Essays on Public Policy and Consumer Choice: Applications to Welfare Reform and State Lotteries

by

Melissa Schettini Kearney

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2002

Author ....
Department of Economics
May 15, 2002

Certified by ....
Jonathan Gruber
Professor of Economics
Thesis Supervisor

Certified by ....
Joshua Angrist
Professor of Economics
Thesis Supervisor

Accepted by ....
Peter Temin
Elisha Gray II Professor of Economics
Chairman, Department Committee on Graduate Studies

Essays on Public Policy and Consumer Choice: Applications to Welfare Reform and State Lotteries

by

Melissa Schettini Kearney

Submitted to the Department of Economics
on 15 May 2002 in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis investigates individual decision-making in response to government policies, in particular, state lotteries and the welfare "family cap." Despite considerable controversy surrounding the use of state lotteries as a means of public finance, little is known about their consumer consequences. Chapter one investigates two central questions about state lotteries and consumer behavior. First, do state lotteries primarily crowd out other forms of gambling, or do they crowd out non-gambling consumption? Second, does consumer demand for lottery games respond to expected returns, as maximizing behavior predicts, or do consumers appear to be misinformed about the risks and returns of lottery gambles?

Analyses of multiple sources of micro-level gambling data demonstrate that lottery spending does not substitute for other forms of gambling. Household consumption data suggest that household lottery gambling crowds out approximately $43 per month, or two percent, of other household consumption, with larger proportional reductions among low-income households. Demand for lottery products responds positively to the expected value of the gamble, controlling for other moments of the gamble and product characteristics. This suggests that consumers of lottery products are not misinformed and are perhaps making fully-informed purchases.

Chapter two investigates the nature of consumer choice under risk in the context of state lottery betting. Economists have traditionally modeled consumer preferences according to expected utility theory, but a recent body of literature challenges this model. An empirical test of the expected utility hypothesis finds that, in general, it is a reasonable description of observed consumer choices. However, the data offer some evidence in support of non-linear probability weighting in consumer preferences.

The second application studied in this thesis is welfare reform. A number of states have recently instituted family cap policies, under which women who conceive a child while receiving cash assistance are not entitled to additional cash benefits. Chapter three investigates how fertility behavior responds to this change in government expenditure policy. The analysis takes advantage of the variation across states in the timing of family cap implementation to determine if these policies are discouraging women from having additional births. The data consistently demonstrate that the family cap does not lead to a reduction in births. This finding of no effect is robust to the incorporation of lead and lag effects, to considering separately total and higher-order births, and to limiting the sample to demographic groups with high welfare propensities.

Thesis supervisor: Jonathan Gruber
Title: Professor

Thesis supervisor: Joshua Angrist
Title: Professor

Acknowledgements

I am extremely grateful to the MIT Department of Economics faculty. They are not only brilliant researchers, they are also very generous teachers. Jon Gruber has been a constant source of encouragement and advice. His love of economics is inspiring. I have learned a great deal from Josh Angrist and consider myself very fortunate to be his student. David Autor and Daron Acemoglu have also been wonderful advisors to me, both extremely helpful and encouraging. Jim Poterba is a true mentor, and I am very grateful for his guidance.

Many other faculty members offered helpful comments on this thesis, including Victor Chernozhukov, Peter Diamond, Sendhil Mullainathan, and Whitney Newey. The Bureau of Labor Statistics CEX division granted me access to confidential data used in chapter one. Thanks to Dean Gerstein of NORC and Brett Toyne of the Multi-State Lottery Association for assistance compiling data. I could not have completed chapter two without the help of Chris Hansen, who taught me how to program in MATLAB. Phil Levine and Diane Whitmore offered helpful comments on chapter three. Bob Schoeni provided welfare policy data.

I am indebted to the National Science Foundation for financial support during my first three years of graduate school, and to the Harry S. Truman Foundation for support in my final year. I remain indebted to my undergraduate professors at Princeton University - especially Beth Bogan, David Card, Anne Case, and Hank Farber - who introduced me to economics and provided me with a foundation of understanding.

My fellow graduate students have been amazing, both in brilliance and humor. I am grateful to my first-year study partners, Mara Berman, Barrett Kirwan, and Manual Amador, for all they taught me, but most of all for our shared laughter and friendship. I have had the good fortune of working alongside Tom Davidoff, Amy Finkelstein, Jon Guryan, Mark Lewis, Sean May, Robin McKnight, Cindy Perry, Stephanie Plancich, Jon Reuter, and Jeff Wilder. I look forward to having them all as my colleagues in the years to come.

My family has been a source of constant love, prayers, and joy. Throughout my graduate program, my parents lovingly reminded me not to work *too* hard, and cheered me on for every exam, presentation, and deadline. The support and friendship of my sisters, Allison, Christina, and Victoria has been, as always, invaluable. A huge thanks to my dear Nonny, for all the prayers and novenas on my behalf, and to my in-laws, Gloria and Dan, for being so interested in my work and supportive of my endeavors.

And finally, I am grateful to my husband, Dan, for editing and technical support; for helping and challenging me to think deeply; for always being there with a smile and hug; for four years of waiting patiently for me to finish my work; and above all, for promising to be with me always.

In loving memory of my "Uncle" Gene,
who would be so proud of me and would actually read this.
I miss him dearly.

Contents

Introduction

This thesis investigates individual decision-making in response to government policy. The first two chapters study the impact of state lotteries on consumer choices. State lotteries are a controversial method of state government finance. Yet, despite considerable debate about their effect on social welfare, prior to this research there have been no thorough studies of their consumer consequences. The first two chapters of this thesis examine state lotteries and consumer behavior. The third chapter evaluates how fertility behavior responds to a change in government expenditure policy. It investigates whether birth rates are affected by the introduction of the "family cap," a welfare reform policy that eliminates additional cash assistance to women who conceive a child while receiving cash welfare.

Chapter one investigates how the introduction of a state lottery impacts household consumption. In the past three decades, state lotteries have expanded dramatically, and lottery tickets now constitute a major consumer good. In 1999, consumers spent a total of $37 billion on lottery tickets, which amounts to an average of $226 per adult living in a lottery state or $370 per household nationwide. Opponents of state lotteries typically view lotteries as a regressive tax on minorities, the poor, and the misinformed. Supporters argue that lotteries are merely entertainment and substitute for alternative forms of gambling. Despite the controversy, there is virtually no empirical research into the validity of the claims on either side of the debate.

This paper addresses two central questions about state lotteries and consumer behavior. First, do lotteries simply crowd out other gambling expenditures, or do they lead to a reduction in non-gambling household consumption? Second, does consumer demand for state lottery tickets respond to variation in the expected return of the gamble? In other words, do consumers appear to be making well-informed, utility-maximizing choices?

The study first investigates how household gambling behavior responds to the introduction of a state lottery. Household-level data on gambling expenditures demonstrate that households increase their gambling expenditures in the presence of a state lottery. Total household gambling expenditures rise with the introduction of a state lottery, which implies that households are not financing lottery gambling by substituting away from other forms of gambling. A complementary analysis of gambling participation confirms that adults do not reduce their participation in previously-existing forms of gambling after a state lottery is introduced.

If consumers respond to the presence of a state lottery with new gambling, then they must substitute away from other consumption. An analysis of household consumption data finds that household spending on lottery tickets is financed completely by a reduction in non-gambling consumption. The introduction of a state lottery is associated with a decline of $128 per quarter in non-gambling consumption. This figure implies a monthly reduction of $27 in per-adult consumption, which compares to average monthly sales of $18 per lottery-state adult. These numbers suggest that households are completely financing their lottery gambling with reductions in non-gambling consumption, and perhaps even crowding in other gambling expenditures. The response is most pronounced for low-income households, which on average reduce non-gambling consumption by three percent.

The final analysis of chapter one is an evaluation of whether lottery consumers appear to be making informed choices over lottery products. This question is important to determining whether the shift in household consumption is consumer-welfare enhancing. If consumers are making informed purchases of lottery products, then consumer demand for lottery products should depend positively on its expected return, controlling for other characteristics of the game. To evaluate whether this prediction holds, I analyze weekly data on big-prize lotto game sales and characteristics for a sample of 91 lotto games from 1992 to 1998. The analysis suggests that sales are positively driven by the expected value of a gamble, controlling for higher-order moments of the gamble, the advertised top prize amount, and non-pecuniary characteristics. This finding suggests that consumers are at least partly – and potentially fully – informed in recognizing the wealth value of a bet.

Chapter two builds on this reduced-form demand analysis with a structural investigation of consumer choices over lotto gambles. It uses the context of state lotteries to evaluate the nature of consumer choice under risk. Understanding how economic actors make decisions under risk is crucial to public finance applications. Attitudes toward risk are fundamental to the optimal design of social insurance programs and to understanding individuals' savings and investment decisions. Economists have traditionally modeled decisions under risk according to the tenets of expected utility theory. But a recent body of literature challenges this model and proposes new explanations of how economic agents make decisions in the face of uncertainty. This chapter provides an empirical evaluation of the descriptive power of expected utility theory using data on the real-world choices of state lottery bettors. While the experimental literature on choices under uncertainty is abundant, there is a need for studies of non-expected utility models that use real-world data.

The chapter develops a model of lottery gambling that considers it to be part investment and part entertainment. First, the consumer optimization problem is specified under the hypothesis of expected utility. Consumer demand is derived as an implicit function of the characteristics of the lottery gamble and the parameters of the choice problem. This yields a structural equation linking observed demand and game characteristics to the parameters of the choice problem. I estimate the parameters of this equation using *generalized method of moments – instrumental variables* techniques and weekly data on lottery game sales and characteristics. The data fail to reject the expected utility specification of the model. This failure to reject the expected utility hypothesis is robust to specifying the utility function as either constant relative risk aversion or constant absolute risk aversion utility. Second, the consumer optimization problem is described with a more general specification suggested by Kahneman's and Tversky's (1992) *cumulative prospect theory.* In particular, the probabilities are allowed to enter the problem non-linearly. I re-derive consumer demand as an implicit function of lottery game characteristics and consumer choice parameters and estimate the choice parameters. Under this less restrictive specification, the data reject the expected utility restriction that probabilites enter the consumer problem linearly, but not in the way suggested by the experimental literature.

Chapter three shifts attention away from state lotteries to government cash assistance. It investigates how consumer behavior responds to a change in the generosity of a major government expenditure program. Over the past decade, states across the country have been experimenting with welfare reform and revising the policies of the federal Aid to Families with Dependent Children (AFDC) program. One of the most controversial reform policies, the "family cap" or "child exclusion," is motivated by the notion that an incremental increase in cash assistance for each additional child increases a woman's propensity to bear additional children. Eighteen states have instituted legislation ending the traditional practice of providing families on welfare with additional cash benefits when a new child is born into the family. An additional five states have implemented policies that alter the form of the additional benefit but fall short of eliminating it entirely.

This chapter examines whether the availability of more resources at the margin increases a woman's propensity to bear additional children. The policy of eliminating the marginal increase in resources raises the price of an additional child and might thus deter women from having additional births. The expected direct effect of the policy is to reduce higher-order births. The policy might also deter women from becoming first-time mothers insofar as it signals that welfare is not as reliable and generous as it was previously. These hypotheses are tested using *Vital Statistics Natality Data, Public-Use Data Files* from 1989 to 1998. The incremental change in births associated with a family cap is identified off variation in the timing of implementation across the 23 family cap states.

The analysis of vital statistics birth data finds no evidence of a reduction in births to women ages 15 to 34 associated with the implementation of family cap policies. The data reject a negative effect of more than 0.5 percent at the 95 percent confidence level when the analysis controls for state effects, month effects, and state specific linear time trends. Numerous specification checks, including lead and lag effects, do not alter this finding. In an attempt to increase the power of the analysis, the regression equation explaining births is estimated for subgroups with relatively high welfare propensities and focuses on additional births born to women in these groups. The data offer no evidence that the family cap is deterring unmarried high school dropouts or teenage women from having additional births. This is true for both blacks and whites.

Chapter 1

State Lotteries and Consumer Behavior

1 Introduction

In the past three decades, the prevalence and scale of state lotteries have expanded dramatically. The first modern state lottery was introduced in New Hampshire in 1964. By 1973, seven states operated state lotteries and consumers spent a total of \$2.1 billion on lottery products (in year 2000 dollars).[1] By 1999, there were 38 state lotteries in operation, and consumers spent a total of \$37 billion. This total represents an annual average of \$226 per adult living in a lottery state, or \$370 per household nationwide. This is more than the average household spent in 1999 on alcoholic beverages or on tobacco products and supplies. It is more than twice the amount households spent on reading materials. And it is roughly equal to what the average household spent on life and other personal insurance.[2]

As the expansion of state lotteries continues, there is enormous public controversy surrounding the use of lotteries as a means of raising public funds. Opponents argue that state lotteries prey on minorities and the poor and that spending on state lotteries displaces consumption and savings. Some worry that governments are "tricking" people with a "sucker's

[1] Clotfelter et al. (1999), p. 100. Their figures are in year 1997 dollars.
[2] United States Bureau of Labor Statistics (2001), Table A.

bet,"[3] exploiting misinformation on the part of consumers. Supporters of state lotteries counter that people from all demographic groups play the lottery. They argue that people demand gambling products and a state lottery capitalizes on that demand by providing a product that substitutes for other forms of gambling. Some characterize lottery sales as voluntary purchases of entertainment goods.

Despite the public controversy, there is virtually no empirical research into the validity of the claims on either side of the debate. This chapter fills that gap by addressing two central questions. First, do lotteries simply crowd out other gambling expenditures, or does the presence of a state lottery lead to a reduction in other forms of household spending? Second, does consumer demand for lottery games respond to expected returns, as maximizing behavior predicts, or do consumers appear to be misinformed about the risks and returns of lottery gambles?

The study first investigates how household gambling behavior responds to the introduction of a state lottery. I conduct two different analyses to answer this question. The first is an analysis of micro-level data on household gambling from confidential Bureau of Labor Statistics (BLS) *Consumer Expenditure Survey (CEX) - Diary Survey* files from 1982 to 1998,. During this time 21 states implemented a state lottery. I exploit the variation across states in the timing of state lottery introduction to compare the change in gambling expenditures among households in states that implement lotteries to the change among households in states that do not. The data demonstrate that total household gambling is increased after a state lottery is introduced, which implies that households are not financing lottery gambling by substituting away from other forms of gambling. A complementary analysis looks at data on adult gambling behavior from two national surveys, a 1998 survey conducted by the National Opinion Research Council (NORC) and a 1975 survey conducted by researchers at the University of Michigan. These data confirm that adults do not reduce their participation in previously-existing forms of gambling after a state lottery is introduced.

[3] To cite two opponents: "In fact, state lotteries ... are mechanisms by which the state seduces its citizens with the promise of riches, suckering them into gambling away their income and their unemployment checks on games that offer an almost infinitesimal chance of winning big." Robyn Gearey in *The New Republic*, May 1997; "The lottery may seem like 'funny money', but it is in effect taxation, taken through a con-trick." *The Economist*, Nov 18, 2000, on Britain's National Lottery.

If consumers respond to the presence of a state lottery with new gambling, then they must substitute away from other consumption. I analyze BLS *CEX - Interview Survey* data from 1984 to 1998 to investigate to what extent this is true. I exploit the variation across states in the timing of state lottery introduction to compare the change in household expenditures among households in states that implement lotteries to the change among households in states that do not. The analysis finds that household spending on lottery tickets is financed completely by a reduction in non-gambling consumption. The introduction of a state lottery is associated with a decline of $128 per quarter in non-gambling consumption. This figure implies a monthly reduction of $27 in per-adult consumption, which compares to average monthly sales of $18 per lottery-state adult. The response is most pronounced for low-income households, which on average reduce non-gambling consumption by three percent. Among households in the lowest income third of the CEX sample, the data demonstrate a statistically significant reduction in expenditures on food eaten in the home (3.3 percent) and on home mortgage, rent, and other bills (6.7 percent).

The final analysis of the chapter is an evaluation of whether lottery consumers appear to be making informed choices. The answer to this question is important to determining whether the shift in household consumption is consumer welfare enhancing. Lottery gambling is part investment, as consumers are making choices over risky assets, and it is part entertainment. Assuming that the entertainment and pecuniary components of the lottery gamble are separable, maximizing behavior predicts that consumer demand for lottery products should depend positively on its expected return, holding constant game characteristics. To evaluate whether this prediction holds, I analyze weekly sales and characteristics data from 91 lotto games from 1992 to 1998. The analysis suggests that sales are positively driven by the expected value of a gamble, controlling for higher-order moments of the gamble and non-wealth creating characteristics. This finding is robust to alternative specifications, including controlling for unobserved product fixed effects. In addition, I find that consumers respond to non-wealth creating, "entertaining" game features. Together, these two findings suggest that consumers are at least partly – and potentially fully – informed, rational consumers. It is consistent with these findings to claim that consumers derive an entertainment equal to the price of the gamble (one minus expected value), and then, insofar as they are making investments, they are informed evaluators of gambles.

The paper proceeds as follows. Section 2 presents an overview of state lotteries in the United States. It briefly discusses the history and operation of state lotteries and then presents micro-level evidence about lottery gambling. The section concludes with a theoretical discussion about the market for lottery products. Section 3 reviews related evidence. Section 4 discusses the impact of state lotteries on household expenditures. It looks first at gambling behavior and then at household non-gambling consumption. Section 5 investigates consumer demand for lottery products as a function of game characteristics. And finally, section 6 provides concluding comments.

2 State lotteries in the United States

2.1 History and operation

The state of New Hampshire ushered in the era of the modern lottery by introducing a state lottery in 1964.[4] Inspired by New Hampshire's lead, New York and New Jersey soon introduced their own state lotteries. Cross-border lottery sales place pressure on neighboring states to implement their own state lottery.[5] Accordingly, the spread of lotteries primarily followed a geographical pattern, spreading first across the Northeast, then to the West, and finally to the Midwest and South. By 1996, 37 states and the District of Columbia operated a state lottery. Appendix Table 1 lists implementation dates.

[4] Previously, lotteries played a role in raising money for such notable projects as Harvard College, the Continental Army, and public works undertakings throughout the Colonial period. A scandal involving the Louisiana Lottery in 1894 led to the prohibition of lotteries for seven decades.
[5] This explanation finds empirical support in Berry and Berry (1990), which finds that the probability that a state will adopt a lottery increases in the number of its neighbors that have previously adopted lotteries even controlling for internal characteristics. There is anecdotal support as well. Both Governor Don Siegelman of Alabama and Governor Jim Hodges of South Carolina campaigned in 1998 on pro-lottery platforms. Sigelman argued, "Hundreds of millions of Alabama dollars have left Alabama to buy lottery tickets in Florida and Georgia. I say it's time for us to keep that money here so that our schools can have pre-kindergarten, our schools can have computers, and our children can go to college tuition-free."

In each case the state ended its former prohibition of lotteries and established a state agency as the sole provider of lottery products. [6] All states use the profits from the state lottery operation as a source of revenue. Ten of the 38 state lotteries allocate lottery revenues to general funds; 16 earmark all or part of lottery revenues to education; and the remainder earmark for a wide variety of uses, some specific and others broad. On average, a dollar wagered on a state lottery game returns 33 cents of profit to the state. This profit can be likened to an excise tax levied at a certain rate on the purchases of a particular product. Assuming a five percent average state income tax, the implicit tax rate on state lotteries in 1997 was approximately 61 percent.[7] In spite of this, the lotteries' contributions to state budgets are modest. In 1997, the contribution of state lottery funds to total own-source general revenues ranged between .41 percent in New Mexico to 4.07 percent in Georgia.[8]

2.2 Lottery gambling: micro-level evidence

Consumer spending on state lottery products in 1999 totaled $37 billion in year 2000 dollars. Furthermore, the *2000 National Gaming Survey* reports that 72 percent of American adults purchased some kind of lottery product during the year, 28 percent played at least once a week, and 14 percent played more than once a week.

Micro-level evidence is available from two independent surveys: the 1975 *National Survey of Adult Gambling* conducted by Kallick et al. at the University of Michigan and the 1998 *National Survey on Gambling* conducted by the National Opinion Research Council (NORC) under contract with the National Gambling Impact Study Commission. The Kallick et. al. (1975) data consist of 1,749 completed interviews covering participants' lifetime and past-year gambling behavior. The NORC (1998) data contain information about the gambling behavior of

[6] The early state lotteries offered passive drawings and instant games. Over the years, states discontinued passive drawings and began to offer games that allowed players to choose their own numbers. These include daily numbers, lotto, and keno products. In recent years some states have adopted video lottery terminals (VLTs), which offer immediate feedback on whether the player won. See Clotfelter et al. (1999) for a more complete discussion on the operation of state lotteries.
[7] Clotfelter and Cook (1989) calculate that the average excise tax on four products in 1985, including federal, state, and local taxes was as follows: beer - 15 percent, wine - 17 percent, liquor - 43 percent, and tobacco products - 49 percent.
[8] National Gambling Impact Study Commission (1999), pp. 2-4.

2,417 adults from a random-digit dial sample.[9] In order to develop estimates of annual lottery expenditures from the information obtained by the NORC survey, I adopt a set of assumptions used by Clotfelter and Cook (1999).[10] Clotfelter and Cook (1999) calculate that estimates of national expenditures based on the NORC (1998) survey and this set of assumptions amount to only 86 percent of recorded sales. The reader should keep in mind that actual expenditures exceed the amounts discussed in this section. The reported expenditure differences across groups reflect true differences under the assumption that groups do not under-report lottery expenditures differentially.

Table 1 presents descriptive information from the NORC survey. The data reveal four general facts. First, people in all demographic groups participate in lottery gambling, where participation is defined broadly as any gambling during the year. Fifty-five percent of males and 47 percent of females report participation. The reported participation rate is 52.4 percent among whites, 42.3 percent among blacks, and 58.8 percent among Hispanics. Table 1 also shows that participation extends across all income groups.

Second, black respondents spend nearly twice as much on lottery tickets as do white or Hispanic respondents. The average reported expenditure among blacks is $200 per year, $476 among those who participate. Black men have the highest average expenditures. Average annual expenditures among the fifteen respondents in this demographic group are over $1,000; among the ten who participated in lottery gambling during the year, annual expenditures are over $2,000. In the 1999 *Current Population Survey* March file, mean income among this group is $10,400. Black women report higher average expenditures than white and Hispanic women as well as white and Hispanic men, in all income groups.

Third, average annual lottery spending in dollar amounts is roughly equal across income groups. Reported annual expenditures are $125, $113, and $145, respectively. This implies that

[9] Clotfelter and Cook (1999) use the NORC combined survey which includes the RDD sample and a gambling patron sample. To preserve the representativeness of the survey sample, I only use the random sample for my analyses.

[10] These assumptions first require assigning discrete values to the reported frequencies: 300 to "about every day", 100 to "1 to 3 times per week," 18 to "once or twice a month," 8 to "a few days all year," and 1 to "only one day in the past year". Second, if a respondent reports playing multiple types of games, it is assumed they played lotto no more than once per week.

on average, low-income households spend a larger percentage of their income on lottery tickets than wealthier households.

Fourth, lottery participation and spending is much higher in states with state lotteries than in states without lotteries. As shown in Table 1, participation in lottery gambling among adults living in lottery states is 54.7 percent, versus 25.2 in non-lottery states. (The difference is statistically significant with a t-statistic of 12.0.) Average annual lottery expenditures are estimated to be \$128 among residents of lottery states and \$47 among residents of non-lottery states. The difference is statistically significant, with a t-statistic of 4.62. By 1998, every continental state without a lottery bordered at least one state with one, making out-of-state lottery gambling feasible for a sizeable number of adults. The difference is much more pronounced in the 1975 survey when only 12 states operated lotteries: 50 percent of adults living in states with lotteries participated compared to only 7 percent of adults in non-lottery states.

2.3 Market conditions: theory

2.3.1 The product market and prices

In a perfect market, characterized by full competition and complete information, gambling products are supplied competitively by private firms and priced at marginal cost. For simplicity, assume that all gambles with the same expected value (EV) are valued equally among consumers. There is no differential entertainment value, nor utility over risk. Define the relevant price to be the price of a gamble with an EV of \$1. Consumers take the private market price as given, $P_p = MC$, and products are allocated efficiently. Contrast this environment to one in which there is only one gambling product and it is supplied by a monopolistic state lottery agency at the monopoly price P_s. Households face a higher price of gambling, $P_s > P_p$, and therefore purchase fewer gambles.

Historically, states have not established state lottery monopolies in a previously competitive environment. The gambling environment in a state *pre*-state-lottery can be described as one in which all lottery games are illegal within the state, but households are offered a limited supply of alternative gambling forms: illegal "numbers" betting, legal casinos, horse tracks or charitable gambling, or out-of-state lottery products. In this "limited" market, the price of gambling faced by household h is

$$P_{0h} = min\{P_n + \alpha_{nh}, P_c + \alpha_{ch}, P_b + \alpha_{bh}\}$$

where P_{0h} is the minimum price of gambling among the three available options. P_n is the average price of a \$1 EV gamble offered by numbers bookkeepers; P_c is the average price of a \$1 EV gamble offered by casinos or other legal venues; and P_b is the average price of a \$1 EV gamble offered by lotteries operated in bordering states. The second component $\alpha_{\cdot h}$ is the transaction cost to the household of the particular gambling type, which includes any transportation cost as well as any stigma associated with the particular form of gambling.

The establishment of a monopolistic state lottery introduces a new gamble at a price to household h of $P_{sh} = P_s + \alpha_{sh}$. The relevant price of a \$1 EV gamble for household h becomes $P_{1h} = \min\{P_{sh}, P_{0h}\}$. If P_{sh} is time-invariant, $P_{1h} - P_{0h} <= 0$, since alternatives remain available. In many cases the difference will be less than zero as lottery gambling itself involves minimal transportation and arguably stigma. (We might suspect that P_{sh} will change; alternatives could become less costly if the introduction of a lottery reduces the stigma of gambling, thereby reducing α_{nh}, α_{ch}, and/or α_{oh}.)

If consumers prefer a corner solution of no gambling or some fixed level of gambling losses, there will be no effect on consumer behavior. However, under the usual assumptions regarding consumer utility, the price and income effects work in the same direction for gambling, and consumers will increase their gambling expenditures. Because the magnitude of the price change varies across households, the response will be heterogeneous. (Once we acknowledge that gambles have differential entertainment values, the household response to state lotteries becomes more varied.) For consumption, the price and income effects work in opposite directions; depending on preferences, spending on non-gambling consumption will fall, rise, or stay the same. If consumers are rational and informed, and externalities are not relevant, then the reallocation of the household budget induced by the introduction of a state lottery will increase household welfare.

2.3.2 Consumer rationality and information

Among the 38 operating state lotteries in 2000, the average pay-out rate was 52 percent, ranging from a low of 26 percent in Delaware to a high of 71 percent in Nebraska.[11] When a lotto jackpot grows sufficiently large through rollovers accumulating from a series of drawings in which no one wins, it may be possible to place a bet with a positive return (Thaler and Ziemba, 1988). But such occasions are rare, and most lottery bets placed are on unfavorable gambles. Why would a risk-averse consumer purchase such a gamble?

The first explanation is that consumers know state lotteries offer unfair gambles but derive entertainment value from playing them. In this case, consumers are fully rational and informed decision makers and the only concern for economists is that the price is set inefficiently high at the monopoly price. An alternative explanation is that consumers are misinformed. In some instances, the odds of winning the jackpot might not be clear. Moreover, the advertised prize is typically the undiscounted prize amount, not the present discounted value of the annuity prize.[12] In addition, it might be the case that consumers know that the odds of winning are very small, but they do not actually understand the implications. Psychologists have documented an "illusion of control," whereby agents deny the operation of chance, believing that they can choose winning numbers through skill or foresight (Langer 1975, 1978). According to Kahneman and Tversky's (1979) prospect theory, agents overweight small probabilities and underweight large probabilities. In this line of thought, the agent is rational, but his objective function is not the objective function of expected utility theory.[13] If consumers are not making informed decisions, the welfare consequences of raising government revenue from lottery purchases is ambiguous.

---

[11] LaFleur's *2001 World Lottery Almanac.*

[12] For example, when the Powerball jackpot was advertised to be $266 million, the present discounted value of the 25-year annuity was $147 million (assuming a six percent interest rate.)

[13] An additional concern not addressed in this chapter is addiction. It is widely argued that gambling is addictive for some people, and lottery gambling is no exception. Becker and Murphy (1988) and Gruber and Koszegi (2000) argue that addiction does not necessarily imply irrationality. But, Gruber and Koszegi (2000) also argue that addiction amplifies the effects of irrationality. If lottery players are addicted consumers, the welfare consequences of state lotteries are ambiguous.

2.3.3 Intra-household externalities

The above discussion focuses on whether the consumer makes choices that unknowingly harm him, either because of irrationality or misinformation. An additional concern is whether the agent makes choices that harm those around him, in particular, other members of his household. Traditionally, economists have considered the family or household as a single unit that maximizes a common objective function subject to the family budget constraint. But recent evidence suggests that the household is a collective, not a unitary, entity and that expenditures depend in part on who controls the household income (Duflo (2000), Browning and Chiaporri (1998), Udry (1996)). If the members of the household do not share a common utility function, any increase in gambling expenditures might come at the expense of the well-being of those not in control of the household finances.

