The Demand for Lottery Products and Their Distributional Consequences

Linda Ghent
Eastern Illinois University
lsghent@eiu.edu

Alan P. Grant
Eastern Illinois University
apgrant@eiu.edu

February 2007

Abstract

This study examines the distributional impact of three types of lottery games offered for sale by the South Carolina Education Lottery (SCEL). We significant sales variation by game type across age and race. We also find each of the three types of game offered by the SCEL to be regressive, as measured by the Suits Index, but we find substantial differences in the degree to which each game is regressive. By modeling income as a distribution rather than in levels, regression analysis shows sales to be flat in the tails of the income distribution, but significantly higher for upper-middle-income residents, and significantly lower for lower-middle-income residents. Together, this evidence suggests that the lottery may not be as regressive as the body of literature suggests it once was, and indicates that when estimating lottery sales with regression analysis, it may be better to model the distribution of income rather than its level.

Introduction

Currently, 42 states (along with the District of Columbia) have state-operated lotteries. The growth of lottery creation throughout the United States has been paralleled by the growth of economic research concerning lotteries. Three important issues have been examined: (1) the factors that lead to the creation of a lottery, (2) the determinants of the demand for lottery play, and (3) the distributional impact of this type of public finance on various demographic and economic groups.

This paper examines two of these issues within the context of the South Carolina Education Lottery (SCEL).[1] The SCEL was established by voter referendum in the fall of 2000, and allows individuals 18 years or older to purchase three separate products: instant scratch-off tickets, fixed-odds online games such as Pick 3 and Pick 4, and rollover games with progressive jackpots (Lotto) such as Powerball or MegaMillions. Our focus will be to determine the factors that influence the purchases of these products separately and to examine their distributional impacts across income and demographic factors.

The Demand for Lottery Products

While there have been a multitude of studies examining lottery sales in general, only a handful of studies have examined sales by type of lottery product. Most state lotteries consist of several products from instant scratch off tickets to Lotto games with large jackpots. Garrett and Sobel (2004) point out the importance of game characteristics in the examination of the demand for lottery products. They find that ticket sales are significantly related to the size of the largest

[1] Ghent and Grant (forthcoming) examines the relationship between the determinants of the voter referendum that created the SCEL and the determinants of aggregate lottery sales.

prize and the odds of winning it.[2] Given that these two factors vary greatly across different types of lottery games, their results suggest that estimating the demand for each lottery product separately is more appropriate than examining total lottery sales. As mentioned above, the SCEL offers three distinct types of lottery product. To better understand the determinants of sales of each, a seemingly unrelated regession model has been estimated for the following three equations:

(1) $INSTANTPC_i = X_i\alpha + \rho \mathbf{M} INSTANTPC_i + \eta_i$,

(2) $FIXEDPC_i = X_i\alpha + \rho \mathbf{M} FIXEDPC_i + \eta_i$, and

(3) $LOTTOPC_i = X_i\alpha + \rho \mathbf{M} LOTTOPC_i + \eta_i$,

where $INSTANTPC_i$, $FIXEDPC_i$, and $LOTTOPC_i$ are per capita sales of instant games, fixed number games, and Lotto tickets in county *i*, X_i is a vector of economic and demographic characteristics in county *i*, and **M** is a ($N \times N$) spatial weights matrix.[3] Rho, the spatial autoregressive coefficient, reflects positive spatial correlation if $\rho > 0$, negative spatial correlation if $\rho < 0$, and no spatial correlation if $\rho = 0$. A significant value of ρ indicates that omitted variables are correlated across neighboring counties.

County characteristics included in X_i consist of age, race, income, education, religious affiliation, and indicators for the possibility of cross-border shopping from neighboring states. **AGE65** is the proportion of a county's residents ages 65 and older. According to Clotfelter and Cook (1990), individuals ages 25 to 64 are more likely to play the lottery than those who are age 65 or older. However, Jackson (1994) reports mixed results in his study of the Massachusetts

[2] In their examination of the UK National Lottery, Forrest et al. (2004) also find that own-game characteristics are the most important determinant of the sales of each product.

