Dreaming Big: Why Do People Play the Powerball?

Emily Oster [1]
Harvard University

Senior Honors Thesis, March 14, 2002

[1]I am grateful to Al Roth for advising and to the Multi-State Lottery Commission and the Connecticut Lottery for the data. Helpful comments and information were provided by Douglas Adams, Chris Avery, Charles Clotfelter, Ben Edelman, Ray Fair, Stephen Fair, John Friedman, Tricia Glynn, David Laibson, Gregor Matvos, Markus Mobius, John Oster, Sharon Oster, Jenna Robins, Gene Schaller, Michael Schwarz and Jesse Shapiro.

Abstract

The question of why individuals participate in gambling activities has been puzzling to economists at least since Friedman and Savage (1948). The increase in state-run lotteries in recent years has reopened this issue alongside the issue of whether lotteries exploit consumers. This paper addresses the question of whether lotteries are exploitive by connecting it to the question of why people play at all. I use two large new datasets on a lotto game (the Powerball lottery) to analyze lottery behavior in detail and provide insight into the validity of three existing theories about lottery play. This paper reveals four new facts about the lottery. First, sales increase more with increases in jackpot in richer areas. Second, sales increase more with increases in jackpot size when the odds of winning are better. Third, sales vary systematically over days of the week and time of the year. Fourth, people will purchase tickets for future drawings when they are available. The facts are most strongly consistent with the idea of an additive utility of lottery gambling. This result suggests that state-run lotteries are actually providing utility to consumers and concerns about exploitation are therefore less important.

Contents

1 Introduction

For thousands of years humanity has been fascinated by games of chance. One of the earliest Babylonian games was a dice game; in the New Testament of the Bible, Jesus' disciples cast lots to determine Judas's successor (Acts 1:26).[1] Lotteries were run routinely in colonial America to raise money. In fact, many of the earliest universities (including Harvard, Yale, and Princeton) were founded partially using money raised through lotteries.[2] In the mid-twentieth century legal lotteries were largely banned, but they were often replaced by illegal numbers games. In cities like New York these numbers games have been running for decades, often finding ingenious ways to operate.[3]

Despite the pervasiveness and obvious appeal of gambling, twentieth century governments in the United States have been squeamish about entering this market. It was only in the 1960s and 1970s that states reintroduced lotteries as a revenue source after a hiatus of more than 60 years in which state-run gambling was outlawed in most state constitutions (Clotfelter and Cook (1989)). Since then, despite some objections, more and more states have started to run their own lottery games. These games range from instant tickets (or "scratchers") to daily games that mimic the illegal numbers games to big jackpot lotteries like MassMillions or Powerball. Most states run multiple types of lotteries at any one time. The big jackpot games have received particular focus, with heavy media attention focused on the largest jackpots and extensive betting across all segments of society.

Economists have long been puzzled by the appeal of gambling to cash-maximizing consumers. Even more puzzling is the fact that the same consumers who purchase insurance (implying risk-aversion) also gamble (implying risk-loving). Early discussions of the gambling/insurance paradox include Friedman and Savage

[1]Michael Coogan, ed. *The New Oxford Annotated Bible* (Oxford: Oxford University Press, 2001), 187

[2]www.library.ca.gov/CRB/97/03/Chapt2.html

[3]As a prime example, in one Harlem newspaper the only financial statistic reported each day was the daily trading volume on the New York Stock Exchange. The last three digits of this number were the winning daily numbers for the illegal games

(1948), Markowitz (1952) and Kwang (1965). Although it is easy to rationalize the draw of gambling venues like casinos or horse racing by arguing that there is a non-monetary appeal, lottery betting seems more difficult to understand. In a purely economic sense, lotteries have an expected value lower than their cost at most times and purchasing a lottery ticket is in expected value terms a bad deal. In addition, playing the lottery is just purchasing a piece of paper; there is not the obvious recreation value that there is, for example, in watching horse races or playing slots in Atlantic City.

Independent of the academic questions of why people play the lottery, there has been significant popular discussion about state-run lotteries in the United States. People often suspect that lotteries exploit consumers, particulary those who are less well-educated. This phenomenon has been referred to in the media as "taxing stupidity."[4] This paper addresses the question of whether lotteries exploit consumers by testing three existing theories about why people play the lottery. I argue that the concern that the lottery is exploitive is founded in the impression that consumers play because they do not understand that the odds of winning are infinitesimal. I test this theory of incorrect probability processing against two theories in which lottery players are utility-maximizing without being cash-maximizing. I conclude that the data are most consistent with a theory in which the lottery actually creates utility for consumers. This result suggests that concerns about exploitation are less important.

Given that seemingly risk-averse agents (individuals being risk-averse according to most prevailing economic theory) continually defy expectations by purchasing lottery tickets, it is not surprising that there is a significant body of research about lotteries in economics as well as in sociology and psychology. In general this literature is separated into two questions: who plays the lottery, and why do they play?

To a large extent, the literature in economics has focused on the question of who plays the lottery. This is viewed as the relevant public policy question, in large part because lotteries are run by the state. The "takeout" rate on most lotteries is at

[4]http://www.cnn.com/CNN/bureaus/chicago/stories/9807/pball/TMP901741369.htm

least 50% (meaning 50 cents of each dollar spent goes to the prize and 50 cents goes to the state) and as a result there is an implicit tax on tickets. This problem is made worse by the states' monopoly on lottery sales, which allows them to have a higher takeout rate than might be determined competitively. For example, racetracks and casinos usually take less than 15% and the old illegal numbers games had a rate lower than 50%.

The money received by the state from lottery revenues is spent for the state as a whole (in most states the money goes either into the general state fund or toward education), so it is often argued that lotteries unfairly tax those who play more frequently and then use the money for the benefit of all. In addition, the overall effect of lotteries on state revenue is ambiguous. In states with no income tax but high sales and excise taxes, lotteries decrease other forms of revenue to the state (Borg, Mason and Shapiro (1993)). Further, in some states that use lottery revenues for educational funding there is evidence that this lottery revenue funding substitutes for general funding for education so the overall increase is small (Spinler (1995)).

A number of papers have used a variety of datasets to untangle the question of who plays the lottery. Clotfelter and Cook (1989) did some of the earliest work on this problem. Using individual data from Maryland they find that men play more often than women and that African-Americans play more often than whites. Education is negatively correlated with play and income has no nominal effect on amount spent. The income result means that poorer spend, on average, a larger percentage of their income on lottery tickets. In the literature this fact is referred to as the "regressivity" of the lottery. Other papers have largely replicated these results. Canadian lotteries are found to be less regressive than those in the United States (Kitchen and Powells (1991)), but Australian lotteries are more so (Worthington (2001)).

Brown et al. (1992) separate instant and lotto expenditures for consumers in Oregon and find that regressivity is much stronger for instant tickets, although increased education decreases purchases on all products. Instant tickets are those that you scratch off to find out the results instantly; lotto is a pari-mutual game in which you pick numbers and the winning jackpot is a function of the number of tickets purchased. Powerball is a type of lotto game. Additionally, Brown et al. (1992) find that while the poor (annual incomes less than $5000) spend the largest fraction of their income on lottery tickets, the majority of the funds coming into the lottery come from the middle income groups (annual incomes between $15,000 and $35,000). Priog-Good and Mikesell (1995) find using an Indiana lottery that as that lottery has aged it has become increasingly reliant on heavy bettors and regressivity has increased. Lotteries are somewhat substitutable with other forms of gambling. Siegel and Anders (2001) find that the introduction of casino gambling on Indian reserves somewhat decreases lottery sales in Arizona.

Although much of the literature on who play the lottery relies on individual survey data, it is also possible to address this question using aggregate data by area. Results from three Texas games suggest that all lottery purchases are regressive in income but find slightly different results on education. Controlling for income, a higher proportion of college-educated people in the area increases sales of lotto and daily numbers games, but decreases sales of instant tickets (Price and Novak (1999,2000)).

Several papers have recognized that there may be an econometric distinction between the determinants of choice to play the lottery and the determinants of how much to spend. Scott and Garen (1994) and Stranahan and Borg (1998a) both estimate a binary model on the choice play/no play and a continuous model on amount spent. Both papers find that African-American consumers are equally likely to play, but they have higher expenditures conditional on playing. Income seems to affect only the probability of play, not the amount of expenditure.

In general, this literature on lottery players suggests that the poor spend a larger fraction of their income on the lottery, even if the results do not unequivocally suggest that they spend a nominally larger amount. Further, more education is generally a negative predictor of play, providing support for the "tax on stupidity" hypothesis.

From a decision theory perspective, who plays the lottery is less important than why they play. This question has been approached by psychologists and sociologists, as well as by economists. Friedman and Savage (1948) suggest that individuals may have non-concave utilities over some range of outcomes; for example over prizes that move them between socio-economic classes. This could explain the appeal of long-odds, huge-prize lotteries like the Powerball.

Alternatively, early work in psychology shows that individuals tend to overweigh small probabilities, which could lead to lottery play (Preston and Baratta (1948), Griffith (1949)). More recently, Kahneman and Tversky (1979) incorporated this idea into prospect theory as a "probability weighting function." Thaler (1991) points out that behavior under this weighting function is consistent with lottery play. Camerer (2000) concurs, and notes that prospect theory can explain gambling behavior when expected utility theory fails.

Although economic theory primarily models the monetary aspects of the lottery, psychology and sociology suggest other reasons for its appeal. Adams (1997,2000) uses survey data from Arizona to explore people's reported motives for playing. He finds that while winning money is one reported reason, anticipation and fun also play a role. Further, he finds a significant portion of people have social reasons for playing – they can discuss the lottery with friends, coworkers, etc. Economics has not been immune to the suggestion that something other than monetary concerns motivate lottery play. Several papers in the economic literature have formalized the idea that adding a utility of gambling to the expected utility function would predict that risk-averse consumers will gamble (Fishburn (1980), Conlisk (1993), Le Menestrel (2001)).

This paper connects the public-policy question of whether the lottery is exploitive with these decision theory issues. I present and test three theories about lottery play using a new dataset. The first theory is prospect theory: individuals play the lottery because they do not understand that the odds of winning are very small. This theory leaves open the possibility that state lotteries are exploitive. This is not necessarily a guarantee that state lotteries should be discontinued; the benefit from money raised by state lotteries has to be weighed against the exploitation. However, the validation of this theory would provide fuel for the argument against state-run lotteries. The second theory is that people are risk-loving over some range of the utility function, causing them to gamble. The third theory is that people get some additive utility of gambling from playing the lottery. The last two theories imply that playing the lottery is utility-maximizing if not cash-maximizing; if this is true, it provides less support for concerns about exploitation.

In general, we note that the three theories tested here all have similar positive implications – namely that people will play state lotteries – but different normative implications. This type of analysis mirrors that of addiction models. Gruber and Köszegi (2001), for example, demonstrate that both a rational addiction model and a hyperbolic discounting model are consistent with cigarette consumption. They note, however, that while a rational addiction model suggests that (from a policy standpoint) there should be no extra taxes on cigarettes, a hyperbolic discount model suggests there should be high taxes. In the case of the lottery, a prospect theoretic model suggests that state lotteries should be reconsidered but the either of the other two models is generally supportive of state-run gambling.

This paper tests the three theories of lottery play by presenting an in-depth analysis of patterns of play using two new datasets. The data allow exploration of patterns of lottery play across zip codes or states and over time. Previous literature has generally been constrained to look either at cross-sectional predictors of play using demographic data or at the relationship between sales and prize using overall lottery data. This paper extends the analysis considerably by exploring, among other things, how the relationship between sales and prize differs across demographic characteristics.

This paper focuses on the Powerball lottery. The Powerball is the ultimate example of a big jackpot lottery. Over its history the jackpots have varied from $2 million to almost $300 million. Both datasets used provide information on the same game. The first dataset contains Powerball sales for each drawing, by state, since the inception of the Powerball in 1992. The second has two years of daily Powerball sales by zip code for the state of Connecticut.

This work differs from existing literature on the lottery both in the question it asks and the quality of data used. Previous work has not generally recognized the importance of the decision theory questions to the public policy issues in this context. In addition, the dataset used here is larger than what has been used in the past, and enables analyses that has not previously been possible.

I find that the theory in which individuals get an additive utility from gambling is most consistent with the data. The implication of this finding is positive for state-run lotteries; it suggests they may be providing a service as well as earning money for state coffers.

The rest of this paper is structured as follows: section 2 contains background information about the Powerball lottery and provides additional motivation for the issues identified in the introduction. Section 3 presents the data. Section 4 presents the three theories of why people play the lottery. Section 5 tests the implications of these theories for differences in sales across demographic characteristics, differences in sales when the odds of winning change, and differences in sales as people's feelings about their luck change. Section 6 concludes.

2 Background

2.1 Rules of the Game

The Powerball is a multi-state lotto game run by a consortium of 22 states.[5] Individuals in each state purchase tickets and there is a twice-weekly drawing of numbers. A winner can come from any of the participating states, and their prize money comes from the overall revenue. However, individual states keep a percentage of the money (about 50%) from the ticket sales in their own state

Powerball is a simple lotto game. Each ticket allows the purchaser to pick six

[5] The participating states are: Arizona, Connecticut, Washington DC, Delaware, Georgia, Iowa, Idaho, Indiana, Kansas, Kentucky, Louisiana, Minnesota, Missouri, Montana, Nebraska, New Hampshire, New Mexico, Oregon, Rhode Island, South Dakota, Wisconsin and West Virginia.

numbers. The first five numbers are chosen from a pool of 49 (without replacement), and the last – the "powerball" – from a separate pool of 42. The drawing is done by two machines that mix up the balls and then drop the appropriate number out a trap door. One machine dispenses 5 white balls and the other dispenses one red one. The chance of getting all six numbers correct is:

$$\left(\frac{5!44!}{49!}\right) \times \frac{1}{42} = \frac{1}{80,089,128}$$

The Powerball also offers smaller prizes for matching fewer than six numbers. Column 1 of figure 1 shows the complete prize schedule for the current game, which began in 1997. The black ball is the powerball. The smaller prizes are an important feature of the game. Some literature on gambling has suggested that having small prizes which people are more likely to win will keep them coming back to the game even if they never win the big one (Haruvy, Erev and Sonsino (2001)).

Prior to 1997, the game was structured somewhat differently. The basic idea was the same, but instead of choosing 5 from 49 and 1 from 42 the choice was 5 from 45 and then 1 from 45. The odds of winning in this structure are about 1 in 55 million. The decision to change the odds was made by the lottery primarily to increase the average size of the jackpots.[6]

The prizes for matching fewer than six balls were slightly smaller in the older form of the game. They are shown in the second column of figure 1.

