Gamblers Favor Skewness, Not Risk: Further Evidence from United States’ Lottery Games

Thomas A. Garrett
Russell S. Sobel

Department of Economics
West Virginia University
Morgantown, West Virginia 26506

JEL Code: D81

Keywords: Risk, Skewness, Uncertainty, State Lotteries

Abstract

Theoretical models of risk have attempted to explain why risk-averse individuals take unfair gambles. Using all United States’ lottery games, we find theoretical and empirical evidence that skewness of prize distributions explains why risk averse individuals may play the lottery.

Corresponding Author: Thomas A. Garrett, Department of Economics, West Virginia University, Morgantown, West Virginia 26506
Phone: (304) 293-5721
Fax: (304) 293-5652
E-mail: tgarrett@wvu.edu
Gamblers Favor Skewness, Not Risk: Further Evidence from United States’ Lottery Games

I. Introduction

The popularity of state lotteries in America has grown over the last several decades. In 1986, revenues from all games offered in the United States totaled 12.5 billion dollars, whereas 1996 revenues from all lottery games topped 34 billion dollars. Although the popularity of state lotteries has increased, economic theory has said little about why people play lotteries, given their low payout rates and remote odds of winning. Provided that the average expected lottery payout is only 50 cents, Thaler and Ziemba (1988) state that for every dollar bet, the lottery customer is “paying 50 cents for a fantasy.” Several authors, however, have attempted to explain why risk-averse individuals favor such unfair gambles. Friedman and Savage (1948) suggest that risk averse people may indulge in unfair gambles if winning will significantly improve their standard of living. In a later paper, Kahneman and Tversky (1979) suggest that players place decision weights on the probabilities of each outcome. An over-weighting of low probabilities, especially those associated with multi-million dollar jackpots, may explain the attractiveness of state lotteries. Finally, Quiggin (1991) uses a rank-dependent utility

---


2 The Friedman-Savage (1948) model has frequently been used to explain why the same individual buys insurance (suggesting risk aversion), yet places bets on unfair gambles (suggesting risk loving).
function to explain why individuals play lottery games. He theorizes that it is utility maximizing to play the lottery if smaller prizes are offered besides the jackpot.

Golec and Tamarkin (1998) provide fresh insight into why risk-averse bettors may indulge in unfair gambles. Focusing on horse track gamblers, Golec and Tamarkin find that bettors’ behavior at horse tracks is explained by expected utility functions that not only consider the mean and variance (risk) of returns, but also the skewness of returns. The work of Golec and Tamarkin has its basis in what is called the “long shot bias,” where high-probability, low variance bets provide relatively high average returns, and low-probability, high variance bets provide relatively lower average returns. Long shots are thus overbet and favorites are underbet. Prior to Golec and Tamarkin, Quandt (1986) suggested that this long shot bias is consistent under the assumption that bettors are risk lovers with mean-variance utility functions. Golec and Tamarkin, however, find evidence that this long shot bias may be explained by bettors’ preferences for positive skewness, rather than risk. Bettors are risk-averse, but are attracted to the positive skewness of returns offered by low probability, high variance bets.

Our paper extends the work of Golec and Tamarkin by modeling and empirically testing an expected utility model for a lottery player that incorporates the skewness of returns. We question whether the findings of Golec and Tamarkin may also provide insight as to why risk-averse individuals play lotteries. Using state lotteries to examine individuals’ preferences for risk and skewness has several advantages over the use of horse race bettors. State lotteries command a greater customer base than horse racing. In 1992, per capita racing wagers totaled $66.39, whereas per capita lottery
wagers totaled $120.92. Given that a greater number of people play lotteries than wager on horse racing, individuals’ preferences for risk and skewness can be more generalized through the use of state lotteries than horse racing. Also, unlike horse racing, the prize structure of state lotteries is much more skewed, as lottery games have a higher probability of winning nothing and a much smaller probability of winning a jackpot. Thus, if individuals truly favor skewness of returns, state lotteries provide a more preferable arena for modeling and testing individuals risk preferences.