[3] The elements of the spatial weights matrix, $\mathbf{M}_{ij}$, initially assume a value of 1 if county *i* and county *j* abut. Elsewhere, the elements are assigned a value of zero. Once constructed, the rows of the matrix are normalized such that the entries sum to one.

state lottery. In 1983, the proportion of the population age 65 or older was inversely related to per capita lottery sales; by 1990, this relationship had reversed itself. This result holds when Jackson examines each lottery game separately with one exception: the population age 65 or older has no significant effect on Lotto sales in 1990. In addition, Price and Novak (1999) find that median age is inversely related to sales of Lotto and Pick 3 tickets, but positively related to the sales of instant games.

The proportion of county residents who are African American (**BLACK**) is also included. Several studies have found that African Americans play the lottery more than whites do (see for example, Clotfelter and Cook, 1987; Borg and Mason, 1988; Rubenstein and Scafidi, 2002; and Ghent and Grant (forthcoming). Giacopassi et al. (2006) find no effect of race on total lottery sales in Tennessee, but when sales are sorted by game type, they find that African Americans play significantly more online games than their white counterparts.

Because we are interested in examining the distributional effects of the lottery games, we measure income in terms of the income distribution rather than in levels. **LOWINC** is the proportion of county households earning less than $15,000 per year, **LMINC** is the proportion of lower-middle income households earning between $15,000 and $35,000, and **UMINC** is the proportion of upper-middle income households earning between $35,000 and $50,000 per year.[4] Similar measurement has been applied in Hersch and McDougall (1989) and Giacopassi, et al.(2006).

The estimated effect of income on lottery expenditures in previous studies has been mixed. Price and Novak (1999, 2000) find that sales of instant games are negatively correlated

---

[4] These income distribution variables omit the percentage of high income households earning more than $50,000 per year. Inclusion of this category would result in a perfectly collinear matrix of explanatory variables. Thus, the LOWINC, LMINC, and UMINC coefficients are interpreted as incremental purchases of lottery games relative to the baseline higher income household.

with income. However, Clotfelter and Cook (1987) find that lottery expenditures are flat across the income distribution, which implies a regressive lottery tax. Scott and Garen (1994) find income to have a positive, but declining effect on the probability an individual plays the lottery. Interestingly, once they control for the probability that an individual plays the lottery, income has no significant effect on lottery expenditures. Rubenstein and Scafidi (2002) find similar results.

The level of education in a county is measured by two variables, the proportion of the population age 25 or older without a high school diploma (**NOHS**) and the proportion of the population age 25 or older with at least a bachelor's degree (**EDUCBS**). Both Scott and Garen (1994) and Rubenstein and Scafidi (2002) find an inverse relationship between education and the probability of lottery play. In their analysis of the Tennessee Education Lottery, Giacopassi et al. (2006) report a negative relationship between the proportion of a county's residents with a college degree and that county's lottery sales. Ghent and Grant (forthcoming) confirm the role of education in determining lottery sales by finding that total sales depend positively on the proportion of a county's residents without a high school diploma in South Carolina. Finally, Price and Novak (2000) find that the percent of a particular county's residents with a bachelor's degree is positively associated with Lotto sales, but negatively associated with scratch-off instant games.

Because lottery expenditures are likely to be lower for those who are morally opposed to gambling, we include the proportion of a county's residents who are members of an evangelical protestant or traditional black church (**RELIGION**) in our vector of county characteristics. Previous studies have shown religion to have varying effects on lottery sales. Rubenstein and Scafidi (2002), for example, find that individuals who reports regular church attendance have a lower probability of playing the lottery. In contrast, Giacopassi et al. (2006) report higher lottery sales for counties with greater proportions of evangelical church members. Ghent and Grant (forthcoming) find that while religion plays an important role in shaping the vote to establish a lottery, it has no significant effect on lottery sales.

Finally, we include several variables to account for the distribution of South Carolina's counties along the state's borders. Stover (1990) and Tosun and Skidmore (2004) find border-state competition to be an important determinant of lottery sales. Garrett and Marsh (2002) also find that cross-border lottery shopping is significant, and that the amount of cross-border shopping depends on the size of the retail sector in the relevant border county.

South Carolina abuts North Carolina to the north and Georgia to the south and west. For the time period covered by our data, Georgia had a state lottery and North Carolina did not. Thus, we construct both **NCBORDER** and **GABORDER** as indicator variables, weighted by the percent of the county population employed in retail.