It is possible (even likely) that for a given drawing, no one will win the jackpot: between the inception of the Powerball in April 1992 and the end of 2000 there were 908 drawings and only 114 winners. When no one wins, the jackpot is "rolled over" to the next drawing. The jackpot for any given drawing, therefore, is determined by the rollover from the previous drawing and the sales for the current drawing. The takeout rate for the Powerball is 50% so for every dollar of sales, the jackpot rises by \$0.50. For example, if the jackpot in the current drawing is \$40 million and no one wins, the next jackpot is \$40 million plus half of the sales.

[6]This information comes from personal correspondence with Gene Schaller at the Multi-State Lottery Corporation. The corporation is a non-profit that oversees the Powerball on behalf of the states.

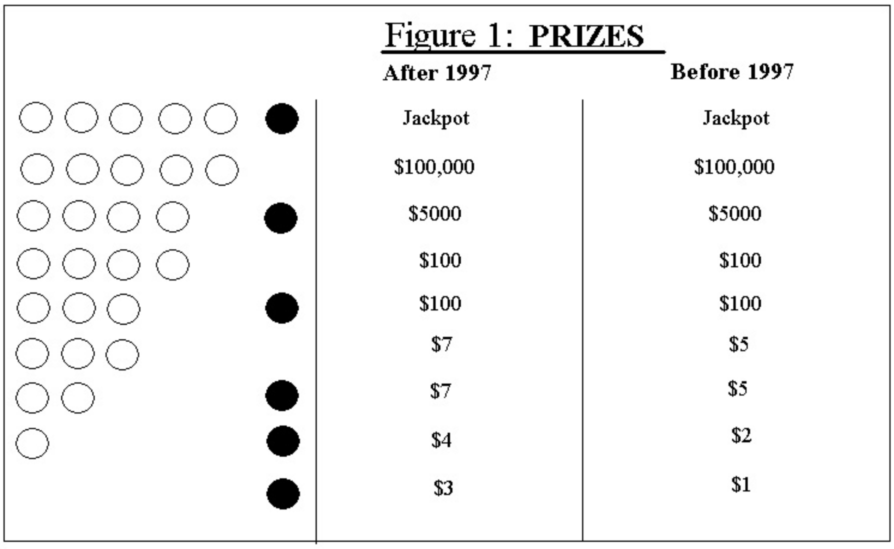


Figure 1: PRIZES

Theoretically, the increase in jackpot size is a direct function of the sales between drawings. In reality there are several deviations from this rule. First, at the lowest jackpots (Powerball currently starts each cycle with a jackpot of $10 million) the sales are not high enough to cover the prize. That is, the Powerball sells fewer than $20 million in tickets for the $10 million jackpot, so sales alone do not cover the prize. In addition, the lottery administration has committed itself to increasing the jackpot by at least $2 million between each drawing. It is not until a jackpot of $16 or $18 million that the sales fully cover the prize. If someone wins the jackpot before this level then the Powerball still pays the full advertised prize (there is a $7.5 million reserve held by the lottery corporation that is used to make up any shortfall).[7]

Powerball drawings are held on Wednesday and Saturday evenings. Shortly after the drawing the new jackpot for the following drawing is announced. If a winning ticket has been sold, the new jackpot is $10 million and the cycle starts again. If no winning ticket is sold, the lottery corporation estimates the new jackpot based on the rollover and the expected sales, and announces that. Except in rare cases in which the jackpot is so large (above $150 million) that the corporation has difficulty estimating the expected sales, they do not change the advertised jackpot between the initial

[7] Information from Gene Schaller, personal correspondence

announcement and the drawing.

Once an individual wins the jackpot, the prize is paid out in one of two ways. Winners can take half of the value in cash, or a 25-year annuity. The lottery corporation advertises on their website that the annuity is a better deal financially, but 49 of 53 winners since 1998 have taken the cash (for information on what lottery winners choose to do with their windfall, see Imbens, Rubin, Sacerdote (1999)). In fact, in present discounted value terms, assuming a 5% nominal interest rate, the PDV of the annuity is slightly more than the cash option. For example, the cash option on the $12 million prize is $6.6 million and the PDV is $7.1 million.

It is possible to have multiple winners since more than one person can purchase the same numbers. In this case the jackpot is split evenly among the winners and they can each choose their preferred method of payment.

There have been several changes in the structure of the lottery over time. As noted above, the odds of winning changed in 1997 from approximately 1 in 55 million to 1 in 80 million. Additionally, there have been changes in the minimum jackpot size. When the lottery began in April 1992, the minimum jackpot was $2 million. In December 1994 the minimum was raised to $3 million; in July 1995 it was raised again to $5 million. Finally, when the odds were changed in November 1997 the minimum jackpot was also raised to $10 million.

The final point about lottery structure is that it is possible to purchase "futures" – tickets for drawings after the current one. Individual states set the rules for how far ahead tickets can be purchased, but most states allow ticket purchase 5 to 10 weeks in advance.

The Powerball lottery was conceived ten years ago as a way of increasing lotto sales in smaller states. In the 30 years since states began to reintroduce lotteries – the first reintroduction was in New Hampshire in 1964 – organizers had observed a drop off in lotto sales. Consumers seemed to only respond to very large jackpots in lotto games. Big states like California could be successful in this market because their consumer base was large enough that they could offer reasonable odds and still expect to have large jackpots. For small states like Connecticut, however, the size of the population made this impossible. The Powerball connected several of these small states to make it possible for them to offer very large jackpots and to collect the correspondingly large revenue.

It is interesting to note that the existence of the scale economies that encourage lotteries like the Powerball to operate already indicate that something odd is going on in this market. If consumers responded only to the expected value of a ticket then small jackpot lotteries with higher odds would be just as effective as large jackpot lotteries with smaller odds. The fact that they are not – that large jackpot, small odds lotteries dominate – is notable.

Although the details of the lottery are important, for the purposes of this paper there are a few major points to keep in mind. The Powerball is a large jackpot lottery that offers big prizes and a very small chance of winning. The odds of matching the winning number are constant and do not change with the size of the jackpot or the number of other tickets sold. Finally, if no one wins the jackpot at a given drawing the current jackpot is "rolled over" and any new sales in the next period are added to that. The rollover aspect of the lottery is what allows for the very large jackpot sizes.

2.2 August 2001

In the last week of August 2001 the Powerball jackpot hit $280 million. It was the second largest jackpot in the history of the Powerball; indeed, it was among the largest in the history of lotteries in the United States. Connecticut is one of the 22 states that participate in the Powerball lottery, but New York is not. As a result, gambling-minded New Yorkers often travel over the border into Connecticut to purchase their tickets. Whether you travel Interstate 95 or the Metro-North railroad line, the first exit or stop upon entrance to Connecticut is Greenwich, one of the state's wealthiest towns.

The increased traffic and long lines of New Yorkers waiting to purchase tickets finally forced the Connecticut lottery to bow to the demands of Greenwich residents and put a moratorium on Powerball sales in the town. This did little to discourage New Yorkers from purchasing tickets; they simply moved their purchases one more exit or one more stop on the train.

What could possibly motivate people to go to such extremes to purchase these tickets? After all, the chance of winning the jackpot is 1 in 80 million, a long shot even for people who believe they can increase their luck by rubbing the ticket with a rabbit's foot or sleeping with it under their pillow. The media asked many players what motivated their purchases. To a large extent, the motivation seemed to be the things about life that would change with the winnings. One customer from New York, thwarted in his effort to buy tickets in Greenwich, noted, "You also have to understand, it's such a large jackpot. Maybe if we hit the jackpot we can afford to live in Greenwich and complain like everyone else."[8] Another said, "I'm going to buy a Jaguar. I've always wanted a Jaguar."[9] Many people cite the age-old adage: you have to play to win.

The jackpot was won by four people, none of whom were from Connecticut or New York. In 1998, when the Powerball jackpot was its highest ever at $295 million, Greenwich experienced a similar influx of purchasers. No one from Connecticut or New York won in that drawing either. In fact, since the Powerball was conceived in 1992, only three people from Connecticut have ever won a jackpot, the last in 1997. And yet people keep coming back to purchase the tickets, and not just at the highest jackpots. People play at the minimum jackpot ($10 million), too. Why do individuals start playing the lottery? Why don't they learn that the lottery is a poor bet? It is my hope that exploring patterns of play will allow me to shed some light on these issues.

3 Data

This paper uses two original datasets. The first was obtained from the Multi-State Lottery Corporation (MSLC) which is the non-profit organization which

[8] New Haven Register, Wednesday August 22, 2001. pg. A1
[9] *ibid*

oversees the Powerball. This dataset covers the Powerball (from its inception in 1992 until the end of the year 2000) and it reports sales for each state for each drawing as well as the advertised (and actual) jackpot for that drawing. For a shorter period of time the same data are available by day.

These data were obtained from the MSLC in the spring of 2001. They were originally provided in two datasets. One contained sales for each state reported by date. The other contained information about the prize by date. The datasets were merged so it is possible to observe the effect of the jackpot size on sales. In this paper I use two forms of the data. For overall analysis I have summed the sales across states for a given drawing or day to get overall sales for the lottery for that day. For analysis across states I use individual state sales. Some descriptive statistics about these data are contained below in table 1. The top section contains general information about the data; the bottom contains some summary statistics about the states drawn from 1990 census data.

Table 1

Descriptive Statistics for Overall Powerball

number of states	22	
dates covered	Apr 22, 1992 to Dec 30, 2000	
prize range	$2 million to $250 million	
sales range (for a single drawing)	$2.24 million to $210.85 million	
State Summary Statistics		
	Mean	**Std. Deviation**
total state pop	2,733,414	1,762,423
median HH income(ten thousands)	27,945	4,998
percent college graduates	.081	.016

The second dataset was obtained directly from the Connecticut Lottery office in July 2001. This contains Powerball sales for each retailer in the state for each day between August 1999 and May 2001. This dataset was also in two parts. Sales were reported for each day, indexed by retailer number (a five digit alpha-numeric code). In a separate dataset the retailer numbers were connected to addresses for each retailer. As with the overall data, I merged these datasets so the sales can be referenced by address. Consistency was checked by comparing the sales reported by Connecticut for a given drawing with the sales reported for Connecticut by the overall lottery. The correlation between the two was 0.9999.

For the purposes of the analysis in this paper I have aggregated the Connecticut data by zip code. There are two reasons for this. First, I am interested in looking at differences across demographics and this is best done across zip codes, for which demographic information is readily available. Second, it is likely that people do not purchase tickets exclusively at the retailer closest to their home. It is therefore not sensible to do demographic analysis by census block, which would assume that people were purchasing at the closest retailer. Of course, it may also be the case that people do not always purchase tickets within their zip code. Unfortunately, I do not have individual data on purchases, so I assume that people generally buy tickets within their zip code.

Table 2 presents descriptive statistics for the Connecticut data. Again, the top of the table shows general facts and the bottom provides summary statistics for zip codes. The summary statistics are drawn from the 1990 census data.

The average population per zip code is about 13,600 people, 9095 of them over 25. Most zip codes are either 100% urban or 100% rural, so the 55.8% in urban areas indicates that slightly less than half the state is rural. Urban areas have on average many more people than rural areas (22,515 average for areas that are all urban, 3596 for areas that are all rural).

Connecticut has a wide range of unemployment levels across zip codes – from 0% (Canton, Centerbrook) to almost 20% (Waterbury). The same is true for the percent of households on public assistance. Although the mean is only 4.3%, the range is from 0 to 50%.

Connecticut is a state with some very wealthy areas but also some areas of enormous poverty. The median household income is $45,933. The lowest average household income for a zip code is in the $9000 range; the highest is over $105,000 (the poorest zip code is once again in Waterbury; the richest is in Stamford, which serves in

Table 2

Descriptive Statistics for Connecticut

number of zip codes	302			
dates covered	August 1, 1999 to July 2, 2001			
prize range	$10 Million to $150 Million			
state sales range	$12,732 to $3.08 million			
Zip Code Summary Statistics				
	Mean	**Std. Deviation**	**Min**	**Max**
total population	13590.966	12758.361	83	60640
population 25+	9094.577	8545.264	46	40959
median household income	45933.901	13977.976	8942	105409
percent in urban areas	.558	.427	0	1
percent with college degree	.168	.075	.020	.411
percent African-American	.052	.115	0	.823
percent unemployed	.050	.030	0	.199
percent HH with public assistance	.043	.054	0	.455
average public assistance	4678.982	1797.343	0	10264
percent in poverty	.056	.067	0	.517

part as a bedroom community for New York City). The range of the variable measuring percent in poverty is 0 to 50%.

This result is mimicked by the percent of college graduates, which varies from 2% to 41% and the percent African-American, which varies from 0 to 82%. One concern about using data from Connecticut is the existence of New Haven, which has a large number of college graduates (because of the academic community) but is likely to be poor because graduate students and college students are “artificially” poor during their education. Although this is an important concern, we note that the Yale zip codes in New Haven are not outliers in either income level or in percent college graduates. In addition, when the regressions in table 3 below are run without the observations from New Haven the results are essentially identical.

All of these statistics tell a similar story about Connecticut. Connecticut is currently the second wealthiest state in the US, but the three major cities (New Haven, Bridgeport and Hartford) are dramatically poor. As a result it contains a lot of zip codes with poor, unemployed people on welfare and also a lot of zip codes filled with New York investment bankers and multi-million dollar houses. From this perspective, Connecticut is an ideal state for the analysis in this paper. I am particularly concerned with how demographic differences across zip codes speak to why people play the lottery. It is therefore extremely helpful to use a state with a wide variation in income and other demographic measures.

Throughout this paper the overall dataset from the MSLC and the Connecticut data will both be used extensively. To clear up potential confusion about which data are being used, each table defines the data used at the bottom.

As discussed in section 1, there is a large literature on the demographic correlates of lottery purchases. Because this paper uses an original dataset it is important to establish that while this dataset is new it is not inconsistent with previous findings about the lottery. In order to do this, table 3 below shows several simple regressions of Powerball sales on demographic information in Connecticut. The aim here is not to do a detailed analysis of the information but rather to demonstrate that the results using this data are not wildly different that those using other datasets.

The regressions in this table are clustered by zip code. Clustering will be used extensively throughout the paper. The dataset is in panel data format, which means for each zip code there is an observation for each day (this is why there are nearly 170,000 observations – 701 days and 302 zip codes). In the regressions in table 3, all of the explanatory variables are constant across days for a single zip code. For this reason it is necessary to control for correlation across days within zip code. If this is not done, the standard errors will be biased downward. Clustering by zip code adjusts the standard errors to correct them.

