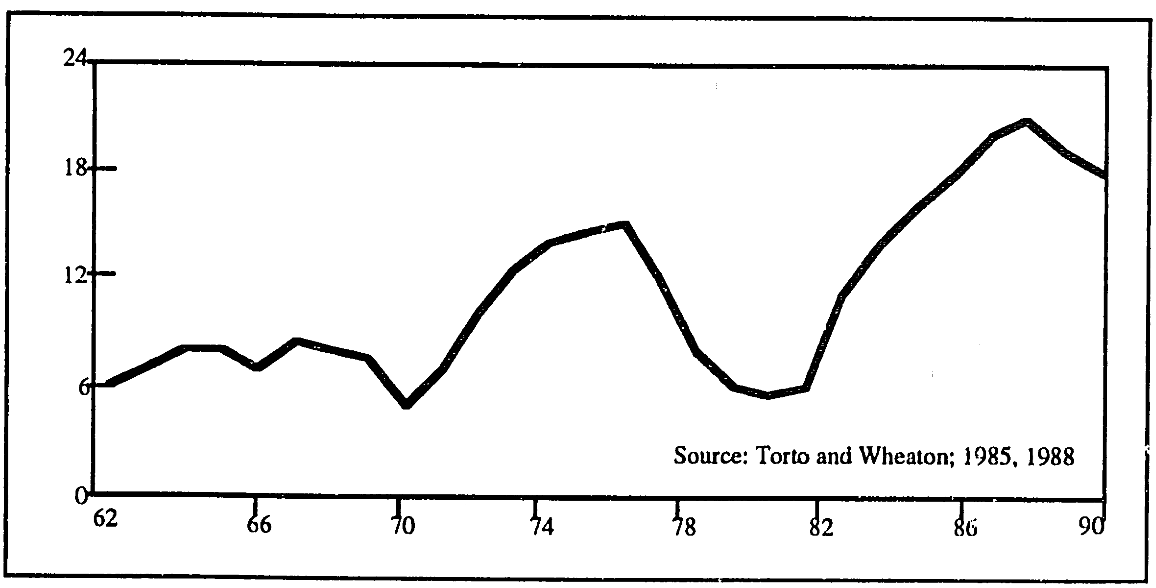


Figure 4.3.2.1: Fluctuations in national vacancy rates.

Separately, Boston vacancy data (BRA, 1989) show similar vacancy dynamics for the Boston market.

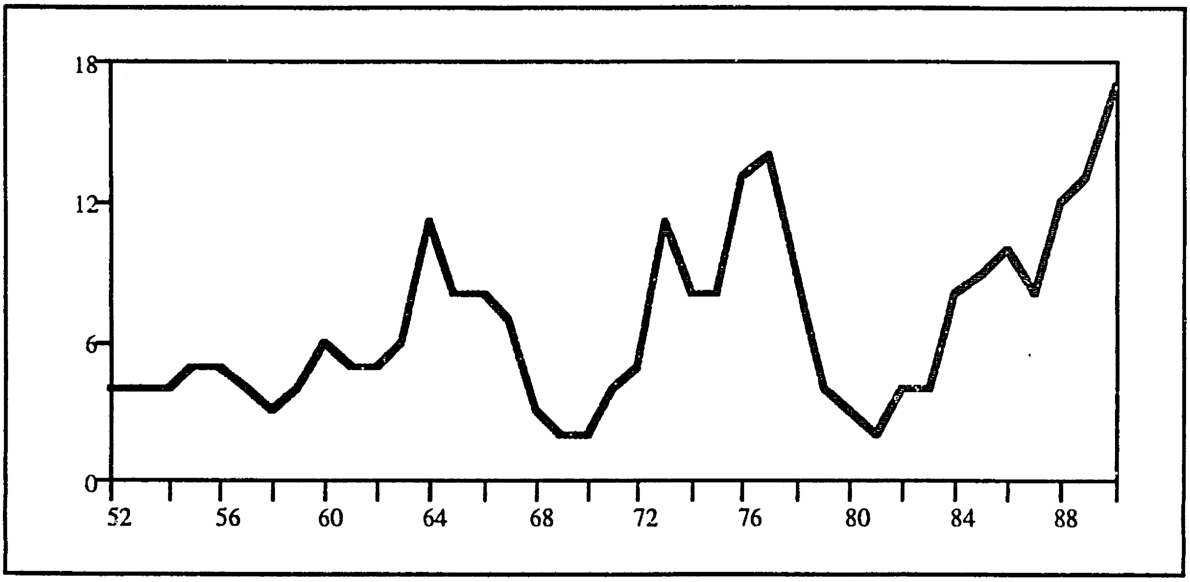


Figure 4.3.2.2 Office vacancy rates, Boston (Boston Redevelopment Authority, 1989)

Interestingly, and mirroring the frantic nature of the oil tanker industry in the early 1970's, discussed in the previous section, the Boston Redevelopment Authority published (1989) a series of increasing demand scenarios for the need for office space in Boston 1989 to 1995. The predictions were all simple trend extrapolations using regression of leased space (the operationalization of demand) as a base, differing only in historical time sampling. Seven years of data gave a moderate future demand, while five and three year historical sampling periods gave raise to median and high demand predictions. No account of the mushrooming of office space in the areas surrounding Boston was integrated in the analysis, nor was there any analysis of whether average space per person had increased as the perverse consequence of the higher leases in that period as will be explained in the next subsection.

The behavior shown in figures 4.3.2.1 and 4.3.2.2 can be produced by a second-order stock management feedback system, which will indeed be shown in section 4.4. The causal structure of an expanded model of such a system is explained in the following subsection. Interviews with real estate developers serve as the main information source. The model can also be used to explain the oil tanker market, as the supply "stock management" structure, the role of the financial community and the perverse demand function ("lock in low prices now before they increase") are similar in both markets. Yet the institutional structure of the markets are different as will be explained in section 4.3.3.

The second-order stock management structure is at the center of the causal loop description depicted in 4.3.2.3 termed the Supply subsystem. The supporting sectors of Demand, Finance and Cost of Exit appear to destabilize the system. I.e. deviations from equilibrium are accentuated rather than smoothed. The causal nature of the system is depicted by tracing around the loops .

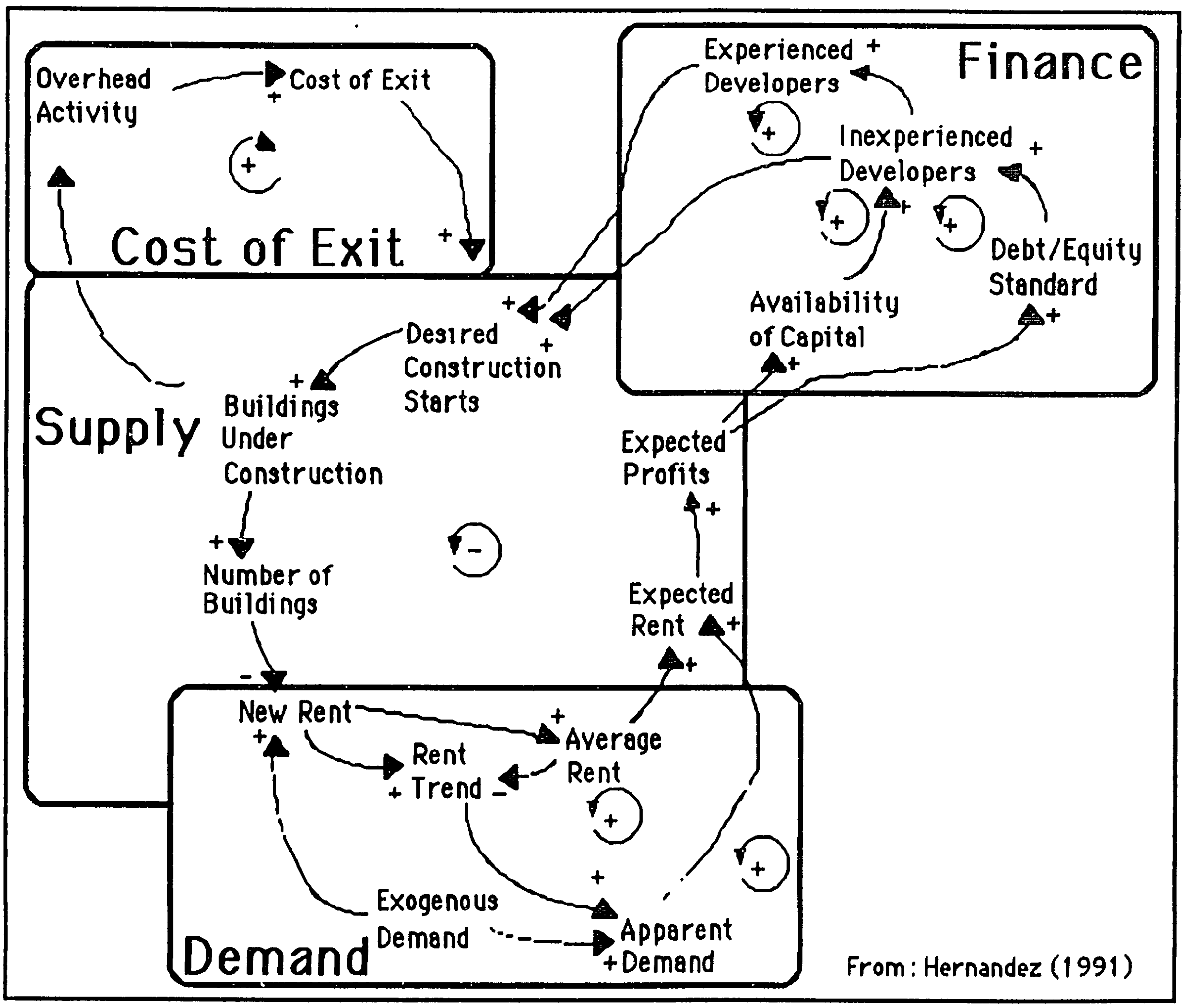


Figure 4.3.2.3: Diagram of the Real Estate Development System; plus signs indicate positive (same direction) causality, while negative signs indicate opposite causality.

Supply

Imagine a demand shock that perturbs the system from an initial equilibrium. Rents begin to increase since supply is fixed in the short run. Expected rents will increase with time, too. Given higher rent levels, more projects now become profitable and desired construction starts increase. Buildings under construction later increase and as they become available, the number of buildings in the market increase. As the supply is now approaching demand, new rents begin to fall. The center loop is self-correcting.

However, by not incorporating the supply line of buildings under constructions, decision makers over-invest and create instabilities. The length of time lags is so substantial and supply dynamics so complex that developers may give up and have no causal model of supply, just as was the case in the oil tanker business. The lack of formalized decision criteria and lack of attention to supply lines (future supply) is evidenced in the interview below:

> A project I am currently working on consisting of a golf course with luxury residential units surrounding it has already taken three years of my time and only a few houses have been built. I am still trying to obtain many of the permits necessary to complete the project.... Additionally, even though the luxury-end housing market is doing poorly currently, it may be very different by the time the project actually opens in the future. Who can tell?
>
> We never did a formal or thorough analysis of what supply may be in works in competition with one of our developments. In analyzing future supply and demand, I think it's too unpredictable to put a lot of emphasis and time in trying to figure it out.[4]

Demand

In figure 4.3.2.3, apparent demand is influenced by two factors. Exogenous demand includes factors outside the real estate system itself, such as economic attractiveness of an area. Apparent demand is, however, also influenced by changes in rents in the following perverse way:

---

[4]Interviewee 1, interview by Karen Hernandez, Philadelphia, PA, December, 1990

If tenants believe rents will increase in the future, they tend to forecast their own office needs into the future to lock in the new leases before rents increase. Tenants thereby increase apparent demand. As one interviewee commented:

> In analyzing market information, we would make trend line projections to determine if growth would continue and the space that we were contemplating building would be absorbed. The same as last years' benchmark, greater or less. How much greater or how much less would be the question. And the way you would determine that is to say most people are bearish, rents are coming down, so we won't do as much. Or the conclusion would be that most people are bullish, rents are firm, so we will do a little more.
>
> We assumed the change in rents would never be more or less than 10% of last years. However, frequently in my experience I saw swings greater than 10% in real rent - not the quoted or lease rent but the rent after it was adjusted for all free lease periods and tenant finishes. The real economic rent tends to be very elastic. After making projections, we would then look at what major competitors were doing.[5]

The point was made previously that short-term supply is inelastic. In the longer term, supply is quite elastic. Demand, however, is very inelastic (with exception of the perverse demand effect mentioned above). The expense of real estate is a small fraction of businesses' expenses and transaction costs for a tenant, i.e. moving expenses, are high.

In turn, an increase in apparent demand increases both expected rents and expected demand. An optimistic increase in expected demand was described by Hoyt (1933) in his description of the Chicago real estate cycle:

> At this phase of the real estate cycle [the boom period], the rapid rate of increase in the population of the city that has recently taken place is projected far into the future in the rosy calculations that are broadcast by real estate men. A city that will surpass in size any metropolis the world has ever known before is erected in these speculative dreams, and facts and figures are collected by business men of the community and by "distinguished scholars" to buttress these "castles in Spain" and to make them seem tangible to the lay mind.

Finance

Figure 4.3.2.3 showed that as expected profits increase, desired construction starts increase since lenders are more willing and eager to provide construction financing. Therefore, the availability of capital increases. Lenders' equity requirements will also

---

[5] Interviewee 2, interview by Karen Hernandez, Massachusetts Institute of Technology, December, 1990.

decrease as the perceived risks by lenders are reduced and banks are eager to get into the growing market. The combination of these two factors decreases barriers to industry entrance and new developers will enter the system with the hope of sharing in these expected profits.

With increased expected profits and banks offering attractive financing, people will enter the market to take try to obtain these increased profits. There are low barriers to entry, and as profit opportunities are perceived, marginal developers tend to enter. During a boom period a construction loan will fund over 90% of the construction cost. The rest is the equity portion provided by the developer. The developer, however, by joint venturing a project, may in the end contribute no real equity in a development project:

> I believe that the another major reason for oversupply occurring is the great amount of leverage developers can obtain. By banks requiring so little equity to be contributed by developers into projects, many developers are developing marginal projects and contributing to the possibility of over-built markets occurring. In the glory days percentages (*of equity participation, BEB*) are thrown away all together and I've done projects in which I contributed practically no money. Even if there are percentages required, I can still joint venture the project, and still contribute no capital.
>
> Additionally brokers take weak projects to weak banks to obtain financing and bankers are paid on business they produce. Both these factors help those marginal projects get developed.[6]

Many developers believe the key to success in real estate is financing. By entering deals with small equity requirements, developers shift much of the risk to the bank. As one interviewee said:

> When times are good I think it's important to prepare for the bad, for example, by obtaining financial partners and good financial relationships with companies with deep pockets. Thereby, when the tough times come, I will still be able to obtain financing for projects.[7]

The burden of refusing to finance a project falls upon the financier. Banks, however, like developers, have pressures to produce and grow, similar to that of developers. The following three comments were made in reference to Texas in the 1980's:

---

[6]Interviewee 1, interview by Karen Hernandez, Philadelphia, PA, December, 1990.
[7]Interviewee 1, interview by Karen Hernandez, Philadelphia, PA, December, 1990.

> The lenders are partially responsible for the problems we have.... they loaned lots of money to people who had no experience and little judgement. (Trammel Crow, Trammel Crow Company in Melody and Wagley, 1989)

> Real estate lenders at commercial banks complained about the risks they were taking but closed their eyes and made deals to earn fees and contribute to corporate earnings, which was their real mission. (Larry Melody, President of L. J. Melody and Company in Melody and Wagley, 1989)

> Developers are programmed to build and lenders are programmed to lend. Everyone felt that their case was special and for one reason or another they would succeed. (Bill Cooper, President of the Paragon Group in Melody and Wagley, 1989)

The last comment is particularly interesting in that developers believe that the market can be segmented into sub-markets with their own particular demand and that no substitution between markets exist. However, rents in different geographical locations and quality classes are certainly correlated.

In the recent real estate boom, The Comptroller of the Currency found that many banks had "ignored or compromised" basic principles "to increase volume and achieve higher levels of interest and fee income" (Boston Globe, 1989). Other problems pointed out include many loans that lacked borrower equity and a failure to obtain accurate, independent appraisals (Wall Street Journal, 1990).

While most developers create financial prospectuses of projects, the assumptions used in these are often questionable. For example, in a ten-year plan, it is common for a developer to show rents increasing at 5% for all ten years. Rents, however, are also cyclical and not always increasing but, at times, decreasing.

Many developers do compute best, worst, and most realistic scenarios. As was the case with the BRA predictions, however, these are very naive and sample a very short time period relative the system's very long resonant period. In addition, the incentive system is biased so that everyone participating has the same interest in making rosy predictions.

Some developers can be found at fault, on the other hand. For example, if a bank requires an internal rate of return of 15% and a project's projected return is currently 12.5%, a developer may manipulate the plan to get the desired result. Additionally, lending institutions have also been guilty of such manipulation.

> Many of the big lending institutions urged the builders on. Despite evidence of severe over-building, the life insurance companies poured money into real estate until 1985. In many cases they were investing for pension funds in return for a fee partially based on how much they put out. "I can recall meetings with an insurance company when we had rent projections of, say, $24 a foot on a project," says a former Lincoln official. "The insurance guys told us to change to projections of $28 so they could get the loan through their committee. That was common then."(Taylor, 1989)

But, in general, both lenders and developers believe there is great uncertainty involved in projections but are pressured to keep developing and loaning.

> Most people in the business lived under the assumption that there were wide possible swings in results regardless of financial projections. But whether your projections were accurate or not, you doubted greatly. However, you had to play the game with the lenders that you felt very confident about the projections. And what we learned later on was that the lenders were also uncomfortable with the probability of the projections but they also had to play the game with their leadership or directors because they also wanted to book loans and they did not want to be in a situation in which they did not meet quota that year.[8]
>
> The appraisers who were supposedly independent also got involved in this match by doing appraisals that allowed a lot of latitude and effectively rubber stamped anything that the lender and developer agreed were reasonable. No risk factor was involved in the appraisal - no beta factor - was factored into analysis. It was one big complicitous circle - no one wanted to say no or they would lose business.[9]

Another interviewee said:

> Additionally, developers are promoters and must motivate people. It's difficult to be realistic. You're always selling and after awhile you start believing your own delusions of grandeur. Developers are a bunch of optimists without a lot at stake which also creates over-built markets.
>
> We would use fifteen page spreadsheets (large sheets with small print) which would compute net present values and internal rates of return of projects by using costs of the project (in great detail) and the revenues and expenses and resulting interest expense computed monthly till the project was either all sold (if housing development) or ten years if it was a project we would hold. Rents were assumed to increase yearly by five percent or the inflation rate. At the time I thought that was a conservative estimate and it wasn't really questioned by lenders.
>
> I spend much time and human resources in tweaking the numbers on these spreadsheets to get the internal rate of return which the bank wanted to be willing to finance the project

---

[8]Interviewee 2, interview by Karen Hernandez, Massachusetts Institute of Technology, December, 1990.
[9]Interviewee 2, interview by Karen Hernandez, Massachusetts Institute of Technology, December, 1990.

> and also for our own internal analysis. In looking back though, I don't think one can know the future with the accuracy that those spreadsheets displayed. We would run best and worst case scenarios but no worst case scenario every approached what the markets are like today.[10]

Just as there is an incentive system within the financial community, there are incentives within the developer community through their compensation systems. Employees responsible for new development, especially at lower levels, typically contribute equity of their elder colleagues to projects and therefore bear little risk of their own. If the project is financially successful, however, they are typically given a percentage of the profits. This type of compensation system promotes development.

Trammel Crow's compensation system was structured in this way.

> In their haste to expand around the country, these companies hired a host of young partners, often M.B.A.s or former leasing agents with no development experience. These partners knew they had to build to make money because, by tradition, they are paid paltry base salaries, $18,000 to $24,000 a year. Their income picks up only when and if the project gets under way and development and leasing fees start pouring in. The ultimate payoff, of course, occurs when a project is sold at a huge profit. (Taylor, 1989)

Exit Costs

Developers typically earn development fees equal to about 5% of the total development cost. These development fees, funded from the construction loan, are certain, safe, and used to pay overhead costs comprised mostly of the salaries of employees. As markets become stronger, fixed cost typically increases as the firm hires more employees. As overhead increases, developers have extra pressure to start new development projects.

The need to cover overhead costs is thought to have contributed to the over-building in Texas (as well as other parts of the country):

> Too many ill-conceived projects were built for fees rather than the profits the buildings would generate once the projects were completed and sold. As on Wall Street, there's been too much money chasing too few deals (Taylor, 1989).

---

[10] Interviewee 1, interview by Karen Hernandez, Philadelphia, PA, December, 1990.

> In 1986 ... Lincoln (Property Company) and Trammel Crow (two of the nation's largest property developers) had a total of $3.5 billion in construction under way. Total development fees: about $120 million (Taylor, 1989).

A developer can do marginal projects just to collect fees to cover the cost of the (bloated) organization. The alternative to development is to find ways to use employees in other businesses or to let them go. Both choices can be expensive and stressful in the short run, so it is tempting to take on marginal activities.

The length of the development process also inhibits developers from exiting the development business. It appears difficult to end a project that has already received a go from city hall. Architect fees have also been paid as well as other non-recoverable costs. A developer becomes financially committed to a project when he buys the site and even more so when he begins construction. It is difficult to suffer this financial loss by walking away from the project and admitting defeat before even entering the game.

> ... At this point, I am not about to walk away from this project given the time and money I have already invested in it.... I will be a developer no matter how tough it may get. It's a big ego thing. It's not like you're producing a homogeneous, mass-merchandised product. Development is a more personal thing like creating a work of art. Being a developer I identify with developer groups and it's like being in a fraternity. It's the majority of my identity that I just can't walk away from.[11]

Another developer said:

> The hardest thing about defaulting on a loan is the phone calls that you have to make to lenders who have counted on you and whom you have relationships with. By making the phone calls, you admit to yourself that you can't do it and then you let the rest of the world know (Melody and Wagley, 1989).

In sum, the real estate market, and in particular lenders and developers appears myopic and institutional factors create biased incentives. Investors will be optimistic when the market looks promising and negative when less rosy prospects prevail.

4.3.3 Differences and commonalities between oil tanker and real estate markets

[11] Interviewee 1, interview by Karen Hernandez. Philadelphia. PA, December, 1990.

Though real estate and oil tanker markets share many features they are different in other respects. One main difference resides in the fact that while oil tanker markets are global i.e. the value of an asset is independent of its location, real estate markets, due to lack of asset mobility, are local.

There are commonalities in the slow physical flow and often biased feedback structure of the markets. These factors all contribute to low causal transparency and make learning difficult. Long assets life times combined with long delivery delays create inertia, making the markets unstable. The existence of demand dynamics that are shorter than, yet interact with, supply dynamics also make causal attribution difficult. Investors confuse demand dynamics induced by a business cycle with the longer supply dynamics.

Both markets contain demand dynamics that are linked in a complex economic picture, a picture that receives much public attention. This public attention focuses on issues that are contextually rich, such as oil embargoes and recessions. This further draws limited cognitive attention away from longer to shorter term dynamics.

Both markets are characterized by the lack of substitutes but are quite competitive. This, in conjunction with low demand elasticity and a "marriage" with a financial sector that often allows investors to build with no equity, creates asymmetric risk perceptions that help perpetuate instabilities.

4.4 Experimental markets

In order to study influences on learning in experimental markets, certain aspects of task complexity should be maintained. A key design objective has been to portray a system with an intermediate degree of transparency; transparent enough so that relationships could be learned within the 4 hour experimental design, but complex enough that some subjects would not learn the relationships. A wealth of information and purported realism has enabled use of sophisticated subjects, such as industry professionals and MBA students. The following are characteristics of tasks that fulfill these experimental requirements:

i. Learning is non-trivial, yet achievable in the experimental setting;

ii. The conditions for learning in the real world are problematic;

ii.1 The strategies that yield good performance in the simulated environment also do well in the real world;

ii.2 Important non-linearities and motivations from real world markets, such as financial rewards and bankruptcies, are replicated;

iii. The task can be varied along the desired dimensions;

iv. Decision making can be monitored unobtrusively;

v. Good and poor decision making strategies can be discriminated;

Two simulated markets, representing commercial real estate and oil tankers, were designed. These markets can be termed stock management tasks, and they share information feedback and physical stock-and flow-structure. The experimental market consists of a decision making interface on top of an underlying model, resembling a "management flight simulator" (Graham et al., 1992). This section explains and shows the model and the simulator interface.

4.4.1 The model

The model embodies the feedback structure of the underlying causal relationships. An experimental subject never sees this model shown in figure 4.4.1. Instead, subjects interact with a simulator in the spirit of a flight simulator (Bakken, et al., 1992) and has access to decision variables that determine his performance. Subjects also have access to information that may support their decision making. The implementation of the simulator allows unobtrusive measurement of decisions and information access behavior. It is also quite user-friendly so that subjects can learn to interact with the system with a minimum of written description.

The model reflects the description of the generic stock management problem depicted in figure 4.2.2 above. A graphical rendition and equations of the full model are found in Appendix 1. Figure 4.4.1 below provides an overview of the model's feedback structure. The dynamic behavior in the model arises from the link between market conditions, i.e. price of services, and ordering new assets. The price of services is determined by the relationship between the total supply of services and demand for them. With surplus capacity relative to demand, prices fall and new construction of tankers or buildings is reduced. Conversely, construction activity is increased by high prices of the services provided (transportation and office space respectively). The demand is the only exogenous variable in the model. It has a weakly autorcorrelated, pink noise, formulation so as to mimic a business cycle fluctuation in demand as well as randomness in that cycle. The demand formulation is stationary.

The instability in the system arises because the competition does not adequately adjust for the pipeline of ordered, yet undelivered, assets. Thus, construction starts increase until prices cover costs and normal profits. At that time, a substantial building activity is under way and leads assets to enter the market thus creating oversupply and downward price pressure.

The decision maker can buy and sell assets from competitors, and may also build new assets. He is constrained in doing so by his net worth. This net worth consists in liquidity as well as unrealized capital gains on assets. The net worth increases by operating profits as well as by interests paid on his bank balance. However, the most spectacular profits arise from selling assets for a higher price than their original purchase value. The relationship between prices and competitor behavior is deterministic.

Initially, market behavior is determined by the competitions' actions. The decision maker's one-percent initial market share makes decisions unimportant in the determination of market behavior. Shrewd decision behavior, however, may increase the player's share and such decision makers may gain control over the market.

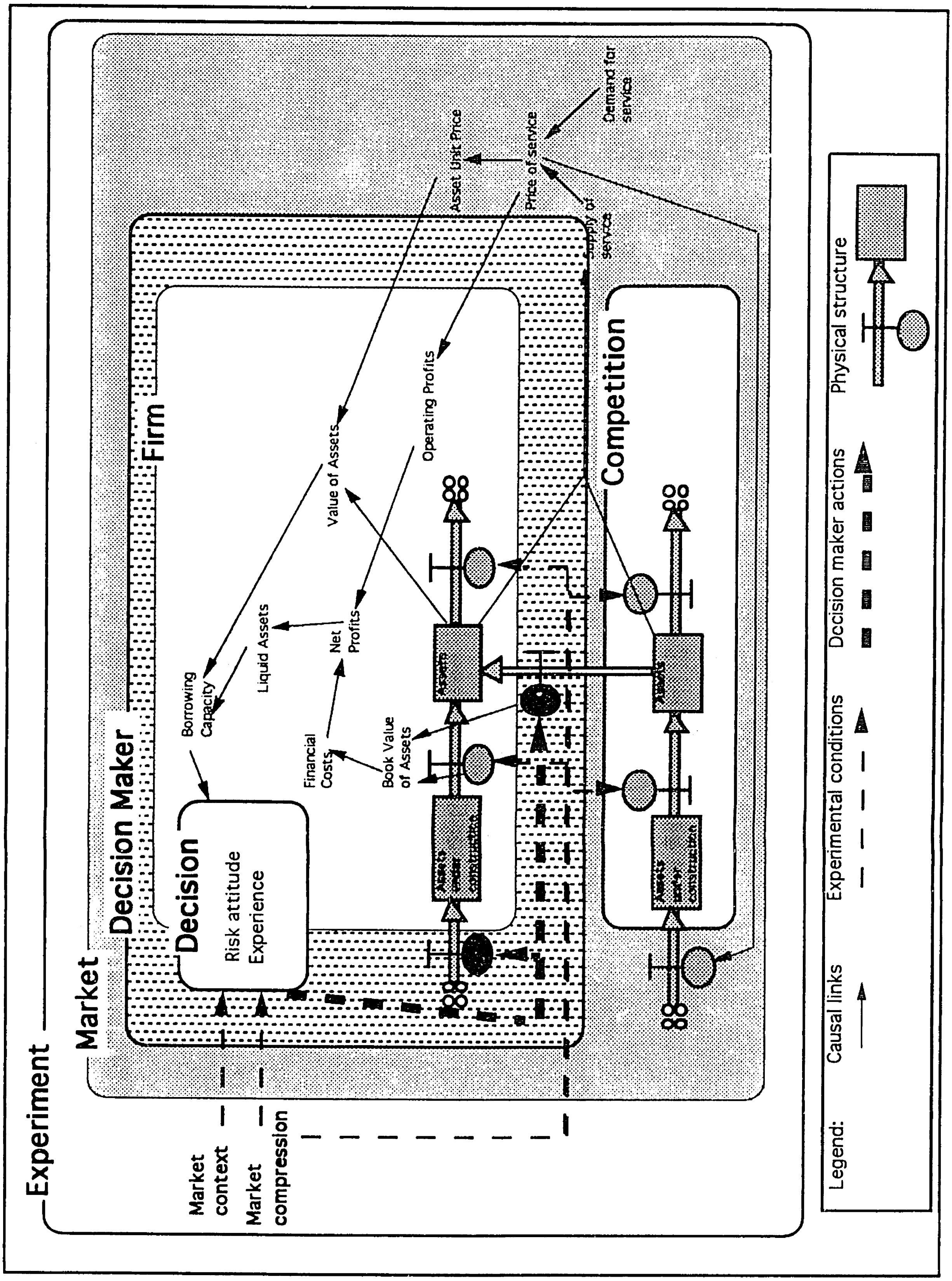
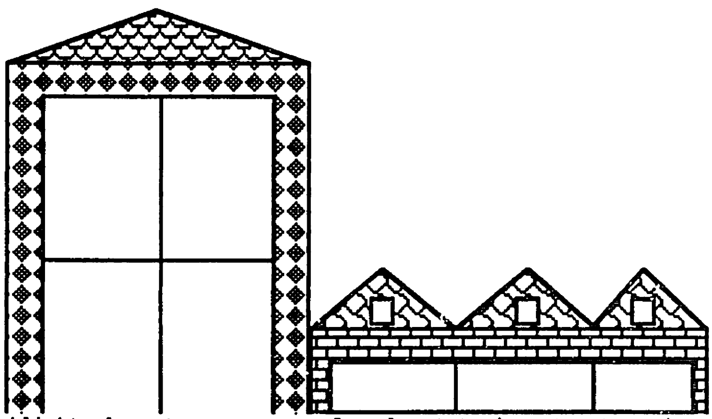


Figure 4.4.1: Experiment and model of market

4.4.2 The simulated decision environment

Below, in exhibit 4.4.2, follows the game description handed out to subjects in the real estate case. The oil tanker description was almost identical and is provided in the appendix.

The downtown office Real Estate market.
A participatory simulation game



This game highlights long term aspects of real estate investments and operations that tends to get little attention in everyday management. Underlying the game is a model that has been constructed using Boston data from reliable sources for the entire post-war era. Of course, the model is not the same as the real thing; like any map it serves to give an overview into issues that are hard to detect in the real world. In the game, you play the role of a president of a real estate investor/developer. The firm is initially small and you control about one percent of the market, but with your understanding and shrewd piloting, it can grow tremendously. The game gives you financial resources and opportunities for making big money, ...and for going bankrupt.

1. Game overview

Figure 1 portrays the overall structure of the game. You operate in a market for prime office developments. You can buy and sell existing, occupied buildings as well as develop new space.

All space is identical and so your lease costs and rental rates and those of your competitors are the same. Likewise, your vacancy rate is the same as the market's. The computer simulates your competition in a very straightforward manner; when expected profits go up, competing developers see opportunities and starts constructing new space. The value of existing space booms, too. Conversely, little development takes place when poor market prospects prevail and building prices fall.

There are some real investment and transaction limitations in this market... You can only invest as long as you have financial leverage to do so. In addition, if you should grow to control the market, you can never sell or buy more than the equivalent of ten per cent of your competitors buildings. Buying up or selling more space than that would destroy the market.

Below you see the overall structure of the game.

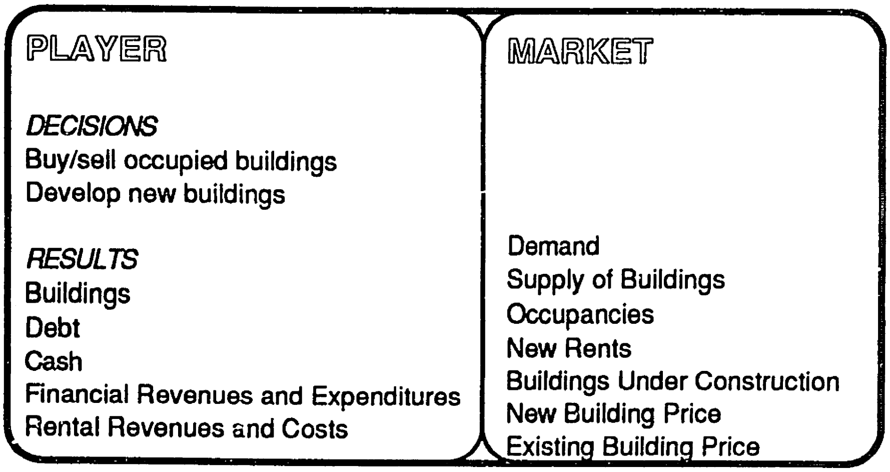


Figure 1.1: Structural Overview of the Office Real Estate game

Below you see the decision making screen as it appears in the game.

	Results for Year	**1989**	
My Buildings	**Existing Building Purchas**	**0**	Buildings/year
	New Permits	**1**	Buildings/year
	Blds u/Cstr (next year: 33%)	3	Buildings
	Buildings	30	Buildings
Prices	New Rents	5.0	$ m/Bld/year
	Existing Building Price	21	$ mill/Bld
	New Construction Price	19	$ mill/Bld
	Maintenance Unit Costs	2.0	$ m/Bld/year
	Occupancy Rate	90%	Percent
Rental Activity	Rental Income	135	$ mill/year
	Rental Costs	54	$ mill/year
	Rental Profits	**81**	$ mill/year
Capital	Interest on Bank Balance (5%)	6	$ mill/year
	Interest paid on Loans (10%)	58	$ mill/year
	Deprec'n {Demol'n} (3.3%)	19	$ mill/year
	Appreciation realized	0	$ mill/year
	Transaction fees (10%)	0	$ mill/year
	Net Financial Gain	**-72**	$ mill/year
	Net Profit	**9**	$ mill/year
Balance Sheet	Bank Balance	117	$ mill
	Loans	583	$ mill
Market Buildings	Mkt Buildings under Constr	300	Buildings
	Mkt Buildings	3000	Buildings
	Demand	2700	Buildings

Figure 1.2: The game interface

2. Running the game

The game requires you to make decisions about real estate transactions. The unit you transact in is Buildings. One Building contains about 200,000 square feet. As you will see below, its price is about $22 million, and brings $5 million per year in Rental Revenues. You initially have 30 buildings (bought at $19.44 million each and financed 100 %. Thus your initial loans are $ 583 million.)

Measurement Unit

My Buildings	**Existing Building Purchas**	0	Buildings/year
	New Permits	1	Buildings/year
	Blds u/Cstr (next year: 33%)	3	Buildings
	Buildings	30	Buildings

You start the game by clicking the mouse on the "Make Decisions" button in the bottom of your screen. Having done that, a new scrolling window (figure 2.2) appears in the bottom left corner of the screen. You first have to indicate your expectation for the value of a building 2 years into the future. You can either scroll using the mouse or type in your decisions of leased space to buy or sell.

Figure 2.1 Make Decisions

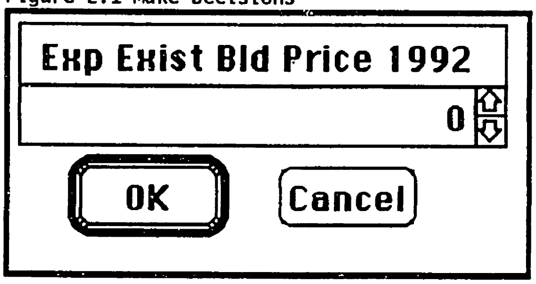


Figure 2.2: Forecast the selling price for an existing building in new shape.

Next you enter a similar forecast for what you believe rents will be two years down the road.

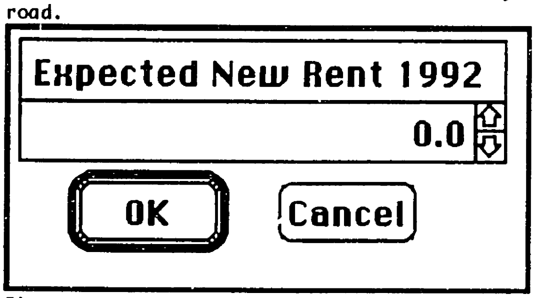


Figure 2.3: Forecast new rent.

Then you enter your decisions concerning space to buy, sell and develop. Upward scrolling translates into purchasing. Downward scrolling translates into selling. If you attempt to type decisions exceeding your financial limits or attempt to buy or sell too much space, you will be prevented from doing so. When you invest in Existing Buildings, the tenants remain in the building.

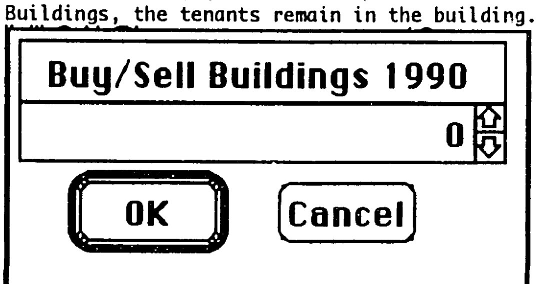


Figure 2.4: Investments(/disinvestment) in existing, leased space

After clicking on "OK" or using the "Return" keyboard button, a new window will appear in the same location. You are then ready to enter your permit requests in the same way. These permits are invariably turned into actual construction. "Cancel" at any time and you are back to clicking the "Make Decisions" button again. Figure 2.5 shows this dialog button.

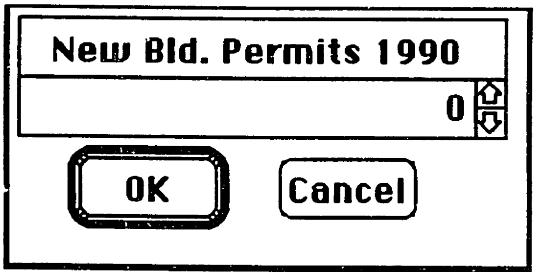


Figure 2.5: Submission of Permits to Develop new buildings, which will eventually be constructed.

Investments are 100 % lender financed, but the lender requires at least 10 % guarantee from the investor. This guarantee is based on your Bank Balance and hidden reserves. The latter source derives from the difference between Book Value and liquidation value of your buildings. Thus you are strongly limited in your transactions by your cash and hidden reserves.

The value of existing space depends only on expected operating profits no matter how much existing space is bought and sold. However, there is a transaction limit; No more than 10 % of competitor space can be bought in one single year. Likewise, your competitors will never increase their space in a single year by more than 10 % (by buying from you). However, they can develop space in addition the these percentages, ... and so can you.

Exhibit 4.4.2: The game description.

The decision maker was instructed to maximize realized net worth. This implied a maximization of the bank balance since the physical assets (Buildings or Ships) were always 100 % externally financed. In the game, financial institutions lend money as long as the player has a 10% equity to back up the investment project. This amount is not used in the project, but as a guarantee for the lender. Both liquid assets and unrealized capital gains in Buildings/Tankers count towards this percentage. Since the unrealized gains inflate in good times, the player's ability to finance is better when prices are rising than when values are depressed.

Bankruptcy occurs when a decision maker does not have enough cash-flow to meet obligations. In such a case, the decision maker was reset with the initial assets, but continued playing with the same market environment. The net liquidation loss (or in the far rarer case; the gain) was put into an "invisible" account and subjected to the same interest rate as other loans. However, this accumulation of "invisible" loans was assumed to be done by an offshore bank that did not communicate with the regular financial community. The invisible loan therefore did not directly influence later decision making. The invisible loan, including accrued unpaid interest payments, was deducted from the sum of profits in the final decision round.

Subjects made two decisions every period; they could develop new properties/order new tankers or buy or sell properties/tankers from competitors with no delay. New construction was paid when it became productive after a production lag.

The new construction price was constant, whereas the value of existing assets varied as a function of the average operating income in the market.

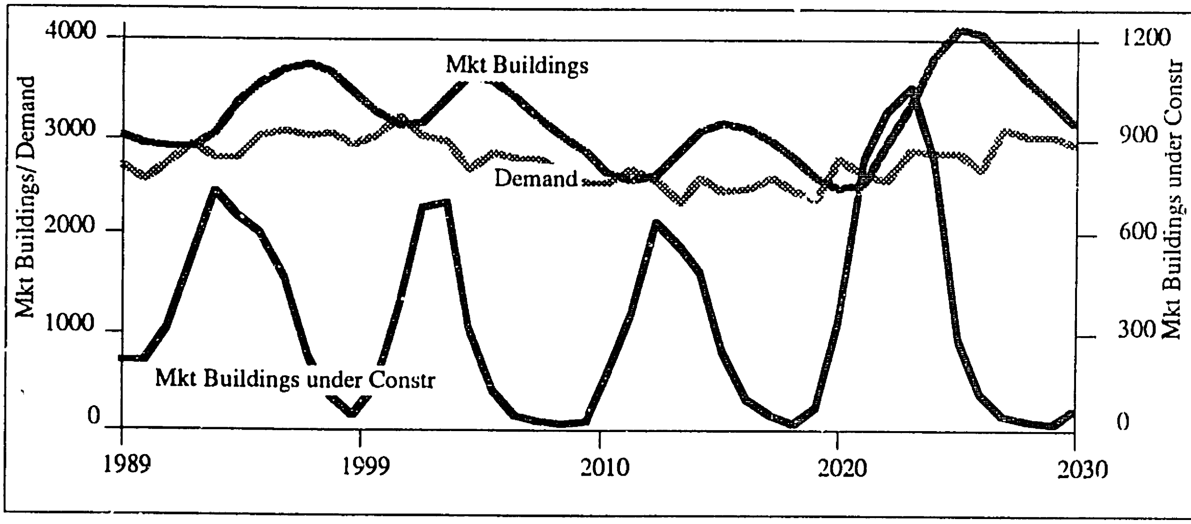
Implementation of high and low market compression

Compressed market dynamics were created by having a resonant period of about 9 years and the less compressed market had a resonant period of about 21 years. Differences in compression were implemented through the simultaneous change of three parameters: Completion time, Average life time and the price elasticity of orders, or Effect of price on orders (EPOR). EPOR was modified so as to ensure reasonable stability congruence between the two frequency environments.

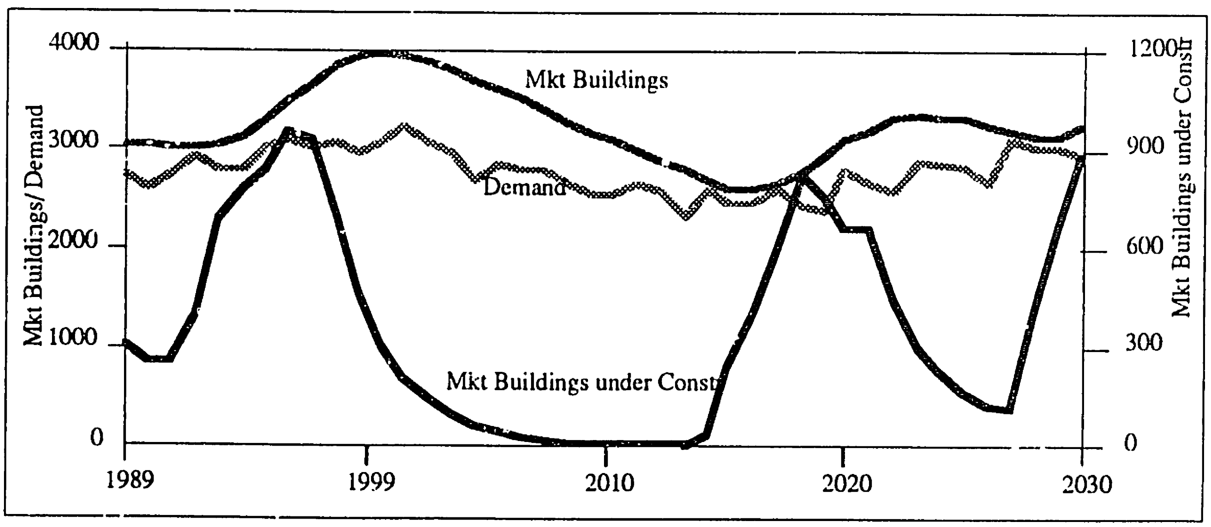
Frequency	Completion Time constant	Average Asset Life Time	= Loan Repayment Schedule	Price elasticity of Orders	→Cycle Period
High	1.5 years	15 years	15 years	6	≅ 7-11 years
Low	3.0 years	30 years	30 years	4	≅17-24 years

Table 4.4.3: Input parameters and output behavior in high and low frequency markets

Figure 4.4.4 and 4.4.5 below show the behavior of the simulated markets under the assumption of high and low frequency. Both markets have been subjected to the same demand pattern so the difference in system frequency is caused by the different time constants and EPOR values.



4.4.4: High frequency market



4.4.5: Low frequency market

Demand patterns

Demand was completely exogenous. To ensure that subjects were not helped by memorizing the exact timing of peaks and valleys, two different demand patterns were introduced. Figure 4.4.6 shows a fast market with the second demand pattern. Comparing the market response to the two demand patterns by comparing figure 4.4.6 to figure 4.4.4, the markets appear qualitatively similar but not identical.