3 Related evidence

This chapter provides to the author's knowledge the first empirical test of the consequences of state lotteries for consumer behavior. Imbens et al. (1999) estimate the effect of lottery winnings on players' subsequent earnings, labor supply, consumption, and savings; this is a distinct question from the impact of lottery exposure on consumption. Clotfelter and Cook's 1989 book provides a comprehensive description of the legalization, provision, marketing, and implicit taxation of state lotteries. Clotfelter et al. (1999) provide a more recent overview of lottery operations, with particular attention to who plays the lottery, how the lotteries are marketed, and what kinds of policy alternatives exist for state and federal policymakers. It discusses survey evidence on lottery gambling based on the 1998 NORC survey discussed in the previous section. Worthington (2001) documents demographic predictors of lottery gambling in Australia and concludes that the implicit lottery tax is regressive.

There has been some limited previous investigation into the sales of lottery products. Clotfelter and Cook (1990a) provide a cursory look at the effect of changing prices and payoffs on lottery ticket sales. The authors observe 170 consecutive drawings of the Massachusetts lotto game in the mid-1980s and find that for each \$1,000 increase in the predicted jackpot due to "rollover", sales increase by \$333. Garrett and Sobel (1999) analyze the demand for lottery games using a 1995 cross-section of 216 lottery games in the United States. The authors make a series of assumptions, including indifference across lottery games, that yield the following result: the expected utility for any lottery player in a state can be represented by equating the odds ratio of winning the top prize in games G and g to the utility of winning the top prize in game g. The authors use the cubic approximation of Golec and Tamarkin (1998) to estimate a model of expected utility; they estimate the odds ratio as a linear function of the top prize, the square of the top prize, and the cube of the top prize. The estimated coefficients on the prize and cubic prize are significantly greater than zero, and the coefficient on the square of the prize is significantly less than zero. The authors interpret this as evidence of a cubic utility function, similar to that proposed by Friedman and Savage (1948) and found by Golec and Tamarkin (1998) in the context of betting at horse tracks.

In addition to the stringency of the identifying assumptions underlying Garrett and Sobel (1999), the empirical analysis of the paper has three major limitations. First, both instant and lotto games are included in the estimation sample. The result thus relies on the very strong assumption of a representative agent across game types. Second, the authors do not control for non-wealth creating characteristics of games. If consumers enjoy playing lottery games for reasons other than the gamble itself, omitting game features from the estimation is problematic. And finally, the key variable in their analysis, jackpot prize, is measured with systematic error. For games with variable jackpots, the authors estimate average prize using annual sales data and the percent of sales that is allocated to the prize. This approach does not incorporate the weekly variation in jackpot size within a game for games with rolling jackpots, but it uses the true jackpot amount for fixed jackpot games.

Gulley and Scott (1993) and Forrest, Gulley, Simmons (2000) analyze the demand for lotteries from the perspective of revenue maximization, rather than consumer preferences. Gulley and Scott (1993) examine drawing level sales data from four lotto games in three states from the late eighties to early nineties. The authors estimate demand as a function of price, defined as one minus the expected value, without controlling for higher-order moments or non-wealth creating characteristics. The resulting price elasticities suggest that two games are setting price close to the revenue maximizing value, one is setting price too low and the other too high. Forrest, Gulley, Simmons (2000) similarly examine sale patterns in the first three years of the UK National Lottery to estimate the price elasticity of demand. Their long-run estimate is close to minus one, which they interpret as evidence that the UK government is maximizing lottery revenue.

4 The impact of state lotteries on consumer expenditures

Lottery betting is widespread and substantial, as documented in Section 2.2 above. This raises the question: does the introduction of a state lottery induce new gambling expenditures and thereby crowd-out non-gambling consumption? Or does it merely cause substitution away from existing gambling alternatives? I answer these questions with three separate analyses. First, I investigate how total household gambling expenditures respond after to the introduction of a state lottery. Second, I analyze how participation in various types of gambling changes. And third, I investigate how household non-gambling expenditures shift in response to the introduction of a state lottery. I investigate the impact on gambling activities and non-gambling consumption separately because there is no single data source containing detailed information about both household gambling and non-gambling consumption.

4.1 How do state lotteries affect total household gambling? Evidence from consumer diaries

I investigate whether the introduction of a state lottery leads to increased household gambling using confidential Bureau of Labor Statistics (BLS) *Consumer Expenditure Survey (CEX) - Diary Survey* data files from 1984 to 1999. These files were accessed under an agreement with the BLS. The BLS CEX program consists of the quarterly Interview Survey and the two-week Diary Survey, each with its own independent sample of approximately 5,000 households (7,500 after 1998). The Diary Survey collects information about weekly household expenditures on frequently purchased small-item goods, including gambling expenditures.

Unfortunately, lottery gambling is drastically underreported in the CEX Diary Survey.[14] Based on 1998 sales data compiled by LeFleurs Inc., adults living in lottery states averaged $226 annually on lottery tickets. In contrast, CEX Diary respondents living in lottery states report an average of $0.71 for the two-week interval. Assuming smooth annual expenditures, this implies mean annual lottery expenditures of only $36. The underreporting is so severe that magnitudes implied by an analyses of this data are not reliable. However, we can infer from the analysis that total gambling expenses increase when a state lottery is introduced, even if we can not precisely say by how much. Furthermore, if underreporting is proportional across demographic groups, the CEX Diary data can reveal differential effects across groups.

Is total gambling higher in lottery states than in non-lottery states? The CEX Diary data suggest that both the unconditional probability of engaging in *any* type of gambling and *total* household gambling expenses are greater among residents in states with state lotteries than among residents in non-lottery states.[15] It appears that these differences are not entirely due to differences in preferences: mean household gambling expenditures are higher *post*-lottery ($2.17) than *pre*-lottery ($0.87) among states that ever adopt lotteries; the *t*-statistic of the difference is 10.4. This provides preliminary evidence that lottery gambling is not completely financed by substitution away from other forms of gambling.

To corroborate this initial finding, we turn to regression analysis. The analysis exploits the variation across states in the timing of state lottery introduction to evaluate whether the presence of a state lottery is associated with a change in household gambling. I use the same empirical strategy in the analysis of non-gambling consumption below. The strategy is to compare the change in expenditures among households in states that implement lotteries to the change in expenditures among households in states that do not make the lottery transition in the same period. Relative to states that have not yet implemented a state lottery, or that did so in the past, this analysis identifies the incremental change in expenditures associated with the

[14] Starting in 1996, the data files record lottery expenditures separately.

introduction of the lottery. The analyses cover the years 1982 to 1998. All dollar values are adjusted to year 2000 dollars using the BLS Consumer Price Index. During this time, 21 states switch status from non-lottery to lottery state; 16 states and the District of Columbia have lotteries in place the entire period; and the remaining 13 states are without a state lottery the entire period.[16]

The estimating equation takes the following form:

$$(1) \qquad y_{ijt} = \alpha + \lambda(LOTSTATE)_{jt} + X_{ijt}\beta_1 + Z_{jy}\beta_2 + M_{ijt}\beta_3 + \gamma_{jt} + \omega_y + \upsilon_j + \varepsilon_{ijt.}$$

In the first analysis, y_{ijt} is defined as gambling expenditures for household i in state j in the two-week time period t. In subsequent analyses, γ_{jt} is defined as total non-gambling consumption and then as spending on particular categories of goods, for household i in state j in reference period t. The regressor of interest is the *LOTSTATE* indictor. It is equal to one if there is a state lottery in the household's state of residency j during the reference period t, and zero otherwise. (For quarterly observations, it is based on the presence of a lottery in the first month of the quarter.) The coefficient on *LOTSTATE* is interpreted as the causal effect of the presence of a state lottery on the dependent variable.

The vector X_{ijt} consists of household level controls for family size, household income, urban status, number of persons less than 18 and over 64, and the sex, race, marital status, and education of the household head. The vector Z_{jy} consists of controls for the state level of cigarette, beer, and gasoline taxes, which vary by year. This controls for differences in the prices of these goods that are not captured in either year or state effects. The vector M_{ijt} consists of a series of dummy variables indicating the months of the year during which the household is observed; it is included in the estimation equation to control for seasonal spending effects. Finally, γ_{jt} is the state unemployment rate averaged over the quarter; ω_y is a binary indictor for the year, which controls for any nationwide shocks to spending; and υ_j is a dummy that captures fixed effects associated with state j.

---

[15] The mean two-week gambling participation rate is 8.5 percent in states with a lottery at the time versus 1.9 percent in non-lottery states; the *t*-statistic of the difference is 50.3. Unconditional mean two-week gambling expenditures are \$2.17 in lottery states versus \$0.71 in non-lottery states; the *t*-statistic of the difference is -14.1.

[16] The set of switching states consists of CO, CA, IO, OR, MO, WV, MT, KS, SD, VA, FL, WI, ID, IN, KY, MN, LA, TX, NE, GA, NM; the always-lottery states are NH, NY, NJ, CT, MA, MI, PA, MD, IL, ME, OH, RI, DE, VT, AZ; and the never-lottery states are AL, AK, AR, HI, ID, MS, NC, NV, OK, SC, TN, UT, WY.

The identifying assumption of equation (1) is that the implementation of the 21 state lotteries during this time period does not coincide with other state-level changes that are not controlled for in the regression but that might affect household expenditure behavior. An obvious candidate is changes in the legalization of other forms of gambling. Fortunately, changes in the availability of other forms of gambling does not coincide with the timing of state lottery introduction.[17]

Table 2 displays the results from estimating equation (1) for gambling behavior using CEX Diary data.[18] Mean gambling expenditures and participation among households in states that do not have a lottery in place at the time are listed in columns 1 and 3, respectively. Column 2 reports coefficients from an OLS regression of equation (1) with expenditure level as the dependent variable. As expenditures constitute a limited dependent variable, interpreting the regression coefficient is not entirely straightforward. When Ordinary Least Squares (OLS) regression is used to estimate the equation for expenditure levels, observations with zero spending are included in the analysis. The estimated impacts combine the extensive and intensive margins. These effects are reported separately in columns 4, 5, and 6. Column 4 lists the coefficients from OLS estimation of equation (1) with the dependent variable defined to be "any gambling expenditures"; column 5 lists marginal effects from a Probit specification. The final column reports the coefficient on $LOTSTATE$ when the dependent variable is the natural logarithm of expenditures. The coefficient necessarily captures changes on the intensive margin as the sample is conditioned on positive spending. To the extent that the introduction of a state lottery affects the extensive margin of gambling, the set of households with positive gambling expenditures is changed and the estimated effect on intensity is contaminated.[19]

[17] The legalization of casino gambling substantially lags the spread of state lotteries. Before the early 1990s, legal casinos only operated in Nevada and Atlantic City, New Jersey. Now they are legal in 28 states. Similarly, riverboat casinos did not begin operating legally until the first one opened in Iowa in 1991. Most Native American tribal gambling started after 1987, when the United States Supreme Court issued a decision confirming the inability of states to regulate commercial gambling on Indian reservations.

[18] With the exception that state unemployment rate is not controlled for in the analyses. State unemployment data were not available when the confidential BLS CEX Diary were accessed at BLS.

[19] Tobit and sample-selection models provide alternatives but have serious drawbacks. Perhaps the most pertinent in this context is conceptual: these models interpret the dependent variable as the censored observation of an underlying continuously distributed latent variable. Since we are interested in the question of how households actually respond, it makes little sense, if any, to say that they reduce spending by more than they actually can. The latent index coefficients have no predictive value for observed spending amounts. The two-part model (2PM) introduced by Cragg (1971) explicitly combines the participation and intensity effects. As discussed in Angrist

The results in Table 2 confirm that the introduction of a state lottery leads households to increase total gambling expenditures and participation. For the overall sample, the estimated coefficient on *LOTSTATE* in the OLS levels specification reveals that two-week gambling expenditures increase by a reported \$1.43, off a mean of \$0.71. OLS estimation of equation (1) for participation in any gambling finds that the introduction of a state lottery leads to an increase in the two-week gambling participation rate of approximately 0.07 percentage points. A Probit specification confirms the general result from OLS. Finally, column 6 reports the estimated effect of the introduction of a state lottery on the intensity of spending. The negative coefficient on the lottery state indicator suggests that new, less-committed gamblers are being brought into the gambling sample. Estimation of a Tobit specification, which includes non-gamblers in the estimation sample, corroborates the finding that gambling expenditures increase significantly in response to the presence of a state lottery.

Table 2 also displays results separately by income group, where households are divided into three strata (thirds of the income distribution) in the CEX survey data. Households in all income groups respond to a state lottery with increased gambling participation and expenditures. (Due to sample size limitations, estimating the equation separately by race is uninformative.)

4.2 How do state lotteries affect participation in various forms of gambling? Evidence from national gambling surveys

The analysis of CEX Diary data finds that household gambling expenditures rise when a state lottery is introduced. This suggests that lottery spending is not totally financed by a reduction in expenditures on previously existing gambling alternatives. But are they partly financed by substitution away from other gambling? To answer this question, I analyze the NORC (1998) and Kallick et. al (1975) data. Relative to the CEX Diary data, these data sources offer the

---

(2001), researchers using this model simply pick a functional form for each part, e.g. linear probability or probit for the first part and a linear or log-linear model for the second part. This has the advantage over the Tobit and other sample-selection models is that it does not impose restrictions on the latent index structure. Functional forms can also be chosen that impose nonnegativity. However, the 2PM does not attempt to solve the sample selection problem and the second part can not be interpreted as causal.

advantage of recording participation by type of gambling, but they have the disadvantage of not containing expenditure amounts. The analysis of this data is thus limited to observing effects on the extensive margin of various types of gambling.

I conduct a regression-adjusted difference-in-difference (DD) analysis on the combined data to determine how the introduction of a state lottery impacts participation in various forms of gambling. The DD analysis compares the mean change in gambling participation between 1974 and 1997 among states that implement a lottery in the intervening years to the mean change in gambling participation among states that did not. The comparison group consists of the set of states that either never have a lottery or have a lottery as early as 1974. The effect of interest is captured in the coefficient on *LOTST7597*year1997* — the interaction between an indicator variable for the year 1997 and an indicator variable for residing in a state that adopted a lottery between 1975 and 1997.[20] All regressions control for the following individual demographics: sex, race, marital status, education, and regular attendance at religious services. They also control for main year effects and a full set of state effects.

Results from the DD analysis of the effect of introducing a lottery on gambling participation are displayed in Table 3. The introduction of a state lottery leads to a statistically significant 50.4 percentage point increase in the probability that an adult participates in gambling of any kind during the year. Not surprisingly, the introduction of a state lottery leads to an increased probability of lottery gambling. More interestingly, the introduction of a state lottery does *not* have a negative effect on participation in track, bingo, private, or unlicensed gambling. The estimated coefficients on the independent variable of interest – *LOTST7597*year1997* – are remarkably close to zero in each of the four regressions. Again, we see that adults in all income groups respond to the introduction of a state lottery with increased

[20] While a DD strategy "differences out" *ex ante* differences, it is still interesting to know whether such differences exist. Are there differences *ex ante* in gambling participation rates, conditional on individual demographics, between states in 1974 that eventually adopt a lottery and those that do not? Regression results suggest there are not. *Lotst7597* is a binary indicator for whether the state implements a lottery between the two survey years. The coefficients on *lotst7597* (standard errors in parenthesis) in regressions with binary dependent variables indicating participation in the various forms of gambling are as follows: lottery .055 (.028), track .044 (.039), bingo .045 (.035), private .105 (.081), and unlicensed .073 (.071). These results suggest that there is no *ex ante* statistically significant difference in gambling participation between residents of *never*-lottery states and residents of states that eventually adopt lotteries.

gambling participation. For no income group do we see a substitution away from other types of gambling.

4.3 How do state lotteries affect household consumption? Evidence from consumer interviews

The analyses of household gambling behavior found no evidence that household lottery spending is financed by substitution away from previously existing forms of gambling. State lottery expenditures must therefore displace non-gambling expenditures. In this section, I investigate to what extent household non-gambling consumption is decreased when a state lottery is introduced. I analyze BLS *Consumer Expenditure Survey (CEX) - Interview Survey* data from 1982 to 1998. The CEX Interview Survey collects information on major items of expense and household characteristics.[21] Households are asked about expenditures for up to three consecutive quarters. The BLS estimates that 90 to 95 percent of expenditures are covered by the Interview survey, but gambling expenditures are excluded. The analysis therefore asks a reduced-form question: does the introduction of a state lottery lead to declines in non-gambling consumption.[22]

I estimate equation (1) for non-gambling consumption. Table 4 lists the results. Column 1 lists mean spending among households in states that do not have a lottery in place at the time. Column 2 reports coefficients from an OLS regression of equation (1) with spending level as the dependent variable. (All households have positive spending so composition-bias is not an issue.) Column 3 lists the implied percentage change from the non-lottery mean. The final column reports the coefficient on $LOTSTATE$ when the dependent variable of equation (1) is the natural logarithm of expenditures. Specifying the function as log-linear has two relevant properties: one, the effect of outliers on the estimated coefficient is mitigated, and two, the

---

[21] The public use CEX Interview files do not include records from Rhode Island and Montana. Furthermore, the BLS public files suppress the state of residence for some records in order to meet the Census Disclosure Review Board's criterion that the smallest geographically identifiable area have a population of at least 100,000. The consequence is that approximately 17 percent of records do not have state identified: state is left blank for all records from Mississippi, New Mexico, Maine, and South Dakota, and for some records from other states. The consumption analysis sample therefore includes observations from 42 states and the District of Columbia.

[22] The unreliability of gambling magnitudes found in the analysis of CEX Diary data preclude the construction of a two-sample IV estimate of the effect of increased gambling on non-gambling consumption.

coefficients are interpreted as percentage changes. This allows us to observe the proportional decline in different categories of spending.

For the overall sample, total quarterly spending falls by $128, implying an average decrease of $43 in monthly household consumption expenditures. The average number of adults in a CEX household is 1.57; from this we calculate a monthly consumption reduction of $27 per-adult. How does this offset of consumption compare to state lottery ticket sales? Based on the LeFleurs sales data, monthly sales per-adult average $18 across the 38 state lotteries. We thus conclude that household lottery gambling is *completely* financed by a reduction in non-gambling consumption. Furthermore, these numbers suggest that lottery gambling crowds-in additional gambling expenditures.

The decrease of $128 in consumption expenditures represents a decline of 1.7 percent relative to mean total spending in the absence of a state lottery. The log-linear specification finds a decline of 2.0 percent (with an associated standard error of 0.7). This latter estimate might be preferred since the effect of outliers is mitigated. The implication is that on average, households displace two percent of their quarterly consumption expenditures with state lottery ticket purchases.

The bottom panel of Table 4 presents the results from two specification checks on the model. Recall from the discussion in Section 2.3 that the introduction of a state lottery has a non-positive effect on the price of gambling. The magnitude of the price decrease varies by household, depending on the availability of alternative gambling forms and the associated transportation or stigma costs. The theoretical implication is that if a neighboring state already offers a state lottery, the introduction of one will have less of an effect on the price. The further implication is that the household response in terms of gambling and non-gambling consumption expenditures will be smaller.

The bottom of Table 4 reports the regression-adjusted effect of the introduction of a state lottery when a bordering state already operates one. The coefficient on the *LOTSTATE* indicator captures the "pure" effect of introducing a state lottery on total non-gambling consumption. The coefficient on *LOTSTATE*BORDER* captures the additional effect of introducing a lottery when a neighboring state already operates one. (To be clear, this interaction term equals zero if the state lottery is introduced before any neighboring states introduce one; it does not switch to one if and when a neighboring state finally does introduce a state lottery.) For the overall sample, the analysis finds that households reduce quarterly consumption by $279.5 when a state lottery is introduced, as shown in column 1. If the lottery is introduced when a neighboring state already operates a lottery, the effect is mitigated by $180.8, as shown in column 2, though the point estimate is not statistically significant. Columns 3 and 4 report the coefficients from a log-linear specification. These estimates suggest that the "pure" effect of introducing a state lottery is a decline in quarterly household spending of 3.4 percent; if a border state previously operated a lottery, the decline is only 1.8 percent.

An additional question is whether the shift in expenditures is temporary. The bottom panel of Table 4 confirms that the reduction in consumption is sustained in the long run. In the first two years after a state lottery is introduced, households respond with an average decline in quarterly non-gambling consumption of 1.9 percent (standard error of 0.80). This response is sustained: the average decline in consumption among households in states with lotteries that have been operating for at least two years, relative to households residing in states without lotteries, is 1.8 percent (standard error of 0.7).

4.4 How do state lotteries effect the consumption of low-income households? Evidence from consumer interviews

Among households in the lowest income third, total quarterly spending is reduced by $135 (see Table 4), implying a decrease of $45 in monthly household consumption expenditures. Using the average number of adults in a CEX household, we calculate a monthly consumption reduction of $29 per-adult. How does this reduction compare to lottery ticket purchases? Sales data are not available by income group, but we can compare this decline in consumption to reported lottery gambling in the NORC (1998) survey data. Lottery-state adults in the lowest income third report an average of $139.5 in lottery spending; adjusting this figure for known underreporting (see Section 4 above) yields average yearly spending of $162.2, or $14 per month. These numbers suggest that low-income households are financing their lottery gambling *completely* by a decline in consumption. Again, the data suggest that perhaps lottery gambling crowds-in other gambling expenses.

Low income households experience the most pronounced percentage decline in consumption spending: 2.9 percent (standard error of 1.2). As reported in Table 4, OLS estimation of the log-linear specification suggests that the average response among households in the middle income group is a decline of 0.5 percent (standard error of .8); 1.6 percent (standard error of .8) among those in the highest income group. The data reject the hypothesis that the proportional decline for the middle income group is the same as for the lowest income group, but we can not reject the hypothesis for the highest income group. Again, households are divided into three strata based on the income distribution in the CEX Interview Survey sample.

Table 5 offers a more detailed picture of how low-income households change their consumption in the presence of a state lottery. Equation (1) is estimated separately for 11 categories of goods: *food at home; medical drugs; rent, mortgage, other household bills; alcohol; smoking products; food out of the home; entertainment; education; household repairs, services, and furnishings; clothes (children and adult); transportation and cars.* The table reports estimates for the levels, participation, and log-linear specifications. It is difficult to obtain precise estimates in this excercise, but the analysis does offer a few interesting insights. First, the decline in consumption appears to be spread across expenditure categories. Point estimates are negative for 10 of the 11 categories. Statistically significant reductions are observed in spending categories that might be classified as "necessities:" *food at home* and *home expenditures including mortgage, rent, and other household bills.* In terms of within-household externalities, it is interesting to note that lottery spending appears to be a substitute for the "adult" good alcohol; on the other hand, there is no evidence that spending is reduced on children's clothing, but statistical power is potentially a problem.

5 Consumer demand for lottery products

The above section provides unambiguous evidence that households respond to the introduction of a state lottery by increasing their gambling expenditures at the expense of a reduction in other forms of consumption. If consumers are fully-rational and fully-informed, and externalities are not relevant, then these behavioral responses are welfare enhancing. However, if the oft-raised concern that consumers are making misinformed choices is true, then the effect on consumer welfare is not clear. This section provides an initial exploration of consumer choices over lottery products and investigates whether consumers of lottery products appear to make informed choices.

As outlined in the introduction, the hypothesis that lottery consumers are being deceived implies that consumer demand for lottery tickets does not respond to the expected value of a gamble, conditional on other features of the game. If consumers are fully misinformed, their demand for lottery gambles might respond to the top prize, but it should not respond to the expected value of the bet. An additional test of consumer rationality and information is whether consumers derive entertainment value from lottery gambling. Assuming consumers are risk-averse, then participation in gambles with an average return of 52 cents on the dollar reflects a fully-rational, fully-informed decision only if there is some entertainment value to participation. I test whether consumers derive entertainment value by observing whether their demand responds to variation in non-wealth creating characteristics of lottery games, such as the number of drawings per week or the number of digits chosen. I perform these two tests simultaneously.

5.1 Data and empirical strategy

To investigate the nature of consumer demand, I combine game level sales data with detailed information about the corresponding lottery game. The analysis is conducted at the level of state, game, and week. To the best of my knowledge, I am the first to compile a comprehensive data set of lottery game characteristics, and this is therefore the first analysis of its kind. I limit the empirical analysis to lotto games, to the exclusion of other types of lottery products including numbers games, instant scratch-off, keno, bingo, and VLT products.[23] Relative to

[23] I include multi-state lotto games in the sample because the two types of products have the same essential structures; they differ only in scale. Multi-state lotto games pool sales across states to engender larger jackpots. There are six unique multi-state lotto products: Wildcard, Powerball, Cash 4 Life, and Daily Millions, which are run

other products, lotto games vary substantially in prize amounts and structure. There is both variation across games and over time within a game as jackpot amounts frequently "rollover" and accumulate. Additionally, to draw conclusions about individual behavior from aggregate sales data we rely on a representative agent assumption; limiting the analysis to a single type of lottery product makes this assumption substantially less stringent.

The structure of a lotto game is defined by the number of digits the bettor chooses and the size of the field. For example, in a lotto game with a 6/44 game matrix, a bettor chooses 6 numbers without replacement from a field of 44; the odds of picking the winning numbers are 1 in 7,059,052. Some lotto games have fixed jackpot amounts; others have "rolling" jackpots such that if the jackpot is not won on a given draw, the jackpot (minus the prize payments for partially correct bets) is rolled over into the jackpot for the next drawing. Some lotto games pay the jackpot as a cash prize, others as a long-term annuity, and others offer a choice. Lotto games also differ in the number of draws per week. [24]

I obtained weekly sales data from 1992 to 1999 from Lefleurs Inc., a group that collects weekly sales data from state lottery agencies. (Appendix Table 2 describes the sales data.) I obtained information about game characteristics from state lottery websites and from lottery game brochures provided by state lottery agencies. For games with rolling jackpots, I obtained times series data on the advertised jackpot amounts from various state lottery agencies. The sample excludes games for which only realized jackpot data is available; in games in which the jackpot rolls over, the actual jackpot amount is a function of both the rollover amount and the induced additional sales. Using the advertised amount avoids incorporating this latter portion

---

by the Multi-State Lottery Association; and The Big Game and Megabucks, which are not. I consider the state version of a multi-state product a unique game; for example, Powerball in Minnesota is considered a different game than Powerball in Montana. This seems appropriate as states run individual advertising campaigns.

[24] I offer two examples. First, a resident of Maryland playing the "Cash in Hand" game can purchase a ticket from any Maryland State Lottery location any day of the week. There are three drawings per week. He pays the retail agent $1 and picks 7 out of 31 numbers, or marks "quick pick" and lets the machine pick the numbers for him. If the 7 numbers on his gameboard match the 7 winning numbers (with odds of 1:2,629,575), and he claims his prize within 182 days from the date of drawing, he is paid $500,000 cash. The state of Maryland will pay each game board with the winning numbers $500,000. (In the unlikely event that more than 5 game boards win, all winning boards will receive an equal share of a $2,500,000 pool.) Second, a resident of Florida playing Florida Lotto pays $1 and picks 6 numbers out of 53, or marks "quick pick". She can place bets on up to 26 consecutive drawings in advance. If the 6 numbers on her ticket match the 6 winning numbers (with odds of 1:22,957,480), and she claims her prize within 180 days, she wins the jackpot amount. The actual prize depends on sales and the number of winners for the draw. If there is no ticket with the winning number, the jackpot rolls over and the cash available for that jackpot is added to the next jackpot prize pool.

into the independent variable. For state-game-week cells that have more than one advertised jackpot (because there are multiple drawings per week and the jackpot is not a fixed amount), I take the maximum advertised jackpot during the week. The final sample used in the empirical analysis consists of nearly 15,000 observations at the game-week level. These observations are from a sample of 91 lotto products from 33 states.

The empirical analysis estimates how weekly sales of lotto tickets respond to changes in the statistical moments of the gamble as well as to differences in game characteristics. The estimating equation takes the following form:

$$y_{sgw} = \alpha + \lambda_1(\textit{expected value})_{sgw} + \lambda_2(\textit{variance})_{sgw} + \lambda_3(\textit{skewness})_{sgw} + \\ + \lambda_5(\textit{nominal top prize})_{sgw} + X_{sgw}\beta_1 + Z_{sy}\beta_{2} + \zeta_s + \omega_w + \upsilon_g + \varepsilon_{ijt.}$$

where y_{sgw} is the natural logarithm of per adult sales from game g, in state s, in week w. The vector X_{sgw} includes non-wealth creating characteristics of the game. The vector Z_{sy} includes controls for the proportion of the state population in seven age-sex demographic groups, observed at the year level. All regressions control for state and week effects, $\zeta_s + \omega_w$. In some specifications, the equation is estimated with a game dummy υ_g to control for unobserved product fixed effects. The equation is estimated using OLS, weighted by state population. Standard errors are robust standard errors, adjusted for clustering at the state-year level to flexibly control for correlation of the error terms.

The moments of a one dollar gamble depend on several factors: the structure of the game, the value of previous rolled-over jackpots, and the number of tickets bought in the current drawing. The moments are calculated using the "real top prize," which is the present discounted value of the advertised jackpot (assuming a six percent interest rate), and all lower prize tiers offered by a game. All prize amounts are adjusted to year 2000 dollars. I make the simplifying assumption that the probability of multiple winners, which depends on the number of tickets bought and the numbers chosen by bettors, is negligible. Hence, the expected value is not adjusted for the probability of having to share the jackpot. The mean expected value of a \$1 bet among the sample of all lotto games is 0.53.