Table 3[a]

Demographic Correlates of Sales: Test for Reliability

Dependent Variable: Log of Per Capita Sales in Zip Code

Regression Type:	OLS	OLS	OLS
Explanatory Variables:			
total population[b]	-1.342^{***}	-1.401^{***}	-1.39^{***}
	(.316)	(.427)	(.356)
adult population[b]	1.932^{***}	2.00^{***}	2.014^{***}
	(.459)	(.626)	(.519)
median household income[c]	$-.891$	-1.355^{*}	$-.593$
	(1.471)	(.716)	(.751)
percent urban	$.505^{***}$	$.499^{***}$	$.56^{***}$
	(.161)	(.154)	(.149)
percent college graduates	$-.194$		
	(2.194)		
percent African-American	-1.281^{**}		
	(.519)		
percent unemployed		$-.077^{**}$	
		(.035)	
percent in poverty		.007	
		(.02)	
percent on public assistance			-1.899
			(1.39)
average public assistance income			-1.361^{***}
			(.452)
constant	5.813^{***}	6.297^{***}	6.257^{***}
	(.34)	(.385)	(.303)
Number of Observations	166474	166474	166474
Number of Zip Codes	302	302	302
R^2	.06	.06	.09

Data used: Connecticut Powerball; Daily August 1,1999 through July 2, 2001

* significant at 10%; ** significant at 5%; *** significant at 1%

[a] Standard errors are in parenthesis, adjusted for correlation within zip code.

[b] Population reported in ten thousands.

[c] Income in hundred thousands.

The demographic correlates of lottery sales here are roughly consistent with earlier work. The coefficient on median household income is negative in all regressions, although significant only in the second. The lottery is regressive as long as this coefficient is not significantly positive. Thus, I conclude that there is regressivity in this lottery, as has been widely demonstrated for other lotteries in the past. An increase in adult population increases lottery purchases, which is also not surprising.

People in urban areas purchase more tickets. The sign on percent with a college degree is negative, which is what we would expect, but it is not significant. This lack of significance may be due to heavy correlation with median household income and, to an extent, percent urban. The coefficient on percent African-American is negative and significant. This negative effect is consistent with other results. In particular, several papers (Scott and Garen (1993), Stranahan and Borg (1998), Oster (2001)) have found that African-Americans are less likely to play the lottery at all. We note also that an increase in unemployment decreases sales, as does an increase in public assistance income.

In general, one of the primary advantages of this dataset over others that have been used in the past is the ability to compare across zip codes. The question of why the lottery is regressive can be analyzed in significantly more detail if it is possible to break down the behavior of different demographics. The size and complexity of this dataset, as well as the simplicity of this lottery product, make this possible here where it has not been in the past.

4 Theory

This section discusses and elaborates on three theories about why people play the lottery that have been suggested elsewhere in the literature. The first theory is prospect theory, which suggests that individuals play because of a cognitive problem in processing probabilities. This theory is consistent with popular concerns about exploitation. The other two theories suggest that individuals are utility-maximizing without being cash-maximizing. These theories are less consistent, if at all, with the concern that lotteries exploit people.

In outlining these theories I keep in mind several empirical facts about lottery play that must be explained. First, people purchase at least one ticket, but not an infinite number. Any theory that explains why people play the lottery must be consistent with this. Second, the lottery is regressive. This is an empirical regularity in the work done on the lottery – rich people do not purchase more lottery tickets, despite being richer. A complete theory of the lottery would also explain this fact.

The data used in this paper suggests some empirical tests that have not previously been possible that may provide insight into the decision to play the lottery. The theories are therefore discussed with an eye to the testable implications. These fall into three general categories: implications for behavior across income groups, implications for behavior when the structure of the game is changed and implications for variation in sales over time.

Subsection 4.1 discusses prospect theory. Subsection 4.2 discusses a theory in which utility functions of wealth are convex over some range. Subsection 4.3 discusses a theory with an additive utility of gambling.

4.1 Prospect Theory

The basic idea behind prospect theory (Kahneman and Tversky (1979)) and the lottery is that people play because they do not correctly process the odds of winning. More specifically, this theory suggests that rather than acting under the true odds of winning, consumers act under their perceived odds, which are larger than the true odds for small probabilities and smaller than the true odds for large probabilities. The idea of incorrect probability processing is not new to this theory: Preston and Baratta (1948) and Griffith (1949) demonstrate the same phenomenon.

More formally, prospect theory suggests that people analyze the probability of a gamble and the value of the reward in a way that differs from expected utility theory. Specifically, they evaluate the true probability incorrectly, and instead act under a "probability weighting function." They value the reward based on a "value" function rather than a conventional utility function. Figure 2 below shows the value function; figure 3 shows the probability weighting function.

Figure 2: Value Function

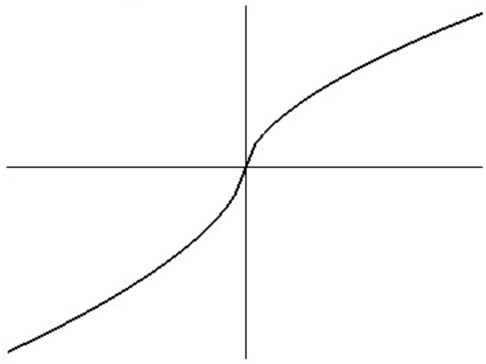
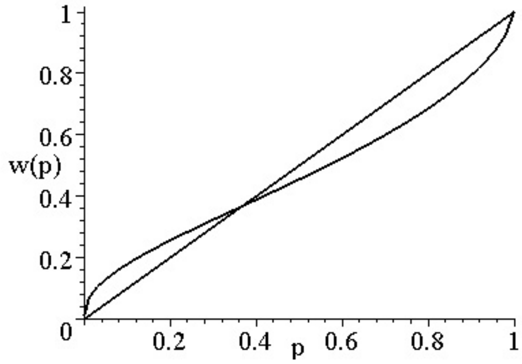


Figure 3: Probability Weighting Function



Unlike in expected utility theory, the value function (figure 2) suggests that gains or losses are measured in changes from the current level of income rather than in levels, so the origin of this function is current wealth. The value function shows that people are risk-averse in gains but risk-loving in losses. Figure 3 shows prospect theoretic weights on the curved line, and the true probability on the straight line. Probabilities less than about .4 are over-weighed; those above that are under-weighed.

Before discussing prospect theory and the Powerball, it is important to note that this theory originated as a description of behavior. As a result, we know that people do behave this way in laboratory situations and it is intuitively consistent with many real-life situations. The question here is not whether prospect theory is "true," but whether it is the best explanation for why people play state lotteries.

The prospect theoretic utility of n lottery tickets (assuming tickets cost \$1) is:

$$PTU_n = w(1-(1-p)^n)v(J) + w((1-p)^n)v(0) + PT_{small,n}$$

where $w(\cdot)$ is the probability weighting function, $v(\cdot)$ is the value function, p is the probability of winning the jackpot, J is the jackpot size and $PT_{small,n}$ is the prospect theoretic utility of the smaller prizes when n tickets are bought.[10] This utility is compared with the certainty equivalent of having the value of n tickets, or $v(n)$ when

[10]The utility of the smaller prizes is less interesting in this analysis because the smaller prizes do not vary with jackpot size. However, a full statement of the prospect theoretic utility, including smaller prizes is shown in appendix A.

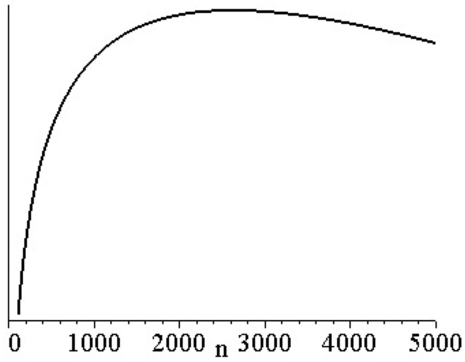
tickets cost \$1. Individuals solve the maximization problem:

$$\max_n [PTU_n - v(n)]$$

In order to discuss predictions of this theory it is useful to adopt an explicit formulation for $v(\cdot)$ and $w(\cdot)$. Here, I use the functional form suggested by Prelec (2000): $v^+(x) = x^{.6}$, $v^-(x) = -1.2(-x)^{.6}$, $w(p) = e^{-(-ln(p))^{.65}}$, where $v^+(\cdot)$ is the value function for gains and $v^-(\cdot)$ is the value function for losses. The difference between $v^+(\cdot)$ and $v^-(\cdot)$ define the convexity of gains and the concavity of losses. In the analysis here we deal only with the gain domain, so $v^-(\cdot)$ is not used.[11]

The first question is whether the theory with this functional form predicts that an individual will purchase at least one ticket; the second is whether their purchases will be bounded. The prospect theoretic utility of a single ticket (at the lowest jackpot of \$10 million) is 27.05, which is compared to $v(1) = 1$. This suggests the individual will purchase at least one ticket, even at the lowest jackpot level. To find the predicted number of tickets purchased, it is necessary to solve the maximization problem. Figure 4 below graphs $PTU_n - v(n)$ against n for a jackpot of \$10 million.

Figure 4: Utility of n Tickets, $10mill Prize



It is clear from this graph that the solution to the maximization problem is around 2500 tickets. This suggests a potential problem with this theory in that, both intuitively and quantitatively, it is not clear why ticket purchases are bounded from

[11]Although the analysis of only one functional form is shown here, a similar analysis was done with the functional form used in Kahneman and Tversky(1992) in which $w(p) = \frac{p^{.61}}{(p^{.61}+(1-p)^{.61})^{\frac{1}{.61}}}$ and $v^+(x) = x^{.88}$, $v^-(x) = -2.25(-x)^{.88}$. The results using this functional form are not qualitatively different.

above at a reasonable level. One solution is to imagine, perhaps realistically, that people are acting under a rule of thumb in which they purchase only up to a certain percentage of their income in lottery tickets.

Having established this theory we move now to a discussion of predictions. The predictions about income are generally ambiguous. In its original form, prospect theory makes no distinction between people in terms of their degree of probability weighting. Most of the experiments were done on college students, which suggests that these cognitive errors are not limited to groups with less education. Preston and Baratta (1948) used professors of math, statistics and psychology in their experiments on probability processing and found similar results to those from their "unsophisticated" experimental subjects. This evidence suggests that there would be little difference across income and education groups.

It is, however, possible to imagine a version of this theory in which income affects behavior. On one hand, if we imagine that people have a rule of thumb telling them not to purchase more than a small percent of their income in tickets, then the rich will purchase more tickets because their incomes are higher. On the other hand there is some evidence that the probability weighting function *does* vary across individuals by income. Donkers, Melenberg and Van Soest (2001) estimate parameters for the probability weighting function and find that, among other things, richer people have less transformed probabilities. If that is true, then we expect the rich to purchase fewer tickets.

Given that we know that as a percentage of income the rich purchase fewer tickets, this second strand of analysis deserves some elaboration. Specifically, we have demonstrated that once individuals enter the market for tickets their ticket purchases are bounded not by the prospect theoretic functional form but by some other assumptions (such as the rule of thumb). For this reason, if richer people purchase less it must be because they do not play the lottery at all jackpots. This implies that it must be the case that for some people the prospect theoretic utility at a jackpot of \$10 million is less than $v(1)$. Mathematically, we can use the functional form used by

Donkers, Melenberg and Van Soest (2001) to solve for the parameters of the probability weighting function that would produce this prediction. They use $v(x) = x^{a_i}$, $w(p) = e^{-(-ln(p))^{b_i}}$. Their paper tests for variation in a_i and b_i. Using their average value of a_i for the value function (about .65) we can test for the value of b_i that would entice someone not to purchase a lottery ticket. We find that $b_i = .833$. It is worth noting that this is quite high, and that the highest value of b_i that Donkers, Melenberg and Van Soest (2001) estimate is $b_i = .52$.

In general, we note that it will be difficult to reject prospect theoretic behavior in the lottery based only on observations about income. When I discuss the detailed income results in section 5 I will discuss the additional assumptions about prospect theory that would produce theory consistent with the empirical results.

The second prediction deals with the odds of winning. When the odds of winning were decreased from 1 in 55 million to 1 in 80 million (a decrease of 31%) in 1997, the weighted probabilities decreased from .00149 to .00137(a decrease of about 8%). This means that the prospect theoretic utility decreases by 8% at any given jackpot when the odds change. The predictions about changes in sales are, again, ambiguous. Under either the better odds or the worse odds the "utility-maximizing" number of tickets is in the thousands, as discussed above. Given this, we assume in either case that ticket purchases must be limited by a rule of thumb or common sense. It is not clear, however, how this rule of thumb changes, if at all, when the odds change. As a result we can make only some general predictions about the effect of a change in odds in this theoretical context. First, we expect a small decrease in sales when odds decrease; certainly smaller than the actual 31% decrease in probability. We also expect this decrease to be a constant percentage across all jackpot levels.

Finally, prospect theory does not make any predictions about changes in sales over time or across individual feelings. If people are purchasing tickets because they think they are going to win, they should be willing to play the lottery every time the chance is offered. As will also be true in section 4.2, both prospect theory and convex utility functions assume that a lottery ticket is a two-attribute good, and they suggest ways that the two attributes may vary from what conventional economic theory assumes. This suggests that events that do not affect the two attributes – jackpot and odds of winning – will not affect sales.

From a public policy perspective, we reiterate here that evidence supporting a prospect theoretic explanation for lottery play is not a good sign for state lotteries. The prospect theoretic explanation suggest that people believe they are cash-maximizing, but mathematically they are not.

4.2 Convex Utility Functions

One of the earliest discussions of the gambling/insurance paradox is the now-classic paper by Friedman and Savage (1948). In essence, this paper explains that people play lotteries and also purchase insurance because they have a "wiggly" utility function that is concave, then convex, then concave. They limit their analysis to low-income individuals, because those are the only people at the area of the utility function to which their analysis applies.

Markowitz (1952) elaborated on their theory and suggested from some early experimental evidence that people think about changes in income rather than levels (as in prospect theory), and therefore he extends the Friedman/Savage framework to medium and high-income individuals. The basic idea behind both Markowitz and Friedman/Savage is the same: people play the lottery because there is an area over which they are risk-loving and will therefore take fair (or slightly unfair) gambles. Although this theory originates in older literature, more recent work continues to consider Friedman/Savage utilities as one possible explanation for why people gamble (Wu (1979), Garret and Sobel (1999)). From a public policy perspective, this theory suggests that people are utility-maximizers without being cash-maximizers, which would be a significantly less negative signal about state lotteries.

Formally, I analyze here the implications of a Markowitz utility framework. Figure 5 below shows the Markowitz utility function; figure 6 shows only the positive quadrant, which is what is relevant for the analysis of gambling.