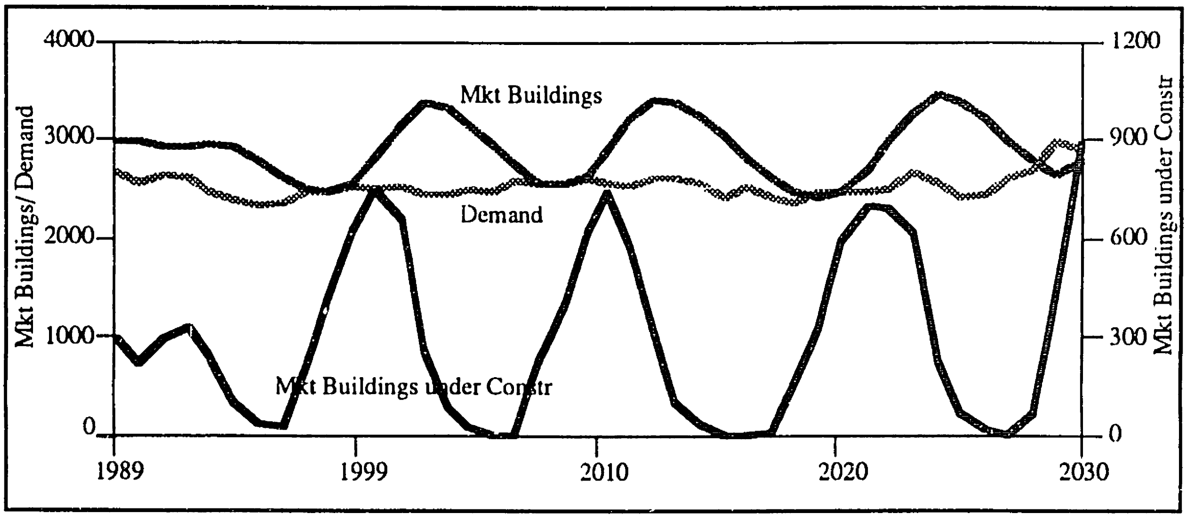
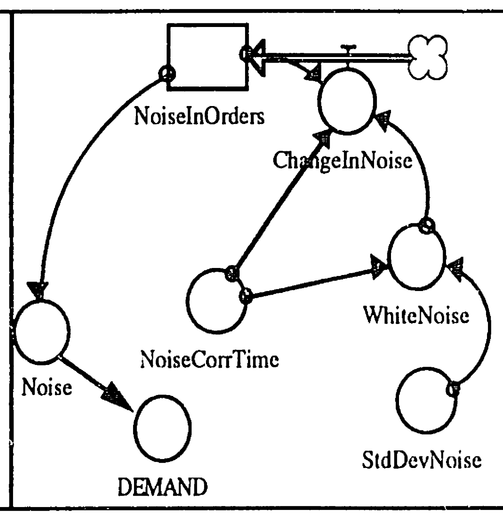


Figure 4.4.6: High frequency market; second demand pattern

Table 4.4.7 below shows the formulation for the demand pattern. Demand is a pink noise, i.e. the demand in a period is correlated with the demand in the previous period. This formulation was run twice in a row. The first run produced "demand 1" and the second "demand 2".

NoiseInOrders = NoiseInOrders + dt * (ChangeInNoise)
INIT(NoiseInOrders) = 0
{Correlated Noise in Orders (dimensionless)}
ChangeInNoise = (WhiteNoise-NoiseInOrders)/NoiseCorrTime
{Change in correlated Noise (1/years)}
DEMAND = K*Noise
Noise =(1+NoiseInOrders)
NoiseCorrTime = 4
{Correlation time for Noise in orders (years)}
StdDevNoise = .05
{Standard Deviation of Noise in Orders (dimensionless)}
WhiteNoise = StdDevNoise*((24*NoiseCorrTime/DT)^.5)*(RANDOM-.5)
{White Noise in Orders (dimensionless)}



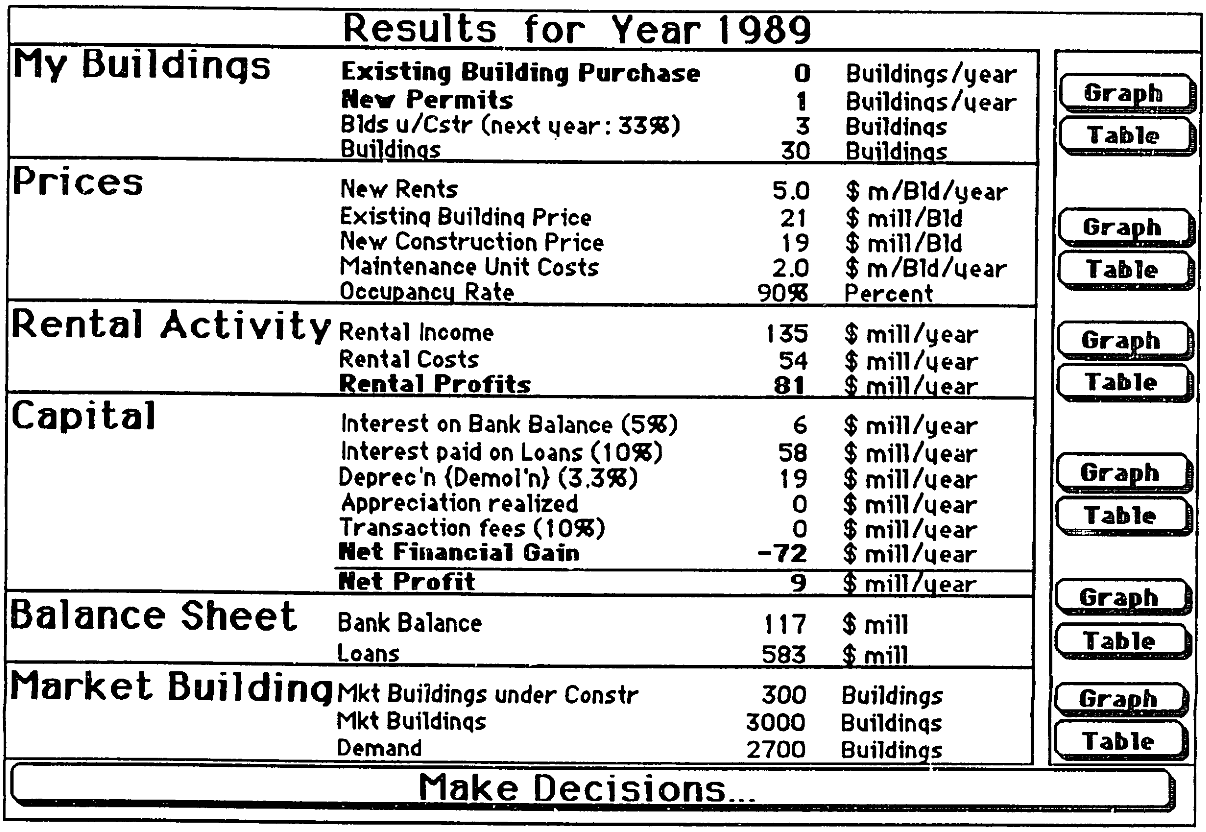
K real estate = 2700

K tanker = 10800

Table 4.4.7 The demand noise formulation, using a weak auto-correlation

Information display

The simulated decision environment filled a 9" Apple Macintosh display. It is shown below in approximately 50% reduction (30% width and 30% height). The two decision variables **Existing Building Purchase** and **New Permits** were at the top and all information grouped in six logical clusters. The decision screen contained information about the current year only. Information about prior simulated history was obtained by a mouse-click in the buttons called Graph or Table to the right of variable clusters



Results for Year 1989

My Buildings	**Existing Building Purchase**	0	Buildings/year	Graph / Table
	New Permits	1	Buildings/year	
	Blds u/Cstr (next year: 33%)	3	Buildings	
	Buildings	30	Buildings	
Prices	New Rents	5.0	$ m/Bld/year	Graph / Table
	Existing Building Price	21	$ mill/Bld	
	New Construction Price	19	$ mill/Bld	
	Maintenance Unit Costs	2.0	$ m/Bld/year	
	Occupancy Rate	90%	Percent	
Rental Activity	Rental Income	135	$ mill/year	Graph / Table
	Rental Costs	54	$ mill/year	
	Rental Profits	**81**	$ mill/year	
Capital	Interest on Bank Balance (5%)	6	$ mill/year	Graph / Table
	Interest paid on Loans (10%)	58	$ mill/year	
	Deprec'n (Demol'n) (3.3%)	19	$ mill/year	
	Appreciation realized	0	$ mill/year	
	Transaction fees (10%)	0	$ mill/year	
	Net Financial Gain	**-72**	$ mill/year	
	Net Profit	**9**	$ mill/year	
Balance Sheet	Bank Balance	117	$ mill	Graph / Table
	Loans	583	$ mill	
Market Building	Mkt Buildings under Constr	300	Buildings	Graph / Table
	Mkt Buildings	3000	Buildings	
	Demand	2700	Buildings	

Make Decisions...

Figure 4.4.8: Main screen, real estate market

By clicking on Graph, a figure similar to the one below filled the screen. The graph contained information up to the current decision period (i.e. year). Figure 4.4.9 below shows the graph available about buildings and construction activity in the simulated year 2011. Note that this and all graphs went to 2039, though all games, went only to 2029 to prevent end-of-game effects.

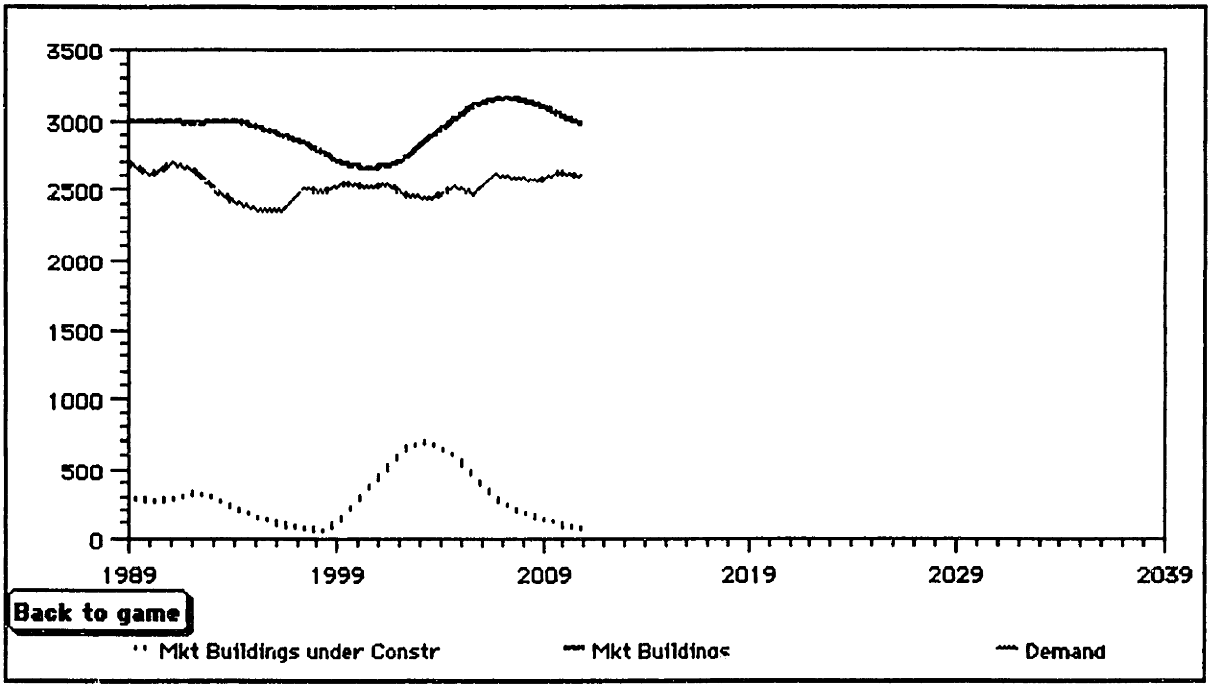


Figure 4.4.9: Graphs over time

Adjustments to represent the oil tanker industry

The difference between the two market contexts consists of context-specific naming of all but "Capital" and "Balance Sheet" variables. The two contexts also differed in terms of absolute scaling. Relative scaling of variables within each context was identical for the two contexts. All time constants and accounting were implemented identically. The decision making screens were also identical in all aspects but the naming and scaling difference. Though the scales of demands were different, their patterns were identical in the two contexts. The initial endowment, relative to total market size, was identical (120 Tankers and 30 Buildings, and $93 and $117 million in bank balances respectively) and represented 1% of the total market in each context. Figure 4.4.10: The oil tanker game screen is shown in figure 4.4.10 below.

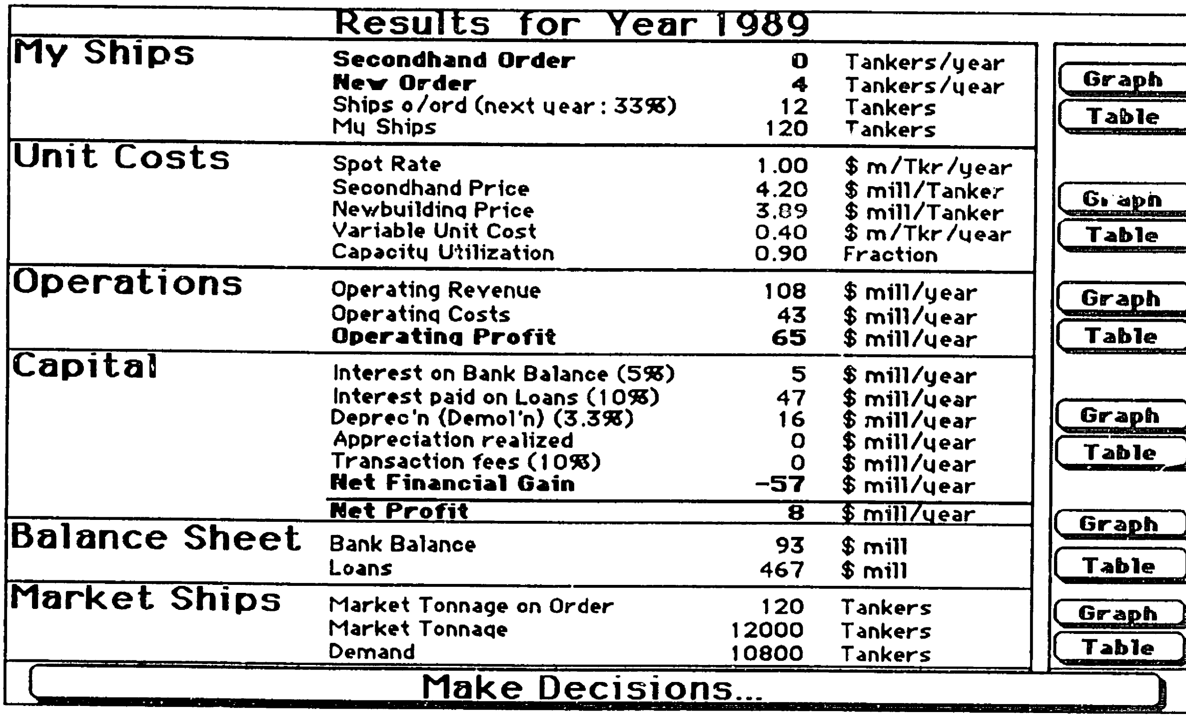


Figure 4.4.10: The Tanker decision making screen

4.5 Benchmarks

In order to assess performance, learning and transfer in the experiments, optimal performance or some other, heuristic, method of comparing the subjects in the various conditions must be established. A benchmark decision rule can gauge performance in decision tasks. This section explains the plausibility and robustness of such a rule.

In cyclical markets, "Buy-Low, Sell-High" (BLSH) rules tend to do well (Marcus, et al., 1991). Such rules assume that if an asset price is lower than some equilibrium value, (usually the combined cost of input factors and transaction expenses) then a purchase should be made. In the "dead zone" between buying and selling, the decision maker should neither buy nor sell. Once the price reaches the equilibrium plus some transaction cost, the decision maker should liquidate his assets. Simple as such a rule might seem, it nevertheless outperforms other reasonable rules in similar markets (Marcus, et al., 1991). The problem of specifying such a decision rule resides in making it

a) Psychologically plausible;

b) Robust with regard to the external environment regardless of exact rule parameters;

c) Dependent only on available information.

Rules need to be psychologically plausible. The one used below requires that subjects understand that they operate in an unstable market. Subjects should know that they do: Before playing the first trial, subjects receive information about the market' structure through the reading of a context specific newspaper article (see appendix 3). In these articles, subjects are presented with indications of market instability and structural explanations of how the markets work.

In addition, subjects have access to the behavior shown in figure 4.4.9 as it unfolds in the game. Market instability should be obvious. Thus, a benchmark tacitly assuming instability is fair for subjects who are able to infer instability from the provided background materials and for subjects who access available simulation time history.

Marcus et al. (1991) found that BLSH rules are robust with regard to "window of opportunity" parameters. Yet, preliminary analysis in the present task showed that a simple buy low, sell high rule based on the current asset price was extraordinarily sensitive to the "window of opportunity" parameter values; high profits in one frequency environment gave poor profits in another. The cause of poor robustness of a simple rule is related to the fact that the experimental markets remain depressed for a long time. A simple rule would kick in under circumstances with low prices in combination with dismal future prospects and bankrupt the oil tanker or real estate player. Consequently, another rule was chosen.

This decision rule incorporate the quoted real estate analysts' "20/20 hindsight" as well as the reflections of the thoughtful shipping investor, in particular with respect to market cues that often are overlooked. Both quotes refer (Conway and McKinley, 1981; Bakken, 1990b) to an "early warning indicator" incorporating insight about the system's delay structure in the form of competitor assets under construction. If perceived, assets under construction, i.e. supply lines, have high diagnostic value.

Buying takes place in a market that is still dismal and getting worse, but with signs of future improvement. This sign of improvement consists in the rate of decline in the assets under construction is decreased. Similarly, it sells at a time when the supply line is still increasing, but at a diminishing pace. The rule, using the changes in the supply line, contains a number of desired features. First, it reflects the rule of a decision maker who has system insight: It is likely that decision making in this dynamically complex market will be more robust, i.e. better in the long run, if rules reflect the markets' deep structure more than their apparent surface cyclicality.

Second, though the rule is easy transferable, transfer will be problematic because the derivative is an extensive concept (Bassok, 1991), not readily available in subjects' problem space. Yet the rule is easy to compute since the inflexion point of the supply line (which is the cue to transact) is readily available to subjects in a graphical format.

Specifically, the timing rule was

Order new assets, ΔKn, when $\Delta SL_t > \Delta SL_{t-1}$ and $P < (1-\alpha_{new})\ NC$ (1)

Buy old assets, ΔKe, when $\Delta SL_t > \Delta SL_{t-1}$ and $P < (1-\alpha_{buy})\ NC$ (2)

Sell assets, $-\Delta Ke$, when $\Delta SL_t < \Delta SL_{t-1}$ and $P > (1+\alpha_{sell})\ NC$ (3)

where

SL = Supply Line of ordered, yet undelivered assets in the market and

$\Delta SL_t = SL_t - SL_{t-1}$

ΔK_n = Investment in new, time delayed, assets

ΔK_e = Investment in existing, immediately available, assets

P = Price of existing assets

NC = Cost of New assets

α_i = Transaction margin in percent of New Cost. Transaction costs were 10 %, thus $.1<\alpha_i<1$

NC, P, and SL were pieces of information about the market available on the screen.

NC was a constant; P and SL were variables.

Though the α_i 's need not be equal, using $\alpha_{new} = \alpha_{buy} = \alpha_{sell}$ ensures a symmetrical timing rule.

Hence, (1), (2) and (3) above each defines decision timing unequivocally as a function of α_i.

The benchmark rule system has a psychological interpretation if the supply line signal is sufficiently smooth (which it is in this task). In the real world the supply line signal is noisy and must be averaged over a sufficiently long time horizon. The use of a rule based on the derivative of a noisy signal may cause unstable system behavior, as discussed in the control engineering literature (Franklin, et al., 1986; Friedland, 1986).

Second, it may be argued that it is hard for subjects to **calculate** the derivative. However, calculating the rate of change in the supply line is easy, if the supply line itself is available. In this experiment supply line information is as prominent as demand and supply. Real supply lines, however, tend to be less available than the stock of productive assets. As an example, shipyards do not make public their order books. Shipping investors must often infer supply lines from other sources, such as labor statistics.

Such inferences are problematic, and supply line information often biased: Tanker builders have an incentive to tell prospective clients that future markets will be rosy. They do so by providing low estimates of current world supply lines (i.e. order books). Underestimation helps shipbuilders sell new ships.

The lack of unambiguous supply line information in the real world may be why decision heuristics in experimental markets are systematically under-weighted (Sterman 1989a, b; Bakken 1990b).

Decision timing is given by the above equations (1), (2) and (3). (4), (5) and (6) below define purchase or sale amount as a function of the aggressiveness of the transaction, β_k, the fraction of potential maximum investment executed. The subjects could not transact more than 10% of the total market value in any single year.

$$\Delta K_n = \text{Min.}((\beta_{new} * L/NC), 0.1*MA) \qquad (4)$$

$$\Delta K_e = \text{Min.}((\beta_{buy} * L/P), 0.1*MA) \qquad (5)$$

$$-\Delta K_e = \text{Max}((-\beta_{sell} * OA), 0.1*MA) \qquad (6)$$

where

MA = Total number of assets in the market,

OA = Total number of assets operated by the decision maker,

L/NC, L/P, 0.1*MA and OA were limits calculated by the game and constrained decision making.

$0 < \beta_k < 1$, k = new, buy, sell reflect the aggressiveness of investment and divestment,

L = Leverage; the maximum amount that the player may borrow,

ΔK_n = Investment in new, time delayed, assets,

ΔK_e = Investment in existing, immediately available, assets,

- ΔK_e = Sales of ones assets,

β_{sell} = 1 in the following, since there are no risks involved in selling as long as the selling price is higher than the purchase price was. Selling out everything becomes a safe rule. Consequently, the performance of the BLSH rule was only investigated for various values of β_{buy} and β_{new}.

With a β_{new} or β_{buy} equal to unity there will be no liquidity in instances where operating profits stay low. β_{new} and β_{buy} may, however, be different, but for simplicity they are identical in the following.

By definition, supply lines in second order systems, such as cyclical markets, peak before stocks. Figure 4.5.1 shows the phase relationships between variables if the system is subjected to a 15 year sine wave input. The benchmark decision rule produces investments when markets are weak and still falling, and divests when markets are still rising. Furthermore, such a timing rule does not require knowledge of specific numeric values and is therefore directly transferable across contexts.

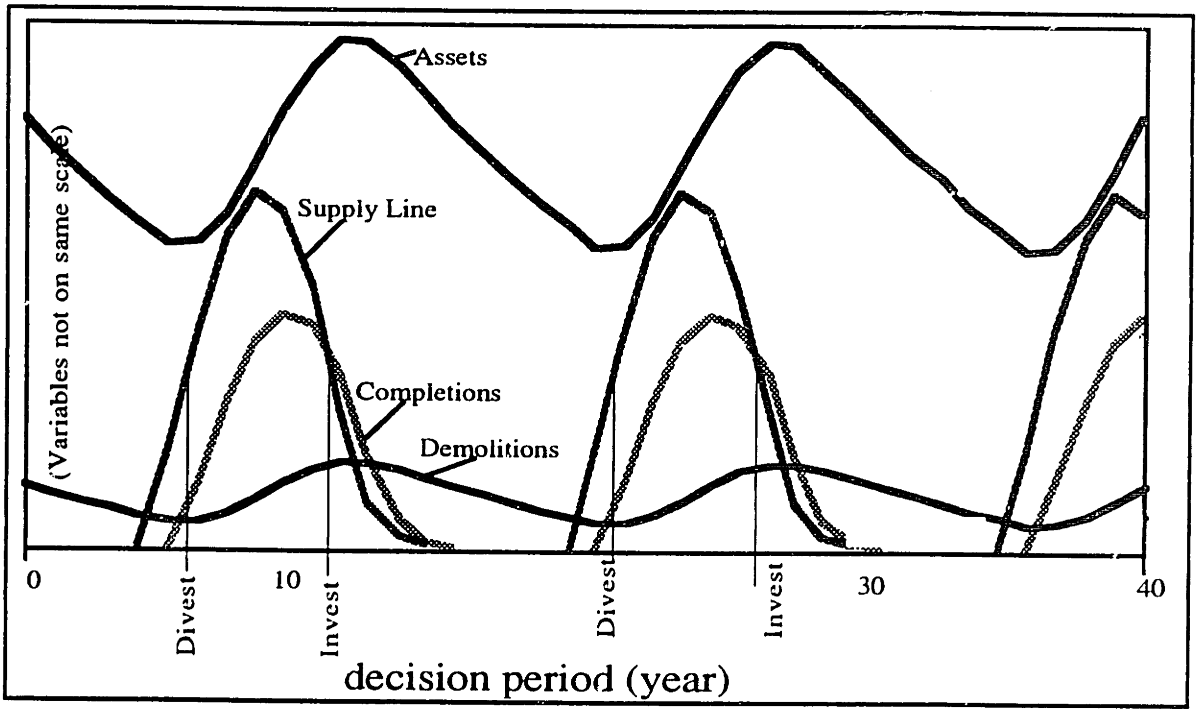


Figure 4.5.1: Decision timing as a function of a rule that reflects the structural knowledge that supply lines peak before stocks do.

Figure 4.5.2 shows the average outcome, as well as the robustness of the decision rule. Context only matters as a scaling difference between real estate and oil tankers. Real estate markets had initial and final profits 25.8% higher than tanker markets.

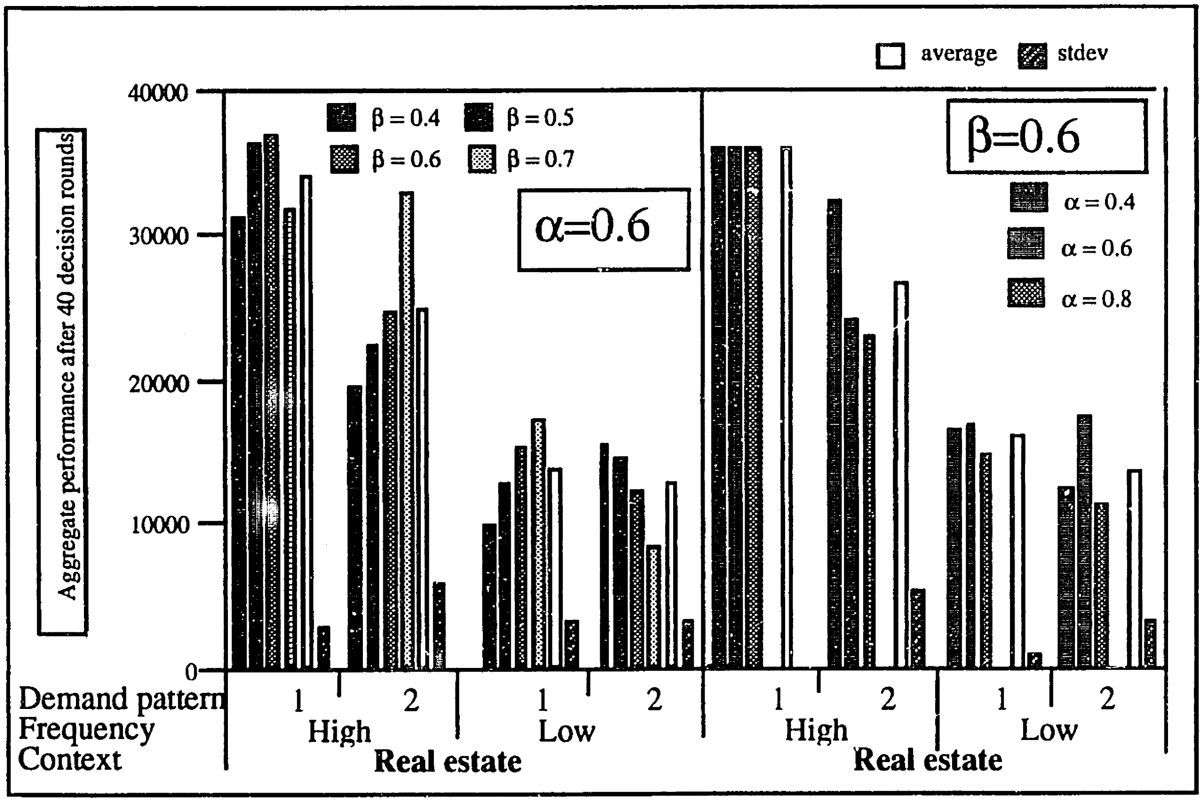


Figure 4.5.2: Aggregate profits produced by the benchmark rule as a function of varying α's and β's

The relatively flat profit surface as a function of α and β indicates that the exact α and β values do not matter much within the limits above and $\alpha = 0.6$ and $\beta = 0.6$ were used in the following.

4.6 Summary

This chapter has provided accounts of stock management task characteristics and shown that both oil tanker and real estate markets can be interpreted as stock management tasks. Moreover, the focus on stock system properties, such as long lead times, feedback reliability etc. help explain poor conditions for learning in these markets. Experimental markets, based on these two contexts, have been designed and benchmark rules formulated, tested, and shown to be robust.

5 Performance and transfer results

5.1 Introduction and overview

As outlined in chapter 1, simulated decision environments may help decision makers learn when real environments are not conducive to learning. Simulated environments are also laboratories where researchers can investigate causes of problematic decision making (Paich and Sterman, 1992). At the same time, their learning effect can also be estimated. Investigations can be made unobtrusively since learning labs are, as the name suggests, laboratories where people make decisions and hopefully learn.

This chapter provides results from experiments in real estate and oil tanker contexts. Remember that there were 41 student subjects. They played a sequence of two trials, each of which comprised 40 decision periods. Compression of the markets was manipulated by changing time constants so that two market frequencies appeared.

The chapter is built up in the following way: The first section illustrates the differences between decisions made by one subject and those made by the decision rule. Though subject 11 was chosen because he is representative, only the statistical analysis provided later in the chapter gives a complete representation of subjects' performance and transfer.

The following section expands on how performance is influenced by the main treatments of context and market frequency. The last section integrates and discusses performance and transfer hypotheses and findings.

5.2 Performance example

Subject 11 has been chosen because his decisions are illustrative and informs the reader of the differences between decisions made by the rule and by the subject. The example also illustrates the action feedback nature of the task.

Starting in the high frequency oil tanker environment, the subject and the rule both begin with 1% of the market assets. They also face an identical demand pattern and to a large extent face the same market conditions: The rest of the market, i.e. competitor ship owners, follow identical decision rules. Yet the subject, due to the consequences of his timid decision making, faces a financial decision environment that less and less resembles that of the rule. This difference resides mainly in the fact that while the rule accumulates profits when selling many tankers when they are highly valued, the subject fails to sell many ships when prices are high. The rule uses this huge bank balance to invest during times of depressed asset values to further build up a fleet that is sold again when prices are high.

Since the subject fails to sell ships, the subject develops less liquidity cushion and hence cannot buy as much when prices are low. Not taking advantage of inflated prices during boom years, subject 11 sails many newly built and expensive ships when operating income hardly covers operating costs. The subject thus periodically experiences large financial losses.

The profits are very different for the subject and the benchmark rule. In addition, the total market is affected differently by the subject and by the rule: The rule is more aggressive and so develops a higher market share that again influences market behavior. At the extreme, had the rule controlled the market, the task would have become a pure stock management task(1992), but even the rule never controls more than 20 % of the market. The lack of dominance in the market is mostly due to the constraint that no more than 10 % of the competitors' asset can be purchased in a single year.

These differences between the subject and the decision rule underline what is meant by action feedback: The decision environment responds to the actions of the decision maker. It is the same difference that lets the rule accumulate assets and so end with over $26 billion, whereas the subject never really gets off the ground and ends up with only 5 % of that. The 41 subject average was similarly 4.4 % of the benchmark on first trial performance.

The subject and his environment are shown left in figures 5.1.1 to 5.1.4 below. The benchmark rule is shown right in figures 5.1.5 to 5.1.8. Figures 5.1.1 and 5.1.5 below are not identical but very similar.

Subject 11 is conservative. He never buys more than 20 tankers in the second-hand market in any single year and never orders more than 25 tankers before 2028 when he orders 50 tankers 3 years in a row. Likewise, he never sells more than 20 tankers before 2023 when he unloads 80 tankers onto the market.

The benchmark rule orders a maximum of 800 tankers already in 2003, and provides massive orders above 600 4 more times. Likewise, the benchmark rule sells more than 500 tankers four times. In generai, subject 11 sells **later** than the rule. Figure 5.1.2 indicates that lateness in selling may stem from the use of price as the **only** signal to time decisions. Consequently, he sells after prices peak. The benchmark rule, on the contrary, sells when the increase in tanker orders is slowing down. Because it takes one year to implement a decision in the second-hand market, and 1 to 4 years in the new building market, the rule generally gets higher selling prices and lower buying prices.

In addition to later timing, that may be caused by lack of supply line considerations, subject 11 does poorly because he buys less aggressively than the rule. As shown by comparing figures 5.1.3 and 5.1.7, the subject loses compared to the rule because he keeps a substantial part of his fleet intact even during boom years, when the rules sells everything at high prices. Subject 11' s β_{sell} was computed to be 0.24 (sd = 0.22), which was lower than the 0.32 trial 1 subject average, and well below the unity β_{sell}. used by the rule. The rule's β_{buy} and β_{new} were both equal to 0.6, whereas subject 11 like other subjects had extremely low parameter values (< 0.01). As will be shown in table 5.4.8, 41 subject average was 0.03 for β_{buy} and 0.06 for β_{sell}.

Though the rule is not influenced by its own past decision behavior, it appears that the subject uses past decisions as an anchor. Subject 11 keeps steady stream of orders for new ships every year. The rule orders and buys ships during short spurts, but at these times a large number of ships are transacted and orders for new boats placed.

Note, however, that subject 11's selling decisions become less timid with experience. A t-test of the mean scores for the first half of the game (20 decision periods of which the subject sold 6 times) and the second half (20 decision periods of which the subject sold 12 times), revealed that the mean β_{sell} increased from 0.08 (sd = 0.07) to 0.32 (sd = 0.23) significant at the $p < 0.005$, level (investigated as separate variances).

In sum, the decision strategy used by subject 11 differs from the rule in two important aspects. The subject acts too late and misses peaks and troughs. In addition, the subject is more timid in his decision making, something that perhaps reflects an anchoring and adjustment decision making process (see Kahnemann and Tversky, 1973 for a further discussion of anchoring and adjustment) where initial decisions act as anchors.

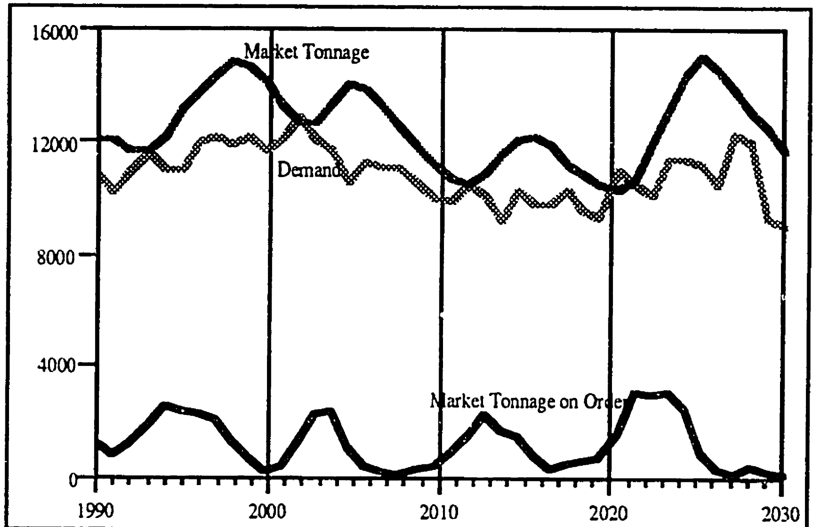


Figure 5.1.1: Demand, Supply (Tonnage) and
Supply line (Tonnage on order) faced by subject 11

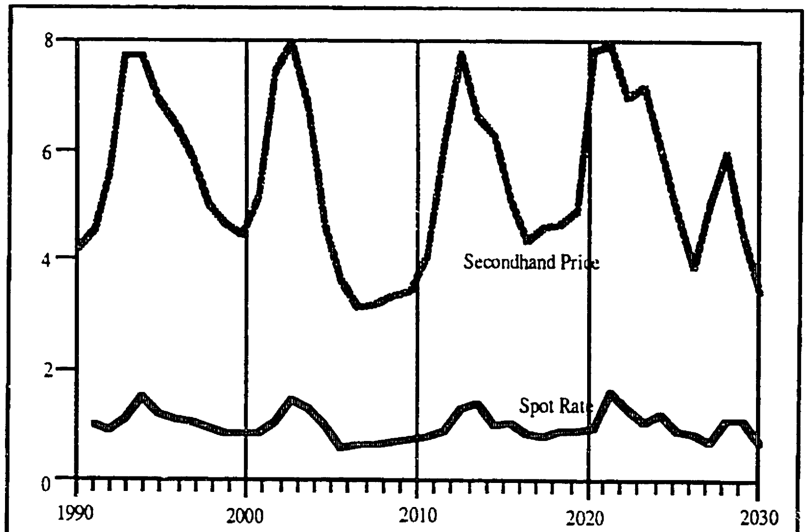


Figure 5.1.2: The price of assets to be traded, secondhand price;
as well as operating income faced by subject 11

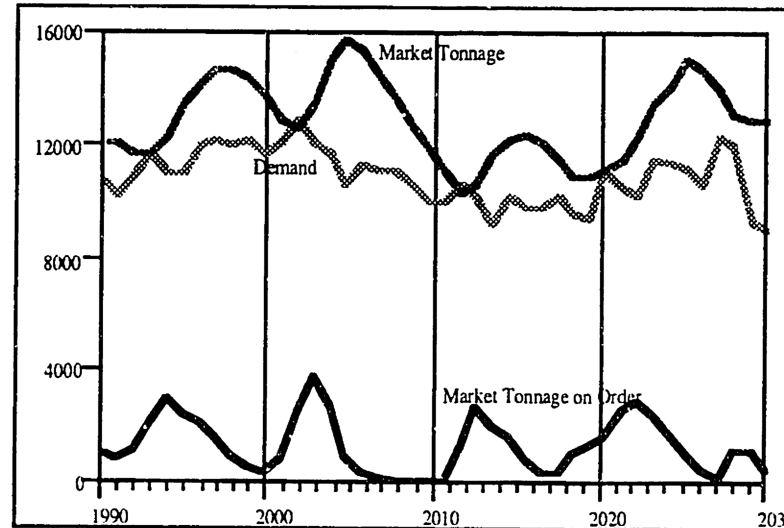


Figure 5.1.5: Demand, Supply (Tonnage) and
Supply line (Tonnage on order) faced by decision rule

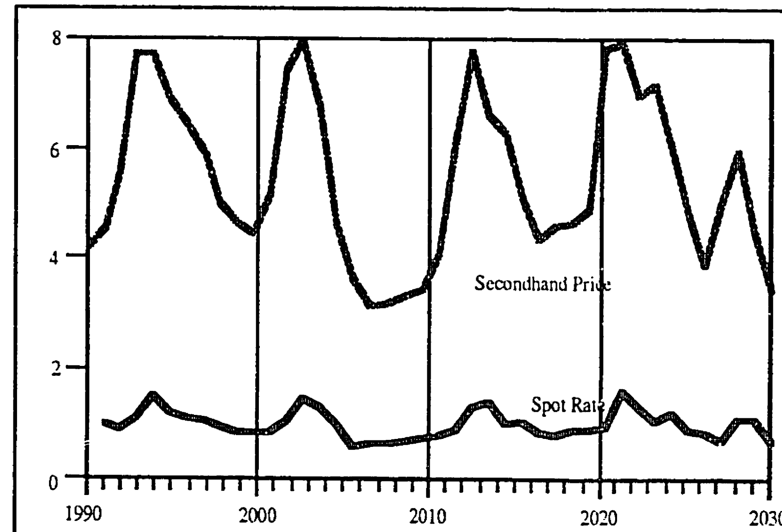


Figure 5.1.6: The price of assets to be traded, secondhand price;
as well as operating income faced by decision rule

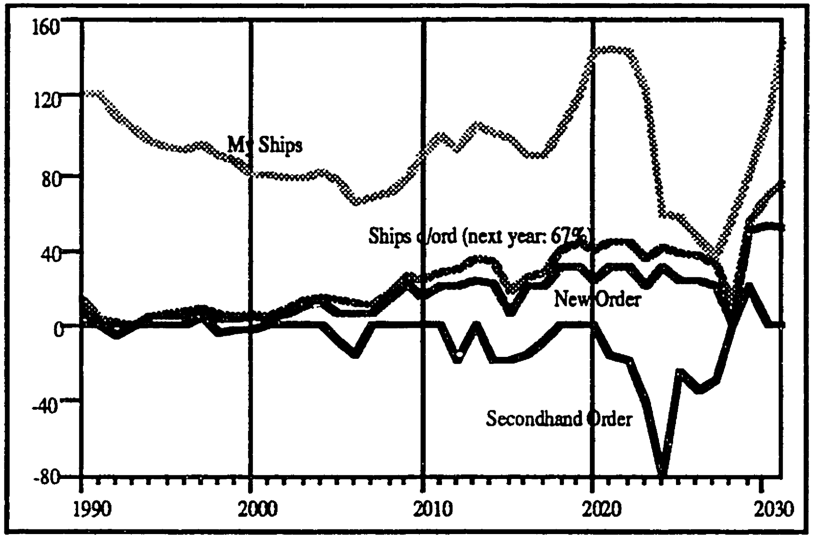


Figure 5.1.3: Decisions made by subject 11

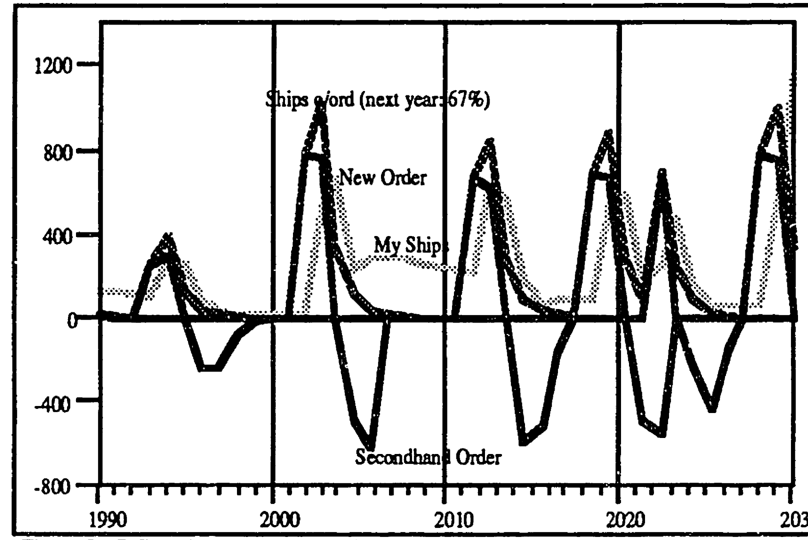


Figure 5.1.7: Decisions made by decision rule

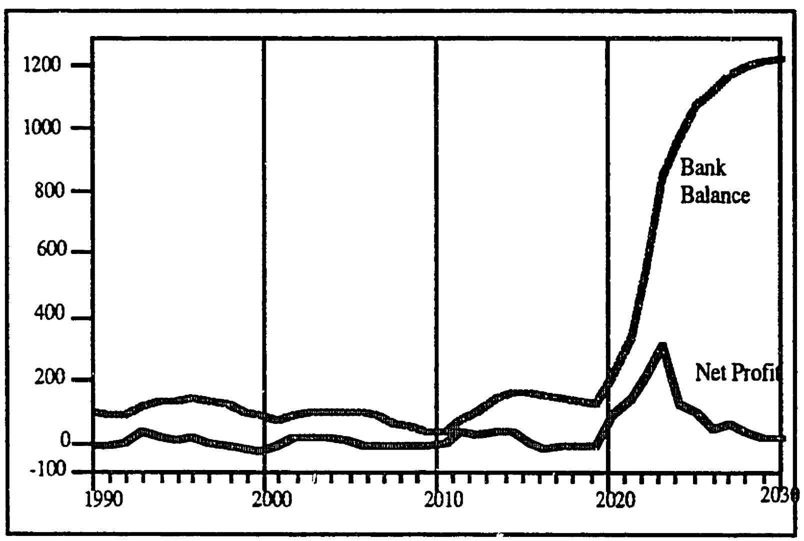


Figure 5.1.4: Profits and Bank Balance achieved by subject 11

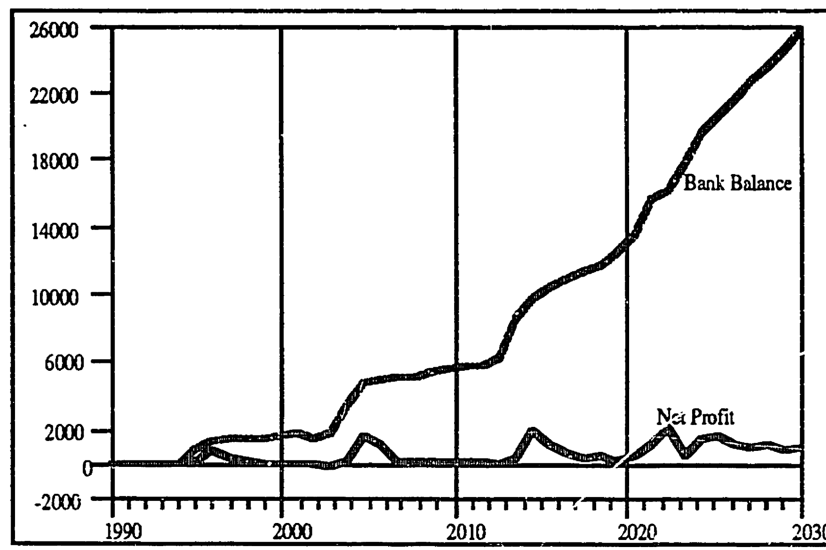


Figure 5.1.8: Profits and Bank Balance achieved by decision rule

Figures 5.1.9 and 5.1.10 show the second trial, when the subject faces a low frequency, real estate environment. The rule and subject decisions are more similar in the second trial. Subject 11's β_{sell} is now over 0.50 (sd = 0.32), which is significant improvement from trial 1 (paired t-test @ $p < 0.02$) reflecting that sample average β_{sell} increases from 0.32 (sd = 0.22) to 0.47 (sd = 0.28) which is also significant (paired t-test @ $p < 0.01$). Just like other subjects, subject 11 increases his β_{buy}, to 0.05 (average of subject's $\beta\beta\upsilon\psi$ = 0.07). Similarly, his β_{order} is decreased from 0.007 to 0.002 (paired t-test @ $p < 0.04$) reflecting that subject average also decreases from 0.025 to 0.016 (ns).

Also in this second trial is there evidence, however, that subject 11 anchors on early decisions, while the rule is more jagged. The benchmark rule makes one or a couple of transactions before it "rests" and is inactive for long periods. The rule, then, behaves similarly to the account of the decision model used by the thoughtful shipping investor in chapter 4. Subject 11, on the other hand, appears to make decisions more like the bulk of the interviewed decision makers and anchors on past decision behavior.

