

NBER WORKING PAPER SERIES

ESTIMATING THE EFFECT OF UNEARNED  
INCOME ON LABOR SUPPLY, EARNINGS,  
SAVINGS, AND CONSUMPTION: EVIDENCE  
FROM A SURVEY OF LOTTERY PLAYERS

Guido W. Imbens  
Donald B. Rubin  
Bruce Sacerdote

Working Paper 7001  
<http://www.nber.org/papers/w7001>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
March 1999

We are grateful for support by the National Science Foundation through grants SBR 9423018 and SBR 9812057, a Sloan Fellowship, and for support and cooperation from the Massachusetts State Lottery commission in general, and Deirdre Coyle in particular. We are also grateful for comments by Joshua Angrist, Gary Chamberlain, Kei Hirano, Larry Katz, Alan Krueger, and participants in seminars at MIT, Princeton University, Columbia University, and the NBER summer institute. Finally we wish to acknowledge superb and tireless research assistance from David Grossman, Aditi Shrikhande, and Eduardo Fajnzylber. We alone are responsible for errors. The views expressed in this paper are those of the authors and do not reflect those of the National Bureau of Economic Research.

© 1999 by Guido W. Imbens, Donald B. Rubin, and Bruce Sacerdote. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Estimating the Effects of Unearned Income on Labor  
Supply, Earnings, Savings and Consumption:  
Evidence from a Survey of Lottery Players  
Guido W. Imbens, Donald B. Rubin, and Bruce Sacerdote  
NBER Working Paper No. 7001  
March 1999  
JEL No. F1, D23, O17

### **ABSTRACT**

Knowledge of the effect of unearned income on economic behavior of individuals in general, and on labor supply in particular, is of great importance to policy makers. Estimation of income effects, however, is a difficult problem because income is not randomly assigned and exogenous changes in income are difficult to identify. Here we exploit the randomized assignment of large amounts of money over long periods of time through lotteries. We carried out a survey of people who played the lottery in the mid-eighties and estimate the effect of lottery winnings on their subsequent earnings, labor supply, consumption, and savings.

We find that winning a modest prize (\$15,000 per year for twenty years) does not affect labor supply or earnings substantially. Winning such a prize does not considerably reduce savings. Winning a much larger prize (\$80,000 rather than \$15,000 per year) reduces labor supply as measured by hours, as well as participation and social security earnings; elasticities for hours and earnings are around -0.20 and for participation around -0.14. Winning a large versus modest amount also leads to increased expenditures on cars and larger home values, although mortgages values appear to increase by approximately the same amount. Winning \$80,000 increases overall savings, although savings in retirement accounts are not significantly affected.

The results do not vary much by gender, age, or prior employment status. There is some evidence that for those with zero earnings prior to winning the lottery there is a positive effect of winning a small prize on subsequent labor market participation.

Guido W. Imbens  
Department of Economics  
8256 Bunche Hall  
UCLA  
405 Hilgard Ave.  
Los Angeles, CA 90095  
and NBER  
imbens@econ.ucla.edu

Donald B. Rubin  
Department of Statistics  
Science Center 709  
Harvard University  
Cambridge, MA 02138

Bruce Sacerdote  
Department of Economics  
6106 Rockefeller Hall  
Dartmouth College  
Hanover, NH 03755

## 1. INTRODUCTION

Knowledge of the effect of income on economic behavior in general, and on labor supply in particular, is of great importance to policy makers. For example, in his introduction to a discussion of the negative income tax experiments, Morrill, Assistant Secretary for Planning and Evaluation in the Department of Health, Education and Welfare during the Nixon administration, wrote regarding the debate over effects of extending cash assistance to the working poor: “Central to this debate has been the question of labor–supply of such families. Would the receipt of assistance payments cause them to work less, or in some cases, quit work altogether?” (Morrill, 1974, p. 156). Estimation of income effects, however, is complicated by the fact that realistic amounts of income are almost never randomly assigned and exogenous changes in income are difficult to identify. In practice, researchers have often taken spousal or property income as exogenous for the purposes of estimating income effects.

In this paper we address the problem of identifying exogenous variation in unearned income by exploiting the randomized assignment of large amounts of money over long periods of time through lotteries. We surveyed individuals who played the lottery in Massachusetts in the mid–eighties, including both winners of large prizes and people who won small, one-time prizes. We investigate the relationship between the amount won and economic behavior as measured by subsequent earnings, labor supply, consumption and savings.