Figure 5: Total Utility Function

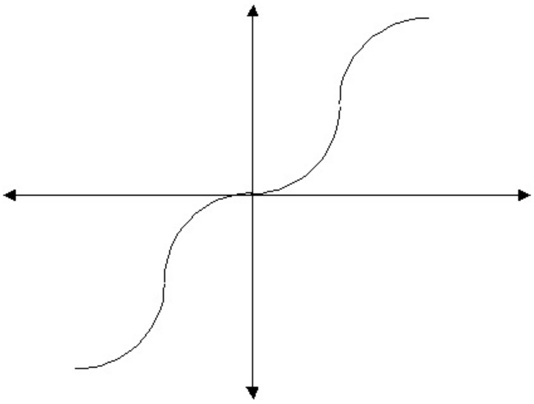
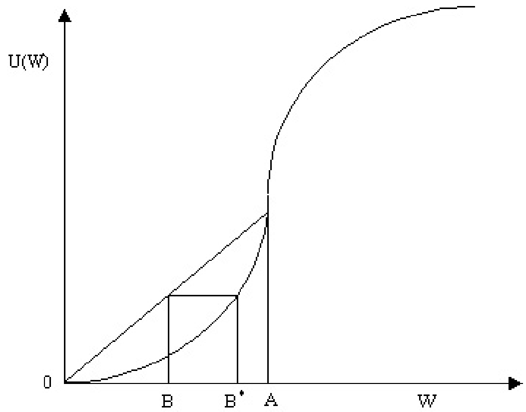


Figure 6: Positive Quadrant



The overall picture begins with a convex segment, followed by a concave segment, then convex and then concave. Markowitz refers to the origin as "customary wealth," which is equivalent to the statement that people consider gambles in terms of changes in their wealth rather than levels. This overall curve suggests that people are risk loving over large losses (in his example, people prefer a one in ten chance of losing \$1,000,000 rather than a certainty of losing \$100,000). They are risk-averse over small losses, risk-loving over small gains and risk-averse over large gains.

Figure 6 is simply the positive quadrant of figure 5. It is simple to see why this would explain lottery expenditures. Consider the chord drawn between A and the origin. The shape of the utility curve implies that people would prefer, for example, a 50-50 gamble between A and 0 rather than a certainty of B, halfway between the two. Indeed, they would be willing to pay $B^* - B$ for the privilege of gambling.

This model explains the puzzle of why people purchase some tickets but not an infinite number. Depending on their degree of risk-loving, people will pay some amount for the ability to gamble – that is they will purchase a lottery ticket for a small amount of money. However, people are risk-averse over large gambles (shown by the concave segment in the upper part of the positive quadrant), so they will purchase only a bounded number of tickets.

Markowitz (1952) makes specific suggestions about how the utility function might differ across income classes. He suggests that the inflection points (both in the positive and negative quadrants) would be further away from the origin for richer people. That is, the rich would be risk-loving over a larger set of values, presumably because they have more resources.

With respect to the lottery, this suggests that richer people would be willing to gamble more. As a result, this model predicts that ticket purchases should increase with income. This is inconsistent with the regressivity that is observed in most studies of the lottery; one possible modification would be to reject the difference in utility curves across income levels and assume everyone behaves as if they have the same curve. In the results section we test whether this is consistent with more detailed facts about income.

In this model, when the odds of winning decreased by 31% the expected utility also decreases by 31%. Roughly, then, we can we expect the sales to decrease by a similar percentage. The shape of the utility function makes people willing to pay for the privilege of gambling. When the odds are better they are willing to pay more; specifically 31% more. Like the prospect theoretic model, this model suggests that there should be a constant percentage decrease in sales when the odds change. It is distinct from the prospect theoretic model in that it suggests that the decrease will be one-for-one with the actual decrease.

Similar to the prospect theory model in section 4.1, this model does not predict that sales should change over time. Once again here the lottery is a two-attribute good and only changes in the jackpot or odds of winning should change sales.

4.3 Additive Utility of Gambling

The third theory in this paper explores the concept of an additive utility of gambling. The basic idea is that the utility from a lottery ticket is the conventional expected utility plus some additional utility from playing the game. This theory has been suggested in psychology (Adams (1997,2000)), as well as in economics (Fishburn (1980), Conlisk (1993), Le Menestrel (2001)).

In this model, individuals maximize utility over ticket purchases (n), where:

$$U = \left[\begin{array}{l} u(X) \text{ if } n = 0 \\ (1-(1-p)^n)u(J+X-n) + (1-p)^n u(X-n) + EU_{s,n} + V(J,X,\delta) \text{ if } n > 0 \end{array} \right.$$

where n is the number of tickets, $u(\cdot)$ is a normal concave utility function, J is the jackpot, X is the individual's wealth, p is the probability of winning, $EU_{s,n}$ is the expected utility of the smaller prizes when n tickets are bought and δ is a free parameter capturing heterogeneity in feelings about the lottery.

In essence, these equations indicate that the utility of having no tickets is simply the utility of current wealth, while the utility of purchasing n tickets is the expected utility of the tickets plus the fun of playing the lottery (the $V(\cdot)$ function). We note first that the solution to this maximization problem is always either 0 or 1 tickets. Since the expected utility of a ticket is negative, we know that $\frac{\partial EU}{\partial n} < 0$. However, since $V(\cdot)$ is not dependent on n, the overall partial is also negative. Therefore, purchasing two tickets will never dominate purchasing one ticket, so the choice of n will either be 0 or 1. Obviously, this theory explains why people purchase one ticket, and also why their purchases are bounded. In this simplest form the theory does not explain why people may purchase two or three tickets. However, it would be a trivial extension to build this into the model. For example, we could imagine that fun is a rapidly decreasing function of the number of tickets, so the first ticket gives a lot of fun, the second less and so on. Formally, this would suggest the following expression for utility:

$$U = \left[\begin{array}{l} u(X) \text{ if } n = 0 \\ (1-(1-p)^n)u(J+X-n) + (1-p)^n u(X-n) + EU_{s,n} + \sum_{j=1}^{n} \frac{1}{j} V(J,X,\delta) \text{ if } n > 0 \end{array} \right.$$

At a certain small level of ticket purchases the added fun will not be enough to motivate more buying. This version of the model is identical to the previous version except that it allows for multiple ticket purchases and suggests that any given individual may purchase more tickets when the jackpot is higher.

Before moving on to predictions made by this theory it is necessary to discuss the structure of the additive part of the utility function. $V(\cdot)$ is dependent on jackpot size, current wealth and δ. $V(\cdot)$ is increasing in jackpot size, meaning that the fun of playing is larger when the jackpot is bigger. This is a logical assumption that is made by both Conlisk (1993) and Le Menestrel (2001).

In addition, we assume that $V(\cdot)$ is decreasing in wealth. At a constant jackpot size and δ, richer people get less additive utility from playing the lottery. The argument for this is based on relative valuations of income and the hypothesis that much of the fun of gambling comes from the ability to "dream" about what you could do if you won. There is significant evidence that the ability to dream is a large motivating factor in playing the lottery. In an Australian survey, for example, 59% of the respondents said that the dream of winning was the motivating factor in play (Productivity Commission (1999)). Adams (1997) found that of 135 lottery players who reported that playing the lottery was "fun," 33.8% said the fun stemmed from the dream or anticipation of winning.

Intuitively, a jackpot of $10 million will provide significantly more dreaming for someone with wealth of $1000 from inner city New Haven than for someone with wealth of $10 million from Greenwich. There is evidence to support the idea that people measure happiness from additional income relative to their current level of income (Easterlin (2001)). In addition, there is evidence that people measure their income relative to the income of those around them (Easterlin (1995), Clark and Oswald (1996)). This suggests that people who live in richer areas would get less benefit from winning than those in poorer ones.

Finally, we assume that $V(\cdot)$ varies in δ. δ is a free parameter in the model that allows for heterogenous feelings about the lottery across time. Essentially, this parameter allows for people to "feel lucky" on a given day; or purchase a ticket as a last-ditch effort to get out of debt. It has been well-documented in the lottery that people feel they can affect their luck by picking particular numbers (Clotfelter and Cook (1989,1991b)). In addition, there is evidence that people's anticipation of events increases when the vividness of the event is greater (Lowenstein (1987)). Taken together, this suggests that lottery consumers may not behave the same way on all days; their purchases may be affected by their feelings about luck, the strength of their feelings about particular numbers, etc. In section 5.3 of the results I consider the effect of things that might change δ for the whole population.

I turn now to the predictions made by this theory about wealth and odds of winning. I have already mentioned that this theory may allow for changes in sales that are not dependent on jackpot or odds of winning by allowing for changes in δ, which is not the case in either prospect theory or the theory of convex utility functions.

The predictions about behavior across wealth levels in this model follow directly from the observation that individuals enter the lottery at different jackpot levels. People purchase 0 tickets until the sum of the expected utility and fun is high enough that they enter the game. The level of entry will be higher for richer people because $V(\cdot)$ is decreasing in X. This predicts a specific pattern of play across zip codes by income: in poor zip codes more people should play at the lowest jackpot levels, and there should be a smaller increase in sales from the lowest to the highest jackpot than there is in the richest zip codes. This is because as the jackpot increases there is little entry in the poorest zip code (they are mostly playing already) but a lot of entry from the richest people, who enter more slowly. The per capita sales at the highest jackpots, however, should be roughly the same because all individuals who are going to play have entered the game by this point.

The predictions about differences in sales when the odds of winning change in this model are also dependent on the differences in times of entry. When the odds decrease from 1 in 55 million to 1 in 80 million, there is no change in the $V(\cdot)$ function, but the expected utility decreases. The effect that this has on behavior is to change the jackpot at which an individual start playing; under the worse odds people will enter the game at higher jackpot levels. In addition, the number of new consumers entering the game for each $1 million increase in prize is larger under the better odds (consider: those who enter between $40 and $50 million in the better odds enter between $52 and $65 million in the worse odds. There are therefore more entries for each $1 million increase in prize under the better odds). This means that the slope of the relationship between sales and prize may be larger under the better odds; the result is that it may be the case that the percentage increase in sales when the odds change is larger under the better odds than the worse odds. This differs from the results of either convex utility theory or prospect theory.

The key features of this model are simple. The first major feature is the prediction that people enter the lottery at different jackpots. As discussed, this differs from either convex utility theory or prospect theory. The difference in entry jackpots produces the specific predictions about effect of income and odds. The second major feature of the model is the existence of an additive utility feature which makes participation in the lottery fun in and of itself and allows this fun to vary over people and over time.

From a public policy perspective this theory, like the theory with convex utility functions, suggests that people are utility-maximizing without being cash-maximizing. In this case, the theory actually suggests that state-run lotteries create utility for consumers, which would encourage rather than malign their existence.

5 Results

This section tests predictions made by the three models in section 4. Section 5.1 analyzes the overall sales response to jackpot size in the lottery. This will establish a functional form for the relationship between sales and prize. Section 5.2 discusses income effects; section 5.3 discusses effects of changes in the odds of winning; section 5.4 discusses systematic differences in sales across time, and potential differences due to "feelings" about the lottery. Finally, section 5.5 discusses the phenomenon of futures in the lottery (the ability to buy ticket for later drawings) and the implications. The chart below summarizes the predictions of the three theories and their public policy consequences.

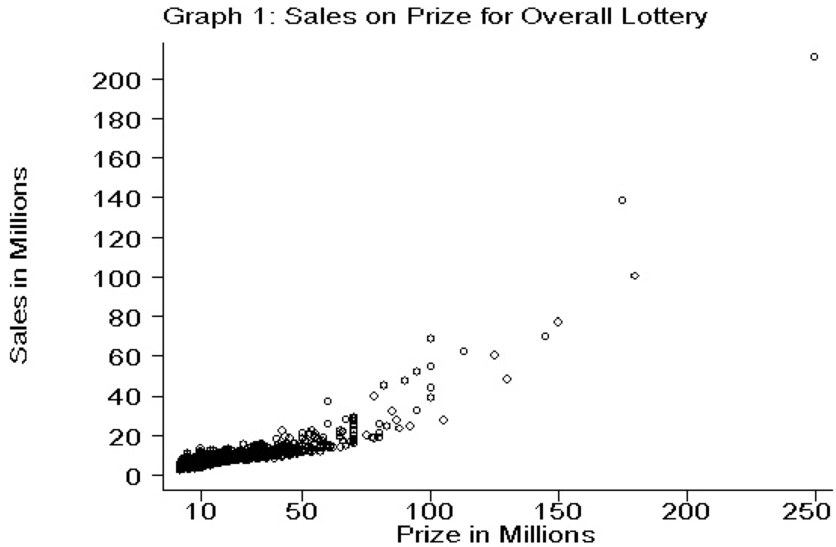
	Prospect Theory	Convex Utility	Additive Utility
Income	Results are ambiguous: in the simplest version, everyone purchases the same number of tickets. If we assume the weighting function varies with income then the rich purchase fewer tickets; if we assume that purchases are bounded by a percentage of income, the rich purchase more.	Richer people purchase more tickets or, in a slightly modified version of Markowitz (1952), all people purchase the same amount.	The jackpot level at which consumers first purchase tickets varies across income levels. Richer people enter the lottery only at higher jackpot levels. As a result in richer zip codes sales should change more when prize increases.
Odds	A decrease in odds leads to a less than one-for-one decrease in sales. This decrease should be a constant percentage across all jackpot levels.	A 31% decrease in odds leads to a 31% decrease in sales.	A decrease in odds causes everyone to change their jackpot of entry. We expect therefore that a decrease in odds causes a larger percentage decrease in sales at higher jackpots.
Feelings	Sales depend only on jackpot and odds.	Sales depend only on jackpot and odds.	Sales may change over time due to factors unrelated to either prizes or odds.
Public Policy	State lotteries take advantage of people who are making a cognitive mistake.	State lotteries are only recognizing that people are risk-loving, and people are expected utility maximizing without being cash-maximizing.	State lotteries are actually creating additive utility for consumers through "fun" utility.

5.1 Sales Response to Prize

Past papers have considered the sales response to prize magnitude in lotteries. This is generally viewed as an "elasticity of demand" for tickets (increases in jackpot size are decreases in price). These papers have attempted to use the rollover phenomenon discussed in section 2 to estimate an elasticity. There are a number of problems with these analyses, most of which are driven by a lack of data. Farrell and Walker (1999) use two instances of rollovers in an English lottery to estimate the demand response to such events. It is difficult to calculate an elasticity with only three data points and, additionally, in a lottery with so few rollovers there is likely to be a large media response to a big rollover that will skew the results.

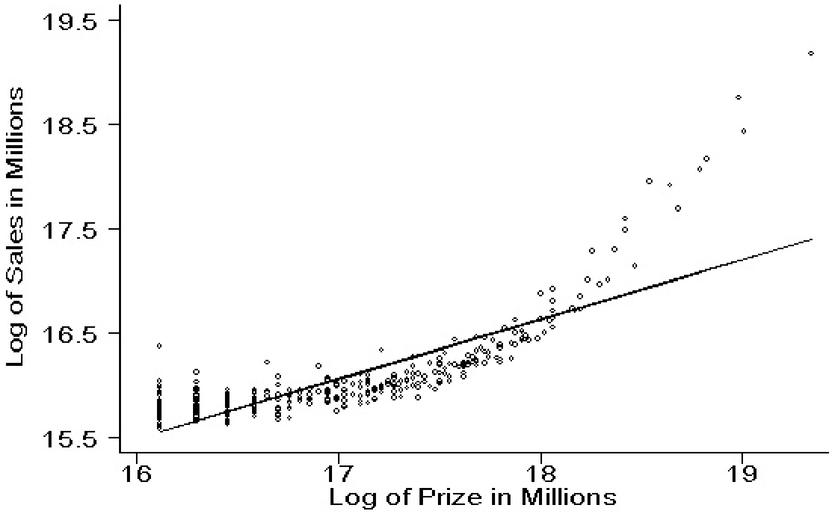
Mason et al (1997), Cook and Clotfelter (1993) and Bartsch and Paton (1999) all use lotteries with more data points to address the same question. However, the papers use different functional forms for the relationship, and none of them present any analysis of why that functional form was chosen. Mason et al. (1997) regresses log sales on log prize; Cook and Clotfelter (1993) and Bartsch and Paton (1999) regress lotto sales on jackpot size in a linear specification.