II. Modeling a Lottery Player’s Expected Utility

The first step in determining a lottery player’s behavior is to model a player’s expected utility. We follow Golec and Tamarkin’s (1998) methodology, which is based on the work of Ali (1977). We assume that all players have identical utility functions, bet their entire wealth, and a losing lottery bet returns zero to the bettor, that is $U(0) = 0$. We also assume that lottery player $j$’s expected utility only depends on the top prize payouts of each lottery game in state $i$, and a player $j$ only plays those lottery games available in his or her state of residence. A lottery player’s expected utility is:

$$E(U_{ji}) = P_{Gi} U_{ji}(X_{Gi}) \prod_{g^*}^{n} P_{gi} U_{ji}(X_{gi})$$  

The first term is the probability of winning the highest top prize game $G$ in state $i$ multiplied by player $j$’s utility from winning the top prize, $X_{Gi}$, in game $G$. The final term considers all other lottery games $g$.

---


4 The possibility of cross-border shopping is ignored for mathematical simplicity.
offered in state $i$ except the highest top prize game $G$ in state $i$. This term is simply the probability of winning the top prize in any game $g$ multiplied by player $j$’s utility from winning the top prize, $X_{gi}$, in any game $g$, summed over all $n$ of $g$ games. The addition of the two terms thus gives the expected utility for player $j$ for all lottery games available in state $i$.

As in Ali (1977) and Golec and Tamarkin (1998), if we assume $U_{ji}(X_{Gi}) = 1$ and lottery players are indifferent between bets on any game $g$ or $G$, then equation (1) can be expressed as:

$$E(U_{ji}) = P_{Gi} = P_{i1} \cdot U_{ji}(X_{i1}) = P_{i2} \cdot U_{ji}(X_{i2}) = \ldots = P_{ni} \cdot U_{ji}(X_{ni}) \quad (2)$$

Thus, for any lottery game $g$ or $G$ in state $i$, $E(U_{ji}) = P_{Gi} = P_{gi} \cdot U_{ji}(X_{gi})$, or

$$\frac{P_{Gi}}{P_{gi}} \cdot U_{ji}(X_{gi}) \quad (3)$$

Our equation (3) is identical to that obtained by Golec and Tamarkin (1998), but our interpretation of each parameter is consistent with lottery games rather than horse racing.\(^5\) Our final formulation suggests that the expected utility for any lottery player in state $i$ can be represented by equating the odds ratio of the highest top prize game $G$ and any other lottery game $g$ to player $j$’s utility from winning the top prize in any game $g$.

Golec and Tamarkin (1998) describe two ways of empirically estimating the expected utility model. Ali (1977) utilized a power function, such that $\frac{P_{Gi}}{P_{gi}} \cdot \hat{a} \cdot \hat{a}$, where the size of $\hat{a}$ determines bettors’ risk preferences. A $\hat{a} > 1$ implies risk loving, $\hat{a} = 1$ implies risk neutrality, and $\hat{a} < 1$ implies

\(^5\) Note that for the highest top prize game in each state $i$, this ratio is simply one.
risk aversion. However, Golec and Tamarkin (1998) suggest there is no prior justification for choosing this functional model form, and it provides no direct measure of skewness. Golec and Tamarkin use the expectations of a Taylor series truncated to three terms, arriving at a polynomial approximation of Ali’s power function. This allows for a more direct test of bettors’ preferences for skewness, as it provides separate, estimable terms for the mean, variance and skewness of returns. To empirically test our expected utility model of a lottery player, we choose the cubic approximation developed by Golec and Tamarkin. We estimate:

\[
\frac{P_{Gi}}{P_{gi}} = \hat{\alpha}_0 + \hat{\alpha}_1 g_{i1} + \hat{\alpha}_2 g_{i2} + \hat{\alpha}_3 g_{i3} \quad (4)
\]