Complicating our analysis is the finding of differences by size of prize pre-dating the time of winning. Ruling out non-randomness of the lottery itself, we consider three sources for these differences. First, differences in the number of tickets bought are correlated with the likelihood of winning a prize and are also likely to be correlated with economic behavior. Second, our control sample of small, one-time prize winners consists solely of season-ticket holders whereas our sample of winners contains both season-ticket holders and single ticket buyers. These groups may differ in economic behavior leading to differences between the two groups preceding winning the lottery. Third, non-response to the survey may be correlated

with both the amount of prize money and with economic behavior. We attempt to adjust for these differences using a variety of statistical methods including ordinary least squares and propensity score methods. We provide some evidence that suggests these methods in combination suffice to remove biases due to pre-lottery differences.

We find that winning a modest prize (\$15,000 per year for twenty years) does not affect labor supply or earnings substantially. Winning such a prize does considerably reduce savings in a number of categories. Winning a much larger prize (\$80,000 rather than \$15,000 per year), reduces labor supply as measured by hours, as well as participation in the labor force and social security earnings; elasticities for hours and earnings are around -0.20, and for participation around -0.14. Winning a large amount relative to a modest amount leads to increased expenditures on cars, and larger home values, although mortgages values appear to increase by approximately the same amount. Winning \$80,000 also increases overall savings, although savings in retirement accounts are not significantly affected.

The results do not vary much by gender, age or prior employment status. There is some evidence that for those with zero earnings prior to winning the lottery there is a positive effect of winning a small prize on subsequent labor force participation.

Based on the statistical evidence we interpret the results as estimates of the causal effect of lottery prizes on labor supply, earnings, savings and consumption. Two caveats should be kept in mind, however. Responses to lottery prizes need not be typical of responses to other forms of unearned income. Of primary interest, as indicated in the first paragraph in this section, is the response of individuals to government-provided cash assistance. In general responses to unearned income may well differ by the source of the unearned income – what Thaler (1990) refers to as fungibility. It is likely, however, that the response to lottery prizes is indicative of the response to other types of unearned income such as cash assistance. In addition, lottery players need not be representative of the population of interest. In buying lottery tickets, they reveal an attitude to risk that differs from that of the population at large. However, it is again likely that their response is at least suggestive of the response of

the general population.

## 2. LITERATURE

There is a large literature concerned with estimating the effect of unearned earnings on labor supply. See Pencavel (1986), and Blundell and MaCurdy (1998) for surveys for men, and Killingsworth and Heckman (1986) for women. Most of the literature utilized data from large surveys such as the Panel Study of Income Dynamics (PSID), the National Longitudinal Survey (NLS), or the Current Population Survey (CPS). A major theme of this literature is the difficulty in constructing exogenous measures of unearned income. Often capital income, or spousal labor income is used. In all cases the assumption that this source of unearned income is exogenous to labor supply decisions of the individual is tenuous at best.

Another strand of the literature on estimation of income effects has analyzed experimental data with clearly exogenous determinants of unearned income. In the early seventies, several negative income tax experiments (NIT) were conducted where selected individuals received randomly assigned tax schedules characterized by a guarantee level combined with a tax rate. See for example Ashenfelter (1978), Rees (1974), Bishop (1980), Stafford (1985), Moffitt and Kehrer (1981), Watts et al (1974), and Hausman and Wise (1977). The NIT experiments were limited in the duration of the income supplement, ranging from three to five years. It is therefore possible that responses to the different tax regimes do not represent long run responses to a permanent change in regime. In fact, this limited duration has been pointed to as a reason for the findings of small or no effects on labor supply for men. In addition, the amounts of income randomly assigned were relatively modest. In contrast, the lottery pays out substantial prizes over extended periods of time.

A third strand of the literature consists of a number of case studies in which large amounts of money were allocated using distribution rules that were arguably independent of preferences and other determinants of economic behavior. Examples of these so-called natural experiments (see Angrist and Krueger (1998) for a survey of this literature) are

Kreinin (1961) and Landsberger (1963, 1970), who looked at one-time war reparations paid to Israeli citizens by the German government, Bodkin (1959, 1963), who looked at one-time payments by the US government to selected service men after World War II, and Holtz-Eakin, Joulfaian and Rosen (1993) who looked at the effects of inheritances on employment. In all these cases, the exogeneity of the income supplements, that is, their independence of preferences and constraints, is still subject to debate, but much more plausible than the assumption of exogeneity of, for example, spousal labor income or capital income.