The "nominal top prize" of a game is the advertised dollar amount. This is the undiscounted sum of the game-specific number of annual payments. In the analysis, the "nominal top prize" is adjusted to year 2000 dollars using the Consumer Price Index, but it is not discounted to present terms. In most instances, it is nearly twice as large as the "real top prize." The highest single-state lotto prize in the sample is associated with the Texas Lotto in January, 1994: a nominal top prize of $18 million, with a present discounted value of $10 million. The largest prize among multi-state games is associated with the Powerball game in July 1998; the nominal prize amount is $266 million, with a present discounted value of $147 million. (The actual jackpot won on this game was $295.7 million, in year 2000 dollars.) The vector X_{sgw} includes the following non-wealth creating game characteristics: number of draws per week, age of game, age of game squared, how many numbers the bettor picks, and the jackpot type (cash, annuity, or a choice).

5.2 Results

Table 6 displays the estimation results. All regressions control for state unemployment rate, state fixed effects, week fixed effects, and demographic composition. Column 1 displays the results of estimating demand as a function of only the statistical moments of the gamble. The results provide preliminary evidence that consumers respond positively to the expected value of a gamble, but the point estimate is not statistically significant. This specification suggests that consumers like variance, but dislike skewness. Note that this finding contradicts the finding of Garrett and Sobel (1999) that consumers respond negatively to variance and positively to skewness. Column 2 adds entertainment characteristics as independent variables. The positive coefficient on expected value increases in absolute value to 0.683 and is statistically significant (standard error of 0.113). This finding rejects the hypothesis that lottery players are misinformed evaluators of gambles.

Column 2 shows that consumer purchases are also driven by non-wealth creating characteristics of lottery products. This implies that consumers are deriving entertainment value from playing the lottery. For example, consumers appear to prefer picking more numbers to fewer and demand more of a game as it ages. The specification reported in column 3 adds the nominal top prize as an independent variable. Not surprisingly, it enters positively and is statistically significant. The interesting result in this column is that the estimated positive effect of expected value is maintained and even strengthened. The point estimate is 0.757, with a standard error of 0.108. Replacing "expected value" with the natural logarithm of one minus the expected value in this specification, yields an estimated price elasticity of -0.39.

The specifications reported in columns 4 and 5 incorporate product fixed effects into the model. The estimation now controls for differences in sales across games that are driven by fixed game characteristics not explicitly captured by the regressors in the model. Again, the data demonstrate that sales are positively driven by the expected value of a gamble and that demand responds to the non-wealth creating characteristics of lotto games. The specification in column 5 yields an estimated price elasticity of -0.17.

It is consistent with these findings to claim that consumers are fully rational: they derive an entertainment value from participating in the lotto gamble that equals the price of the gamble (one minus expected value), and then, insofar as they are making investments, they recognize which gambles are better investments. On the other hand, it is also consistent to argue that consumers are at least partially irrational, believing that the non-wealth characteristics bear on the likelihood of winning positive returns. Though the analysis does not allow us to discriminate between the two scenarios, it does imply that consumers are at least partly – and potentially fully – informed in recognizing the wealth value of a bet.

6 Conclusion

This chapter has offered two main contributions to the public debate regarding the consumer consequences of state lotteries. The first contribution is an empirical investigation of how households shift their spending in response to the introduction of a state lottery. I have used the variation across states in the timing of state lottery introduction to compare the change in expenditures among households in states that implement lotteries to the change in expenditures among households in states that do not. The analyses are based on consumer expenditure data from 1982 to 1998, during which time 21 states implemented lotteries.

The evidence on household gambling expenditures demonstrates that households increase their gambling expenditures in the presence of a state lottery. Total gambling after a lottery is introduced exceeds previous gambling expenditures, which implies that households are not financing lottery gambling completely by substituting away from other forms of gambling. A complementary analysis of participation in various forms of gambling finds that there is no substitution away from participation in other forms of gambling when a lottery is introduced. In fact, my analysis of household non-gambling consumption suggests that household spending on lottery tickets is financed completely by a reduction in other forms of household consumption. The introduction of a state lottery is associated with a decline of $128 per quarter in non-gambling consumption. This figure implies a monthly reduction of $27 in per-adult consumption, which compares to average monthly sales of $18 per lottery-state adult. The response is most pronounced for low-income households, which on average reduce non-gambling expenditures by approximately three percent. The impact of a state lottery is found to be more pronounced if no bordering state previously implemented a lottery. In addition, the decline in non-gambling consumption is sustained in the long run.

The second major contribution of the paper is an evaluation of whether lottery consumers appear to be making informed choices. To evaluate this question I analyze lottery sales data from 91 lotto games from 1992 to 1998 as a function of lottery product attributes, including the statistical moments of the gamble, the advertised undiscounted top prize, and the non-wealth creating characteristics of the game. The analysis suggests that sales are positively driven by the expected value of a gamble, controlling for other characteristics. This finding is robust to alternative specifications, including controlling for unobserved product fixed effects. The NORC (1998) survey offers supporting evidence that agents understand that lotteries are not fair bets. The survey asks respondents how much of the ticket price of their favorite game do they think is returned as prize money. Only 7.5 percent of the respondents thought the pay-out was above the actual average pay-out rate. This finding suggests that consumers are at least partly – and potentially fully – informed in recognizing the wealth value of a bet.

Two things should be kept in mind when interpreting the results of this chapter. First, the analysis has identified average effects, but due to data limitations, can not sufficiently examine the heterogeneity of household response. While the average household reduces consumption by $43 a month in response to the introduction of a state lottery, there are likely to be some households in the tail of the distribution who forego much greater amounts of consumption. Second, intra-household externalities are a potential issue that can not be sufficiently addressed with available data. For example, there is some anecdotal evidence to suggest that some members of lottery-gambling households would rather not spend household money on lottery tickets. Future work examining these issues would lead to a more thorough understanding of the welfare implications of state lotteries.

References

[1] Ali, M.M. 1977. "Probability and utility estimates for racetrack bettors." *Journal of Political Economy* 85, 803-815.

[2] Angrist, J. 2001. "Estimation of Limited Dependent Variable Models With Dummy Endogenous Regressors: Simple Strategies for Empirical Practice." *Journal of Business and Economic Statistics*, 19, No. 1.

[3] Becker, G. and K. Murphy. 1988. "A Theory of Rational Addiction," *Journal of Political Economy.*

[4] Berry, F.S. and W.D. Berry. 1990. "State Lottery Adoptions as Policy Innovations: An Event History Analysis," *American Political Science Review*, 84(2).

[5] Browning, M. and P. Chiappori. 1998. "Efficient Intra-Household Allocations: A General Characterization." *Econometrica*, 66(6), 1241-1278.

[6] Cragg, J.G. 1971. "Some Statistical Models for Limited Dependent Variables With Application to the Demand for Durable Goods," *Econometrica*, 39, 829-844.

[7] Clotfelter, C.T., P.J. Cook, J.A. Edell, and M. Moore. 1999. *State Lotteries at the Turn of the Century: Report to the National Gambling Impact Study Commission.* Duke University.

[8] Clotfelter, C.T. and P.J. Cook. 1990a. "On the Economics of State Lotteries," *The Journal of Economic Perspectives*, 4(4), 105-119.

[9] ———. 1990b. "Redefining Success in the State Lottery Business," *Journal of Policy Analysis and Management* 9.

[10] ———. 1989. *Selling Hope: State Lotteries in America.* Cambridge, MA: Harvard University Press.

[11] Duflo, E. 2000. "Grandmothers and Granddaughters: The Effects of Old Age Pension on Child Health in South Africa." MIT Department of Economics working paper.

[12] Forrest, D.O., D. Gulley, and R. Simmons. 2000. "Elasticity of Demand for UK National Lottery Tickets," *National Tax Journal* 53(4), 853-863.

[13] Friedman, M., and L.J. Savage. 1948. "The Utility Analysis of Choices Involving Risk." *Journal of Political Economy* 56, 279-304.

[14] Garrett, T.A. and R.S. Sobel. 1999. "Gamblers favor skewness, not risk: Further evidence from United States' lottery games." *Economic Letters* 63, 85-90.

[15] Geary, R. 1997. "The Numbers Game: State lotteries – A Ticket to Poverty," in *The New Republic*, May 19.

[16] Gerstein, D. and M. Toce. 1999. *Fast Track Codebook for the Gambling Impact and Behavior Study –Adult Surveys.* National Opinion Research Council.

[17] Golec, J., and M. Tamarkin. 1998. Bettors love skewness, not risk, at the horse track. *Journal of Political Economy* 56, 279-304.

[18] Gruber, J. and B. Koszegi. 2000. "Is Addiction 'Rational'? Theory and Evidence." NBER Working Paper No. W7507.

[19] Gtech Corporation. 2000. *The Vital Signs of Legalized Gaming in America: Gtech's 8th Annual National Gaming Survey.* West Greenwich, RI.

[20] Gulley, O. D., and F. A. Scott. 1993. "The Demand for Wagering on State-Operated Lotto Games," National Tax Journal, 45(1): 13-22.

[21] Imbens, G.W., D.B. Rubin, and B. Sacerdote. 1999. "Estimating the Effects of Unearned Income on Labor Supply, Earnings, Savings, and Consumption: Evidence from a Survey of Lottery Players." NBER Working Paper No. W7001.

[22] *International Gaming & Wagering Business*, April 1998, 44;

[23] *International Gaming & Wagering Business*, June 1998, 48-49.

[24] Kahneman, D., and A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47, 263-291.

[25] Kallick, M., D. Suits, T. Dielman, and J. Hybels. 1976. *Gambling in The United States* (ICPSR 7495) University of Michigan Institute for Social Research.

[26] Kearney, M.S. 2001. "State Lotteries and Consumer Behavior." unpublished MIT mimeo, November.

[27] Spindler, C.J. 1995. "The Lottery and Education: Robbing Peter to Pay Paul?" *Public Budgeting and Finance*, 15, 54-62.

[28] *LaFleur's World Lottery Almanac.* 2001. LaFleur's Inc. www.lafleurs.com.

[29] Mullahy, J. 1997. "Instrumental-Variable Estimation of Count Data Models: Applications to Models of Cigarette Smoking Behaviour," *Review of Economics and Statistics*, 11, 586-593.

[30] National Gambling Impact Study Commission. 1999. *Final Report.*

[31] National Public Radio, *All Things Considered*, April 14, 1999.

[32] National Public Radio, *All Things Considered*, June 22, 1999.

[33] Pulley, B. 1999. "Waiting for riches: a special report - living off the daily dream of winning a lottery prize," *The New York Times*, May 22, A1.

[34] Thaler, R.H. and W.T. Ziemba. 1988. "Anomalies. Parimutuel betting markets: Racetracks and lotteries. *Journal of Economic Perspectives* 2(2), 161-174.

[35] Udry, C. 1996. "Gender, Agricultural Production, and the Theory of the Household," *Journal of Political Economy*, 101 (5), 1010-1045.

[36] United States Department of Labor, Bureau of Labor Statistics. 2001. "Consumer Expenditures in 1999." Report 949.

[37] Worthington, A. 2001. "Implicit Finance in Gambling Expenditures: Australian Evidence on Socioeconomic and Demographic Tax Incidence," *Public Finance Review*, 29 (4), 326-342.

Table 1
Lottery Participation Rates and Expenditures

1998 *NORC* Survey Data

	Overall			*Lottery States*			*Non-lottery states*		
	n	% who played last yr	avg yrly spending, all adults	*n*	% who played last yr	avg yrly spending, all adults	*n*	% who played last yr	avg yrly spending, all adults
overall	2,417	51.3	107.3 (470.7)	2,047	55.7	128.4	357	25.2	47.3 (240.9)
male	1,152	55.8	143.2 (525.7)	981	51.8	153.4	163	30.1	82.5 (331.1)
female	1,265	47.2	91.8 (494.6)	1,066	59.9	105.3	194	21.1	17.8 (114.7)
white	1,769	52.4	107.9 (510.0)	1,059	57.0	119.3 (544.3)	251	24.3	41.4 (215.8)
black	291	42.3	200.1 (711.9)	237	46.0	230.0 (770.5)	53	24.5	67.0 (333.8)
hispanic	170	58.8	108.4 (214.9)	154	61.0	107.5 (208.0)	14	28.6	86.7 (288.0)
other	180	47.2	74.9 (257.3)	141	51.8	81.8 (263.1)	38	28.9	45.5 (238.1)
Household income									
< 27,000	353	45.0	125.4 (560.5)	287	50.5	139.5 (610.0)	63	17.5	53.0 (245.5)
27,000 to 54,000	445	56.2	113.4 (455.0)	368	63.0	127.1 (485.2)	76	22.4	48.0 (261.0)
>54,000	635	59.5	145.8 (554.3)	550	62.9	158.9 (584.1)	83	36.1	59.9 (286.8)
hs drop out	326	46.3	170.2 (716.4)	257	54.0	197.2 (794.0)	65	13.8	63.9 (261.4)
hs graduate	613	52.4	137.5 (573.8)	527	57.3	155.1 (613.2)	82	19.5	28.8 (175.2)
some college	736	55.6	109.1 (504.0)	624	58.8	120.0 (538.2)	110	36.4	47.3 (231.3)
college grad	742	48.4	82.2 (310.6)	639	52.0	86.7 (315.3)	100	.25	51.8 (283.0)

notes:
1. All expenditure amounts are adjusted to year 2000 dollars using the Consumer Price Index.
2. Standard errors in parenthesis.

Table 2
Effects of a State Lottery on Two-week Gambling Participation Rate and Expenditures: Coefficient on *LOTSTATE*

CEX Diary Data

	(1) Mean expenses (no lottery)	(2) OLS Level		(3) Mean participation (no lottery)	(4) OLS Any		(5) Probit Any		(6) OLS Ln	
overall	.714	1.43	***	.019	.069	***	.061	***	-.282	***
(n=79,064)	(10.3)	(.353)		(.136)	(.006)		(.004)		(.089)	
income1	.487	.438		.011	.032	***	.035	***	-.395	**
(n=25,538)	(8.67)	(.242)		(.106)	(.005)		(.003)		(.186)	
income2	.561	1.32	***	.019	.069	***	.070	***	-.207	
(n=27,394)	(6.94)	(.309)		(.136)	(.007)		(.003)		(.156)	
income3	1.12	2.45	***	.027	.102	***	.109	***	-.308	**
(n=26,132)	(14.3)	(.863)		(.161)	(.009)		(.004)		(.131)	

notes:
1. Data are from confidential BLS CEX Diary data files from 1984 to 1999.
2. Standard errors are White's robust standard errors adjusted for clustering within a state-year cell.
3. *** indicates significance at 99 percentile ** at 95 percentile
4. A Tobit specification for levels suggests the same patterns. The coefficients are as follows: overall 58.2; income1 31.7; income 2 46.0; income 3 72.9.
5. All regressions include controls for the following household demographics: family size, before-tax income, urban status, number of persons less than 18 and over 64, the sex and educational attainment of the household head, the race of the household head (when it is not the conditioning variable). All regressions also include controls for state, year, month of year, and state cigarette, beer, and gasoline tax levels.

Table 3
Effects of a State Lottery on Gambling Participation: Difference-in-Difference Estimates

Dep Variable	any		lottery		track	bingo	private	unlicensed
overall	.504	***	.429	***	.011	-.002	.009	-.023
(n=2,572)	(.117)		(.036)		(.027)	(.027)	(.034)	(.033)
Income1	.526	**	.448	***	-.008	-.003	.057	-.031
(n=629)	(.217)		(.072)		(.045)	(.053)	(.071)	(.056)
Income2	.836	***	.469	***	.053	.064	.065	.002
(n=991)	(.203)		(.063)		(.043)	(.045)	(.054)	(.055)
Income3	.413	*	.392	***	-.001	-.091	-.056	-.049
(n=952)	(.230)		(.066)		(.059)	(.052)	(.060)	(.066)

notes:
1. Data is from 1998 NORC survey and 1975 Kallick et al. survey.
2. "Any" gambling is not equal to the sum of the five types of gambling displayed because the 1998 file separately categorizes participation in casino, charitable, card, bar/restaurant, internet, and indian reservation gambling.
3. All regressions control for sex, race, marital status, education, and regular attendance at religious services. They also control for main year effects and include a full set of state dummies.
4. Standard errors are White's robust standard errors adjusted for clustering within a state-year cell.
5. *** indicates significance at 99 percentile ** at 95 percentile

Table 4
Effects of a State Lottery on Quarterly Consumption: Coefficient on *LOTSTATE*

CEX Interview Data

	(1) Mean spending (no lottery)	(2) OLS Level		(3) change/ mean	(4) OLS Ln	
overall	7,447.2	-128.1	**	-.017	-.020	***
(*n*=251,239)		(58.3)			(.007)	
income 1	4,707.2	-135.4		-.029	-.029	**
(*n*=81,752)		(82.4)			(.012)	
income 2	6,221.9	-3.75		-.001	-.005	
(*n*=86,332)		(65.4)			(.008)	
income3	11,208.9	-186.5		-.017	-.016	**
(*n*=83,155)		(109.6)			(.008)	

Does bordering a lottery state matter?

	OLS Level			OLS Ln		
	Lotstate		*Lotstate*border*	*Lotstate*		*Lotstate*border*
overall	-279.5	**	180.8	-.034	**	.016
(*n*=251,239)	(115.0)		(121.3)	(.015)		(.016)

Are there short-term and long-term effects?

	OLS Level			OLS Ln			
	Years 1 or 2		*Year 3 & beyond*	*Years 1 or 2*		*Year 3 & beyond*	
overall	-162.2	**	-62.1	-.019	**	-.018	**
(*n*=251,239)	(67.4)		(61.4)	(.008)		(.007)	

notes:
1. Data is from 1982 to 1998 CEX Interview Surveys.
2. Standard errors are White's robust standard errors adjusted for clustering within a state-year cell.
3. *** indicates significance at 99 percentile ** at 95 percentile
4. The lowest income third in the sample distribution is characterized by annual household income <=$9337.4; the highest is >=$26,151.
5. All regressions include controls for the following household demographics: family size, before-tax income, urban status, number of persons less than 18 and over 64, the sex and educational attainment of the household head, the race of the household head (when it is not the conditioning variable). All regressions also include controls for state, year, month of year, and state cigarette, beer, and gasoline tax levels.

Table 5
Effects of a State Lottery on Quarterly Consumption: Coefficient on *LOTSTATE,* by Expenditure Category

CEX Interview Data – Households in the Lowest Income Third

	Mean spending (no lottery)	OLS Level		OLS Any		OLS Ln (c.o.p.)	
total spending	4,707.2	-135.4		—		-.029	**
		(82.4)				(.012)	
1. food at home	751.0	-11.8		-.004		-.033	**
		(10.3)		(.002)		(.016)	
2. medical drugs	68.0	-.459		.018	**	-.053	
		(4.57)		(.009)		(.031)	
3. home – mortgage, rent,	1,427.8	-77.0	**	-.001		-.067	***
other bills		(30.8)		(.002)		(.021)	
4. alcohol	65.0	-6.10	**	-.022	**	-.038	
		(2.67)		(.009)		(.028)	
5. smoking	64.7	-2.98		-.014		-.011	
		(2.22)		(.008)		(.022)	
6. food out	248.3	-7.86		-.003		-.026	
		(9.16)		(.008)		(.024)	
7. entertainment	236.8	4.10		.004		-.028	
		(15.6)		(.007)		(.021)	
8. education	119.5	2.50		.007		.029	
		(11.2)		(.007)		(.066)	
9. house - repairs,	374.8	-10.7		-.00002		-.056	
services, furnishings		(19.2)		(.010)		(.033)	
10. clothes	264.7	-10.08		-.011		-.022	
		(8.75)		(.008)		(.026)	
10a. kids	39.9	-.485		-.001		-.009	
		(1.97)		(.006)		(.033)	
10b. adult	224.8	-9.60		-.010		-.023	
		(7.74)		(.008)		(.026)	
11.transportation/cars	1,086.5	-15.0		-.008		-.005	
		(42.8)		(.005)		(.023)	
sample size	81,752						

notes:
1. Data is from 1982 to 1998 CEX Interview Surveys.
2. Standard errors are White's robust standard errors adjusted for clustering within a state-year cell.
3. *** indicates significance at 99 percentile ** at 95 percentile
4. The lowest income third in the sample distribution is characterized by annual household income <=$9337.4.
5. All regressions include controls for the following household demographics: family size, before-tax income, urban status, number of persons less than 18 and over 64, the sex and educational attainment of the household head, the race of the household head (when it is not the conditioning variable). All regressions also include controls for state, year, month of year, and state cigarette, beer, and gasoline tax levels.

Table 6
Weekly Ln Lotto Sales per Adult as a Function of Game Attributes

dep var: ln(pasales)	(1)		(2)		(3)		(4)		(5)	
Expected value	.377		.683	***	.757	***	.299	***	.346	***
	(.406)		(.136)		(.126)		(.060)		(.060)	
Variance/1M.	.040	***	.003		-.006		.010	***	.004	**
	(.010)		(.006)		(.004)		(.001)		(.002)	
skewness /1T	-.0002	***	.000008		.00003		-.00004	***	-.00001	*
	(.00005)		(.00004)		(.00003)		(.00001)		(.000007)	
nominal top prize/1M.	-		-		.007	***	-		.004	***
					(.002)				(.0008)	
no. draws per week	-		-.059	**	-.052	**	-		-	
			(.024)		(.024)					
age of game	-		-.133	***	-.126	***	-.201	***	-.206	***
			(.041)		(.041)		(.076)		(.076)	
(age of game)^2	-		.020	***	.020	***	.022	***	.023	***
			(.004)		(.004)		(.003)		(.003)	
pick 5	-		.828	***	.785	***	-		-	
			(.154)		(.155)					
pick 6	-		.398	***	.401	***	-		-	
			(.151)		(.149)					
pick 7	-		.857	***	.823	***	-		-	
			(.182)		(.177)					
cash jackpot	-		-1.13	***	-1.11	***	-		-	
			(.149)		(.143)					
choice (cash/ann)	-		.290	*	.214		-		-	
			(.156)		(.157)					
state unemployment rate	-.030		-.029		-.030		-.029	*	-.029	*
	(.025)		(.019)		(.019)		(.015)		(.015)	
product fixed effects	no		no		no		yes		yes	
state fixed effects	yes		yes		yes		yes		yes	
week fixed effects	yes		yes		yes		yes		yes	
*demog. controls**	yes		yes		yes		yes		yes	
constant	-380.8	***	25.1		11.7		-129.1		-127.7	*
	(142.9)		(75.2)		(72.9)		(66.7)		(66.8)	
sample size	14,669		13,930		13,930		13,930		14,669	
R^2	.61		.89		.89		.92		.91	

notes:
1. Unit of observation is state-week-game.
2. The sample includes 91 lotto products from 33 states.
3. Standard errors are adjusted for clustering at the state-year level, to flexibly account for correlations among errors.
4. Lottery sales data are from Lefleurs inc.
5. Data on game characteristics is compiled by author using information provided by state lottery associations.
6. Monthly state unemployment data are from the Bureau of Labor Statistics.
7. All regressions are population weighted. All regressions control for the proportion of the state population in the following categories: females age 18-24, 25-44, 45-64, 64+, males age 18-24, 25-44, 65+. Yearly state population figures are from the U.S. Census Bureau.

Appendix Table 1
State Lottery Implementation, by Year

1964	New Hampshire
1967	New York
1970	New Jersey
1972	Connecticut, Massachusetts, Michigan, Pennsylvania
1973	Maryland
1974	Illinois, Maine, Ohio, Rhode Island
1975	Delaware
1978	Vermont
1981	Arizona
1982	District of Columbia, Washington
1983	Colorado
1985	California, Iowa, Oregon
1986	Missouri, West Virginia
1987	Montana, Kansas, South Dakota
1988	Virginia, Florida, Wisconsin
1989	Idaho, Indiana, Kentucky
1990	Minnesota
1991	Louisiana
1992	Texas
1993	Nebraska, Georgia
1996	New Mexico

Appendix Table 2
Lottery Sales (in Year 2000 Dollars)

	Mean state sales		All states		
	Monthly total (in millions)	Monthly per adult sales	Yearly total (in millions)	No. of states (inc. DC) with lotteries	No. of states reporting sales
Overall	78.8	18.3	33,409	-	-
1992	67.8	16.0	24,207	35	32
1993	80.8	17.5	31,574	37	34
1994	86.3	18.9	34,158	37	33
1995	78.1	18.9	34,671	37	37
1996	81.0	18.5	34,981	38	36
1997	78.7	18.3	34,951	38	37
1998	77.1	18.9	34,287	38	38

notes:
1. Lottery sales data is from Lefleurs inc., who collects information from state lottery agencies.
2. Population figures used for per adult calculations are BLS census population numbers.
3. These figures reflect sales on all lottery games, including lotto, multi-state lotto, numbers, instant, keno, sports, bingo, and VLT products.

Chapter 2

Preferences under Risk: The Case of State Lottery Bettors

1 Introduction

The attitude toward risk of economic agents has long been the object of considerable attention. Understanding how economic actors make decisions under risk is crucial to applications of microeconomics and macroeconomics alike. For example, attitudes toward risk are fundamental to the optimal design of social insurance programs and to explaining the behavior of stock markets. Economists have traditionally modeled decisions under risk according to the tenets of expected utility theory. But a recent body of literature challenges this traditional model and proposes new explanations of how economic agents make decisions in the face of uncertainty.

This chapter provides an empirical evaluation of the descriptive power of expected utility theory. It investigates whether observed consumer demand for state lottery gambles is described by the predictions of expected utility theory, or whether the data reject the expected utility hypothesis in favor of some of the generalizations proposed by *cumulative prospect theory.* In Chapter one, I estimate reduced form demand equations as a function of the characteristics of lottery games. This chapter takes that reduced-form analysis further by putting structure on the problem in order to link consumer demand to the parameters governing consumer preferences.

I examine demand for a sample of U.S. state lottery games offered during 1992 to 1999. The majority of U.S. adults participate in lottery gambling, spending nearly 40 billion annually on state lottery tickets. The choices we observe thus reflect decisions made in a "real-world" context by a general population. Observations on this type of decision making are arguably more useful for the purpose of drawing conclusions about how economic actors behave than is information obtained from laboratory or classroom experiments. While the experimental literature on choices under uncertainty is abundant, there are virtually no studies of non-expected utility models that use real-world data. Jullien and Salanie's (2000) study of British horse races is a notable exception.

To investigate how consumers make choices under risk, I specify the consumer optimization problem over lottery ticket purchases. The consumer is assumed to maximize the sum of the expected utility of the gamble and the entertainment value of playing the game. To be clear, lottery gambling is modeled as both investment and consumption. I begin by specifiying the problem under the hypothesis of the expected utility representation of consumer preferences. In particular, I assume probabilities enter the problem linearly. From the optimization problem, I derive consumer demand as an implicit function of the observed characteristics of the gamble and the unknown parameters describing consumer preferences. This demand equation structurally links observed demand and game characteristics to the parameters of the choice problem. I estimate the parameters of this equation using generalized method of moments-instrumental variables (GMM) techniques and weekly data on lottery game sales and characterisitcs.

The parameter vector is over-identified and a test of the over-identifying restrictions provides a test of the hypothesis of the model. The over-identification test fails to reject the expected utility specification of the model. The failure to reject the expected utility hypothesis is robust across three common specifications of utility: *constant relative risk aversion* (CRRA) utility, *constant absolute risk aversion* (CARA), and cubic utility. Under expected utility theory and the assumption of CRRA utility, agents are found to be slightly risk-averse; point estimates for the coefficient of relative risk aversion are 0.460 and 1.12, depending on the functional specification of the entertainment value.

As an additional test of the descriptive value of expected utility theory, I derive parameter estimates under an alternative representation of preferences suggested by cumulative prospect theory and directly test the restrictions imposed by expected utility theory. The probabilities of gains and losses are allowed to enter the problem non-linearly and differently from one another. In addition, value is defined over gains and losses, rather than final wealth. I re-derive consumer demand as an implicit function of lottery game characteristics and consumer choice parameters. I then similarly estimate the choice parameters using GMM techniques. Under this less restrictive specification, the data reject the expected utility restriction that probabilities enter the consumer problem linearly. Under the assumption of either CARA or CRRA utility, the data suggest a convex weighting function of the probability over gains. This supports the finding of Jullien and Salanie (2000), who estimate a convex weighting function of win probabilities using data on British racetrack betting. But it presents us with a puzzle. Assuming CRRA utility, point estimates for the coefficient of relative risk aversion are .262 and 2.26. If economic agents are risk averse and apply a convex transformation to small win probabilities, then why do they play the lottery at all? It must be the case that participation in lottery gambling is not driven by an erroneous or inflated idea about winning, but rather by the entertainment value of playing the game.

The chapter proceeds as follows. Section 2 provides a background discussion of expected utility theory and the debate about whether it accurately describes consumer choice under risk. Section 3 discusses related evidence from previous empirical work. Section 4 describes the model of lottery betting that underlies the emprical analysis. Section 5 solves the model and describes the estimation procedure. Section 6 presents the results. And finally, Section 7 concludes with suggestions for future work.

2 Background

For most of the twentieth century economists have modeled consumer choice under risk as though it were governed by the tenets of von Neumann-Morgenstern expected utility theory. The reach of expected utility theory in economics can not be overestimated. Public finance economists rely on it to model life-cycle savings decisions and insurance purchases. Macroeconomists use it to model portfolio investment decisions. Development economists base models of village money-lending on it. These are just a few of the numerous varied applications. However, in the last decade, the dominance of the expected utility paradigm in economics has been challenged.

This challenge is largely based on a growing body of experimental evidence that individuals do not maximize expected utility. A number of alternative so-called non-expected utility models of decision making have been developed. This section begins with a review of the basic axioms and predictions of expected utility theory. It then describes some of the main violations of expected utility theory that have been observed by experimental researchers. And finally, it describes an influential alternative theory.