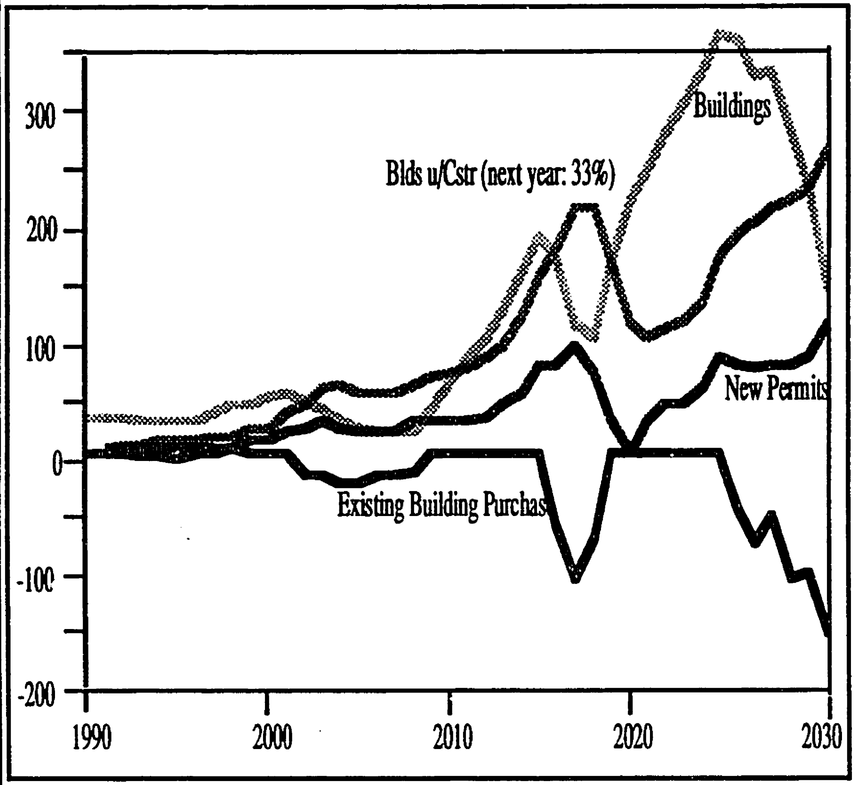


Figure 5.1.9: Subject 11's second trial

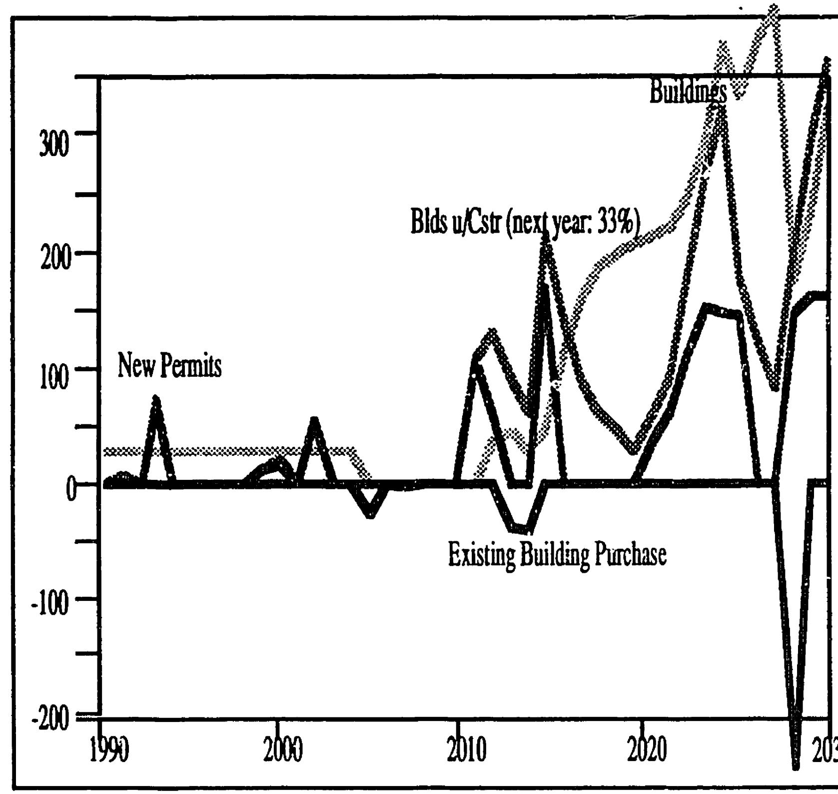


Figure 5.1.10: Decision rule for same market

The next section will look into more representative, statistical, accounts of the differences between benchmark and subject decisions. In particular, the influence of the treatments on performance and transfer is investigated.

5.3 Performance and transfer results

In this sub-section the results will be presented briefly. The next subsections will explain and discuss the findings in more depth. In this and subsequent presentation, the subscripts of context and frequency corresponds to the initial hypotheses. Thus, context = 1, refers to the real estate, i.e. familiar, context (where subjects are expected to do well). Context = 2 represent unfamiliar, i.e. oil tanker, markets. Similarly, frequency = 1 refers to high and frequency = 2 refers to low frequency.

For clarity of presentation, the output from the computer program have been used directly in the text. Thus, effect sizes are β-values from regression equations. For simplicity, all effect sizes are shown, though the insignificant ones have been shaded. Similarly, the variable names from the computer program have been retained and should be self-explanatory. In cases of ambiguity, variable names are explained, however. For example, CONT1, FREQ1, CONT2, FREQ2 refers to context and frequency in trial 1 and trial 2 respectively.

The ANOVA tables have been simplified somewhat compared to the customary design, however, to avoid redundancy: The sum-of-squares column has been deleted as it can be computed using degrees of freedom and mean-squares. In general, effects up to 0.1 level of significance are discussed (two-tail).

The analysis will chiefly investigate relationship between how well subjects do, as explained by the treatment conditions. One performance measure is used throughout the next two chapters. It is explained below.

5.3.1 Performance measure

The performance measure, PM has the following formulation:

$$PM_s = \frac{\sum_t [(\pi_{s,t}|E_{s,t}) - (\pi_{b,t}|E_{s,t})]}{\sum_t (\pi_{b,t}|E_{b,t})}$$

Where

PM_s = Performance Metric for subject s

$\pi_{s,t}$ = subject s' profit for decision t in the following year

$\pi_{b,t}$ = benchmark profit for decision t in the following year

$E_{b,t}$ = information and leverage environment available to benchmark before making decision t

$E_{s,t}$ = information and leverage environment available to subject s before making decision t

t = 1, 2, ...,40 (number of decision rounds for one trial)

s = 1, 2, ..., n

PM can be explained in the following way:

Every decision period, the benchmark profit is subtracted from the subject's, using the subject's environment for both. The subject's final performance, PM, is the accumulation over 40 periods of this difference, normalized using the raw benchmark score presented in chapter 4.

The next section will look into how the treatment conditions can explain PM in the first trial.

5.3.2 First trial performance

The model of performance in the first trial can be stated as

$$PM_1 = \alpha_0 + \beta_1 C_1 + \beta_2 F_1 + \beta_3 C_1 F_1 + \varepsilon$$

where C_1 is the context in the first trial, F_1 the frequency in the same trial, and $C_1 F_1$ the interaction between the two. C_1 is a dummy variable that takes on the value 1 (familiar i.e. real estate) and 2 (unfamiliar, i.e. oil tanker). F_1 can take the value 1 (high frequency) and 2 (low frequency). α_0 is the regression constant and the β_i's the regression coefficients.

Table 5.3.1 below shows cell sizes . Though not identical, the cell sizes are similar.

		FREQ1		
		HIGH	LOW	TOTAL
CONT1	FAMILIAR	8	11	19
	UNFAMILIAR	13	9	22
	TOTAL	21	20	41

Table 5.3.1: Cell sizes in first trial.

The cell means and standard deviations, σ, are shown below in table 5.3.2. Standard deviations reflect effect sizes, and a data transformation, using logarithmic or a square root transformation is suggested. However, since some PM's were negative, neither a square root or a logarithmic transformation could be used. Consequently, the raw data were analyzed. The higher levels of standard deviation in high performance environments is explained in a later subsection, however.

		FREQ1		
		HIGH	LOW	AVERAGE
CONT1	FAMILIAR	0.05 (σ=0.05)	0.36 (σ=0.40)	0.23 (σ=0.34)
	UNFAMILIAR	0.05 (σ=0.06)	0.14 (σ=0.13)	0.09 (σ=0.10)
	AVERAGE	0.05 (σ=0.06)	0.26 (σ=0.32)	0.16 (σ=0.25)

Table 5.3.2: Cell means and standard deviations in first trial

The analysis of variance provided in table 5.3.3 below indicates that high performance comes when the context is familiar and when frequency is low. There is also an interaction effect so that low frequency, familiar context is the one where subjects do far better than in the other cells.

The analysis of variance is shown below

ANALYSIS OF VARIANCE

DEP VAR: PM1 N: 40 MULTIPLE R: .578
SQUARED MULTIPLE R: .335

SOURCE	DF	MEAN-SQUARE	F-RATIO	P
CONT1	1	0.151	3.348	0.076
FREQ1	1	0.469	10.370	0.003
CONT1* FREQ1	1	0.164	3.621	0.065
ERROR	36	0.045		

ESTIMATES OF EFFECTS $B = (X'X)^{-1} X'Y$

		PM1
CONSTANT		0.158
CONT1	Fam	0.063
FREQ1	High	-0.110
CONT1 FREQ1	Fam High	-0.065

Table 5.3.3: ANOVA of Context, frequency and interaction effects in trial 1.

Since cell standard deviations were quite different, a non-parametric test of the context and frequency effects were performed. The two-sample Wilcoxon test was used to compare ranks for the context conditions first and for the frequency conditions next. The levels of significance were slightly lower using this non-parametric test as shown in table 5.3.4

	trial 1	
	context	frequency
Computed T Wilcoxon	438	323
Critical T Wilcoxon	437	342
significant @	$p < 0.1$	$p < 0.005$

Table 5.3.4: Wilcoxon tests in trial 1.

A discussion of the findings is provided in section 5.4.[12]

5.3.3 Performance in and transfer to the second trial.

As shown below in table 5.3.5, subjects do 33 % (0.053/0.158) better on average in the second trial, yet this difference is not significant using a paired t-test of the means.

```
     PAIRED SAMPLES T-TEST ON     PM1    VS     PM2    WITH    40 CASES

MEAN DIFFERENCE =        0.053
SD DIFFERENCE =          0.329
T =                     -1.024      DF =   39      PROB =      0.312
```

Table 5.3.5: Paired t-test of differences between first and second trial

Another way of investigating improvement from trial 1 to trial 2 is to perform a sign test of whether the percentage of subjects who improved is significant., compared to the null-hypothesis of 50% of the subjects will improve. Shown in table 5.3.6, 2/3 of the subjects improve, which is significant at $p < 0.04$ level.

[12] Note that PM1 is a kind measure of subject performance where subjects do better than the rule. This is explained in section 5.4. A simple metric of subject profits after 40 years is only 4.4% of benchmark.

COUNTS OF DIFFERENCES (ROW VARIABLE GREATER THAN COLUMN)		
	PM1	PM2
PM1	0	13
PM2	27	0
TWO-SIDED PROBABILITIES FOR EACH PAIR OF VARIABLES		
	PM1	PM2
PM1	1.00	
PM2	0.04	1.00

Table 5.3.6: Sign test of differences between first and second trial

The lack of significance of the t-test of performance reflects both high performance variance and little learning. The finding is not of grave concern, since the purpose of the investigation is to determine to what extent treatment conditions affect performance in the second trial.

The complete model of how the conditions contribute to transfer and performance in the second trial uses a predictors the performance in the first trial, the treatment conditions, and their two-way interactions is shown below using the notation from subsection 5.3.2.

$$PM_2 = \alpha_0 + \beta_1 PM_1 + \beta_2 C_2 + \beta_3 F_2 + \beta_4 C_1 + \beta_5 F_1 + \beta_6 C_2 F_2 + \beta_7 F_1 C_1 + \beta_8 C_2 C_1 + \beta_9 C_2 F_1 + \beta_{10} F_2 C_1 + \beta_{11} F_2 F_1 + \varepsilon$$

In this model of second trial performance, performance is explained by the treatments in the current as well as in the previous trial. Effects of initial conditions as well as changes in conditions between the trials indicate transfer. Note that only two-way interactions are investigated. Higher-order interaction terms would include extremely small cell sizes and tend to have low reliability. They were therefore included in the error term.

Table 5.3.7 below indicates number of observations per cell.

CONT2		FREQ2 HIGH	LOW	TOTAL
	FAMILIAR	10	9	19
	UNFAMILIAR	14	8	22
	TOTAL	24	17	41

Table 5.3.7: Cell sizes in second trial.

Table 5.3.8 below shows cell means of the trial 2 treatments, while table 5.3.9 further below shows the analysis of variance and effect sizes.

CONT2		FREQ2 HIGH	LOW	AVERAGE
	FAMILIAR	0.09 (σ=0.10)	0.53 (σ=0.53)	0.30 (σ=0.43)
	UNFAMILIAR	0.10 (σ=0.09)	0.16 (σ=0.08)	0.12 (σ=0.09)
	AVERAGE	0.09 (σ=0.09)	0.36 (σ=0.42)	0.21 (σ=0.31)

Table 5.3.8: Cell means and standard deviations in second trial.

ANALYSIS OF VARIANCE

1 CASES DELETED DUE TO MISSING DATA.

DEP VAR: PM2 N: 40 MULTIPLE R: .694
SQUARED MULTIPLE R: .482

SOURCE	DF	MEAN-SQUARE	F-RATIO	P
PM1	1	0.413	5.961	0.021
CONT2	1	0.439	6.343	0.018
FREQ2	1	0.325	4.693	0.039
CONT1	1	0.198	2.858	0.102
FREQ1	1	0.044	0.634	0.433
CONT2* FREQ2	1	0.391	5.643	0.025
CONT1* FREQ1	1	0.004	0.059	0.810
CONT1* CONT2	1	0.015	0.223	0.640
FREQ1* FREQ2	1	0.001	0.011	0.919
FREQ1* CONT2	1	0.059	0.858	0.362
CONT1* FREQ2	1	0.038	0.542	0.468
ERROR	28	0.069		

ESTIMATES OF EFFECTS $B = (X'X)^{-1} X'Y$

		PM2
CONSTANT		0.119
PM1		0.560
CONT2	FAM	0.112
FREQ2	HIGH	-0.097
CONT1	FAM	-0.078
FREQ1	HIGH	0.042
CONT2 FREQ2	FAM HIGH	-0.104
CONT1 FREQ1	FAM HIGH	0.011
CONT1 CONT2	FAM FAM	0.022
FREQ1 FREQ2	HIGH HIGH	0.005
FREQ1 CONT2	HIGH FAM	-0.041
CONT1 FREQ2	FAM HIGH	0.033

Table 5.3.9: ANOVA of performance and transfer effects in trial 2.

As mentioned, trial 2 performance is 31% higher than trial 1. This was shown in table 5.3.3, and may also be computed from table 5.3.9:

PM2 = 0 12 + 0.56 * PM1, where PM1 = 0.16 => PM2 = 0.21; 31 % higher than PM1.[13] The PM1 coefficient, β_1, is positive and indicates that performance in the two trials is correlated.

Just as in the equation for the first trial, the analysis of variance shows that subjects do better in a familiar and in a low frequency environment. Once again, there is an interaction between current frequency and context, reflecting that subjects in the low frequency, familiar environment do much better than in any of three other environments.

None of the changes between the trials or conditions in the first trial carry over to PM2 if a 0.1 level of significance is required. However, at the 0.102 level context 1 matters, albeit in the opposite direction of the one predicted: Later performance is helped by an initial unfamiliar context. As already mentioned in chapter 3, subjects in an initially unfamiliar context (where they do poorly, as shown in the previous analysis) may come to develop a deeper understanding. This contention is discussed further below.

In the regression equation, the C1C2 interaction would show the impact of context change. Though a change was expected to decrease PM2, no transfer effect was found.

Similarly, a transfer effect from frequency change would have given a significant F1F2 interaction. Though expected, no such effect was found. Likewise, the hypothesis stated a positive effect from an initial high frequency. No transfer effect from frequency was found.

---

[13] Note that PM2 is a kind measure of subject performance where subjects do better than the rule. This is explained in section 5.4. A simple metric of subject profits after 40 years is only 5.1 % of benchmark (or a 16% improvement over the same metric of PM1)

Due to the large differences in standard deviation between cell sizes, a less restrictive Wilcoxon test was performed on the main effects. This is shown below in table 5.3.10.

	trial 2	
	context	frequency
Computed T Wilcoxon	440	465
Critical T Wilcoxon	437	466
significant @	$p < 0.1$	$p < 0.1$

Table 5.3.10: Wilcoxon tests in trial 2.

The next section will explain the expected and unexpected effects in more detail.

5.4 Discussion of treatment effects on performance and transfer

This section first provides methodological clarifications with regards to the unexpected, opposite effect of frequency. Further below follows more thorough discussions of the context and frequency effects.

5.4.1 Methodological issues

Though the higher performance in low frequency environments can be partly explained by a beneficial "slow motion effect" that will be discussed further in a later subsection, another, related, reason for higher performance in low frequency environments is portrayed in figure 5.4.1 below.

As shown in chapter four, subjects make a disproportionate part of their profits in the years where asset transactions are advantageous. A subject who only operates and replaces the existing fleet will go bankrupt in the first cyclical downturn; subjects must transact by selling high and buying low to make substantial profits.

While there exist two to three short buying and a similar number of selling opportunities in the low frequency market, the high frequency market contains about four to five. In contrast to the decision rule, subjects must learn the system's causal relationship before they can adequately make transactions. As noted in section 5.2, subjects learn to be more aggressive with experience as far as buying and selling is concerned. It was shown that subject 11 became more aggressive during each trial.

To recognize and take advantage of transaction opportunities, subjects must realize the inherent unstable nature of the system. The low frequency condition allows a longer training period before transactions that distinguish high and low performance must be made. Thus, subjects in the high frequency condition miss out on the first profit opportunity. This is evidenced by the fact that β_{sell} increases significantly in both trials from first to second half of the game. In the first game, β_{sell} increases from 0.26 to 0.42 (paired t-test $p< 0.001$). Similarly β_{sell} increases from 0.45 to 0.55 in second game (paired t-test $p< 0.001$). No systematic differences in the purchase and ordering parameters were found over a trial, however.

In the table below the subject and the rule, which is replaced in the subject's decision environment each year, in the high and low frequency environments.

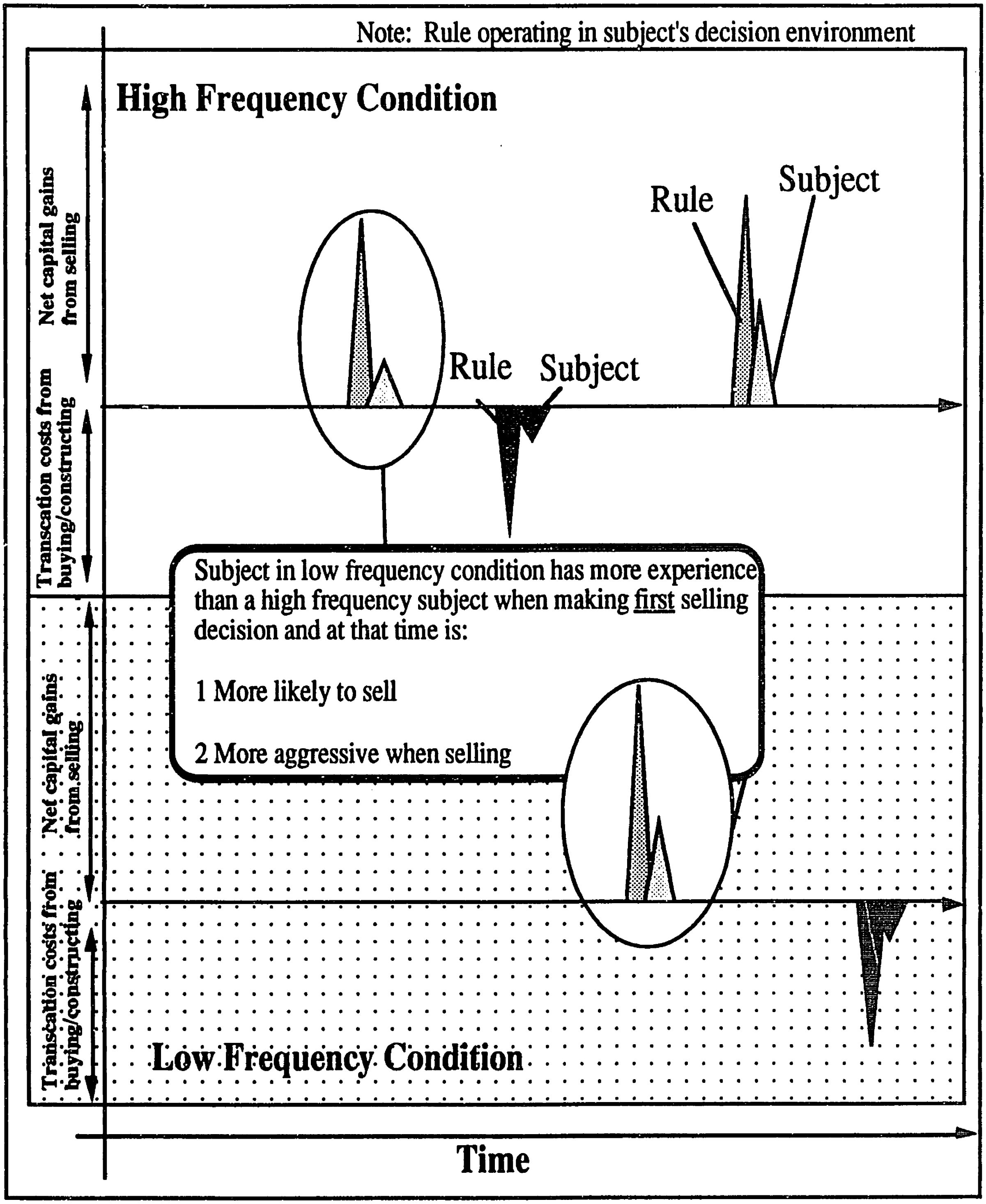


Figure 5.4.1: Subject and rule in high and low frequency conditions.

This figure explains that it will be difficult to infer whether differences in frequency conditions are due to cognitive differences between the two conditions, or whether the differences just reflect that the high frequency subject must achieve faster learning or transfer better than the low frequency subject.

Note, however, the this methodological concern will not influence transfer effects. Similarly, if the first trial had been a good learning environment, then the subjects should have internalized the good decision making and hence no learning to take advantage of transaction opportunities would be required. But we have already seen that the difference between high and low frequencies are as pronounced in the first as in the second trial.

Figure 5.4.1 above also helps explain an apparent anomaly, shown below in figure 5.4.2. Figure 5.4.2 indicates that subjects do better than the rule when the rule is replaced every year, while the rule does better than the subject for longer replacement intervals. This reversal happens because the rule in the short interval is penalized for its far higher β_{buy} and β_{new} than the subjects'. The rule incurs large transaction costs, but does not receive the resulting benefits since the newly purchased assets are "stolen" from the rule. Subjects, being more cautious, do not pay these transaction costs.

When the replacement interval is long, however, then the assets are not taken away before the rule itself can sell them at a profit. The fact that in the very short run, there are substantial transaction costs, so the cautious investor actually outperforms the more aggressive is not only a methodological artifact: It shows that the pay-back of decisions must be judged with a sufficiently long time frame and furthermore indicates that subjects fail to do so..

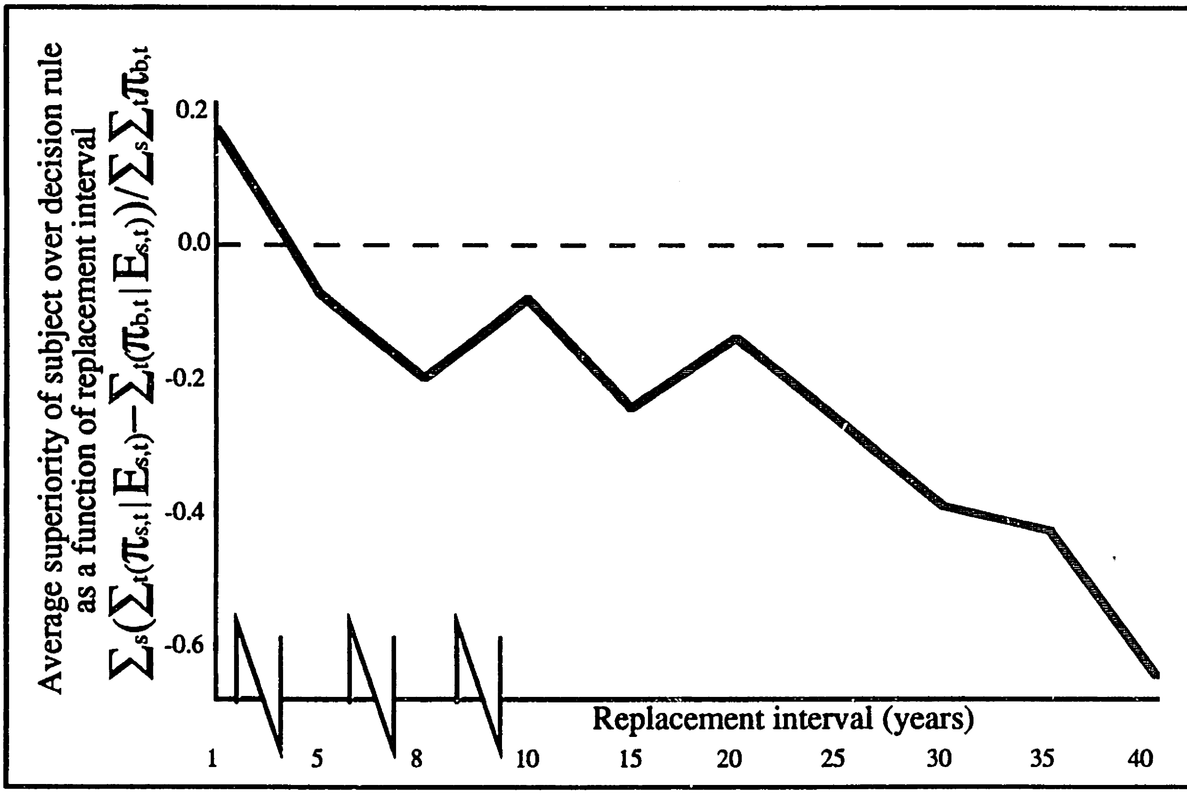


Figure 5.4.2: Superiority of decision rule in a trial as a function of replacement interval.

5.4.2 Performance and transfer

The positive effects of familiar contexts in both trials were expected, while the frequency effects were opposite to expectations. The lack of transfer is indicated by a multitude of observations. Foremost, the same errors induced by treatment conditions in the first condition reappeared in the second. Subjects cannot have transferred much deep insight into the second task.

Moreover, it appears that the treatment conditions in the first trial, as well as changes in conditions between trials play no role (with one possible exception) in the second trial. This again implies that the cognitive impact between conditions have little lasting value. The one exception was that an initial demanding context appears to help transfer performance, although with marginal statistical significance.

5.4.3 Characteristics of high performance settings

There are similarities in high performance settings, regardless of whether the cause is due to context familiarity or low frequency. As shown in figure 5.4.3 and 5.4.4 below, high performance and high variance are correlated.

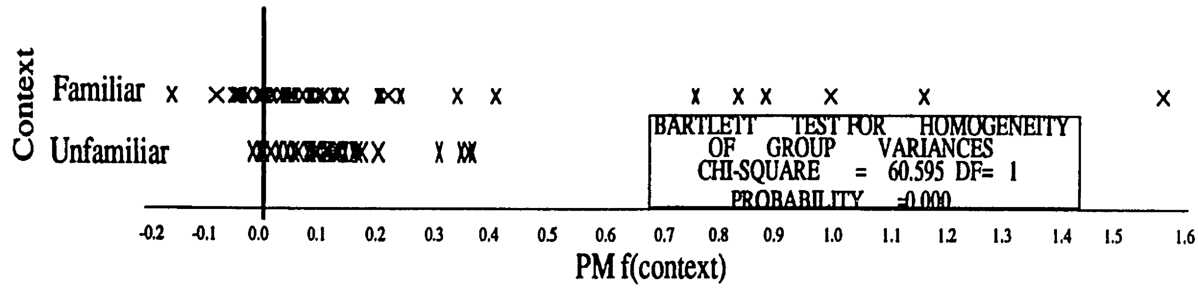


Figure 5.4.3: High variance in familiar context (trial 1 and trail 2 are pooled)

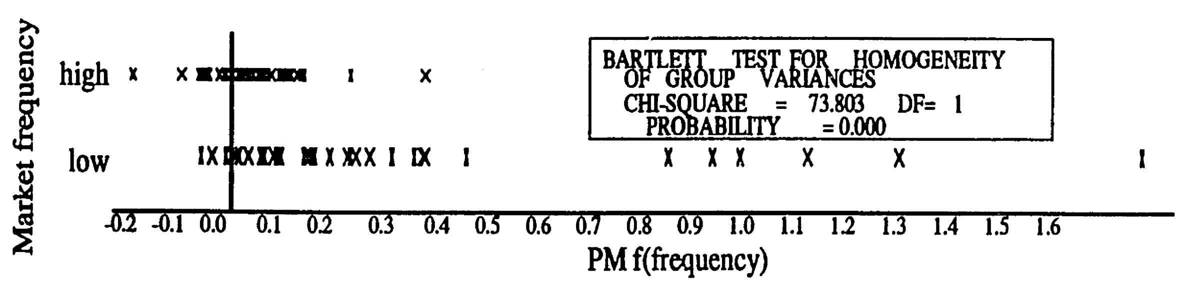


Figure 5.4.4: High variance in low frequency setting (trial 1 and trail 2 are pooled).

Familiar and low frequency environments induce risk taking as evidenced by the fact that the low frequency/familiar context condition consistently yielded higher β_{new}, β_{buy} and β_{sell} than the other conditions. This is shown further below in tables 5.4.8 and 5.4.9.

In section 5.2 it was shown that high performance is due to aggressiveness of buying and of selling. Buying aggressively has the consequence of leading to potentially high profits in operating the assets and/or selling the assets later. However, the chances of bankruptcies will also increase with risk taking, especially as far as investments are concerned.

The fact that low frequency and real estate environments have higher variance should therefore be paralleled with high performance as well as with more bankruptcies. We have already seen that variance and performance correlates. Table 5.4.5 below shows a support for the contention that high performance environments lead to more risky decision making. Subjects who did well in the first trial have more bankruptcies in the second trial (variable name BKRUPT2). Similarly, and as expected, they have more bankruptcies if the current decision environment is familiar.

Subjects who initially face a familiar environment go bankrupt. They also appear to learn from their failure: They are less likely to go bankrupt in the second trial. The latter argument is dependent on grouping bankruptcies in trial 1 into three categories: 0 (when BKRUPT1=1) , 1 or 2 (when BKRUPT1=2), 3 or more (when BKRUPT1=3). This grouping was done since there was an inverted u-shape relationship between initial and later bankruptcies. Subjects who initially go broke once or twice become more cautious, whereas subjects who initially go bankrupt more often are not becoming more cautious. This may be explained by the fact that really poorly performing subjects lose self-confidence and give up, whereas a couple of bankruptcies only sharpen the senses.

Table 5.4.5 below shows the complex relationships between trial 2 bankruptcies, early performance, and context familiarity in both trials.

ANALYSIS OF VARIANCE

1 CASES DELETED DUE TO MISSING DATA.

DEP VAR: BKRUPT2 N: 40 MULTIPLE R: .741
SQUARED MULTIPLE R: .549

SOURCE	DF	MEAN-SQUARE	F-RATIO	P
CONT1	1	1.527	5.742	0.023
FREQ1	1	0.168	0.631	0.433
PM1	1	1.975	7.426	0.011
PM2	1	2.413	9.074	0.005
BKRUPT1	2	1.078	4.052	0.028
CONT2	1	3.544	13.324	0.001
FREQ2	1	0.003	0.010	0.922
CONT1* FREQ1	1	0.659	2.477	0.126
CONT2* FREQ2	1	1.066	4.006	0.055
ERROR	29	0.266		

ESTIMATES OF EFFECTS $B = (X'X)^{-1} X'Y)$

		BKRUPT2
CONSTANT		1.493
CONT1	FAM	-0.222
FREQ1	HIGH	-0.085
PM1		1.287
PM2		-1.131
BKRUPT1	NONE	-0.001
BKRUPT1	1 OR 2	-0.388
CONT2	FAM	0.355
FREQ2	HIGH	0.010
CONT1 FREQ1	FAM HIGH	0.139
CONT2 FREQ2	FAM HIGH	-0.189

Table 5.4.5. Bankruptcies in the second trial are partly explained by high initial performance and by a familiar context.

The fact that high initial trial performance lead to more bankruptcies in the second trial also underlines that high performance gives rise to overconfidence. Since risk taking in general is good, one would also think that late performance can be predicted by early bankruptcies. No effect was found of early bankruptcies on late performance in the present data set (but see Bakken, et al. 1992, for such an effect).

In a prior investigation with a simpler buy-low, sell-high rule (without the supply-line condition), high β's indeed led to performance that was erratic, giving spectacular profits in some frequency scenarios and serious bankruptcies in others. The fact that there is a risk-return relationship between subject bankruptcies and performance indeed supports the contention that subject's decisions are of the simple buy-low, sell-high type that is less stable than the benchmark rule.

It appears that familiar, non threatening, environments induce more risk taking. Similarly, and independently, other stress reducing factors, such as good results in the first trial, also induce more risk taking, and bankruptcies in the second trial.

The fact that subject decisions in lenient environments are more aggressive is also shown in table 5.4.6 for the first, and table 5.4.7 for the second trail. First, it appears that subjects have a hard time incorporating the system's lagged structure into their decision making. Consequently, new orders are negatively correlated to performance. The fact that subject's tend to decrease their new orders in the second trial also help explain the trickiness of the delayed ordering.

ANALYSIS OF VARIANCE

DEP VAR: PM1 N: 40 MULTIPLE R: .725
SQUARED MULTIPLE R: .525

SOURCE	DF	MEAN-SQUARE	F-RATIO	P
CONT1	1	0.241	6.631	0.015
FREQ1	1	0.391	10.783	0.002
βORDER1	1	0.160	4.414	0.044
βSELL1	1	0.331	9.127	0.005
βBUY1	1	0.090	2.487	0.125
CONT1* FREQ1	1	0.186	5.121	0.031
SELL1* BUY1	1	0.098	2.687	0.111
ERROR	32	0.036		

ESTIMATES OF EFFECTS $B = (X'X)^{-1} X'Y$

		PM1
CONSTANT		-0.220
CONT1	FAM	0.088
FREQ1	HIGH	-0.103
βORDER1		-0.140
β SELL1		0.086
β BUY1		0.156
CONT1 FREQ1	FAM HIGH	-0.076
β SELL1 β BUY1		-0.021

Table 5.4.6: Performance as explained by treatment conditions and subjects aggressiveness in first trial.

ANALYSIS OF VARIANCE

DEP VAR: PM2 N: 41 MULTIPLE R: .776
SQUARED MULTIPLE R: .602

SOURCE	DF	MEAN-SQUARE	F-RATIO	P
CONT2	1	0.216	4.730	0.037
FREQ2	1	0.818	17.912	0.000
βORDER2	1	0.000	0.002	0.968
βSELL2	1	0.729	15.974	0.000
βBUY2	1	0.097	2.121	0.155
CONT2* FREQ2	1	0.259	5.666	0.023
SELL2* BUY2	1	0.665	14.562	0.001
ERROR	33	0.046		

ESTIMATES OF EFFECTS $B = (X'X)^{-1} X'Y)$

		PM2
CONSTANT		-0.047
CONT2	FAM	0.079
FREQ2	HIGH	-0.158
βORDER2		-0.007
βSELL2		0.197
βBUY2		0.265
CONT2 FREQ2	FAM HIHG	-0.090
βSELL2 βBUY2		-0.145

Table 5.4.7: Performance as explained by treatment conditions and subjects aggressiveness in first trial

The two tables above indicate that there is a relationship between subject aggressiveness and treatment condition. Measuring aggression as β, i.e. the degree to which subject's trade as a fraction of the potential trade, the tables below indicate that aggression is higher in the low frequency, familiar context cell.

First trial	β Mean	St dev.	β in low frequency/high familiarity condition...	
New	2.6 %	3.2	Higher	$p < 0.06$ (using/sqroot of β)
Buy	6.3 %	9.8	Higher	ns
Sell	32.1 %	22.9	Higher	ns

Table 5.4.8: Subject's aggressiveness in trial 1.

Second trial	β Mean	St dev.	β in low frequency/high familiarity condition...	
New	1.6 %	2.3	Higher	$p < 0.05$ (using/sqroot of β)
Buy	7.7 %	13.7	Higher	ns
Sell	47.8 %	28.4	Higher	ns

Table 5.4.9: Subject's aggressiveness in trial 2.

The next two sections will further explore the relationship between treatment conditions, cognition, transfer and aggressiveness.

5.4.4 Context

Context familiarity has a positive influence on performance. The cognitive relief provided by a familiar context ensures higher performance. Yet this learning does not endure or transfer. On the contrary, subjects do marginally better in the second trial if the initial trial contained an unfamiliar context.

If familiarity increases performance, then one must ask if this relationship is monotonic or curvilinear: Too high a context familiarity might be a burden. A related pilot study of professional and student performance reported in appendix 1, (see also Bakken et al., 1992)[14], indicated that students learn more easily than professionals and together with the present study propose an inverted u-shape relationship between context familiarity and performance. This form of a relationship is also found in other areas of performance, e.g. the relationship between stress and performance (Hockey and Hamilton, 1983) and is shown in figure 5.4.9 below.

[14] See also appendix 1

The pilot study used mostly MBA trained professionals with 10-20 year's experience, mostly in the oil tanker and real estate industries depicted. Hence they were more familiar with the context than students who were MIT Sloan School graduate management students. Comparing 16 professional 2-person teams to 17 single students, it was found that students initially went bankrupt more often. Yet trial 2 performance was higher for students than for professionals.

In cognitive terms, one can explain that very high context familiarity increases the probability that a subject's problem space contains domain information: Subjects have pointers to experiences outside the laboratory. Such pointers may help solving the problem at hand (as in the main experiment). Pointers are, however, only valuable if the representation they point to is valid for the task at hand.

Because pointers tend to use concrete (i.e. surface) features (Northcraft and Neale, 1986) they may point to irrelevant, or incongruent context features. High familiarity will only help performance if positive transfer is likely to be triggered, i.e. when the accessed information is *representative of the problem* (Kahnemann and Tversky, 1982).

As pointed out in chapter 4, the oil tanker and real estate industries are markets where work experience has problematic diagnostic value: Meaningful and timely outcome feedback is simply not available. Since the simulated time horizon in the experimental task was much longer than time horizons typically emphasized by professionals (Bakken, 1990; Fuglseth, 1989), it is likely that contextual features trigger irrelevant information. Both the limited set of variables in the game, and the total absence in the simulated task of issues that are prominent in professionals' concepts, may have contributed to this. Concepts such as geography, marketing and product quality are indispensable for successful operations in both real estate and shipping markets, yet had no bearing on the simulated task.

As mentioned, professionals have schemas available that focus on issues that are irrelevant to the simulated task at hand. In addition, they may also have a mental model which is at odds with the lessons to be gained from the experiment. Whatever mental models of the long term dynamics exist among professionals, they are embedded in dense cognitive networks (Bottom, et al., 1989; Neale and Northcraft, 1988; Novick, 1988). In the case that subjects prior schemas are incongruent with the laboratory environment, subjects must unlearn inferences. The resolution of inconsistencies in mental models is exactly what makes (un)learning so hard among successful professionals (Argyris, 1991). One reason for this is that successful professionals seldom get corrective feedback and so have no well developed meta-schema for handling erroneous inferences: People pay attention to conceptual issues they have seen before and know how to act upon (March and Olsen, 1974; De Geus, 1988). Another reason for problematic learning is that even when people know about their own cognitive dissonances, these may still be hard to resolve (Akerlof and Dickens, 1982; Jacoby, et al., 1987).

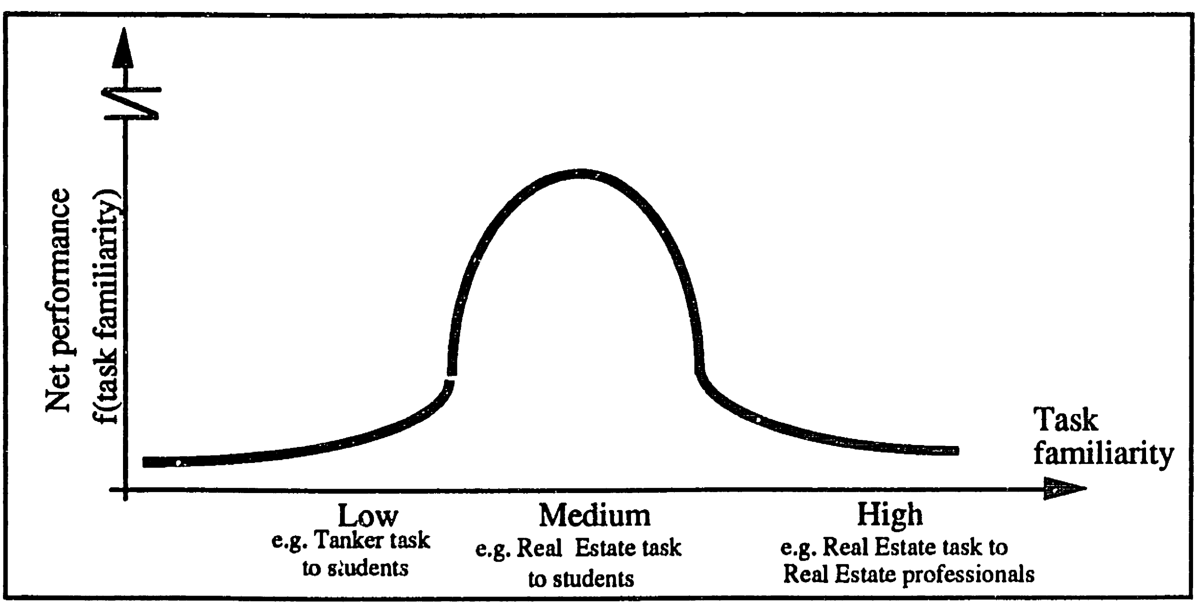


Figure 5.4.9: Performance as a function of context familiarity.

In static tasks, cognitive density is called expertise and its quality is seldom questioned (Holyoak, 1991). However, in environments with delayed or uncertain cause-and-effect relationships, subjects may well perpetuate incorrect causal models (Hogarth, 1981). The conditions for on-the-job validation are problematic when experience in one aspect of a market, i.e. the short term, are invalid for task dynamics of a different, more elongated, nature (Fuglseth, 1989).

The fact that experience and task familiarity can hinder learning has been recognized by many (see e.g. Argyris, 1978). As a consequence, learning among people with deeply embedded understanding can be improved when moving "to the left" on the familiarity axis on figure 5.4.9. As a cognitive opener to get unstuck from established routines, unfamiliarity may be a boon. Virtual worlds, such as simulated markets, may be helpful in creating learning (Schön, 1983). Indeed Papert (1981) suggests that computer simulations may facilitate learning by creating transitional objects (though computer languages as metaphors for transferable thinking is problematic, see e.g. Kirkland, 1984). Transitional objects are characterized by being easier to mentally manipulate than real decision objects. A simulated tanker market may be a transitional object to real estate professionals as it has these desired properties. Indeed this was done in pilot study 2 (see appendix 1).

Though the impact of context familiarity on performance in the current task was positive both in the first and the second trial, an interesting impact from the initial context on the second trial performance occurred. The finding that an initial unfamiliar context helps performance suggests that a hard task may induce between-trial reflection and search for underlying causal relationships. The finding also suggests that questionnaire scores should be higher after a hard initial task, but as will be shown in the next chapter, no such relationship was found.

Though an unfamiliar environment may increase transfer, there are obvious limits to such a finding: If the learning environment is incomprehensible, then, of course, no learning and hence no transfer of understanding can take place. The finding of higher transfer from an unfamiliar environment must be interpreted so that if cognitive complexity is induced by the task, the little learning that takes place has more transfer value.

Similarly, the results in this subsection also indicate that one must be careful in analyzing current game performance as an indication of deep learning: In the present study, those contextually familiar environments that produced high performance also appeared to reduce transfer of performance (see also Teach, 1990).

On the other hand, the experiments indicated that poor context familiarity may reduce the value of a learning environment since they become cognitively burdened with having to understand new concept with little surplus mental resources to deal with complex dynamics.

The next subsection will investigate the role of frequency on performance and learning.

5.4.5 Frequency

The finding that a low frequency market leads to better performance was counter to the hypothesis described in chapter 3. One aspect of better performance was the learning speed issue outlined in figure 5.4.1. There are, however, competing explanations: The high frequency environment was characterized by shorter asset life times (17 vs. 34 years) and thus shorter loan repayment schedules. The high frequency market required a 6.7% repayment of all loans every year. This burden was only 3.3% in the low frequency environment. This may have added stress to the compressed environment, as already shown by less performance variance in the high frequency environment and lower β in the unfamiliar/high frequency conditions . In theory, the benchmark rule, which also faced this higher financial burden, should have been equally fair to both frequency conditions. But contrary to the rule, subjects are affected by stressful situations where decision repertoire is limited and complex thinking similarly hurt (Gladstein and Reilly, 1985).

This finding that lower frequency yields stronger performance goes counter to prior experimental work about the link between stability properties of markets, and human performance. Such studies have defined structural complexity different from the experiments here, and they have focused on one or more of the following:

i) Instability
ii) Delay between action and consequence
iii) Delay between consequence and feedback about that consequence
iv) Self-reinforcing processes.

Diehl (1992) found that as delays increased, performance suffered. Self-reinforcing side-effects also reduced performance. The same weakening of the learning effect by longer delays was found by Brehmer (1988). Sterman and Paich (1992) similarly, in a product life cycle experiment, found that performance relative to benchmark worsened as the strength of word-of-mouth and repurchasing interval increased. Kampmann (1992) used an experimental set-up where 4-5 decision makers, each running their own firm, participated in a market. Just as expected, markets with short lags between desired and actual production, and without self-reinforcing side-effects (from production on subsequent demand) converged towards price and production equilibrium faster than if the structure was less transparent.

It must be noted, however, that the above tasks were all control tasks. In such tasks, strong supply line and weaker stock control parameters tend to do well. As shown in section 5.2, the present task performance is improved with aggressive decisions. In addition, the experimental markets were also characterized by immediate and high leverage secondary controls, through the buying and selling of used assets. This again made the control problem trivial as long as the player controlled little of the market. The chief impact of delay in the present task has been different than in most control studies.

Thus, the present findings underline that structural features in addition to influencing control may influence system frequency and thereby lead to environments that are more or less conducive to learning.

The present tasks were thus more of "understanding" and "risk strategy" tasks. Poor performance may be attributed to subjects' reluctance to sell assets. Similarly, the rule invests heavily during depressed times. Subjects, however, appear to anchor on prior investments and fail to be aggressive enough.

In the present study, there are three transparency effects with respect to frequency that go counter each other. The first, treated in the psychology literature (See e.g. Skinner, 1974, Thirney et al, 1987) states that with longer action to outcome delay, subjects lose ability to interpret and use feedback information.

The second effect, related to the previous one, is that time constants influence system's oscillatory tendencies at the same time as these constants influence the immediacy of feedback and degree of control. Diehl (1992) and Kampann (1992) have focused on subject control as a function of these time constants. The oscillatory tendencies also influence performance as was shown by low frequency markets giving higher performance than high frequency markets.

The third issue, unique to the present tasks, is the link from delay parameters to how much cash-flow a decision maker must earn to break even. A higher debt burden may influence stress among subjects and lead to conservative decision strategies and consequently poorer performance in short delay, i.e. high frequency, conditions. High frequency settings are more exacting (Larrick et al, 1989); subjects are closer to bankruptcies at least initially, something that is more stressful and so reduces the available decision repertoire.

Low frequency leads to higher current-trial performance. As noted this relationship held true for both trial 1 and trial 2. Since the research design had only two data points, it cannot by itself describe anything but a linear relationship between frequency and performance. A shown in chapter 4, however, the slow motion argument has its limits since at the extreme, a low enough frequency will appear to subjects as a market in a either a steady growth or a decline phase. One must therefore expect an inverted u-shape of performance in various frequency environments: Subjects do poorly in high and extremely low frequency environments, yet in moderately low frequency environments they do well as shown by the present experiment:

Indeed, several of the real estate investors interviewed in chapter 4 showed evidence of assumptions of stable and infinite growth that may explain their poor performance when it appeared that the market indeed was unstable.