The sales data examined using equations (1) through (3) are from the South Carolina Education Lottery over the period of January 2002 through March 2003. Summary statistics and data sources for all of the variables used in our analysis are reported in Table 1. Mean per capita sales of instant lottery tickets are more than three times greater than mean per capita sales of other lottery products. Lotto sales, however, vary most across counties (as measured by the coefficient of variation).

The results from the OLS estimations are in Table 2. There are several striking differences in the estimated coefficients for each of the three products.[5] The proportion of the county population age 65 or older is significantly and positively related to the sales of instant lottery tickets, but has no effect on the sales of fixed-number or Lotto games. Jackson (1994)

---

[5] An F-test allows us to reject the null hypothesis that the coefficients are equal across the three equations at the 1% significance level.

finds similar results using 1990 data for Massachusetts. Additionally, Price and Novak (1999) report that median age is inversely related to Lotto and Pick 3 sales in Texas, but has a positive impact on the sale of instant games.

Race also provides mixed results, with the proportion of the population that is African American having a significant positive impact on sales of instant and fixed-number games, but no significant effect on Lotto sales. This result is also consistent with the findings of Jackson (1994) and Price and Novak (1999).

Although Giacopassi et al. (2006) find that the proportion of households affiliated with an evangelical faith is positively related to higher lottery sales, we find **RELIGION** to be insignificant in all three of our estimations. Our results are similar to those of Scott and Garen (1994) and Rubenstein and Scafidi (2002) who find religion to be unrelated to lottery expenditures, and confirm Ghent and Grant's prior finding that religion is uncorrelated with aggregate lottery product purchases in South Carolina.

The signs of the estimated coefficients on the income distribution variables are consistent across the three products. Lower-middle income households appear to spend less on all lottery products than do high income households. This result is almost identical to that of Giacopassi, et al. (2006). Interestingly, counties with a large proportion of low income households have similar per capita lottery sales as those with a large proportion of high income households. Most striking is the coefficient on upper-middle income. A one-percentage-point increase in the proportion of upper-middle income households raises per capita sales of instant lottery games by almost $59. The effects of a one percentage point change in this variable on per capita sales of fixed and Lotto games are $15 and almost $12, respectively. Thus, it appears that household income does affect lottery sales, but the direction and magnitude of that effect depends on the level of the household's income. Lower-middle income households buy the fewest lottery products; upper-middle income households purchase the most. These results suggest that estimates of lottery demand that use only levels of income without accounting for its distribution may cause researchers to overlook important information contained in the data. There seems to be particular value in accounting for income distribution for those studies oriented toward determining the lottery's distributional burden.

Next, we want to consider the effects of the border variables. **NCBORDER** is positive and significant in all three equations, indicating that counties along the northern border of the state experience greater than average sales of all lottery games than other counties. For example, a county with the average retail employment situated along the North Carolina border will experience $130 higher per capita instant game sales when compared to its counterpart elsewhere.[6] We hypothesize that this is due to cross-border shopping by North Carolina residents. Just as interesting are the insignificant **GABORDER** coefficients – as suggested by a more competitive environment, counties that border Georgia (which also has a lottery) have sales statistically equivalent to those in non-border counties.

Each of the OLS equations includes a test for spatial correlation. The results in Table 2 indicate that instant games sales are not spatially correlated. However, both fixed games and Lotto sales show significant positive spatial correlation. This implies that these purchases by households in abutting counties are related in some way that is not captured by the regressors included in our estimation.

---

[6] Because our border indicators are weighted by retail sales, regression coefficients do not directly state the effect of being on the state's border. The $130 figure has been backed out of the estimated coefficient by evaluating sales at the mean value of retail employment along the northern border.

The Distributional Effects of Lottery Products

Early research on the economic effects of lotteries concentrated on the regressive nature of the lottery tax (Clotfelter, 1979; Clotfelter and Cook, 1987). In fact, one of the basic criticisms of state-run lotteries is that lottery sales have a regressive distributional impact. Most research has shown lottery products to be regressive in nature (see, for example, Scott and Garen (1994) and Rubenstein and Scafidi (2002)). A more recent study by Oster (2004) suggests that the degree of regressivity of lottery products depends on the size of the prize.