2.1 Expected utility theory

In the theory of von Neumann-Morgenstern expected utility, uncertain prospects are modeled as probability distributions over a given set of prizes and probabilities are objective. A simple probability distribution p on X is specified by (i) a finite subset of X, called the support of p and denoted by supp(p), and (ii) for each x of supp(p) a number $p(x) > 0$, with $\sum_{x \in supp(p)} p(x) = 1$. Let the set of simple probability distributions on X be denoted by P. The outcomes $x_i, ...x_n$ could represent alternative final wealth levels, alternative changes from the individual's current wealth level, or alternative non-monetary outcomes. In the terminology of probability theory, the term "lottery" refers to a simple probability distribution - for example, the lottery p gives prize x with probability 1/3 and prize y with probability 1/5. The reader should not confuse the use of the term in this section with the empirical application of state lotteries discussed elsewhere in the chapter.

As in standard consumer theory, an individual is assumed to have preferences over the set P of all simple probability distributions on X given by a relation $\succ$ that expresses strict preference. The following three axioms are assumed for the relation $\succ$:

> Axiom 1: $\succ$ must be asymmetric and negatively transitive.

We also construct a weak preference $\succeq$ and an indifference relation $\sim$. So, if p and q are two simple lotteries, we have either $p \succ q$, $p \prec q$, or $p \sim q$.

> Axiom 2: Suppose p and q are two probability distributions such that $p \succ q$. Suppose a is a number from the open interval $(0, 1)$, and r is some other probability distribution. Then $ap + (1 - a)r \succ aq + (1 - a)r$.

Kreps (1990) calls this the *substitution axiom.* The consumer's preference between p and q is not affected by the addition of an alternative contingency.

> Axiom 3: Suppose that p, q, and r are three probability distributions such that $p \succ q \succ r$. Then numbers a and b exist, both from the open interval $(0, 1)$, such that $ap + (1 - a)r \succ q \succ bp + (1 - b)r$

This axiom - referred to by Kreps (1990) as the *Archimedean axiom* - says that since p is strictly preferred to q, no matter how undesirable r is, there is some a close to one such that $ap + (1-a)r$ is preferred to q. And furthermore, no matter how much p is preferred to q, there is some b close to zero such that $bp + (1-b)r \prec q$.

These three axioms lead to the following proposition:

> *Expected utility theory proposition*:
>
> A preference relation $\succ$ on the set P of simple probability distributions on a space X satisfies axioms 1, 2, 3 above if and only if there is a function $u : X \to R$ such that $p \succ q$ if and only if $\sum_{x \in supp(p)} u(x)p(x) \succ \sum_{x \in supp(p)} u(x)q(x)$. Moreover, if u provides a representation of $\succ$ in this sense, then v does as well if and only if constants $a > 0$ and b exist such that $v(.) = au(.) + b$, i.e. v is an affine transformation of u.

As Kreps (1990) explains, this proposition establishes the existence of a numerical representation for preferences on p; i.e., there is a function $U : P \to R$ such that $p \succ q$ if and only if $U(p) > U(q)$. Furthermore, this function U takes the form of expected utility over prizes: $U(p) = \sum_{x \in supp(p)} u(x)p(x)$, for some $u : X \to R$.

Economic applications of expected utility theory often specialize to the case where the prizes are dollar values. When talking about preferences over money, economists use some basic vocabulary that I use throughout the chapter. I define some of the terms here, using the definitions provided by Kreps (1990). It is commonly assumed that individual preferences exhibit *risk aversion,* which is equivalent to assuming concave u. It implies that a consumer prefers the certainty of the expected value of a gamble to the risky gamble itself. In contrast, a *risk-loving* consumer prefers the risky gamble. A consumer is *risk-neutral* if he is indifferent between the two. The *certainty equivalent*, denoted $C(p)$, for a lottery p is defined as the prize x such that the consumer is indifferent between receiving x with certainty and accepting the gamble p. Note that in the case of risk aversion, $C(p) < Ep$, where Ep denotes the expected value of p. The difference $Ep - C(p)$ is called the *risk premium* and denoted $R(p)$. For a fixed consumer with utility function u, if $R(p)$ is nonincreasing in the final wealth level, the consumer is said to be *nonincreasingly risk averse* (or less-formally *decreasingly risk averse*); if $R(p)$ is constant in the final wealth level, the consumer is said to have *constant risk aversion*; and if $R(p)$ is nondecreasing in the final wealth level, the consumer is said to have *nondecreasing risk aversion.*

2.2 Violations of expected utility theory

Economists have long noted that certain behaviors do not conform to the joint hypothesis of utility maximization and diminishing marginal utility. In their classic 1948 paper, Friedman and Savage address the phenomena of people both buying insurance and participating in gambling. Friedman and Savage propose that this seemingly-contradictory behavior could be rationalized by placing a convex segment in the middle range of an otherwise concave utility function. This implies that gamblers place a high value on the chance to increase their wealth greatly and thereby move into a new class of wealth. People in the first concave segment are predicted to purchase low probability, high payoff gambles that reach into the convex segment, while simultaneously insuring against wealth-decreasing risks. The model predicts that poor people will play the lottery and rich people will abstain. The model also has the unlikely prediction that middle income people will gamble as much as possible. Markowitz (1952) addresses this unfortunate prediction by placing the convex segment at current wealth, and treating gambling as an exploitation of local risk preference.

A weakness of the Friedman-Savage thesis is that it has been unable to explain other well-documented gambling phenomena. For example, the observed *long-shot bias* - which leads bettors to demand a higher expected return for a bet on a favored horse relative to a horse with a low win probability - is accentuated at the end of the day. However, these and other such tendencies can be rationalized under *cumulative prospect theory*, which is discussed below.

There is a growing body of experimental evidence that individuals' preferences over lotteries do not conform to the fundamental assumptions leading to the expected utility representation of preferences. One of the most well-known violations of expected utility theory is the *Allais paradox.* Machina (1989) borrows the following example from Allais (1953). Consider the following pair of decision problems:

$a1$: 1.00 chance of \$1 million

versus

$a2$: .10 chance of \$5 million; .89 chance of \$1 million; .01 chance of \$0

and

$a3$: .10 chance of \$5 million; .90 chance of \$0

versus

$a4$: .11 chance of \$1 million; .89 chance of \$0

As Machina (1989) reports, in Allais' original study as well as subsequent studies that have given this problem to subjects, the modal choice is invariably *a1* in the first pair and $a3$ in the second pair. But the substitution axiom requires that if *a1* is preferred in the first pair, *a4* must be preferred in the second pair, and vice versa. The substitution axiom says that an individual's preference between a .11 chance of $1 million versus a .10 chance of $5 million and a .01 chance of $0 is independent of what happens in another contingency. But the choices suggest otherwise. Algebraically, a preference for *a1* over *a2* implies

$$u(\$1\text{M}) > .10u(\$5\text{M}) + .89u(\$1\text{M}) + .01u(\$0)$$

or equivalently

$$.11u(\$1\text{M}) + .89u(\$0) > .10u(\$5\text{M}) + .01u(\$0) + .89u(\$0\text{M}).$$

This says that the consumer likes substituting a .11 chance of winning $1 million for a .10 chance of $5 million and a .01 chance of $0. But, a preference for $a3$ over $a4$ implies

$$.11u(\$1\text{M}) + .89u(\$0) < .10u(\$5\text{M}) + .01u(\$0) + .89u(\$0),$$

which says that the consumer does not like making that substitution. Kahneman and Tversky (1979) explain this violation in terms of a *certainty effect*: people overweight outcomes that are considered certain, relative to outcomes that are merely probable.

In addition to the Allais paradox class of violations, experimental researchers have documented a *framing effect*, whereby choices made under uncertainty are affected by a shift in the "status quo" (see Kahneman and Tversky (1979)). Kreps (1990) uses the following example to illustrate one type of framing. Consider the following two pairs of compound gambles:

$b1$: .34 chance of the lottery: 33/34 chance of $27,500; 1/34 chance of $0
.66 chance of $24,000

versus

$b2$. .34 chance of $24,000
.66 chance of $24,000

and

$b3$: .34 chance of the lottery: 33/34 chance of $27,500; 1/34 chance of $0.
.66 chance of $0

versus

$b4$: .34 chance of \$24,000

.66 chance of \$0

The consistent choices are either $b1$ and $b3$ or $b2$ and $b4$, but subjects tend to focus on the "differences" between the two in each pair. Whether or not the "status quo" is perceived to be \$0 or \$24,000 influences preferences, and it is not uncommon for subjects to select $b2$ and $b3$. Again, this is a violation of the substitution axiom.

Subsequent experiments have found additional departures from expected utility. Psychological experiments have routinely documented what Kahneman and Tversky (1979) term *loss aversion*: people are more averse to losses than they are to attracted to same size gains. While this finding is consistent with concave utility under expected utility theory, the additional finding is that the value function kinks at the reference level, so that people are significantly risk averse for even small amounts of money. People dislike losing \$10 more than they like gaining \$11, and hence prefer their status quo to a 50/50 bet of losing \$10 or gaining \$11. This suggests that something more than simple risk aversion is taking place in people's choices. As a clear illustration of this point, Rabin (1999) develops a calibration theory showing that expected-utility theory is "an utterly implausible explanation for appreciable risk aversion over modest stakes."

There is also ample evidence that people make systematic errors in their assessments of probabilities, thereby biasing their judgment in situations of uncertainty. The literature distinguishes between risk and uncertainty: the probabilities assigned to outcomes in decisions of risk are objective, and in decisions of uncertainty are subjective. The examples discussed above have focused on departures from expected utility theory in risky situations. The other broad set of experiments illustrating departures from expected reveal that people do not evaluate subjective probabilities as assumed. Rabin (1996) provides a review of biases in judgement under uncertainty, showing that people make systematic errors in their attempts to maximize their preferences. Because the empirical application of this paper investigates the nature of choices in a risky environment (the probability of selecting the winning lottery numbers is an objective probability), I have not focused on these departures from expected utility theory.

2.3 *Cumulative prospect theory*: an alternative to expected utility theory

Motivated by some of the aforementioned violations of expected utility theory, researchers have developed nonlinear ("non-expected utility") functional forms for individual preference functions over lotteries. Arguably the most influential of these is Kahneman and Tversky's (1979) *prospect theory*, later developed into *cumulative prospect theory*(CPT) in Tversky and Kahneman (1992). This alternative theory of choice assigns value to gains and losses, rather than final assets. As they explain, our perceptions are sensitive to the evaluation of changes in differences, rather than to exact magnitudes. This is true for brightness, loudness, or temperature, and as they explain, for levels of wealth as well. They point out that the same level of wealth may imply poverty for one person and great riches for another, depending on their original position. (See Rabin (1996) for a more extensive discussion of *reference-dependent preferences.*)

Prospect theory makes the additional claim that the value function is concave for gains and convex for losses, and generally steeper for losses than gains. This relates to the observation mentioned above that people are significantly risk averse for even small amounts of money. These preferences suggest a kink in the utility function at the reference point. Allowing for concavity above the kink and convexity below reflects the observed diminishing sensitivity of preferences; that is, the marginal change in perceived well-being is greater for changes that are closer to one's reference level than those that are further away.

The other major feature of cumulative prospect theory is that it replaces the objective probabilities of the expected utility representation with decision weights. Decision weights are assumed to be lower than the corresponding probabilities in the case of high probabilities and higher than corresponding low probabilities. The weighting function w is a strictly increasing function from [0,1] to [0,1] with $w(0)=0$. Kahneman and Tversky (1979) speculate that overweighting of low probabilities may contribute to the attractiveness of both insurance and gambling.

3 Related evidence from previous empirical studies

There is a large quantity of research devoted to studying consumer choice under risk. However, there is scant research testing competing theories of risk on real-world experiments. Research on the predictive power of alternative theories of choice has been almost exclusively conducted with laboratory or classroom experiments. The non-experimental research on choice under risk has focused on estimating degrees of risk aversion under the assumption of expected utility theory.

Jullien and Salanie (2000) is a notable exception. Their paper moves beyond the basic expected utility framework and explores various alternative models of choice under uncertainty in a nonlaboratory environment. They estimate a multinomial model on a data set of 34,443 British horse races. Under the assumptions of von Neumann-Morgenstern expected utility theory, they find empirical support for the CARA specification of utility and find that bettors are slightly risk loving. Maintaining the assumption of CARA utility, they find that cumulative prospect theory has higher explanatory power than expected utility theory. However, their findings for the probability weighting functions do not corroborate the experimental evidence. The weighting function appears to be slightly convex for gains and highly concave for losses. They find little evidence for changing concavity of the probability weighting functions. Though experimental evidence has documented a "certainty effect," their estimated weighting function for loss probabilities does not become convex for probabilities close to one. Nor does the estimated convex weighting function for win probabilities become concave close to zero, as overweighting of low probabilities would imply. (Jullien and Salanie (2000) pp. 516-17)

Jullien and Salanie (2000) build on a tradition of studying the behavior of racetrack bettors, though previous studies are limited to the expected utility framework. Ali (1977) analyzes the betting behavior in 20,247 harness horse races run in New York. He defines the subjective win probability as an average of the proportion of the total money bet on a horse and the objective win probability as the proportion of times a horse finishes first. His analysis shows that the subjective win probability for a horse with a high objective win probability is understated and the subjective win probability for a horse with a low objective probability is overstated. This is the *long-shot bias*, alluded to in Section three above. Ali concludes that the representative bettor is risk-loving, and that the degree of risk affinity increases in wealth. Previously, Weitzman (1965) had documented the long-shot bias in his study of over 12,000 thoroughbred races.

Though it is not the main contribution of my paper, when estimating the parameters describing consumer preferences, I estimate risk aversion parameters. I therefore review some previous estimates here to put my estimates in context of previous findings. In a fairly recent study, Cicchetti and Dubin (1994) analyze market data on whether individuals purchase insurance against the risk of telephone line trouble in the home. They find support for the hyperbolic absolute risk aversion (HARA) class of utility functions and a small degree of absolute risk aversion. Earlier studies suggest a greater degree of consumer risk aversion. Farber (1978) models union behavior as the maximization of the expected utility of the median-aged member of the union. Using maximum likelihood techniques and data from 1948 to 1973, he estimates a coefficient of relative risk aversion greater than 2.5. Friend and Blume (1975) examine cross-sectional survey data from 1962 and 1963 on household asset holding and find that constant relative risk aversion (CRRA) utility fits the data relatively well, and better than log utility in particular. They estimate a coefficient of relative risk aversion in excess of one, and "probably in excess of two." Szpiro's (1986) analysis of times series data on U.S. insurance premiums supports Friend and Blume (1975). His analysis supports the hypothesis of CRRA utility and estimates the coefficient of relative risk aversion to be between 1.2 and 1.8.

4 Model and data

The real-world experiment of state lotteries provides us with thousands of observations on consumer choice under risk. This chapter uses a sample of these observations to estimate the parameters describing consumer preferences. In Chapter one, I estimated reduced-form demand equations as a function of the characteristics of lottery games. This chapter extends that reduced-form analysis by putting structure on the problem and thereby identifying the parameters of underlying preferences. I start with the consumer optimization problem and derive consumer demand as a function of gamble characteristics - namely, the jackpot prize and probability of winning - and the parameters of the choice problem. This yields an equation linking observed demand to the unknown parameters of the choice problem. I then empirically estimate the choice parameters and draw conclusions about how consumers are making decisions under risk.

State lotteries provide a natural laboratory for studying consumer choice under risk. An important advantage of the state lottery context over other "real world" contexts - including racetrack betting - is that the odds of winning are known equally across bettors. The probability of selecting the winning numbers is fixed and is listed on game tickets, state lottery websites, and in game brochures. In contrast, bettors on horse races have subjective probabilities that can differ from the bookmakers odds, and differentially so across consumers. Another important advantage is that sales information is available by game and week. I can therefore observe how the amount bet varies with the jackpot amount. Jullien and Salanie (2000) do not have information on the amounts bet in different horse races, so they must assume that all bettors bet the same amount in all circumstances and that they do not spread their bets among several horses.

Unfortunately, even the state lottery context is not ideal. State lottery sales data are not available for individual bettors. This paper therefore has the same major weakness as Jullien and Salanie's study of racetrack bettors: it relies on a representative agent assumption to draw conclusions about individual preferences.

4.1 Framework of the model

I model the consumer optimization problem of lottery ticket purchases in two ways. First I specify the optimization problem under the assumption that preferences are governed by the axioms underlying the expected utility representation. I solve the optimization problem and derive an equation that implicitly determines lottery ticket demand. This equation provides us with an empirical test of the fit of the model. As discussed below, I estimate this equation using GMM techniques and data on state lottery sales and gamble parameters. A chi-squared test is used to establish whether the data reject the hypothesis of the model specification. The expected utility equation is estimated under alternative assumptions of the functional form of the utility function $u(\cdot)$ to insure that the analysis is testing the restrictions of the expected utility representation of preferences and not a particular functional form assumption of utility.

After solving implicitly for lottery demand under the expected utility hypothesis, I specify the optimization problem under the assumption that preferences are described by cumulative prospect theory (CPT). From the consumer optimization problem, I derive an equation that implicitly determines lottery ticket demand under this alternative hypothesis. This equation is similarly estimated with GMM techniques. As CPT is a generalization of expected utility theory, failure to reject the expected utility hypothesis should imply failure to reject the CPT hypothesis. The data could, however, reject expected utility and fail to reject CPT. In addition, estimating the equation in this more general form provides a direct test of the expected utility assumption of linear probabilites.

As with the reduced-form demand analysis of Chapter 1, I limit this study to demand for lotto games to the exclusion of other types of lottery products. Lotto games vary substantially in prize amounts and structure, much more so than other types of lottery products. There is variation across games as well as within games over time. In this study, I only focus on lotto games that "rollover" and accumulate when the winning numbers are not chosen in a particular week. The reasons for this exclusion are discussed below. An additional advantage to looking only at lotto games is that it makes the representative agent assumption much less stringent. The assumption of a representative consumer playing big-prize lotto games is a much more realistic assumption than a representative consumer playing all types of lottery games, including, for example, both big-prize lotto games and instant scratch-off tickets. In addition, a single function $e(\cdot)$ - defined below - mapping the entertainment value for lotto games is more plausible than a single function mapping the entertainment value for various game types. In the remainder of the paper, the use of the word "lottery" should be understood to refer to state lotto games.

I focus on the consumer choice between two games offered concurrently, rather than on the total allocation decision or on the decision to bet x dollars on a particular game. In a state that offers two lotto games in the same week, X_1 and X_2, a player purchases x_1 tickets on lotto X_1 and x_2 tickets on lotto X_2, such that on the margin he is indifferent between the two games. A one dollar lottery ticket purchase is tantamount to a one dollar bet. In the data, we observe this choice for a particular pairing of games over a series of many weeks as the lottery jackpots accumulate and "reset." There is also variation across states in the set of games offered in a particular week.

In reality, consumers tend to play multiple games, so the consumer problem is not a discrete choice problem but rather an allocation problem. Specifying the problem as an allocation decision between two competing lotto games, as opposed to a total or single game allocation decision, facilitates the identification of model parameters. In particular, any factors that drive total lottery sales and are unrelated to the lottery games themselves, are "differenced out" of the problem. For example, we do not need to be concerned that in a particular state and week total lottery gambling is depressed because of bad weather or because people are betting instead on a big college basketball game. In these and other similar situations, the allocation of dollars between the two games is unaffected; the consumer purchases tickets for the two games such that on the margin he is indifferent between them. This is true whether he is spending more or less than usual on lottery tickets. Additionally, the consumer optimization problem is considered to be a one-period game. This is important because the extension of non-expected utility models to dynamics is still controversial (Jullien and Salanie (2000), Machina (1989)).

The model allows a lottery ticket to constitute both a risky asset and an entertainment good. The consumer is endowed with preferences over risk given by the functional $U(F, \beta)$,where F is a risky distribution and β is a parameter vector describing the bettor's preferences. The distribution of returns F_j associated with game X_j, where $j \in \{1, 2\}$, is characterized by the game's jackpot and probability of winning the jackpot. Let R_j denote the jackpot in game X_j and let p_j denote the probability of betting the winning numbers on a single ticket. I assume that if a player purchases x tickets, then he plays x numbers. When a consumer chooses x_j, he is essentially controlling his probability of winning R_j; the probability of betting the winning numbers is $p_j x_j$.

Consumer utility is specified over the lotto jackpot and not over any lower prize tiers. Though obviously incomplete, this is a reasonable simplification. The justification for this simplified characterization is that the jackpot tends to be orders of magnitude greater than lower prize tiers. Among the single-state lotto games for which I collected jackpot and odds data, the average jackpot has a present discounted value of \$1.1 million; the average second prize among these games is \$1,474. Among the multi-state lotto products, the average jackpot has a presented discounted value of \$11.7 million, compared to an average second prize of \$77,000.[1] For the sake of completeness, this simplification should eventually be abandoned and lower prize tiers should enter into the characterization of F_j.

In addition to the value $U(F_j, \beta)$ associated with playing game j, the consumer derives an additive entertainment value e_j that is independent of potential wealth creation. The value of a lottery bet on game X_j is thus $U(F_j, \beta) + e_j$. The results of the reduced-form demand analysis in Chapter 1 suggest that it is important to allow for an entertainment value of the gamble. Recall that consumer demand responds to game characteristics irrelevant to wealth creation, such as the number of drawings per week and the number of digits chosen by the bettor. The function e_j captures the "thrill" of gambling that we might think comes from such pleasures as picking one's lucky numbers or thinking about how one would spend a million dollars.

Let us define $e(x_j)$ such that $e(0) = 0$; $e(1) > 0$; $\frac{\partial e(x_j)}{\partial x_j} > 0$; $\frac{\partial^2 e(x_j)}{\partial x_j^2} < 0$. The first condition says that the consumer derives no entertainment from the lottery game until he purchases a ticket (i.e., bets one dollar). The second condition says that there is positive entertainment associated with a single dollar bet. (This condition rules out the specification of e as the natural logarithm of dollars bet.) The third condition says that the marginal entertainment value of a dollar bet is declining in the number of dollars bet. This accords with our intuition. In addition, it guarantees that the optimization problem has an interior solution. If the marginal entertainment value were not declining, a corner solution could be optimal - the consumer could optimally place all his lottery bets on the game with the highest expected utility, $U(F_j, \beta)$. A useful extension to the model would be to relax the assumption of additivity of $U(F_j, \beta)$ and e_j. A second useful extension would be to model e_j explicitly using game characteristics or game fixed effects, for example, by using the estimates from the reduced-form demand analysis of Chapter 1.

[1] These averages are for the sample used in the demand analysis of Chapter 1 which is a larger sample than that used in this chapter.

4.2 Data

The empirical analysis of this paper looks at cross-sectional time-series data. Each observation represents a pairing of two lotto games in a week. I observe sales, jackpot, and win probabilities for these pairings across states and over time. The sample of games comprising these pairings are a subset of the lotto games used in the reduced-form demand analysis of Chapter 1. I describe that dataset only briefly here. The sales data were purchased from Lefleur's Inc., a group that collects sales data from state lottery agencies. I compiled historical data on advertised jackpot amounts from sources provided by various state lottery agencies. I exclude from the sample observations for which only realized jackpot data is available, since the realized amount is a function of both the rollover amount and the induced additional sales. I obtained information about the odds of winning from state lottery websites when available, and directly from state lottery agencies in cases of discontinued games. All dollar amounts are adjusted to year 2000 dollars using the Consumer Price Index.

I place two additional restrictions on the sample used in this analysis. First, I only include lotto games that have "rolling" jackpots. The rules of these games are such that when the winning numbers drawn by the lottery agency do not match the numbers on any of the tickets bet, the jackpot is rolled over into the jackpot for the next drawing (minus the prize payments for partially correct bets). Since the drawing of winning lottery numbers is random, limiting my sample to rolling jackpot games assures the exogeneity of the jackpot to consumer preferences. This will be important for identification. I define Rj as the present discounted value of the advertised jackpot amount. (State lottery agencies advertise a game's jackpot as the undiscounted sum of the annual payments.)

Second, the sample is limited to state and week observations that have positive sales on at least two lotto games. My dataset has sales, jackpot, and odds data for 2,926 such state-week cells: 2,308 with two games, 559 with three games, and 59 with four lotto games. For state-week cells with complete data for more than two lotto games, I construct an observation for each pairing. The final sample includes 4,339 observations from 23 states. Each observation represents a pairing of game X_1 and X_2 in a particular week.

5 Solving the model

In this section, I model the consumer optimization problem of lottery ticket purchases in two ways. First I specify the optimization problem using the expected utility representation of preferences. I solve the optimization problem and derive an equation that implicitly determines consumer demand for lottery tickets as a function of F, the risky distribution of the gamble, and β, the parameter vector describing the bettor's preferences. Second, I specify the optimization problem under the cumulative prospect theory representation of consumer preferences and similarly derive an equation that implicitly determines lottery ticket demand as a function of F and β. These two equations identify β in terms of the observed gamble characteristics under the two competing theories. They provide the basis for the empirical tests of how well each of the theories describe the consumer choices observed in the data.

5.1 Solving the model under the hypothesis of expected utility

Assume that the functional $U(F_j, \beta)$ takes the expected utility representation:.

$$U(F_j, \beta) = \int u(\cdot, \beta) dF_j(\cdot)$$

The value to the consumer of betting on game $j \in \{1, 2\}$ is $U(F_j, \beta) + e_j$. This consumer is assumed to be a representative consumer who maximizes his utility in each "market" $i \in \{1, ..., 4339\}$, where a market is defined as a pair of lotto games in a particular week. Recall that the jackpot amount varies weekly. The probability of betting the winning numbers on a single ticket does not necessarily change weekly, but state agencies do change the structure of games over time, so there is some time-series variation in the win probability associated with a particular game.

The optimization problem facing the representative consumer in a market i can be written as follows:

$$\begin{aligned} &\max_{x_{1i}, x_{2i}} \; p_{1i} x_{1i} u(M_i + R_{1i} - x_{1i} - x_{2i}) + p_{2i} x_{2i} u(M_i + R_{2i} - x_{1i} - x_{2i}) + \\ &(1 - p_{1i} x_{1i} - p_{2i} x_{2i}) u(M_i - x_{1i} - x_{2i}) + e_1(x_{1i}) + e_2(x_{2i}) \qquad (1) \end{aligned}$$

where M_i is the consumer's wealth when he faces the i^{th} pair of lotto games.

This equation says that the consumer chooses the amount of tickets he buys on games X_1 and

X_2 by maximizing the sum of the expected utility over the two games plus the entertainment value of betting x_1 and x_2 dollars on the games. The expected utility is defined over final wealth, which is initial wealth M, minus the amount spent on lottery tickets $(x_1 + x_2)$, plus the jackpot in the event of a win (R_1 or R_2). Data on inital wealth levels are from the U.S. Census Bureau and are three-year moving averages of median household income by state. The data is adjusted to year 2000 dollars using the Consumer Price Index. Ideally, initial wealth would be observed for each of the 4,339 markets. Unfortunately, the data is only available at the level of state and year, not state and week and is thus a crude measure of wealth.

Equation (1) incorporates the fundamental assumption of expected utility theory that probabilities enter the consumer optimization problem linearly. As seen in the equation, a consumer controls his probability of winning jackpot R_j through his selection of x_j; the probability of betting the winning numbers on game X_j is $p_j x_j$. Equation (1) incorporates two simplifying assumptions about winning. The first is that the probability of winning both jackpots is negligible. The second is that the probability of sharing the jackpot - i.e., of multiple winning tickets in a single drawing - is negligible. Relaxing these assumptions, particularly the second, would be useful extensions to the model.

Approximating $R - x_1 - x_2$ with R and setting $\frac{\partial L}{\partial x_1} = \frac{\partial L}{\partial x_2}$ yields the following[2]:

$$0 = p_{2i}u(M_i+R_{2i})-p_{1i}u(M_i+R_{1i})-(p_{2i}-p_{1i})(1-p_{1i}x_{1i}-p_{2i}x_{2i})u(M_i-x_{1i}-x_{2i})+\frac{\partial e_2}{\partial x_{2i}}-\frac{\partial e_1}{\partial x_{1i}} \quad (2)$$

Equation (2) implicitly defines the parameters β that describe consumer preferences as a function of lotto game sales (x_1 and x_2), jackpot amounts (R_1 and R_2), and win probabilities (p_1 and p_2) under the assumptions of von Neumann-Morganstern expected utility. In the expected utility specification, the vector β of unknown parameters that describe consumer preferences is comprised of the parameters of the utility function and the entertainment function.

In order to estimate this model empirically, we need an "error" term. (Without one the equa-

[2]The *first order condition* of (1) with respect to x_j is the following:

$$\frac{\partial L}{\partial x_j} = p_j u(M + R_j - x_j - x_{\text{-}j}) - p_j x_j u_{x_j}(M + R_j - x_j - x_{\text{-}j}) - p_{\text{-}j} x_{\text{-}j} u_{x_j}(M + R_j - x_j - x_{\text{-}j})$$

$$-p_j(1 - p_j x_j - p_{\text{-}j} x_{\text{-}j})u(M - x_j - x_{\text{-}j}) - (1 - p_j x_j - p_{\text{-}j} x_{\text{-}j})u_{x_j}(M - x_j - x_{\text{-}j}) + \frac{\partial e_j}{\partial x_j} = 0$$

tion presents a mathematical certainty for all 4,339 observations.) I introduce random error into the equation by parameterizing the entertainment value e into a functional component and an idiosyncratic component. To begin, I assume the entertainment functional can be approximated with the square root of the number of tickets played, multiplied by a constant for flexiblity: $a_j\sqrt{x_j}$. This function is chosen because it is increasing, concave, and twice-differentiable. I test whether parameter estimates are robust to alternative functional forms of e. The second component of the entertainment value is a multiple of the number of tickets played that is determined by a random "draw," ζ_i, that is particular to the market observation on the consumer. The entertainment value of betting on a game in any market i is represented as follows:

$$e(x_{ji}) = a_j\sqrt{x_{ji}} + \zeta_i x_{ji},$$

where ζ_i is an i.i.d. error term across markets. It is an idiosyncratic component of utility that influences the consumer's preference for a particular lotto game in a particular market. It is independent of observed characteristics of the lotto game. If $a_j > 0$ and the absolute value of ζ_i is within the necessary bounds the four conditions are satisfied: $e(0) = 0$, $e(1) > 0$, $\frac{\partial e(x_j)}{\partial x_j} > 0$; $\frac{\partial^2 e(x_j)}{\partial x_j^2} < 0$.