The coefficient \(\hat{\alpha}_1\) captures bettors’ preferences over the mean of returns, whereas risk aversion is determined by the size of \(\hat{\alpha}_2\): \(\hat{\alpha}_2 > 0\) suggests risk loving, \(\hat{\alpha}_2 < 0\) suggests risk aversion, and \(\hat{\alpha}_2 = 0\) suggests risk neutrality. A player’s preferences for skewness is measured by \(\hat{\alpha}_3\). A positive \(\hat{\alpha}_3\) would imply a favorable preference for skewness, whereas a negative \(\hat{\alpha}_3\) would reflect a dislike for skewness. We expected \(\hat{\alpha}_1 > 0\), \(\hat{\alpha}_2 < 0\), and \(\hat{\alpha}_3 > 0\). That it, lottery players are risk-averse, but choose to play those lottery games having a higher skewness of returns. This simply suggests that although lottery players may be risk-averse, the large top prize jackpots available entice risk-averse individuals to play the lottery.
A listing of United States' lottery games can be found in Public Gaming, 1996. There are two types of lottery games: instant and on-line. Instant lottery games, also termed "scratch-offs," require the player to scratch the play area to reveal a winning prize. Lottery agencies change their instant game portfolio every several months. On-line lottery games are a more permanent part of a state’s lottery portfolio. To play an on-line game, the player is required to fill out a play slip and submit the slip for processing. The player wins if she matches some or all of her numbers with those drawn.

On-line games can have either a parimutuel or a fixed prize structure. Fixed prize games offer a fixed top prize that is independent of sales and the number of top prize winners. The prizes for fixed prize games thus remain the same for every drawing. Parimutuel prize games, however, have a top prize that is dependent upon the sales for that drawing and the number of players that win the top prize. The top prize for a parimutuel game thus varies for each drawing. Because actual top prize figures for parimutuel prize games are not available because they are different at each drawing, we estimated the average top prize per drawing for all parimutuel games by using annual sales data and the percent of sales that is allocated to the top prize. 1995 sales information for all lottery games was obtained from Public Gaming, September 1996.

III. Empirical Methodology and Results

We estimate equation (4) using data from all 216 on-line lottery games offered in the United States during 1995. For each game, the top prize and the odds of winning the top prize were obtained from each state’s lottery commission. The dependent variable, \( \frac{P_{Gi}}{P_{g_i}} \), is calculated for every lottery game in our sample. We then performed several regressions. First,
we included all 216 lottery games in the sample. We then divided the sample between those games offering top prizes less than $10,000 and those games offering top prizes greater than $10,000 and performed regressions on each sub-sample. This allows us to examine whether risk and skewness preferences differ if one considers different magnitudes of top prize payouts. Finally, a set of state dummy variables were included in all three regressions. This provides the opportunity to test whether preferences within a state differ from risk preferences across all states. The results from our empirical estimations are shown in Table 1.

[Table 1 about here]

In the full-sample specification, the coefficient estimates on \( \hat{\alpha}_1 \) and \( \hat{\alpha}_3 \) are significantly greater than zero, whereas \( \hat{\alpha}_2 \) is significantly less than zero. A graph of our cubic utility function is shown in Figure 1.

[Figure 1 about here]

The shape of the utility function is similar to that proposed by Friedman and Savage (1948) and found by Golec and Tamarkin (1998). At low odds (a higher probability of winning) where skewness is small, gamblers behave as though they are risk-averse. However, at higher odds (a lower probability of winning) and thus higher skewness, gamblers behave as though they are risk-loving. The evidence from the full-sample specification suggests that lottery players, like horse race bettors, are risk averse but favor positive skewness. This conclusion also holds true for both sub-samples. The inclusion or

\[ \text{Further note that the adjusted } R^2 \text{ is larger for those games having a top prize greater than } $10,000 \text{ (worse odds of winning) compared to those offering a top prize less than } $10,000 \text{ (better} \]
exclusion of state dummy variables in each specification to test for differences within a state does not change the sign or significance of any coefficients. Whether we consider lottery games within a state or games across all states, we find that lottery players are risk-averse, but prefer positive skewness.