Price and Novak (1999, 2000) examine the level of regressivity for three Texas lottery products: Lotto, Pick 3, and instant games. By examining the coefficient on log income in sales equations and computing a Suits Index (Suits, 1977), they find all three types of lottery products to be regressive. Instant games are shown to be the most regressive; Lotto is shown to be least regressive.

The results from the OLS estimations described above do not imply that the South Carolina Education Lottery games are regressive. In fact, the households with the greatest lottery sales are those in the upper-middle income category (between \$35,000 and \$50,000 in annual household income). To examine this issue more directly, we follow Price and Novak by estimating Suits Indices for the three lottery games sold in South Carolina.

A Suits Index is a common measure used to measure the progressivity or regressivity of various taxes. Similar in nature to a Gini coefficient, the Suits Index measures the relative sizes of cumulative tax burden and cumulative income. The calculation of the Suits Index can best be seen by examining Figures 1 through 3, which show "Lorenz" curves for the three lottery games sold in South Carolina. The horizontal axis in each of these diagrams measures the percentage of cumulative income[7], while the vertical axis measures the percentage of cumulative sales. The Suits Index, which can vary from +1 to -1, is calculated as

(4) $$S = (K - L)/K = 1 - (L/K),$$

where K is the area of the triangle below the diagonal line and L is the area beneath the "Lorenz" curve. When the tax is progressive, area L is smaller than area K, making S positive. A value of S equal to zero means that a tax is proportional, while a value less than zero implies the tax is regressive.

Note that, since the "Lorenz" curves shown in Figures 1 through 3 each generally lie above the diagonal line, the Suits Indices for these three games are all negative. This indicates that each of South Carolina's three types of lottery game are regressive. Fixed-odds online games are the most regressive, with a Suits Index of -0.266 (compared with comparable values of -0.179 for instant games and -0.139 for Lotto).[8] Our results for the relative regressivity of Lotto games versus instant games are comparable to those of Price and Novak (1999, 2000), who find Lotto to be less regressive than instant games.

One point of interest is the "Lorenz" curve for Lotto games. In the middle of the income distribution, the curve actually crosses the diagonal line and appears to be slightly progressive over some levels of income. This result is similar to our regression finding that lower-middle income households appear to buy the fewest lottery tickets, while upper-middle income households purchase the most.

Though the point estimates for each estimated Suits Index are negative, it is unsure to what extent one can rely on the finding of regressivity. To be more precise about the values of the Suits Indices for these three lottery products, 95-percent confidence intervals were estimated

[7] Total income earned in each county is found by multiplying per capita income by county population.

[8] We also calculated a Suits Index for total lottery sales. Its value is -0.186.

following a bootstrapping procedure similar to that outlined in Anderson et al. (2003). The results are presented in Table 3. Each type of lottery game displays a different degree of regressivity, and a different degree of variability with respect to the calculated indices' certainty. Online games are clearly the most regressive, with an upper bound to the 95% confidence interval that is near the point estimates for the other game types. Instant games are less regressive, and Lotto games least regressive of all.[9] Because the potential jackpots offered by Lotto games are generally much larger than the prizes available to players of the other game types, this result may confirm what is suggested by Oster (2004)—that the degree of regressivity may depend positively on the size of the prize. Likewise it may simply reflect the fact that individuals with different characteristics find different types of games appealing.

Conclusions

This study examines the distributional impact of three types of lottery games offered for sale by the South Carolina Education Lottery during 2002 and 2003. We find both consistencies and disparities in the sales of different products. In particular, we find that of the proportion of the population age 65 or older is positively related to sales of instant games, and our analysis points to significantly higher sales of instant and online fixed-odds games to African Americans. There is also evidence of significant cross-border shopping for all three games.