Evaluating $\frac{\partial e_1}{\partial x_{1i}}$ and $\frac{\partial e_2}{\partial x_{2i}}$, equation (2) can be rewritten as follows:

$$\begin{aligned} \zeta_i^{EU} \equiv\ & \zeta_{1i} - \zeta_{2i} = p_{2i}u(M_i + R_2 i) - p_{1i}u(M_i + R_1 i) - \\ & (p_{2i} - p_{1i})(1 - p_{1i}x_{1i} - p_{2i}x_{2i})u(M_i - x_{1i} - x_{2i}) + \frac{1}{2}a_2 x_{2i}^{-\frac{1}{2}} - \frac{1}{2}a_1 x_{1i}^{-\frac{1}{2}} \end{aligned} \qquad (3)$$

Equation (3) defines ζ_i as a market-specific error term that captures the difference in entertainment value between games X_1 and X_2 in a particular week. It is the part of relative entertainment that is not determined by the game jackpots or win probabilities, but rather by unobservable, random features of the particular pairing. It is common in the Industrial Organization literature to assume product-specific errors that are independent of observed product characteristics. In this specification, I define the independent error term as being specific to a pair i of lotto games in a particular week. This seems to be a better description of reality. For example, a local newspaper might run a compelling story about a previous game X_1 winner that increases the relative entertainment value of betting on game X_1 over X_2 in that week.

5.2 Solving the model under the hypothesis of cumulative prospect theory (CPT)

Under CPT, the optimization problem facing the representative consumer in market i is the following:

$$\max_{x_{1i},x_{2i}} \; G(p_{1i}x_{1i})u^W(R_{1i}-x_{1i}-x_{2i})+G(p_{2i}x_{2i})u^W(R_{2i}-x_{1i}-x_{2i})+ \qquad (4)$$
$$H(1-p_{1i}x_{1i}-p_{2i}x_{2i})u^L(-x_{1i}-x_{2i})+e_1(x_{1i})+e_2(x_{2i})$$

where G is the weighting function of win probabilities, H is the weighting function of loss probabilities, u^W is the value function over wins, and $u^L$is the value function over losses.

This representation of the consumer choice problem deviates from the expected utility representation in three ways. First, the expected utility specification imposes the following restrictions: $G(p_jx_j)=p_jx_j$ and $H(1-p_jx_j-p_{\text{-}j}x_{\text{-}j})=1-p_jx_j-p_{\text{-}j}x_{\text{-}j}$. In contrast, CPT allows the probabilties to enter non-linearly. Second, M_i is normalized to zero since CPT models risks relative to the consumer's initial position. Third, the value function is allowed to be different for gains and losses. The value function over losses, u^L, is thought to be convex and steeper than u^W, reflecting the phenomenon that Kahneman and Tversky refer to as "loss aversion".

Unfortunately, the losses in the lottery context are too small and have too little variation to allow us to empirically distinguish between u^W and u^L. The changes in assets we observe are essentially at and above the reference point. I therefore focus on the CPT specification of decision weights and do not attempt to test loss aversion in this paper: I replace u^W and u^L with u, a continuous increasing function such that $u(0)=0$.

Approximating $R-x-y$ with R and setting $\frac{\partial L}{\partial x_1}=\frac{\partial L}{\partial x_2}$ yields the following[3]:

$$\begin{aligned} 0 \;=\; & p_{2i}\frac{\partial G(p_{2i}x_{2i})}{\partial x_{2i}}u(R_{2i})-p_{1i}\frac{\partial G(p_{1i}x_{1i})}{\partial x_{1i}}u(R_{1i})- \\ & u(-x_{1i}-x_{2i})[p_{2i}\frac{\partial H(1-p_{1i}x_{1i}-p_{2i}x_{2i})}{\partial x_{2i}}-p_{1i}\frac{\partial H(1-p_{1i}x_{1i}-p_{2i}x_{2i})}{\partial x_{1i}}]+ \end{aligned}$$

[3]The *first order condition* of (4) with respect to x_j is the following:

$$\frac{\partial L}{\partial x_j}=p_1\frac{\partial G(p_jx_j)}{\partial x_j}u(R_j-x_j-x_{\text{-}j})-G(p_jx_j)\frac{\partial u(R_j-x_j-x_{\text{-}j})}{\partial x_j}-G(p_2x_2)\frac{\partial u(R_{\text{-}j}-x_j-x_{\text{-}j})}{\partial x_j}-$$
$$p_j\frac{\partial H(1-p_1x_1-p_2x_2)}{\partial x_j}u(-x_j-x_{\text{-}j})-H(1-p_1x_1-p_2x_2)\frac{\partial u(-x_j-x_{\text{-}j})}{\partial x_j}+\frac{\partial e_j}{\partial x_j}=0$$

$$\frac{\partial e_2}{\partial x_{2i}} - \frac{\partial e_1}{\partial x_{1i}} \tag{5}$$

Equation (5) implicitly defines the vector β of unknown parameters, which describe consumer preferences, as a function of observed lotto game sales and characteristics. In the CPT specification of preferences, β is comprisėd not only of the parameters of the utility function and the entertainment function, but also of the parameters of the probability weighting functions G and H.

As above, to introduce an error term into the equation, I parameterize entertainment. I assume initially that the entertainment value of betting on a game in a particular market can be represented as follows:

$$e(x_{ji}) = a_j\sqrt{x_{ji}} + \zeta_i x_{ji}$$

Equation (5), which governs the model parameters under CPT, can now be rewritten as

$$\begin{aligned} \zeta_i^{CPT} &\equiv \zeta_{1i} - \zeta_{2i} = p_{2i}u(M_i + R_{2i}) - p_{1i}u(M_i + R_{1i}) - \\ & (p_{2i} - p_{1i})(1 - p_{1i}x_{1i} - p_{2i}x_{2i})u(M_i - x_{1i} - x_{2i}) + \frac{1}{2}a_2 x_{2i}^{-\frac{1}{2}} - \frac{1}{2}a_1 x_{1i}^{-\frac{1}{2}} \end{aligned} \tag{6}$$

As above, ζ_i as a market-specific error term that captures the relative entertainment value of X_1 over X_2 in a particular week.

5.3 Empirical estimation of the model parameters

Define Z to be a $T * L$ vector of known lottery game characteristics, where T is the sample size. Define ζ to be the $T * 1$ vector of error terms composed of the 4,339 market observations on ζ_i. To identify the parameters of the model, I exploit the independence of Z and ζ - $\mathrm{E}[\zeta|Z] = 0$ - which implies that $\mathrm{E}[\zeta' f(Z)] = 0$. This independence provides the *moment conditions*, $m(\beta)$, needed to identify the unknown parameters in β. This estimation strategy is commonly referred to as *generalized method of moments - instrumental variables* (GMM-IV).[4]

The basic idea of GMM-IV is to choose the parameters of the model that fit the moments of the model most closely. GMM chooses the parameters which minimize the quadratic

$$Q_T = \bar{m}(\beta)' W \bar{m}(\beta)$$

[4] I can not use nonlinear least squares estimation because x and y are jointly determined. GMM, on the other hand, is equiped to deal with this type of problem. Another advantage of GMM over other estimation procedures (e.g. MLE) is that the statistical assumptions required for hypothesis testing are quite weak.

where β is a K-vector of parameters, $m(\beta)$ is an L-vector of orthogonality conditions, and W is an $L*L$ positive definite weighting matrix. The moment conditions come from the orthogonality of the errors ζ and the instruments Z. In particular, I make use of the exogeneity of the rollover jackpots; the vector Z consists of functions of R_1 and R_2 The sample analog of $\bar{m}(\beta)$ is the following:

$$\hat{m}(\beta) = \frac{1}{T} Z' \zeta(x, y, p, q, R_1, R_2; \beta).$$

Identification requires that the number of moment conditions excedes the number of unknown parameters, i.e. that $L \geq K$. When $L = K$, the model is exactly identified and there is a single solution to the moment equations. Hansen (1982) demonstrates that when $L > K$, the optimal weighting matrix is the inverse of the asymptotic covariance matrix of $\hat{m}(\beta)$. The various specifications of the model in this paper are always estimated using the optimal W.

6 Results

6.1 Estimation of parameters under the expected utility hypothesis

Do the data reject the hypothesis of the expected utility representation of preferences? To answer this question, I empirically estimate the parameters of the implicit demand equation derived from the expected utility representation. I assume the utility function u is described by constant relative risk aversion (CRRA): $u(c) = \frac{1}{1-\gamma} c^{1-\gamma}$, where c is consumption and γ is the *coefficient of relative risk aversion*, defined by $\frac{-u''(c)}{u'(c)} c$.

Assuming CRRA preferences, equation (3) is the following:

$$\begin{aligned} \zeta_i^{EU-CRRA} = {} & p_{2i} \frac{1}{1-\gamma} (M_i + R_{2i})^{1-\gamma} - p_{1i} \frac{1}{1-\gamma} (M_i + R_{1i})^{1-\gamma} - \\ & (p_{2i} - p_{1i})(1 - p_{1i}x_{1i} - p_{2i}x_{2i}) \frac{1}{1-\gamma} (M_i - x_{1i} - x_{2i})^{1-\gamma} + \\ & \frac{1}{2} a_2 x_{2i}^{-\frac{1}{2}} - \frac{1}{2} a_1 x_{1i}^{-\frac{1}{2}} \end{aligned} \tag{7}$$

The GMM objective function $\hat{Q}_T = \hat{m}(\hat{\beta})' W \hat{m}(\hat{\beta})$ is constructed with the moment conditions $\hat{m}(\hat{\beta}) = \frac{1}{T} Z' \zeta^{EU-CRRA}(x, y, p, q, R_x, R_y; \hat{\beta})$. By definition, $\zeta_i^{EU-CRRA}$ is the i.i.d. component of the relative entertainment value of betting on game X_1 over X_2 in market i and is independent of the vector Z of observed lotto game characteristics.

I exploit the independence of the error term and the jackpot amount, which implies the orthogonality of the error term and any function of the jackpot amount: $\mathrm{E}[\zeta|Z] = 0 \Rightarrow \mathrm{E}[\zeta' f(Z)] = 0$. I define

$$Z = [R_1 \ R_2 \ ln(R_1) \ ln(R_2) \ R_1^2 \ R_2^2 \ (ln(R_1)^2 \ (ln(R_2))^2].$$

Using functions of the jackpots as instruments increases the number of moment conditions, which increases the efficiency of the estimation. The sample size of the data is sufficiently large that eight moment conditions can be used - and potentially many more - without sacrificing the ability to obtain valid inference via GMM about the structural parameters of the model. (Koenker and Machado (1999), Andrews (2001)).

Table 1a displays the results from GMM-IV estimation of equation (7). The equation has three unknown parameters ($K = 3$): γ, a_1, and a_2. The point estimate for γ is 0.460 with a standard error of 0.003. Neither of the entertainment parameters a_1 and a_2 is statistically different from zero. The vector β is over-identified with eight moment conditions. Greene (1997) explains that if the moment equations over-identify the parameters, then they imply substantive restrictions. He further explains, "As such, if the hypothesis of the model that led to the moment equations in the first place is incorrect, at least some of the sample moment restrictions will be systematically violated. This provides the basis for a test of the overidentifying restrictions." The GMM objective function Q_T is a Wald statistic distributed as χ^2 with $L - K$ degrees of freedom. The data fail to reject the expected utility representation of preferences: Q_T is $7.76 * 10^{-12}$.

To confirm that the results are not being driven by the particular specification of entertainment, I assume alternative functional forms for the entertainment value of lottery betting and re-estimate equation (7). First, I relax the restriction on the exponent characterizing the entertainment value and parameterize $e_j(x_{ji})$ as follows:

$$e_j(x_{ji}) = a_{ji} x_{ji}^c + \zeta_i x_{ji}.$$

The revised equation (7) has four unknown parameters - γ, c, a_1 and a_2 - and the model continues to be over-identified. The results are listed in the second row of Table 1a. The data again fail to reject the expected utility representation of preferences: Q^T is $1.71 * 10^{-12}$. However, consumers now appear to be more risk averse: the point estimate for the coefficient of relative risk aversion is 1.12, with a standard error of 0.006..

Second, I parameterize the entertainment value with a quadratic:

$$e_j(x_{ji}) = x_{ji} + a_j x_{ji}^2 + \zeta_i x_{ji.}$$

The revised equation (7) has three unknown parameters: γ,a_1,and a_2.The results are virtually unchanged, as shown in the third row of Table 1a. The data fail to reject the expected utility representation of preferences: Q^T is $1.72 * 10^{-7}$. And the point estimate of γ is slightly above one. Blanchard and Fischer (1989, p. 44) note that substantial empirical work has been devoted to estimating the elasticity of substitition, which is defined as $1/\gamma$, and that the estimates "usually lie around or below unity". I thus conclude that this analysis is generating estimates of the coefficient of relative risk aversion that are consistent with previous estimates.

The failure to reject the expected utility hypothesis is not sensitive to the specification of utility. I estimate the parameters of equation (3) under two alternative functional forms for utility. I assume CARA preferences, instead of CRRA preferences, with a *coefficient of absolute risk aversion* defined by $\delta = \frac{-u''(c)}{u'(c)}$. Blanchard and Fischer (1989, p. 44) note that CARA is "usually thought of as a less plausible description of risk aversion than constant relative risk aversion; the CARA specification is, however, sometimes analytically more convenient than the CRRA specification and thus also belongs to the standard tool kit [of economists]." For this reason, CARA preferences are assumed for the purpose of a specification check on the model rather than as the base, or preferred, specification.

Assuming CARA preferences, equation (3) is the following:

$$\begin{aligned} \zeta_i^{EU-CARA} &= -p_{2i}\frac{1}{\delta}\exp(-\delta(M_i + R_{2i}) + p_{1_i}\frac{1}{\delta}\exp(-\delta(M_i + R_{1i}) - \\ & (p_{2i} - p_{1i})(1 - p_{1i}x_{1i} - p_{2i}x_{2i})\frac{1}{\delta}\exp(-\delta(M_i - x_{1i} - x_{2i})) + \\ & \frac{1}{2}a_2 x_{2i}^{-\frac{1}{2}} - \frac{1}{2}a_1 x_{1i}^{-\frac{1}{2}} \end{aligned} \tag{8}$$

The GMM objective function $\widehat{Q}_T = \hat{m}(\widehat{\beta})'W\hat{m}(\widehat{\beta})$ is constructed with the moment conditions $\hat{m}(\widehat{\beta}) = \frac{1}{T}Z'\zeta^{EU-CARA}(x, y, p, q, R_x, R_y; \widehat{\beta})$, where $\zeta_i^{EU-CARA}$ is the i.i.d. component of the relative entertainment value of betting on game X_1 over X_2 in market i. β is comprised of three unknown parameters: δ, a_1, and a_2.I use GMM-IV techniques to estimate equation (8) as written, as well as under the two alternative specifications of entertainment defined above.

Table 1b displays the results. The data fail to reject the expected utility representation of preferences in all three specifications. In the first specification, Q_T is $1.12{*}10^{-8}$; in the second Q^T is

$4.05*10^{-24}$; and in the third, Q^T is $6.37*10^{-170}$. In other words, the objective function is essentially zero in all specifications, corroborating the failure to reject the expected utility representation of preferences under the assumption of CRRA preferences.

A secondary result from this analysis is that under the assumption of CARA utility, estimated risk preferences are sensitive to the specification of the entertainment value. This is a troubling finding regarding the robustness of the CARA functional form. Parameterizing the entertainment function as a power function (row 2 of Table 1b), lottery players appear to be risk averse. But, when a flexible quadratic form is assumed (row 3 of Table 1b), lottery players are found to be risk loving. (In the base specification, δ is estimated imprecisely.)

The second robustness check on utility specification is shown in Table 1c. I evaluate ζ_i^{EU} - defined by equation (3) - under the assumption of cubic utility. This functional form of utility corresponds to the Friedman and Savage (1948) suggestion that people both gamble and purchase insurance because they place a high value on the chance to increase their wealth greatly, i.e. that utility is described with a covex segment in the middle range of an otherwise concave function. The point estimates of the utility parameters have the signs predicted by Friedman-Savage utility; the point estimate is negative for squared consumption, or wealth, and positive for cubed. Unfortunately, the parameters are not precisely identified so the alternative hypothesis can not be rejected. Furthermore, the data continue to fail to reject the expected utility hypothesis: the objective function has a value of $8.60 * 10^{-9}$.

6.2 Estimation of parameters under the cumulative prospect theory hypothesis

This section presents the results of a second test of expected utility theory. I estimate the parameters governing consumer choices under the less restrictive assumptions of cumulative prospect theory. I then directly observe whether the data reject the restrictions imposed by expected utility theory.

In the CPT formulation of individual decision making, agents weigh probabilities in a systematic way, and furthermore, they weigh probabilities differently for wins and losses. (See the CPT representation of the consumer optimization problem in equation (4) above.) The weighting functions G (over win probabilities) and H (over loss probabilities) are strictly increasing functions from $[0,1]$ to $[0,1]$ with $G(0) = H(0) = 0$ and $G(1) = H(1) = 1$.

I assume the common power specification for both G and H:

$$G(p) = p^{\alpha}\text{and } H(p) = p^{b}$$

The power specification forces the weighting function to be either wholly concave (exponent below one) or wholly convex (exponent above one). This restriction is reasonable in the present context because the data do not allow us to trace out G and H curves on the entire [0,1] interval.The CPT literature points to a weighting function that is concave over small probabilities and convex over high probabilities. If we were to trace out G and H curves over the entire interval, the theory predicts changing curvature. Tversky and Wakker (1995) present a theoretical analysis of decision weights that is motivated by observed patterns of choice in experimental data; they claim that the observed pattern suggests an S-shaped weighting function that overweights small probabilites and underweights moderate and high probabilties. Tversky and Fox (1994) observe the choices of 40 subjects and puts the crossing between 0.3 and 0.4. In the context of state lottery jackpots, the win probabilites are always substantially below this switching point and the loss probabilites are always substantially above. In the data, we observe the lower portion of G, and expect it to be concave: $\alpha < 1$; and we observe the upper portion of H and expect it to be convex: $b > 1$.

The second main difference in the CPT specification of the consumer problem is that the value functon is defined over gains and losses. Initial wealth is taken to be the reference point, and therefore M does not enter the value function. Recall that CPT has the added generality of allowing for different value functions over wins and losses. Unfortunately, the state lottery experiment does not provide a rich source of variation in losses and it is consequently infeasible to estimate distinct value functions.

As above, I begin by assuming utility can be desribed by the common CRRA specification. I flexibly parameterize the entertainment value of lottery betting with a quadratic functional:

$$e_j(x_{ji}) = x_{ji} + a_j x_{ji}^2 + \zeta_i x_{ji.}$$

Assuming CRRA utility with a coefficient of relative risk aversion γ and the above form of entertainment, ζ_i is defined by (re)evaluating equation (6), which defines ζ_i^{CPT}, as follows:

$$\zeta_i^{CPT-CRRA} = p_{2i}^2 \alpha (p_{2i} x_{2i})^{\alpha-1} \frac{1}{1-\gamma} (R_{2i})^{1-\gamma} - p_{1i}^2 \alpha (p_{1i} x_{1i})^{\alpha-1} \frac{1}{1-\gamma} (R_{1i})^{1-\gamma} -$$

$$(p_{1i}^2 - p_{2i}^2)\beta(1 - p_{1i}x_{1i} - p_{2i}x_{2i})^{\beta-1}\frac{1}{1-\gamma}(-x_{1i} - x_{2i})^{1-\gamma} +$$
$$2a_2x_{2i} - 2a_1x_{1i} \tag{9}$$

The GMM objective function $\widehat{Q}_T = \hat{m}(\widehat{\beta})'W\hat{m}(\widehat{\beta})$ is constructed with the moment conditions $\hat{m}(\widehat{\beta}) = \frac{1}{T}Z'\zeta^{CPT-CRRA}(x, y, p, q, R_x, R_y; \widehat{\beta})$. By definition, $\zeta_i^{CPT-CRRA}$ is the i.i.d. component of the relative entertainment value of betting on game X_1 over X_2 in market i and is independent of the vector Z of observed lotto game characteristics. As above, I exploit the orthogonality of the error term and any function of the jackpot amount, $E[\zeta' f(Z)] = 0$, and define

$$Z = [R_1\ R_2\ ln(R_1)\ ln(R_2)\ R_1^2\ R_2^2\ (ln(R_1)^2\ (ln(R_2))^2].$$

Table 2a displays the results from GMM-IV estimation of equation (9). The equation has five unknown parameters $(K = 5)$: γ, α, b, a_1, and a_2. With eight moment conditions, the parameter vector β is over-identified. As discussed above, the GMM objective function Q_T is a Wald statistic distributed as χ^2 with $L - K$ degrees of freedom and is the test statistic for the over-identifying restrictions. In this specification, it is $8.16 * 10^{-6}$, and we therefore fail to reject this generalized representation of preferences. This is not surprising, since the data failed to reject the more restrictive specification in Table 1a, row 3. It is not completely implied though, because under the CPT hypothesis, initial wealth M does not enter the consumer optimization problem.

The estimation of equation (9) is uninformative about the degree of risk aversion and the nature of probability weights: γ, α, and b are all estimated imprecisely. To increase the power of the estimation, I impose the restriction $a1 = a2 = 0$; as shown in Table 2a, row 1, the data suggest this is the case. This says that the entertainment value of lottery betting is proportional to the number of tickets purchased. The estimation results are listed in Table 2a, row 2. The point estimate of the coefficient of relative risk aversion γ is 0.262 with a standard error 0.018: consumers are only slightly risk-averse. The power function describing H, the weighting function of loss probabilities, is imprecisely estimated. The power function describing G, the weighting of win probabilities, is estimated to be convex: $\widehat{\alpha} = 3.22$, with a standard error of 0.007. This is the first notable deviation from expected utility theory: the data reject the restriction that probability of winning enters the consumer optimization problem with an exponent of one. However, the rejection is in the opposite direction than predicted by the theory.[5]

[5]Estimating equation (9) with the restriction that $\alpha = b$, which represents the case of "reflection", yields a more

This finding is not sensitive to the specification of utility. I estimate the parameters of equation (6) under the assumption of CARA preferences, instead of CRRA, maintaining the quadradtic functional form for the entertainment value. The coefficient of absolute risk aversion is defined by $\delta = \frac{-u''(c)}{u'(c)}$. The GMM objective function is composed of the eight moment conditions defined by $Z\prime\zeta^{CPT-CARA}(x, y, p, q, R_x, R_y; \beta)$, where $\zeta_i^{CPT-CARA}$ is defined by (re)evaluating equation (6) as follows:

$$\begin{aligned}\zeta_i^{CPT-CARA} &= -p_{2i}^2\alpha(p_{2i}x_{2i})^{\alpha-1}\frac{1}{\delta}\exp(-\delta R_{2i}) + p_{1i}^2\alpha(p_{1i}x_{1i})^{\alpha-1}\frac{1}{\delta}\exp(-\delta R_{2i}) + \\ &\quad (p_{1i}^2 - p_{2i}^2)\beta(1 - p_{1i}x_{1i} - p_{2i}x_{2i})^{\beta-1}\frac{1}{\delta}\exp(\delta(x_{1i} + x_{2i})) + 2a_2x_2 - 2a_1x_1 \quad (10)\end{aligned}$$

Table 2b displays the results. With eight moment conditions, the parameter vector $\beta^{CPT-CARA} = [\delta\ \alpha\ b\ a_1\ a_2]'$ is over-identified. The GMM objective function Q_T is a Wald statistic distributed as χ^2 with three degrees of freedom and is the test-statistic for the over-identifying restrictions. In the first specification of Table 2b, Q_T is $1.57 * 10^{-18}$: the data fail to reject the model. The estimation of equation (10) is uninformative about the degree of risk aversion and the nature of probability weights: δ, α, and b are all estimated imprecisely.

As in Table 2a, to increase the power of the estimation, I impose the restriction $a1 = a2 = 0$, which says that the entertainment value of lottery betting is proportional to the number of tickets purchased, which is suggested by the data.The data fail to reject the model: Q^T is $5.49 * 10^{-16}$. The power function describing H, the weighting function of loss probabilities, is imprecisely estimated. The power function describing G, the weighting of win probabilities, is estimated to be slightly convex: $\widehat{\alpha} = 1.91$, with a standard error of .063. This corroborates the deviation from expected utility theory obtained under the assumption of CRRA preferences: the data reject the restriction that probability of winning enters the consumer optimization problem with an exponent of one. Again, the rejection is in the opposite direction than predicted by the theory.

This finding is strikingly similar to the finding of Jullien's and Salanie's (2000) analysis of data on British horseraces. Assuming CARA preferences and a power weighting function, they estimate a slightly convex (albeit insignificant) weighting function for gains: $\widehat{\alpha} = 1.16$ with a standard error of 0.143 (Jullien and Salanie 2000, p. 517). They estimate a concave weighting function for losses: $b = 0.318$, with a standard error of 0.272. In addition, their data do not support the hypothesis of changing concavity of the weighting functions.

positive point estimate of γ (2.26, standard error of 0.002) but an imprecise estimate of the power function exponent.

7 Discussion

This chapter has used the real-world "experiment" of state lotteries to empirically test the usefulness of expected utility theory for describing consumer choices. I have tested expected utility theory two ways. First, I specified the consumer optimization problem under the hypothesis of expected utility theory. I then derived demand for lottery gambles as an implicit function of observed lotto game characteristics and unknown consumer choice parameters. GMM-IV methodology was used to estimate the unknown parameters of the model. The parameters were over-identified by the GMM-IV moment conditions, which provided an over-identification test of the model. I concluded from the over-identification test that the data fail to reject the hypothesis of expected utility. This failure to reject was robust to alternative specifications of the functional form of utility and entertainment value.

Second, I derived demand for lottery gambles as an implicit function of observed lotto game characteristics and unknown consumer choice parameters under a more general representation of preferences, as suggested by cumulative prospect theory. In particular, probabilites were allowed to enter the consumer problem non-linearly and value was defined over changes in wealth, rather than final wealth level. Again, GMM-IV methodology was utilized to estimate the unknown and over-identified parameters of the model. I concluded that the data fail to reject the specification of the model; this was true under the assumption of CRRA preferences and under the assumption of CARA preferences. Furthermore, under both specifications of utility, the data suggested that consumers weight the (very small) probability over winning with a convex weighting function. This is a notable deviation from expected utility theory, and it corroborates the finding of Jullien and Salanie (2000). However, the deviation is the opposite of that found in the experimental literature, which has suggested that small probabilities are overweighted.

The finding of a convex weighting of "win" probabilities, even when agents are observed to be risk-averse, is notable. The overweighting of win probabilities has been previously offered as an explanation as to why risk-averse agents would participate in lottery gambling. The findings of this paper imply that lottery players participate in lotto gambling not because they overweight the chance of winning, nor because they are risk-loving. Rather, these findings suggest that consumers play the lotto because it is entertaining. This is consistent with the demand analysis of chapter one.

This chapter has relied on a representative agent assumption to draw conclusions about consumer preferences from market data. This is a weakness of the analysis, but can not be avoided given available data. A second weakness of the analysis, and one that can potentially be remedied, is that many of the parameter estimates are imprecise. A way to increase the power of the GMM-IV estimation would be to to incorporate lower prize tiers into the model by defining utility not only over the jackpot of the lotto game, but also over the second and third place prizes. This would add complexity to the optimization problem, but would have the benefit of providing additional GMM-instruments to be used in the GMM-IV estimation. A final weakness is that by defining demand implicitly as a function of lotto game characteristics, it is not transparent to compare the results of this analysis with the reduced form analysis of Chapter 1. Due to the complexity of the problem, I was unable to derive closed-form solutions for $x_1$and x_2. Additional work in this direction might prove fruitful.

References

[1] Allais. 1953. "Le comportement de l'homme rationnel devant le risque, critique des postulats et axiomes de l'ecole Americaine." *Econometrica* 21, 503-546.

[2] Ali, M.M. 1977. "Probability and utility estimates for racetrack bettors." *Journal of Political Economy* 85, 803-815.

[3] Andrews, Donald W.K. and Biao Lu. 2001. "Consistent model and moment selection procedures for GMM estimation with application to dynamic panel data models." *Journal of Econometrics* 101, 123-164.

[4] Blanchard, Olivier and Stanley Fischer. 1989. *Lectures on Macroeconomics.* Cambridge, MA: The MIT Press.

[5] Cicchetti, Charles J. and Jeffrey A. Dubin. "A Microeconometric Analysis of Risk Aversion and the Decision to Self-Insure," *Journal of Political Economy* 102(1), 169-186.