The interview with one tanker investor, however showed that even in situations where commonly used market data only show growth, there exist cyclical tendencies that can be inferred from structural relationships. This investor's inferences appeared not to be widely shared; people focus on concrete data (Northcraft and Neale, 1988) and inferences about underlying structure appear not to be made (see also Paich and Sterman, 1992).

5.4.6 Discussion summary

The findings that low frequency and familiar environments are helpful, must be interpreted with caution. Though conditions inducing good performance in the first trial are transferred via the fact that performance in the two trials are positively correlated, subjects are likely to transfer particularly well from those environments where they do poorly, as from an unfamiliar context. Likewise the fact that high initial performance cause many bankruptcies in the second trial also underlines that early performance and transfer may go in opposite directions, something that was also shown in Bakken et al., (1992). In the terms of (Senge, 1990) the "video-game syndrome" can happen in simulated decision environments where subjects who focus on current performance tend not to learn much.

The finding that hard decision environments are related to high transfer was not supported by the main analysis in more than one out of six possible indicators: There were no significant interaction effects to indicate that changes in context or frequency environments between trials improved performance.

As noted in appendix 1, insights may be improved by a different format than the pure gaming environment used here. Kim (1989) suggests that reflective exercises may help subjects address structural issues. Discussions can moreover force hidden assumptions out into the open (Bakken et al., 1992). Without such exercises, transfer is poor. Chapter 7 will further speculate on their potential value.

5.5 Summary

The table below summarizes the major effects of the treatment conditions. Since the first trial results are essentially replicated by the second trial, only second trial results are shown.

Significance	Effect	Level	Coefficient
Significant	REGRESSION CONSTANT		0.12
	Treatment effects		
	Second Market Context	FAM	0.11
	Second Market Frequency	HIGH	-0.10
	Interaction effects within trial		
	Context * Frequency	FAM*HIGH	-0.10
Marginally Significant	Effect of first trial performance		0.56
	Transfer effects		
	Carry-over from first trial		
Not Significant	First Market Context	FAM	-0.08
	First Market Frequency	HIGH	
	Interaction between trials		
	No significant interactions		

Context	Frequency: High	Low	Mean
Familiar	-0.09	0.31	0.11
Unfamiliar	-0.11	-0.10	-0.11
Mean	-0.10	0.10	0.00

Table 5.5.1: Summary of effects in trial 2

Note that performance is better with lower system frequency, a finding that goes counter to the expectation. The reasons for this were explained partly as the increased transparency of a "slow motion" environment, and partly as the result of a psychologically more lenient depreciation and amortization scheme that makes the decision environment less stressful and so induces more aggressive decision making.

Context familiarity improves performance. This was explained by the reduced strain on cognitive resources in a familiar environment. However, by drawing on a pilot study to the main analysis provided in this thesis, it was also argued that context familiarity does not increase monotonically. Context interference, it was argued, makes learning harder for subjects who have had substantial exposure to a certain decision environment.

In absence of any guided process to induce reflection, transfer of insights is helped marginally by an initial context that lacks familiarity. The present data indicate that no other treatment conditions help or hinder transfer. Thus, little can be said about whether game designers should use several different contexts to further learning and whether compressed or less compressed environments enhance transfer from workshop to workplace.

Subjects who encounter a tanker (unfamiliar) environment do better in trial two; not because they remember the tanker task (they do no better in the second task if this also is a tanker task), but because an unfamiliar environment may induce reflection between trials. Chapter 6 will look more closely into supplementary questionnaire and other data that may support the reflection argument.

6. Other factors affecting performance and transfer

6.1 Introduction and overview

As noted in the previous chapters, theories about causes of decision making performance should be grounded in information about subjects' underlying cognition. Hitherto cognitive mechanisms have been inferred, rather than measured. This chapter will specifically investigate process data.

The need to augment performance measures with data that highlight cognitive processes has been underlined by many, see e.g. (Einhorn et al., 1979; Payne, 1976; Huber, 1986). However, cognitive processes are not easily measured (Ericsson and Simon, 1980; Vennix, 1990). Though concurrent verbalizations may mirror underlying thinking processes (Ericsson and Simon, 1980), some have suggested that more easily collected retrospective verbalizations may be disjoined from actions (Broadbent, 1986). This is especially true for tasks that are highly automated: Behavior in such tasks is commonly called overlearned and does not require cognitive activity. Without any cognitive activity to report, subjects will have to invent reasons for their decisions and verbal accounts will differ from actions.

In fact, the relationship between questionnaire scores and decision performance is sometimes negative (Berry, 1984), thus suggesting that people may have two separate mental databases, one for verbalization, another one for decision making. Moreover, if the two databases compete for limited cognitive resources, one must expect that high performance on one dimension leads to lower performance on the other. Precisely because the relationship between action and inference processes are unclear, decision making research should compare action, outcome and cognitive processing data (Vennix, 1990; Kleinmuntz and Thomas, 1987; Kleinmuntz, 1990).

There also exist other ways of testing cognitive theories underlying decision making processes. As noted by Einhorn et al. (1979), decisions can be simulated by assuming various information processing strategies: Decision rules can be inferred from decision data. Estimates of subjects' decision rules enable testing of hypotheses of specific cognitive limitations and errors. Sterman (see especially Paich and Sterman, 1992) and Hammond (see e.g. Hammond and Summers, 1972) have used assumptions about how subjects integrate information cues as well as decision data to numerically estimate and calibrate decision rules.

There are several limitations to rule estimation. Though econometric and other estimation methods can establish that misperceptions occur, the question of why is not easily addressed with such models (Kleinmuntz, 1990). Since decision rules often produce flat error surfaces, interpretations about decision behavior can be problematic (Dawes, 1979; Bakken, 1989a) and the robustness with regards to alternative parameters may be poor. For these reasons, in addition to the estimation of subject aggressiveness, β, done in the previous chapter; a more exploratory use of the process data is applied in this chapter .

For explanatory reasons, a conceptual model of the various sources of research inference is provided in figure 6.1.1. It shows the relationships between subject background, understanding, treatment conditions, and experimental experience and performance as well as the measurements used to monitor these relationships. Chapter 5 has already been devoted to explaining how treatment conditions affect performance. Some arguments for performance differences have likewise been discussed. In this chapter, hypotheses and explanations of process differences in the different treatment conditions will be further refined.

The depicted model suggests that subjects go into the experiment with academic and work experience. These experiences enable subjects to understand the task at hand.

Understanding must again be related to performance. Figure 6.1.1 highlights the conceptual model and, in bold, the corresponding measurements.

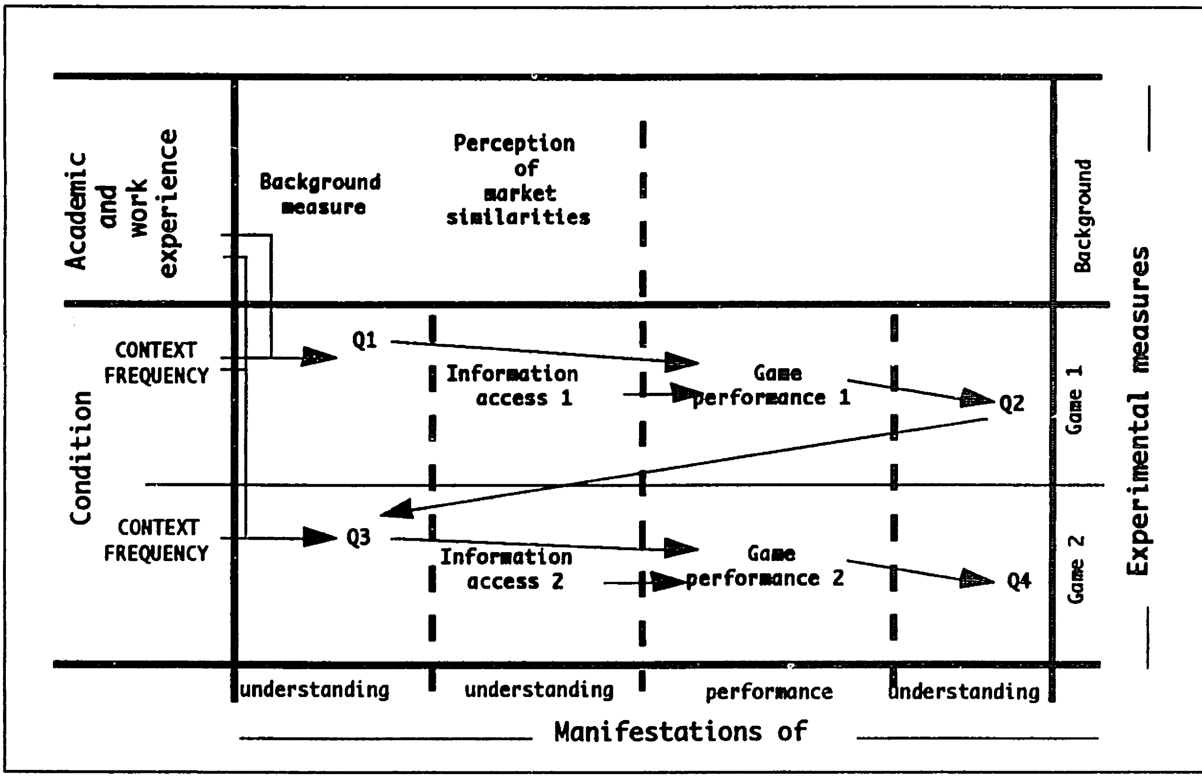


Figure 6.1.1: The underlying model of interrelationships (metrics shown in bold typeface)

Subjects rated their own academic and work experience. A related questionnaire about how subjects judged similarities between different markets likewise indicated subjects' perceptions. The understanding of the tasks' causal structure was monitored by questionnaires used before and after each game play. Subjects' use of information reflects their understanding of the task. However, at the same time good use of information also leads to increased insight into the task.

This underlines the circular causality of performance and learning; good understanding increases the likelihood of good information use that further increases understanding. Another example of the circular causality was provided in chapter 5: Poor performing subjects tend to limit their decision repertoire thus further reducing the opportunities for improvements in performance.

This chapter is divided into five sections. The next section shows how subject background, i.e. academic and work experience, influences performance. The following section shows how subjects view similarities between markets; their information acquisition process is also described. Information access is measured by how many times subjects use look-up tables and graphs.

Questionnaires where subjects indicate direction and degree of delay between causal relationships were used before and after game playing. They monitored how subjects' mental models influenced performance, and how mental models were influenced by exposure to the decision environment as well as by the treatment conditions.

Each section starts with a recollection of why and how the data were collected. Then follows a summary of findings. Each section concludes with a discussion of the implications of the findings for performance and transfer processes.

6.2 Academic training and work experience

Training in general domains of scientific inquiry should help subjects apply inference and decision strategies to judgment and decision tasks (Fong, 1986; Fong and Nisbett, 1991; Nisbett, et al., 1987; Davis and Hogarth, 1992). However, when taught at universities, especially at professional schools such as management schools, scientific approaches are often integrated in more applied domains. When teaching system dynamics, for example, the mechanisms of "worse-before-better behavior", positive and negative feedback loops, and the attractiveness principle are all, for pedagogical reasons, taught by examples. As shown in chapter 5, the use of somewhat familiar concepts indeed help performance in a dynamically complex task. However, it may well be that students take away domain specific structural understanding, or worse, that they bring with them decision heuristics with low robustness.

The use of a sequence of two tasks should facilitate a distinction between effects that help initial framing of the first task and effects that help subjects understand the underlying task structure and help them transfer understanding to the second task.

In the academic background questionnaires, subjects rated their experience in various business school and engineering disciplines. Some of these, like accounting, business strategy and computer science, provide frameworks for understanding the current task. Computer and business language literacy should improve decision making for several reasons. Familiarity with the computer environment should free cognitive resources that may be used for decision making, which may also alleviate stress and thereby induce more risk taking. Familiarity with accounting should help subjects understand the financial variables and depreciation schemes. Prior exposure to business strategy should similarly help students be attentive to a fluctuating business environment.

Economics (and finance) and system dynamics[15] (including control engineering) should help subjects' performance. These fields differ from the above in that they are more general. Thus, these areas of academic knowledge should help subjects both frame the current task, as well as help them transfer understanding to a similar task. Since the teaching of principles was done at a management school, this should further help subjects draw inferences in other business tasks (Nisbett, et al., 1987). As noted by Davis and Hogarth (1992), people need both scientific (what they call conceptual) as well as application (what they call domain) skills in order to act successfully and be able to learn from experience.

---

[15] It appears that some subjects did not make a distinction between system dynamics and control engineering.

In the previous chapter it was noted that professional skills can be a double-edged sword in experimental tasks. Just like professional experience can cause expertise blindness, academic training may also hinder performance. To that effect, White (1991) suggests that control engineering, which is viewed as a unifying framework at several of MIT's engineering departments, in fact may limit thinking processes. White describes an engineering competition between MIT and Berkeley students where the MIT team lost because they failed to identify a hidden energy source in a "perpetual motion" machine. They used a control engineering framework, while an electrical analogy was called for. When the depressed team reflected back on the causes of their defeat (p. 190), one of the students said:

> ...too bad Pepper and I took controls [control engineering] so recently; if we hadn't maybe we wouldn't have tried to find the controller in the machine... if only we knew more about electricity and electrical testing...

The questionnaires should enable to investigate whether background would help or hinder performance. Figure 6.2.1 below shows the self-rating system.

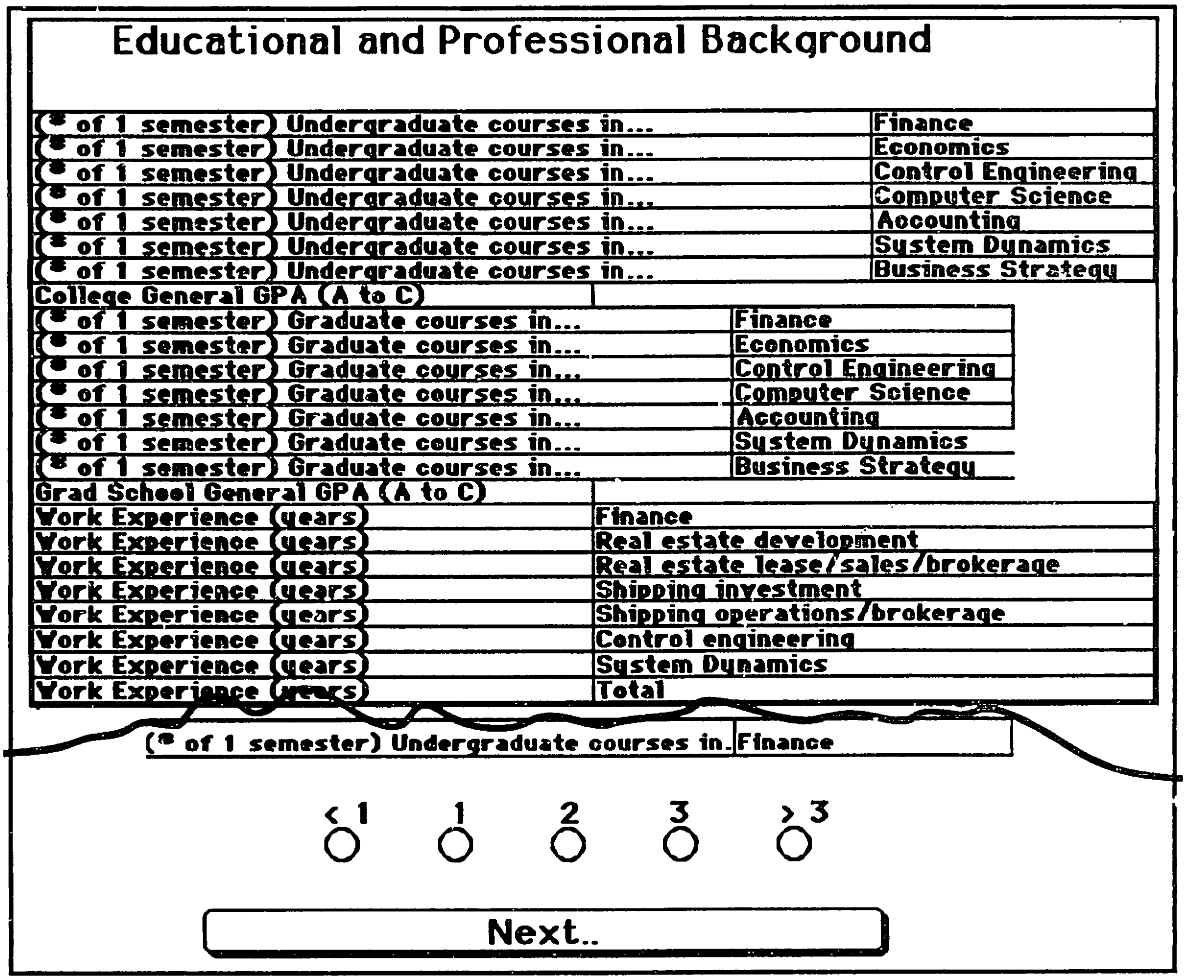


Figure 6.2.1: The questions used to measure subject background. Only one line was shown at the time as depicted in the bottom part of the figure.

The results were first computed using simple correlations between the various subject background variables and PM1 and PM2. There were significant and positive effects from background in undergraduate finance, computer science, system dynamics and business strategy classes on initial performance. Graduate system dynamics classes also influenced initial performance positively. It appears that economics, system dynamics, business strategy and computer science may contribute to performance.

The number of courses in graduate system dynamics and control engineering classes were added arithmetically and called "System Dynamics Factor". Likewise, all classes a subject had taken in economics, finance and accounting were added to one single measure, "Economics Factor". The "Computer Science Factor" added graduate to undergraduate classes. Likewise, "Business Strategy Factor" added graduate and undergraduate coursework. All four scales, called factors, explained performance in the first, but not in the second trial as shown below in table 6.2.2.

		Undergraduate Courses								Graduate Courses											
		Finance	Economics	Control Engineering	Computer Science	Accounting	System Dynamics	Business Strategy	Grade	Finance	Economics	Control Engineering	Computer Science	Accounting	System Dynamics	Business Strategy	Grade	System Dynamics Factor	Economics Factor	Computer Science Factor	Business Strategy Factor
	PM1	0.33			0.44		0.33	0.43							0.36			0.38	0.28	0.43	0.28
	PM2																				
number of classes taken	Average	1.3	2.5	1.1	2.4	1.7	0.8	0.8	3.0	1.0	1.7	0.8	1.3	1.1	1.2	1.4	3.5	4.1	5.5	3.8	2.3
	Stdev	1.1	1.2	1.0	1.4	1.2	0.9	0.6	1.2	0.6	1.0	1.0	1.2	0.5	1.1	0.8	1.1	2.2	1.7	1.5	1.0

Table 6.2.2: Correlation coefficients between performance indices and academic background. All p values < 0.1

Regressing each academic subject on PM is problematic, since the academic classes taken are not independent. For instance, subjects who take finance have also taken economics classes: Simple regressions overstate the number of relationships. A multiple regression equation using all four academic factor background variables was run. As shown in table 6.2.3, only System Dynamics of the four helps explain a significant amount of performance variance in the first trial.. No effects were found for the second trial.

DEP VAR: PM1 N: 40 MULTIPLE R: .526 SQUARED MULTIPLE R: .277
ADJUSTED SQUARED MULTIPLE R: .194 STANDARD ERROR OF ESTIMATE: 0.225

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2 TAIL)
CONSTANT	-0.491	0.211	0.000	.	-2.326	0.026
BSFACT	0.023	0.034	0.112	0.740	0.673	0.505
CSFACT	0.028	0.017	0.262	0.766	1.595	0.120
SDFACT	0.020	0.012	0.264	0.869	1.711	0.096
ECONFACT	0.012	0.015	0.124	0.794	0.766	0.449

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
REGRESSION	0.677	4	0.169	3.347	0.020
RESIDUAL	1.769	35	0.051		

Table 6.2.3: Only System Dynamics help explain the variance.

There are two major reasons that help explain why academic background helps initial performance. The first is that background variables enable subjects to understand the unfolding of the dynamics by being a "cognitive" booster, i.e. freeing resources to deal with the task dynamic and its implications. This first explanation is similar to the reasons why context familiarity will help performance.

A competing explanation is that appropriate background give some subjects a better pre-game understanding of the underlying forces. To investigate the competing explanation, the relationship between background variables and the first questionnaire score, Q1 was examined. Though graduate grades are positively correlated with Q1 at the $p < 0.08$ level, there is no general relationship between those background variables and Q1.

One may argue that certain background variables may help in framing the subsequent task, so that background variables should help later questionnaire scores. However, no meaningful relationships were found between background and such scores, even when regressing background variables independently with each of the questionnaire scores.

This section indicates that initial performance is positively influenced by certain undergraduate background factors, such as finance, computer science and system dynamics. Courses in business administration and engineering disciplines may help because they free up cognitive resources so that subjects better understand the unfolding task. The background influence on understanding does, however, not last after the first task is performed. The lack of persistent effects suggests that the task experience itself provides enough familiarity with the business concepts to attenuate the (admittedly narrow range of) background effects investigated and appears to be the main reason why subjects do better in trial 2, though increased aggressiveness of selling also contributes, as shown in chapter 5.

This finding, namely that the first task provides subjects with a experimental familiarity, may help explain why there is a general practice effect, but no change in subjects' ability to deal with the cognitive complexity of an unfamiliar task or a high frequency: The first trial is harder than the second since subjects have to learn to interact with the computer and decipher financial information. These aspects, helped initially by various backgrounds, were identical in the two contexts and general task learning was thus readily transferable. The problematic behavior, however, was not.

Work experience

Work experience in system dynamics had a positive impact on initial performance. Neither later performance nor transfer was influenced by this work experience as shown in table 6.2.1. No other work experience influenced performance or transfer. The lack of effect from work experience on performance is apparently a result of low variance in the experience data. Subjects tended to have less than one years' experience in any of the indicated fields.

	Work experience							
	Finance	Real Estate Development	Real Estate Operations	Shipping Investment	Shipping Operations	Control Engineering	System Dynamics	Total
PM1							.41	
PM2								
Average years	<2	<.5	<.5	<.5	<.5	<.5	<.5	<4
Std deviation	.33	.14	.03	.00	.00	.07	.03	.75

Table 6.2.3: Correlation coefficients between performance indices and work experience (top), p value < 0.05. Lack of variance in work experience (bottom).

6.3 Rating of market similarity

In chapter 5 it was argued that subjects' prior conception of how markets behave may interfere with the experimental task. In order to investigate these interferences a questionnaire was designed to measure to what extent subjects regard various markets as similar. The prediction was that subjects who rate markets that have physical resemblance erroneously as also being dynamically similar, will have a higher probability of interference between concrete and dynamic schemas and do poorly. In cases where such interference is a good thing, for instance if subjects regard commercial real estate and oil tankers as physically similar, transfer performance should increase.

A test was designed to determine whether such interference could explain poor initial and later performance. Subjects rated similarities of different markets along three dimensions of similarity: Physical appearance, Industry structure and Market dynamics. The rating was done before any preparatory material was distributed.

Figure 6.3.1 shows the response options available to subjects (only one rating line was shown at a time).

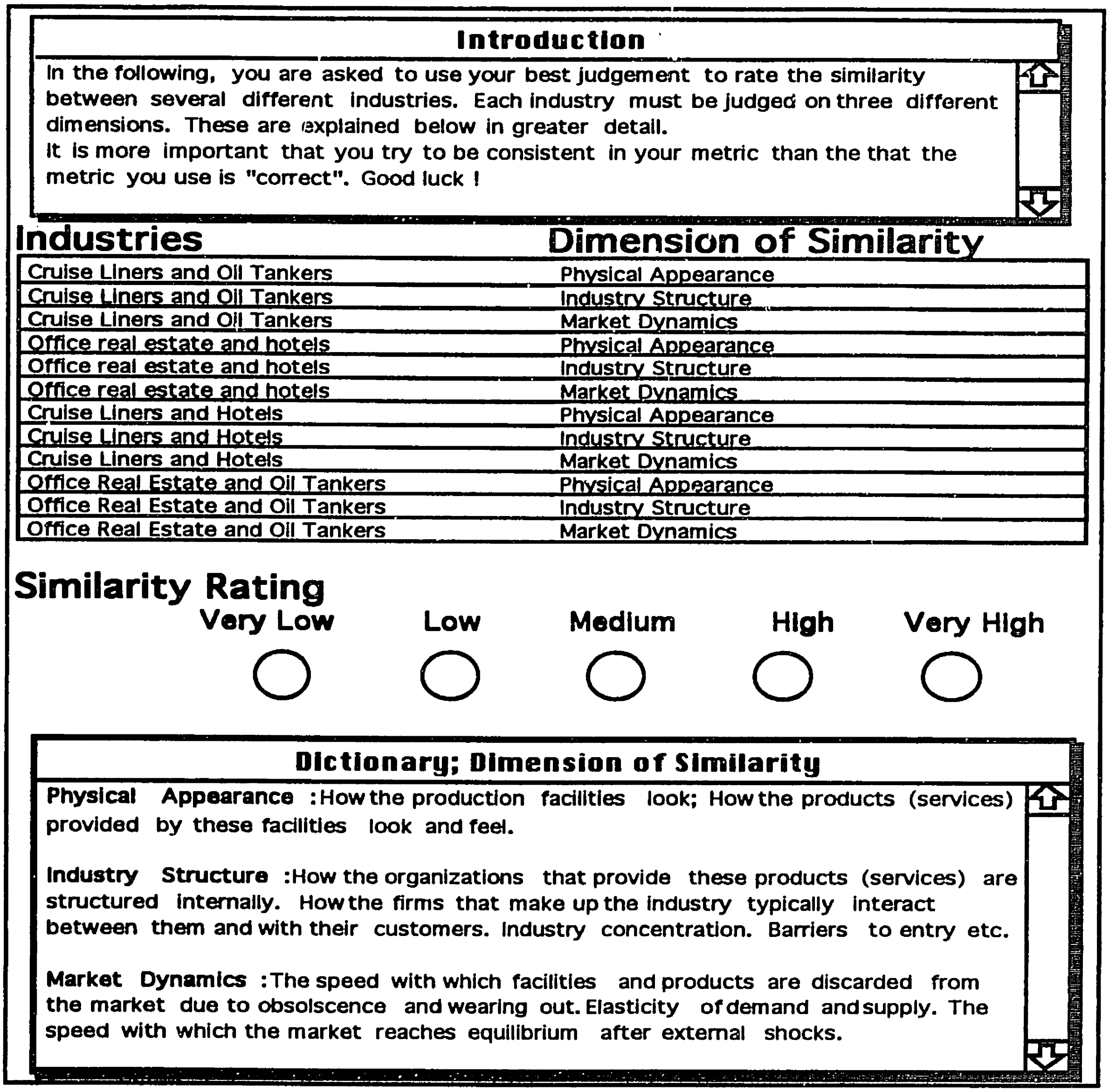


Figure 6.3.1: Similarity rating set-up.

Simple correlations between the 12 ratings and performance revealed no significant effects. Figure 6.3.2 shows rating averages. Note that subjects regard cruise liners and hotels much as more similar than office real estate and oil tankers (paired t-test, $p < .001$), at the same time they regard cruise liners and oil tankers with the same high similarity as office real estate and hotels. This may be explained by the fact that oil tankers and cruise liners have in common being ships; office real estate and hotels are both buildings; hotels and cruise liners both have rooms where people sleep and restaurants where one can eat.

Subjects appear to realize that physical similarity and market dynamics are uncorrelated, however.

Performance index	Cruise Liners and Oil Tankers: Physical Similarity	Cruise Liners and Oil Tankers: Industrial Organization	Cruise Liners and Oil Tankers: Market Dynamics	Cruise Liners and Hotels: Physical Similarity	Cruise Liners and Hotels: Industrial Organization	Cruise Liners and Hotels: Market Dynamics	Office Real Estate and Hotels: Physical Similarity	Office Real Estate and Hotels: Industrial Organization	Office Real Estate and Hotels: Market Dynamics	Office Real Estate and Oil Tankers: Physical Similarity	Office Real Estate and Oil Tankers: Industrial Organization	Office Real Estate and Oil Tankers: Market Dynamics
PM1 PM2	–	–	–	–	–	–	–	–	–	–	–	–
Average	4.00	3.18	3.95	4.35	3.70	4.20	3.73	4.53	4.70	2.15	3.15	3.68
St. dev	1.11	0.90	1.06	1.12	1.16	1.11	1.32	0.91	0.94	0.53	0.99	1.11

Figure 6.3.2: Similarity rating averages and standard deviations (note: 2 is very low; 6 is very high).

It thus appears that the physical similarity rating did not reveal much about the nature of the relationship between how subjects perceive the markets and subsequent performance or transfer.

6.4 Information acquisition

As discussed above, subjects' use of information is self-reinforcing in the sense that use of good sources gives deeper insight into the task, improves performance, and allows more risky decisions which enable even more insight which further increases performance. Conversely, poor use of information sources may perpetuate poor performance. In dynamic tasks, subjects often fall into decision traps but, unlike the animal caught in the trap, decision makers often have no way of assessing that they are doing poorly. In the experiment, subjects did not receive any information about the optimality of potential performance. Of course, subjects who go bankrupt translate that information into an indication that they can do better. In the absence of such information, however, outcome feedback by itself has a problematic diagnostic value, even as a static information source. Subjects can calculate their operating and net profit margins as a consequence of total turn-over or net asset position. Few subjects appeared to do that, and it would not have helped much either. Though net borrowing cost was 10 % of loans, it was unclear whether a doubling of that to 20 % should be considered great or whether a higher or lower return on equity should be the goal. Note that high and low frequency benchmarks differed; the high frequency market that started with about $ 0.1 billion a final sum of $ 30 billion, thus an geometric mean annual return of 15.3 %. ({30/0.1)^(1/40) = 1.153} i.e. 15.3 % per year), while the $10 billion made in the low frequency environment corresponded to 12.2 % per year.

Thus, outcome, i.e. performance, feedback was not very useful to subjects and there should be no effect on performance from using performance feedback. However, as the decision rule, subjects who used information highlighting the markets' dynamic process should be able to increase performance if they were able to understand that process. As an indication of understanding, subjects' use of information clusters was monitored.

Figure 6.4.1 represents an a priori classification of information usefulness. Time history is only important for those variables where history shows patterns that can reveal how the future will look. Most information clusters contained information where the current year contained enough information. The design also assured that the static value of these information clusters were low. Thus, the correlation of low static and dynamic usefulness was high.

As noted in chapter 4, information about the last years' market supply line, i.e. Buildings under construction/Tonnage on order, is required to calculate the high performance benchmark rule. The same graph of time history showed evidence that the market fluctuated. Similarly, the unit cost graphs/tables showed the corresponding fluctuations of lease/transportation prices. Subjects who accessed either information sources would get an efficient view of market dynamics, but history of the entire market behavior gave better "early warning" signals about the future.

Since the design ensured that the information sources that were dynamically important also had high static value, whereas the sources having low dynamic value also were statistically unimportant, high performing subjects should access clusters of "Unit Costs" and "Market Ships". using figure 6.4.1 terminology.

Note that since graphs fitted on one computer screen, while tables required scrolling to see more than 10 years at the time, the graphs represented a more efficient access to information.

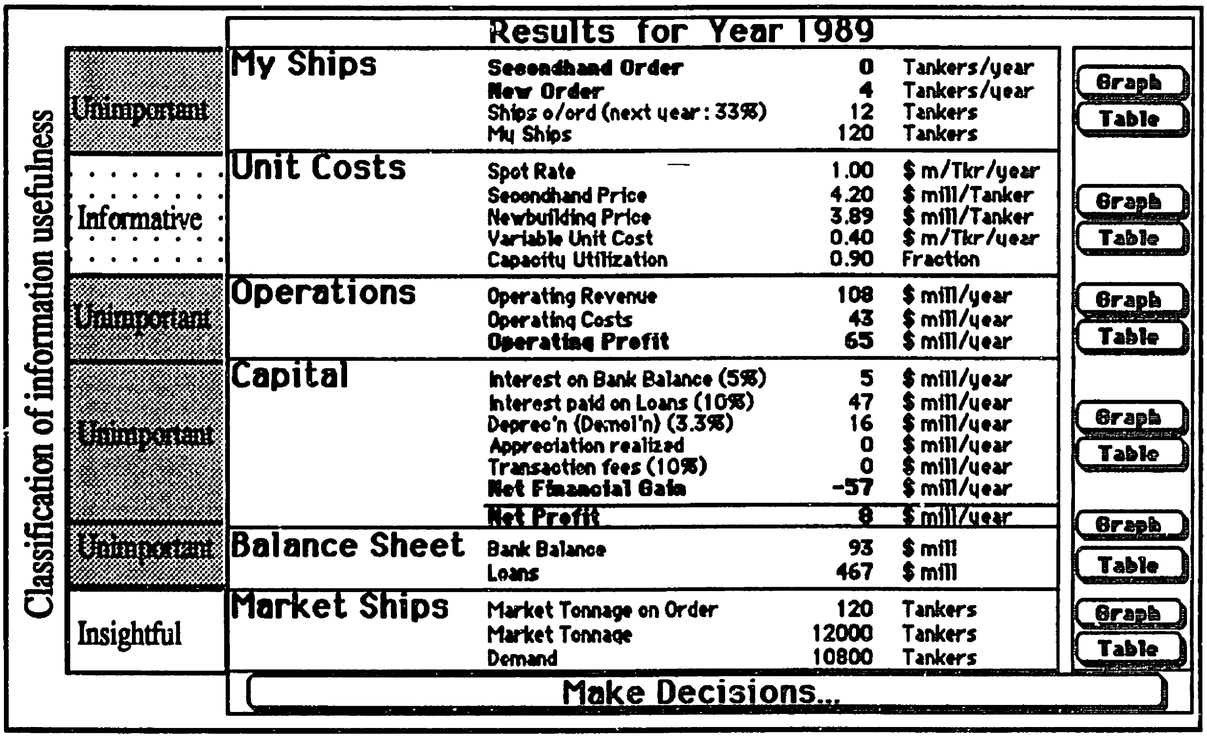


Figure 6.4.1: Classification of information usefulness

The findings, shown below in figure 6.4.2 and .3, revealed a complex picture. First, average use of the informative graph and table containing information about demand, supply and supply line (Market) are the only clusters of information whose access increased from trial 1 to trial 2 (19 subjects accessed the cluster more in trial 2, 10 accessed the source less, sign test $p < 0.13$ ns). Access to other information sources is reduced over trials, indicating some movement towards better mental models with regards to information sources (Kleinmuntz and Schkade, 1988).

		Use of information in Market 1											
		Graphs						Tables					
		Decisions	Unit Costs	Operations	Income and expenses	Balance Sheet	Market	Decisions	Unit Costs	Operations	Income and expenses	Balance Sheet	Market
PM1													
PM2							.28					.42	.35
Look-up	Avg	1.4	5.9	0.7	1.1	1.1	7.9	1.0	2.4	0.3	0.7	0.5	1.0
	Stdev	1.6	7.4	1.1	1.4	2.1	10.0	1.8	4.3	0.9	1.6	1.3	2.3

Table 6.4.2: Correlation between performance indices and use of information, first trial. (p values < 0.05).

		Use of information in Market 2											
		Graphs						Tables					
		Decisions	Unit Costs	Operations	Income and expenses	Balance Sheet	Market	Decisions	Unit Costs	Operations	Income and expenses	Balance Sheet	Market
PM1							.28						.37
PM2													.64
Look-up	Avg	0.2	5.7	0.2	0.4	0.9	8.2	0.1	1.3	0.0	0.2	0.2	0.9
	Stdev	0.4	9.3	0.9	0.8	2.3	12.4	0.4	2.3	0.3	0.8	0.6	2.6

Table 6.4.3: Correlation between performance indices and use of information, second trial. (p values < 0.05).

Table 6.4.4 shows the correlation matrix for correlations > 0.4, which corresponds to p < 0. 05. The table indicates co-linearity in information used, especially in the first trial. In addition, the diagonal in the middle sub-table also indicates a particularly high degree of cluster re-use over trials.

CORRELATION MATRIX

Submatrix: First Trial

	DECIG1	UNCOSTG1	OPSG1	FINANG1	BALANG1	MARKTG1	DECIT1	UNCOSTT1	OPST1	FINANT1	BALANT1	MARKTT1
DECIG1	1.000											
UNCOSTG1	0.400	1.000										
OPSG1			1.000									
FINANG1				1.000								
BALANG1	0.424		0.480	0.570	1.000							
MARKTG1						1.000						
DECIT1						0.502	1.000					
UNCOSTT1							0.620	1.000				
OPST1		0.452						0.758	1.000			
FINANT1							0.399	0.821	0.563	1.000		
BALANT1							0.429	0.492			1.000	
MARKTT1							0.458	0.590	0.529	0.488	0.425	1.000

Submatrix: First and second trial

	DECIG1	UNCOSTG1	OPSG1	FINANG1	BALANG1	MARKTG1	DECIT1	UNCOSTT1	OPST1	FINANT1	BALANT1	MARKTT1
DECISG2	0.513		0.434									
UNCOSTG2		0.825							0.407	0.278		
OPSG2												
BALANG2		0.553	0.411	0.430	0.747							
MARKTG2												
DECIST2						0.880	0.570					
UNCOSTT2							0.737	0.495	0.395	0.435		
OPDT2								0.600	0.416			
FIANT2								0.619		0.806		
BALANT2										0.454		
MARKTT2											0.613	0.719

Submatrix: Second Trial

	DECISG2	UNCOSTG2	OPSG2	BALANG2	MARKTG2	DECIST2	UNCOSTT2	OPDT2	FIANT2	BALANT2	MARKTT2
DECISG2	1.000										
UNCOSTG2		1.000									
OPSG2			1.000								
BALANG2				1.000							
MARKTG2					1.000						
DECIST2					0.559	1.000					
UNCOSTT2							1.000				
OPDT2								1.000			
FINANT2								0.563	1.000		
BALANT2							0.471			1.000	
MARKTT2											1.000

Table 6.4.4: Correlation matrix for information look-up. Single $p < 0.05$, i.e $r > 0.4$ used as cut-off criterion

To address the co-linearity issue, a search for common factors was initiated, using the principal components approach with varimax rotation. A scree test of the original factors indicated that five factors explained respectively 24.7, 15.1, 11.9, 8.6 and 7.5% of the variance. Three were retained and rotated. The first three rotated factors explained 20.9, 15.7 and 15.1% of the variance. The component loadings of the rotated factors is shown below in table 6.4.5.

ROTATED LOADINGS

		Factor	1:smart	2:random	3:stupid	Usefulness
Use of information in trial 1	Graphs	Decisions	0.16	0.72	-0.17	
		Unit Costs	0.44	0.54	0.09	+
		Operations	-0.20	0.56	0.33	
		Income and expenses	0.06	0.57	0.10	
		Balance Sheet	0.02	0.77	0.10	
		Market	0.73	0.16	-0.07	+++
	Tables	Decisions	0.83	0.08	0.20	
		Unit Costs	0.51	0.09	0.78	+
		Operations	0.42	0.11	0.58	
		Income and expenses	0.25	0.05	0.89	
		Balance Sheet	0.48	-0.15	0.31	
		Market	0.79	-0.11	0.13	++
Use of information in trial 2	Graphs	Decisions	0.02	0.58	0.38	
		Unit Costs	0.28	0.56	0.02	+
		Operations	0.20	0.17	0.14	
		Balance Sheet	0.04	0.82	-0.10	
		Market	0.78	0.18	-0.10	+++
	Tables	Decisions	0.70	0.09	0.12	
		Unit Costs	0.53	0.01	0.12	+
		Operations	0.12	-0.00	0.92	
		Income and expenses	0.03	0.11	0.59	
		Balance Sheet	0.32	0.09	0.01	
		Market	0.60	-0.26	0.01	++

Table 6.4.5: Rotated component loadings with indicated usefulness..

These loadings were used on the original information access data set.

Though none of the factors could explain performance in either trial by themselves, regressing the three factors on PM1 first and PM2 second showed that the first factor which loads heavily on the important information sources and negatively on some of the unimportant explains performance in both trials as shown in table 6.4.6. The fact that "stupid" explains negative performance in the first trial was also expected. The fact that "random" information use reduces performance in the second trial is also noticeable. The bottom part of the table also indicates that PM1 plays an insignificant role in explaining PM2 when the information access is controlled. This indicates that the first trial serves to orient subjects in the information access environment (though learning to leave dear information sources appears to be hard, as indicated by 6.4.4).

First trial

1 CASES DELETED DUE TO MISSING DATA.

DEP VAR: PM1 N: 40 MULTIPLE R: .431 SQUARED MULTIPLE R: .185
ADJUSTED SQUARED MULTIPLE R: .118 STANDARD ERROR OF ESTIMATE: 0.235

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	P(2 TAIL)
CONSTANT	0.143	0.054	0.000	0.012
smart	0.005	0.002	0.498	0.020
random	-0.005	0.004	-0.257	0.216
stupid	-0.012	0.006	-0.295	0.070

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
REGRESSION	0.453	3	0.151	2.731	0.058
RESIDUAL	1.992	36	0.055		

Second trial

DEP VAR: PM2 N: 40 MULTIPLE R: .484 SQUARED MULTIPLE R: .234
ADJUSTED SQUARED MULTIPLE R: .146 STANDARD ERROR OF ESTIMATE: 0.286

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	P(2 TAIL)
CONSTANT	0.209	0.072	0.000	0.006
PM1	0.238	0.203	0.192	0.249
smart	0.006	0.003	0.436	0.053
random	-0.011	0.005	-0.456	0.033
stupid	-0.006	0.008	-0.125	0.447

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
REGRESSION	0.876	4	0.219	2.672	0.048
RESIDUAL	2.868	35	0.082		

Table 6.4.6: Information access patterns help explain performance.

This section has shown that information use indeed can explain performance. Subjects who have a pattern of smart information use indeed do better than those who access unimportant information or those that access information randomly. Moreover, it appears that subjects tend to stay with their information sources as indicated by the fact that correlation between information sources across trials was high. This indicates that subjects who do not understand what information to use, and so do poorly, have a hard time finding the more useful information clusters. This again suggests that subjects who enter the experiment with the wrong frame of mind will continue to do poorly as we saw in section 5.3. As shown in table 6.4.7 below, right information use is related to subjects' experience in economics.

DEP VAR:"Smart" N: 41 MULTIPLE R: .356 SQUARED MULTIPLE R: .127
ADJUSTED SQUARED MULTIPLE R: .030 STANDARD ERROR OF ESTIMATE: 23.456

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	P(2 TAIL)
CONSTANT	7.946	21.920	0.000	0.719
SDFACT	-0.393	1.240	-0.053	0.753
ECONFACT	3.483	1.572	0.388	0.033
BSFAC	-4.984	3.587	-0.253	0.173
CSFAC	0.002	1.804	0.000	0.999

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
REGRESSION	2875.843	4	718.961	1.307	0.286
RESIDUAL	19807.081	36	550.197		

Table 6.4.7: Subjects with training in economics appear to access the "smart" information cues.

We have seen, however, that subject academic background did not explain performance in the second trial. Thus, it appears that subjects' economics background helps them frame the probiem in terms of accessing the right information, but apparently a deeper understanding is required in addition, to be able to use this information successfully.

The next session will look into the role of causal understanding and investigate the degree to which questionnaire data can explain performance

6.5 Causal understanding

The previous chapters have indicated that causal understanding tends to be problematic in real world decision environments, especially those characterized by long delays between action and its feedback consequences. Causal understanding indicates the set of relationships held in a decision maker's head about causal links between variables in the market.

Chapter 5 showed that treatment conditions affected performance. Presumably the complexity of unfamiliar and high frequency environments consume cognitive resources and so make feedback links more opaque. This presumption can be tested by monitoring how subjects rate the direction and delay of causal links in the experimental markets. In this section relationships between the quality of the mental model, treatment conditions and decision performance will be described using data derived from a causal questionnaire.

The design of this questionnaire needs to be explained further, since there exist a multitude of causal links that may be monitored. A comprehensive description of subjects' mental models would be prohibitive, even in a simple experimental decision environment. Recall that the experimental decision environment was presented to the subjects by a game interface. The user interface showed 24 variables (See appendix 2 for the full equation list). Between the variables on the screen there are therefore 24 x 23/2 (276) possible relevant uni-directional relationships that may be presented in a decision maker's problem space (not taking the relationships between a variable and itself over time, nor the 276 reverse relationships, into account). The fact that professionals' content rich mental models interfere with the task at hand to decrease performance and hinder learning, indicates that many more relationships may be present in subject's mental models.

All 276 causal links in a subject's mental model may be monitored before and after each decision. Considering the 80 decision rounds and two decisions per round, and that each causal link would have to be qualified by a scale with at least 3 ratings (positive influence from variable a to b, negative influence and no influence) the measurement task would entail 276 x 80 x 2 x 3 (132480) responses by each subject.

Consequently, an alternative way to monitor mental models was chosen. A sample of 28 questions was selected. Five response alternatives existed and the full questionnaire was administered both before and after each of the two games were played. Insert 6.5.1 below shows the introductory text.

Introduction

In the following, you are asked to use your best guess to rate relationships between variables in the model. Imagine that the left bold face variable below increases. What will happen to the right variable? Will it increase immediately or after a delay? Will it remain unchanged? Will it decrease immediately or after a delay? An immediate increase implies that an increase in the left variable is followed by an increase in the right variable in the same year. If the left variable stays constant after an initial increase, so will the right variable.

In a delayed relationship, however, the right variable continues to increase for some time even after the left variable is constant. As an example, if the left variable is "Incoming Students" and the right variable is "Graduations", the realtionship is "delayed increase", because "Graduations" continue to increase even after "Incoming Students" have stopped increasing. Students remain in the pipeline for some time and "Graduations" thus are a delayed function of the "Incoming Students".

Do not dispair if you find the task difficult or the alloted time short. Just do your best. If you know whether the relationship is "increase" or "decrease", but uncertain as to whether it is "immediate" or "delayed", just choose what feels best of the two latter.

You have all the time you need to familarize yourself with the screen. When you feel you are ready to make the ratings, push on the "Now Make Choices" button. You make choices by clicking in the appropriate button. Having done that, you have about 25 seconds to make the next choice. A timer will appear in the upper right corner of this screen once you have finished this introduction.

After about 24 questions about cause and effect relationships, you will be asked to rate how quickly the market reacts to disturbances. More about that later...

Insert 6.5.1: Introductory text.

Table 6.5.2 shows the response alternatives.

..Relative to Tankers

An increase in...

Mkt Tonnage	leads to...		in... **Newbuilding Starts**
Average Life Time of Tankers	leads to...		in... **Depreciation**
Mkt Tonnage u/Constr	leads to...		in... **Secondhand Price**
Spot Rate	leads to...		in... **Mkt Tonnage**
Mkt Tonnage u/Constr	leads to...		in... **Mkt Tonnage**
Secondhand Price	leads to...	O **immediate increase**	in... **Spot Rate**
Operating Profits	leads to...	O **delayed increase**	in... **Operating Costs**
Mkt Tonnage u/Constr	leads to...	O **no change**	in... **Spot Rate**
Interest on Bank Balance	leads to...	O **immediate decrease**	in... **Bank Balance**
Newbuilding Starts	leads to...	O **delayed decrease**	in... **Capacity Utilization**
Transaction fees	leads to...		in... **Loans**
Average Life Time of Tankers	leads to...		in... **Secondhand Price**
Newbuilding Starts	leads to...		in... **Secondhand Price**
Price Elasticity of Newbuilding	leads to...		in... **Mkt Tonnage**
Mkt Tonnage	leads to...		in... **Mkt Tonnage u/Constr**
Appreciation realized	leads to...		in... **Transaction fees**
Loans	leads to...		in... **Interest paid on Loans**
Operating Costs	leads to...		in... **Operating Profits**

Table 6.5.2: Causal questionnaire

The questionnaire existed in two forms; --tanker and real estate-- and contained questions that were presented line by line about the context/frequency condition that had been finished just before the questionnaire was presented or was about to be played. A 25 second time limit existed for each question, after which the next question appeared until all questions were asked. A new question would not appear before the allotted 25 seconds had run out, so the subject could change his mind during that time.