One unique result of this study is that we find each of the three types of game offered by the SCEL to be regressive, as measured by the Suits Index, but find substantial differences in the degree to which each game is regressive. Regression analysis sheds additional light on the regressivity question; we find that the tails of the income distribution purchase about the same

---

[9] Though we reject a null hypothesis of "no regressivity" for Lotto games at the 95% level of confidence, we cannot do so at the 99% level. The 99% confidence interval for Lotto games spans zero; 99% confidence intervals for the other two game types do not.

number of tickets per capita, but we also find significant variation in the middle of the income distribution. Lower-middle-income residents purchase significantly fewer tickets, while upper-middle-income residents purchase significantly more. Together, these pieces of evidence suggest that the lottery may not be as regressive as the body of literature suggests it once was, or that it may not be equally regressive in all places. Finally, our analysis highlights the role that specification may play in reaching conclusions about the lottery's regressivity: when modeling lottery sales with regression analysis, it may be better to account for the distribution of income than simply its average level.

References

Anderson, J. E., A. G. Roy, and P. A. Shoemaker. 2003. Confidence Intervals for the Suits Index. *National Tax Journal* 56(1): 81-90.

Borg, M. O. and P. M. Mason. 1988. The Budgetary Incidence of a Lottery to Support Education. *National Tax Journal* 41(1): 75-85.

Clotfelter, C. T. 1979. On the Regressivity of State-Operated Numbers Games. *National Tax Journal* 32(4): 543-547.

Clotfelter, C. T. and P. J. Cook. 1987. Implicit Taxation in Lottery Finance. *National Tax Journal* 40(4): 533-546.

Clotfelter, C.T. and P.J. Cook. 1990. On the Economics of State Lotteries. *Journal of Economic Perspectives* 4(4): 105-119.

Forrest, D., O.D. Gulley, and R. Simmons. 2004. Substitution between Games in the UK National Lottery. *Applied Economics* 36(7): 645-651.

Garrett, T. A. and T. L. Marsh. 2002. The Revenue Impacts of Cross-Border Lottery Shopping in the Presence of Spatial Autocorrelation. *Regional Science and Urban Economics* 32(4): 501-19.

Garrett, T. A. and R. S. Sobel. 2004. State Lottery Revenue: The Importance of Game Characteristics. *Public Finance Review* 32(3): 313-330.

Ghent, L.S. and A.P. Grant. Forthcoming. Are Voting and Buying Behavior Consistent? An Examination of the South Carolina Education Lottery, *Public Finance Review*.

Giacopassi, D., M. Nichols, and B. Stitt. 2006. Voting for a Lottery. *Public Finance Review* 34(1): 80-100.

Greene, W. 2000. *Econometric Analysis, Fourth Edition*. Upper Saddle River, New Jersey: Prentice Hall, Inc.

Hersch, P. L. and G. S. McDougall. 1989. Do People Put Their Money Where Their Votes Are? The Case of Lottery Tickets. *Southern Economic Journal* 56(1): 32-38.

Jackson, R. 1994. Demand for Lottery Products in Massachusetts. *Journal of Consumer Affairs* 28(2): 313-25.

Oster, E. 2004. Are All Lotteries Regressive? Evidence from the Powerball. *National Tax Journal* 57(2): 179-187.

Price, D.I. and S. E. Novak. 1999. The Tax Incidence of Three Texas Lottery Games: Regressivity, Race, and Education. *National Tax Journal* 52(4): 741-51.

Price, D. I. and S. E. Novak. 2000. The Income Redistribution Effects of Texas State Lottery Games. *Public Finance Review* 28(1): 82-92.

Rubenstein, R. and B. Scafidi. 2002. Who Pays and Who Benefits? Examining the Distributional Consequences of the Georgia Lottery for Education. *National Tax Journal* 55(2): 223-38.

Scott, F. and J. Garen. 1994. Probability of Purchase, Amount of Purchase, and the Demographic Incidence of the Lottery Tax. *Journal of Public Economics* 54(1): 121-43.

Stover, M. E. 1990. Contiguous State Lotteries: Substitutes or Complements? *Journal of Policy Analysis and Management* 9(4): 565-68.

Suits, D.B. 1977. Measurement of Tax Progressivity. *American Economic Review* 67(4): 747-752.

Tosun, M.S. and Skidmore, M. 2004. Interstate Competition and State Lottery Revenues. *National Tax Journal* 57(2): 163-178.