Finally, as in the current paper, Kaplan (1985) analyzes a survey of lottery winners. Kaplan, however, only collected data on economic behavior immediately prior, and immediately subsequent, to the lottery winning, and he is therefore unable to estimate long term effects whereas we have twenty years of accurate earnings data from the social security administration. In addition, Kaplan only surveyed people who won amounts large enough that the prizes were paid in yearly sums, whereas we also have a control group of small, one-time prize winners. Finally, as in many of the other natural experiment studies referenced above, Kaplan does not have detailed data regarding the individuals' economic circumstances and behavior prior to the infusion of income. Given that his response rate is considerably lower than ours (approximately 30% versus 50% in our sample), and given our results, it appears likely that these shortcomings bias his results.

### 3. THE SURVEY

Our data set consists of two samples, the “winners” sample and the “losers” sample. The relevant population for the winners sample consists of people playing the Megabucks lottery in Massachusetts during the years 1984 through 1988 and winning a major prize. Major prizes for the purposes of this study are prizes that are paid out in yearly installments over twenty years. They range from \$22,000 to \$9,696,000, with the sample mean and median equal to \$1,104,000 and \$635,000 respectively.<sup>2</sup> One implication of the yearly payment

---

<sup>2</sup>All prizes are in 1986 dollars.

feature of the lottery is that the lottery administration has a very accurate record of the current address of these people, which enabled us to send out surveys to essentially all people in this population.<sup>3</sup>

The “losers” sample comes from the population of season ticket holders between 1984 and 1988 who have won at least one small, one-time prize, ranging from from \$100 to \$5,000; we do not know for these individuals the exact amount of the prize.<sup>4</sup> The people in this sample are referred to as the “losers”, although it should be stressed that they did actually win small, one-time, prizes.

The survey questionnaire, provided in Appendix A, consists of three sets of questions. The first set concerns the current (time of survey) situation of the respondent and his or her household. These include questions regarding the labor market status of the respondent and spouse, their financial assets, their housing situation and expenses on their children’s education. Because of the findings in the NIT’s concerning the increase in divorce rates as a result of the payments, we also inquired about marital status changes since the date of winning the lottery. These are the outcome variables of primary interest.

Second, there are a number of questions regarding the situation at the time they won their prize in the lottery (the big prize for the winners, and the small, one-time prize for the losers). In principle, there is no substantive interest in the answers to these pre-lottery questions. The reason for including them in the survey is threefold. First, and most important, they assist in making the inferences more credible. Although in principle the randomization ensures that the different groups of winners and losers are comparable at the time of playing the lottery, in other words that the distribution of pre-lottery variables should be independent of the amount won, for the aforementioned reasons of differential ticket buying behavior, the season-ticket/single-ticket buyer differences, and differential nonresponse, there may actually

---

<sup>3</sup>See Sacerdote (1996) for more detail on the lottery and the survey.

<sup>4</sup>The lottery does not have historical records for people winning small prizes who bought single tickets, or for season ticket holders who did not win anything.

be differences between the various groups predating the prize. Similar concerns, in that case primarily due to nonresponse, arose in the analysis of the NIT experiments (Hausman and Wise, 1979, 1985). The variables describing the situation of the respondent at the time of playing the lottery can be used to help adjust for such differences, and allow us to evaluate the procedures used for these adjustments. The second reason for including the questions concerning the individual's situation at the time of winning is to investigate the heterogeneity of the income effects by gender, pre-lottery labor market status, and age. Finally, the inclusion of control variables or regressors can improve the precision of the estimates, just as in a randomized experiment.

Third, we asked respondents to authorize the release of their social security earnings records to us. We deliberately chose to ask for social security earnings records, rather than inquire about earnings from the respondents directly, in order to get more accurate records over a longer period of time. For those who signed the social security release forms we have accurate earnings records for at least six years preceding and six years following the time of winning. Although we did include questions concerning spousal labor market status both current and at the time of lottery playing, we did not ask for spousal earnings information, out of concern for the response rates.

The survey was conducted in three stages. In July 1995, we sent out by regular mail pilot surveys to 50 winners and 50 losers to assess response rates and various approaches to increasing them. In July 1996 we sent out, again by regular mail, surveys to 752 winners and 637 losers. Finally, in September 1996 we send out reminders to 297 nonresponding winners and 297 nonresponding losers. The reminders were sent by Federal Express to increase the likelihood of potential respondents paying attention to the survey. In the pilot survey and the main mailing, respondents were offered the choice between lottery tickets with a nominal cost of 100 dollars or gift certificates in major department stores with a nominal cost of 50 dollars. In the follow-up part of the survey, 49 winners and 49 losers received with the survey a 10 dollar note, and were offered a check for an additional 40 dollars in exchange for

returning the survey. The other 248 winners and 248 losers approached in the follow-up were offered a check for 50 dollars for returning the survey. Incentive schemes where potential respondents are paid prior to responding have been previously implemented in Philipson (1997), who discusses the merits of such schemes in detail.