The dataset used here contains enough rollovers to test various relationships between sales and jackpot size. Graph 1 shows sales graphed on prize for the overall lottery – all states summed together.

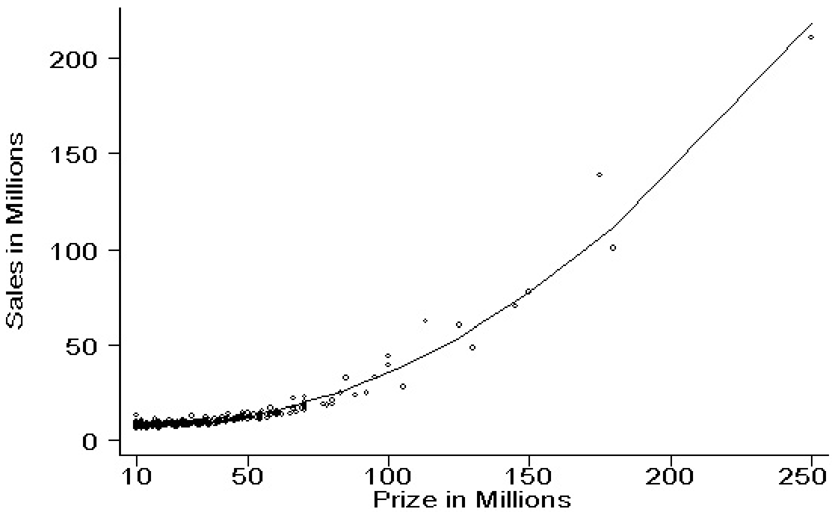


The relationship between sales and prize appears to be non-linear, particularly at the larger prize sizes. Graph 1 suggests that a linear regression of sales on prize will explain less variation than a regression with a functional form that allows for some curvature. There are several natural candidates for the relationship of sales to prize. Graphs 2 through 4 below show graphs of sales on prize for different functional forms. Graph 2 is a log-log relationship, graph 3 a quadratic specification and graph 4 is log-linear. In each graph the line shown is the fitted values. These graphs use the same data as graph 1 above.

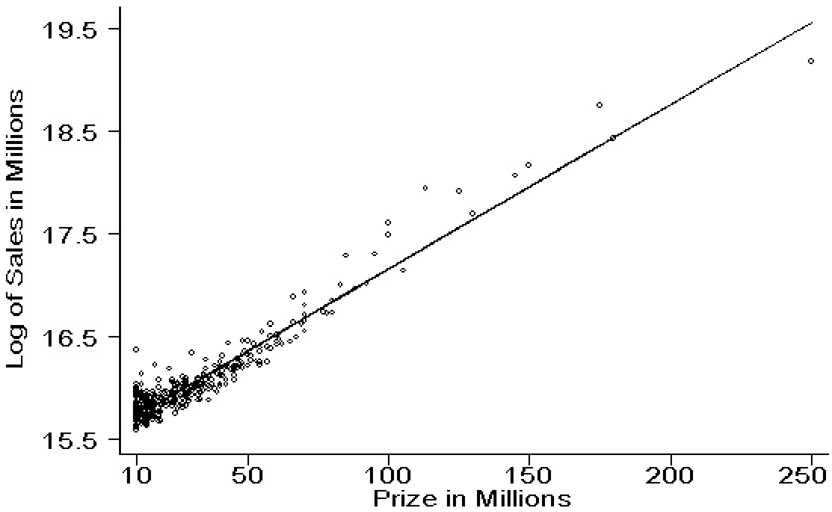
Graph 2: Log Sales on Log Prize



Graph 3: Quadratic Specification



Graph 4: Log Sales on Prize



Both the quadratic and the log-linear specification fit the data closely. One potential concern, however, is that the results may be driven by several outlying points. Table 4 below reports regression results for the relationships shown above (linear, log-log, quadratic and log-linear). The right hand side of the table limits the data to prizes below $150 million, which will eliminate the few very high outliers. The equations estimated (s is sales, p is prize) are:

$$\text{Linear: } s_i = \alpha + \beta_1 p_i \tag{1}$$

$$\text{Log-Log: } ln(s_i) = \alpha + \beta_1(ln(p_i)) \tag{2}$$

$$\text{Quadratic: } s_i = \alpha + \beta_1 p_i + \beta_2 p_i^2 \tag{3}$$

$$\text{Log-Linear: } ln(s_i) = \alpha + \beta_1 p_i \tag{4}$$

Table 4[a]

Regressions of Sales on Prize, Different Functional Forms

	All Observations				**Below $150 Million**			
Functional Form	*Linear*	*Log-Log*	*Quadratic*	*Log-Linear*	*Linear*	*Log-Log*	*Quadratic*	*Log-Linear*
Dependent Variable	Sales	Log Sales	Sales	Log Sales	Sales	Log Sales	Sales	Log Sales
Explanatory Variables:								
prize	.478***		−.109***	.016***	.273***		−.119***	.016***
	(.017)		(.012)	(.0003)	(.009)		(.014)	(.000)
log prize		.573***				.486***		
		(.022)				(.019)		
prize squared			.004***				.004***	
			(.000)				(.000)	
constant	−3.593***	6.317***	8.238***	15.558***	1.712***	7.770***	8.461***	15.556***
	(.714)	(.38)	(.318)	(.011)	(.353)	(.32)	(.278)	(.012)
Number of Observations	330	330	330	330	326	326	326	326
R^2	.71	.67	.97	.92	.72	.67	.93	.89

Data used: Overall Powerball; Drawings November 5, 1997 through December 30, 2000

* significant at 10%; ** significant at 5%; *** significant at 1%

[a] Standard Errors in parenthesis

Table 4 and graphs 1-4 suggest that the functional forms most consistent with this data are the log-linear or quadratic forms. In table 4, we can compare the R^2 for the log-linear with log-log, and for linear with quadratic. It is easy to see that the log-linear and quadratic explain much more of the variation in sales than the other two specifications. Although both log-linear and quadratic explain much of the variation in sales, the log-linear specification is less sensitive to outliers. Table 4 demonstrates that the relationship in the log-linear form between sales and prize is essentially identical for the regression using all observations and the regression limited to prizes below \$150 million. The coefficient on prize in the quadratic form, however, changes between the two regressions. For the rest of the analysis in this paper I therefore use the log-linear functional form.

The coefficient on prize in table 4 in the log-linear functional form is .016. Since prize is measured in millions of dollars, this suggests that a \$1 million increase in prize will produce a 1.6% increase in sales. As can also be seen from the graph, this means that the absolute increase in sales when the jackpot moves from \$100 to \$110 million is larger than when the jackpot moves from \$40 to \$50 million. Since the sales at \$100 million are larger than the sales at \$40 million, an increase of \$10 million in the first case yields an increase in sales of 16% over the initial level at \$100 million; this will be larger than the 16% increase over initial sales at \$40 million. Formally, $\frac{\partial sales}{\partial prize}$ is increasing in prize (because, with the parameters estimated, $\frac{\partial sales}{\partial prize} = .016e^{15.58+.016prize}$, which is increasing in prize). This is not an artifact of the choice of log-linear as the functional form. In the quadratic functional form, $\frac{\partial sales}{\partial prize} = -.109 + .008prize$, which is also increasing in prize.

There are two possible reasons that this could be true. One possibility is that the pool of buyers stays the same size at all jackpots, but they increase their purchases more when the jackpot is higher. Alternatively, it may be the case that increased sales are due to new individuals entering the lottery, and more new individuals enter at higher jackpot levels. In reality, it is likely that both are happening. Walker (1998) finds in data from the United Kingdom that at the lower jackpots 63% of individuals report playing and they spend on average £2.40. In contrast, when the jackpot gets very large in a "double rollover," 73% report playing and the average spending is £3.10.

In the context of the theories in section 4, the functional form of the relationship between sales and prize is only marginally informative. One important contrast between the theories is that prospect theory and, to an extent, convex utility theory suggests that people who play will play at all jackpots. In prospect theory, for example, the size of the parameters for the weighting function under which some people would not play at the lowest jackpot are larger than is generally assumed from laboratory experiments. Additive utility theory, on the other hand, assumes increased entry as the jackpot increases. Under the second version presented in section 4.3, additive utility theory also predicts increases in sales for each individual as the prize increases.

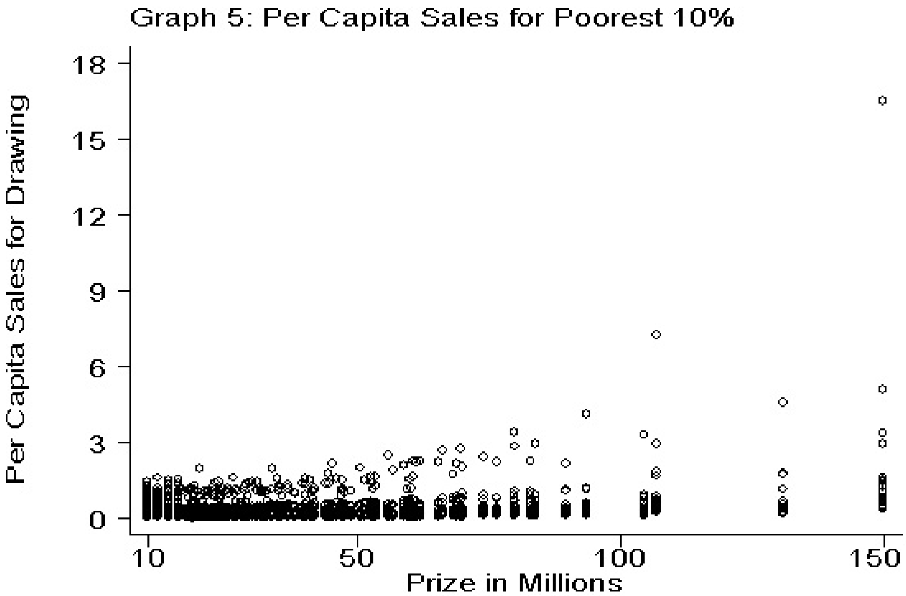
Were it possible to show conclusively that the shape of the relationship between sales and prize is due to entry by additional players as the jackpot increases, this would be support for additive utility theory over either of the other theories. In fact, it seems likely that some of the increase in sales is due to increases in the number of players but we have no conclusive proof. In general, the discussion of functional form is not able to rule out any of the theories as the potential explanation for lottery play.

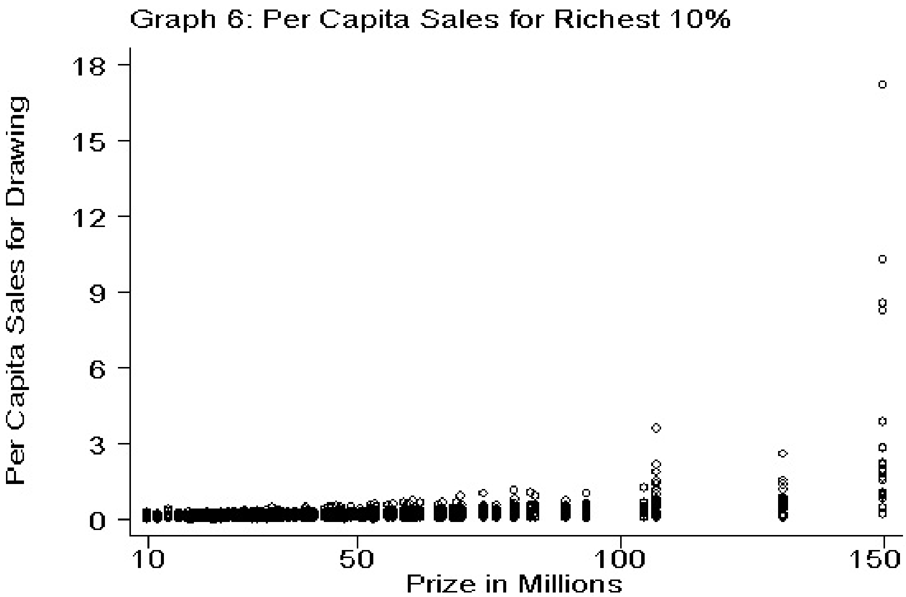
5.2 Income Effects

Disentangling lottery behavior across income levels has long been one of the primary concerns of the economic literature on this topic. A number of papers have established that the lottery is regressive – the poor spend a larger fraction of their income on lottery tickets – and a few have suggested that the poor spend absolutely more money on tickets. The literature has not, however, been able to go further in analyzing how patterns of play may differ across demographics.

The models in section 4 make different predictions about lottery sales in richer and poorer areas, as noted in the chart at the start of section 5. Convex utility theory suggests that there should be little difference across income, or that the rich should play more often. Prospect theory makes ambiguous predictions: the rich may play less if they are less susceptible to probability mistakes, or they may play more if people are acting under a rule of thumb telling them to purchase only a small percentage of their income in lottery tickets. Additive utility theory suggests that richer people enter the market for the first time only at higher jackpots.

Graph 5 below shows per capita sales (by drawing) plotted against jackpot size for the poorest 10% of zip codes in Connecticut. Graph 6 shows the same graph for the richest 10%.





It is clear from these graphs that the poorest zip codes purchase more tickets at the lowest jackpot levels. However, at the highest jackpots the sales are about the same (slightly over $16 per capita in the poorest zip codes and around $17 per capita in the richest). This evidence is consistent with additive utility theory: the elasticity of sales with respect to prize seems to be larger in the richer zip codes. The graphs are generally not consistent with the predictions of convex utility theory. In particular, to make this empirical fact consistent with convex utility theory it would be necessary to assume that the area over which people are risk-loving changes when the jackpot changes and, additionally, changes differently for richer and poorer people. We cannot reject this evidence as inconsistent with prospect theory if richer people have different probability weighting functions and enter therefore the lottery later. This would produce the pattern seen in these graphs; as discussed in section 4, however, there are some problems with that modification because it requires parameterizing the weighting function with values that are not generally consistent with what is observed in laboratory experiments.

Although the difference in sales response to prize across income levels is well-illustrated graphically for the extreme income levels, table 5 tests whether this is true across all levels of income by estimating a panel regression. Each observation is a zip code-day. That is, sales are indexed by both zip code and the day they are observed.