Table 1: Summary Statistics and Data Sources

Variable Name	Variable	Mean	Standard Deviation	Minimum	Maximum
AGE65	% age 65 and older	12.70	1.77	7.9	16.5
BLACK	% Black	37.38	16.39	6.8	71.0
RELIGION	% Evangelical Protestant	41.60	9.68	21.07	60.65
NOHS	% without HS degree	28.55	6.72	12.1	40
EDUCBS	% with college degree	15.57	6.23	8.3	33.2
LOWINC	% income < $15,000	22.70	6.23	11.7	40.4
LMINC	% income $15,000 - $35,000	29.15	1.91	23.9	32.7
UMINC	% income $35,000 - $50,000	17.40	1.60	13.2	20.2
NCBORDER	% retail employment * NC border	2.26	4.39	0	17.76
GABORDER	% retail employment * GA border	1.39	2.88	0	10.99
INSTANTPC	Instant game sales per capita	244.50	152.74	40.65	1119.41
FIXEDPC	Fixed game sales per capita	76.66	56.81	13.97	308.52
LOTTOPC	Lotto ticket sales per capita	60.44	51.37	5.04	231.50

INSTANTPC, FIXEDPC, and LOTTOPC provided by the South Carolina Education Lottery. Religion provided by the American Religion Data Archive. All remaining data provided by the Bureau of the Census.

Table 2: Lottery Demand Estimates

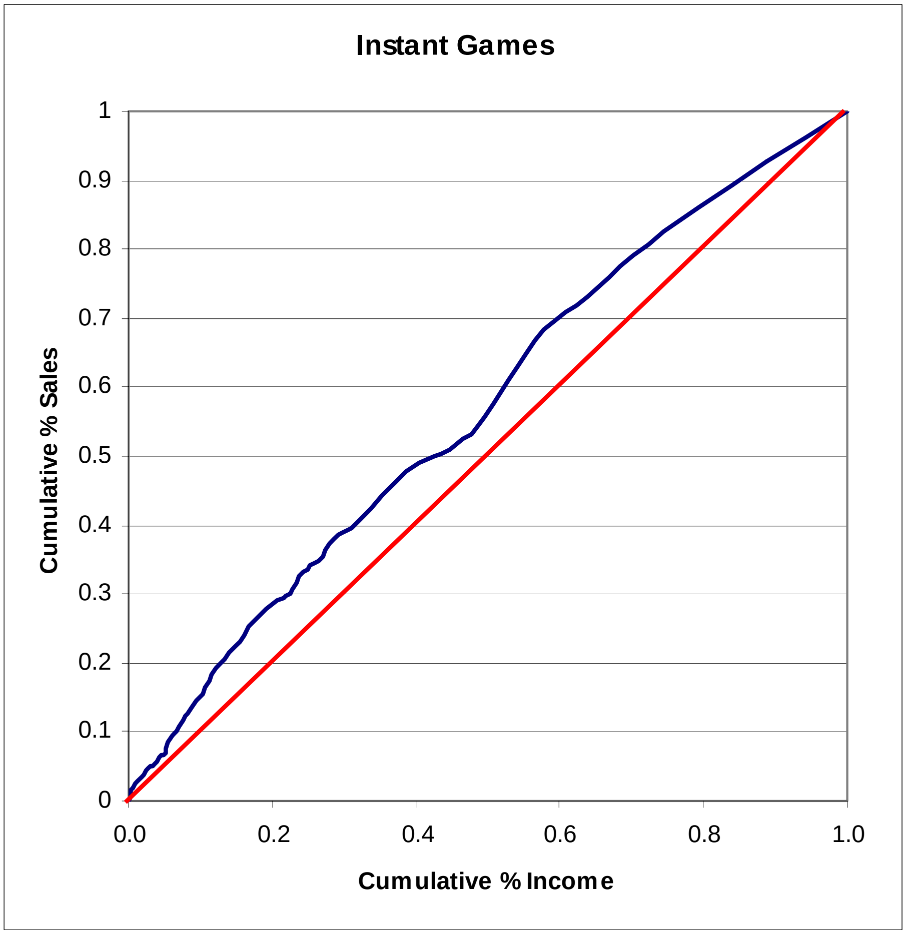
	Coefficients (T-statistics)		
Variable	**INSTANTPC**	**FIXEDPC**	**LOTTOPC**
Demographic Characteristics:			
AGE65	21.187*	5.301	-3.303
	(1.69)	(1.35)	(-0.80)
RELIGION	164.340	67.651	97.726
	(0.57)	(0.47)	(1.03)
BLACK	8.120**	3.185***	1.374
	(2.40)	(2.99)	(1.23)
NOHS	-0.406	-1.260	1.151
	(-0.06)	(-0.57)	(0.50)
EDUCBS	-1.934	-0.788	1.709
	(-0.26)	(-0.33)	(0.69)
Income Measures:			
LOWINC	4.508	2.055	2.423
	(0.50)	(0.73)	(0.81)
LMINC	-43.765***	-15.143***	-16.468***
	(-2.64)	(-2.92)	(-2.42)
UMINC	58.913***	15.122**	11.944*
	(2.85)	(2.30)	(1.75)
Other Variables:			
NCBORDER	1482.275**	405.610*	533.845**
	(2.20)	(1.92)	(2.50)
GABORDER	171.484	211.992	38.344
	(0.19)	(0.75)	(0.13)
ρ	0.272	0.572***	1.011***
	(1.05)	(3.46)	(4.51)
CONSTANT	-306.725	-13.252	96.824**
	(-0.54)	(-0.07)	(0.51)
N	46	46	46
Adjusted R^2	0.3250	0.5546	0.4092