The score is built on the assumption that when subjects indicate causal relationships, the causal map resulting from a subject's rating reflects the causal understanding that underlies his or her decisions. This assumption is itself problematic: Subjects may have separate cognitive domains for action and inference (Broadbent and Broadbent, 1986).

As noted by Ericsson and Simon (1984), however, the issue of dual problem spaces may be resolved by having elicitation methods tapping the same cognitive processes as those used during problem solving. Concurrent verbalizations is one way of achieving congruence between different problem spaces. The present set-up was designed so as to bring together the questionnaire mode of knowledge elicitation with the decision mode:

The time to answer was quite short to facilitate subjects' providing their intuitive answers. Furthermore, the questionnaires involved the same variable names as the experimental markets.

As mentioned the questionnaire was used 4 times: At the beginning and end of each trial. Of course, the sequence of questions was randomized each time the questionnaire was used.

There were 5 response alternatives, each giving 1 point if right, 0.5 point if "half right" and 0 points if wrong. The score 0.5 was given if the relationship in question was correct, but had an inverted delay classification (immediate/delayed). Thus, the likelihood of scoring correctly at random was 1/5 or 20%, plus the likelihood of scoring 0.5 points that applied to the four out of five cases that yields (0.5*4/5)/5 or 8%, for a total of 28 %. As will be shown below, average subject response was 49 % of correct.

The main prediction was that treatment conditions cause performance differences by making the task more or less transparent. Thus, subjects in less transparent conditions will score lower on the questionnaire. Since the market context was presented to subjects in form of a context dependent newspaper article, and a context <u>and</u> frequency dependent game description before the questionnaire was used, one can explain questionnaire score by the immediately following and preceding context and frequency conditions. This has been done in equation 6.1. All 40 subjects' 4 questionnaire scores have been assigned a questionnaire number (1, 2, 3 and 4), a context, and a frequency. The following model was used

$$Qscore_{1,2,3,4} = \alpha_0 + \beta_1 Qn + \beta_2 C_{1,2} + \beta_3 F_{1,2} + \beta_4 C_2 F_{1,2} + \varepsilon \quad 6.1$$

Recall from chapter 5 that Q_n refers to questionnaire number, F_i refers to frequency in trial i, C_i refers to context in trial i, α is the intercept, and β_j the direction and strength of the variables in question. Qscore is the questionnaire scores from all 4 questionnaires.

Table 6.5.1 below shows questionnaire score means and standard deviations.

	Q1	Q2	Q3	Q4
N OF CASES	41	40	39	39
MEAN	0.450	0.497	0.483	0.548
STANDARD DEV	0.116	0.158	0.119	0.159

Table 6.5.1: Questionnaire score means and standard deviations.

Subjects improve their understanding from during the course of the experiments, as evidenced by the significance of a paired t-test of the Q1 and Q4 means ($p < 0.005$).

Table 6.5.2 shows that the same conditions that cause high performance also cause better understanding of causal links. Statistical significance is overstated using 152 degrees of freedom since questionnaire scores are not independent but correlated. Consequently, the analysis was performed with 34 instead of 152 degrees of freedom in the denominator. The p-values are interpolated using an F-table.

ANALYSIS OF VARIANCE

DEP VAR: QSCORE N: 159 MULTIPLE R: .305
SQUARED MULTIPLE R: .093

SOURCE	DF	MEAN-SQUARE	F-RATIO	P
QN	3	0.046	2.501	0.072
FREQ	1	0.062	3.343	0.080
CONTEXT	1	0.062	3.346	0.080
FREQ* CONTEXT	1	0.009	0.478	0.500
ERROR	34	0.019		

ESTIMATES OF EFFECTS $B = (X'X)^{-1} X'Y$

		QSCORE
CONSTANT		0.505
QN	1	-0.041
QN	2	0.004
QN	3	-0.004
FREQ	High	-0.020
CONTEXT	Familiar	0.020
FREQ CONTEXT	High Familiar	-0.008

Figure 6.5.2: ANOVA of questionnaire score as function of questionnaire number and treatment conditions.

Logically, the game experience is a much more powerful conveyer of the frequency effect than two simple variables in the game description. Consequently, figure 6.5.3 below shows only q2 and q4 frequency results, but all context results.

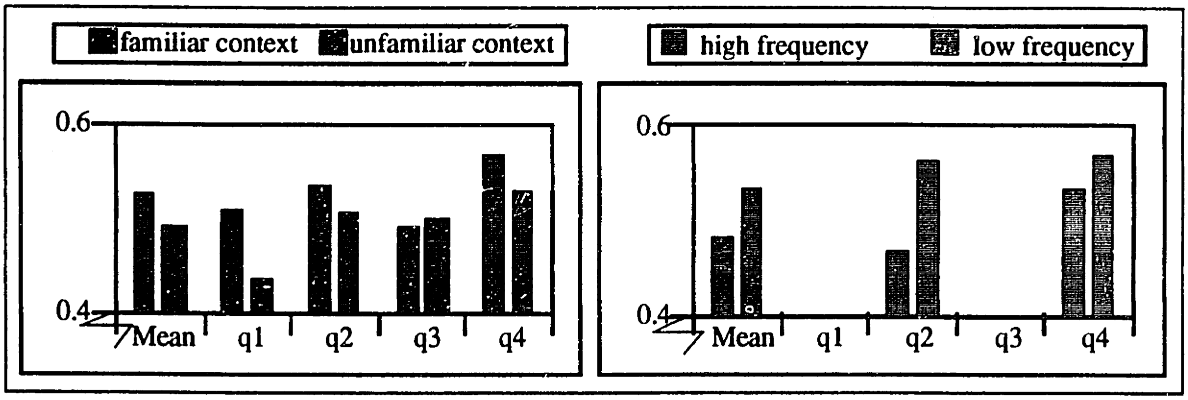


Figure 6.5.3: Context familiarity and high frequency lead to better mental model of the task.

Current conditions

The main prediction emanating from chapter 5 is that the initial mental model, combined with the understanding that takes place during the game playing, will explain the mental model of a subject at the end of a trial session. The question is, will the treatment conditions explain the residual variance ? This was tested by equation 6.2

$$Q2= \alpha_0 + \beta_1 C_1 + \beta_2 F_1 + \beta_3 C_1 F_1 + \beta_4 PM_1 + \beta_5 Q1 + \varepsilon \qquad 6.2$$

The analysis of variance in table 6.5.3 shows that questionnaire scores after the trial reflect performance as well as initial understanding (Q1), but also that when performance and quality of the initial mental model are controlled, the context and frequency do not matter significantly to the quality of the initial mental model.

ANALYSIS OF VARIANCE

2 CASES DELETED DUE TO MISSING DATA.

DEP VAR: Q2 N: 39 MULTIPLE R: .580
SQUARED MULTIPLE R: .336

SOURCE	DF	MEAN-SQUARE	F-RATIO	P
CONT1	1	0.018	0.962	0.334
FREQ1	1	0.001	0.064	0.803
PM1	1	0.126	6.568	0.015
Q1	1	0.141	7.323	0.011
CONT1* FREQ1	1	0.006	0.315	0.579
ERROR	33	0.019		

ESTIMATES OF EFFECTS $B = (X'X)^{-1} X'Y$

		Q2
CONSTANT		0.190
CONT1	FAM	-0.024
FREQ1	HIGH	-0.007
PM1		0.284
Q1		0.584
CONT1 FREQ1	FAM HIGH	-0.014

Table 6.5.4 ANOVA of Q2 as a function of treatment conditions, earlier causal understanding and early performance.

However, we know that PM1 is influenced by treatment conditions. By allowing the treatment conditions to explain all variance, an effect may be found according to equation 6.3:

$$Q2= \alpha_0 + \beta_1 C_1 + \beta_2 F_1 + \beta_3 C_1 F_1 + \varepsilon \qquad 6.3$$

In the ANOVA table 6.5.5 below, there is evidence that subjects' understanding is poorer after a high frequency condition than after a low frequency condition. However, there is no significant impact on scores from context familiarity.

ANALYSIS OF VARIANCE

1 CASES DELETED DUE TO MISSING DATA.

DEP VAR: Q2 N: 40 MULTIPLE R: .335
SQUARED MULTIPLE R: .112

SOURCE	DF	MEAN-SQUARE	F-RATIO	P
CONT1	1	0.005	0.201	0.657
FREQ1	1	0.089	3.717	0.062
CONT1* FREQ1	1	0.009	0.390	0.536
ERROR	36	0.024		

ESTIMATES OF EFFECTS $B = (X'X)^{-1} X'Y)$

		Q2
CONSTANT		0.511
CONT1	FAM	0.011
FREQ1	HIGH	-0.048
CONT1 FREQ1	FAM HIGH	-0.015

Table 6.5.5 ANOVA of Q2 as a function of treatment conditions only.

Independently, Q4 should be explained by the treatments, prior mental model and second trial performance. The following model was tested, reflecting equation 6.2:

$$Q4 = \alpha_0 + \beta_0 PM2 + \beta_1 Q3 + \beta_2 C_2 + \beta_3 F_2 + \beta_4 C_2 F_2 + \varepsilon \qquad 6.4$$

As shown below in table 6.5.6, no such relationships between Q4 and the treatments were found. Q4 is almost uniquely determined by Q3 (r=0.79).

ANALYSIS OF VARIANCE

2 CASES DELETED DUE TO MISSING DATA.

DEP VAR: Q4 N: 39 MULTIPLE R: .737
SQUARED MULTIPLE R: .543

SOURCE	DF	MEAN-SQUARE	F-RATIO	P
PM2	1	0.008	0.600	0.444
CONT2	1	0.000	0.023	0.881
FREQ2	1	0.000	0.000	0.995
Q3	1	0.336	25.288	0.000
FREQ2* CONT2	1	0.040	3.045	0.090
ERROR	33	0.013		

ESTIMATES OF EFFECTS B = (X'X) X'Y)

		Q4
CONSTANT		0.120
PM2		0.057
CONT2	1	0.003
FREQ2	1	-0.000
Q3		0.848
FREQ2	1	Not interpretable
CONT2	1	0.036

Table 6.5.6: ANOVA of questionnaire scores.

Subsequently, a formulation parallel to equation 6.3 was run where both PM2 and Q3 were taken out. This equation could not, however, explain differences in treatment conditions either. As it adds nothing new to table 6.5.5, it is not shown.

This section has looked into the relationship between current conditions, questionnaire scores and performance. Chapter 5 showed links between current performance and current conditions. Above, a similar, though less statistically significant, relationship was found between questionnaire scores and treatments. Below we will look into the role of treatment conditions in explaining transfer effects.

Transfer

In chapter 5 it was indicated that an initial unfamiliar context improves later performance because subjects in that condition get an improved understanding as a consequence of the treatment. Since it has already been shown that Q2 is not positively affected by an unfamiliar context, it remains to be investigated whether there is a delayed effect so that the better understanding is manifested in an improvement from questionnaire 2 to 3 (remember that the general effect from Q2 to Q3 is negative as shown by figure 6.5.3 above). Using the score change from Q2 to Q3 and its interaction with the initial context, equation 6.6 shows the formulation:

$$PM_2 = \alpha_0 + \beta_1(Q3\text{-}Q2) + \beta_2 C_2 + \beta_3 F_2 + \beta_4 C_1 + \beta_5 F_1 + \beta_6 C_2 F_2 + \beta_7 F_1 C_1 + \beta_4 C_1(Q3\text{-}Q2) + \varepsilon \qquad 6.6$$

Table 6.5.6 below shows that though performance improves from trial 1 to trial 2, the improvement in questionnaire scores related to the first context does not explain PM2 scores. Changing the (Q3-Q2) difference scores to Q3 scores does not modify this conclusion.

The effects of current context and frequency as well as the interaction between them is the same in table 6.5.7 as it was in chapter 5.

ANALYSIS OF VARIANCE

2 CASES DELETED DUE TO MISSING DATA.

DEP VAR: PM2 N: 39 MULTIPLE R: .594
SQUARED MULTIPLE R: .352

SOURCE	DF	MEAN-SQUARE	F-RATIO	P
Q3MINQ2	1	0.003	0.044	0.836
CONT1	1	0.035	0.448	0.509
FREQ1	1	0.005	10.057	0.813
CONT2	1	0.321	4.056	0.053
FREQ2	1	0.462	5.842	0.022
CONT1* FREQ1	1	0.017	0.209	0.650
CONT2* FREQ2	1	0.249	3.152	0.086
CONT1* Q3MINQ2	1	0.005	0.066	0.799
ERROR	30	2.372	0.079	

Table 6.5.7: ANOVA of PM2 as a function of improvement in questionnaire scores.

As noted in the previous chapter, an unfamiliar context may lead to improved transfer since it requires a deeper search in the problem space for causal relations. The same argument should yield improved transfer as a consequence of changes in the experimental environment. No such changes were noted in the performance data.

This may, however, be due to the fact that when subjects encounter a changed environment, two concurrent phenomena that cancel each other out

1. There is increased cognitive search for causal connections (which should increase deep understanding, but decrease performance in the changed environment)

2. There is less familiarity with the experimental environment (which should decrease surface understanding and performance in the task, but have little impact on deep understanding).

Chapter 5 showed that an initial unfamiliar environment increased performance in the second trial. This provides evidence for contention 1: In a less familiar environment, subjects must define new cognitive concepts. This requires effort that detracts from the task they are performing. Consequently, current performance is poor. Yet, subjects in the unfamiliar condition have had a more difficult task. The construction of a new schema shows up as increased performance in trial 2.

Similarly, subjects who experience a changed environment are under a higher cognitive load. Though the load is different since it is less a question of creating a new schema than modifying an old, it is suggested that the increased load will detract from performance. This task of modifying the old schema should take place during the reading of the background materials and during the early parts of the game. Later parts of the game should yield increased performance. However, there is an additional burden related to the first one. The familiarity with the experimental environment is lower for changed condition subjects. This will decrease surface understanding, and reduce attentive resources so that performance is hindered during the task (just like any unfamiliar game). Consequently, there are two opposing factors with regards to performance. In chapter 5 we consequently saw no effect on performance after a changed context.

Yet the added challenge of modifying the old schema to also explain the new information should improve the quality of that schema significantly. Investigating the particular role of context change, we have equation 6.7:

$$Q4 = \alpha_0 + \beta_1 PM2 + \beta_2 C_1 C_2 + \beta_3 F_1 F_2 + \varepsilon \qquad 6.7$$

Table 6.5.8 below shows that the mental model is better when subjects are exposed to a changed context, thus supporting the argument that a harder environment actually yields a better mental model.

ANALYSIS OF VARIANCE

2 CASES DELETED DUE TO MISSING DATA.

DEEP VAR: Q4 N: 39 MULTIPLE R: .336
SQUARED MULTIPLE R: .113

SOURCE	DF	MEAN-SQUARE	F-RATIO	P
PM2	1	0.033	1.366	0.250
CONT1* CONT2	1	0.080	3.284	0.079
FREQ1* FREQ2	1	0.001	0.040	0.844
ERROR	35	0.024		

ESTIMATES OF EFFECTS B = (X'X)$^{-1}$ X'Y)

		Q4
CONSTANT		0.523
PM2		0.096
CONT1 CONT2	FAM FAM	-0.045
FREQ1 FREQ2	HIGH HIGH	0.005

Table 6.5.8: Questionnaire score, performance, context and frequency changes.

Note that the poor performance in the high frequency environment was explained as the effects of increased perceived risk in addition to a higher cognitive load in that condition. In fact, the hypotheses stated in chapter 3 even suggested a positive impact on performance from a high frequency environment. The lack of effect on the mental model from a changed frequency is therefore not very surprising.

The questionnaire scores suggest a complex picture where those environments that improve current-trial performance are less conducive to transfer of understanding. This finding indicates that task environments leading to high cognitive load among subjects may be more efficient in helping subjects create a generalizable causal web, similar to Gentner and Toupin's (1986) notion of an exemplar.

In terms of designing effective learning labs, this suggests that though *performance* in the lab is increased by a moderately familiar context, the *learning in more taxing environments is more likely to endure.* As noted in chapter 5, increased cognitive complexity can stem from an unfamiliar task or high frequency environment.

Likewise, a task where subjects have strong preconceptions, such as professionals have when encountering a learning lab of a well-known context with unexpected emphasis and conclusions, may also be characterized as complex. The finding in this subsection of changes in context improving understanding, suggests that learning is harder to achieve in complex environments. Once achieved, however, such learning is more likely to endure.

6.6 Summary and discussion

Treatment conditions appear to yield differential performance effects by making the experimental market in unfamiliar environments opaque. One would hope that academic training would significantly help subjects see through information opaqueness and understand underlying relationships better so as to improve performance. And indeed, subjects trained in business and engineering do better in the first trial. However, this initial advantage has no carry-over effect. The exposure to the task itself puts all subjects on an even basis for trial two: background differences wash out. For accounting, finance, computer science and business strategy, the fact that training did only help subjects in the first trial was expected. Background in system dynamics and economics, however, should help subjects see the underlying phenomenon (an unstable market with highly elastic prices where there exist profit opportunities from buying low and selling high). These backgrounds would help subjects transfer better, especially from the cognitively taxing environments. The above hypotheses of better task insight and transfer from system dynamics and economics and the finding that cognitive complexity might help transfer of insights were re-examined, using the following formulation:

$$Q4 = \alpha_0 + \beta_1 C_2 + \beta_1 F_2 + \beta_3 \Delta C^* SD + \beta_4 C_1^* EC + \varepsilon \qquad 6.7$$

Where SD and EC denote the aggregated number of courses taken in system dynamics and economics respectively. Table 6.7.1 below shows the findings

ANALYSIS OF VARIANCE

SOURCE	DF	MEAN-SQUARE	F-RATIO	P
CONTZ	1	0.068	3.016	0.092
FREQ2	1	0.007	0.330	0.569
SDFACT* DELTACON	1	0.037	1.642	0.209
ECONFACT* DELTACON	1	0.009	0.390	0.537
ERROR	34	0.023		

Table 6.7.1: The interaction between subject background and contexts

As shown in table 6.7.1, however, no such interactions effects were found.

As noted by Kleinmuntz and Thomas, (1987) it appears that subjects have some poor prior mental model of the experimental market and no good heuristics to help improve it.

Kleinmuntz and Schkade (1988) and Dawes (1979) have argued that a main cause for deficient decision making is that subjects cannot discriminate helpful from unhelpful information sources. Since they do not know the outcomes they could have obtained using different decision and information acquisition strategies, judging the usefulness of various information sources is problematic to subjects. There is a movement from unimportant to more important information sources, but it is slow.

Section 6.5 revealed that current performance is related to the quality of understanding. The tasks appear more transparent when they are familiar and appear in the low frequency conditions. Transparency leads to less cognitive burden on the subjects and increases mental resources to deal with the task at hand. At the same time, high transparency also induces subjects to take more risks which should yield improved transfer from transparent environments (Bakken et al., 1992). No positive transfer effect from high transparency on understanding was shown, however.

On the contrary, chapter 5 indicated that a less transparent initial task context helps transfer performance. Questionnaire data corroborated this finding by showing that a changed context in trial 2, though not improving current task (i.e. trial 2) performance, led to an improvement in the mental model at the end of that trial. One can only speculate what effects such an improved mental model will have on subsequent performance. Gentner and Toupin (1986) have shown, however, that when subjects encounter several instances of isomorphic tasks in various contextual disguises, they build exemplars. The fact that late questionnaire scores are significantly improved by a change in context suggests that, during or after trial 2, the rudiments of an exemplar are constructed in subjects' problem space. The questionnaire can indicate such an exemplar, since it emphasizes the cause-and-effect relationships that would underlie such an exemplar.

Treatment conditions produce delayed consequences in terms of transfer of understanding and performance. These effects are opposite of the immediate effects of the same treatments and suggests that causal understanding operates on at least two levels. The mental model in terms of correctness of causal relationships helps subjects identify the task characteristics and improves decision making for the current task. On a different level, though, challenging conditions, such as an unfamiliar context may activate deeper processing. This processing appears to have a dual effect of establishing deeper understanding that is reflected in increased later questionnaire scores and later performance, but the side-effect is that attention is taken away from understanding and performing in the task at hand. Since these two effects cancel out, there are no positive performance influences from cognitively taxing tasks.

Figure 6.6.2 summarizes the findings and the discussion and extends the findings to future research. The figure shows that the treatment effects on performance and understanding are similar. Yet, while cognitively easy tasks are conducive to high performance, more complex tasks further transfer. Consequently, the expectation is that later tasks, if transfer is needed, will be improved by cognitively complex environments. Note that the findings and the example in the figure are only supported by the context dimension of the experiment.

The cognitive effort implied by the frequency dimension is different than the one suggested by context. This will be one of the conclusions discussed in the next chapter.

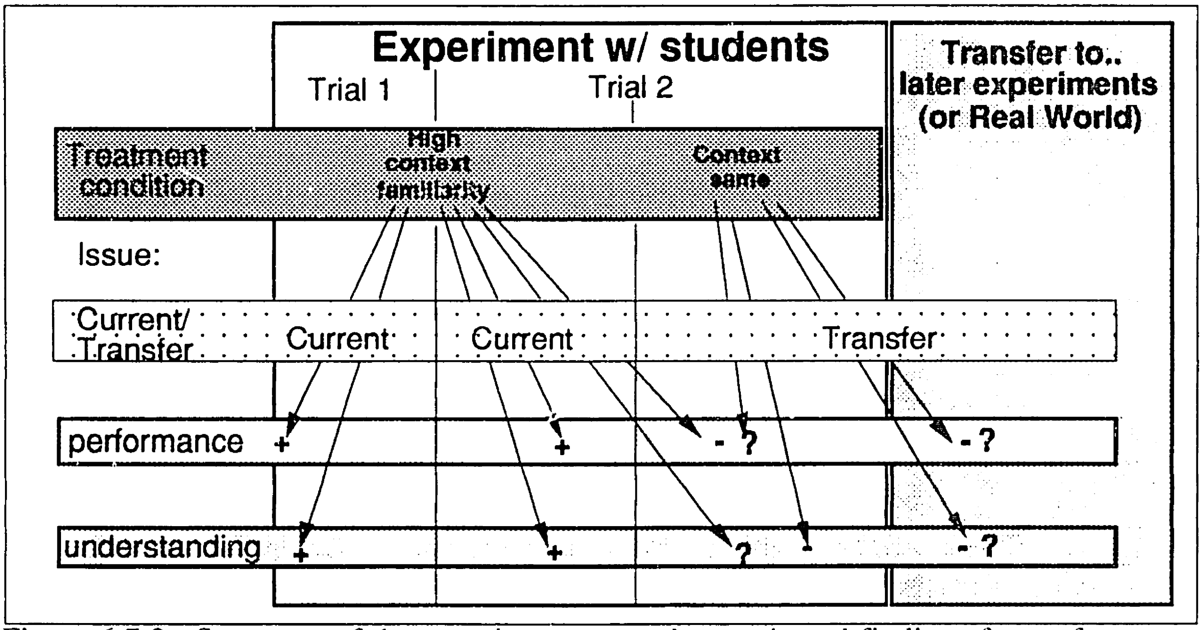


Figure 6.7.2: Summary of the most important and unequivocal findings for performance and transfer among student subjects. Transfer to the real world suggested.

Figure 6.7.2 suggests that a future study could extend the present design by adding a third trial so as to mimic the real world. It should investigate whether context changes between trial 1 and 2, that we know improve Q4, also will improve later performance and understanding. This suggestion for future research and more general conclusions and implications are put forward in the next chapter.

7. Conclusions, implications and future research

7.1 Introduction

This chapter outlines methodological problems before it draws conclusions and implications for related research in dynamic decision making and problem solving. Likewise, the experiments make several suggestions for how to conduct learning exercises with a particular emphasis on how simulations such as these may help performance in unstable markets. Future research is also outlined.

7.2 Conclusions

The hypotheses stated in chapter 3 and the findings are reported in this section.

7.2.1 Methodological cautions

The findings were not very strong and should be interpreted with caution: First, the subject data did not entirely conform to a normal distribution. Though there were no single outliars, the upper decile of the data appears to carry much of the variance. For instance, while three out of four main performance conditions (Unfamiliar/High; Unfamiliar/Low; Familiar/High) were indistinguishable, one cell, containing about 10 performance observations in each trial, explained almost all systematic variance. Though this finding was replicated, and so provides some indication of robustness, the fact that the some data cells had far higher variance than others must remain in the readers' mind.

Questionnaire data were better behaved. However, the difference between high and low score cells was quite small, and the statistical significance of findings consequently marginal. Likewise, variances in subject background data were too small to enable strong interpretations.

The data have been discussed thoroughly as if the results were reliable rather than random noise. This has been done, because prior findings and supplementary data support marginal performance findings. Similarly, as long as these cautions are made, it appears that the rigorous data analysis and discussion will benefit future studies of similar issues as well as being an eye-opener for those with more applied interests.

7.2.2 Review of findings

Performance is weak: Subjects in the first and second trial perform on average at about 5% of the simple decision rule. The findings are sobering in other aspects, too. Treatment conditions do influence performance and understanding, but have very limited effects on transfer: Practice effects exist, but with a few notable exceptions treatment conditions do not significantly improve transfer of insight and performance.

If the treatment conditions do not ensure transferable learning, is there evidence that other factors do? The appendix 1 reports pilot studies that were also training sessions. In these, post-game discussions played a major role and, according to participants, made the experience more meaningful. Discussions and other reflective exercises (Kim, 1989; Senge 1992) may improve participants' receptiveness to new ideas (Schön, 1983). On the other hand, though reflections by definition increase verbal activity and cognitive effort related to that activity, it remains to be proven that increased reflectiveness has any influence on decision making in environments that are not by themselves reflective.

Sobering in terms of prospects for improving the value of simulations as learning devices, these weak effects document performance, learning and transfer in dynamic environments. The figures 7.2.1 and 7.2.2 below reformulate the research questions, their operationalization and the findings. The next three sections explain in more detail the findings relative to context and frequency and in particular address the issue of why the cognitive complexity induced by an unfamiliar task is different from the poor system transparency involved in the high frequency condition.

Underlying issue	Operationalized Issue	Treatment conditions	Underlying metric	Operationalized metric	Trial 1 metric	Trial 2 metric
System transparency	Dynamic compression	10/20 year system period	Performance	Subject performance compared to benchmark rule's (PM1 & PM2)	PM1	PM2
			Transfer	Subject performance compared to benchmark rule's (PM1 & PM2)	NA	Trial 1 conditions/ changes in conditions; PM1 on PM2
Cognitive complexity	Context familiarity	Familiar (real estate) Unfamiliar (oil tanker)	Performance	Subject performance compared to benchmark rule's (PM1 & PM2)	PM1	PM2
			Transfer	Subject performance compared to benchmark rule's (PM1 & PM2)	NA	Trial 1 conditions/ changes in conditions; PM1 on PM2
Interaction with performance, learning, transfer, and treatment conditions	Help from academic training	None	Expsoure to various contexts and disciplines	Number of undergraduate and graduate classes taken	NA	NA
Interaction with performance, learning, transfer, and treatment conditions	Help from work experience	None	Expsoure to various contexts and disciplines	Years' work experience	NA	NA
Interaction with performance, learning, transfer, and treatment conditions	Information acqusition	None	Information access	Number of look-ups in graphs and tables with varying usefulness	Trial 1 look-ups	Trial 2 look-ups
Interaction with performance, learning, transfer, and treatment conditions	Quality of mental model	None	Understanding of causal links within the task	Questionnaire scores	Q1; Q2	Q1; Q2

Table 7.2.1: Overview of research design and operationalized metrics.

Underlying issue	Operationalized Issue	Expected effect	Achieved effect	Comments
System transparency	**Dynamic compression**	10 yr more transparent than 20; 10 yr higher performance	Opposite effect	Low freq task may be less stressing through less debt burden Low freq task means both "slow motion" and more time to learn before critical decisions
		10 yr more transparent than 20; 10 yr better transfer	No effect on performance	
Cognitive complexity	**Context familiarity**	Real estate less taxing than oil tanker; RE higher performance more transparent than Oil tkr	As predicted No effect on performance from changed context	Lack of transparency may lad to increased search, more info processing; help transfer However, relationship can be reversed with much context familiarity
		Real estate less taxing than oil tanker; RE better transfer more transparent than Oil tkr	Very weak, opposite, effect from initial unfamiliar context	Change in context leads to better Q4 scores Complexity may lad to increased search and processing that help transfer
Interaction with performance, learning, transfer, and treatment conditions	**Help from academic training**	All courses will increase familiarity and PM1; System Dynamics and Economics will increase transfer	Effects on PM1 from undergrad education, no transfer effects	
Interaction with performance, learning, transfer, and treatment conditions	**Help from work experience**	All work experience in RE and Oil tankers will increase familiarity and PM1; System Dynamics and Economics will increase transfer	Effects on PM1 from SD work experience, no transfer effects	
Interaction with performance, learning, transfer, and treatment conditions	**Information acqusition**	Subjects who do well will access important market information	As predicted	
Interaction with performance, learning, transfer, and treatment conditions	**Quality of mental model**	Understanding will be related to conditions in the same way as performance and transfer	Performance effects replicated	Appears that hard tasks have two effects: Negative on current performance and current men.al model. Positive on the quality of future mental model and performance.

Table 7.2.2: Overview of research design and findings.

7.2.3 Context

The results corroborated the main hypothesis that a familiar context leads to higher performance: Subjects performed better in the real estate environment. The cause for poorer performance in the less familiar oil tanker environment is that the unfamiliar contexts reduce market transparency; subjects need to use significant cognitive resources just to integrate the many new concepts (tonnage, spot rate, etc.). In a more familiar environment with concepts such as buildings, rents, etc., there will be more available resources to see interconnection and to master task dynamics. Consequently, subjects also score better on the questionnaires related to familiar environments. The higher performance is related to the more aggressive decision making called for in the experiment.

However, a prior experiment showed that students do better than mostly MBA trained professionals in a similar setting (Bakken, et al., 1992). This suggests that context expertise, an extreme form of context familiarity, decreases performance if such expertise is at odds with the learning objectives of the experimental task. Under such re-learning exercises, professionals need to discard inappropriate schemas and build up new ones, something that requires substantial cognitive resources that detracts from the task at hand.

Consequently, the relationship between context familiarity and performance has the form of an inverted u-shape. Lack of context familiarity diverts cognitive resources to remembering names of variables. With more familiarity, more attention can be directed to the task and so performance improves. Context expertise, if not totally relevant to task, however interferes with the learning task and reduces performance.

The added complexity of low and extremely high contextual familiarity decreases current performance, but leads to a more lasting learning: Transfer performance and understanding is increased by more demanding tasks and subjects perform better in the second trial when they are exposed to an initial (unfamiliar) oil tanker environment. Similarly, questionnaire scores are improved if subjects face a demanding sequence of two different contexts rather than the same context twice.

7.2.4. Market frequency

Market frequency was the operationalization of the experiment's dynamic compression and refers to the time-varying behavior of a system. The high frequency manipulation contained an oscillation period of about 7-11 years and showed four peaks and troughs, about twice as many as low frequency markets.

The hypothesis put forward in chapter 3 stated that subjects would do better in the high frequency environment because more instances of oscillatory behavior make the system's patterns of behavior perceptually more transparent. The results showed, however, that subjects did better in a low frequency environment: Low frequency environments are "slower" and it is easier to comprehend the changes that occur from one period to the next. Slow motion is not the only factor contributing to performance in the low frequency environment: Subjects have more time to learn before they have to make important investment decisions than in high frequency environments.

Another factor contributing to poor high frequency performance is that this market contains a faster, less lenient depreciation schedule. Though this difference in depreciation schedules are also felt by the benchmark rule and should not lead to performance differences between the two frequency conditions, the human subjects are and so prone to differential risk effects while the rule is not affected by such stress. Higher depreciation drains financial resources and appears to induce more risk aversion in high frequency environments that again may cause lower performance.

Performance findings were corroborated by better questionnaire scores in the low frequency environment, thus supporting the contention that low frequency environments are more transparent. However, the frequency treatments had no lasting influence; neither transfer of understanding or performance were affected.

7.2.5 Differences in complexity implied by context and frequency

While the familiar context and low frequency conditions had similar positive effects on performance and current understanding, the transfer effects of context and frequency differed. Complexity induced by context conditions, i.e. unfamiliar initial context and context change, had positive effects on subjects' ability to transfer performance and understanding, respectively. Complexity caused by initial high frequency or a frequency change between trials gave no similar positive transfer effects.

In the context complexity case, an lack of familiarity appears to induce a more extensive cognitive search so that current performance and understanding is degraded at the benefit of a delayed positive effect on *transfer* performance. Similarly, a context change between trials results in a better mental model of the task, but questionnaire scores improve only after the experience has been made.

Mental search results when subjects do not readily find a fitting mental model of the task they encounter. Search must be vigorous if there exist a wealth of existing, yet ill fitting preconceptions of the task (as the case with the professional's irrelevant or wrong expertise). Similarly, unsuccessful search for available schemas and the subsequent creation of new concepts (as the case with unfamiliar task) also imply more extensive search and cognitive activity. Ceteris paribus, the more search, the more imprint the task will have upon cognition.

This account can explain the co-existence of poor performance and good transfer by students in unfamiliar environments. Likewise, it explains the better quality of subjects' mental models after a change in context. It does however, also predict lower performance after a changed context and higher quality of the mental model after the initial unfamiliar context. Neither of these effects were found, which weakens the contention above. As mentioned previously, however, the lack of significant findings in these two cases may be due to small sample sizes, especially in view of the large error component in the data.

The above definition of cognitive complexity does not fit the frequency treatment. In hindsight, there is no reason to believe that subjects had strong prior expectations about either market frequency. Hence, one cannot argue that there was a difference in the two frequency treatments with regards creation of new cognitive categories. Lower current performance in the high frequency environment has thus more of a perceptual than a cognitive explanation, in addition to the risk and "unfair" learning factors mentioned above.

The frequency change treatment should induce more cognitive search than a frequency sameness. Such search difference would require two concurrent conditions to hold. First, it would require that subjects tried to match the market behavior from the first task onto the second. Second, the pattern must be more readily found in the same frequency condition.

Due to the fact that first and second task had different demand streams, however, pattern matching would be unsuccessful regardless of whether the frequency changed or did not. With the benefit of hindsight, one must recognize that while cognitive searches are differently affected by context conditions, frequency conditions do not yield similar cognitive effects. Frequency effects are consequently related to risk, learning and perceptual issues. Though cognitive in nature, these issues are shallower; i.e. have lower probability of inducing cognitive search, than contextual factors.

This explanation again fits with findings in problem solving research described in chapter 2, in that cognitive searches are more readily directed along contextual than structural dimensions: Poor transfer in paradigmatic algebra word problem tasks is also caused by the high salience and importance of contextual factors.

7.3 Implications

This study has implications for theoretic work in problem solving and dynamic decision making. Likewise, the study's use of models of unstable markets as well as the account of how people operate in these markets have implications for how one can improve decision makers' performance. Finally, the findings also have implications for running learning labs. The four areas are treated separately in the four subsections below.

7.3.1. Problem solving

Research on transfer of problems solving has, among other aims, attempted to elucidate cognitive transfer mechanisms (Gholson et al., 1987; Gick and Holyoak, 1983; Hussy, 1984; Kamouri et al., 1986; Kieras and Bovair, 1986; Medin, et al., 1983; Novick, 1988). One finding has been that in order for transfer to take place, subjects need to construct or otherwise internalize an exemplar (Gentner and Tupin, 1986). The present finding of context change enhancing transfer, suggests that a sequence of two isomorphic tasks may help exemplars to develop.

An important question in the problem solving area has been how concrete a subject's experiences must be for exemplars to be formed. Can general inference rules (Ploger and Wilson, 1991) serve as exemplars? Or must exemplars be more domain specific (Proctor, 1988), i.e. concrete? The answer to this question has consequences for whether one should teach rules of inference (Nisbett et al., 1987) and hope that students themselves form exemplars, or whether one should repeat isomorphic problems and their solutions in a multitude of contextual disguises.

If transfer is achieved with fairly abstract exemplars, then subjects with training in abstract ways of handling markets would transfer better. The present study indicated that transfer is not helped by such training and suggests that exemplars are not formed by abstract models. In addition, the fact that the experience itself induce better understanding further supports the claim that changes in contexts appear to help the formation of exemplars, suggests that they must be fairly concrete.

Though exemplars provide frameworks that enable subjects to use prior knowledge during problem solving, they do not guarantee seamless transfer. As shown by the professionals' problematic performance, exemplars may involve a whole schema full of unimportant and sometimes wrong information. Consistent with prior findings of concept formation (Wattenmaker et al., 1986) and cognitive strategies (Klayman and Ha, 1988), the findings here suggest that subjects have poor heuristics for improving on their well-developed schemas. By confirmatory inference strategies, subjects seek evidence to support the existing schema instead of systematically probing into inconsistencies with new findings.

The finding that coursework in several disciplines helps subjects' initial framing of the task, yet fail to help transfer, suggests that transfer mechanisms are complex and that interdomain transfer as defined by algebra word problems are highly problematic: The fact that subjects are able to transfer to an isomorphic laboratory task does not imply that transfer helps them in the more complicated, partly recurring, tasks of realistic decision environments.

7.3.2. Dynamic Decision Making

The tasks investigated in the previous chapters share several features with prior dynamic decision research, yet they also depart from and expand the paradigm. Dynamic decision researchers often investigate decision making with a control task paradigm (Sheridan, 1974; Sterman, 1989a; Diehl, 1992; Paich, 1993). In control tasks, feedback delays tend to reduce task transparency and degrades control quality. The chief consequence of poor transparency of subject's own control and inherent feedback mechanisms has been poor performance (Diehl, 1992) and slower return to equilibrium (Kampmann, 1992) compared to benchmarks. Pattern matching appears to be the mode of learning, as opposed to the more robust strategy of learning underlying causal mechanisms (Paich and Sterman, 1992).

The impact from subject decision making on market behavior was negligible in the present experiments, however, and points to that the experiments reported here departs from the control paradigm. However, the task was still an action feedback task in that subjects' immediate environment, i.e. decision makers' cash and physical assets, were influenced by the combined effects of prior investments and sales. Environments differed in the liquidity cushion provided, and the results indicate that such cushions influenced subjects' decision strategy.

In the tasks, there was a relationship between structural parameters, and thus system frequency, and liquidity cushions. Low cushions, induced by high frequency markets, lead to risk-averse behavior and this overreaction helps explain why subjects in "hard" markets actually go bankrupt less often than in "easy" situations.

In other words, structural parameters that have consequences for performance in dynamic tasks may have their effect less from the control parameters themselves, than from side-effects these parameters have for subject risk-taking. In prior dynamic decision work, where potential profits have varied as a function of treatment conditions, the fact that structural parameters may have introduced differences in risk-taking has not been investigated thoroughly.

The present experiments have in common with control task findings that feedback delays influence market dynamics. Counter to control tasks, however, where the added complexity of long acquisition lags degrade performance, it has been shown that long delays may lead to slower unfolding of problematic dynamics that again may improve performance.

Moreover, it was found that task context matters in performance. This supports (Hussy, 1984) who found that girls do better in isomorphic "kitchen" tasks and boys better in "moon landing" control tasks, presumably because the youngsters where more familiar with the respective domains. Furthermore, the present study showed that feedback and context dimensions interact so as to cause significantly better performance when both the context and the frequency are "lenient": Many of the findings of poor results in the dynamic decision literature may have looked better if, in addition to improving feedback transparency, one had also improved context familiarity. Research using generically formulated tasks is particular prey to this criticism. Anyone having played e.g. the "long wave game" (Sterman, 1985) will testify that the lack of task context induces complexity in its own right that is quite problematic.

The poor transfer in this study was indicated by the fact that low contextual familiarity and high frequency hinder performance also in the second trial. This finding supports prior decision research in that performance remains poor after several trials (Paich and Sterman, 1992) and that learning fails to address structural issues Decision heuristics lack robustness with regards to slight changes in parameter values. Subjects are equally prey to the high frequency environment the second as the first trial, something that supports Paich and Sterman's finding of a "ballistic" pattern-matching heuristic that lacks robustness.

In all, even though the present approach investigated decision making in a system less governed by subjects' action feedback, several major findings in Dynamic Decision Making, such as the failure to perform well in complex environments, the failure to learn, as well as failure to develop robust heuristics have been corroborated. However, the study has also indicated that research should seek higher context precision. Similarly, the findings indicate that the impact of subjects' decision heuristics and system parameters on how systems unfold over time, create different opportunities for learning. These interactions should also be treated with more caution in future research.

7.3.3. Dynamically complex markets

Market instabilities are undesirable. This is what motivated the choice of learning environments in this thesis, and is moreover the contention for much work in macro-economics (see e.g. Blanchard and Fischer (1989) chapter 5). One way to counter instabilities is to let governments control the supply of credit (Johansen, 1983). If, government resource allocation is more effective than that of a market, such control may be an effective way to reduce instabilities. However, government control is becoming increasingly harder to enforce (Tranøy, 1993). In addition, such controls often have devastating consequences for the efficiency of resource allocation. As an example, centrally regulated communist economies tend to gain economic stability by controlling credit use. The inefficiency cost has been devastating.

As mentioned in the previous chapters, learning from experience is problematic as long as decision makers are unable to extract or integrate inherent causal (and uncertainty) relations. Instabilities may indeed perpetuate if decision frequencies are much shorter than the market dynamics: Decision makers fail to become aware of longer term dynamics. The interview data indeed suggest that there exist forward and backward decision myopia (Wheaton and Torto, 1988). However, market instabilities can only persist if subjects fail to understand important underlying causal mechanisms. Conversely, the myopic sources of instability (there also exist others, with different policy recommendations, see e.g. Blanchard and Fischer (1989) will disappear if decision makers obtain insight into underlying forces or if they are able to use robust rules based on the dynamics of such forces.

The benchmark decision rule may help create stability in the markets for commodity-like assets with long life times. It is based on the structure of the markets, on agents' bounded rationality and works by selling assets before they peak and buying assets before their values go up. Consequently the rule is stabilizing. If people could internalize the inherently benign buy-low-sell-high rule, unstable markets would be stabilized. Unfortunately, the study has shown that people are poor in internalizing such a rule; one cannot guarantee high performance by simply exposing subjects to the game. However, the rule could itself well be taught and so improve the behavior of unstable markets.

Similarly, public and industry executives that have obtained expertise in a stable economic environment with open loop control, i.e. with weak feedback signals, will have difficulties dealing with a more complex market feedback environment. It has been shown that such markets have tendencies towards instability, something that further decreases opportunities for feedback interpretation. Learning labs may be the only way to achieve experiential learning in such environments.

Learning labs designed to increase insight are longer and more involved than the experimental markets portrayed here, and may indeed help subjects internalize transferable learning of particular interest in dynamically complex markets. On the other hand, subjects with System Dynamics background, with some exposure to management flight simulators, did not perform better in the unstable markets. Moreover, the present study certainly paints a sobering picture of the potential pitfalls inherent in using simulated environments. The present study notwithstanding, the costs are small and the potentials great as will be explained in the next section.

7.3.4. Learning in simulated environments

The research reported here was partially motivated by an urge to understand the learning that takes place in a learning lab. The results show a complex picture of underlying mechanisms. Subjects' laboratory performance can be manipulated by task context and frequency manipulations. Extending the findings to areas with even lower context familiarity suggests that in such circumstances, learning may perhaps not take place at all. Consequently, such environments will probably not lead to transfer, either.

Poor transfer from unguided game playing, even when augmented by questionnaires that must have induced reflection, must make game designers cautious: The same environments that cause poor performance in the first trial also cause poor performance in the second, and suggest that running learning labs . with two games "back-to-back" yields limited educational value.

Increased transfer, appear to stem from the challenge inherent in the task's contextual disguise. Both the less known oil tanker environment as well as a context change stimulated cognitive processes so as to enhance transfer.

Certainly, there exist other ways to stimulate cognitive activity. Reflective exercises are commonly used (Kim, 1989), and the findings here suggest that if discussions increase cognitive activity, then such exercises might help subjects develop more robust and transferable exemplars, often called system's archetypes or generic structures (Senge, 1990). Yet, discussions and other reflective exercises tend to be verbal and conceptual, and must be of limited value unless the real decision making environment also becomes reflective.

As mentioned, decision environments that are moderately remote from decision reality appears less threatening to subjects. Such environments are conducive to learning since they at the same time induce more risk-loving behavior in the lab and because less threat helps subjects be more inquisitive (Argyris and Schön, 1978). The findings here support that claim and have underlined that both system structure and context may influence learning through the perceived task risk.

The findings also suggest that low stress environments have the side-effect of being less conducive to transfer of understanding. Though the present data have not thoroughly investigated the trade-off between ease of performance and transfer effect, it is probable that a smooth training session where professionals discuss some contextually remote issue will have little profound consequences on decision heuristics: Though a well-known environment where subjects are experts hinders learning in the lab, it may be that the little learning that does take place will be more readily used in the workplace.

If a learning lab and a mental schema differ, a change in the established web of causal webs is hard to establish. A change in heuristics, however, is easier to accomplish in a brief lab setting. Subjects are more likely to remember heuristics like "in dynamic systems where the structure lacks transparency, proceed with caution" than to internalize the system structure or establish a procedure to look for system structure. Consequently, the limits of such decision rules should be made clear to learning lab participants, or they may make negative transfer from the workshop to the workplace.

By providing more examples of the stock-adjustment problem in various contextual disguises, it may be that the unstable stock-adjustment system, its behavior and consequences for profitable decision making can form an exemplar that helps subjects recognize such systems when they see them in the real world. The fact that context change helps understanding indicates such an effect. Yet, the lack of positive transfer effect of the background variables shows that abstract frameworks and formal reasoning schema will not by themselves enable decision makers to make better decisions (Nisbett et al., 1987).

7.4 Future research

The use of simulated decision environments as virtual worlds for exploration of issues that do not lend themselves well to real world experimentation is steadily increasing. There exists little research about the learning effects of such microworlds. This is unfortunate, especially considering that computer environments enable unobtrusive monitoring of decision making and in that sense are ideal research environment. The present study is an example of how such research may be carried out. It is especially troublesome that not only commercial, but also research, environments create learning labs without a more careful monitoring the learning process. Hypotheses of learning that may take place in such environments abound, but research testing those propositions are few (see e.g. Gould, 1993).

As mentioned, it may be that reflective exercises will help make participants more "mindful" to use Solomon's (1990) expression. The degree to which such mindfulness will increase performance in a real environment is open to question, however. The counter argument is that people in most environments, especially experts, base their decision making on heuristics. The degree to which such heuristics can be modified by pure reflection needs to be addressed.

The experiments reported herein should themselves be refined on a number of dimensions. With respect to the frequency treatment, it appears that the treatment confounds at least three effects that may be investigated separately: Fast dynamics let the subjects' see more instances of the problematic behavior and so may lead to better learning. Yet, in the experiments, this effect was confounded with burdensome depreciation schemes, which appeared to reduce riskiness. Similarly, long asset life times lead to a system with a longer "memory" that make environments harder to control (Diehl, 1992), yet long life times also contribute to slower dynamic unfolding that appears to improve performance. Similarly, the research design could be extended to include a third task as already suggested in chapter 6, so as to investigate further transfer effects.