Column 1 estimates the regression of log sales on prize, median household income and median household income interacted with prize (holding several other factors constant). The coefficient of interest is that on the interaction between household income and prize size. The graphs above suggest that richer zip codes should have a larger sales response to prize, so the coefficient should be positive. The regression in column 1 is clustered by zip code. Columns 2 and 3 estimate elasticities for the top 10% of zip codes by income and the bottom 10%, respectively.

The positive sign on the interaction between median household income and jackpot size reflects the facts shown in the graphs – at higher income levels, increases in jackpot size produce larger increases in sales. The negative sign on median household income, although not significant, supports the idea that at the lowest jackpot levels richer communities have fewer sales. Using the estimated magnitudes of coefficients (and assuming for this test that the coefficient on median household income is significant) it is possible to estimate at what level income begins to have a positive effect on sales – that is, at what prize level the total coefficient on income moves from

Table 5[a]

Difference in Prize Elasticity Across Zip Codes by Income

Dependent Variable: Log of Per Capita Sales in Zip Code

Regression Type	OLS All Zip Codes	OLS Top 10% Income	OLS Bottom 10% Income
Explanatory Variables:			
prize	.012***	.018***	.015***
	(.001)	(.001)	(.001)
median HH income[b]	−.942	−.824	1.423
	(.603)	(3.14)	(3.23)
median HH income × prize[b]	.007***		
	(.002)		
total population in zip[c]	−1.721***	.925	−1.352
	(.331)	(6.402)	(.899)
adult population in zip[c]	2.485***	−1.724	1.586
	(.486)	(8.845)	(1.49)
urban percentage	.41***	.366	.886*
	(.155)	(.812)	(.458)
constant	5.341***	5.56***	4.577***
	(.286)	(1.556)	(.911)
Number of Observations	166474	16470	16818
R^2	.12	.10	.21

Data Used: Connecticut Powerball; Daily August 1, 1999 through July 2, 2001

* significant at 10%; ** significant at 5%; *** significant at 1%

[a] Standard errors in parenthesis, adjusted for correlation within zip code

[b] Income in hundred thousands

[c] Population in ten thousands

negative to positive. To do this we find the prize level Y at which $-.942 = .007Y$, which implies that Y=\$134 million.

It is also clear that the elasticity with respect to prize is larger in the richer zip codes. The difference is substantial – an increase of \$1 million in prize implies about a 1.5% increase in sales in the poorest zip codes, and about a 1.8% increase in the richest. The size of the standard errors mean that these differences are significant.

One concern here is the generalizability of these results. That is, we may be seeing a pattern that exists only in Connecticut. Using the overall dataset, however, it is possible to use cross-state variation in income to see if the phenomena observed in Connecticut is true across the lottery as a whole. In this case, we expect states with higher average income levels to have a greater elasticity. Table 6 below shows a regression similar to that in table 5, but using state rather than zip code income measures.

Table 6[a]

Difference in Prize Elasticity across State Income

Dependent Variable: Log of Per Capita State Sales

Regression Type:	OLS
Explanatory Variables:	
prize	.008***
	(.003)
median HH income[b]	.405*
	(.214)
median HH income × prize[b]	.003**
	(.001)
state population[c]	−.139***
	(.039)
percent college graduates	−11.09
	(7.562)
constant	−2.039
	(.452)
Number of Observations	6930
R^2	.65

Data used: Overall Powerball; Drawings November 5, 1997 through December 30, 2000

* significant at 10%; ** significant at 5%; *** significant at 1%

[a] Standard errors in parenthesis, adjusted for correlation within state

[b] Income in ten thousands

[c] State population in millions

Although the variation in income across states is small (from about $20,000 to about $40,000) the coefficient on the interaction between prize and median household income is still significant. This is consistent with the findings in Connecticut. Richer states, like richer zip codes, have a larger sales response to prize.

The results in this section significantly further understanding about the relationship of income and lottery purchases. The simple story of regressivity that has been told in the literature in the past leaves out important details about behavior in the lottery market. It is not the case that the lottery is only played by the poor, but it is true that richer people wait until larger jackpots to play.

In addition to providing details about demographic incidence of the lottery, this section has provided insight into three theories in section 4. The results here are inconsistent with the convex utility theory. That theory suggests that either everyone should purchase the same number of tickets, or the rich should purchase more. Clearly, the empirical pattern is both different and more complex than this prediction. The results potentially, although not unequivocally, consistent with prospect theory. In general, we noted that the prospect theoretic predictions about behavior across income levels were ambiguous. Richer people may be less susceptible to cognitive mistakes, but they may also purchase more tickets because of their higher income level. In short, the theory does not directly predict the results here, but there is a version of prospect theory that would not be inconsistent with the results. The additive utility theory is more clearly consistent with the data. Because this theory assumes that the fun of playing is decreasing with income, it suggests that richer people should enter the lottery later, which produces the prediction of differing elasticities, which is what the empirical evidence shows.

5.3 Lottery Structure

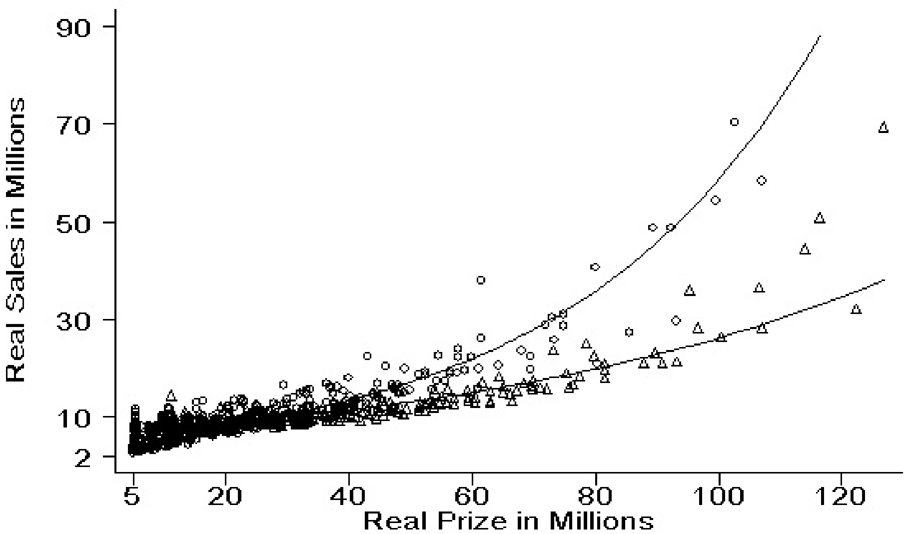
One of the primary advantages of the Powerball lottery from the point of view of this analysis is that the odds of winning the jackpot have changed over time. In particular, in 1997 the odds of the game were decreased by 31% and the minimum jackpot was raised by $5 million. Prior to the change the chance of winning the jackpot was approximately 1 in 55 million and the minimum jackpot was $5 million. Beginning in November 1997 the odds were decreased to 1 in 80 million and the minimum jackpot was raised to $10 million. Both changes were well-publicized by the lottery. Even without publicity, it would be difficult for lottery players not to notice the change since the range of numbers they were choosing from changed.

In this section I explore the effect of the change in odds on lottery sales. As mentioned, when the odds changed the minimum jackpot changed as well. In the rest of this section I do not control for minimum jackpot size. In theory, people should not care what the minimum is; in reality there seems to be little response. Appendix B shows the effect of earlier minimum jackpot changes on lottery sales.

All of the models in section 4 suggest that sales should change when the odds of winning change. They differ, however, in their predictions of magnitude. Specifically, the convex utility theory suggests that sales should decrease by about 31% when the odds change by 31%; prospect theory suggests that the decrease should be smaller than this, but still a constant percentage decrease across jackpot levels. Additive utility theory predicts that the decrease should be a larger percentage at higher jackpot levels – that is, that the relationship between sales and prize should have a larger slope when the odds are better.

The effect of the change in odds can be seen intuitively in graph 7 below. The graph shows the sales at each prize size for the two game structures. The graph excludes some very high jackpots that appeared only in the version of the game with worse odds so it is easier to see the relevant data points. In this case, "worse" odds is the current game (1 in 80 million chance of winning) and "better" odds refers to the older game (1 in 55 million chance of winning). The graph here shows real sales and real prize to adjust for inflation over the 8 years.

Graph 7: Real Sales on Real Prize
Triangles: Worse Odds; Circles: Better Odds



This graph demonstrates that sales at a given jackpot do seem to be higher at the better odds. In addition, the relationship between sales and prize at the better odds does seem to be slightly steeper than at the worse odds. This can be seen most clearly through the fitted lines in the graph. The line fitted to the earlier odds is steeper than that to the later (worse) odds.

This graph does not control for a variety of factors including the date, the population with access to the lottery and several macro variables. Table 7 below, therefore, explores the effect of odds using a regression format. Column 1 includes only a dummy for odds; column 2 includes both a dummy and the dummy interacted with the prize. The dummy for odds is 1 when the odds are better (1 in 55 million) and 0 when they are worse (1 in 80 million). We therefore expect a positive coefficient on this variable.

Both the convex utility theory and prospect theory from section 4 suggest that the interaction between odds and prize should not have a significant coefficient; the additive utility theory, however, suggests that the interaction will be positive. Both columns control for the time since the inception of the lottery (with the idea that sales may change systematically over time) and the monthly unemployment rate in the U.S. as a general measure of economic well-being. The regression uses the real values of per capita sales and prize (in 1992 dollars). The use of per capita sales implicitly controls for total population with access to the lottery.

Column 1 of this table shows that the odds change has an effect on sales. The coefficient on odds is .14 suggesting that an increase of 31% in the odds of winning leads to about a 15% increase in sales, controlling for other factors. This is a slightly smaller effect than is predicted by convex utility theory and slightly larger than predicted by prospect theory.

Column 2, however, demonstrates that the relationship between sales and prize is different at different odds. Specifically, under the better odds sales increase more with a marginal increase in prize. This coefficient is significant at 1%. This result is not consistent with either prospect theory or convex utility theory but is consistent with the additive utility theory as discussed.

Table 7[a]

Effect of Odds Change in Overall Lottery

Dependent Variable: Log of Per Capita Sales (in 1992 dollars)

Regression Type	OLS	OLS
Explanatory Variables:		
prize (in 1992 dollars)	.017***	.014***
	(.0002)	(.0002)
unemployment rate	−.348***	−.264***
	(.028)	(.023)
time[c]	−.037***	−.028***
	(.004)	(.003)
dummy for odds	.142***	−.124***
	(.026)	(.025)
dummy for odds × prize		.011***
		(.001)
constant	.21	−.25
	(.214)	(.178)
Number of Observations	898	898
R^2	.84	.89

Data Used: Overall Powerball; Drawings April 22,1992 through December 30, 2000

* significant at 10%; ** significant at 5%; *** significant at 1%

[a] Standard errors in parenthesis

[b] Time measured in hundreds of days

In general, this analysis suggests that changing the odds of winning does have a significant effect on sales. When the lottery corporation decreased the odds of winning by 31%, they faced a significant decrease in sales at the lower jackpots. The enormous response to very large jackpots, however, at least theoretically outweighed this enough to warrant the change in odds. This is a testimony to the degree to which large jackpots in this type of lottery capture the imagination so the increase in sales is disproportionate to the increase in expected value. In and of itself this is supportive of the additive utility theory: in either prospect theory or the convex utility theory people are responding to only the odds of winning and the amount you win. The fact that enormous jackpots seem to garner disproportionate sales may indicate that things other than these factors matter in lottery play. The next section explores this idea further by considering times during the year when people may be more likely to play and their effects on lottery sales.

5.4 Are You Feeling Lucky?

One of the major distinctions between the models in section 4 is that the additive utility model introduces a free parameter δ. At the most basic level this parameter suggests that sales will vary over time for reasons that are unrelated to the expected value of winning. This parameter may be understood as feelings about "luckiness" or, following the ideas of Lowenstein (1987) on anticipation, it may be the vividness of the dreams created by the possibility of winning.

Overall, this parameter suggests that a wide variety of non-jackpot, non-odds related events will affect lottery sales. This section tests the validity of any model of the lottery as a two-attribute good. Both prospect theory and convex utility theory assume that the lottery has two attributes – the odds of winning and the prize. In that sense, both theories consider the lottery ticket as a type of investment. Additive utility theory, on the other hand, assumes there is something else about the lottery that affects people's desire to play.

I find that lottery sales differ systematically in a number of ways that are discussed in more detail below. These fall roughly into three categories: deadline effects, calendar anomalies and time trends in sales.

5.4.1 Deadline Effects

One of the more interesting phenomenon about the Powerball is the existence of strong deadline effects in ticket sales. I have adopted this phrase to refer to the fact that as the drawing grows closer daily sales increase. The Powerball holds drawings on Wednesdays and Saturdays. Sales on Sunday are the lowest, Monday and Thursday are higher than Sunday; Tuesday and Friday are even higher and the highest sales are on Wednesday and Saturday. Table 8 below summarizes sales figures by day for the entire lottery (using only the current form of the game, since 1997, although the phenomenon is not limited to this form).

Table 8[a]

Sales by Days of Week

Day of Week	Mean
Sunday	756.8
Monday	1626.1
Tuesday	2651.7
Wednesday	6282.3
Thursday	1443.8
Friday	2767.5
Saturday	6267.0

Data used: Overall Powerball; Daily November 2, 1997 through December 30, 2000

[a] All sales figures reported in thousands

In general, we see that sales increase up to Wednesday, decrease on Thursday and increase to Saturday. Table 9 below controls for prize in examining this phenomenon. Column 1 controls only for days to the drawing (all dummies are compared to the baseline of Sunday, 3 days before the drawing). Column 2 controls for days to drawing and adds dummies for the end of the week drawing in order to compare the two weekly drawings.

This table confirms the evidence from the summary statistics. As the drawing grows closer the sales increase significantly. The magnitude of the effect is very large.[12] The sales on the day of the drawing are 2.18 log points or 784% higher than those three days before.

The drawing at the end of the week is more popular (controlling for size of the prize). In particular, there is a very large Friday effect (29% average increase in sales over Tuesday, compared to a small Thursday effect at 4% increase over Monday and an only modest Saturday effect at 12% increase over Wednesday). This is somewhat supportive of the idea that feelings about luckiness and well-being positively influence lottery play. Friday – the first day of the weekend, the end of the work-week – is likely to be a time when people would purchase tickets. They may think, for example, about how nice it would be if they won on Saturday and could forget about work on Monday.

[12] The t-statistics are also very large, between 30 and 120; this reflects the fact that this phenomenon is true in every drawing for which I have data except for drawings on Christmas, the fourth of July and New Years Day.