Table 1 summarizes the response rates for the different mailings. The overall response rate is approximately 46%, slightly higher for losers at 49% than for winners at 42%. It should be noted, however, that the follow-up mailing did not include all nonrespondents from the previous mailing for budgetary reasons. Had we followed up on all nonrespondents in the main mailing using the 10-dollar-cash/40-dollar-check incentive scheme, the expected overall response would have been  $(0.38 + (1 - 0.38) \times 0.23) \times 100\% = 53\%$ . Consistent with Philipson's (1997) findings, the incentive scheme with \$10 up front and a promise of \$40 more rather than a promise of \$50 did lead to a higher response rate (23% versus 16%). A test of the null hypothesis that the two response rates are equal gives a t-statistic of 1.81, with a p-value of 0.08.

#### 4. THE DATA

We divide the sample into three groups. First, the "losers", those who won a one-time prize between 100 and 5,000 dollars. Second, the "small winners", who won a total amount less than 500,000 dollars (in 1986 dollars), or a yearly amount less than 25,000 dollars for twenty years: on average the people in this group won yearly prizes of 15,000 dollars (with a standard deviation of 6,000 dollars). Third, the "big winners", who won more than 500,000 total, that is, more than 25,000 yearly for twenty years: on average this group won approximately 80,000 dollars per year (with a standard deviation of 70,000 dollars).<sup>5</sup> We shall, however, also carry out analyses using the exact value of the prize using the entire sample of winners.

---

<sup>5</sup>The cutoff point for the small vs big winners sample was chosen so the two groups were of comparable size.

The reasons for dividing the winners sample into two subsamples are twofold. First, for policy purposes, the big winners are of less interest. Few social policies involve income transfers of this magnitude. The comparison of the small winners and the losers, on the other hand, is potentially relevant for social policy because it involves yearly transfers that are not out of line with existing income transfer programs. Figures 1 through 3 illustrate this. The first figure presents a histogram of the yearly prize for all winners. For this group the yearly prize ranges from close to zero to \$500,000. Figure 2 restricts the sample to small winners. Figure 3 presents a histogram of social security earnings in the year prior to winning the lottery for the subpopulation for whom these earnings were positive. For this group the average earnings were \$22,000 (\$16,000 if we average also over those with zero earnings), comparable to the \$15,000 average yearly prize for the small winners. The second reason involves the sample selection. In the introduction three sources of association between prizes and pre-lottery variables were discussed. The first and second of these were differences in number of tickets bought, and differences between the population of season ticket holders from whom the losers were drawn and the population of winners, which includes both one-time ticket buyers and season ticket holders, respectively. These differences do not affect the comparison between different groups of winners, and by comparing big and small winners we therefore eliminate both these complications and need only be concerned with biases from differential non-response by prize.

Our basic sample for the analyses presented below consists of individuals with complete answers to the questions on pre-lottery conditions (that is, number of tickets bought, age, years of high school, years of college, gender, whether the individual was working at the time of playing the lottery), and who authorized the release of their social security earnings. This leaves us with a sample of 496 observations, 259 losers, 93 small winners and 144 big winners. For analyses here involving additional variables (e.g., weekly hours, or savings) we select subsamples of this basic sample with complete answers to the questions regarding the additional variables involved. In doing so we discarded individuals who responded to some

of the questions, and therefore possibly introduced biases or at least lost some precision. In future work we intend to follow a more principled approach to missing data involving modelling the nonresponse and multiply imputing the missing data (Little and Rubin, 1987; Rubin, 1987).

Tables 2 and 3 present summary statistics for the variables used in the analyses. For each variable the mean and standard deviation for the entire sample are given. In addition, we report the mean for the three groups (losers, small winners, and big winners) as well as t-statistics for the null hypotheses that the averages for the loser and small winner subpopulations and the small and big winner subpopulations are identical.

Overall there are substantial differences between the three groups, even prior to winning the lottery. This highlights the need for the statistical adjustment procedures described in Section 5. It also underlines the importance of collecting detailed pre-treatment information when designing studies such as ours. Although the source of variation in the treatment of interest appears completely random, complications arising from, for example, non-response can invalidate simple comparisons.