[6] Cliff, Michael T. 2000. "GMM and MINZ Program Libraries for MATLAB." Mimeo, Krannert Graduate School of Management, Purdue University.

[7] Farber, Hank. 1978. "Individual Preferences and Union Wage Determination: The Case of the United Mine Workers," *Journal of Political Economy* 86 (5), 923-942.

[8] Fishburn, Peter. 1988. *Nonlinear Preference and Utility Theory.* Baltimore: Johns Hopkins University Press, 1988.

[9] Friedman, Milton and L.J. Savage. 1948. "The Utility Analysis of Choices Involving Risk," *Journal of Political Economy* 56(4), 279-304.

[10] Friend and Blume. 1975. "The Demand for Risky Assets,"*American Economic Review* 65(5), 900-922.

[11] Greene, William H. 1997. *Econometric Analysis* 3rd edition. New Jersey: Prentice Hall.

[12] Harless, D.W. and C.F. Camerer. 1994. "The Predictive Utility of Generalized Expected Utility Theories," *Econometrica* 62(6), 1251-1289.

[13] Hansen, L. 1982. "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica* 50, 1029-1054.

[14] Jullien, B. and B. Salanie. 2000. "Estimating Preferences under Risk: The Case of Racetrack Bettors,"*Journal of Political Economy 108, 503-530.*

[15] Kahneman, Daniel, and Amos Tversky. 1979. "Prospect theory: An Analysis of Decision under Risk," *Econometrica* 47, 263-291.

[16] Koenker, Roger and Jose A.F. Machado. 1999. "GMM Inference When the Number of Moment Conditions is Large," *Journal of Econometrics* 93, 327-344.

[17] Kreps, David M. 1990.*A Course in Microeconomic Theory.* Princeton University Press.

[18] MacCrimmon, Kenneth and Stig Larsson. 1979. "Utility Theory: Axioms versus 'Paradoxes'," in Maurice Allais and Ole Hagan, eds. *Expected Utility Theory and the Allais Paradox.* Dordrecht, Holland: D. Reidel Publishing Co., 333-409.

[19] Machina, Mark. 1983a. "Generalized Expected Utility Analysis and the Nature of Observed Violations of the Independence Axiom," in Bernt Stigum and Fred Wenstop, eds., *Foundations of utility and risk theory with applications.* Dordrecht, Holland: D. Reidel Publishing Co., 263-95.

[20] ——. 1983b. "The Economic Theory of Individual Behavior Toward Risk: Theory, Evidence and New Directions," Manuscript, Institute for Mathematical Studies in the Socials Sciences Technical Report, Stanford University.

[21] ——. 1987. "Choice under Uncertainty: Problems Solved and Unsolved," *Journal of Economic Perspectives* 1(1), 121-54.

[22] ——. 1989. "Dynamic Consistency and Non-Expected Utility Models of Choice under Uncertainty," *Journal of Economic Literature* 27(4), 1622-1668.

[23] Newey, Whitney and Daniel McFadden. 1994. "Large Sample Estimation and Hypothesis Testing," in ed. R.F. Engle and D.L. McFadden *Handbook of Econometrics IV.*, Chapter 36, 2111-2176.

[24] Rabin, Matthew. 1996. "Psychology and Economics," Mimeo, University of California at Berkeley, September.

[25] ——. 1999. "Risk Aversion and Expected-Utility Theory: A Calibration Theorem." Mimeo, University of California at Berkeley, May.

[26] Sugden, Robert. 1985. "Regret, Recrimination, and Rationality," *Theory and Decision* 19(1), 77-99.

[27] Szpiro, George G. 1986. "Measuring Risk Aversion: An Alternative Approach," *Review of Economics and Statistics* 68(1), 156-159.

[28] Thaler, R.H. and W.T. Ziemba. 1988. "Anomalies. Parimutuel betting markets: Racetracks and lotteries," *Journal of Economic Perspectives* 2(2), 161-174.

[29] Tversky, Amos and C. Fox. 1994. "Weighting Risk and Uncertainty," *Psychological Review.*

[30] Tversky, Amos and Daniel Kahneman. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty* 5, 297-323.

[31] Tversky, Amos and Richard H. Thaler. 1990. "Anomalies: Preference Reversals," *Journal of Economic Perspectives* 2(2), 201-211.

[32] Tversky, Amos and Peter Wakker. 1995. "Risk Attitudes and Decision Weights," *Econometrica* 63 (6), 1255-1280.

[33] Quandt, Richard. 1986. "Betting and Equilibrium," *Quarterly Journal of Economics* 101(1), 201-208.

[34] Varian, Hal R. 1988. "Estimating Risk Aversion from Arrow-Debreu Portfolio Choice," *Econometrica* 56(4), 973-979.

[35] Weber, Martin and Colin Camerer. 1987. "Recent Development in Modeling Preferences uncer Risk," *OR Spektrum* 9, 129-51.

Table 1a
Optimization under Expected Utility Theory

CRRA: $(1/1-\gamma)*C^{(1-\gamma)}$

γ is *coefficient of relative risk aversion*

1. $e_{ji}=a_j x_{ji}^{.5}+\zeta_{ji}x_{ji}$

γ	-	a_1	a_2
.460	-	-.0005	-.0007
(.003)		(31.6)	(40.2)

Q: $7.76*10^{-12}$

2. $e_{ji}=a_j x_{ji}^{.c}+\zeta_{ji}x_{ji}$

γ	c	a_2	a_1
1.12	1.00	1.58	1.58
(.006)	(150.0)	(355.9)	(378.68)

Q: $1.71*10^{-12}$

3. $e_{ji}=xij + a_j x_{ji}^{.2}+\zeta_{ji}x_{ji}$

γ	-	a_1	a_2
1.28	-	-.000005	-.000002
(2.73)		(.022)	(.015)

Q: $1.72*10^{-7}$

note: The models are estimated in MATLAB Version 6.1.0. The GMM-IV objective function is specified as an unconstrained multivariable function and minimized using MATLAB's simplex search method.

Table 1b
Optimization under Expected Utility Theory

CARA: $(-1/\delta)*e^{(-\delta*C)}$

δ is *coefficient of absolute risk aversion*

1. $e_{ji}=a_j x_{ji}^{.5}+\zeta_{ji}x_{ji}$

δ	-	a_1	a_2
$-7.84*10^{-10}$	-	.047	.063
$(1.43*10^{-9})$		(419.4)	(331.5)

Q: $7.76*10^{-12}$

2. $e_{ji}=a_j x_{ji}^{c}+\zeta_{ji}x_{ji}$

δ	c	a_1	a_2
$1.19*10^{-7}$	1.06	9.77	9.78
$(1.71*10^{-9})$	$(1.16*10^{10})$	$(1.31*10^{10})$	$(2.53*10^{9})$

Q: $1.71*10^{-12}$

3. $e_{ji}=xij + a_j x_{ji}^{2}+\zeta_{ji}x_{ji}$

δ	-	a_1	a_2
$-2.83*10^{-8}$	-	-.905	-1.24
$(2.60*10^{-9})$		$(2.14*10^{84})$	$(3.51*10^{84})$

Q: $1.72*10^{-7}$

note: The models are estimated in MATLAB Version 6.1.0. The GMM-IV objective function is specified as an unconstrained multivariable function and minimized using MATLAB's simplex search method.

Table 1c
Optimization under Expected Utility Theory

Cubic Utility: $C+b_1*C^2+b_2*C^3$

$e_{ji}= x_j + a_j x_{ji}^{2}+\zeta_{ji}x_{ji}$

b_1	b_2	a_1	a_2
-5.76	.00002	-1.36	-1.07
$(2.63*10^{8})$	(2.30)	$(1.01*10^{15})$	$(8.19*10^{14})$

Q: $8.60*10^{-9}$

note: The models are estimated in MATLAB Version 6.1.0. The GMM-IV objective function is specified as an unconstrained multivariable function and minimized using MATLAB's simplex search method.

Table 2a
Optimization under Cumulative Prospect Theory
Probability weight functions: $G=p^{\alpha}$, $H=p^{b}$

CRRA: $(1/1-\gamma)*C^{(1-\gamma)}$

γ is *coefficient of relative risk aversion*

1. $e_{ji}=xij + a_j x_{ji}^{2} + \zeta_{ji} x_{ji}$

γ	α	b	a_1	a_2
-.215	3.24	.00003	$-9.31*10^{-18}$	$-4.75*10^{-18}$
$(1.93*10^{10})$	$(7.86*10^{12})$	$(3.79*10^{10})$	(.026)	(.025)

Q: $8.16*10^{-8}$

2. $e_{ji}=xij + a_j x_{ji}^{2} + \zeta_{ji} x_{ji}$

$a_1 = a2=0$

γ	α	b	a_1	a_2
.262	3.22	.0000002	-	-
(.018)	(.007)	(40.1)		

Q: $8.34*10^{-17}$

note: The models are estimated in MATLAB Version 6.1.0. The GMM-IV objective function is specified as an unconstrained multivariable function and minimized using MATLAB's simplex search method.

Table 2b
Optimization under Cumulative Prospect Theory
Probability weight functions: $G=p^{\alpha}$, $H= p^{b}$

CARA: $(-1/\delta)*e^{(-\delta*C)}$

δ is *coefficient of absolute risk aversion*

1. $e_{ji}=xij + a_j x_{ji}^{2} + \zeta_{ji} x_{ji}$

δ	α	b	a_1	a_2
-.000007	27.1	.000001	$-4.31*10^{-11}$	$-6.97*10^{-12}$
(.0004)	$(1.92*10^{8})$	(40.7)	(.038)	(.026)

Q: $1.57*10^{-18}$

2. $e_{ji}=xij + a_j x_{ji}^{2} + \zeta_{ji} x_{ji}$

$a_1 = a2=0$

δ	α	b	a_1	a_2
-.0000001	1.91	.00003	-	-
$(1.38*10^{-11})$	(.063)	(2181.0)		

Q: $5.49*10^{-16}$

note: The models are estimated in MATLAB Version 6.1.0. The GMM-IV objective function is specified as an unconstrained multivariable function and minimized using MATLAB's simplex search method.

Chapter 3

Welfare Reform and Fertility Behavior: A Look at the Family Cap

1 Introduction

Over the past decade, states across the country have been experimenting with welfare reform. One of the most controversial reform policies, the "family cap" or "child exclusion," is motivated by the notion that an incremental increase in cash assistance for each additional child increases a woman's propensity to bear additional children. With the intention of heightening personal responsibility, 18 states have responded to this concern by instituting legislation ending the traditional practice of providing families on welfare with additional cash benefits when a new child is born into the family. An additional five states have implemented quasi-family cap policies that alter the form of the additional benefit, but fall short of eliminating it entirely. This chapter uses the variation across states in the timing of family cap implementation to identify whether the denial of incremental benefits leads to a reduction in births.

A woman's decision to give birth is part of a complex series of decisions influenced by social, religious, economic, and other demographic and personal factors. The question of how welfare benefits affect this decision focuses only on the economic factors determining this choice. The primary economic question is whether the availability of more resources at the margin increases a woman's propensity to bear additional children. The policy of eliminating the marginal increase in resources raises the price of an additional child and might thus deter women from having additional births. The direct effect of the policy is to potentially reduce higher order births. Additionally, the policy sends a message that welfare is not as reliable and generous as it used to be. It might thus deter poor women from becoming mothers until they have enough resources and skills to avoid dependence on welfare. This would have the additional effect of reducing first births.

The economic theory underlying this question is that of rational choice, and in particular the role of incentives as important determinants of behavior. There is an extensive literature on various potential incentive effects of the welfare system. Econometric studies generally show that labor supply is reduced by the Aid to Families with Dependent Children (AFDC) program and that higher potential benefits induce greater participation in this program. The evidence regarding the effects of AFDC on family structure is more mixed, but recent studies have found a weak effect.[25] In sum, there are both theoretical and empirical reasons to believe that women recognize the incentives and disincentives of the welfare program, and respond to them.

Identifying the causal effects of welfare on fertility decisions is not entirely straightforward. A regression of the number of births a woman has on the welfare benefits she receives confounds the direction of causality. The amount of cash assistance a woman receives is determined by the number of children she has. Many studies have tried to identify the causal relationship by exploiting cross-sectional variation in state benefit levels and birth rates. The main weakness of this latter strategy is that there are fixed differences in birth rates across states that can not be controlled for in a cross-sectional analysis.

This chapter addresses these problems by using a plausibly exogenous source of variation in incremental benefits and data from a panel of states. The nineties was a decade of unprecedented welfare reform during which all states were encouraged to experiment with new policies. The implementation of family caps does not appear to be driven by movements in birth rates. Rather, welfare reform has been a political movement during the time period being studied, and state policies have been adopted based on the politics and priorities of the state. For this reason, the legislative "quasi-experiment" is reasonably considered exogenous.

Furthermore, the variation in timing across the 23 states that implement family caps in this period provides us with multiple "quasi-experiments" from which we can identify an effect. The use of state level panel data allows me to identify the effect of the family cap by comparing the change in birth rates for a state that implements a family cap in a given period, relative to states that do not implement a family cap in the given period. Level differences between states are controlled for, as well as level differences between years that are common to all states.

I find no evidence that family cap policies are leading to a reduction in births to women ages 15 to 34. When state effects, month effects, and state specific linear time trends are controlled for, a negative effect of more than 0.5 percent can be rejected at the 95 percent confidence level. (The upper bound of the confidence interval is an increase in births of 0.8 percent.) Numerous specification checks, including lead and lag effects, do not alter this conclusion. In an attempt to increase the power of the analysis, the regression equation explaining births is estimated for subgroups with relatively high welfare propensities and focuses on additional births born to women in these groups. Again, I find no evidence that the family cap is deterring unmarried high school dropouts or teenage women from having additional births. This is true for both blacks and whites.

2 Background

2.1 Family cap policies

In August of 1996 the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) replaced the AFDC program with a block grant program called Temporary Assistance for Needy Families (TANF) and gave states flexibility to create new cash assistance programs for families with dependent children, effective July 1997. Many states actually received waivers from the federal government that allowed them to experiment with the rules of welfare even before the passage of PRWORA. Starting with New Jersey in 1992, nineteen states

[25] For a survey of the literature, see Robert Moffitt, "Incentive Effects of the U.S. Welfare System: Review," Journal of Economic Literature, Vol. XXX (March 1992), pp. 1-61.

received approval to implement family cap policies under waivers. An additional four have implemented family cap policies as part of their state TANF programs.

Under AFDC, the addition of a child to a family on welfare was automatically accompanied by additional cash benefits. In contrast, under family cap policies additional benefits are either eliminated or reduced for a child who is conceived while the mother is receiving assistance. These policies are implemented in four ways: provide no additional benefits to a family receiving welfare for a child born ten months after the family begins receiving assistance; provide only a partial increase in benefits for an additional child; provide no additional cash benefits but do give in-kind benefits of equal monetary value; and give the incremental increase in cash benefits to a third party (e.g., church) to act on behalf of the child. (Appendix Table 1 lists approval and implementation dates and state policy types.)

As an example of how the policy works, consider the effect of New Jersey's family cap on a woman on welfare with one eligible child.[26] Under AFDC, this family would receive $322 per month in cash assistance. If the woman gave birth to another child, the family would receive an additional $102 per month, and an additional $64 per month for any additional births. Food stamp benefits would also increase, but by less than the maximum due to the incremental income from AFDC benefits. Under the family cap, the family would continue to receive $322 with the birth of any additional child. The food stamp benefit would increase by more, however, though not enough to offset the decline in cash assistance.

2.2 Other recent welfare reforms

In the time period being studied, AFDC waivers were granted liberally to states as the federal government became increasingly interested in new ways to run the current welfare program. Since 1962, the Social Security Act has authorized the Secretary of Department of Health and Human Services to waive specified requirements of the act in order to enable a state to carry out any experimental, pilot, or demonstration project that the Secretary deems in accord with

[26] Example taken from M. Camasso, C. Harvey, R. Jagannathan, and M. Killingsworth, "New Jersey's Family Cap and Family Size Decisions: Some Findings from a 5-year Evaluation," presented at the NBER 1999 summer conference in Cambridge, MA.

the objectives of AFDC. The Reagan, Bush, and Clinton administrations all espoused a liberal policy of granting waivers. By mid-August 1996, the Clinton administration had approved more than 70 waivers for more than 40 states, many of which were for statewide reforms. Many of these states have incorporated provisions of their AFDC waiver projects into their initial TANF plans.[27]

Many AFDC waiver projects were aimed at encouraging labor force participation and human capital development. Restrictive reforms in this vain include time limits on benefit durations, tightened work requirements, and benefits linked to school attendance or performance. Liberalizing reforms include a more generous treatment of earnings and resources and an increased vehicle asset limit. Some states received waivers designed to encourage recipients to leave welfare by expanding transitional medical and childcare benefits. Some waivers expanded eligibility for two-parent (unemployed) families, mitigating the discriminatory affect of AFDC against dual parent families.

(Appendix Table 3 lists approval and implementation dates of states' first major waivers.)

Tightened work requirements - though not explicitly designed to affect the reproductive behavior of welfare mothers, as are family caps - might affect fertility decisions. The relative cost of having a child is higher when the child does not qualify the mother for an exemption from work requirements. Less generous child exemptions might therefore lead to decreased fertility. Under AFDC, non-exempt adult recipients receiving cash assistance were required to participate in the Job Opportunities and Basic Skills Training (JOBS) program, but a recipient was exempt from this requirement if she met certain age or caretaker criteria. The exemption criterion responsible for the majority of exemptions was for primary caretaker relatives of children up to six years of age, or under three if the state guaranteed childcare (Gallagher 1998). Thirteen states altered their JOBS requirements under waivers, typically lowering the age of the youngest child that qualifies a recipient for a work exemption and, in some cases, removing the caretaker exemption altogether. Additionally, PRWORA eliminated the federal JOBS requirements and exemption rules, and imposed the requirement that all adult recipients participate in work activities within two years of program enrollment. The federal legislation

[27] Information in this section is taken from the 1998 *Green Book*, Section 7, From the U.S. Government Printing

requires states to meet specific work participation rates and imposes restrictions on who can be exempt from these calculations and for how long. States responded by imposing stricter child age requirements under TANF than were in place under AFDC. The empirical estimation in this chapter controls for changes in work exemption rules so that any response to tightened work requirements is not interpreted as an effect of the family cap. (Appendix Table 2 lists work exemption type and date of implementation.)

Benefit duration *time limits* might have an effect on reproductive decisions insofar as they deliver the message that welfare assistance is only temporary and women perceive children to be a hindrance to achieving financial self-sufficiency.[28]. Under AFDC there were no restrictions on the length of time a family could receive welfare assistance. Twenty-four states received waivers to implement welfare time limits and PRWORA made time limits mandatory. PRWORA also mandated that state TANF plans subject teen parents to stay-in-school and live-at-home provisions. In order to isolate the effect of a family cap policy from any effect these restrictive provisions might have had on fertility outcomes, I will test whether the inclusion of controls for the implementation of time limits and TANF changes the estimated effect of a family cap policy. (Appendix Table 3 lists TANF and time limit implementation dates.)

3 Previous literature

3.1 Studies on the incentive effects of the welfare system

The 1980s saw the rise of a literature studying the determination of welfare participation at a point in time. This literature relies on the economic reasoning that welfare participation occurs if the utility from being on the program is greater than from being off the program. The probability of participation is thus theoretically positively related to the guarantee level, negatively related to the benefit reduction rate, and negatively related to the relevant market

office Online via GPO Access. http://frwebgate.access.gpo.gov/.../105_green_book.

[28] A time limit refers to a length of time for which a family can receive cash assistance and after which a family's benefit is either terminated or reduced or the family is required to participate in work requirements

wage. Empirical tests lend confirmation to these basic predictions. Almost all of the 16 studies surveyed by Moffitt (1992) report positive and significant effects of the guarantee level and negative and significant effects of the benefit reduction rate. More recently, welfare caseload studies evaluate how the welfare reform of the mid-nineties has affected the number of people receiving welfare.

Schoeni and Blank (2000) find strong evidence that recent policy changes have reduced public assistance participation. Shoeni and Blank are particularly interested in the impact of AFDC waivers implemented between 1992 and 1996 and the impact of the 1996 PRWORA. Their study analyzes the outcomes of adult women from the March Current Population Surveys (CPS) from 1977 through 1999 and concludes that among all women with less than a high school education, waivers decreased welfare participation by 0.9 percentage points, or about 10 percent. The introduction of TANF is estimated to have had twice as large of an effect on welfare participation, leading to a decline of 1.90 percentage points. The 1999 study of the effects of the 1996 reforms on caseloads conducted by the Council of Economic Advisors concludes that one-third of 33 percent decline between 1996 and 1998 was due to welfare reform. The economic literature on the subject seems to be in general agreement that program participation responds to program generosity.[29]

Unlike the literature on participation, the literature on the effects of the welfare system on family structure offers mixed results. This literature has mainly been concerned with whether the welfare system encourages single motherhood. Because benefits are typically paid to female heads of family with children but no spouse present, the program provides an incentive to have children outside a marital union. Research in the area can be traced party to Gary Becker's models of economics of the family (Moffitt, 1992; Becker, 1981).

The literature on this question from the 1970s offers mixed evidence and shows no consistent pattern. (See Moffitt 1992 for a brief review.) The literature from the 1980s uses more recent data, but still offers mixed evidence. Ellwood and Bane (1985) make use of cross-state and within-state variation in welfare benefit levels and a sample of women from the 1976 Survey of Income and Education. They find no significant effect of welfare benefit amount on the probability that an unmarried woman age 18-44 had a child last year. Duncan and Hoffman (1990) analyze data on 900 black teenage girls from the PSID (1968-1985) and find statistically insignificant positive effects of AFDC benefit levels on the probability of an out-of-wedlock birth. Lundberg and Plotnick (1990) use a three-stage nested logit model to estimate the determinants of a hierarchical series of outcomes for teenagers: pregnancy, abortion, and marriage. Their study uses data on 1,181 white females in the NLSY aged 14-16, for the years 1979-1986. Their findings suggest strong policy effects on pregnancy probabilities and on pregnancy resolution decisions. Lundberg and Plotnick note that their findings contrast with those of Ellwood and Bane (1985) and suggest that the difference in time period might be the reason; they also note that their findings are contrary to Duncan and Hoffman (1990), and suggest that the difference in sample race might be why.

Some recent work looks at the impact of welfare at the margin. Argys and Rees (1996) use an NLSY sample of 1,344 unmarried women, all of whom received welfare payments at some point between 1979 and 1991 to test the relationship between welfare generosity and additional births. Based on cross-sectional variation in welfare benefits, they conclude that a $100 increase in marginal benefit is associated with a 3.9 percentage point increase in the probability of conception leading to birth. However, once they include state fixed effects in their model, neither the welfare guarantee variable nor the marginal benefit variable is significant.

Robins and Fronstin (1996) examine CPS data from 1980 to 1998 to estimate the effect of basic AFDC benefit level and the effect of incremental benefit on family size decisions among never-married women. They find that both the benefit level and the incremental benefit for a second child positively affect family size decisions of African-American and Hispanic women, but not of white woman. Fairlie and London (1997) estimate the probability of a higher-order birth among a sample of AFDC recipients and comparison non-AFDC women using SIPP data. They find an insignificant positive relationship bet ween the probability of another birth and the incremental impact on benefits for AFDC women, but a larger positive correlation among non-AFDC women. They conclude that the relationship for AFDC women is spurious.

[29] Note that some of the response to recent reforms is "mechanical", as opposed to "behaviorial" as it is possible that some recipients reached newly imposed time limits within the time period being studied.

In sum, the question of whether welfare affects family formation and fertility decisions remains an open question.

3.2 Studies on family cap policies

Two states, Arkansas and New Jersey, have released evaluations of their family caps based on experimental designs. In both states, women receiving welfare were randomly assigned to a treatment group that was subject to a cap, and a control group that was not. Turturro et. al. (1997) report no statistically significant difference between births to control and experimental groups in Arkansas for program years 1994-1997. Camasso et. al. (1999) use two analytical approaches to evaluate New Jersey's experience: an experimental design as mentioned and a quasi-experimental pre-post analysis of the entire welfare caseload over a 6-year period than includes the implementation of the family cap. Both analyses suggest that the family cap did influence the propensity of women on AFDC to bear children. The study reports that both pregnancies and births among women on welfare declined after program implementation, and that the number of abortions increased.

Unfortunately, there seems to have been severe implementation problems with the experimental design in both states, making the results difficult to interpret. Many members of both the treatment and control groups report not knowing which policy applied to them. (Loury 2000) For example, in the Arkansas project, 46 percent of women in the treatment group and 52 percent in the control group indicated to evaluators that they did not know how much more money that would receive if they had an additional child. Additionally, in New Jersey, more than one-quarter of case workers admitted to evaluators that they used discretion when making treatment or control assignments, thereby negating the randomness of the assignment.

Another potentially large problem with the methodology of these studies is the restriction of the analyses to births to women on welfare. The sample restriction introduces sample selection problems that potentially bias the findings. As discussed above, the literature on the incentive effects of welfare has confirmed that generosity affects participation. Furthermore, the participation response seems to be nearly immediate. Both Blank and Shoeni

(2000) and the 1999 CEA report find measurable effects of PRWORA on participation within the first two to three years after its passage. It is thus imprudent to assume that the welfare population is not changing in response to the family cap policy.

To illustrate the point, consider women who desire more than one child. The existence of a family cap might encourage such women to find alternative means of financial support or discourage them from incurring the stigma cost of enrolling for welfare assistance. Under such a scenario, the presence of a family cap would change the composition of the welfare population; there would be fewer women on welfare who desire multiple children. Birth rates might remain completely unchanged, but looking only at welfare births would suggest a decline. In this sense, the strategy is likely biased toward finding a negative effect on birth rates.

The direction of the bias is not unambiguous, however. Suppose the existence of a family cap sends the message that welfare is not a reliable or generous source of resources. Then some poor women who might have had a baby in the pre-family cap world will now be more committed to staying in school and/or delay becoming a mother until they are more financially secure. They will not have a birth and will not enroll in the welfare program. Birth rates would actually be lower as a consequence of the family cap, but looking only at women on welfare would suggest no change. In this case, the strategy is biased toward finding no effect on birth rates.

Due to this concern, I do not condition on welfare receipt, and thus am able to identify the pure effect of the family cap on birth rates. The strategy of the present paper does not confound the participation decision with the birth decision. In addition, both the New Jersey and Arkansas studies analyze the effect of a single family cap policy. Such targeted evaluations are always open to questions about how generalizable the results are to other contexts. It is therefore useful to synthesize the various family cap "experiments" into one coherent analysis.

In a contemporaneous working paper, Horvath-Rose and Peters (2000) look across states and analyze the effect the family cap has had on the non-marital birth ratio[30]. Their study concludes that the family cap waiver has decreased non-marital fertility for all race and age groups. Their analysis uses aggregate state level data based on Vital Statistics birth records

[30] The non-marital birth ratio is defined as the number of non-marital births divided by the total number of births.

from 1984 to 1996. The dependent variables in their study are log of odds ratio transformations of race and age group specific ratios of non-marital births for each state. The main regressions of the study include state and year fixed effects, high school completion by 18 to 24 year olds, proportion of urban population, proportion of fundamentalist adherents, and the following indicator variables: minor parent provision waiver, time limit waiver, work requirement waiver, AFDC-UP waiver, child support waiver, expanded income disregard and asset limit waiver, school attendance and performance requirement waiver, parental consent requirement for an abortion, requirement for sex education in schools.

The specification used in Horvath-Rose and Peters (2000) is not ideal for identifying the effect of the family cap on birth rates. By focusing the analysis on the ratio of non-marital to total births, their analysis confounds marriage responses and fertility responses to the family cap. Furthermore, though marital status is key to their study, the authors do not account for the changes in the reporting of marital status in vital statistics data that occurred during the time period covered by their analysis.[31] There 4also are potential problems with their waiver codings; they appear to differ substantially from those used in the CEA report and listed in the Urban Institute report.

Horvath-Rose and Peters (2000) find large and immediate reductions in non-marital births in response to the family cap. For women ages 20 to 49, unlagged regressions suggest decreases of 1.4 percentage points for white women, off a base rate of .21, and 3.1 percentage points decline for black women, off a base rate of .51. Lagged regressions suggest decreases of 2.4 and 4 percentage points, respectively. For births born to teens, unlagged regressions suggest that the ratio has decreased more than 3 percentage points for black teens (mean of .87) to almost 5 percentage points for all teens (mean of .68). For regressions incorporating 9-month lags, the magnitudes are 2 percentage points for black teens and 6 percentage points for white teens. The

[31] As reported in the technical appendix of the 1997 vital statistics report, birth certificates in 46 states and Washington D.C. include a direct marital status question. Nevada collects marital status information from the electronic birth registration process, though it is not included on the birth certificate. This procedure was started in Nevada in 1997, after 1995-1996 procedures overestimated the number of births to unmarried women. The remaining three states of Connecticut, Michigan, and New York make marital status determinations based on whether a paternity acknowledgement was received, the father's name is missing, and lastly, whether the father's and mother's surnames are different. A direct question was not added in Texas and California until 1994 and 1997, respectively.

results of the unlagged regressions suggest an immediate decrease in births; such an effect would require an immediate marriage response, third trimester abortions, or large anticipation effects. Additionally, it is curious that such a large effect is observed for the entire population of women aged 20 to 49, most of whom will never be at risk of welfare receipt.