Table 9[a]

Deadline Effects and Day of Week Effects

Dependent Variable: Log of Per Capita Sales

Regression Type:	OLS	OLS
Explanatory Variables:		
prize	.017***	.017***
	(.000)	(.000)
two days before drawing	.674***	.634***
	(.018)	(.019)
one day before drawing	1.218***	1.088***
	(.018)	(.019)
day of drawing	2.18***	2.118***
	(.018)	(.019)
Thursday		.081***
		(.019)
Friday		.26***
		(.019)
Saturday		.125***
		(.019)
constant	12.742***	12.741***
	(.016)	(.014)
Number of Observations	1155	1155
R^2	.96	.96

Data used: Overall Powerball; Daily November 2, 1997 through December 30, 2000

* significant at 10%; ** significant at 5%; *** significant at 1%

[a] Standard errors are in parenthesis.

Another way to think about this is that the lottery may be part of people's weekend entertainment package.

Although the deadline effect is strong in these data, it is possible that this is not a deadline effect at all, but simply a normal pattern of gambling through the week. One way to test this is to examine sales of instant tickets across days of the week. Since the jackpot for instant tickets does not change, we would not expect to see any deadline effects here.

For a short period of time, instant ticket sales are available from the Connecticut lottery. Table 10 below shows the summary of sales for each day of the week. The pattern of instant ticket sales through the week does not mirror the Powerball pattern. On the contrary, sales are highest on Monday and lowest on

Saturday.

Table 10[a]

Instant Sales by Days of Week

Day of Week	**Mean**
Sunday	1399.1
Monday	1733.5
Tuesday	1482.9
Wednesday	1495.3
Thursday	1473.7
Friday	1483.6
Saturday	1172.8

Data used: Connecticut Instant Sales; Daily July 1, 1999 through October 10, 2001

[a] All sales figures reported in thousands

Although the pattern of deadline effects is strong and regular, it is not clear why it should exist in this context. Roth and Ockenfels (2000) find evidence of deadline effects in Ebay auctions – 50% of the sales happen in the last 10% of the auction time. Although it is from this work that the phrase "deadline effect" is drawn, the nearly linear increase in sales up to the day of the drawing is not consistent with their power law pattern.

In this case, unlike in the case of auctions, people are required to invest their money first to find out the outcome later. Thus, our baseline prediction is actually that people would want to keep their money as long as possible and everyone would purchase on the last day. Given this, we consider what factors could move people away from purchasing on the final day. One possibility is that people get a small amount of additional anticipation utility from having the actual ticket in their hand. Therefore they would like to purchase the ticket on the first day, but procrastination forces them to keep moving their ticket purchases back.[13] This explanation suggests both a reason why all ticket purchases would not happen on the last day, and a reason for the increase up to that point. An alternative explanation is based on transaction costs. People may move ticket purchases away from the last day because it is cheaper in terms of time to purchase tickets on a weekday rather than on Saturday. Further research

[13] I thank David Laibson for this suggestion.

with individual consumers is necessary to understand this phenomenon more fully.

5.4.2 Calendar Anomalies

One way to test for heterogeneity in lottery sales is to consider times during the year when people may feel particularly drawn to the lottery. An obvious example is a birthday or anniversary; of course, it is not possible to test that hypothesis using aggregate data. Instead, I consider times when many people may be influenced simultaneously. This includes times when people may feel particularly lucky and as well as times when they may be particularly in need of money. The theory behind this is that in both of these times the dream of winning may be more vivid because they are more specific. Lowenstein (1987) suggests that the increased vividness may produce heightened levels of anticipation.

This section suggests four times during the year when higher sales might be expected. The first, which was tested in the previous section, is Friday. As mentioned, people may enjoy beginning their weekend with the potential for a large win, and the end of the week may be a time they feel lucky. The second time is the last day of the month. We expect that this is likely to be a time when people have finished paying their bills and perhaps are more vividly wishing for a windfall of money. The third time is the three weeks in December before Christmas. The holiday season is expensive, and people are likely to have specific things that they can imagine purchasing with extra cash. Finally, the week between Christmas and New Years may evidence higher sales. People are normally in a good mood, off from work and celebrating. Further, this may be their last chance to win in the current year.

Table 11 below tests the effect of the last day of the month, the early part of December and Christmas week.

All of the coefficients on the calendar measures are positive and significant. Sales on the last day of the month are about 4.7% higher than the average other day. Sales in Christmas week are 8.9% higher and sales in early December are about 6% higher.

Table 11[a]

Effect of Calendar Anomalies on Sales

Dependent Variable: Log of Per Capita Sales

Regression Type	OLS	OLS	OLS
Explanatory Variables:			
prize	.017***	.017***	.017
	(.0002)	(.0002)	(.0002)***
Monday	.66***	.66***	.66***
	(.016)	(.016)	(.016)
Tuesday	1.123***	1.123***	1.123***
	(.016)	(.016)	(.016)
Wednesday	2.169***	2.169***	2.169***
	(.016)	(.016)	(.016)
Thursday	.739***	.739***	.739***
	(.016)	(.016)	(.016)
Friday	1.392***	1.392***	1.392***
	(.016)	(.016)	(.016)
Saturday	2.309***	2.309***	2.309***
	(.016)	(.016)	(.016)
last day of month	.047**		
	(.024)		
early December		.059***	
		(.016)	
Christmas week			.087***
			(.031)
constant	12.698***	12.695***	12.698***
	(.012)	(.012)	(.012)
Number of Observations	1526	1526	1526
R^2	.97	.97	.97

Data used: Overall Powerball; Daily October 27, 1996 through December 30, 2000

* significant at 10%; ** significant at 5%; *** significant at 1%

[a] Standard errors are in parenthesis.

The fact that we observe systematic (and large) differences in sales across time is strong evidence that people are influenced in their choice to purchase tickets by something other than the size of the jackpot or the odds of winning. The types of differences specifically confirmed here are evidence for the existence of some δ that measures imaginability, vividness, or feelings of luck in playing the lottery. An interesting coincidence is that at least three of these effects – the Friday effect, the Christmas week effect and the end of the month effect – are mimicked in the behavior of the stock market. This is not surprising if both behaviors are motivated by the same feelings of hope or despair, although the phenomenon deserves more investigation before conclusions can be drawn.

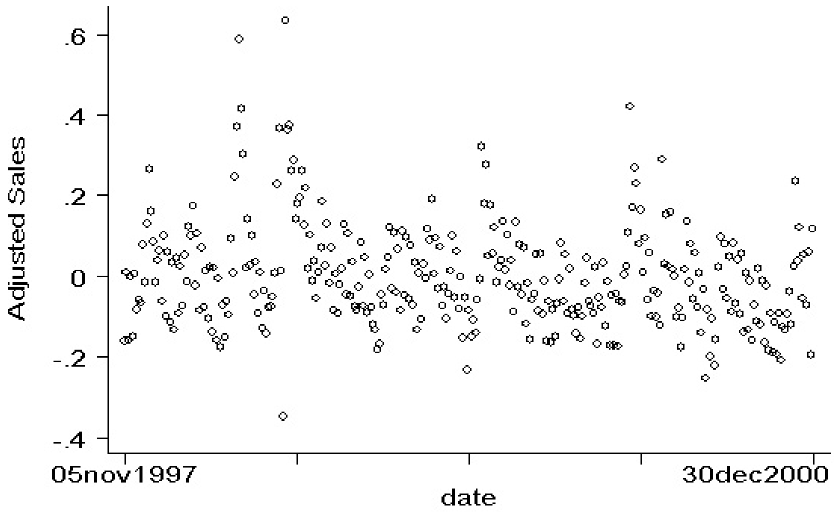
5.4.3 Time Trends

In section 5.3, when I considered the effect of changes in the lottery structure, a time trend was included in the analysis. I noted then that sales seemed to be decreasing over time. This part of the analysis looks more carefully at that effect and the implications.

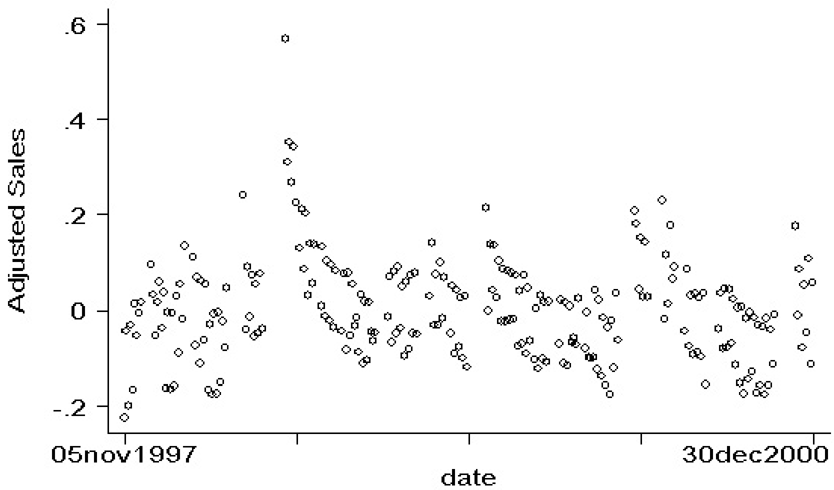
As has been emphasized before, one of the primary differences among the models in section 4 is that both the convex utility model and the prospect theoretic model treat the lottery as a two-attribute good where the additive utility theory does not. One of the more obvious ways to test for this is to test for the existence of a time trend in sales. If the lottery is truly a two-attribute good then there should be no changes in sales over time.

Showing changes in sales over time graphically is difficult because of changes in jackpot size. Graphs 8, 9 and 10 below use sales adjusted for jackpot size graphed against date. To create the adjusted sales variable the log of sales was regressed against the jackpot and the adjusted sales is the residual. Graph 8 shows the time trend for all jackpot sizes, graph 9 shows it for only jackpots below 40 million and graph 10 shows jackpots equal to and above 40 million. All graphs are limited to the current form of the game and use sales and prize in 1992 dollars.

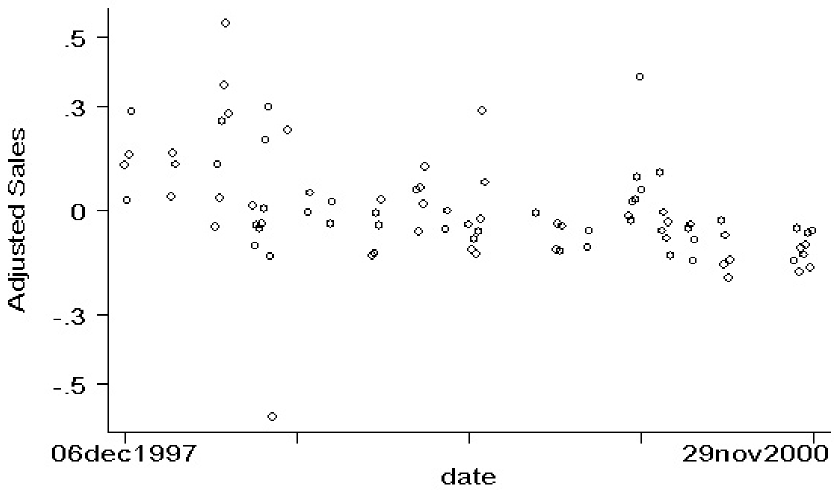
Graph 8: Sales Adjusted for Prize Graphed over Time



Graph 9: Sales Adjusted for Prize Graphed over Time
Jackpots Under $40 Million



Graph 10: Sales Adjusted for Prize Graphed over Time
Jackpots Over or Equal to $40 Million



These graphs suggest some changes in sales over time. Sales seem to decrease over time in particular at the higher jackpots. These graphs, however, do not control for other macro variables that may affect sales. Table 12 below shows a regression of real sales on real prize and time controlling for jackpot and unemployment rate. Column 2 is limited to prizes under 40 million and column 3 is limited to those 40 million and over. Time is measured in hundreds of days.

Table 12[a]

Changes in Sales over Time

Dependent Variable: Log of Per Capita Sales (in 1992 dollars)

Regression Type	OLS All Jackpots	OLS Under $40 Million	OLS $40 Million and Up
Explanatory Variables:			
prize (in 1992 dollars)	.014***	.009***	.015***
	(.0002)	(.001)	(.0003)
time[b]	−.007	.009	−.047***
	(.006)	(.006)	(.011)
unemployment rate	.017	.20**	−.429***
	(.089)	(.091)	(.159)
constant	13.497	12.409	16.337
	(.531)	(.544)	(.949)
Number of Observations	330	243	87
R^2	.93	.4	.96

Data Used: Overall Powerball; Drawings November 5, 1997 through December 30, 2000

* significant at 10%; ** significant at 5%; *** significant at 1%

[a] Standard errors in parenthesis

[b] Time measured in hundreds of days.

Time is significant and negative only for jackpots above $40 million; below that there seems to be no effect. One likely explanation for this is the fact that some people choose only to play when the jackpot is very high. Their subjective estimation of what is a "high" jackpot will be influenced by jackpots they have seen in the past. As more high jackpots accumulate, any given jackpot seems smaller so there is a decrease in sales. An alternative theory would be that sales decrease over time because people are learning that the lottery is not a good bet. Were this true we would expect to see a decrease over time for all jackpots, not just the largest ones. The existence of a time trend, like the other facts presented in this section, argue strongly against a theory in which the lottery is simply a two-attribute good.

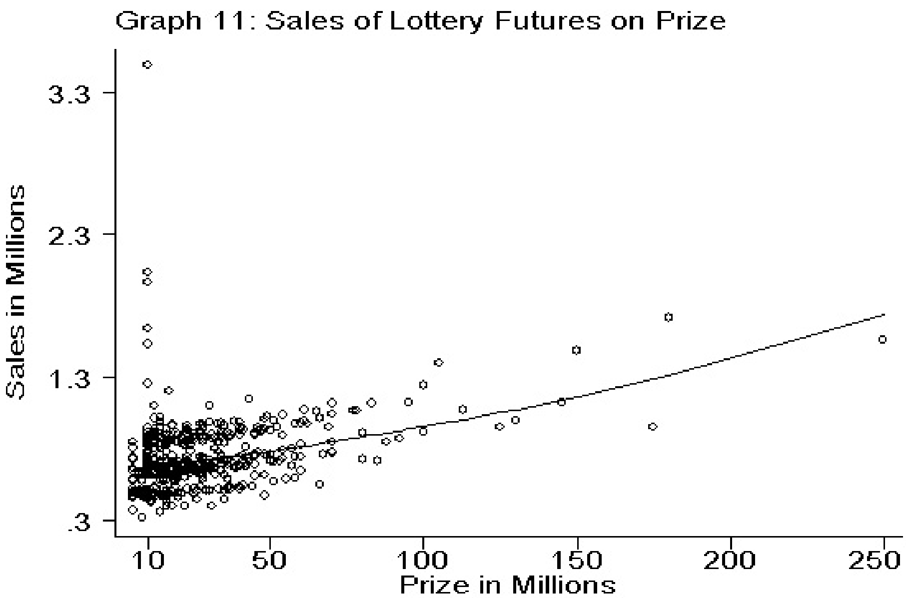
5.5 Futures

One of the more interesting features of the Powerball lottery is that it is possible to purchase "futures" – tickets for drawings other than the current one. The number of weeks ahead that you can purchase is determined by individual states, but most states allow purchases for 5 to 10 drawings in the future. The sales of futures are non-negligible. The average percentage of sales for a given drawing from futures are about 8.4%.