The small sample size has been mentioned several times. In conjunction with other refinements, one should also try to replicate the findings. If they are replicated, then one should pool the results so as to be able to state findings, especially the transfer claims, with more confidence.

It may be hard to generalize the findings of familiar contexts being more conducive to performance. A first step towards validating the findings will be to use the research design where the two markets are equally familiar to subjects. Graduate students in shipping economics programs will have more familiarity with the tanker context, and an extension of the present research is currently under way to replicate these findings in such an environment.

As this thesis was put to print, an old oil tanker whose engines stopped outside the Shetland Islands was crushed and left tens of million gallons of crude oil to the unsuspecting birds. This indicates that the nature of these markets have important societal consequences. More research into the dynamics of complex markets is called for, too.

References

Aftenposten (1992, Dec. 1). Bankkrisen vil ingen ende ta. Aftenposten, p. 33.

Akerlof, G. A. and Dickens, W.T. (1982). "The economic consequences of cognitive dissonance." *American Economic Review.* **72**: 307-19.

Allen, J.A., Hays, R.T., Buffardi, L.C. (1986). "Maintenance Flight Simulator Fidelity and Individual Differences in Transfer of Training." *Human Factors.* **28**(5): 497-509.

Andreassen, P.B. (1991). "Causal Prediction versus Extrapolation: Effects of Information Source on Judgmental Forecasting Accuracy." mimeo, MIT, Cambridge, MA.

Andreassen, P. B. (1988). "Explaining the price-volume relationship: The difference between price changes and changing prices."*Organizational Behavior and Human Decision Processes,* 41, 371-89.

Andreassen, P.B. (1990). "Judgmental Extrapolation and Market Overreaction: On the Use and Disuse of News." *Journal of Behavioral Decision Making.* **3**: 153-174.

Andreassen, P.B. and Kraus, S.J. (1990). "Judgmental Extrapolation and the Salience of Change." *Journal of Finance.* **9**: 347-372.

Argyris, C. (1991). "Teaching smart people how to learn." Harvard Business Review (May-June): 99-109.

Argyris, C. and Schön, D. (1978). *Organizational learning: A theory of action perspective.* Addison-Wesley.

Arkes, H.R. and Blumer, C. (1985). "The psychology of sunk cost." *Organizational Behavior and Human Decision Processes.* **35**: 124-40.

Ashford, S.J. and Tsui, A.S. (1991). "Self-regulation for managerial effectiveness: The role of active feedback seeking." **34** (2): 251-280.

Bailey, D. (1991).*A warning for New England Banks.* Boston Globe. Boston, MA: January

Bakken, B. E., Gould-Kreutzer, J. M., and Kim, D. H. (1992). Management Flight Simulators and Organizational Learning: Some experimental evidence. *European Journal of Operational Research, 59, 167-182.*

Bakken, B.E. (1989a). "High and Low Performers: Differences in Information Processing Strategies and Learning." Working paper no. D-4047, MIT, Cambridge.

Bakken, B.E. (1989b). "Learning in dynamic simulation games: Using performance as a measure." In Millin, P.M. and Zahn, E.O.K., ed. *Computer-based management of complex systems: Proceedings of the 1989 International Conference of the System Dynamics Society.* Stuttgart, W. Germany: Springer-Verlag.

Bakken, B.E. (1990a). "An Interview With a Thoughtful Shipping Investor." System Dynamics Group Working Paper, MIT, Cambridge, MA.

Bakken, B.E. (1990b). "Learning and transfer by students and professionals in cyclical markets." Working paper no. D-4166, System Dynamics Group, M.I.T., Cambridge, MA.

Bar-Hillel, M. (1980). "The base-rate fallacy in probability judgments." *Acta Psychologica.* **44**: 211-33.

Bassok, M. (1990). "Transfer of Domain-Specific Problem-Solving Procedures." *Journal of Experimental Psychlogy; Learning, Memory and Cognition* **16**(No 3): 522-533.

Bassok, M. and Holyoak., K.J. (1989). "Interdomain Transfer Between Isomorphic Topics in Algebra and Physics." *JEP; L, M and C* **15**(No 1): 153-166.

Berry, D.C., Broadbent, D.E. (1984). "On the Relationship between Task Performance and Associated Verbalizable Knowledge." *Quarterly Journal of Experimental Psychology.* **36A**: 209-231.

Blanchard, J. B., and Fischer, S. (1989). "Lectures on Macroeconomics." Cambridge, MA: MIT Press.

Bottom, W.P., Bontempo, R.N. and Holtgrave, D.R. (1989). "Experts, novices, and the St. Petersburg Paradox: Is one solution enough?" *Journal of Behavioral Decision Making.* **2**.

Brehmer, B. (1980). "In One Word: Not From Experience." *Acta Psychologica.* **45**: 233-41.

Brehmer, B. (1987). "Systems Design and the Psychology of Complex Systems." In Rasmussen, J. and Zunde, P., ed. *Empirical Foundations of Information and Software Science.* Plenum Publ. Corp.

Brehmer, B. (1988).*Strategies in real time dynamic decision making.* University of Chicago, Chicago IL: Einhorn Memorial Conference.

Brehmer, B. and Kuylenstierna, J. (1980). "Content and Consistency in Probabilistic Inference Tasks." *Organizational Behavior and Human Performance.* **26**: 54-64.

Broadbent, D. (1987). "Structures and Strategies: Where are we now ?" *Psychological Research.* **49**: 73-79.

Broadbent, D. and Aston, B. (1978). "Human Control of a Simulated Economic System." *Ergonomics.* **21**(12): 1035-43.

Broadbent, D.E., Fitzgerald, P., Broadbent, M.H.P. (1986). "Implicit and Explicit Knowledge in the Control of Complex Systems." *British Journal of Psychology.* **77**: 33-50.

Cheng, P. and Nisbett, R. (1985). "Pragmatic Reasoning Schemas." *Cognitive Psychology.* **17**: 391-416.

Clairborne, W. and Day, K. (1992). "Reichmann Aura: Fading Fast". *International Herald Tribune* April 27: 9-10.

Cohen, J. (1982). "Are People Programmed to commit fallacies? Further thoughts about the interpretation of experimental data on probability judgment." *Journal of the Theory of Social Behavior.* **12**: 251-274.

Cohen, J. (1990). "Things I Have Learned (So Far)." *American Psychologist.* **45**(12): 1304-1312.

Cooksey, R. W., Freebody, P. (1986). "Social Judgment Theory and Cognitive Feedback: A General Model for Analyzing Educational Policies and Decisions. "*Educational Evaluation and Policy Analysis,* 8(No 1), 17-29.

Corcoran, E. and Wallich, P. (1991). "The Analytical Economist: Bursting bubbles." (May): 88.

Cyert, R. and March, J. (1963). *A Behavioral Theory of the Firm.* Prentice Hall.

Davis, H. L., and Hogarth, R. M. (1992), "Rethinking Management Education: A View from Chicago. "*Selected papers of the Graduate School of Business at The University of Chiocago,* No 72.

Dawes, R. M. (1979). "The robust beauty of improper linear models in decision making." American Psychologist, 34, 571-82.

Dawes, R.M. and Corrigan, B. (1974). "Linear models of decision making." *Psychological Bulletin.* **81**: 95-106.

de Geus, A.P. (1988). "Planning as Learning." *Harvard Business Review.* **66**(2): 70-74.

Diehl, E. (1988). "Learning through participatory simulations--A study based on the market growth model." System Dynamics Group, Sloan School of Management, MIT, Cambridge, MA.

Diehl, E.W. (1989). "A study of human control in stock-adjustment tasks." In Milling, P.M. and Zahn, E.O.K., ed. *Computer-based management of complex systems: Proceedings of the 1989 International Conference of the System Dynamics Society.* Stuttgart, W. Germany: Springer-Verlag.

Diehl, E.W. (1992). "Effects of Feedback structure on Dynamic Decision Making" Ph.D Dissertation, Sloan School of Management, Massachusetts Institute of Technology

Dörner, D. (1980). "On the difficulties people have in dealing with complexity." *Simulation and Games.* **11**(1): 87-106.

Dörner, D. (1989). "The Logic of Failure." Lehrstuhl Psychologie II, Universität Bamberg, Bamberg, Germany.

Drewry Shipping Consultants, L. (1990). *Tanker Spot Rate Trends.* Shipping Statistics and Economics, # 236. June. London:

Edwards, W. (1962). "Dynamic decision theory and probabilistic information processing." *Human Factors.* **4**: 59-73.

Einhorn, H. J., Kleinmuntz, D., and Kleinmuntz, B. (1979). Linear Regression and Process-Tracing Models of Judgment. Psychological Review, 56(5), 465-85.

Einhorn, H.J. (1982). "Learning from experience and suboptimal rules in decision making." In Kahneman, D., et al., ed. *Judgment under uncertainty: Heuristics and biases.* Cambridge Univ. Press.

Einhorn, H.J. and Hogarth, R.M. (1978). "Confidence in Judgment: Persistence of the Illusion of Validity." *Psychological Review.* **85**(5): 395-416.

Einhorn, H.J., Kleinmuntz, D. and Kleinmuntz, B. (1979). "Linear Regression and Process-Tracing Models of Judgment." *Psychological Review.* **56**(5): 465-85.

Ericsson, K.A. and Simon, H.A. (1980). "Verbal reports as data." *Psychological Review.* **87**: 215-51.

Ericsson, K.A. and Simon, H.A. (1984). *Protocol Analysis: Verbal Reports as Data.* Cambridge, Mass: MIT Press.

Estes, W.K. (1986). "Array Models for Category Learning." *Cognitive Psychology.* **18**: 500-549.

Estes, W.K. (1988). "Human learning and memory." In Stevens, ed. *Handbook of experimental psychology.*

Fearnleys (1983-1988). *Fearnleys Review.* Oslo, Norway: Fearnresearch.

Fischoff, B. (1975). "Hindsight ≠ Foresight: The Effect of Outcome Knowledge on Judgment Under Uncertainty." *Journal of Experimental Psychology.* **1**(3): 188-99.

Fong, G.T. and Nisbett, R.E. (1991). "Immediate and Delayed Transfer of Training Effects in Statistical Reasoning." *Journal of Experimental Psychology.* **120**(1): 34-45.

Fong, G.T., Krantz, D.H., Nisbett, R.E. (1986). "The Effects of Statistical Training on Thinking about Everyday Problems." *Cognitive Psychology.* **18**: 253-292.

Forrester, J. W. (1970). "Counterintuitive Behavior of Social Systems." In *Collected Papers of Jay W. Forrester* Wright-Allen Press.

Forrester, J.W. (1961). *Industrial Dynamics.* Cambridge: MIT Press.

Forrester, J.W. (1968). "Market growth as influenced by capital investment." *Industrial Management Review.* **9**(2): 83-106.

Franklin, G.F., Powell, J.D. and Emani-Naeini, A. (1986). *Feedback Control of Dynamic Systems.* Reading, MA: Addison-Wesley.

Friedland, B. (1986). *Control System Design.* New York: McGraw-Hill.

Fuglseth, A.M. (1989). "Beslutningsstøtte: metode for diagnose av lederes informasjons- og situasjonsoppfatninger. (Decision Support; method for diagnosing managers´ assessments of information and situations)." Ph D Thesis, Norwegian School of Economics and Business Administration, Bergen.

Funke, J. (1991). "Solving Complex Problems: Exploration and Control of Complex Systems." In French, R.J.., ed. *Complex Problem Solving: Principles and Mechanisms.* Hillsdale, NJ: Erlbaum.

Gentner, D. and Toupin, C. (1986). "Systematicity and Surface Similarity in the Development of Analogy." *Cognitive Science.* **10**: 277-300.

Gholson, B.,Eymard, L. A., Morgan, D., and Kamhi, A. G. (1987). "Problem-Solving, Recall, and Isomorphic Transfer among Third-Grade and Sixth-Grade Children." *Journal of Experimental Child Psychology*, 43, 227-243.

Gick, M.L. and Holyoak, K.J. (1980). "Analogical Problem Solving." *Cognitive Psychology.* **12**: 306-355.

Gick, M.L. and Holyoak, K.J. (1983). "Schema Induction and Analogical Transfer." *Cognitive Psychology.* **15**: 1-38.

Gilden, D.L. and Proffitt, D.R. (1989). "Understanding Collision Dynamics." *Journal of Experimental Psychology.* **15**(No 2): 372-383.

Gladstein, D. L., and Reilly, N. P. (1985). "Group Decision Making Under Threat: The Tycoon Game." Academy of Management Review, 28(No 3), 613-627.

Gladstein, D.L. and Reilly, N.P. (1985). "Group Decision Making Under Threat: The Tycoon Game." *Academy of Management Review.* **28**(No 3): 613-627.

Globe, The Boston , Tuesday, January 16, 1990, "New England Banks Face Heavy Scrutiny," Pg. 1

Golucke, U. (1987). "The Ship Investment Calculator." Future Mappers, Oslo.

Gould, J. (Ed.). (1993). "Special issue of the System Dynamics Review: Pre-college System Dynamics education." System Dynamics Review, In Press. John Wiley et sons.

Graham, A.K., Senge, P.M., Sterman, J.D. and Morecroft, J.D.W. (1992). "Computer-based case studies in management education and research." Working paper no. D-4039, System Dynamics Group, MIT, Cambridge, MA.

Greller, M.M. (1980). "Evaluation of Feedback Sources as a Function of Role and Organizational Level." *Journal of Applied Psychology.* **65**(No 1): 24-27.

Hammond, K. R., and Summers, D. A. (1972). "Cognitive Control." *Psychological Review* 79(No 1), 58-67.

Hammond, K.R. (1978). "Toward Increasing Competence of Thought in Public Policy Formation." In Hammond, K.R., ed. *Judgement and Decision Making in Public Policy.* Denver: Westview.

Hammond, K.R., Summers, D.A. and Deane, D.H. (1973). "Negative effects of outcome-feedback in Multiple-cue probability learning." *Organizational Behavior and Human Performance.* **9**: 30-34.

Hayes, J.R. and Simon, H.A. (1974). "Understanding Written Problem Instructions." In Gregg, L.W., ed. *Knowledge and Cognition.* Potomac, MD: Lawrence Erlbaum Ass.

Henderson, R. (1987). "The Failure of Established Firms in the Face of Radical Change." Draft paper, Harvard University, Cambridge, MA.

Hernandez, K (1991). "Learning in real estate: The role of the development system in creating oversupply". MS thesis, Sloan School of Management. Massachusetts Institute of Technology.

Hoch, S.J. and Loewenstein, G.F. (1989). "Outcome Feedback: Hindsight and Information." *Journal of Experimental Psychology.* **15**(No 4): 605-619.

Hockey, R., and Hamilton, P. (1983). "The cogntive pattern of stress states. "In R. Hockey (Eds.), *Stress and fatigue in human performance* (pp. 331-362). New York: Wiley.

Hogarth, R.M. (1981). "Beyond Discrete Biases: Functional and Dysfunctional Aspects of Judgmental Heuristics." *Psychological Bulletin.* **90**(2): 197-217.

Hogarth, R.M. and Reder, M.W. (1987). *Rational Choice: The Contrast between Economics and Psychology.* Chicago: Univ. of Chicago Press.

Hogarth, R.M., Gibbs, B.J., McKenzie, C.R.M. and Marquic, M.A. (1989). "Learning to make decisions: When incentives help and hinder." Center for Decision Research, University of Chicago, Chicago, IL.

Holyoak, K.J. (1990). "Problem solving." In Osherson, D.N. and Smith, E.E., ed. *An Invitation to Cognitive Science. Volume III: Thinking.* Cambridge, MA: MIT Press.

Hoyt, H. (1934). *100 years of Chicago Land Values.* Chicago: University of Chicago Press.

Huber, O. (1986). "Mental representation in multistage decision making. "In R. Tietz,W. Albers, and R. Selten (Eds.), *Bounded rational behavior in experimental games and markets: Proceedings of the fourth conference on experimental economics* (pp. 66-81). Bielefeld, W. Germany: Springer-Verlag.

Huber, O., Debeutz, A., Pratscher, J. and Quehenberger, I. (1992). "Perceived Control in a Multistage Decision Task." *Journal of Behavioral Decision Making.*

Hussy, J.F.a.W. (1984). "Komplexes Problemlösen: Beiträge zu zeiner Erfassung sowie zur Frage der Bereichs- und Erfahrungsabhängigkeit." *Zeitschrift für experimentelle und angewandte Psychologie* **XXXI**(1): 19-38.

Hylton, R.D. (1991).*Trammell Crow Plans Revamping.* New York Times. New York:

Isaacs, W., and Senge, P. (1992). Overcoming limits to learning in computer-based learning environments. *European Journal of Operational research*, 59, 183-196.

Ivanchevich, J.M. and McMahon, J.T. (1982). "The Effects of Goal Setting, External Feedback, and Self-Generated Feedback on Outcome Variables: A Field Experiment." *Academy of Management Review.* **25**(No 2): 359-372.

Jacoby, J., Jaccard, J., Kuss, A., Troutman, T. and Mazursky, D. (1987). "New Directions in Behavioral Process Research: Implications for Social Psychology." *Journal of Experimental Social Psychology.* **23**: 146-175.

Jacoby, J., Kuss, A., Troutman, T. and Mazursky, D. (1984). "When Feedback is Ignored: Disutility of Outcome Feedback." *Journal of Applied Psychology.* **69**(No 3): 531-545.

Jones, E.E., Kanouse, D. E., Kelley, H. H., Nisbett, R. E., Valins, S., and Weiner, B. (1987). *Attribution. --Perceiving the causes of behavior*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Kahneman, D. and Tversky, A. (1974). "Judgment under uncertainty: Heuristics and biases." *Science*. **185**: 1124-31.

Kahneman, D., Slovic, P. and Tversky, A., ed. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge Univ. Press.

Kamouri, A.L., Kamouri, J. and Smith, K.H. (1986). "Training by exploration: facilitating the transfer of procedural knowledge through analogical reasoning." *International Journal of Man-Machine Studies*. **24**: 171-192.

Kampmann, C.E. (1990). "Feedback complexity, market dynamics, and performance: Some experimental results." Working paper no. D-4193, System Dynamics Group, M.I.T., Cambridge, MA.

Kampmann, C.E. (1992). "Feedback complexity and Market Adjustment in Experimental Economics". Ph.D dissertation, Sloan School of Management, Massachusetts Institute of Technology.

Kieras, D. E., and Bovair, S. (1986). "The Acquisition of Procedures from Text: A Production System Analysis of Transfer of Training." *Journal of Memory and Language*, 25, 507-524.

Kim, D.H. (1989). "Learning laboratories: Designing a reflective learning environment." Working paper no. D-4026, System Dynamics Group, MIT, Cambridge, MA.

Kindleberger, C. P. (1988). The financial crises of the 1930's and the 1980s: Similarities and differences. *Kyklos, 41*(Fasc. 2), 171-186.

Klayman, J. and Ha, Y.W. (1987). "Confirmation, disconfirmation, and information in hypothesis testing." *Psychological Review*. **94**: 211-28.

Klayman, J. and Ha, Y.W. (1989). "Hypothesis testing in Rule Discovery: Strategy, Structure and Content." *Journal of Experimental Psychology*. **15**(No 4): 595-604.

Kleinmuntz, D. (1985). "Cognitive Heuristics and feedback in a dynamic decision environment." *Management Science*. **31**: 680-702.

Kleinmuntz, D. and Schkade, D.S. (1988). "The cognitive implications of information displays in computer-supported decision making." Working paper no. 2010-88, Sloan School of Management, MIT, Cambridge, MA.

Kleinmuntz, D.N. (1990). "Information processing and the misperception of feedback in dynamic decision making." University of Illinois at Urbana-Champaign, Champaign, IL.

Kleinmuntz, D.N. and Thomas, J.B. (1987). "The Value of Action and Inference in Dynamic Decision Making." *Organizational Behavior and Human Decision Processes*. **39**: 341-64.

Kluwe, R.H., Haider, H. and Misiak, C. (1990). "Learning by doing in the control of a complex system." In Mandl, d.C., Bennett and Frierich, ed. *Learning and instruction in an international context*. Oxford: Pergamon Press.

Kolodner, J.L. (1985). "Experiential Processes in Natural Problem Solving." Georgia Institute of Technology, Atlanta, GA.

Kolodner, J.L. and Kolodner, R.M. (1985). "Using Experience in Clinical Problem Solving." Georgia Institute of Technology,

Kotovsky, K., Hayes, J.R. and Simon, H.A. (1985). "Why are some problems hard ? Evidence from Tower of Hanoi." *Cognitive Psychology*. **17**: 248-294.

Kreutzer, W. B.,Kreutzer, D. P., and Gould, J. M. (1992). "The Quahog Oil Production Simulator." In J. e. a. Vennix (Ed.), *The International System Dynamics Conference*, (pp. 329-336). Utrecht, The Netherlands:

Kunda, Z. and Nisbett, R.E. (1986). "The Psychometrics of Everyday Life." *Cognitive Psychology*. **18**: 195-224.

Larrick, R.P., Morgan, J.N. and Nisbett, R.E. (1989). "Who uses the normative rules of choice ?" University of Michigan, Ann Arbor, MI.

Laughunn, D.J. and Payne, J.W. (1984). "The impact of sunk outcomes on risky choice behavior." *Canadian Journal of Operations Research and Information Processing*. **22**: 151-81.

Laurent, G. (1970) "A Dynamic Analysis of the Housing Market in Paris." MS thesis , Sloan School of Management, MIT.

Lichtenstein, S., Fischoff, B. and Phillips, L. (1977). "Calibration of probabilities: The state of the art." In Jungerman, H. and deZeeuw, G., ed. *Decision making and change in human affairs*. Amsterdam: Reidel.

Luce, R.D. and Raiffa, H. (1957). *Games and Decisions*. New York: John Wiley and Sons.

Mackinnon, A.J. and Wearing, A.J. (1980). "Complexity and Decision Making." *Behavioral Science*. **25**: 285-296.

Mackinnon, A.J. and Wearing, A.J. (1985). "Systems Analysis and Dynamic Decision Making." *Acta Psychologica*. **58**: 159-72.

Maidment and Drummond (1989). "Tanker profitability." Masters thesis written for Shell Int'l., London Business School, London.

March, J.G. (1978). "Bounded rationality, ambiguity, and the engineering of choice." *Bell Journal of Economics*. **9**: 87-108.

March, J.G. and Olsen, J.P. (1974). *Ambiguity and choice in organizations*. Bergen: Universitetsforlaget.

Marcus, H.S., Glucksman, M.A., Ziogas, B.O. and Meyer, K.L. (1991). "The Buy-Low, Sell-High Investment Methodology: The Case of Bulk Shipping." **21**(2): 8-21.

Meadows, D.J. (1969). "The Dynamics of Commodity Production Cycles: A Dynamic Cobweb Theorem." Ph.D. thesis, Sloan School of Management, MIT, Cambridge, MA.

Medin, D.L., Dewey, G.I. and Murphy, T.D. (1983). "Relationships Between Item and Category Learning: Evidence That Abstraction is not Automatic." *Journal of Experimental Psychology.* **9**(No 4): 607-625.

Melody, T and Wagley, J (1989) "Managing for Survival: Real Estate Development in Texas 1980-1989," mimeo, MIT Center for Real Estate development,

Metcalfe, J. (1986). "Feeling of knowing in memory and problem solving." *Journal of Experimental Psychology.* **12**: 288-294.

Moissis, A.A. (1989). "Decision making in the insurance industry: A dynamic simulation model and experimental results." M.Sc. thesis, Sloan School of Management, MIT, Cambridge, MA.

Montgomery, H. and Svenson, O., ed. (1989). *Process and Structure in Human Decision Making.* New York: Wiley.

Morecroft, J.D.W. (1983). "System Dynamics: Portraying Bounded Rationality." *Omega.* **11**(2): 131-42.

Morecroft, J.D.W. (1985). "The feedback view of business policy and strategy." *System Dynamics Review.* **1**(No 1): 4-21

Morris, N.M. and Rouse, W.B. (1985). "The Effects of Type of Knowledge upon Human Problem Solving in a Process Control Task." *IEEE Transactions on Systems, Man and Cybernetics.* **SMC-15**(No 6): 698-706.

Munthe, P. (1992). *Bankkrisen. Årsaker og konsekvenser* NOU report. University of Oslo, Economics Department.

Nadler, D. (1979). "The Effects of Feedback on Task Group Behavior: A Review of the Experimental Research." *Organizational Behavior and Human Performance.* **23**: 309-338.

Neale, M.A. and Northcraft, G.B. (1988). "Experience, Expertise, and Decision Bias in Negotiation: The Role of Strategic Conceptualization." In Sheppard, et al., ed. *Research on Negotiating in Organizations.* Greenwich, CT: JAI Press.

Newell, A. and Simon, H.A. (1972). *Human problem solving.* Englewood Cliffs, N.J.: Prentice-Hall.

Nisbett, R., Fong, G.T., Lehman, D.R. and Cheng, P.W. (1987). "Teaching Reasoning." *Science.* **238**: 625-31.

Nisbett, R.E. and Wilson, T.D. (1977). "Telling more than we can know: Verbal reports on mental processes." *Psychological Review.* **84**: 231-259.

Northcraft, G. B., and Neale, M. A. (1986). "Opportunity costs and the framing of decisions." *Organizational Behavior and Human Dceision Processes*, 37, 348-356.

Northcraft, G.B. and Earley, P.C. (1989). "Technology, Credibility and Feedback Use." *Organizational Behavior and Human Decision Processes*. **44**: 83-96.

Novick, L.R. (1988). "Analogical Transfer, Problem Similarity, and Expertise." *Journal of Experimental Psychology*. **14**(No 3): 510-520.

Novick, L.R. and Holoyoak, K.J. (1991). "Mathematical Problem Solving by Analogy." **17**(3): 398-415.

Özveren, C.M. and Sterman, J.D. (1989). "Control theory heuristics for improving the behavior of economic models." *System Dynamics Review*. **5**(No 2): 130-147.

Paich, M (1993). "Boom, bust and failures to learn in experimental markets". Ph.D Dissertation, Sloan School of Management, Massachusetts Institute of Technology.

Payne, J. W. (1976). "Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis." *Organizational Behavior and Human Performance*, 16, 366-87.

Payne, J., Braunstein, M.L. and Carroll, J.S. (1978). "Exploring predecisional behavior: An alternative approach to decision research." *Organizational Behavior and Human Performance*. **22**: 17-44.

Ploger, D. and Wilson, M. (1991). "Statistical Reasoning: What is the role of Inferential Rule Training? Comment on Fong and Nisbett." *Journal of Experimental Psychlogy: General* **120**(2): 213-214.

Porter, M. (1980). *Competitive Strategy*. Cambridge, MA: Free Press.

Porter, M.E. (1990). "The Competitive Advantage of Nations." Harvard Business Review, (March-April): 73-93.

Potter, M.C. (1990). "Remembering." In Osherson, D.N. and Smith, E.E., ed. *An Invitation to Cognitive Science. Volume III: Thinking*.

Proctor, R. W. (1988). The Acquisition of Task-Specific Productions and Modifications of Declarative Representations in Spatial-Precuing Tasks. *Journal of Experimental Psychology*, 117(No 2), 182-196.

Proffitt, D.R. and Gilden, D.L. (1989). "Understanding Natural Dynamics." *Journal of Experimental Psychology*. **15** (No 2): 384-393.

Putnam, R. (1990). "Transfer of Action Inquiry Skills." Unpublished Ph.D. dissertation, Harvard University

Raff, I. (1960). "Dynamics of the Tankship Industry." MS thesis, MIT, Dept of Ocean Engineering, Cambridge, MA.

Raia, A.P. (1966). "A Study of the Educational Value of Management Games." *Journal of Business*. **39**(No 3): 339-352.

Raiffa, H. (1968). *Decision Analysis*. Reading, MA: Addison-Wesley.

Randers, J. (1984a). "Prediction of Paper Pulp Prices: A Review Two Years After." wp 84/6, Norwegian School of Management, Oslo.

Randers, J. (1984b). "The Tanker Market." Working paper, Norwegian School of Management, Oslo.

Rapoport, A. (1967). "Variables affecting decisions in a multistage inventory task." *Behavioral Science*. **12**: 194-204.

Rapoport, A. and Wallsten, T. (1972). "Individual Decision Behavior." *Annual Review of Psychology*. **23**: 131-76.

Rasmussen, J. (1974). "Outlines of a hybrid model of the process plant operator." In Sheridan, T.B. and Johanssen, G., ed. *Monitoring Behavior and Supervisory Control.*

Rawlinson, R. and Porter, M.E. (1983). "The Oil Tanker Shipping Industry in 1983." Case 9-384-034, Harvard Business School, Cambridge, MA.

Reed, S.K., Dempster, A. and Ettinger, M. (1985). "Usefulness of Analogous Solutions for Solving Algebra Word Problems." *Journal of Experimental Psychology*. **11**(No 1): 106-125.

Reitman, W. (1964). "Heuristic decision procedures, open constraints, and the structure of ill-defined problems." In Selley, W. and Bryan, D.S., ed. *Human judgments and optimality.* New York: Wiley.

Reve, T. (1992). "Organisatoriske forklaringsmodeller til bankkrisen" Working paper. Norwegian School of Economics and Business Administration.

Richardson, G. (1991). *The Feedback Thought in Social Science*. University of Pennsylvania Press, Philadelphia, PA

Ross, B.H. and Kennedy, P.T. (1990). "Generalizing From the Use of Earlier Examples in Problem Solving." *Journal of Experimental Psychology*. **16**(No 1): 42-55.

Salomon, G. (1987). "AI in Reverse: Computer Tools that turn Cognitive." Tel Aviv University, Tel Aviv.

Salomon, G. and Globerson, T. (1987). "Skill may not be enough: The role of mindfulness in learning and transfer." *Journal of Education* **11**(Special Issue): 623-637.

Salomon, G. and Perkins, D. (1989). "Rocky Roads to Transfer: Rethinking Mechanisms of a Neglected Phenomenon." *Journal of Education* **24** (2): 113-142.

Salomon, G., Guterman, E. and Globerman, T. (1989). "The Computer as a Zone of Proximal Development: Internalizing Reading-Related Metacognitions From a Reading Partner." *Journal of Education* **81**(4): 620-627.

Salomon, G., Perkins, D.N. and Globerson, T. (1991a). "Partners in Cognition: Extending Human Intelligence with Intelligent Technologies." *Journal of Education* **20**(3): 2-9.

Salomon, G., Zellermayer, M., Globerson, T. and Givon, H. (1991b). "Enhancing Writing-Related Metacognitions Through a Computerized Writing Partner." *Journal of Education* **28**(No 2): 373-391.

Salpukas, A. (1990).*After a Long Slump, World Shipping is Embarking on a Modest Recovery.* New York Times. New York:

Schein, E. H. (1992). *How can Organizations Learn Faster? The Problem of the Green Room.* World Economic Forum, Davos, Switzerland.

Schön, D. (1983). *The Reflective Practitioner: How Professionals Think in action.* New York: Basic Books.

Senge, P.M. (1990). *The Fifth Discipline: The art and practice of the learning organization.* New York, NY: Doubleday.

Senge, P.M. and Sterman, J.D. (1990). "Systems thinking and organizational learning: Acting locally and thinking globally in the organization of the future." Working paper no. D-4158, System Dynamics Group, Sloan School of Management, M.I.T., Cambridge, MA.

Sheridan, T.B. and Johanssen, G. (1974). *Monitoring Behavior and Supervisory Control.*

Simon, H.A. (1956). "Rational Choice and the Structure of the Environment." *Psychological Review.* **63**(2): 129-38.

Simon, H.A. (1973). "The Structure of Ill Structured Problems." *Artificial Intelligence.* **4**: 181-201.

Simon, H.A. (1976). *Administrative Behavior.* New York: Free Press.

Simon, H.A. (1978). "Rationality as process and product of thought." *American Economic Review.* **68**: 1-16.

Simon, H.A. (1979). "Rational decision making in business organizations." *American Economic Review.* **69**(4).

Simon, H.A. and Hayes, J.R. (1976). "The Understanding Process: Problem Isomorphs." *Cognitive Psychology.* **8**: 165-190.

Simon, H.A. and Hayes, J.R. (1977). "Psychological Differences Among Problem Isomorphs." *Cognitive Psychology.* .

Slovic, P. (1971). "Hypothesis testing in the learning of positive and negative linear functions." **11**: 368-376.

Slovic, P., Fischoff, B. and Lichtenstein, S. (1977). "Behavioral Decision Theory." *Annual Review of Psychology.* **28**: 1-39.

Smith, E.E. (1990). "Categorization." In Osherson, D.N. and Smith, E.E., ed. *An Invitation to Cognitive Science. Volume III: Thinking.* Cambridge, MA: MIT Press.

Smith, V. L. (1982). Microeconomic Systems as an Experimental Science. *American Economic Review,* 72(5), 923-55.

Smith, V. L.,Suchanek, G. L., and Williamns, A. W. (1987, May). Bubbles, crashes and endogenous expectations in experimental spot asset markets. p.

Snow, C. (1976). "A Comment on Business Policy Teaching Research." *Academy of Management Review*. (October 1976): 133-13' .

Staw, B.M. (1976). "Knee-deep in big muddy: A study of escalating commitment to a chosen course of action." *Organizational Behavior and Human Performance*. **16**: 27-44.

Steinmann, D. O. (1976). The Effects of Cognitive Feedback and Task Complexity in Multiple-Cue Probability Learning. *Organizational Behavior and Human Performance*, 15, 168-179.

Sterman, J. D. (1985, STRATEGEM-2: A microcomputer simulation game of the Kondratieff cycle (User's Guide). p.

Sterman, J.D. (1988). "Deterministic chaos in models of human behavior: Methodological issues and experimental results." *System Dynamics Review*. **4**(1-2): 148-178.

Sterman, J.D. (1989a). "Misperceptions of Feedback in Dynamic Decision Making." *Organizational Behavior and Human Decision Processes* **43**: 301-335.

Sterman, J.D. (1989b). "Modeling managerial behavior: Misperceptions of feedback in a dynamic decisionmaking experiment." *Management Science*. **35**(3): 321-339.

Sterman, J.S. and Paich, M. (1990). "Boom and Bust Markets." Paper presented at the 1990 Int'l System Dynamics Conference, Boston, MIT, Cambridge.

Sweller, J., Mawer, R.F. and Ward, M.R. (1983). "Development of Expertise in Mathematical Problem Solving." *Journal of Experimental Psychology*. **112**: 639-661.

Taylor, J (1989) "The Dinosaurs are Dying," *Forbes*, May 1, , pg. 92.

Teach, R. (1990). Profits: The False Prophet in Business Gaming. *Simulation and Games*, 21(No 1), 12-26.

Teach, R. (1990). "Profits: The False Prophet in Business Gaming." *Simulation and Games*. **21**(No 1): 12-26.

Teger, A.I. (1980). *Too much invested to quit*. New York: Pergamon Press.

Thaler, R. and Johnson, E. (1986). "Hedonic framing and the break-even effect." Cornell University, Johnson Graduate School of Management, Ithaca, NY.

Tierney, W.M., Hui, S.L. and McDonald, C.J. (1986). "Delayed Feedback of Physician Performance Versus Immediate Reminders to Perform Preventive Care." *Medical Care*. **24**(No 8): 659-666.

Timmers, H.a.W., W.A (1977). "Inverse statistics and misperception of exponential growth." **21**: 558-562.

Toda, M. (1962). "The design of a Fungus-Eater: A model of human behavior in an unsophisticated environment." *Behavioral Science*. **7**: 164-83.

Tranøy, B. S. (1992) *Bankkrisen. Årsaker til og betydning av endrede politiske rammebetingelser.* MA thesis, University of Oslo, Department of Political Science.

Tversky, A. and Kahneman, D. (1987). "Rational choice and the framing of decisions." In Hogarth, R.M. and Reder, M.W., ed. *Rational choice: The contrast between economics and psychology.* Chicago: Univ. of Chicago Press.

Tzelgov, J., Tsach, U. and Sheridan, T.B. (1985). "Effects of indicating failure odds and smoothed outputs on human failure detection in dynamic systems." *Ergonomics.* **28**(No 2): 449-462.

Vennix, J. (1990). *Mental Models and Computer Models.* Ph.D thesis. Nijmegen, The, Netherlands: Nijmegen Institute for Cognition Research and Information Technology.

Vennix, J. A. M., and Geurts, J. L. A. (1987). Communicating insights from complex simulation models , *Simulation and Games,* 18(No 3), 321-343.

Vennix, J.A.M. and Geurts, J.L.A. (1987). "Communicating insights from complex simulation models." *Simulation and Games.* **18**(No 3): 321-343.

Wagenaar, W.A.a.S., et al. (1975). "Misperception of exponential growth." *Perception and Psychophysics* **18**: 416-422.

Wagenaar, W.A.a.T., H (1978). "Extrapolation of exponential time series is not enhanced by having more data points." *Perception and Psychophysics* **43**: 182-184.

Wagenaar, W.A.a.T., H. (1979). "The pond-and duckweed problem: Three experiments in the misperception of experimental growth." *Acta Psychlogica* **43**: 239-251.

Wall Street Journal (1990) "Nation's Banks Cautioned about Real-Estate Lending," February 13, 1990, A2,2.

Wattenmaker, W.D., Dewey, G.I., Murphy, T.D. and Medin, D.I. (1986). "Linear Separability and Concept Learning: Context, Relational Properties, and Concept Naturalness." *Cognitive Psychology.* **18**: 158-194.

Weick, K.E. (1977). "Enactment Processes in Organizations." In Staw, B.m. and Salanchick, G.R., ed. *New Directions in Organizational Behavior.* Chicago: St. Clair Press.

Wheaton, W. and Torto, R. (1988). "Vacancy Rates and the future of Office Rents." *AREUEA Journal.* **Summer**.

Wheaton, W. C. (1990). Vacancy, Search, and Prices in a Housing Market Matching Model , *Journal of Political Economy,* 98(61), 1270-1292.

Wiener, N. (1961). *Cybernetics, or Control and Communication in the Animal and in the Machine.* 2nd. Cambridge: MIT Press.

Wolfe, T. (1976). "The Effects and Effectiveness of Simulations in Business Policy Teaching Applications." *Academy of Management Review.* **1**(No 2): 47-56.

Wolfe, T. (1985). "The Teaching Effectiveness of Games in Collegiate Business Courses." *Simulation and Games.* **16**(No 3): 251-288.

Womack, J., Jones, D. and Roos, D. (1990). *The Machine That Changed The World.* New York, NY: Macmillan Publishing Co.

Zannettos, Z. (1960) *A Theory of Tank-ship rates*. NY, NY

Appendix 1: Pilot studies

Pilot study 1

Procedure

Subjects were familiar with accounting and finance; most were current or future MBA's. Seventeen were graduate students at MIT's Sloan School of Management and had completed at least one semester-long course in modeling social systems and commercial markets using the system dynamics framework. They participated voluntarily and were paid according to performance. The other 32 subjects were managers from major corporations. Many had working experience in one of the two markets. Most had 2 to 10 years' professional experience in addition to an MBA degree. Whereas students played individually, professionals played in teams of two, and the results of the 16 professional teams are compared to the 17 individual students.

Participants first read a 2 page newspaper article about current market conditions (appendix C). All subjects (also professionals) worked individually at this point. A seven page briefing book about the simulator was provided before the subjects started interacting with the computer . Participants were then instructed to play until they had finished at least a full, i.e. 40 period, trial of the game. They were instructed that they would be reset to the beginning year in case of bankruptcy and that they should continue until the screen had said "game over". The exact length of a trial was not communicated to subjects.

After a break that lasted from half an hour to a week, subjects were presented with the other market and the instructions to continue playing until "game over" was shown were repeated. The participants filled out questionnaires about their professional and educational experience at the end of the second session.

Students were paid a flat, $4 per hour pay for the about 4-5 hours the experiment lasted. In addition, they were given a bonus that was a linear function of their game performance. This bonus averaged an additional 4 $ per hour and amounted to a total $110 for the participant with the highest score. The professionals had to announce their results publicly to their peers after the sessions.

The performance metric was a modified M1 so that a trial ended in the bankruptcy year or year 40, whichever came first. In the case of no bankruptcy, the accumulated score for the first 40 decisions is reported. In case of bankruptcy, M1 is reported

$$M1_s = \sum_t p_{s,t}\Omega E_{s,t} \Big/ \sum_t p_{b,t}\Omega E_{b,t}$$

Where

$M1_s$ = measure 1 for subject s
$p_{s,t}$ = subject s' profit for decision t
$p_{b,t}$ = benchmark profit for decision t
$E_{s,t}$ = information and leverage environment available to subject s before making decision t
$E_{b,t}$ = information and leverage environment available to benchmark before making decision t

$t = 1, 2, \ldots,$ trial end
trial end = min. (bankruptcy year, 40)
$s = 1, 2, \ldots, n$

Figure A.1 shows that the pilot study 1 interface contained information about the last 6 decision periods. This interface contained so much time series information that the bottom information buttons were not used. Thus, information acquisition data were not meaningful. Consequently, the main study was implemented using an interface where only information about he current year was presented on the screen, and where simulated history was obtained by requesting a "graph" or "table". The later change was made in order to make information acquisition explicit. Other than enabling to monitor how subjects accessed information, no decision or learning influence from the design change was expected.

Make Decisions	1989	1990	1991	1992	1993	1994
Buy/Sell Existing	0					
Permit to Develop	10					
Under Construction	15					
Leasable space	1000					
New Rents	30					
Existing Price	144					
Development Price	150					
Management Cost	14					
Vacancies/1000	120					
Rental Revenue	27					
Financial Result	-20					
Total mgt Expenses	13					
Transaction Fees	0					
Profits	-6					
Cumulative Profits	15					
Book Value	150					
Total U/construction	750					
Total Market Space	50000					
Demand	44000					

Our sq ft Chart

Lease rates&Prices

Profits & Assets

Total Market

All data table

Figure A.1: User interface of pilot study 1.

Results

By looking at the average number of decisions before a full 40 period trial was completed, we tap into the decision making approach of the player. Many decisions before a full trial may indicate that a player has an exploratory attitude; she goes bankrupt many times and so suffers in the short run with the potential of gaining deeper understanding that might be beneficial in the long run. However, many bankruptcies may also indicate that the person does not understand the game dynamics. If bankruptcies are indicative of poor understanding, then subjects who do poorly in the initial scenario trial should also do poorly in the transfer trial.

No difference was found between subject performance in the two markets. Recall, however that there were only two markets, one familiar/low frequency (real estate) and another unfamiliar/high frequency (oil tanker). This design thus combined high dynamic complexity with context familiarity in the one market and low dynamic complexity with lack of context familiarity in the other. The results showed a between-group difference of student and professional subjects in terms of number of bankruptcies in market 1 in that students went bankrupt more often. In the transfer trial there was no difference between students and professionals in terms of number of bankruptcies, meaning that students significantly reduced their bankruptcies. Consequently student performance among students in the transfer market increased significantly, while the increase in professional performance was insignificant.

Figure A.2 shows that when presented with the second market, the students outperform the professionals. The differences in change in student bankruptcies and performance were significant at the 0.05 level (two-tailed tests)[16]. So were all between group differences, excluding the number of bankruptcies in the transfer market and initial performance.

[16] Due to an accident, the original performance data were lost. No statistical measurements could be obtained. The scores in pilot 1 come from Bakken, 1990b.

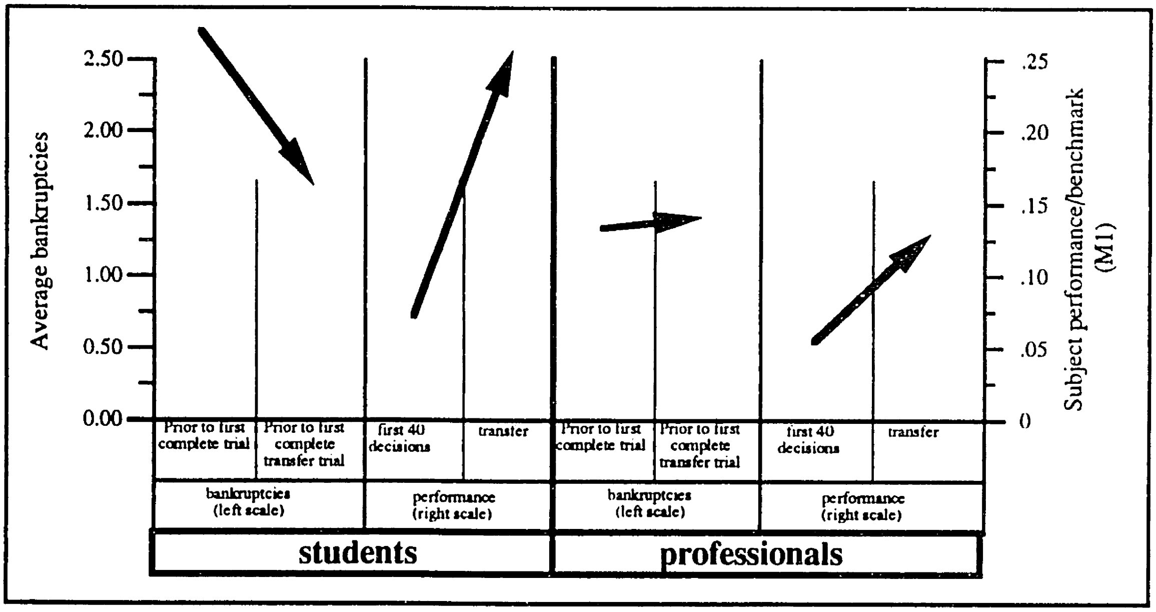


Figure A.2: Students go bankrupt more often early than later, but professionals go broke with equal frequency regardless of experience.

Note that average student performance in the transfer trial was about 25% of benchmark Not surprisingly, the subject factor was a significant contributor to performance. This is not shown for study 1, but figure 6.2.2.3 shows the analog picture for study 2.

Discussions

Why do professionals use a less exploratory decision strategy than students ?

Students played individually and professionals in teams of two. This introduced a confounding factor of risk attitudes. The group polarization phenomenon assures that team decision making exaggerates risk attitudes in groups compared to average risk attitude among group members. Results indicate that although every effort was made to ensure a non-threatening decisions environment, professionals have a deeper real world experience and so will feel more naturally constrained in their decision repertoire. The same imbedded routines that hinder open-mindedness and exploration in the real world may also prevent exploration and learning in the simulated environment. The simulated environment is probably more threatening to professionals than to students, especially to those that participated in environments where they had years of experience. Some of the professional teams' internal discussions were recorded during game playing. A typical group discussion was:

Person 1: "Buildings are cheap, let's buy up to our financing limit."

Person 2: "What do you mean. Buy ? This market is like the one I *remember* from the Texas market in 1985. Let's wait four or five more years. Now it is too risky to do anything."

The professional subjects with market experience clearly disliked that they didn't find their day to day decision environment reproduced in the market, and a typical comment would be:

Person 3 to experimenter: "We are bankrupt, this is not fair. How can we do well when you have put such a heavy depreciation burden on us. And don't you know that we are a company committed to quality, while in this game we only operate in a commodity market."

However, during the debriefing discussion, participants agreed that the simplifications were good, since they helped focus on dynamic processes that were partly disentangled from and longer the business cycle. Several professionals were unaware of the fact that capital cycles necessarily are of lower frequency than business cycles. In hindsight, professionals agreed that their experience had made it hard for them to accept the markets portrayed by the simulation game and that instead of exploring the market, they had tried to avoid bankruptcy.