IV. Conclusion

The conclusion reached by Golec and Tamarkin (1998) is important, as they provide new insight as to why risk-averse individuals might gamble. Because Golec and Tamarkin’s results are based solely on bettors at horse tracks, however, we have modified their expected utility model to represent the expected utility of the more common lottery player. Using data from all on-line lottery games in the United States, the estimates of our model suggest that lottery players may be risk-averse but favor positive skewness of returns. As state lotteries are the most prevalent form of gambling in the United States, our results provide a more general and definitive conclusion regarding the risk preferences and behavior of gamblers.

---

odds of winning). This finding is consistent with that of Golec and Tamarkin (1998).
References


Table 1 - Cubic Estimation of the Expected Utility Model

Dependent Variable: $\frac{P_{Gi}}{P_{gi}}$

<table>
<thead>
<tr>
<th>All On-Line Lottery Games</th>
<th>$\hat{a}_0$</th>
<th>$\hat{a}_1$</th>
<th>$\hat{a}_2$</th>
<th>$\hat{a}_3$</th>
<th>State Dummies Included</th>
<th>Adj. $R^2$</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.020</td>
<td>1.023***</td>
<td>-0.323***</td>
<td>0.030***</td>
<td>No</td>
<td>0.742</td>
<td>216</td>
</tr>
<tr>
<td></td>
<td>(1.306)</td>
<td>(11.004)</td>
<td>(6.836)</td>
<td>(5.209)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.021</td>
<td>1.032***</td>
<td>-0.317***</td>
<td>0.030***</td>
<td>Yes</td>
<td>0.760</td>
<td>216</td>
</tr>
<tr>
<td></td>
<td>(0.191)</td>
<td>(10.458)</td>
<td>(6.246)</td>
<td>(4.691)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>On-Line Games Having A Top Prize &gt; $10,000</th>
<th>$\hat{a}_0$</th>
<th>$\hat{a}_1$</th>
<th>$\hat{a}_2$</th>
<th>$\hat{a}_3$</th>
<th>State Dummies Included</th>
<th>Adj. $R^2$</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.060</td>
<td>0.922***</td>
<td>-0.279***</td>
<td>0.026***</td>
<td>No</td>
<td>0.636</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td>(1.418)</td>
<td>(5.700)</td>
<td>(3.585)</td>
<td>(2.754)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.037</td>
<td>0.892***</td>
<td>-0.247**</td>
<td>0.023*</td>
<td>Yes</td>
<td>0.692</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td>(0.203)</td>
<td>(4.661)</td>
<td>(2.576)</td>
<td>(1.918)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>On-Line Games Having A Top Prize &lt;= $10,000</th>
<th>$\hat{a}_0$</th>
<th>$\hat{a}_1$</th>
<th>$\hat{a}_2$</th>
<th>$\hat{a}_3$</th>
<th>State Dummies Included</th>
<th>Adj. $R^2$</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.012**</td>
<td>0.500***</td>
<td>-0.136***</td>
<td>0.089***</td>
<td>No</td>
<td>0.128</td>
<td>112</td>
</tr>
<tr>
<td></td>
<td>(2.033)</td>
<td>(4.350)</td>
<td>(4.139)</td>
<td>(3.967)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.014</td>
<td>0.428***</td>
<td>-0.115***</td>
<td>0.074***</td>
<td>Yes</td>
<td>0.014</td>
<td>112</td>
</tr>
<tr>
<td></td>
<td>(0.754)</td>
<td>(3.439)</td>
<td>(3.218)</td>
<td>(3.031)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*** Denotes significance at 1%, ** at 5%, and * at 10%.
Absolute $t$-statistics in parentheses.

* The dependent variable is the probability of winning the highest top prize of game $G$ in state $i$, divided by the probability of winning the top prize in any other game $g$ offered in state $i$. 
Figure 1 - The Cubic Utility Function

Utility

Area of Risk Aversion

Area of Risk Loving

High $P_g$ Low $P_g$