In general, the existence of futures (and the fact that people purchase them) argues strongly for the additive utility of gambling theory. Unless there is some added utility garnered from having a ticket there is no reason that people would prefer one ticket for each of the next ten drawings over ten tickets for this drawing. In fact, the strategy of buying one ticket for each of the next ten drawings is clearly dominated (from an expected value perspective) by other strategies. Imagine that going to purchase tickets is costly – if you live in a state with no Powerball, for example, and must travel to a neighboring state to buy tickets. It may only be possible to travel once every month to purchase tickets. One strategy would be to go at the beginning of the month and buy one ticket for each drawing. However, a dominant strategy would be to wait until a higher jackpot and go to purchase 10 tickets for that drawing. Even though it is not always possible to know if the jackpot will increase at the next drawing, it is nearly always the case that the jackpot will get up to at least $20 million before it is won, and usually much higher. For this reason unless there is utility from participating in the drawing there is no reason to systematically buy futures.

The existence of futures seems generally supportive of the additive utility theory. However, it might be the case that people purchase futures in a more rational way. For example, perhaps people only purchase futures when the jackpot is very, very high because they are anticipating the possibility that it will get even higher, and they do not want to travel or wait in line to purchase tickets again. In the context of the large drawing in August 2001 that was discussed in section 2, it may be that people from New York City purchased tickets for several drawings in the future to avoid travelling to Connecticut and waiting in line yet again. This theory would predict that most sales of futures would be at the highest jackpots, or at the $10 million jackpots (when people have purchased tickets anticipating that the jackpot may get even higher and instead someone wins and it goes back to $10 million).

Graph 11 below shows sales of futures for a given drawing against the prize in that drawing; the line represents the fitted values.



This graph demonstrates that although we do see very high sales of futures at $10 million, this is far from the only time that people purchase futures. There is actually little sales response to prize. Overall, this graph suggests a pattern of people buying, for example, one ticket for each of the next ten drawings. This is strong evidence in support of the additive utility theory of lottery play.

6 Conclusion

State lotteries are interesting to economists and policy makers for a variety of reasons. For economists, the appeal of lotteries (and other forms of gambling more generally) to consumers contradicts traditional expected utility theory. One of the most widely accepted theoretical points in economics is diminishing marginal utility of money – a concave utility function. And yet that theory is inconsistent with risk-taking behavior like playing the lottery. For policy makers, the existence and popularity of state lotteries raises questions about whether states are duping consumers and encouraging them to waste their money. Lottery foes commonly lament that states tax the poor and uneducated with no compensatory benefit. Both of these concerns – those of economists and policymakers – are fundamentally based on the question of why people play lotteries. Economics asks that question directly; policymakers need the answer to understand whether state lotteries are really as bad as some people fear.

This paper uses new data to explore three existing theories about why people play the lottery. One of these theories suggests that lotteries take advantage of a cognitive error by consumers; one suggests that diminishing marginal utility of money over all gambles is an incorrect assumption; one suggests that lotteries provide extra utility through "fun." The goal in this paper is not only to answer the decision theory questions about the lottery but to connect the theories with public policy concerns about exploitation. In this way this paper departs from earlier work on the lottery and is among the first to make policy recommendations.

I find that the data are most consistent with a theory in which the lottery provides additional utility to consumers. This theory is supported by data on differences in elasticity of sales with respect to jackpot across income levels. In addition, it is the only theory that is consistent with a change in the relationship between sales and prize when the odds change. Most importantly, however, the additive utility theory is supported because it has become increasingly clear in the data that a lottery ticket is a good with more than two attributes. That is, ticket sales are determined by more than just the odds of winning and the size of the jackpot. In the results section I showed that the day of the week, time of the year and the time since the lottery inception all have an effect on ticket sales. In addition, people purchase futures in this lottery, meaning they purchase tickets for drawings other than the current one. This is strong evidence that people like to be involved in the drawing, which indicates some extra utility. These facts are inconsistent with either the prospect theoretic explanation or the convex utility theory.

I conclude that there is at least a strong element of fun and entertainment inherent in playing the lottery, similar to that which we think motivates casino gambling or betting on horses. This counters the suggestion among some policymakers that lotteries take advantage of people. In fact, this suggests that lotteries may be a positive force. States get money to spend on something useful like education, and consumers get a little chance to dream big.

Appendix A: Full Prospect Theory Functions

The full prospect theoretic utility of n tickets is:

$$
\begin{aligned}
PTU_n &= w((1-(1-p)^n))v(J) + w((1-(1-p_1)^n))v(100,000)\cdots \\
&\quad + w((1-(1-p_2)^n))v(5000) + w((1-(1-p_3)^n))v(100)\cdots \\
&\quad + w((1-(1-p_4)^n))v(100) + w((1-(1-p_5)^n))v(7)\cdots \\
&\quad + w((1-(1-p_6)^n))v(7) + w((1-(1-p_7)^n))v(4)\cdots \\
&\quad + w((1-(1-p_8)^n))v(3)
\end{aligned}
$$

where J is the size of the jackpot, p is the probability of winning the jackpot and p_j is the probability of winning the jth smaller prize.

The p_j probabilities with the corresponding prizes are listed in the table below.

Probability	**Prize**
1 in 80,089,128	Jackpot
1 in 1,953,393	\$100,000
1 in 364,041	\$5000
1 in 8,879	\$100
1 in 8,466	\$100
1 in 206	\$7
1 in 604	\$7
1 in 117	\$4
1 in 74	\$3

Appendix B: Changes in Minimum Odds

One thing we are concerned about is that the change in the minimum jackpot when the odds changed would change the sales independently of the odds of winning changing. Fortunately, the minimum jackpot has changed more than once. Table B1 below shows the effect of the change of minimum jackpot from $2 million to $3 million (column 1) and then from $3 million to $5 million (column 2). Note that in both cases the jackpots are limited to the jackpots that are shared in both game structures.

Table B1[a]

Effect of Changes in Minimum Jackpot

Dependent Variable: Log of Per Capita Sales (in 1992 dollars)

Regression Type	OLS	OLS
Explanatory Variables:		
prize (in 1992 dollars)	.026***	.019***
	(.001)	(0)
dummy for minimum $3 million	−.016	
	(.042)	
dummy for minimum $5 million		−.003
		(.039)
total population[b]	−.036***	.012***
	(.007)	(.003)
time[c]	.006	−.036***
	(.017)	(.006)
unemployment rate	−.246***	−.018
	(.076)	(.049)
constant	.731	−2.089***
	(.651)	(.315)
Number of Observations	329	299
R^2	.9	.92

Data used: Overall Powerball; Drawings April 22,1992 through December 30, 2000

* significant at 10%; ** significant at 5%; *** significant at 1%

[a] Standard errors in parenthesis

[b] Population in millions

[c] Time in hundreds of days

It can easily be seen here that in neither case did the change in jackpot significantly change sales, once population and time are controlled for.

References

[1] Adams, D., 1997. Playing the Lottery: Social Action, Social Networks and Accounts of Motive. Dissertation, University of Arizona.

[2] Adams, D., 2000. Games People Play (and Why They Play Them): The 'Meaning' of Lottery Participation. Humanity and Society 24.

[3] Badad, E., Katz, Y., 1991. Wishful Thinking - Against All Odds. Journal of Applied Social Psychology 21:1921-1938.

[4] Bartsh, R., Paton, V., 1999. The Presence of Odd Pricing in the Texas State Lottery. Journal of Applied Social Psychology 29:2394-2409.

[5] Borg, M., Mason, P., Shapiro, S., 1993. The Cross Effects of Lottery Taxes on Alternative State Tax Revenue. Public Finance Quarterly 21:123-140.

[6] Brown, D.J., Kaldenberg, D.O., Browne, B.A., 1992. Socio-Economic Status and Playing the Lotteries. Sociology and Social Research 76: 161-167.

[7] Camerer, C., 2000. Prospect Theory in the Wild: Evidence from the Field. In *Choices, Values and Frames*, edited by Daniel Kahneman and Amos Tversky. Cambridge, UK: Cambridge University Press, 2000.

[8] Clark, A., Oswald, A., 1996. Satisfaction and Comparison Income. Journal of Public Economics 61:359-381.

[9] Clotfelter, C.T., Cook, P., 1989. Selling Hope: State Lotteries In America. Harvard University Press, Cambridge.

[10] Clotfelter, C.T., Cook, P., 1991a. The "Gambler's Fallacy" in Lottery Play. National Bureau of Economic Research Working Paper 3769.

[11] Clotfelter, C.T., Cook, P., 1991b. Lotteries in the Real World. Journal of Risk and Uncertainty 4:227-232.

[12] Coogan, M., editor, 2001. *The New Oxford Annotated Bible*. Oxford: Oxford University Press, 2001.

[13] Conlisk, J., 1993. The Utility of Gambling. Journal of Risk and Uncertainty 5: 255-275.

[14] Cook, P., Clotfelter, C., 1993. The Peculiar Scale Economies of Lotto. American Economic Review 83:634-643.

[15] Donkers, B., Melenberg, B., Van Soest, A., 2001. Estimating Risk Attitudes using Lotteries: A Large Sample Approach. Journal of Risk and Uncertainty 22:165-195.

[16] Easterlin, R., 1995. Will Raising the Incomes of All Increase the Happiness of All? Journal of Economic Behavior and Organization 27:35-47.

[17] Easterlin, R., 2001. Income and Happiness: Towards a Unified Theory. The Economic Journal 111:465-484.

[18] Farrell, L., Hartley, R., Lanot, G., Walker, I., 2000. The Demand for Lotto: The Role of Conscious Selection. Journal of Business and Economic Statistics 18 (2), 228-241.

[19] Farrell, L., Walker, I., 1999. The Welfare Effects of Lotto: Evidence from the UK. Journal of Public Economics 72:99-120.

[20] Fishburn, P., 1980. A Simple Model for the Utility of Gambling. Psychometrica 45:435-448.

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

[22] Garrett, T., Sobel, R., 1999. Gamblers Favor Skewness, not Risk: Further Evidence from United States' Lottery Games. Economics Letters 63:85-90.

[23] Griffith, R., 1949. Odds Adjustments by American Horse-Race Bettors. American Journal of Psychology 62:290-294.

[24] Gruber, J., Köszegi, B., 2001. Is Addiction "Rational"? Theory and Evidence. Quarterly Journal of Economics 116:1261-1303.

[25] Haruvy, E., Erev, I., Sonsino, D., 2001. The Medium Prizes Paradox: Evidence from a Simulated Casino. Journal of Risk and Uncertainty 22:251-261.

[26] Imbens, G., Rubin, D., Sacerdote, B., 1999. Estimating the Effect of Unearned Income on Labor Supply, Earnings, Savings and Consumption: Evidence form a Survey of Lottery Players. National Bureau of Economic Research Working Paper 7001.

[27] Kahneman, D., Tversky, A., 1979. Prospect Theory: An Analysis of Decision Under Risk. Econometrica 47: 263-291.

[28] Kahneman, D., Tversky, A., 1992. Advances in Prospect Theory: Cumulative Representation of Uncertainty. Journal of Risk and Uncertainty 5:297-323.

[29] Kitchen, H., Powells, S., 1991. Lottery Expenditures in Canada: a regional analysis of determinants and incidence. Applied Economics 23:1845-52.

[30] Kwang,N.Y., 1965. Why do People Buy Lottery Tickets? Choices Involving Risk and the Indivisibility of Expenditure. Journal of Political Economy 73:530-535.

[31] Le Menestrel, M., 2001. A Process Approach to the Utility for Gambling. Theory and Decision 50:249-262.

[32] Lowenstein, G., 1987. Anticipation and the Valuation of Delayed Consumption. The Economic Journal 97:666-684.

[33] Markowitz, H., 1952. The Utility of Wealth. Journal of Political Economy 60:151-158.

[34] Mason, P.M., Steagall, J.W., Fabritius, M.M., 1997. The Elasticity of Demand for Lotto Tickets and the Corresponding Welfare Effects. Public Finance Review 25:474-90.

[35] Mikesell, J., 1994. State Lottery Sales and Economic Activity. National Tax Journal 47:165-71.

[36] Mikesell, J., Priog-Good, M., 1990. State Lotteries and Crime. American Journal of Economics and Sociology 49:7-19.

[37] O'Donoghue, T., Rabin, M., 1999. Doing it Now or Later. The American Economic Review 89:103-124.

[38] Oster, E., 2001. Who Buys Hope? Determinants of Habitual Lottery Play. Mimeo.

[39] Pirog-Good, M., Mikesell, J., 1995. Longitudinal Evidence of the Changing Socio-Economic Profile of a State Lottery Market. Policy Studies Journal 23:451-465.

[40] Preston, M., Baratta, P., 1948. An Experimental Study of the Auction-Value of An Uncertain Outcome. American Journal of Psychology 63:183-193.

[41] Price, D., Novak, E.S., 1999. The Tax Incidence of Three Texas Lottery Games: Regressivity, Race and Education. National Tax Journal 52:741-751.

[42] Price, D., Novak, E.S., 2000. The Income Redistribution Effects of Texas State Lottery Games. Public Finance Review 28:82-92.

[43] Productivity Commission of Australia, 1999. Australia's Gambling Industries. Report #10, November 1999.

[44] Roth, A., Ockenfels, A., 2000. Last Minute Bidding and the Rules for Ending Second-Price Auctions: Theory and Evidence from a Natural Experiment on the Internet. National Bureau of Economic Research Working Paper 7729.

[45] Scott, F., Garen, J., 1994. Probability of Purchase, Amount of Purchase and the Demographic Incidence of the Lottery Tax. Journal of Public Economics 54:121-143.

[46] Siegel, D., Anders, G., 2001. The Impact of Indian Casinos on State Lotteries: A Case Study of Arizona. Public Finance Review 29:139-147.

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

[48] Stranahan, H., Borg, M., 1998a. Separating the Decisions of Lottery Expenditures and Participation: A Truncated Tobit Approach. Public Finance Review 26:99-117.

[49] Stranahan, H., Borg, M., 1998b. Horizontal Equity Implications of the Lottery Tax. Public Finance Review 51:71-82.

[50] Thaler, R., 1991. Quasi Rational Economics. Russell Sage Foundation, New York.

[51] Walker, I., 1998. Lotteries: Determinants of Ticket Sales and the Optimal Payout Rate. Economic Policy 27:357-392.

[52] Wärneryd, K., 1996. Risk Attitudes and Risky Behavior. Journal of Economic Psychology 17:749-770.

[53] Worthington, A., 2001. Implicit Finance in Gambling Expenditures: Australian Evidence on Socioeconomic and Demographic Tax Incidence. Public Finance Review 29:326-342.

[54] Wu, R., 1979. A Cross-Section Study on the Demand for State Lottery Tickets. American Economist 23:6-12.