$*\alpha < 0.10$ $**\alpha < 0.05$ $***\alpha < 0.01$

Table 3: Confidence Intervals for Suits Indices

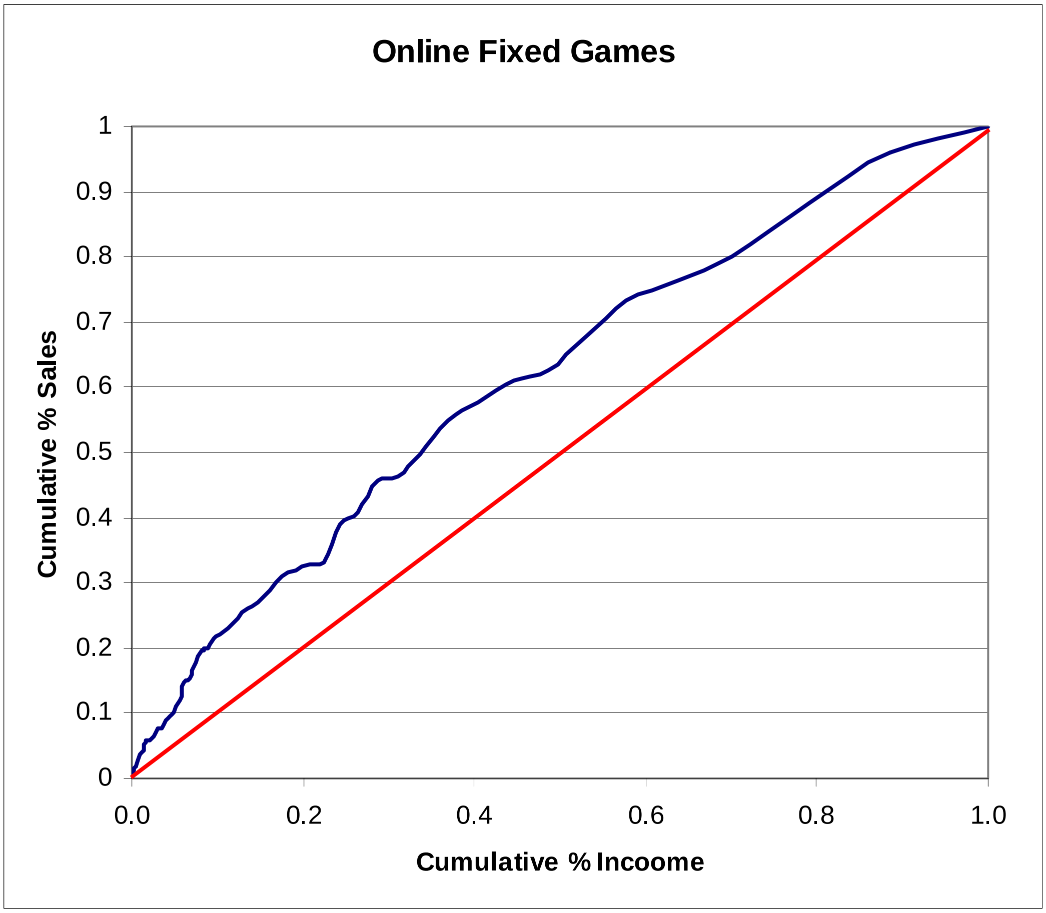
Game Type	Point Estimate	95% Lower Bound	95% Upper Bound
All Games	-0.186	-0.265	-0.097
Lotto	-0.139	-0.266	-0.022
Online	-0.266	-0.367	-0.129
Instant	-0.179	-0.234	-0.105