4 Data and Empirical Strategy

4.1 Data

Data on births are from the *Vital Statistics Natality Data, Public-Use Data Files,* years 1989 to 1998, compiled by the U.S. National Center for Health Statistics (NCHS). The Natality public-use data files include all births occurring within the United States. I limit my sample to births occurring to women ages 15 to 34 in one of the 50 states or District of Columbia. (The age restriction is applied because younger women are more likely to be at risk of welfare dependence and affected by the family cap policy.) The vital statistics data files identify the state of residence and month of birth, as well as mother's education, mother's race, mother's marital status, and live birth order. I use this information to create a data file of state birth counts.

Information on welfare policies is obtained from three sources. The first source is a 1999 technical report of the Council of Economic Advisors (CEA), which relied on experts from the Department of Health and Human Services as well as non-governmental research institutions. The second source is a 1998 Urban Institute report on state TANF programs. And thirdly, I draw on information contained in Crouse (1999), prepared for the U.S. Department of Health and Human Services. Crouse (1999) nicely summarizes information contained in a 1997 report of the U.S. Department of Health and Human Services *Setting the Baseline: A Report on State Welfare Waivers.*

4.2 Empirical strategy

The effect of the family cap on fertility can occur through two channels: conception and abortion. The theoretical predictions are that the existence of a cap on benefits for additional children will raise the price of a child and thereby cause a woman to avoid pregnancy or, once becoming pregnant, avoid a birth by having an abortion. This reasoning finds a foundation in Becker's models of the family. Becker (1981) uses the price of children and the real income to explain, among other things, why a rise in the wage rate of working women reduces fertility and why various government programs - such as AFDC - might significantly affect the demand for children. Assuming that women, on average, respond to a price increase (or benefit decrease) of an additional child, we expect that when we look at aggregate birth counts, the number of higher-order births in a state that has effectively raised the price of additional children will fall, all else being equal.

The analysis of this chapter identifies the sum of the conception and abortion responses and reports the net effect of the family cap on reproductive behavior. It is estimated at the aggregate level, looking at the number of births in a state in a month. The identification strategy of this chapter is to compare the change in the log number of births occurring in a state that becomes a family cap state to the change in the log number of births that occurs in states that do not make the family cap transition in the same period. Relative to states that have not yet passed a family cap, or that did so in the past, this analysis identifies the incremental change in births that is associated with the introduction of the family cap.

The estimation technique applies ordinary least squares (OLS) regression to model fertility in state *s* in month *t*. The base estimating equation takes the following form:

$$y_{st} = \alpha + \beta_1 * famcap_{st} + \beta_2 * wke1_{st} + \beta_3 * wke2_{st} + \beta_4 * wke3_{st} + \beta_5 * ln(welfare\ benefits)_{sy} + \beta_6 * ln(female\ pop\ 15\text{-}34)_{sy} + \beta_7 * (prop15\text{-}19)_{sy} + \beta_8 * (prop20\text{-}24)_{sy} + \beta_9 * (prop25\text{-}29)_{sy} + \beta_{10} * (unemp\ rate)_{s(y-1)} + \gamma_s + \sigma_t + \xi_s * time + \varepsilon_{st}$$

The variables are defined as follows:

y_{st} – log(total number of births in state *s* in month *t*)

$famcap_{st}$ – a binary indicator for whether a family cap has been in place for at least 6 months in state s in month t. The base specification sets the indicator to one for the 18 "full" family caps. (An alternative specification sets the indicator to one additionally for the five "partial" caps.)

$wke1_{st}$ – work exemption 1 – a binary indicator for whether the state exempts mothers with a child as old as six months to three years.

$wke2_{st}$ – work exemption 2 – a binary indicator for whether the state exempts mothers with a child newly born to six months old, but no older.

$wke3_{st}$ – work exemption 3 – a binary indicator for whether the state does not have any exemptions based on the age of the mother's child.

$welfare\ benefits_{sy}$ – the maximum monthly benefit for a family of three on AFDC/TANF, in 1998 dollars, in state s in year y.

$female\ pop\ 15\text{-}34_{sy}$ – the female population age 15 to 34, in state s in year y, according to the U.S. census.

$prop15\text{-}19_{sy}$, $prop20\text{-}24_{sy}$, $prop25\text{-}29_{sy}$ – the proportion of the female population age 15 to 34 in the different five-year age groups, in state s in year y. The proportion age 30 to 34 is the reference category.

$unemp\ rate_{s(y-1)}$ - the unemployment rate in state s in year y-1.

γ_s – a binary indicator for state s, to capture state fixed effects.

σ_t – month fixed effects for month t, to capture month fixed effects.

$\xi_s *time$ - linear time trend specific to state s.

The dependent variable is the log of the total number of births in state s in month t to women age 15 to 34. The distribution of total monthly births in a state is highly skewed, so a log transformation is preferred. Furthermore, the log transformation converts changes to percentage terms which aids in interpretation. There is potential problem, however, with focusing on state birth rates. The appropriateness of this approach relies on the assumption that there is not widespread migration in response to family cap policies. This assumption finds support in Levine and Zimmerman (1999). Their paper evaluates the extent to which differences in welfare generosity across states leads to interstate migration and concludes that welfare induced migration is not a widespread phenomenon.

There are a total of 6,120 observations: 51 states*120 months. In the base specification, all births to women ages 15 to 34 are included in the birth counts. This sample restriction is imposed in order to target a group of women for whom welfare policies might be relevant. In an effort to better target the analysis to the appropriate population - while continuing to avoid the selection problems of conditioning on welfare participation - the model is estimated separately for demographic subgroups with varying welfare proportions. It is also estimated separately for first and high-order births. This latter distinction is instructive since the family cap should have a direct effect on higher-order births, as discussed above, but only an indirect effect, if any, on first births.

The variable of interest is the binary indicator for a family cap, *famcap*. The 18 states that implement policies between 1989 and 1998 eliminating additional cash assistance for a child born to a mother on welfare are referred to as ever-treated states. (A more inclusive specification also considers the five states that alter the form of additional assistance to be treated states.) The other 32 states and the District of Columbia are considered never-treated states. For the never-treated states, the family cap indicator is always equal to zero. There is variation across the ever-treated states in the timing of family cap implementation. For these states, the family cap indicator takes on a value of one if the observation represents a month that occurs at least 6 months after the state's family cap policy was implemented. Allowing a 6-month lag recognizes that conception responses can not take place within 9 months of the policy implementation and that most abortion responses will occur in the first trimester of pregnancy. An alternative specification uses a family cap indicator that incorporates a 9-month lag. In this specification there is assumed to be no immediate abortion response, but subsequent responses can either be through conception or abortion.

Work exemption variables are included in the model to control for the effect that tightened work restrictions might have on fertility decisions. These three variables are mutually exclusive. The omitted category is a traditional AFDC/JOBS exemption policy. As discussed above, under AFDC, primary caretaker relatives of children up to six years of age, or up to three if the state guaranteed childcare, were exempt from the JOBS requirement. Under TANF, all states imposed tightened work requirements with welfare recipients required to work sooner in terms of the age of their youngest child than previous law stipulated; a number of states tightened their exemption rules before the implementation of TANF under AFDC waivers. The relative cost of having a child is higher when the child does not exempt the mother from work requirements. Economic reasoning thus implies the sign of $\beta_2, \beta_3,$ and β_4 to be non-positive, and since *wke1* represents the least strict non-AFDC policy and *wke3* represents the most strict, we expect and $\beta_4 <= \beta_3 <= \beta_2$.

Welfare benefits are controlled for in the model to account for any change in benefit levels that might be correlated with the introduction of family cap policies. The regressor is defined as the natural logarithm of the maximum monthly benefit for a family of three on AFDC/TANF and it varies by state and year.[32] Eleven states have explicitly changed their benefit levels under TANF. Almost all states have seen the inflation-adjusted level of welfare benefits fall during the nineties. All else equal, a higher benefit level makes raising a family on welfare easier; we thus expect birth rates to be positively correlated with benefit levels. The predicted sign on β_5 is thus positive.

The natural logarithm of the female population age 15 to 34 is also explicitly included in the regression model. If the female population is trending non-linearly, then the state-specific linear time trend will not adequately capture population movements, which undoubtedly affect birth counts. The proportion of women in each five-year age group is also controlled for, to account for idiosyncratic demographic shifts that might be spuriously correlated with the implementation of family cap policies. These population variables are based on figures from the U.S. census bureau. The model also explicitly controls for the state unemployment rate lagged one year in order to capture shifts in economic conditions that are not uniform across states nor are adequately described by a linear state trend.

The identification strategy makes the assumption that birth rates in states that implement family caps are not trending differently than birth rates in other states pre-cap. There is some evidence that this is the case, as discussed below. Note that this assumption is

[32] I thank Robert Shoeni for providing me with this data, which was used in the 1999 CEA report.

tantamount to claiming that the introduction of family cap legislation is exogenous to birth rate trends. One would prefer not to rely on this assumption, but there is no obvious way to avoid it. It is somewhat relaxed, however, by the inclusion of state specific time trends, denoted in months. These controls allow fertility rates to trend uniquely, albeit linearly, for each state without undermining the viability of this empirical strategy. It does appear that ever-treated states have, on average, higher birth totals than never-treated states. Individual state fixed effects are included in the regression to control for level differences. Month dummies are included in the model to account for any idiosyncratic movements in birth rates common to all states.

One way to indirectly test for the endogenous introduction of a family cap is to see whether there is a spurious correlation between a family cap policy and birth rates by looking for effects before they would reasonably be expected. For example, the results suggest that the *approval* of a family cap waiver is not positively associated with births. (This is discussed in more detail below.) Another way to potentially address endogeneity is to distinguish between states that adopted the family cap as part of TANF and those that requested family cap waivers.[33] It is a reasonable observation that states that request a waiver are more likely to be responding to shifts in birth outcomes than are those states that implement a cap as part of the national reform of welfare. Unfortunately, 19 out of the 23 states that enact any type of family cap legislation requested waivers to do so. In the spirit of the idea, however, I drop the set of five states that had waivers approved before 1995 with the idea that a family cap was still a novel idea, as opposed to something fairly common. Estimating the base equation for this reduced sample does not alter the results.

5 Results and conclusion

5.1 Mean fertility rates, by treatment status

[33] I thank a thoughtful referee for this suggestion.

Figure 1 plots annual fertility rates, defined as births per 1,000 women in the relevant age group, for women age 15 to 34, women age 20 to 34, and women age 15 to 19. Comparing trends in the early part of the decade for ever-treated and never-treated states offers no evidence of divergent trends. It appears that the two sets of states are on similar paths. Furthermore, at no point do annual fertility trends seem to decline more sharply (or increase less steeply) for states that implement caps relative to states that do not, as we would expect to see if family cap policies reduced births. If there is any break in trend at all between the two groups, it appears that fertility rates for women age 20 to 34 start to pull ahead of never-treated states in the later part of the decade. There is no logical reason to suspect family cap policies to encourage births, so this is presumably spurious. It will be looked at in more detail below. Among teenage women age 15 to 19, annual fertility rates trend together for the two groups of states throughout the entire period 1989 to 1998.

Table 1 also makes the point that relative to never-treated states, ever-treated states do not experience a decline in birth rates. Mean total monthly births and fertility rates are listed for the overall sample of states, for never-treated states pre and post 1996, and for ever-treated states pre and post family cap implementation (with a six-month lag). The year 1996 is chosen as the break-point for never-treated states because only one state implements a family cap after 1996. The table reports that among women age 15 to 34 in ever-treated states, fertility declines from an average monthly rate of 7.7 births per 1,000 women to an average monthly rate of 7.4 births per 1,000 women between years with and without a family cap. However, for this same age group of women in never-treated states, fertility declines from an average monthly rate of 7.7 before 1996 to a rate of 7.3 in the years 1996 to 1998. For women age 20 to 34 there is no decline in fertility among either ever-treated or never-treated states, and for teenage women 15 to 19, the average monthly fertility rate falls by 1.2 births per 1,000 women in both sets of states.

5.2 Regression results - all births to women 15 to 34

Table 2 displays the results from estimating the above equation, and slight variations of it, for births to all women age 15 to 34. In all six specifications displayed in this table, the family cap indicator equals one if the observation state and month has a "full" family cap implemented for at least six months. All control for state and month fixed effects, as well as state-specific linear time trends. Recall that it is important to control for time trends in case state birth rates are trending differently and are spuriously correlated with the implementation of a family cap. The first regression looks at the effect of the family cap, only controlling for population and population composition. The coefficient on family cap is estimated to be zero, with a standard error of .003. The results in column 2 show that the inclusion of an indicator for a time limit does not change the point estimate for the family cap coefficient; it is also clear that the time limit indicator does not belong in the model.

Column 3 reports the results of the preferred specification. The inclusion of controls for the maximum monthly AFDC/TANF benefit and work exemption policies bring the estimated coefficient on the family cap indicator to .002, with a standard error of .003. Furthermore, these controls generally enter the model as expected: benefit levels are positively correlated with birth rates, and indicators for the work exemption rules are non-positive. Augmenting the model does not alter the estimated coefficient on family cap. The specification of column 4 includes an indicator for TANF being implemented for at least six months; column 5 includes an indicator for a major waiver being implemented for at least six months; and column 6 includes indicators for both, with the waiver variable being set to zero when the TANF variable is equal to one. None of these controls matter.[34]

Seven specification checks are run using the preferred model described above (shown in column 3 of Table 2). The results of these specification checks are reported in Table 3. The first five variations on the model redefine the family cap. Previously the family cap indicator was set to one if a state eliminated additional cash benefits entirely, which 18 states ultimately do. Redefining the indicator to equal one if the state has any type of cap in place for at least six

[34] Shoeni and Blank (2000) find no significant effects of TANF on work participation, weeks worked, hours worked, own earnings, or family earnings, though they do find a negative effect on caseloads. The Shoeni and Blank findings, combined with the results estimated here, suggest that in the years immediately following its implementation, TANF affected welfare participation, but not work and family planning decisions. However, it is possible that responses will take longer to become evident.

months does produce the expected negative coefficient. As shown in column 1, the point estimate is .006, with a standard error of .003. Columns 2 and 3 incorporate a longer lag time into the family cap indicator. The base case tests for a response starting 6 months after a cap is instituted. This allows an initial abortion response from women in their first trimester of pregnancy and subsequent responses through either pregnancy avoidance or abortion. A nine-month lag assumes there is no initial abortion response, and a twelve-month lag assumes even more time is needed for either response. Still, we fail to estimate a negative effect of the family cap. Oddly, the coefficient on family cap in these specifications is significantly positive. The point estimates are small, suggesting an increase of 0.6 or 0.9 percent, but still surprising. This issue will be addressed in Table 7.

Columns 4 and 5 incorporate a shorter lag into the family cap indicator. The results suggest that there is no response - either negative or positive - to either the approval of a family cap or the implementation of a family cap in the first month. The point estimate in both cases is .002, with a standard error of .003. This suggests, with 95 percent confidence, that the effect is between a decline in births of 0.4 percent and an increase in births of 0.8 percent, i.e., if it is non-zero, it is tiny. This is comforting in that it suggests there is not a spurious correlation between the adoption of a cap and birth rates. Recall that the efficacy of the empirical strategy assumes that family caps are not adopted in response to shifts in monthly birth totals. A more skeptical interpretation of this result is that the adoption of a family cap is positively correlated with births, but that there is an "anticipation" effect whereby women avoid pregnancy when they hear a family cap is about to be imposed, and the two offset to zero. Because there appears to be no decline in births after the family cap is implementation, this interpretation seems highly unlikely.

Two final checks are reported in Table 3. The regression reported in column 6 estimates the equation for the dependent variable natural log of the fertility rate, instead of the natural log of total births. The mean fertility rate across the 6,120 state-month cells in the sample is 7.6 births per 1,000 women. Defining the dependent variable this way essentially moves the log of the female population age 15 to 34 from the right hand side of the equation to the left hand side, i.e. it restricts the coefficient to be one. It does alter the finding of no decline. The estimated effect lies between a 0.6 percent decline in the fertility rate to a 0.7 percent increase. In other words, if the effect is in fact non-zero, it is less than one percent. The last specification check involves dropping from the sample those states that had a family cap approved before 1995. As discussed above, the point of this exercise is to address potential endogeneity of a family cap policy by removing the states that applied for a family cap waiver when the concept of a cap was still novel. The point estimate in this specification is negative 0.3 percent and is not significantly different from zero.

It should be emphasized that the finding of no effect is not due to imprecisely estimated coefficients. There is enough power in the empirical design to reject estimates the size of those reported in other studies of the family cap. Camasso et. al. (1999) report that births declined by nine percent among longer-term welfare recipients and by twelve percent among their sample of new applicants. The results based on a lagged family cap indicator in Horvath-Rose and Peters (2000) suggest declines of eleven percent for white women age 20 to 49 and approximately eight percent for black women age 20 to 49. The analysis of ten years of state monthly birth totals has enough power to reject a decline of more than 0.5 percent at the 95 percent confidence level, in the preferred specification (reported in column 3 of Table 2). The most negative point estimate found among the full sample of women age 15 to 34 is -.003, which is obtained when states with pre-1995 waiver approvals are dropped from the sample. Even in this specification, the lower-bound on the 95 percent confidence interval is -0.015. That is, the test has enough power to reject an effect larger than 1.5 percent for the full sample of women age 15 to 34.

5.3 Births by demographic groups and birth order

Assume that women at risk of welfare dependence have additional births less frequently in response to the family cap, but that these births comprise only a small fraction of births born to women ages 15 to 34. The power of the empirical analysis might not be strong enough to detect an effect. In order to increase the power of the analysis, I estimate the equation separately for demographic groups that have different welfare propensities and separately for first and higher-order births. The disadvantage of this approach is that the smaller sample size necessarily decreases the test's power. The results are displayed in Tables 4, 5, and 6. As discussed above, since the family cap policy denies additional benefits for additional children, the direct effect is on higher-order births. Any negative effect on first-births is presumably through a welfare reform "threat" effect.

Twelve demographic groups are identified, based on race, marital status, and education. For ease of analysis and exposition, women whose race is not classified as "black" in the natality files are considered "white". Marital status is categorized as unmarried or married. For women ages 20 to 34, I define two education categories, high school graduate and high school dropout. Being a teenager is classified as an education category of its own in order to avoid mislabeling young women who are still in school as high school dropouts. Mean total monthly births for these groups are listed in Table 1.

The highest welfare participation rate is among black, unmarried, high school dropouts (age 20 to 34). In the March 1989 CPS, there are 267 observations in this demographic group and the welfare participation rate is 61.9 percent. The results of estimating the model on this subset of women are shown in the first panel of Table 4. The family cap does not appear to reduce either additional or first births for this group. The standard errors are larger than for the full sample of women age 15 to 34. The 95% confidence interval around the point estimate for higher order births is 0.0002 to 0.086; that is, no decline in births falls within the confidence interval. The positive coefficient is troublesome. It is also surprising that the work exemption variables enter with positive coefficients; AFDC/TANF benefit levels enter with the expected positive sign. White, unmarried, high school dropouts have the next highest welfare participation rate. Of the 820 such observations in the March 1989 CPS, 36.9 percent report receipt of welfare payments. The results are similar in that no decline is detected in either additional or first births (Table 5, panel 1). And again, for additional births, the coefficient on family cap is curiously positive and statistically significant.

A sizeable percentage of black, unmarried high school graduates receive welfare. According to the 1989 CPS, the participation rate is 22.0 percent, with a sample size of 1,077. As with the full sample results, the estimated result is indistinguishable from zero, though the coefficient is now less precisely estimated. The 95 percent confidence interval for additional births extends from a decline of 3.0 percent to an increase of 3.4 percent. It is even wider for first births, encompassing changes between an 7.5 percent decline to a 5.2 percent increase.

The other groups have very low welfare participation rates. Unmarried white high school dropouts (age 20 to 34) have a participation rate of 5.5 percent in the 1989 CPS. Married women are generally ineligible: 12.9 percent of married black high school dropouts report welfare receipt and only 2.7 percent of high school graduates do. For married white women, only 5.3 percent of high school dropouts and 1.2 percent of high school graduates report welfare receipt.

Most teenagers do not have children, so the percentage of teenage females on welfare is an understatement of the percentage of teenage mothers who receive welfare. Furthermore, some of the younger teens reporting welfare receipt are presumably dependents. Nonetheless, the relative proportions are useful to note. In the March 1989 CPS, 8.8 percent of the 642 unmarried black teens report positive welfare receipt during the previous year; 1.7 percent of unmarried white teens do. There are very few married teenagers in the CPS. Table 6 displays the results for teenage women age 15 to 19. Again, there is no statistical evidence that birth rates are declining in response to family caps. Unfortunately, the estimates are imprecise. The estimated effect of the family cap on additional births to black unmarried teens is positive. For unmarried white teens the effect is more reasonable, an estimated 1.1 percent decline, but the 95% confidence interval ranges from a decline of 3.0 to an increase of 0.8.

A potential objection to looking at results for unmarried women is that marriage decisions are potentially affected by the introduction of the family cap. Under this line of reasoning, unmarried women who get pregnant under a family cap respond by getting married. Were this the case, we should see a decrease in births to unmarried women and a corresponding increase in births to married women. This story is not supported by the data. The results shown in Tables 4, 5, and 6 do not indicate that the introduction of a family cap policy is associated with an increase in the number of higher-order births born to married women.

The specifications displayed in Table 7 explore the positive coefficient on family cap found for unmarried high school dropouts age 20 to 34 and for unmarried black teenage women. For completeness, the augmented model is also estimated on the sample of unmarried white teenagers. The model includes a series of seven dummy variables to capture the effect of the family cap three to six months before, zero to two months before, one to three months after, four to six months after, seven to nine months after, ten to twelve months after, and more than a year after. If the positive coefficient is picking up a spurious correlation between birth rates and the introduction of a family cap policy, we might worry that the policy is not "exogenously" implemented. This does not appear to be the case. For unmarried high school dropouts, the positive association does not appear until more than a year after the family cap has been implemented. For unmarried black teenagers the positive association first appears seven months after the policy. There is some spurious association that is not explained by population shifts nor unemployment rates. Future research would be well served by exploring this odd result with additional years of data.

6 Discussion

The finding of no systematic effect of the family cap on fertility rates is in line with previous studies that fail to find a relationship between welfare and family formation decisions. If this empirical result is correct, then the widespread adoption of the family cap as a state welfare policy appears ineffective at best and misguided at worst. Women are not responding by having fewer additional births, and consequently, fewer resources are being provided per child on welfare.

The results of this chapter are unequivocal: the implementation of family cap policies does not appear to affect the decision to have a child. However, there is limited post-family cap data. Most states that eliminated cash assistance for additional children did so in 1995 and 1996 and vital statistics birth data is only available through 1998. It is possible that effects on fertility will not be evident for another few years. Future research should investigate whether the finding of no response holds in the long run.

References

[1] Argys, Laura M. and Daniel I. Rees. 1996. "The Impact of Welfare Generosity on the Fertility Behavior of Recipients," Working Paper #9601. University of Colorado at Denver, Center for Research on Economic and Social Policy.

[2] Becker, Gary. 1981. *A Treatise on the Family.* Cambridge: Harvard University Press.

[3] Bronars, Stephen G. and Jeff Grogger. 1994. "The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment," *American Economic Review,* 84(5): 1141-1156.

[4] Camasso, Michael, Carol Harvey, Radha Jagannathan, and Mark Killingsworth. 1999. *New Jersey's Family Cap and Family Size Decisions: Some Findings from a 5-Year Evaluation.* Rutgers University.

[5] Crouse, Gil. 1999. *State Implementation of Major Changes to Welfare Policies, 1992-1998.* Office of Human Services Policy, Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services. www.aspe.hhs.gov/hsp/Waiver-Policies99/policy_CEA.htm

[6] Council of Economic Advisors. 1999. *Technical Report: Economic Expansion, Welfare Reform, and the Decline in Welfare Caseloads: An Update.* Executive Office of the President of the United States.

[7] Danziger, Sheldon, Robert Haveman, and Robert Plotnick. 1981. "How Income Transfers Affect Work, Savings, and The Income Distribution: A Critical Review," *Journal of Economic Literature*, 19(3): 975-1028.

[8] Duncan, G.J. and S.D. Hoffman. 1990. "Welfare Benefits, Economic Opportunities and Out-of-Wedlock Births Among Black Teenage Girls," *Demography,* 27: 519-535.

[9] Ellwood, David and Mary Jo Bane. 1985. "The Impact of AFDC on Family Structure and Living Arrangements." In R. Ehrenberg (ed.), *Research in Labor Economics*, 7: 137-207.

[10] Fairlie, R.W. and R.A. London. 1997. "The Effect of Incremental Benefit Levels on Births to AFDC Recipients," *Journal of Policy Analysis and Management*, 16; 575-597.

[11] Gallagher, L. Jerome, Megan Gallagher, Kevin Perese, Susan Schreiber, and Keith Watson. 1998. "One Year after Federal Welfare Reform: A Description of State Temporary Assistance for Needy Families (TANF) Decisions as of October 1997," *Occasional Paper Number6.* Washington D.C.: Urban Institute. http://newfederalism.urban.org/html/occas6.htm.

[12] Green Book. 1998. "Section 7: Aid to Families with Dependent Children and Temporary Assistance for Needy families," Ways and Means Committee, from the U.S. Government Printing Office Online via GPO Access. http://frwebgate.access.gpo.gov.

[13] Horvath-Rose, Anne and H. Elizabeth Peters. 2000. "Welfare Waivers and Nonmarital Childbearing," *Joint Center for Poverty Research Working Paper* 128.

[14] Levine, Phillip B. and D.J. Zimmerman. 1999. "An Empirical Analysis of the Welfare Magnet Debate," *Journal of Population Economics*, Vol. 12, No. 3: 391-409.

[15] Loury, Glenn C. 2000. "Preventing Subsequent Births to Welfare Recipients," in Douglas J. Besharov and Peter Germanis, eds, *Preventing Subsequent Births to Welfare Recipients.* College Park, Maryland: School of Public Affairs.

[16] Lundberg, S. and R.D. Plotnick. 1990. "Effects of State Welfare, Abortion, and Family Planning Policies on Premarital Childbearing Among White Adolescents," *Family Planning Perspectives*, 29 (2): 52-60.

[17] Moffitt, Robert. 1992. "Incentive Effects of the U.S. Welfare System: A Review," *Journal of Economic Literature*, 30: 1- 61.

[18] Moffitt, Robert. 1996. "The Effect of Welfare on Marriage and Fertility: What Do We Know and What Do We Need to Know?" Baltimore, MD: Johns Hopkins University, Department of Economics.

[19] Robins, P.K. and P. Fronstin. 1996. "Welfare Benefits and Birth Decisions of Never-Married Women," *Population Research and Policy Review*, 15: 21-43.

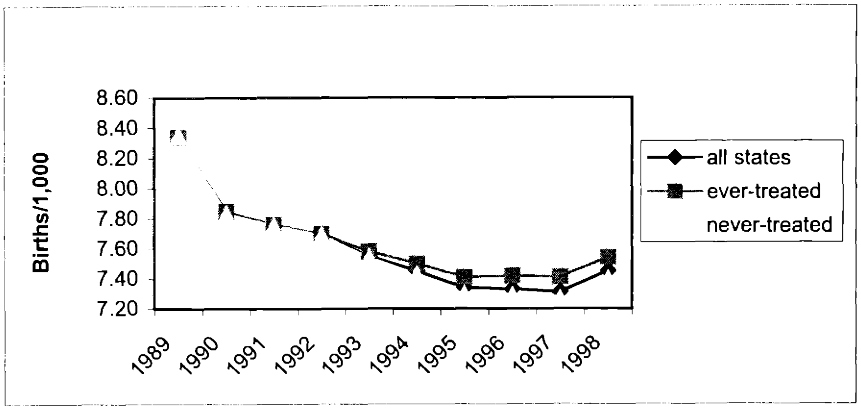
[20] Schoeni, Robert F. and Rebecca M. Blank. 2000. "What has welfare reform accomplished? Impacts on welfare participation, employment, income, poverty, and family structure," *NBER Working Paper* 7627. Cambridge, MA.

[21] "Setting the Baseline: A Report on State Welfare Waivers," 1997. Office of the Assistant Secretary for Planning and Evaluations, U.S. Department of Health and Human Services. http://aspe.hhs.gov/hsp/isp/wavier2/title.htm.

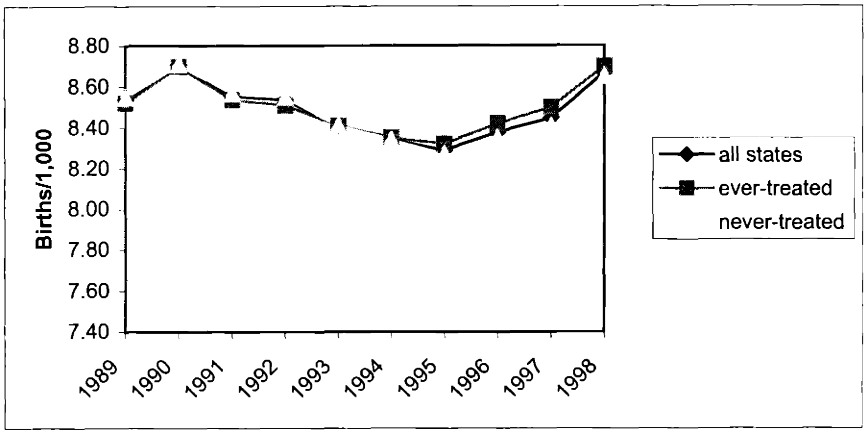
[22] Turturro, Carolyn, Brent Benda, and Howard Turney. 1997. *Arkansas Welfare Waiver Demonstration Project, Final Report (July 1994 through June 1997).* University of Arkansas at Little Rock, School of Social Work.