The simulated environment is threatening enough for professional participants to limit their decision repertoire and make learning and transfer difficult. Threat has the general effect of narrowing options considered (Gladstein and Reilly, 1985). Thus, without imbedding the learning lab in a context that makes the simulator less threatening, the results indicate that real world learning disabilities will be reproduced in the laboratory.

Should professionals mimic students?

Students' game playing strategy seemed more playful and students use information available in the game as a springboard for investigating causal dynamics to a higher extent than professionals. Students take the exercise more as a learning experience. In the very short run they suffer and go bankrupt, but in the long run they improve their performance.

The professionals spent several hours in directed discussions after the experimental game-playing and measurements finished. These discussions remedied some of the transfer shortcomings of the simulated markets. First, experiences were shared and enabled those who had followed conservative strategies to benefit from those who had pursued more adventuresome decisions. In a very compressed time interval, the discussion revealed that markets that seem to be different share many commonalties. Participants could quickly perform and discuss mental what-if analyses due to the intense shared experience. Furthermore, because of the substantial overlap of participants' experiences, sharing these mental simulations was greatly facilitated.

As an example, professionals in one post-discussion session questioned that asset values could swing as much as 40% between peaks and troughs. During the last ten years, they argued, asset values in the northeast US had only gone up and in other parts of the country values had never dropped by more than 10%. Interestingly, there were people in the *same session* with detailed exposure to a different region that also questioned the validity of the simulated asset cycles. They had experienced asset value reductions of

50% over a period of a few years! The discussion then turned towards the preconditions for such a fall and what it meant for the northeast region.

From an organizational learning viewpoint, it is interesting that the colleagues with exposure to the market with the 50% drop *had not previously shared his experiences*. He had, until the simulated experience, *thought that the other region was so unique that it was irrelevant to current colleagues*.

One corporation used the learning lab to motivate changes in its incentive structure. These changes were predicated on expected future difficulties induced by the cyclical nature of their business. Until then, it had been hard to convince junior partners that the future of development was problematic. The lesson taken from three to ten years' professional experience during the upswing of the cycle was "an investment missed today is millions lost in unearned capital gain tomorrow." The robustness of such a lesson is greatly reduced if the time horizon is expanded from the common 3-4 years to 30-40 years! The cyclical experiences and emotionally laden bankruptcies improved the receptiveness for organizational change by making market frequency ranges more salient through participation and bankruptcies.

Pilot study 2

As a result of the first pilot study, it was decided to change the user interface. Measurement units were included. This made the task less taxing on people's memory and following Bassok (1990) would help transfer. Another refinement consisted in separating market frequency and context and the 2x2 design (frequency x context) described in chapter 3 was implemented. The information display was modified so as to only show information about the current decision period. Access to prior decision periods could only be obtained through clicking appropriate graph and table buttons. These clicks were recorded.

One of the corporations where the pilot study 1 learning lab took place, decided to offer it to another US region. This study, called pilot study 2, took place in January 1991 with a group of 12 real estate professionals. They participated in groups of 2. Results from these 6 groups are recorded.

The corporation chose to expand the 3/4 day program in study 1 to a 2 day program in study 2. The apparent success of study 1 in terms of changes in organizational policies had convinced the firm to continue using the games that now was the introduction to a broader program. Otherwise, pilot study 1 for the professionals and study 2 were essentially equal. No formal incentives were provided, but players were highly competitive. Even though they were told that "results are not directly comparable", participants vividly discussed their performance after the game. There is no reason to believe that they did not do their best.

Procedure

Teams started with the unfamiliar context and made decisions in this market for ten periods. This was done in order for subjects to gain market familiarity in an environment where they felt less threatened. By defining the playing field in a different industry, it was hoped that risk aversion would be reduced and insight, learning and transfer improved. No results from these initial decisions in the unfamiliar context are reported. Teams then turned to the familiar real estate context where two teams continued with the 40 decisions in the high frequency market and later switched to the low frequency market. The other four teams played the reversed sequence.

Results

The score metric was M1 as in pilot study 1.

Contrary to study 1/professionals, no trial effect was found as shown in figure A.3. Though there was main effect of frequency, the graph shows that the low frequency market showed higher performance variance. As mentioned above, some teams are good and some are bad. The Team effect was significant , $p<0.05$, two-tailed test..

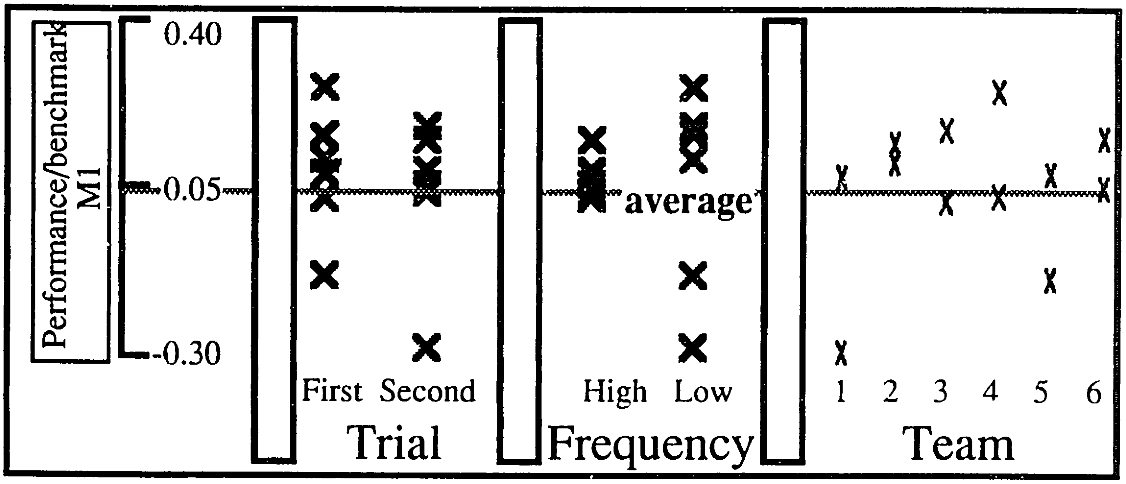


Figure A.3. Trial, Frequency and Team in pilot study 1.

Discussion

Contrary to study 1, no Trial effect was found in study 2. The lack of Trial effect may be due to learning taking place prior to recording during the initial Tanker decisions. As indicated in study 1, an exploratory attitude will enhance learning and the choice of a unfamiliar context may have induced more learning than for professionals in study 1. Study 2 Tanker games contained more early bankruptcies than professionals in study 1, a fact that can be contributed to the reduced threat posed by what was expressly termed "trial games". Study 1 also allowed far more decisions between the first and last 40 decisions than study 2, with an average of 60-70 decisions between the two measures in 1 and no decisions between the two measures in study 2.

Summary of pilot studies

Even though performance in study 1 and study 2 are not directly comparable, it appears that professionals' performance in both studies plateaus out on a very poor level. The performance leap realized by the study 1 students in their transfer trial did not materialize among professionals in either study.

The first experiment showed a transfer effect (mainly due to the students superior performance), whereas the second did not. Now, the sample size in the two studies was small, especially in the last study where only 6 teams participated and a total lack of transfer there might be a statistical aberration and goes against the common sense that people should improve with experience.

Findings of poor transfer performance among professionals corroborates several earlier studies. Brehmer (1988) found that in a real time system, mimicking the situation among the professionals here in that the setting was not very conducive to experimentation and reflection, performance was poor. After a couple of trials, performance did not improve further but converged at a poor level. Once lags were removed from the simulated system performance was higher and furthermore continued to improve after 6 trials.

Hogarth, et al. (1989) similarly suggest that exacting tasks hinder performance. The environments faced by both students and professionals, where experimentation may lead to bankruptcies can be said to be exacting. Though student and professional subjects alike were informed that the purpose of the study was to investigate and enhance learning and the objective context was the same, discussion protocols and decision making data all suggest that the task played a different role for professional than for student subjects.

The studies indicated that professionals replicate real life decision making in the laboratory. (See also Moissis, 1989). Their view of the decision making context will thus be deliberately constrained and not make use of information more readily available in the game than in real life. Supply line information, presented on the screen in the game, is notoriously unavailable in real life (Randers, 1984) but professionals' may limit their information acquisition. Unfortunately, study 1 contained little meaningful information acquisition data and the claim that professionals used supply line information less can not be inferred directly from the results.

A consequence of using established mental models must be that simulated market experience has a harder time entering subjects' problem space. Lack of cognitive effort among professionals may lead to a less thoroughly worked out mental model of the experimental market and subsequent lack of transfer.

Added to the cognitive factor is the experiential one: Due to their constrained decisions, professionals experience a narrower bandwidth of the system than students. In dynamic, unstable (i.e. exacting) tasks, decisions interacts with and govern experimental markets' behavior. As a consequence, when attempting to transfer their understanding to a new frequency (and/or context), professionals meet an environment that is "newer" to them than to the students who have "pushed the system further" and consequently seen the system from more a broader perspective. The students, due to their more varied initial contact with the system, meet a new environment that is less "new" than what professionals experience.

Appendix 2: The detailed model stock- and flow structure

Figures A.4 to A.8 show the stock-and flow structure for the real estate context. In figure A.4 terms, New Permits are requested when Desired Buildings increase as a consequence of high Average Rents. In addition, the competitors realize that there is a need for a pipeline of buildings Under Construction. This need reflects the Demolition Rate and the Construction Time. Because the game is implemented as an interactive decision making simulation, there is a distinction between the Decision Maker and the Competition. The variable names in bold in A.4, i.e. **New Permits** and **Acquisition** represent the only decisions in the game. New Permits first increase the total number of buildings Under Construction, and later on lead to more Buildings. Positive Acquisitions increase the decision maker's Buildings, whereas negative Acquisitions decrease their number in his possession. Acquisitions do not change the total supply of buildings but only move buildings between the player and the competition.

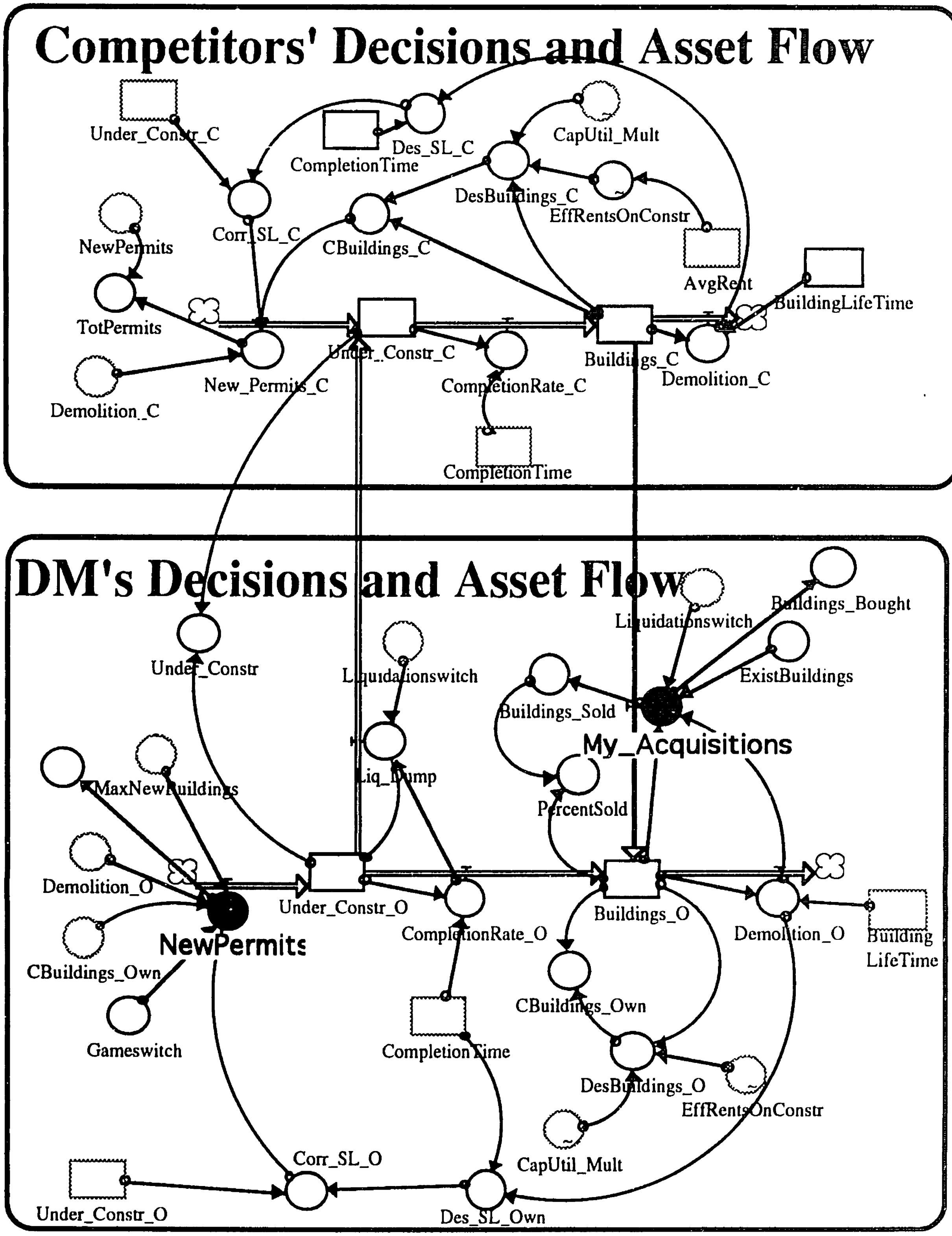


Figure A.4: The flow of assets and decision making structure.

Rent formation is shown in figure A.5. Average Rents are a consequence of total Demand and total Supply in the market. In other words normalized Capacity Utilization determines the Rent base, which again is influenced by Variable Costs. The Average Rent is a simple information smooth of the Rent.

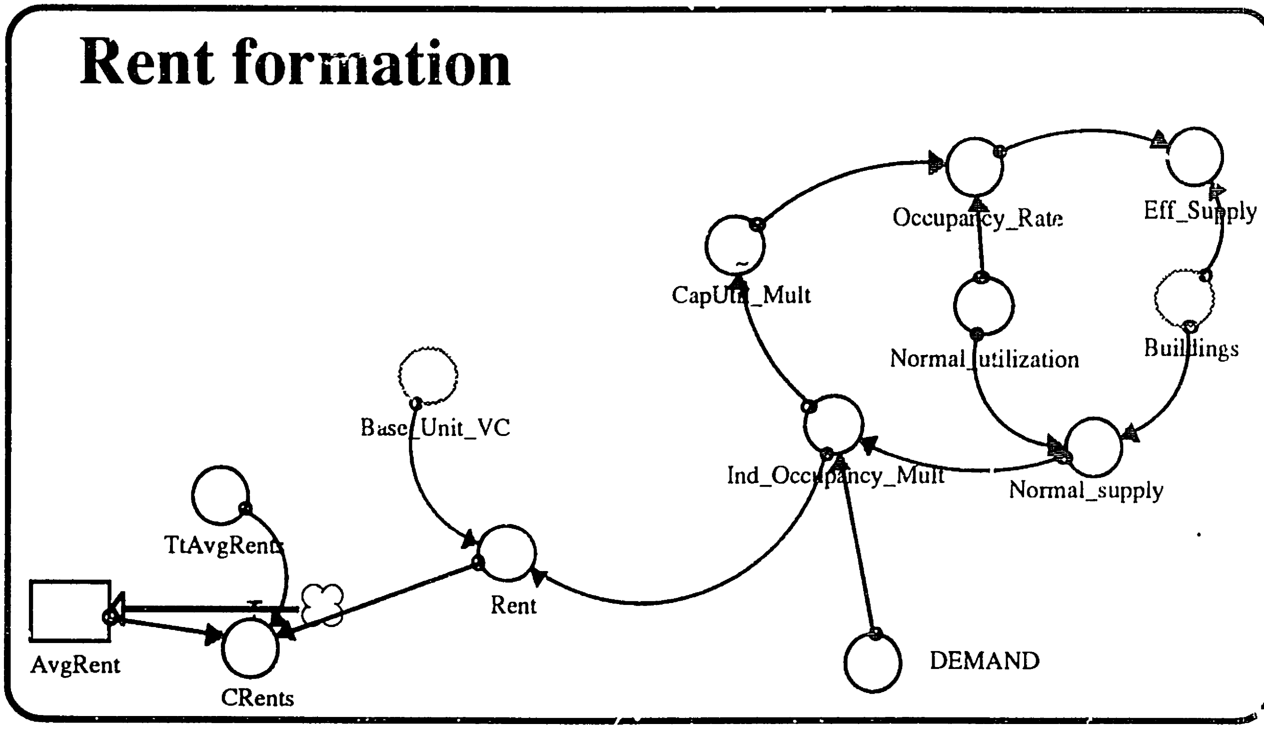


Figure A.5: Rent formation

The rest of the model is basically accounting. Figure A.6 shows profits and loss, figure A.7 shows available funds for investment, figure A.8 shows bookkeeping of loans and assets.

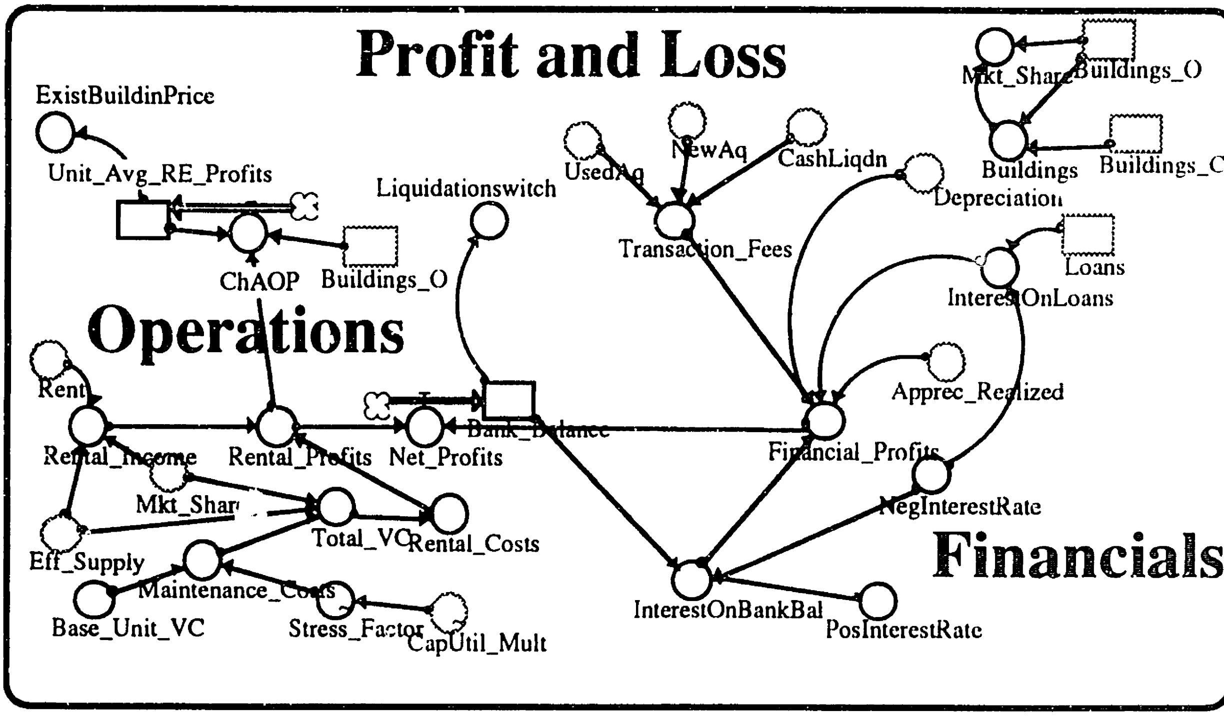


Figure A.6: Profits and loss.

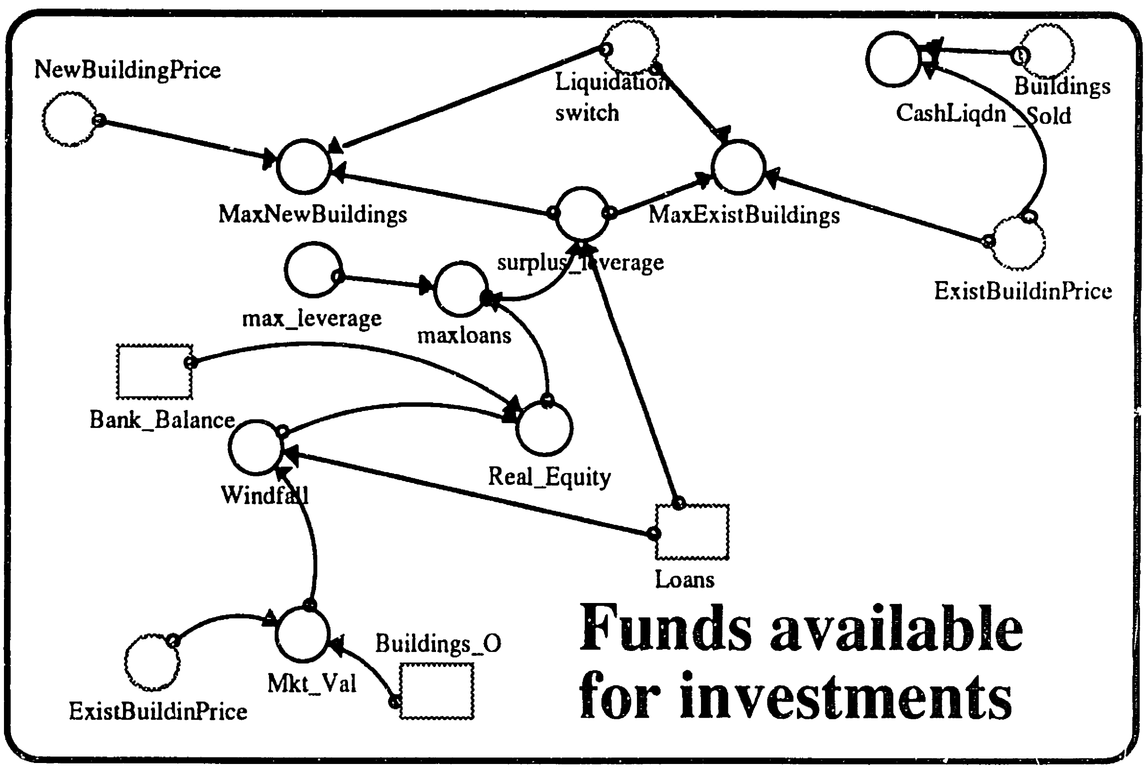


Figure A.7: Funds for investments

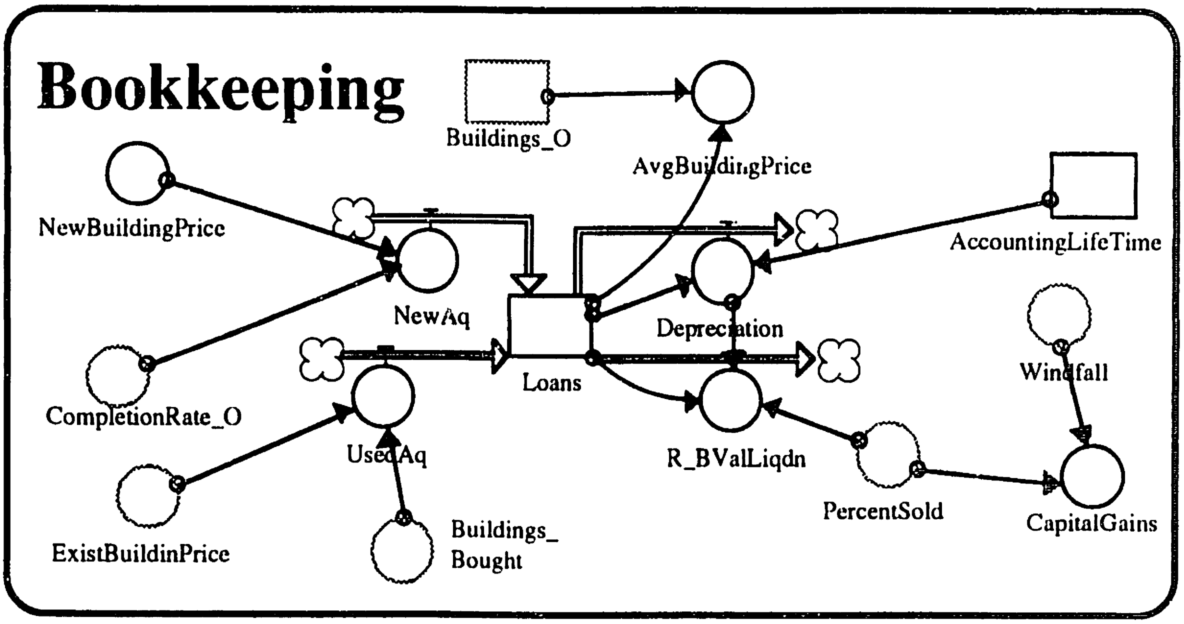


Figure A.8: Bookkeeping of investments

Appendix 3: The model and benchmark equations

The model was implemented in a Wingz® electronic spread-sheet. The initial spreadsheet is shown below in figure A.9 without information acquisition buttons. Line and column numbers have been added to the user interface as well.

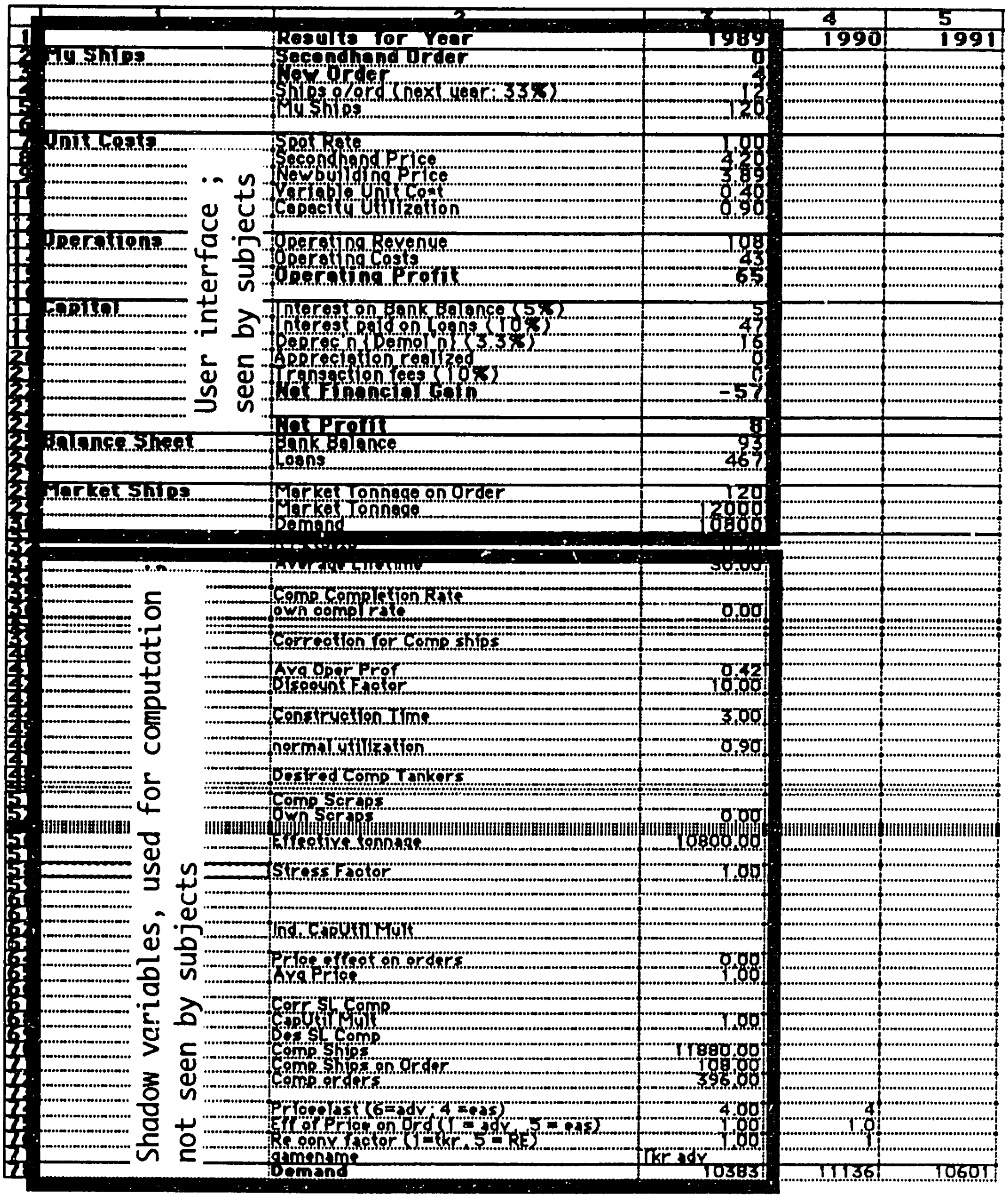


Figure A.9: The line and column addresses used in the model program

Listing A.10 shows the model program linking period t to t+1

```
Define in,out,cin,cout, colon0, colon1, counter, terminate, gamenr, dummy,clock,
Priceelast, EPOR, REconv,gamename, count
Priceelast = c74

EPOR = c75

REconv = c76

gamename = c77

repaint off
terminate = a100
while terminate =0
{nonlinearity EPOR calc}

cin = makecell((col()-2),65)
in = (indirect(cin))/REconv

if in > 5. out = 4
elseif in > 4.5 out = (.1*in +3.5)
elseif in > 3.5 out = (.25*in + 2.825)
elseif in > 3 out = (.3*in + 2.635)
elseif in > 2.5 out = (.6*in + 1.575)
elseif in > 2. out = (.9*in +.675)
elseif in > 1.5 out = (1.15*in +.075)
elseif in > .5 out = (1.35*in - .35)
elseif in >0 out = (.65*in)
else out = 0
end if
put out^EPOR into makecell((col()-1),64)

{recalc column}
put r9c3 into makecell(col()-1,9)
put indirect(makecell(col()-2,78)) into makecell(col()-1,30)

if indirect(makecell(col()-2,34)) < 0

put indirect(makecell(col()-2,34)) * 1.1 into makecell(col()-1,34)
else
put indirect(makecell(col()-2,34)) * 1.05 into makecell(col()-1,34)
end if

put indirect(makecell(col()-2,4))/r44c3 into makecell(col()-1,36)

put indirect(makecell(col()-2,29))*r46c3 into makecell (col()-1,80)

put indirect(makecell(col()-2,4))-indirect(makecell(col()-1,36))
+indirect(makecell(col()-1,3)) into makecell((col()-1),4)

put indirect(makecell (col()-2,5))/r33c3
 into makecell((col()-1),52)

put indirect(makecell (col()-2,5))
+indirect(makecell(col()-1,36))
+indirect(makecell((col()-1),2))
```

```
-indirect(makecell(col()-1,52)) into makecell((col()-1),5)

put (indirect (makecell(col()-1,30))/indirect(makecell(col()-1,80)))
into makecell(col()-1,62)

put ((r79c3/(REconv) + (indirect(makecell(col()-1,62))-(r79c3/(REconv)*.2040353))^(Pricеelast)))
*REconv
into makecell(col()-1,7)

put (indirect(makecell (col()-1,7))- indirect(makecell(col()-2,65)))/2
into makecell((col()-1),32)

put (indirect(makecell (col()-2,65))+indirect(makecell(col()-1,32)))
 into makecell((col()-1),65)

{Capacity utilization multiplier computation}
in = indirect(makecell(col()-1,62))
if in > 2 out = 1.19
elseif in < 1 out = in
elseif in < 1.2 out = .5*in + .5
elseif in < 1.4 out = .25*in + .8
elseif in < 1.6 out = .125*in + .975
elseif in < 1.8 out = .065*in + 1.071
else out = .01*in + 1.17
end if

put out into makecell(col()-1,68)

put indirect (makecell(col()-1,68))*r46c3 into makecell(col()-1,11)

{Stress factor computation}
in = indirect (makecell(col()-1,68))
if in > 1.2 out = 4.1
elseif in < .8 out = .75
elseif in < 1 out = 1.25 * in -.25
elseif in < 1.04 out = 2.5 * in -1.5
elseif in < 1.08 out = 5 * in -4.1
elseif in < 1.12 out = 10 * in -9.5
elseif in < 1.16 out = 20 * in -20.7
else out = 40 * in -43.9
end if
put out into makecell(col()-1,58)


put indirect(makecell(col()-1,58))*r79c3
 into makecell(col()-1,10)

put (indirect(makecell (col()-2,71)) /r44c3 )
into makecell(col()-1,35)

put (indirect(makecell (col()-2,70))/r33c3)
 into makecell(col()-1,51)

put indirect(makecell(col()-1,51))*r44c3 into makecell(col()-1,69)

put (indirect(makecell(col()-1,69))-indirect(makecell(col()-2,71)))/r44c3
into makecell(col()-1,67)
```

```
put indirect(makecell (col()-2,70)) + indirect(makecell(col()-1,35))
 - indirect(makecell(col()-1,51)) - indirect(makecell(col()-1,2))
 into makecell((col()-1),70)

put (indirect(makecell (col()-2,70))*indirect(makecell(col()-1,64))
*indirect(makecell (col()-1,68)))
 into makecell((col()-1),48)

put (indirect(makecell((col()-1),48))-indirect(makecell(col()-2,70)))/r44c3
into makecell((col()-1),39)

put max(0,(indirect(makecell((col()-1),39))+indirect(makecell(col()-1,51))
+(.75*indirect(makecell (col()-1,67)))))
into makecell((col()-1),72)

put indirect(makecell (col()-2,71)) + indirect(makecell(col()-1,72))
-indirect(makecell(col()-1,35)) into makecell((col()-1),71)

put (indirect(makecell((col()-1),4))+indirect(makecell((col()-1),71)))
into makecell((col()-1),28)

put (indirect(makecell((col()-1),5))+indirect(makecell((col()-1),70)))
 into makecell((col()-1),29)

put (indirect(makecell((col()-1),29))*indirect(makecell((col()-1),11)))
into makecell((col()-1),56)


put indirect(makecell((col()-1),7))*indirect(makecell(col()-1,5))
        *indirect(makecell(col()-1,11))
into makecell((col()-1),13)

put
        indirect(makecell(col()-1,5))* indirect(makecell((col()-1),11))*
        indirect(makecell((col()-1),10))
into makecell((col()-1),14)


put indirect(makecell(col()-1,13)) - indirect(makecell((col()-1),14))
into makecell((col()-1),15)

put indirect(makecell(col()-2,25)) * .05
into makecell((col()-1),17)

put indirect(makecell(col()-2,26)) * .1
into makecell((col()-1),18)

put indirect(makecell(col()-2,26))/r33c3
into makecell((col()-1),19)


if indirect(makecell(col()-1,2)) < 0

put -indirect (makecell(col()-2,8))*indirect(makecell(col()-1,2))
        + (indirect(makecell((col()-1),2))/indirect(makecell((col()-2),5)))*
                indirect(makecell((col()-2),26))
into makecell((col()-1),20)
```

```
else dummy = 0

put dummy into makecell((col()-1),20)

end if

if indirect(makecell(col()-1,2)) >= 0

put indirect(makecell((col()-1),2))*indirect(makecell((col()-2),8)) +
        indirect(makecell((col()-1),36))*indirect(makecell((col()-1),9))
        - indirect(makecell((col()-2),26))/r33c3
        + indirect(makecell((col()-2),26))
        into makecell((col()-1),26)

else

put indirect(makecell((col()-1),2))/indirect(makecell((col()-2),5))*indirect(makecell((col()-2),26))
        +indirect(makecell((col()-1),36))*indirect(makecell((col()-1),9))
        - indirect(makecell((col()-2),26))/r33c3
        + indirect(makecell((col()-2),26))
        into makecell((col()-1),26)

end if

put (abs(indirect(makecell((col()-1),2)))*indirect(makecell((col()-2),8)) +
        indirect(makecell((col()-1),3))*indirect(makecell((col()-1),9)))*.1
into  makecell((col()-1),21)

put indirect(makecell((col()-1),17))-indirect(makecell((col()-1),18))
- indirect(makecell((col()-1),19))+indirect(makecell((col()-1),20))
- indirect(makecell((col()-1),21))
into makecell((col()-1),22)

put indirect(makecell((col()-1),22))+indirect(makecell((col()-1),15))
into makecell((col()-1),24)


put indirect(makecell((col()-1),24)) + indirect(makecell((col()-2),25))
into makecell((col()-1),25)


Put indirect(makecell((col()-2),41))
        +(indirect(makecell((col()-1),15))/indirect(makecell((col()-1),5))
        -indirect(makecell((col()-2),41)))/2
                into makecell((col()-1),41)

Put indirect(makecell((col()-1),41))*r42c3 into
        makecell((col()-1),8)

{is count > 50 ? if so, start new game}


        select range makecell((col()-2),1)
        column width 0

        select range makecell((col()+1),1)

        column width 1200
```

```
        select range makecell((col()+1),1)

if indirect (makecell(col(),1)) > 2030
clock = now()
put clock into a90
select controls
clear
save as ""andgamename andcelltext(a90)
close
MESSAGE "saving results"
{count = 1}
end if
terminate = 2


end while
terminate = 0
put terminate into a100
```

Listing A.10: Model program listing

Listing A.11 shows the benchmark script

```
repaint off
define returned_value, doit, count, in, out, cin, cout, gamenr, terminate,
        minsho,maxsho,maxno,clock,purchv,currcol,MktTonnUconstr,
PrevMkttonnUConstr,OldMkttonnUconstr,
        WindFall,UsedPrice,NewPrice,RealEquit,MktVal,MaxLev,MaxInvLiq,LevMarg,CurrLev,
        MaxNewTonn, MaxUsedTonn, OwnTonnUConstr, OwnTonn, CompTonnUConstr, CompTonn,
CumCash,
        BookVal, gamename, lifetime, Demand, expusedprice, expected_demand, expected_supply,
        a, b, purchase, decrule, maxloans, prevusedprice,
        colon0, colon1, counter, dummy,develop,
        Priceelast, EPOR, REconv, K, BIAS, PROFCOL,LOSSCOL, L, M, N, bankbal, liqloss,
supplygradient,
        PrevOwnTonn, PrevComptonn, gamenbr, SUBJECTNBR, simprofit, achievedprofit, totprofit,
sellflag,counting, sellstep

LOSSCOL = 2
PROFCOL = 2
decrule = 2
FOR SUBJECTNBR = 11 TO 11 STEP 1

for gamenbr = 5 to 10 STEP 5

for sellstep = 50 to 50 step 5

New Worksheet ""
window location (-400, -400)
window size (13000, 8000)
save as "ClubMac :BeB:Research Games:Game
results/91:"andSUBJECTNBRand":"andGAMENBRand"c"
OPEN "ClubMac :BeB:Research Games:Game results/91:"andSUBJECTNBRand":"andGAMENBR
select range d1
FOR bias = 1.4 to 1.4 step .2

B=0
A=0
FOR B = .6 to .6 step .1
{B is agressiveness, from 0 to 1; 1 - buy all you can}
a = 0
{A is window of buying/selling from +/- .2 to +/- 1. before buying/selling}
{for a = .05 to .35 step .1}

FOR L = 1 TO 1 {to montecarlo simulate pink noise demand}
{select range D1..AR77
Clear}
cin = makecell(col(),1)
count = indirect(cin)
SELECT range b1..as85
copy

go to window "ClubMac :BeB:Research Games:Game
results/91:"andSUBJECTNBRand":"andGAMENBRand"c"
select range b1
paste values
select range d1
```

```
sellflag=1


FOR K= 5 TO 44

if K = sellstep
sellflag = 1
elseif K = 2*sellstep
sellflag = 1
elseif K = 3*sellstep
sellflag = 1
elseif K = 4*sellstep
sellflag = 1
elseif K = 5*sellstep
sellflag = 1
elseif K = 6*sellstep
sellflag = 1
elseif K = 7*sellstep
sellflag = 1
elseif K = 8*sellstep
sellflag = 1

end if


if sellflag =1
        go to window "ClubMac :BeB:Research Games:Game
results/91:"andSUBJECTNBRand":"andGAMENBR
        select range makerange(k-1, 4, k,85)
                achievedprofit = indirect(makecell(k,24))
        copy
        go to window "ClubMac :BeB:Research Games:Game
results/91:"andSUBJECTNBRand":"andGAMENBRand"c"
                select range makecell (k-1,4)

                paste values
else
        go to window "ClubMac :BeB:Research Games:Game
results/91:"andSUBJECTNBRand":"andGAMENBRand"c"
        achievedprofit = indirect(makecell(k,24))
        end if

sellflag = 0

select range makecell (k,1)

PURCHASE = 0
{a=a1
b=a2}
{decrule=a3
BIAS = A4}
REconv = c76
currcol= col()

gamename = c77
count = indirect(cin)
put count into makecell(col(),150)
```

```
clock = now()
put clock into makecell(col(),151)
terminate = 0

Demand = indirect(makecell(col()-1,30))


lifetime = c33
NewPrice        = indirect(makecell(col()-1,9))
UsedPrice       = indirect(makecell(col()-1,8))
prevusedprice = indirect(makecell(col()-2,8))
OwnTonnUConstr = indirect(makecell(col()-1,4))
OwnTonn         = indirect(makecell(col()-1,5))
CompTonnUConstr = indirect(makecell(col()-1,71))
MkttonnUconstr = CompTonnUConstr+OwnTonnUConstr
PrevMkttonnUConstr = indirect(makecell(col()-2,71))+indirect(makecell(col()-2,4))
OldMkttonnUconstr = indirect(makecell(col()-3,71))+indirect(makecell(col()-3,4))
CompTonn        = indirect(makecell(col()-1,70))
BookVal         = indirect(makecell(col()-1,26))
CumCash         = indirect(makecell(col()-1,25))
MktVal = UsedPrice*OwnTonn

WindFall        = MktVal-BookVal

RealEquit       = Windfall + CumCash

MaxLev          = 10

Maxloans        = RealEquit*10

LevMarg         = Maxloans-BookVal

MaxUsedTonn     = LevMarg/UsedPrice
MaxNewTonn      = LevMarg/NewPrice


if indirect(makecell(col()-1,25)) < 0

put (MktVal*.9 - BookVal)
+ indirect(makecell(col()-1,34)) - (c25) + CumCash
 into (makecell(col()-1,34))
select range c2..c6
select more range c17..c27
copy
select range makecell (currcol-1, 2)
paste VALUES
select range makecell (currcol, 1)

CompTonnUConstr = indirect(makecell(col()-1,71)) + OwnTonnUConstr - indirect(makecell(col()-1,4))
Put CompTonnUConstr into makecell((col())-1,71)
CompTonn        = indirect(makecell(col()-1,70)) + OwnTonn - indirect(makecell(col()-1,5))
put CompTonn into makecell((col())-1,70)

put indirect(makecell((col())-1,7))*indirect(makecell(col()-1,5))
        *indirect(makecell(col()-1,11))
into  makecell((col())-1,13)

put
```

```
        indirect(makecell(col()-1,5))* indirect(makecell((col())-1,11))*
        indirect(makecell((col())-1,10))
into  makecell((col())-1,14)


put indirect(makecell(col()-1,13)) - indirect(makecell((col())-1,14))
into  makecell((col())-1,15)

put indirect(makecell(col()-1,15)) + indirect(makecell((col())-1,22))
into makecell((col())-1,24)

OwnTonnUConstr = indirect(makecell(col()-1,4))
OwnTonn         = indirect(makecell(col()-1,5))

end if


minsho = int(max(-.1*CompTonn, -(1 -(1/lifetime))*OwnTonn))

maxsho = int(max (0,min(.1*CompTonn,MaxUsedTonn)))


{Decision Rule 1: expusedprice = usedprice}

if decrule = 1

expusedprice = usedprice
if expusedprice<(BIAS-a)*newprice {BUY}
purchase = {b}b*maxsho
end if

if expusedprice>(BIAS+a)*newprice {SELL}
purchase = 1*minsho
end if

end if    {END DEC RULE 1}

{Decision Rule 2: sell when gradient of SL starts to fall}

        if decrule=2

                if (MkttonnUconstr>PrevMkttonnUConstr) and
                {second derivative is negative (i.e growth gradient falling)}
                ((MkttonnUconstr-PrevMkttonnUConstr) < (PrevMkttonnUConstr-OldMkttonnUconstr))
                {SELL}
                purchase = {b}1*minsho
                {sellflag=1}
                end if
                if ((MkttonnUconstr<PrevMkttonnUConstr) and usedprice > bias*newprice)
                purchase = minsho
                {sellflag=1}
                end if

{if     (usedprice>bias*newprice) purchase = {b}1*minsho
                end if}

                if ((MkttonnUconstr>PrevMkttonnUConstr) and
                usedprice < (2-bias)*newprice)
```

```
                    {buy }
                    purchase = b * maxsho


                    end if
                    if (newprice < bias* usedprice)  develop = b*maxsho
end if

                    if ((MkttonnUconstr-PrevMkttonnUConstr) < (PrevMkttonnUConstr-
OldMkttonnUconstr)) develop = 0
end if
                    if (usedprice < prevusedprice) develop = 0
end if

                    if newprice>usedprice develop = 0 end if

                    {END DEC RULE 2}

          end if

put purchase into makecell(col(),2)
put develop into makecell(col(),3)

PUT MINSHO INTO makecell(col(),6)
PUT MAXSHO INTO makecell(col(),12)

select range makecell(col()+1,1)

achievedprofit = indirect(makecell(col()-1,24))


{COUNT SCRIPT}

Priceelast = c74

EPOR = c75

REconv = c76

gamename = c77

{nonlinearity EPOR calc}

cin = makecell((col()-2),65)
in = (indirect(cin))/REconv

if in > 5.  out = 4
elseif in > 4.5 out = (.1*in +3.5)
elseif in > 3.5 out = (.25*in + 2.825)
elseif in > 3  out = (.3*in + 2.635)
elseif in > 2.5  out = (.6*in + 1.575)
elseif in > 2. out = (.9*in +.675)
elseif in > 1.5  out = (1.15*in +.075)
elseif in > .5  out = (1.35*in - .35)
elseif in >0  out = (.65*in)
else out = 0
end if
put out^EPOR into makecell((col()-1),64)
```

```
{recalc column}
put r9c3 into makecell(col()-1,9)
put indirect(makecell(col()-2,78)) into makecell(col()-1,30)

if indirect(makecell(col()-2,34)) < 0

put indirect(makecell(col()-2,34)) * 1.1 into makecell(col()-1,34)
else
put indirect(makecell(col()-2,34)) * 1.05 into makecell(col()-1,34)
end if

put indirect(makecell(col()-2,4))/r44c3 into makecell(col()-1,36)

put indirect(makecell(col()-2,29))*r46c3 into makecell (col()-1,80)

put indirect(makecell(col()-2,4))-indirect(makecell(col()-1,36))
+indirect(makecell(col()-1,3)) into  makecell((col()-1),4)

put indirect(makecell (col()-2,5))/r33c3
 into  makecell((col()-1),52)

put indirect(makecell (col()-2,5))
+indirect(makecell(col()-1,36))
+indirect(makecell((col()-1),2))
-indirect(makecell(col()-1,52)) into  makecell((col()-1),5)

put (indirect (makecell(col()-1,30))/indirect(makecell(col()-1,80)))
into  makecell(col()-1,62)

put ((r79c3/(REconv) + (indirect(makecell(col()-1,62))-(r79c3/(REconv)*.2040353))^(Priceelast)))
*REconv
into makecell(col()-1,7)

put (indirect(makecell (col()-1,7))- indirect(makecell(col()-2,65)))/2
into  makecell((col()-1),32)

put (indirect(makecell (col()-2,65))+indirect(makecell(col()-1,32)))
 into  makecell((col()-1),65)

{Capacity utilization multiplier computation}
in = indirect(makecell(col()-1,62))
if in > 2 out = 1.19
elseif in < 1 out = in
elseif in < 1.2 out = .5*in + .5
elseif in < 1.4 out = .25*in + .8
elseif in < 1.6 out = .125*in + .975
elseif in < 1.8 out = .065*in + 1.071
else out = .01*in + 1.17
end if

put out into makecell(col()-1,68)

put indirect (makecell(col()-1,68))*r46c3 into makecell(col()-1,11)

{Stress factor computation}
in = indirect (makecell(col()-1,68))
if in > 1.2 out = 4.1
```

```
elseif in < .8 out = .75
elseif in < 1 out = 1.25 * in -.25
elseif in < 1.04 out = 2.5 * in -1.5
elseif in < 1.08 out = 5 * in -4.1
elseif in < 1.12 out = 10 * in -9.5
elseif in < 1.16 out = 20 * in -20.7
else out = 40 * in -43.9
end if
put out into makecell(col()-1,58)


put indirect(makecell(col()-1,58))*r79c3
 into makecell(col()-1,10)

put (indirect(makecell (col()-2,71)) /r44c3 )
into makecell(col()-1,35)

put (indirect(makecell (col()-2,70))/r33c3)
 into makecell(col()-1,51)

put indirect(makecell(col()-1,51))*r44c3 into makecell(col()-1,69)

put (indirect(makecell(col()-1,69))-indirect(makecell(col()-2,71)))/r44c3
into makecell(col()-1,67)

put indirect(makecell (col()-2,70)) + indirect(makecell(col()-1,35))
 - indirect(makecell(col()-1,51)) - indirect(makecell(col()-1,2))
 into makecell((col()-1),70)

put (indirect(makecell (col()-2,70))*indirect(makecell(col()-1,64))
*indirect(makecell (col()-1,68)))
 into makecell((col()-1),48)

put (indirect(makecell((col()-1),48))-indirect(makecell(col()-2,70)))/r44c3
into makecell((col()-1),39)

put max(0,(indirect(makecell((col()-1),39))+indirect(makecell(col()-1,51))
+(.75*indirect(makecell (col()-1,67)))))
into makecell((col()-1),72)

put indirect(makecell (col()-2,71)) + indirect(makecell(col()-1,72))
-indirect(makecell(col()-1,35)) into makecell((col()-1),71)

put (indirect(makecell((col()-1),4))+indirect(makecell((col()-1),71)))
into makecell((col()-1),28)

put (indirect(makecell((col()-1),5))+indirect(makecell((col()-1),70)))
 into makecell((col()-1),29)

put (indirect(makecell((col()-1),29))*indirect(makecell((col()-1),11)))
into makecell((col()-1),56)


put indirect(makecell((col()-1),7))*indirect(makecell(col()-1,5))
        *indirect(makecell(col()-1,11))
into makecell((col()-1),13)

put
```

```
        indirect(makecell(col()-1,5))* indirect(makecell((col()-1),11))*
        indirect(makecell((col()-1),10))
into makecell((col()-1),14)


put indirect(makecell(col()-1,13)) - indirect(makecell((col()-1),14))
into makecell((col()-1),15)

put indirect(makecell(col()-2,25)) * .05
into makecell((col()-1),17)

put indirect(makecell(col()-2,26)) * .1
into makecell((col()-1),18)

put indirect(makecell(col()-2,26))/r33c3
into makecell((col()-1),19)


if indirect(makecell(col()-1,2)) < 0

put -indirect (makecell(col()-2,8))*indirect(makecell(col()-1,2))
        + (indirect(makecell((col()-1),2))/indirect(makecell((col()-2),5)))*
                indirect(makecell((col()-2),26))
into makecell((col()-1),20)

else dummy = 0

put dummy into makecell((col()-1),20)

end if

if indirect(makecell(col()-1,2)) >= 0

put indirect(makecell((col()-1),2))*indirect(makecell((col()-2),8)) +
        indirect(makecell((col()-1),36))*indirect(makecell((col()-1),9))
        - indirect(makecell((col()-2),26))/r33c3
        + indirect(makecell((col()-2),26))
        into makecell((col()-1),26)

else

put indirect(makecell((col()-1),2))/indirect(makecell((col()-2),5))*indirect(makecell((col()-2),26))
        +indirect(makecell((col()-1),36))*indirect(makecell((col()-1),9))
        - indirect(makecell((col()-2),26))/r33c3
        + indirect(makecell((col()-2),26))
        into makecell((col()-1),26)

end if

put (abs(indirect(makecell((col()-1),2)))*indirect(makecell((col()-2),8)) +
        indirect(makecell((col()-1),3))*indirect(makecell((col()-1),9)))*.1
into makecell((col()-1),21)

put indirect(makecell((col()-1),17))-indirect(makecell((col()-1),18))
- indirect(makecell((col()-1),19))+indirect(makecell((col()-1),20))
- indirect(makecell((col()-1),21))
into makecell((col()-1),22)
```

```
put indirect(makecell((col()-1),22))+indirect(makecell((col()-1),15))
into makecell((col()-1),24)


put indirect(makecell((col()-1),24)) + indirect(makecell((col()-2),25))
into makecell((col()-1),25)


Put indirect(makecell((col()-2),41))
        +(indirect(makecell((col()-1),15))/indirect(makecell((col()-1),5))
        -indirect(makecell((col()-2),41)))/2
                into makecell((col()-1),41)

Put indirect(makecell((col()-1),41))*r42c3 into
        makecell((col()-1),8)


{bankbal= INDIRECT(MAKECELL(col()-1,25))
liqloss = INDIRECT(MAKECELL(col()-1,34))
}

simprofit = INDIRECT(MAKECELL(col()-1,24))
put achievedprofit into makecell(col()-1,87)

put simprofit-achievedprofit into makecell(col()-1,86)

{select range (MAKECELL(col()-1,25))
copy}
END FOR
totprofit= sum(r86c3..r86c46)
put totprofit into r23c45

save as "ClubMac :BeB:Research Games:Game
results/91:"andSUBJECTNBRand":"andGAMENBRand"c"
close
close

go to window "dseries"

if gamenbr = 5 put totprofit into makecell(sellstep/5+1,(subjectnbr))
else
put totprofit into makecell (sellstep/5+10,(subjectnbr))
end if
sellflag = 0
{ THIS IS FOR AGGRAGATE SIM RES;
PUT OUT OF USE 6/15 91 BEB}
{
open "sim res"
PROFCOL = PROFCOL+1
PUT {BANK BALANCE} bankbal INTO
MAKECELL (PROFCOL, 10)
LOSSCOL = LOSSCOL + 1

PUT {LIQUIDATION LOSSES} liqloss INTO
MAKECELL (LOSSCOL, 11)
PUT gamenbr INTO MAKECELL (LOSSCOL, 1)
PUT B INTO MAKECELL(LOSSCOL, 2)
{PUT BIAS INTO MAKECELL(LOSSCOL, 3)}
```

```
PUT DECRULE INTO MAKECELL(LOSSCOL., 4)
}
end for
END FOR

END FOR
END FOR
end for

end for
```