Figure 1: “Lorenz” Curve for Instant Lottery Games



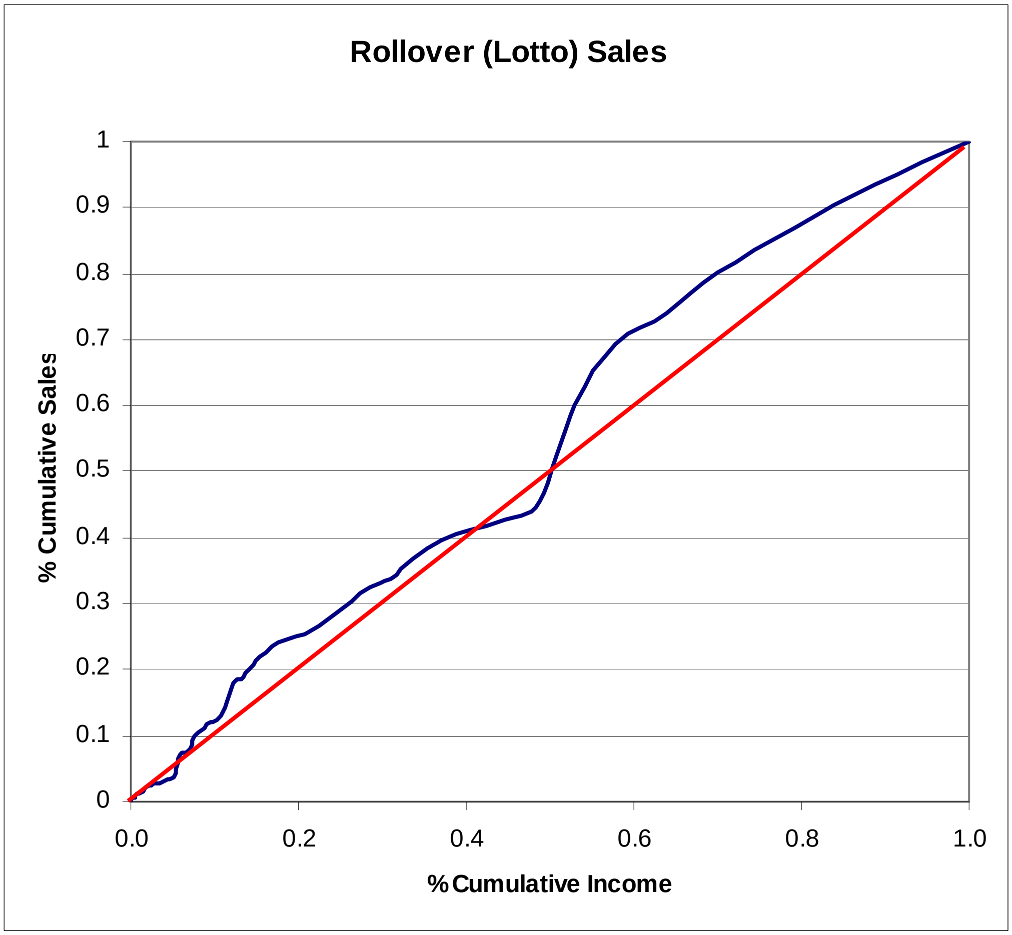
Suits Index = -0.179

Figure 2: "Lorenz" Curve for Fixed Lottery Games



Suits Index = -0.266

Figure 3: "Lorenz" Curve for Lotto Games



Suits Index = -0.139