Figure 1: Monthly Fertility Rates

All women age 15 to 34



All women age 20 to 34



All women age 15 to 19

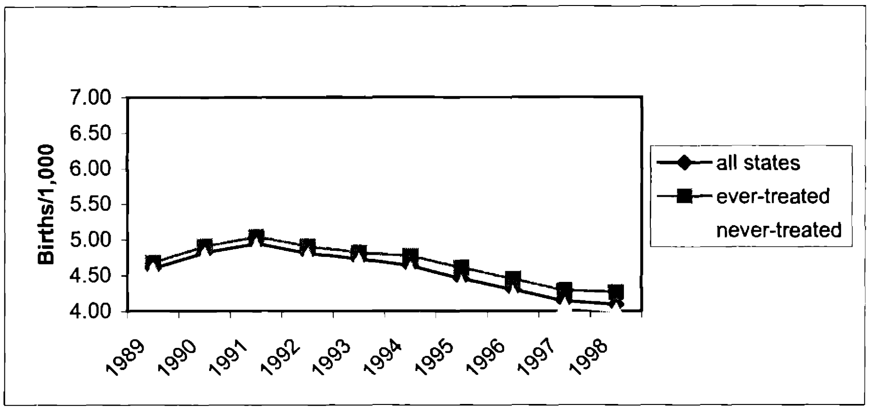


Table 1:
Mean Total Monthly Births,
across states 1989 to 1998

	Overall	Never-treated, 1989-1995	Never-treated, 1996-1998	Ever-treated, pre-family cap	Ever-treated, post-family cap
all women age 15-34	5,809	5,221	4,933	7,362	6,152
per 1,000 women	7.6	7.7	7.3	7.7	7.4
all women age 20-34	4,988	4,486	4,222	6,328	5,288
per 1,000 women	8.5	8.5	8.5	8.5	8.5
all female teens 15-19	820	735	711	1,034	863
per 1,000 women	5.0	5.3	4.1	5.6	4.4
unmarried HS dropouts age 20-34					
black	130	127	105	153	136
white	283	196	228	472	291
unmarried HS grads age 20-34					
black	371	332	316	450	463
white	546	466	554	666	583
married HS dropouts age 20-34					
black	30	32	23	35	25
white	432	346	307	690	392
married HS graduates age 20-34					
black	252	233	200	303	306
white	2,996	2,806	2,530	3,630	3,118
unmarried teens					
black	236	212	191	295	288
white	380	318	378	480	408
married teens					
black	21	21	11	21	12
white	276	209	150	276	167

Source: Vital Statistics Natality Data, Public-Use Data Files, years 1989 to 1998, compiled by the U.S. National Center for Health Statistics (NCHS); annual population estimates by state and age group come from the U.S. census bureau (figures are not available by race.)
Note: "Ever-treated" in this table is defined as ever eliminating additional cash benefits.

Table 2
Dependent variable: log births
Women ages 15 to 34

	(1)		(2)		(3)		(4)		(5)		(6)	
famcap	.000		.000		.002		.002		.001		.001	
	(.003)		(.003)		(.003)		(.003)		(.004)		(.004)	
ln(max monthly benefits)	-		-		.116	***	.114	***	.116	***	.115	***
					(.021)		(.021)		(.021)		(.021)	
work exemption 1	-		-		.001		-.0001		.001		-.001	
					(.003)		(.003)		(.003)		(.004)	
work exemption 2	-		-		-.006		-.007		-.006		-.007	
					(.004)		(.005)		(.004)		(.005)	
work exemption 3	-		-		.000		-.001		.000		-.001	
					(.005)		(.005)		(.005)		(.005)	
ln(female pop 15 to 34)	.600	***	.600	***	.524	***	.522	***	.524	***	.519	***
	(.098)		(.098)		(.098)		(.098)		(.098)		(.098)	
prop1519	.219		.219		.182		.186		.176		.168	
	(.242)		(.242)		(.246)		(.246)		(.247)		(.267)	
prop2024	.624	***	.624	***	.677	***	.680	***	.673	***	.674	***
	(.200)		(.200)		(.199)		(.199)		(.199)		(.199)	
prop2529	1.14	***	1.14	***	1.11	***	1.11	***	1.10	***	1.09	***
	(.236)		(.236)		(.234)		(.234)		(.236)		(.236)	
time limit	-		.000		-		-		-		-.002	
			(.003)								(.003)	
tanf-official	-		-		-		.004		-		.007	
							(.005)				(.005)	
waiver	-		-		-		-		.001		.003	
									(.002)		(.003)	
lagged state unemp rate	-.009	*	-.009	*	-.009	*	-.009	*	-.009	*	-.009	*
	(.005)		(.005)		(.005)		(.005)		(.005)		(.005)	
state effects	yes		yes		yes		yes		yes		yes	
month effects	yes		yes		yes		yes		yes		yes	
state time trend	yes		yes		yes		yes		yes		yes	
constant	-.155		-.155		.530		.567		.534			
	(1.25)		(1.25)		(1.24)		(1.24)		(1.24)			
R^2	.99		.99		.99		.99		.99		.99	
Sample size	6,120		6,120		6,120		6,120		6,120		6,120	

notes:
Standard errors are White's robust standard errors. Significance levels: * 90 percent level, ** 95 percent, *** 99 percent.
The family cap, time limit, tanf, and waiver variables are binary variables that equal one if the particular policy has been implemented for at least six months. In these specifications, *famcap* indicates policies that eliminated cash assistance for any additional child; the indicator is set to one for 18 states at some point in the data. In column (6), the waiver indicator equals zero when the TANF indicator gets set to one.

Table 3
Dependent variable: log births
Women ages 15 to 34

	(1)	(2)	(3)	(4)	(5)	(6) dep. var.: ln(fert rate)	(7)
famcap	-	-	-	-	-	.0002	-.003
						(.003)	(.006)
famcap -any	.006 *	-	-	-	-	-	-
	(.003)						
famcap – 1 mo. lag	-	-	-	.002	-	-	-
				(.003)			
famcap - 9 mo. lag	-	.006 *	-	-	-	-	-
		(.003)					
famcap - 12 mo. lag	-	-	.009 ***	-	-	-	-
			(.003)				
famcap - approved	-	-	-	-	.002	-	-
					(.003)		
ln(max monthly benefits)	.115 ***	.115 ***	.114 ***	.116 ***	.116 ***	.099 ***	.107 ***
	(.020)	(.020)	(.020)	(.021)	(.021)	(.020)	(.022)
work exemption 1	-.0001	.001	.001	.001	.001	.002	.002
	(.003)	(.003)	(.003)	(.003)	(.003)	(.003)	(.004)
work exemption 2	-.007 *	-.007	-.007 *	-.006	-.006	-.005	-.010 *
	(.004)	(.004)	(.004)	(.004)	(.004)	(.004)	(.006)
work exemption 3	-.001	.000	-.0002	.000		.002	-.007
	(.005)	(.005)	(.005)	(.005)		(.005)	(.005)
ln(female pop 15 to 34)	.518 ***	.521 ***	.523 ***	.522 ***	.523 ***	-	.595 ***
	(.098)	(.098)	(.097)	(.098)	(.098)		(.107)
prop1519	.185	.193	.192	.189	.185	-.127	.032
	(.246)	(.246)	(.246)	(.246)	(.246)	(.236)	(.264)
prop2024	.689 ***	.700 ***	.714 ***	.684 ***	.682 ***	.377 *	.769 ***
	(.198)	(.198)	(.197)	(.199)	(.199)	(.191)	(.210)
prop2529	1.13 ***	1.13 ***	1.15 ***	1.11 ***	1.11 ***	.738 ***	1.29 ***
	(.233)	(.234)	(.234)	(.234)	(.234)	(.225)	(.252)
lagged state unemp. rate	-.009 *	-.010 *	-.010 *	-.009 *	-.009 *	-.007	-.017 ***
	(.005)	(.005)	(.005)	(.005)	(.005)	(.005)	(.006)
state effects	yes	yes	yes	yes	yes	yes	yes
month effects	yes	yes	yes	yes	yes	yes	yes
state time trend	yes	yes	yes	yes	yes	yes	yes
constant	.608	.559	.530	.560	.546	1.47 ***	-.482
	(1.24)	(1.24)	(1.24)	(1.25)	(1.25)	(.149)	(1.35)
R^2	.99	.99	.99	.99	.99	.90	.99
Sample size	6,120	6,120	6,120	6,120	6,120	6,120	4,800

notes: Standard errors are White's robust standard errors. Significance levels: * 90 percent level, ** 95 percent, *** 99 percent.
The basic *famcap* variable indicates policies that eliminated cash assistance for any additional child; the indicator is set to one for 18 states at some point in the data. *famcap-any* indicates any policy that altered the amount or form of additional benefits; 23 states have *famcap-any* equal to one at some point in the analysis. *famcap- 1 mo. lag, famcap- 9 mo. lag* and *famcap-12 mo. lag.* refer to a family cap being implemented for at least 1, 9, and 12 months. *famcap-approved* refers to the date the family cap was approved, as opposed to implemented.
In column (6), fertility rate is defined as births per 1,000 women age 15 to 34. Mean fertility rate in the data is 7.6 births per 1,000 women.
In column (7), states that had family cap waivers approved before 1995 are dropped from the estimation sample.

Table 4
Dependent variable: log births
Black women ages 20 to 34

	Unmarried				Married			
	HS dropout (1)		HS grad (2)		HS dropout (3)		HS grad (4)	
	add. births	first births	add. births	first births	add. births	first births	add. births	first births
famcap	.043 *	.012	.002	-.012	.064 *	-.030	.013	.028
	(.022)	(.036)	(.016)	(.032)	(.034)	(.055)	(.021)	(.043)
ln(max monthly benefits)	.418 **	.324	-.244 *	.574 **	-.124	-.169	-.015	-.568 **
	(.176)	(.265)	(.146)	(.256)	(.244)	(.355)	(.143)	(.279)
work exemption 1	.056 **	.029	-.051 ***	-.005	.039	.021	.004	.038
	(.025)	(.037)	(.017)	(.031)	(.033)	(.053)	(.019)	(.042)
work exemption 2	.081 *	.101 *	-.036	.117 ***	-.039	-.043	-.006	.146 **
	(.040)	(.057)	(.024)	(.043)	(.047)	(.075)	(.029)	(.062)
work exemption 3	.072	.145 ***	-.040	.127 ***	-.205 ***	-.135 *	-.088 **	-.066
	(.046)	(.055)	(.043)	(.050)	(.060)	(.073)	(.044)	(.064)
ln(female pop 20 to 34)	1.58 *	1.33	1.73 ***	.881	1.22	-1.84	.919	.318
	(.897)	(1.27)	(.657)	(1.16)	(1.25)	(1.66)	(.714)	(1.39)
prop2024	.555	1.48	1.08	3.41	3.55 *	-2.19	.984	-1.22
	(1.46)	(2.30)	(1.09)	(2.36)	(1.94)	(2.81)	(.984)	(2.35)
prop2529	1.50	6.22 **	1.32	1.74	7.35 ***	-.301	1.96	-.413
	(1.86)	(2.81)	(1.43)	(2.65)	(2.64)	(3.68)	(1.40)	(3.10)
lagged state unemp. rate	.119 ***	-.079	.021	-.009	.043	.081	-.119 ***	-.150 *
	(.045)	(.095)	(.037)	(.074)	(.061)	(.090)	(.038)	(.084)
state effects	yes	yes	yes	yes	yes	yes	yes	yes
month effects	yes	yes	yes	yes	yes	yes	yes	yes
state time trend	yes	yes	yes	yes	yes	yes	yes	yes
constant	-18.2	-19.5	-18.1 **	-13.2	-16.6	27.0	-9.06	-1.35
	(11.4)	(16.5)	(8.35)	(15.2)	(16.1)	(21.6)	(9.10)	(17.7)
R^2	.97	.93	.99	.94	.93	.80	.98	.88
Sample size	4,981	4,608	5,701	4,995	4,812	3,655	5,989	4,677

notes:
Standard errors are White's robust standard errors. Significance levels: * 90 percent level, ** 95 percent, *** 99 percent.
The family cap, time limit, tanf, and waiver variables are binary variables that equal one if the particular policy has been implemented for at least six months.
The proportion of demographic subgroup on welfare (not restricted to mothers), based on weighted means from the 1989 March *Current Population Survey*: (1) 61.9, n= 267; (2) 22.0, n=1,077; (3) 12.9, n=75; (4) 2.7, n=556.
Mean total births is calculated across state and month cells.

Table 5
Dependent variable: log births
White women ages 20 to 34

	Unmarried				Married			
	HS dropout (1)		HS grad (2)		HS dropout (3)		HS grad (4)	
	add. births	first births	add. births	first births	add. births	first births	add. births	first births
famcap	.062 ***	.008	-.013	-.004	.046 ***	.043	-.006	-.001
	(.019)	(.029)	(.009)	(.025)	(.015)	(.028)	(.004)	(.025)
ln(max monthly	.481 ***	.791 ***	-.107 *	.478 ***	.254 **	.461 **	.052 *	.369 **
benefits)	(.129)	(.185)	(.057)	(.170)	(.115)	(.207)	(.027)	(.162)
work exemption 1	-.044 **	-.093 ***	.013	-.091 ***	.004	-.084 ***	.007 **	-.076 ***
	(.020)	(.022)	(.008)	(.019)	(.019)	(.026)	(.003)	(.019)
work exemption 2	.024	.057 *	-.024 **	.025	.018	.086 ***	-.024 ***	.028
	(.025)	(.032)	(.011)	(.028)	(.023)	(.033)	(.005)	(.027)
work exemption 3	.092 ***	.168 ***	.076 ***	.148 ***	.029	.160 ***	-.025 ***	.101 ***
	(.025)	(.029)	(.014)	(.022)	(.021)	(.032)	(.005)	(.022)
ln(female pop 20 to	.763	2.10 ***	-.312	-.067	.432	1.31	.356 ***	-1.11
34)	(.646)	(.749)	(.336)	(.767)	(.564)	(.873)	(.132)	(.693)
prop2024	-5.17 ***	-4.47 ***	-.108	-.270	-7.99 ***	-6.17 ***	-.302	-.857
	(1.21)	(1.48)	(.750)	(1.33)	(1.19)	(1.66)	(.263)	(1.31)
prop2529	-.936	-5.25 ***	-1.12	-5.94 ***	-4.64 ***	-7.28 ***	.339	-4.69 ***
	(1.38)	(1.94)	(.848)	(1.79)	(1.29)	(2.16)	(.309)	(1.76)
lagged state unemp.	.181 ***	-.142 **	.040 **	-.245 ***	.131 ***	-.320 ***	.017 **	-.398 ***
rate	(.038)	(.065)	(.018)	(.063)	(.037)	(.074)	(.007)	(.069)
state effects	yes	yes	yes	yes	yes	yes	yes	yes
month effects	yes	yes	yes	yes	yes	yes	yes	yes
state time trend	yes	yes	yes	yes	yes	yes	yes	yes
constant	-4.38	-23.8 **	10.7 **	4.84	2.50	-11.9	3.06 *	18.9 **
	(8.08)	(9.39)	(4.27)	(9.69)	(6.99)	(10.8)	(1.66)	(8.71)
R^2	.96	.92	.99	.94	.97	.92	.99	.94
Sample size	6,116	5,935	6,120	6,009	6,120	5,900	6,120	6,080

notes:
Standard errors are White's robust standard errors. Significance levels: * 90 percent level, ** 95 percent, *** 99 percent.
Proportion of demographic subgroup on welfare (not restricted to mothers), based on weighted means from the 1989 March *Current Population Survey*:
(1) 36.9, n=820; (2) 5.5, n=5,843; (3) 5.3, n=1,129; (4) 1.2, n=8,240.
Mean total births is calculated across state and month cells.

Table 6
Dependent variable: log births
Teenage women ages 15 to 19

	Black				White			
	Unmarried (1)		Married (2)		Unmarried (3)		Married (4)	
	add. births	first births	add. births	first births	add. births	first births	add. births	first births
famcap	.067 ***	.034	.035	-.048	-.011	.007	.000	.013
	(.016)	(.040)	(.040)	(.073)	(.010)	(.026)	(.017)	(.034)
ln(max monthly	.332 *	1.28 ***	-.511 **	-1.08 **	.020	.580 ***	-.102	.050
benefits)	(.180)	(.282)	(.260)	(.450)	(.075)	(.166)	(.106)	(.256)
work exemption 1	-.037 **	-.017	-.024	.033	.018 *	-.079 ***	-.012	-.043
	(.018)	(.038)	(.040)	(.069)	(.009)	(.019)	(.018)	(.033)
work exemption 2	-.016	.131 **	-.051	-.028	.003	.027	-.016	.068
	(.025)	(.057)	(.052)	(.098)	(.012)	(.028)	(.026)	(.045)
work exemption 3	.067 *	.056	-.313 ***	-.185 *	.053 ***	.064 ***	-.065 ***	-.061
	(.037)	(.057)	(.074)	(.109)	(.015)	(.025)	(.025)	(.044)
ln(female pop 15 to	.009	1.16 **	-.195	.603	.923 ***	-.139	1.75 ***	.617
19)	(.206)	(.458)	(.499)	(.812)	(.249)	(.264)	(.256)	(.431)
lagged state unemp.	-.019	.223 ***	-.023	-.184	.073 ***	-.344 ***	.035	-.295 ***
rate	(.037)	(.073)	(.070)	(.131)	(.024)	(.065)	(.031)	(.080)
constant	3.44	-13.2 **	6.44	-3.66	-4.64 *	2.33	-13.9 ***	-4.33
	(2.43)	(5.18)	(5.62)	(9.52)	(2.65)	(2.84)	(2.76)	(4.69)
state effects	yes	yes	yes	yes	yes	yes	yes	yes
month effects	yes	yes	yes	yes	yes	yes	yes	yes
state time trend	yes	yes	yes	yes	yes	yes	yes	yes
R^2	.99	.87	.90	.44	.99	.93	.98	.91
Sample size	5,493	4,667	4,750	1,824	6,120	5,920	6,107	5,723

notes:
Standard errors are White's robust standard errors. Significance levels: * 90 percent level, ** 95 percent, *** 99 percent.
The family cap, time limit, tanf, and waiver variables are binary variables that equal one if the particular policy has been implemented for at least six months.
Proportion of demographic subgroup on welfare (not restricted to mothers), based on weighted means from the 1989 March *Current Population Survey*:
(1) 8.77, n=642; (2) 40.4 , n=6; (3) 1.72, n=4,394; (4) 5.3, n=203.
Mean total births is calculated across state and month cells.

Table 7
Dependent variable: log births
Higher-order births

	(1) Black, unmarried, HS dropouts 20-34		(2) White, unmarried, HS dropouts 20-34		(3) Black, unmarried teenagers 15-19		(4) White, unmarried teenagers 15-19	
famcap – 3 to 6 mos. before	.007		.021		.008		-.047	***
	(.031)		(.024)		(.021)		(.016)	
famcap – 0 to 2 mos. before	.017		-.029		.032		-.009	
	(.031)		(.031)		(.021)		(.015)	
famcap – 1 to 3 mos. after	.019		.036		.038		-.015	
	(.042)		(.037)		(.028)		(.017)	
famcap – 4 to 6 mos. after	.007		.034		.032		-.028	
	(.033)		(.029)		(.023)		(.018)	
famcap – 7 to 9 mos. after	.001		.019		.053	*	-.019	
	(.044)		(.042)		(.028)		(.016)	
famcap – 10 to 12 mos. after	.025		.049		.084	***	-.032	*
	(.039)		(.035)		(.029)		(.016)	
famcap –more than a year after	.079	***	.095	***	.090	***	-.020	
	(.029)		(.027)		(.021)		(.014)	
ln(max monthly benefits)	.412	**	.469	***	.329	*	-.007	
	(.177)		(.129)		(.177)		(.075)	
work exemption 1	.058		-.040	**	-.036	**	.016	*
	(.025)		(.020)		(.018)		(.009)	
work exemption 2	.076	*	.023		-.016		-.002	
	(.039)		(.026)		(.025)		(.012)	
work exemption 3	.073		.096	***	.067	*	.048	***
	(.047)		(.025)		(.037)		(.015)	
ln(female pop 15 to 34)	1.60	*	.788		-.200		.928	***
	(.912)		(.650)		(.282)		(.238)	
prop2024	.689		-5.05	***	-1.55		-.472	
	(1.48)		(1.22)		(1.38)		(.902)	
prop2529	1.70		-.794		-1.22		.739	
	(1.87)		(1.38)		(1.92)		(.793)	
lagged state unemp. rate	.114	**	.177	***	-.017		.076	***
	(.045)		(.039)		(.038)		(.025)	
state effects	yes		yes		yes		yes	
month effects	yes		yes		yes		yes	
state time trend	yes		yes		yes		yes	
constant	-18.6		-4.76		6.67		-4.77	*
	(-18.6)		(8.12)		(3.96)		(2.67)	
R^2	.97		.96		.99		.99	
Sample size	4,981		6,116		5,493		6,120	

notes:
Standard errors are White's robust standard errors. Significance levels: * 90 percent level, ** 95 percent, *** 99 percent.
The basic *famcap* variable indicates policies that eliminated cash assistance for any additional child; the indicator is set to one for 18 states at some point in the data.

Appendix Table 1
State family cap policies

	Date Implemented	Date Approved	No increase in assistance for add.child	Partial increase in cash assistance for add. child	Increase in assistance for add. child provided as voucher	Increase in cash assistance for add. child to third party
Arizona	11/95	5/95	X			
Arkansas	7/94	4/94	X			
California	9/97	8/96	X			
Connecticut	1/96	12/95		X		
Delaware	10/95	5/95	X			
Florida	10/96	6/96		X		
Georgia	1/94	11/93	X			
Idaho	7/97	-	X			
Illinois	12/95	9/95	X			
Indiana	5/95	12/94	X			
Maryland	3/96	8/95				X
Massachusetts	11/95	8/95	X			
Mississippi	10/95	9/95	X			
Nebraska	11/95	2/95	X			
New Jersey	10/92	7/92	X			
North Carolina	7/96	2/96	X			
North Dakota	7/97	-	X			
Oklahoma	10/96	-			X	
South Carolina	10/96	5/96			X	
Tennessee	9/96	7/96	X			
Virginia	7/95	7/95	X			
Wisconsin	1/96	6/94	X			
Wyoming	1/97	-	X			

Source: Urban Institute (1998) summary of state TANF policies; Crouse (1999) - note these are the same dates used in the 1999 CEA report; Health and Human Services, Assistant Secretary for Planning and Evaluation, *Setting the Baseline: A Report on State Welfare Waivers*

Notes: Nineteen of the 23 states requested family cap waivers. Idaho, North Dakota, Oklahoma, and Wyoming implemented family caps as part of their TANF programs.

Appendix Table 2
Caretaker Work Exemption Policies,
Date of Implementation by Age of Youngest Child

	No exemption	Date implemented (Waiver or TANF)	Up to and including 6 mos.	Date implemented (Waiver or TANF)	Over 6 months	Date implemented (Waiver or TANF)
Alabama					1 year	11/96 (T)
Alaska					1 yaer	7/97 (T)
Arizona					1 year	10/96 (T)
Arkansas			3 mos	7/97 (T)		
California			6 mos	1/98 (T)		
Colorado	county option	(T)				
Connecticut					1 year	10/96 (T)
Delaware			13 weeks	3/97 (T)		
D.C.			1 year	3/97 (T)		
Florida			3 mos	10/96 (T)		
Georgia	no ex.	1/97 (T)				
Hawaii			6 mos	2/97 (W)		
Idaho	no ex.	7/97 (T)				
Illinois					1 year	7/97 (T)
Indiana[1]					1 year	10/96 (T)
Iowa	no ex.	1/97 (T)	3 mos	10/93 (W)		
Kansas					1 year	10/96 (T)
Kentucky					1 year	10/96 (T)
Louisiana					1 year	1/97 (T)
Maine					1 year	11/96 (T)
Maryland			12 weeks	10/96 (W)	1 year	12/96 (T)
Massachusetts			6 mos	9/96 (T)		
Michigan	no ex.	10/94 (W)	3 mos	9/96 (T)		
Minnesota					1 year	9/97 (T)
Mississippi					1 year	9/97 (T)
Missouri					1 year	12/96 (T)
Montana	no ex.	2/97 (W)				
Nebraska			12 weeks 3 mos	3/96 (W) 7/97 (T)	1 year	12/96 (T)
Nevada					1 year	12/96 (T)
New Hampshire					3 years	(T)
New Jersey			12 weeks	7/97 (T)	2 years	10/92 (W)
New Mexico					1 year	7/97 (T)
New York					1 year	11/97 (T)

Appendix Table 2b (cont'd)
Caretaker Work Exemption Policies,
Age of Youngest Child and Date of Implementation

	No exemption	Date implemented (Waiver or TANF)	Up to and including 6 mos.	Date implemented (Waiver or TANF)	Over 6 months	Date implemented (Waiver or TANF)
North Carolina					5 years 1 year	7/96 (W) 1/97 (W)
North Dakota			3 mos	7/97 (T)		
Ohio					1 year	10/96 (T)
Oklahoma					1 year	10/96 (T)
Oregon			3 mos	2/93 (W)		
Pennsylvania					1 year	3/97 (T)
Rhode Island					1 year	5/97 (T)
South Carolina					1 year	10/96 (T)
South Dakota			12 weeks	12/96 (T)		
Tennessee			16 weeks	9/96 (W)		
Texas					4 years	?
Utah	no ex.	10/96 (T)				
Vermont			16 weeks	7/94 (W)	18 mos	9/96 (T)
Virginia					18 mos	10/97 (T)
Washington					1 year	1/97 (T)
West Virginia					1 year	1/97 (T)
Wisconsin			12 weeks	9/97 (T)	1 year	1/96 (W)
Wyoming			3 months	1/97 (T)		
Total						

Notes: Under TANF, 26 states exempt a mother while the youngest child is under 1 year of age; Vermont and Virginia allow an exemption up to 18 months; Texas is the only state to have a higher age limit, set at 4 years, but the exemption may only be used once for each family.
1. Indiana law only allows exemptions for care of a child under 12 weeks if child is conceived while family is on aid.
Sources: Crouse (1999) – note these are the same dates used in the 1999 CEA report; Health and Human Services, Assistant Secretary for Planning and Evaluation, *Setting the Baseline: A Report on State Welfare Waivers;* Urban Institute (1998) summary of state TANF policies.

Appendix Table 3
Welfare policy dates: AFDC waivers, TANF, and Time limits

	AFDC waiver		TANF implemented		Time limit implemented
	Implemented	Approved		Actual	
Alabama			12/96		12/96
Alaska			7/97		7/97
Arizona	11/95	5/95	11/95		11/95
Arkansas	7/94	4/94	-		-
California	12/92	10/92	-	1/98	-
Colorado	-	-	7/97		7/97
Connecticut	1/96	8/94	1/98		1/98
Delaware	10/95	5/95	7/97		7/97
District of Columbia	-	-	3/97		3/97
Florida	10/96	6/96	2/94		2/94
Georgia	1/94	11/93	1/97		1/97
Hawaii	2/97	6/94	2/97		2/97
Idaho	7/97	8/96	7/97		7/97
Illinois	11/93	11/93	2/96		2/96
Indiana	5/95	12/94	5/95		5/95
Iowa	10/93	8/93	10/93		10/93
Kansas	-	8/96	10/96		10/96
Kentucky	-	-	10/96		10/96
Louisiana	-	2/96	1/97		1/97
Maine	-	6/96	11/96		11/96
Maryland	3/96	8/95	1/97		1/97
Massachusetts	11/95	8/95	12/96		12/96
Michigan	10/92	8/92	--		--
Minnesota	-	-	7/97		7/97
Mississippi	10/95	9/95	10/96	7/97	10/96
Missouri	6/95	4/95	7/97		7/97
Montana	2/96	4/95	2/97		2/97
Nebraska	10/95	2/95	11/95		11/95
Nevada	-	-	12/96		12/96
New Hampshire	-	6/96	10/96		10/96
New Jersey	10/92	7/92	4/97	7/97	4/97
New Mexico	-	-	7/97		7/97
New York	-	-	12/96	11/97	12/96

Appendix Table 3 (cont'd)
Welfare policy dates: AFDC waivers, TANF, and Time limits

	AFDC waiver		TANF implemented		Time limit implemented
	Implemented	Approved	Official	Actual	
North Carolina	7/96	2/96	1/97		7/96
North Dakota	-	-	7/97		7/97
Ohio	7/96	3/96	10/96		10/97
Oklahoma	-	-	10/96		10/96
Oregon	2/93	7/92	10/96		7/96
Pennsylvania	-	-	3/97		3/97
Rhode Island	-	-	5/97		5/97
South Carolina	-	5/96	10/96		10/96
South Dakota	6/94	3/94	12/96		12/96
Tennessee	9/96	7/96	10/96		10/96
Texas	6/96	3/96	11/96		6/96
Utah	1/93	12/92	10/96		1/97
Vermont	7/94	4/93	9/96		--
Virginia	7/95	7/95	2/97		7/95
Washington	1/96	9/95	1/97		8/97
West Virginia	2/96	7/95	1/97		1/97
Wisconsin	1/96	6/94	9/96	9/97	10/96
Wyoming	-	-	1/97		1/97
Total		34			47

Sources: Crouse (1999) – note these are the same dates used in the 1999 CEA report; Health and Human Services, Assistant Secretary for Planning and Evaluation, *Setting the Baseline: A Report on State Welfare Waivers;* Urban Institute (1998) summary of state TANF policies.