Listing A.11: The benchmark script

Appendix 4: Preparatory information read by subjects

Article A.12 shows the real estate article

New York Times, Dec 12 1989, page 1

Northeast Banks Face Heavy Losses on Problem Loans

by Michael Quint

Troubled real estate loans are causing heavy losses for bankers in the Northeast.

With office vacancy rates now reaching 25 to 30 percent in places like central New Jersey and Stamford, Conn., and with many condominium developments still only half filled after two years or so on the market, a growing number of developers are having difficulty paying off their loans. Many bankers now concede that they may never collect the full value of their loans to developers.

Among the 10 states whose banks show the highest increase in bad real estate loans, all but two are in the Northeast, with New York, New Jersey and Connecticut all on the list. Banks in those 10 states had increases of at least 55 percent in their delinquent real estate loans from the end of 1988 to mid-1989, according to a report by Sheshunoff Information Services Inc. of Austin, Tex. The report was based on bank reports to the Government.

No Texas-Size Crisis Seen

But despite the problems, analysts do not expect a crisis of the magnitude that led to the failure of many banks and savings and loan associations in energy-producing states. The Northeast is not as devastated as Texas was in the mid-1980's, and banks in the region have a bigger financial cushion than many of the savings and loans that collapsed in Texas.

The difficulties with real estate loans are not confined to the Northeast, where banks are the primary lenders for commercial real estate. Banks in Arizona and Florida are among others with major problems. The real estate troubles of some banks in the energy states, though are beginning to recede.

Aftermath of '87 Crash

"The United States enjoyed an unprecedented building boom during the 1980's, with overbuilding and an oversupply of office space, shopping centers, hotels--just about everything, all over the country," said David Shulman, director of real estate research at Salomon Brothers. "Texas was not an anomaly, it was a precursor."

Where there is trouble the cause is most often overzealous lending by institutions seeking new markets as opportunities to lend to businesses have dwindled and foreign lending has often proved to be a money loser, analysts say. The eagerness to lend has resulted in too many new buildings flooding the market in many areas.

But the problem in the Northeast has also been caused by a weakening in the region's economy, resulting in part from a slowing in growth of military spending, a shrinking of computer firms in the Boston area and cuts by Wall Street firms after the stock market crash in October 1987.

As banks in the Northeast try to stem growing losses by cutting back on new lending, the region's economic woes are increasing. Economists studying the Northeast predict modest rises in unemployment and more delinquencies on loans of all kinds.

Still analysts do not think losses from the $19.7 billion in delinquent real estate loans nationally, about a third of which are held by Northeastern banks, are high enough to touch off a round of collapses comparable to what the savings and loan industry has experienced. Banks around the country have more than $200 billion of shareholder capital and about $50 billion of reserves for loan losses.

Nor do banks have the extensive ties to real estate interests that existed among savings institutions, many of which were owned and operated by developers. Because banks are not allowed to act as developers and own real estate, except to dispose of foreclosed properties, there is less room for the fraudulent lending practices that were so costly to savings and loans, analysts say.

Still, the problem has become a major drain for banks, particularly those in the Northeast. The Bank of New England recently said it expected a large loss in the fourth quarter because of problems in real estate lending. And the Bank of Boston said its third-quarter loss of $125 million was mostly a result of a $370 million increase in its reserve for loan losses, mostly for real estate.

Rising Share of Bad Loans

The Bank of Boston said its delinquent real estate loans had risen by $185 million, to $1.1 billion, or about 13 percent of all its real estate loans. Nationally, bad real estate loans amount to just under 3 percent of all real estate loans held by banks.

Both Bank of New England and Bank of Boston, the two largest banking companies in New England, were recently required by Government regulators to sign agreements to improve their procedures for limiting their risk on real estate loans.

Savings banks in the Northeast have also been hit hard by losses on real estate loans, said Don J. Fauth, an analyst at the First Albany Corporation, a securities firm. He cited the Government takeover last week of City Federal Savings Bank, the largest savings institution in New Jersey, and said several other savings banks in the region would not survive in their present form. He noted that after the savings banks raised more capital earlier in the decade by issuing stock to investors, they tried to increase profits by turning to more risky lending on construction projects rather than their traditional home mortgage lending business.

Effect on Citibank

Among the nation's largest banks, Citibank has also reported a big rise in bad real estate loans in recent quarters and expects more.

Insurance companies are also big lenders on commercial real estate, but generally their loans come only after a project is completed and is at least half occupied. While their profits may be affected by increased vacancies and smaller than-expected rental-income increases, analysts say they are not experiencing the kind of loan losses that banks are.

L. William Seidman, chairman of the Federal Deposit Insurance Corporation, which oversees the health of the nation's banks, predicted last week that real estate losses would increase further in the current quarter. Certain Northeastern areas "have some of the highest commercial vacancy rates in the country," he noted.

What other regions do not share with the energy-producing states, at least so far, is a drop in employment sharp enough to lead to a big rise in defaults on home mortgages. But recent declines in home prices and slower sales of homes in the Northeast and other areas are making bankers worry about bigger problems in the future. In the New York City are, for example, prices have fallen 5.1 percent in the last year, compared with a 14.4 percent gain in the two years ending in 1988.

Dangers of Defaults

When a period of rapid increases in home prices, like ones that occurred in the Northeast, is followed by steep drops, mortgage defaults can be devastating to bankers because the value of the abandoned home can be less than the amount due on the defaulted mortgage.

"We have not seen unusual losses on single-family mortgages so far but if the markets continue to worsen, you have to worry," said Donald McCormick, chairman of the Howard Savings Bank in Livingston, NJ.

But for now, problems with loss for commercial property account for almost all of the banking industry's real estate problems. Loans for construction of residential and commercial properties, and other loans for commercial properties, totaled $ 33 billion at the end of September, accounting for nearly 19 percent of loans and leases held by banks, from about 6 percent in 1982.

It Looked Good at the Time

In Garden City, LI, for example, executives at Howard Savings thought they were being prudent and cautious when they provided half of $100 million loan to finance a development of 300 high-priced condominiums. The bank verified that buyers had made commitments to acquire about two-thirds of the units. What the bank did not foresee was that many buyers are now threatening to walk away from their commitments and trying to recover their deposits which averaged about $50,000.

Today, fewer than two-thirds of the units in the complex have been sold and prospects for that project and others are so bleak that Howard Savings earlier this week increased reserve for losses on real estate loans by $70 million, resulting in a loss for the year of about $45 million.

Article A.12: The real estate article

Article A.13 is the oil tanker article

New York Times, December 5 1989, page D1

For Supertankers, Super Profits

by Agis Salpukas

About five years ago, with the supertanker business in the doldrums, the Loews Corporation bought seven large tankers for about $5 million each--basically, their scrap value. Now those ships are carrying crude oil between the Middle East and the Gulf Coast for up to $3 million for just one voyage, operating at considerable profit. And if the Tisch family, which controls Loews, ever decides to sell them, it could get close to $40 million apiece.

The Tisches' success is but one indication that after more than a decade of decline and shakeout, the supertanker business has been showing signs of strong recovery.

The turnaround can be attributed partly to a substantial shrinking earlier in the decade of a fleet with so much overcapacity that shipping rates were plummeting. While the number of ships was declining, the need for them began to increase sharply as demand for oil rose and OPEC members changed their strategy from seeking to cut production and raise prices to increasing production as a way of winning market share.

Until recently, the growing demand could be met by bringing back ships that had been mothballed. But now most of those ships are in service. And few new ships are being built, analysts, say, mainly because rates are still not high enough and shipowners and investors are still uncertain of whether the recovery will be sustained. Also, building costs are high and oil companies are still squeamish about the Exxon Valdez disaster eight months ago.

So attractive has the business become that investors, who once bought shipping companies simply for the potential increase in value of assets like tankers, are now entering the business for its rising operating profits as well. Two weeks ago, for example, the Belzberg family of Canada acquired Marine Transport Lines, the third-largest shipping company and one of the oldest in the United States.

The turnaround has made a few daring investors look very smart. When the Tisch family bought the seven tankers, shipowners and banks that had repossessed big ships were only too glad to unload them. There was so much skepticism about the merits of the deal that a public offering by Loews failed to raise enough money, and the company had to dig into its own reserves.

But with the recovery, Loews was able to sell one smaller tanker for $10 million last year. The six larger ones it still owns may be worth $35 million to $40 million each.

The Surge in Rates

The rapid rise in value is not as astonishing when the increase in shipping rates is considered.

In 1985 rates for the biggest tankers hovered at an average of $5000 a day, forcing many shipowners to file for bankruptcy, said James L. Winchester, a shipping analyst with Mabon, Nugent and Company. Banks often resorted to seizing the ships and selling them for scrap value, seeing little prospect of a quick turnaround.

But by last year, rates had reached an average of $16,000 a day, Mr. Winchester said. This year they have declined to $13,000 to $14,000 a day, although he thinks they will rise again soon. At some periods this year, when there was a shortage of capacity in the Middle East, rates reached $25,000 to $30,000 a day.

Yet current rates are still far below levels that would spur a boom in orders for new tankers, even though many ships are beginning to age and will soon need to be replaced.

Mr. Winchester estimated that daily rates of $35,000 to $40,000 are needed just to pay off the financing over 20 years.

Nonetheless, new orders for tankers, which cost about $85 million for a large, 280,000-ton vessel, have been coming at a steady rate. And because shipbuilding capacity has shrunk in the downturn that began in the early 1970's, the remaining yards--mostly in South Korea and Japan--are booked for the next two years.

Rates have risen enough that owners of ships in good condition have begun reaping substantial profits from operating them instead of selling them.

To be sure, the volatile world oil situation makes the supertanker business risky. But, betting that the recovery will last, the Tisches are seeking to keep a 49 percent interest in the tankers they are selling, to share in the operating profits.

"In the past five years the big money was made in owning, but now the money will be made from the actual operating income," said James S. Tisch, an executive vice president at Loews and the son of Laurence A. Tisch, the chairman of Loews and president of CBS Inc.

The industry is now attracting other investors interested in buying companies not as an asset play but because they will be profitable on an operating basis.

Belief that the upturn will last led the Belzbergs to pay $128.8 million for Marine Transport Lines, which is based in Secaucus, N.J. The company operates three large oil tankers and owns and operates smaller tankers that carry petroleum products and chemicals. It also has cargo vessels.

The buyout was led by Richard T. duMoulin, the former chief operating officer of the OMI Corporation, a tanker company. He will be the new chairman and the chief executive of Marine Transport.

"The industry had 13 terrible years from 1974 to 1987," Mr. duMoulin said. He said the high cost of labor and materials at shipbuilding yards makes it very expensive to buy new tonnage. Worldwide demand has been rising by 3 percent a year since 1985. Meanwhile, with the Organization of Petroleum Exporting Countries stepping up production, OPEC exports have surged by 7 percent during that period. Because the oil from the Middle East must be moved to distant markets, in Europe, Asia and the United States, the tanker industry plays an important role.

Prospects are that the OPEC production rise will continue as members of the group further increase exploration and production. During the first eight months of this year, OPEC oil shipments were up 14 percent from the period a year earlier.

While there has been a strong recovery in the demand for tankers, the supply, which began to shrink in 1978, is still far below levels earlier in the decade, said Sally H. Smith, a shipping analyst at Alex, Brown and Sons.

Mr. Winchester estimated that about 40 percent of tanker tonnage was scrapped between 1983 and 1986.

For a time, increasing demand was met by bringing back laid-up ships. The active fleet increased by 13 percent from 1985 to 1988. Now most tankers are active, and few new ships are being built.

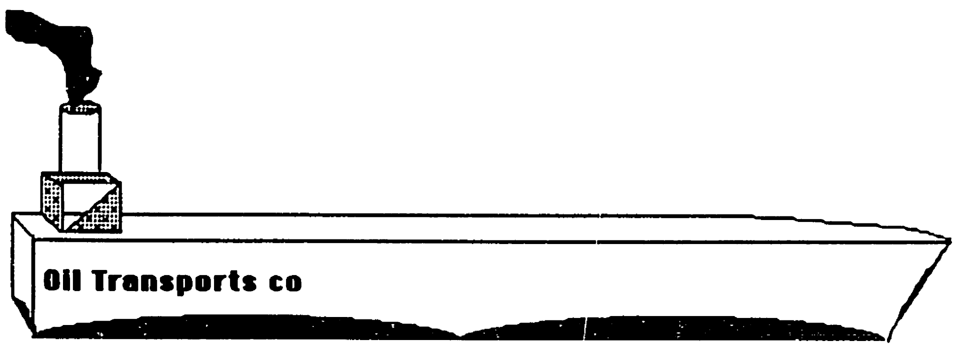
"The outlook is for slower growth in the active fleet in future years," Ms. Smith said.

Article A.13: The oil tanker article

Article A.14 shows the oil tanker market description (the real estate description was provided in chapter 4.)

The Market for Oil Tanker Transportation

A participatory simulation game



Can one understand this market that often has been described as "the biggest poker game in the world" ? In a world where transportation demand is hard to predict, many investors have lost their fortunes and those of some banks in the process. In fact, some banks lost so much in the early eighties that they decided to pull out of tanker lending completely.

This game purports to highlight aspects of tanker investments and operations that get little attention in everyday asset management. Underlying the game is a model that has been constructed using data from reliable sources for the entire post-war era. Of course, the model is not the real thing; like any map it serves to give an overview into issues that are hard to detect in the real world. In addition, it puts you into the presidentship of a ship-owning firm. The firm is initially small and you control about one percent of the market, but with your understanding and shrewd piloting, it can grow tremendously. The game gives you financial resources and opportunities for making big money, ...and for going bankrupt.

Figure 1 portrays the overall structure of the game. You are a player operating in a market for oil transportation services, buying and selling secondhand tonnage as well as ordering new tonnage from shipyards.

All tonnage is identical and so your operating costs and those in the market are identical. Likewise, your capacity utilization is the same as the market's. The computer simulates your competition in a very straightforward manner; when expected profits go up, the market invests in new tonnage and at the same time secondhand prices increase. Conversely, no new orders are placed when poor market prospects prevail.

There are some real investment and transaction limitations in this market... You can only invest as long as you have financial leverage to do so. In addition, if you should grow to control the market, you can never sell or buy more than the equivalent of ten per cent of your competitors' ships. Buying up or selling more space than that would destroy the market.

Below you see the overall structure of the game.

PLAYER	MARKET
DECISIONS Buy/sell Tankers Build new Tankers *RESULTS* Tankers Debt Cash Financial Revenues and Expenditures Operating Revenues and Costs	Demand Supply of Tankers Capacity Utilization Spot Rate Tankers Under Construction Newbuilding Price Secondhand Price

Figure 1.1: Structural Overview of the Oil Tanker game

Below you see the decision making screen as it appears in the game.

	Results for Year	**1989**	
My Ships	**Secondhand Order**	**0**	Tankers/year
	New Order	**8**	Tankers/year
	Ships o/ord (next year: 67%)	12	Tankers
	My Ships	120	Tankers
Unit Costs	Spot Rate	1.00	$ m/Tkr/year
	Secondhand Price	4.20	$ mill/Tanker
	Newbuilding Price	3.89	$ mill/Tanker
	Variable Unit Cost	0.40	$ m/Tkr/year
	Capacity Utilization	0.90	Fraction
Operations	Operating Revenue	108	$ mill/year
	Operating Costs	43	$ mill/year
	Operating Profit	**65**	$ mill/year
Capital	Interest on Bank Balance (5%)	5	$ mill/year
	Interest paid on Loans (10%)	47	$ mill/year
	Deprec'n (Demol'n) (6.7%)	31	$ mill/year
	Appreciation realized	0	$ mill/year
	Transaction fees (10%)	0	$ mill/year
	Net Financial Gain	**-73**	$ mill/year
	Net Profit	**-8**	$ mill/year
Balance Sheet	Bank Balance	93	$ mill
	Loans	467	$ mill
Market Ships	Market Tonnage on Order	1200	Tankers
	Market Tonnage	12000	Tankers
	Demand	10800	Tankers

Figure 1.2: The decision making screen

2. Running the game

The game requires you to make decisions about tonnage transactions. The unit you transact in is "Tankers". Its price about $4.2 million dollars, and it brings in $1 million per year in Operating Revenues. Your initial endowment is 120 Tankers (bought at $ 3.89 million each and financed 100 %. Thus your initial loans are $467 million.)

Measurement Unit

My Ships	**Secondhand Order**	**0**	Tankers/year
	New Order	**8**	Tankers/year
	Ships o/ord (next year: 67%)	12	Tankers
	My Ships	120	Tankers

You start the game by clicking the mouse on the "Make Decisions" button in the bottom of your screen (see figure 2.1). Having done that, a new scrolling window (figure 2.2) appears in the bottom left corner of the screen. You must enter your forecast of secondhand tanker price in 1992.

Figure 2.1 Make Decisions

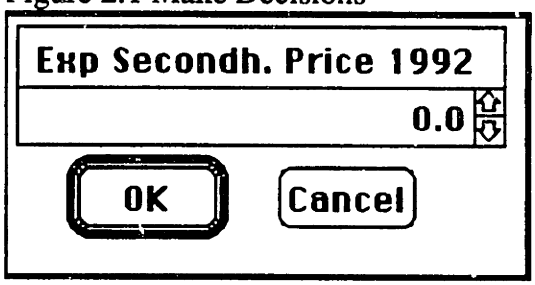


Figure 2.2: Expected tanker price

You then enter the expected Spot Rate, i.e. the going price for leasing a tanker for one year.

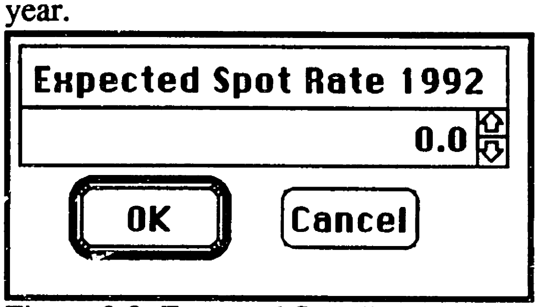


Figure 2.3: Expected Spot Rate

You next enter your decisions of how many existing ships to buy or sell. Positive numbers indicate purchases and negative numbers indiacte thta you sell your own tankers. You can either scroll using the mouse or type in your decisions of Existing tonnage to buy or sell. Upward scrolling translates into purchasing. Downward scrolling translates into selling. If you attempt to type decisions exceeding your financial limits or attempt to buy or sell too much tonnage, you will be prevented from doing so. When you invest in Secondhand Tankers, you retain the one year lease contract on the boat.

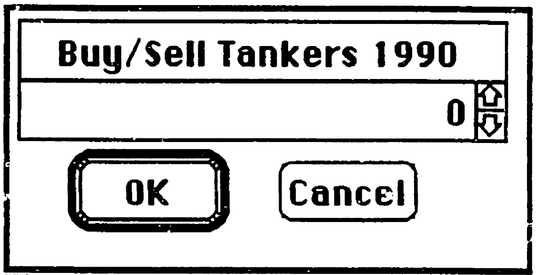


Figure 2.4: Investments in secondhand tonnage

After clicking on "OK" or using the "Return" keyboard button, a new window will appear in the same location. You are then ready to enter your orders to shipyards in the same way. These orders are invariably turned into completion. "Cancel" at any time and you are back to clicking the "Make Decisions" button again. Figure 2.5 shows this dialog button.

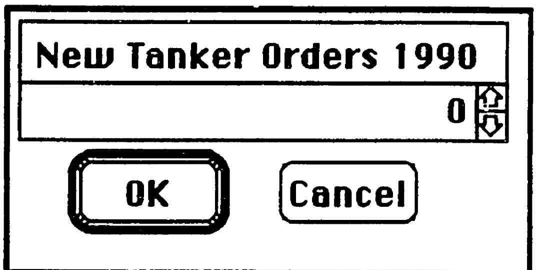


Figure 2.5: Ordering of New Tankers from Shipyards.

Investments are 100 % lender financed, but the lender requires at least 10 % guarantee from the investor. This guarantee is based on your Bank Balance and hidden reserves. The latter source derives from the difference between Book Value and liquidation value of your Tankers. Thus you are strongly limited in your transactions by your cash and hidden reserves.

The value of existing tonnage depends only on expected operating profits no matter how many existingtankers are bought and sold. However, there is a transaction limit; In one year no more than 10 % of competitor tonnage can be bought in one single year. Likewise, competitors will never buy more than a 10 % increase in existing tonnage. However, they can order tankers in addition the these percentages, ... and so can you.

There are two ways of making profits, either speculation (buying low and selling high) and operations (letting operating income cover more than capital costs and management expenses).

Article A.14: The oil tanker market description

Appendix 5: Results

Table A.15 shows subjects, results and conditions

subjnb	-raw1	-raw2	olddel	-pm1	-pm2	cont1	cont2	struct	struct2	dem1	dem2	deltac	deltast
1	-0.36	-0.03	0.18	-0.09	-0.10	2	2	1	1	1	2	1	1
2	0.06	0.51	0.69	-0.04	0.00	2	2	1	1	2	1	1	1
3	0.14	0.37	0.46	-0.12	-0.20	2	2	1	2	1	2	1	2
4	-0.16	0.11	0.25	0.02	-0.11	2	2	1	2	2	1	1	2
5	-1.93	-0.15	*****	-0.09	-0.13	2	2	2	1	1	2	1	2
6	0.22	0.11	0.06	-0.18	-0.13	2	2	2	1	2	1	1	2
7	0.10	0.12	0.12	-0.05	-0.12	2	2	2	2	1	2	1	1
8	0.22	0.17	0.15	-0.16	-0.12	2	2	2	2	2	1	1	1
9	0.02	0.14	0.20	-0.05	0.00	2	1	1	1	1	2	2	1
10	-0.08	0.42	0.64	0.05	-0.35	2	2	1	2	2	1	1	2
11	0.05	2.48	3.02	-0.05	-1.02	2	1	1	2	1	2	2	2
12	0.15	0.18	0.19	-0.15	-0.08	2	1	1	1	2	1	2	1
13	-0.03	0.33	0.48	0.02	-0.13	2	1	2	1	1	2	2	2
14	0.07	-0.11	-0.20	-0.12	-0.06	2	1	2	1	2	1	2	2
15	2.28	0.32	-0.11	-0.31	-0.91	2	1	2	2	1	2	2	1
16	1.33	-6.75	-4.39	-0.36	-0.37	2	1	2	2	2	1	2	1
17	0.01	0.22	0.32	-0.07	-0.16	1	1	1	1	1	1	1	1
18	-0.31	0.07	0.28	-0.02	-0.10	1	1	1	1	1	1	1	1
19	0.38	3.13	3.62	-0.79	-1.59	1	1	2	2	2	2	1	1
20	-0.19	0.13	0.29	0.01	-0.03	1	1	1	2	2	1	1	2
21	-3.48	0.05	-2.42	-0.25	0.13	1	1	2	1	1	2	1	2
22	-0.20	0.49	0.80	-0.17	-0.16	1	1	2	1	2	1	1	2
23	-0.07	0.01	0.04	0.01	-0.03	1	2	1	1	1	2	2	1
24	0.13	0.48	0.61	-0.24	-0.44	1	1	2	2	1	1	1	1
25	-0.12	0.49	0.75	-0.02	-0.11	1	2	1	1	1	2	2	1
26	0.01	-0.07	-0.11	-0.06	-0.07	2	1	1	1	1	2	2	1
27	0.11	0.23	0.29	-0.14	-0.17	1	2	1	2	1	2	2	2
28	0.14	0.18	0.20	-0.09	-0.14	1	2	1	1	2	2	2	1
29	-0.09	0.32	0.51	0.02	-0.16	2	1	1	2	2	1	2	2
30	-0.22	0.09	0.26	-0.86	-0.13	1	2	2	1	2	1	2	2
31	0.22	0.14	0.10	-0.27	-0.12	1	2	2	2	1	2	2	1
32	-0.16	-0.30	-0.39	-1.18	-0.09	1	2	2	2	2	1	2	1
33	0.13	0.79	1.01	-0.08	-0.37	2	2	1	1	2	1	1	1
34	-0.03	0.08	0.14	-0.01	-0.09	2	2	1	1	2	1	1	1
37	-0.71	-0.15	0.34	-0.01	-0.07	2	2	2	1	1	2	1	2
42	0.01	0.56	0.77	-0.09	-0.24	2	1	1	1	2	1	2	1
55	0.02	0.07	0.10	-0.01	-0.24	1	1	2	2	1	2	1	1
56	0.31	-7.38	-5.87	-0.11	-0.05	1	1	2	2	2	1	1	1
57	0.05	0.03	0.01	-0.04	0.00	1	2	1	1	1	2	2	1
61	-0.13	0.12	0.25	-0.07	-0.09	1	2	2	1	1	2	2	2
62	0.33	-0.03	-0.19	.	0.00	1	2	2	1	2	1	2	2

Table A.15: Subjects and conditions

Table A.16 shows results

Resume of results

q1	q2	q3	q4	q12	q34	q23	q14	ugfin	ugeco	ugce	ugcs	ugacc	ugsd	ugbs	ugg
0.47	0.62	0.559	0.588	0.500	0.833	0.278	0.833	2	2	2	2	2	2	2	3
0.29	0.41	0.382	0.471	0.444	0.833	0.389	0.833	2	4	3	3	2	3	2	6
0.41	0.44	0.441	0.382	0.556	0.889	0.556	1.000	2	2	2	6	2	2	3	5
0.32	.	.	.	.	.	.	.	2	2	2	3	2	2	2	6
0.38	0.44	0.471	0.500	0.722	0.833	0.667	0.667	2	4	2	2	3	3	2	4
0.44	0.50	0.412	0.441	0.778	0.722	0.667	0.778	3	3	6	5	3	2	2	1
0.35	0.35	0.324	0.382	0.667	0.833	0.556	0.667	2	3	2	2	2	2	2	3
0.68	0.79	0.706	0.706	0.833	0.722	0.611	0.889	3	6	2	2	4	2	2	5
0.50	0.71	0.618	0.706	0.500	0.778	0.556	0.722	2	5	3	6	3	2	2	6
0.29	0.32	0.324	0.294	0.500	0.278	0.389	0.667	2	4	3	4	2	2	2	6
0.41	0.62	0.618	0.676	0.333	0.667	0.444	0.611	3	4	4	3	3	3	2	4
0.35	0.29	0.412	0.559	0.389	0.500	0.222	0.167	3	4	2	3	4	2	3	6
0.50	0.41	0.618	0.559	0.444	0.722	0.667	0.611	2	3	3	4	2	2	2	5
0.53	0.68	0.382	0.676	0.667	0.444	0.667	0.278	3	6	2	5	2	2	2	4
0.50	0.65	0.471	0.412	0.556	0.611	0.389	0.722	2	4	2	4	2	2	2	5
0.65	0.65	0.441	0.382	0.611	0.389	0.389	0.667	2	4	3	6	3	2	2	5
0.50	0.29	0.324	0.676	0.056	0.222	0.556	0.722	2	2	3	6	3	2	2	6
0.62	0.65	0.647	0.618	0.778	0.889	0.778	0.944	4	4	2	3	6	2	3	6
0.50	0.74	0.559	0.765	0.556	0.667	0.500	0.611	5	5	2	6	4	2	4	4
0.38	0.41	0.471	0.412	0.722	0.556	0.444	0.611	3	3	3	2	2	2	2	5
0.35	0.68	.	.	.	.	.	.	2	3	2	5	2	2	3	4
0.62	0.21	0.676	0.676	0.111	1.000	0.778	0.167	6	4	2	3	6	2	4	6
0.38	0.65	0.412	0.412	0.500	0.944	0.167	0.333	3	4	2	3	4	2	2	4
0.50	0.53	0.559	0.559	0.500	0.944	0.500	0.889	2	4	2	2	3	2	2	2
0.68	0.59	0.676	0.647	0.611	0.833	0.444	0.722	2	3	2	3	4	4	2	4
0.50	0.44	0.529	0.471	0.611	0.944	0.667	0.722	2	3	2	3	4	2	2	5
0.59	0.32	0.529	0.706	0.333	0.722	0.778	0.333	3	4	3	3	2	2	2	2
0.47	0.29	0.471	0.559	0.167	0.889	0.611	0.111	6	6	5	3	6	6	3	6
0.56	0.62	0.618	0.823	0.611	0.667	0.500	0.778	2	4	2	3	4	2	2	4
0.44	0.65	0.471	0.618	0.667	0.778	0.556	0.778	3	5	6	6	4	6	3	6
0.38	0.56	0.500	0.794	0.222	0.667	0.389	0.667	4	4	3	5	2	2	2	4
0.32	0.74	0.588	0.765	0.389	0.778	0.333	0.778	4	5	3	6	5	3	3	4
0.35	0.41	0.353	0.324	0.667	0.722	0.444	0.611	3	3	3	5	3	2	2	4
0.27	0.32	0.441	0.353	0.611	0.778	0.556	0.778	4	6	2	6	4	2	4	5
0.38	0.32	0.471	0.353	0.611	0.778	0.722	0.778	2	4	2	6	3	2	2	5
0.35	0.41	0.735	0.735	0.556	0.778	0.167	0.389	2	3	3	2	2	2	2	4
0.56	0.68	0.618	0.735	0.500	0.778	0.556	0.778	2	6	2	4	4	2	2	4
0.44	0.38	0.353	0.294	0.278	0.333	0.444	0.611	2	2	3	4	2	2	2	4
0.50	0.47	0.265	0.294	0.500	0.778	0.278	0.278	3	4	2	3	3	2	2	5
0.71	0.62	0.471	0.529	0.722	0.611	0.278	0.278	2	6	2	4	4	2	2	4
0.56	0.65	0.412	0.441	0.833	0.778	0.111	0.278	3	6	2	3	3	2	3	4

Table A.16.a: Questionnaire scores and academic undergraduate background

gfin	gecon	gce	gcs	gacc	gsd	gbs	gg	wefin	wered	wereo	wewes	wesho	wece	wesd	wet
3	4	2	2	3	3	4	6	2	2	2	2	2	2	2	6
2	3	6	2	2	6	2	6	2	2	2	2	2	2	2	3
2	3	2	4	3	2	3	4	2	2	2	2	2	2	2	3
3	3	2	3	3	2	3	6	2	2	3	2	2	2	2	3
2	2	2	2	2	1	1	1	1	1	1	1	1	1	2	2
4	5	2	4	3	4	4	5	2	2	2	2	2	3	2	3
3	4	2	3	3	2	4	5	2	2	2	2	2	2	2	5
3	2	2	2	2	2	2	5	6	6	2	2	2	2	2	4
3	5	4	2	3	2	4	6	2	2	2	2	2	2	2	3
2	3	2	2	2	3	3	6	2	2	2	2	2	2	2	3
1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2
1	1	1	1	1	1	4	1	1	1	1	1	1	1	1	1
3	2	2	2	2	2	2	4	2	3	2	2	2	2	2	4
3	6	2	5	2	2	2	4	4	2	2	2	2	2	2	3
3	3	3	2	2	4	2	5	2	2	2	2	2	2	2	2
3	4	2	6	3	2	4	6	2	2	2	2	2	2	2	3
2	3	2	6	3	3	3	6	2	2	2	2	2	2	2	3
3	2	2	2	2	2	3	6	3	2	2	2	2	2	2	4
2	3	2	2	3	3	3	5	2	2	2	2	2	2	2	5
2	2	2	2	2	2	2	4	6	2	2	2	2	3	2	5
2	3	2	2	3	3	3	5	2	2	2	2	2	2	2	4
2	3	2	3	3	3	3	6	4	2	2	2	2	2	2	3
2	3	2	3	3	2	3	4	3	2	2	2	2	2	2	4
3	2	2	2	2	2	2	4	6	5	3	2	2	2	2	5
2	2	2	2	2	2	2	4	2	2	2	2	2	2	2	2
3	2	2	2	2	2	2	5	2	5	2	2	2	2	2	4
2	3	2	2	3	3	3	6	2	2	2	2	2	2	2	4
2	3	3	3	3	3	3	6	5	2	2	2	2	2	2	4
3	3	2	2	2	2	2	4	2	2	2	2	2	2	2	4
1	2	6	6	2	6	3	6	2	2	2	2	2	4	4	3
4	4	2	3	3	4	3	5	2	2	2	2	2	3	2	4
3	4	2	3	4	3	3	4	2	2	2	2	2	2	2	4
2	2	2	3	3	2	3	6	2	2	2	2	2	2	2	3
3	4	2	3	3	3	5	6	2	2	2	2	2	2	2	2
2	2	2	3	2	4	2	5	2	2	2	2	2	2	2	5
3	4	2	2	3	3	3	3	2	2	2	2	2	2	2	3
2	3	2	2	3	2	3	6	2	2	2	2	2	2	2	3
2	3	2	2	3	2	3	4	2	2	2	2	2	2	2	3
2	3	2	2	3	2	3	4	2	2	2	2	2	2	2	3
3	5	2	2	2	3	4	6	5	2	2	2	2	2	2	4
3	4	2	3	3	2	3	6	4	1	1	1	1	1	1	3

Table A.16.b: Academic graduate background and work experience

CLOT	CLOT	CLOT	OREH	OREH	OREH	CLHP	CLHI	CLHM	OREO	OREO	ORET
5	3	4	5	3	4	2	4	4	2	4	4
4	5	3	5	6	5	2	4	6	2	3	3
5	2	3	6	4	5	4	6	5	2	4	5
3	3	4	4	5	4	5	5	5	2	2	2
3	2	4	5	6	5	2	4	4	2	3	3
5	3	3	5	4	5	3	5	4	2	3	4
2	3	4	5	3	4	6	5	6	2	4	5
2	2	5	4	4	6	3	4	6	2	3	4
6	4	3	6	3	5	2	5	5	2	2	3
3	4	5	3	2	4	4	5	5	2	3	3
5	3	3	4	3	3	4	5	5	3	3	4
2	3	4	5	3	4	4	5	5	2	3	4
5	3	4	4	2	3	3	5	5	2	3	3
5	3	3	5	4	4	6	5	5	2	3	2
5	2	3	1	5	3	2	6	5	2	4	4
3	3	5	4	4	6	3	4	5	2	4	5
5	3	4	5	4	4	4	5	5	2	2	4
3	2	6	4	3	5	3	6	6	2	3	5
3	3	5	5	5	6	5	4	4	2	4	4
4	1	2	3	2	3	2	5	4	2	3	2
3	4	3	4	4	5	5	5	5	3	5	5
5	4	5	5	4	3	3	5	4	2	4	5
4	3	2	3	4	3	6	4	3	2	3	4
5	4	4	4	3	3	2	2	4	2	2	3
5	4	5	5	6	6	4	5	6	2	3	4
2	5	6	4	5	5	5	4	5	2	3	4
3	3	4	5	4	5	5	5	4	2	3	5
.	.	.	.	.	.	.	.	.	.	0	0
5	3	4	5	3	5	3	4	4	2	2	3
4	4	3	5	4	4	2	5	5	2	4	4
4	3	5	4	5	5	3	3	2	2	2	3
4	3	4	5	3	4	5	5	5	2	3	3
3	2	5	3	2	5	6	5	5	2	2	4
5	5	3	5	3	3	3	2	2	3	4	2
5	4	6	3	3	5	3	4	5	2	4	4
3	4	4	1	1	1	5	5	5	5	5	5
4	2	5	5	3	5	6	4	6	2	3	5
5	3	2	5	4	3	3	4	5	2	3	3
3	3	4	4	4	3	4	5	4	2	2	4
5	3	4	6	5	4	4	3	5	2	4	3
5	4	3	5	3	3	3	5	5	2	5	5

Table A.16.c: Subject's rating of similarities between aspects of markets

DEG1	UCG1	OPG1	FING	BALG	MKTG	DET1	UCT1	OPT1	FINT	BALT	MKTT	BKRU
0	6	0	1	0	0	1	0	0	0	0	0	1
3	1	0	1	0	7	2	2	0	2	0	4	1
0	0	0	4	1	1	1	1	0	0	0	0	0
2	1	3	1	3	1	0	0	0	0	0	0	3
2	5	0	1	1	2	0	6	0	0	0	1	3
5	20	1	1	2	35	0	1	0	0	0	1	0
0	4	3	2	3	6	2	19	2	9	3	1	0
3	28	0	0	1	28	6	17	5	3	1	10	0
1	3	1	0	0	4	0	0	0	0	0	1	0
2	2	1	1	0	1	1	2	0	0	0	0	3
0	0	0	0	0	20	0	0	0	0	0	0	0
0	1	0	2	1	0	0	6	2	3	0	0	0
2	0	0	4	0	0	0	0	0	0	0	0	1
0	23	3	2	0	13	0	0	0	0	0	0	0
3	8	0	1	0	6	0	0	0	0	0	0	0
1	0	0	0	0	1	1	1	0	0	0	0	0
0	8	2	0	1	11	0	0	0	0	0	0	1
1	2	0	0	1	4	1	0	0	0	0	0	2
0	2	0	0	0	25	2	1	0	0	6	9	0
0	0	0	0	0	0	0	0	0	0	0	0	2
2	7	2	0	0	4	0	0	0	0	0	0	3
0	8	0	1	1	16	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	1	0	6	0	0	0	2	2	0	5
3	7	1	2	5	35	8	7	0	1	5	1	1
1	7	0	1	0	5	1	7	0	1	0	5	0
1	0	0	0	0	0	0	0	0	0	0	0	1
2	8	3	1	1	0	1	4	1	1	2	0	1
1	3	3	6	8	11	0	0	0	1	0	0	1
0	2	1	1	1	19	0	0	0	0	0	0	0
1	1	0	1	0	23	0	1	0	0	0	0	1
4	3	1	1	0	5	1	0	0	0	0	0	0
1	2	0	0	1	1	0	1	0	0	0	0	1
7	18	2	4	10	1	1	2	0	0	0	0	3
0	13	0	0	3	10	1	4	0	1	1	2	1
2	11	1	3	0	19	6	8	0	4	0	5	0
0	5	0	0	0	1	3	5	0	0	0	2	0
2	27	0	2	0	1	0	4	1	1	0	0	1
1	2	0	0	1	0	0	0	0	0	0	0	1
3	3	1	2	0	0	0	0	0	0	1	0	1

Table A.16.d: Use of information and bankruptcy in trial 1.

DEG2	UCG2	OPG2	FING	BALG	MKTG	DET2	UCT2	OPT2	FINT2	BALT	MKT	BKRU
0	11	0	0	0	0	0	0	0	0	0	0	1
0	0	0	0	3	3	0	7	0	0	2	0	0
0	0	0	1	0	0	0	0	0	0	0	0	0
1	2	4	2	6	0	0	0	0	0	0	0	0
0	6	0	0	0	5	0	2	0	0	0	0	3
1	25	0	0	2	31	0	0	0	0	0	0	0
1	1	1	0	0	1	0	1	2	3	0	1	0
1	30	0	0	0	36	1	2	0	0	0	3	0
0	1	0	0	0	2	0	0	0	0	0	0	1
1	4	0	0	0	1	0	0	0	0	0	0	0
0	0	0	0	0	11	0	0	0	0	0	0	0
0	0	0	0	0	0	1	2	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	11	0	2	0	17	0	1	0	0	0	0	3
1	2	0	0	0	6	0	0	0	0	0	1	2
0	0	0	0	0	0	0	0	0	0	0	0	5
0	12	0	0	0	3	0	1	0	0	0	1	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	21	0	0	0	0	0	15	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	2
0	26	0	2	3	28	0	2	0	3	3	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	38	1	8	0	0	1	1	1
0	7	0	0	0	0	0	5	0	0	0	1	0
0	0	0	0	0	0	0	0	0	0	0	0	0
1	5	0	1	1	0	0	4	0	0	1	0	0
0	1	0	1	5	9	0	0	0	0	0	0	0
0	10	4	3	0	23	0	0	0	0	0	0	0
0	1	0	0	0	32	0	0	0	0	0	1	2
0	0	0	0	0	1	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0	0
1	25	0	3	12	1	0	0	0	0	0	0	1
0	3	0	0	0	4	0	0	0	0	0	0	0
0	3	0	1	4	39	2	4	0	0	1	5	0
0	3	0	0	0	6	0	8	0	0	0	5	3
0	35	0	0	0	0	0	5	0	3	0	0	0
0	8	1	0	1	17	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	0	0	0	0	2

Table A.16.e: Information access and bankruptcies in trial 2