#### 4.1 PRE-LOTTERY CHARACTERISTICS

On average the individuals in our basic sample are 50 years old at the time of winning, which, for the average person was in 1986. They have 3.7 years of high school and 2.0 years of college. Of the 496 individuals 63% are male, and 78% were working at the time they won the lottery. In principle these characteristics should not differ between the three groups, losers, small and big winners. However, losers are significantly more educated than both groups of winners, and they are also older. This is likely to reflect the differences between season ticket holders and single ticket buyers. The fact that the differences between losers and small winners are much larger than those between small and big winners suggests that differential nonresponse is not as important a source of bias as the difference between season-ticket holders and single ticket buyers. A final question regarding the situation at the time of winning the lottery is the number of tickets bought in a typical week. Because there were

some extremely large numbers (up to 200 tickets per week), we transformed this variable by taking the minimum of the number reported and ten. As expected, the number of tickets bought is considerably higher for winners (five per week for the transformed variable, nine per week for the raw measure) than for losers (two per week according to both the raw measure and the transformed variable).

#### 4.2 LABOR SUPPLY AND EARNINGS

We observe for each individual in the basic sample social security earnings for six years preceeding and following winning. Average earnings, in terms of 1986 dollars, rise over the pre-winning period from \$13,930 to \$16,330, and the decline back to \$13,290. For those with positive social security earnings, average earnings rise over this period from \$20,180 to \$24,300. From the 496 individuals in our basic sample, 53% claimed to be currently working ("what were you doing most of last week?"). We also asked for hours worked ("in a typical week, how many hours per week do you currently work"). The average number of hours was 30.6, or 40.3 for the 76% with positive hours. Note that a higher percentage claimed positive hours ("in a typical week") than claimed to be working ("most of last week"). There was, in fact, no-one who claimed to be working while reporting zero hours, but 12% claimed positive hours in a typical week, while at the same time professing not to be working most of last week. This group reported an average of 9.3 hours, compared to an average of 43.5 hours for those who reported to be working. Regarding their spouse at the time of the survey 40% claimed to have a spouse currently working.

#### 4.3 CONSUMPTION AND SAVINGS

We included questions about housing and cars in the questionnaire. In both cases we asked for the value of the houses and cars, as well as the amount of debt associated with them. On average the value of all cars was \$18,200, with an average debt of \$3,000. For housing the average value was \$166,300, with an average mortgage of \$48,800. Note that this is averaged over the entire sample, with zeros included for the 7% respondents who reported not owning their homes. Only 12 respondents out of the basic sample of 496 did not reply

to the housing question, so it is unlikely that the high percentage of homeowners is due to confusing not owning with non-response.

We asked a number of questions concerning current wealth. We aggregated these into three categories. The first concerns retirement type accounts, including IRA's, 401K plans and other retirement related savings. The second consists of stocks, bonds and mutual funds. The third savings category is savings in general savings accounts and CD's.<sup>6</sup> We construct an additional variable adding up the three savings categories as "total financial wealth".<sup>7</sup> For the losers, wealth in the various savings accounts is considerably higher than wealth in housing, \$176,000 versus \$144,000. The distributions of these financial wealth variables are very skewed, with for example wealth in mutual funds for the 414 respondents ranging from zero to \$1.75 million, with a mean of \$53,000, a median of \$10,000 and 35% zeros. In the analyses below we therefore transform the four savings categories as  $s = \log(S + 1)$  where  $S$  is savings in 1000's of 1986 dollars.<sup>8</sup>

Finally, we asked respondents for savings rates out of "your and your spouse's total after-tax income" both actual for the current year and planned for the next year.<sup>9</sup> Average reported savings rates are 11.8% for the current and 13.0% for the next year. Thus on average respondents planned to save 1.2 percentage points more next year than this year, significantly different from zero with a t-statistic of 4.8. Sixteen percent intended to increase savings rates for next year, against only two percent who planned to reduce savings. We

---

<sup>6</sup>See the appendix with the questionnaire for the exact formulation of the questions.

<sup>7</sup>To reduce the effect of item-nonresponse for this last variable, total financial wealth, we added zeros to all missing savings categories for those people who reported positive savings for at least one of the categories. That is, if someone reports positive savings in the category "savings accounts and CD's", but did not answer the question for mutual funds, we impute a zero for mutual funds in the construction of total financial wealth. For the 462 observations on total financial wealth, zeros were imputed for 27 individuals for retirement savings, for 48 individuals for mutual funds, and for 27 individuals for general savings. As a result, the average of the three savings categories does not add up to the average of total savings, and the number of observations for the total savings variable is larger than that for each of the three savings categories.

<sup>8</sup>An alternative would be to model separately the probability of savings being equal to zero and the distribution of savings conditional on it being positive.

<sup>9</sup>The question about next year's savings plans was suggested to us by David Laibson.

generally interpret these intended savings rates as a less than perfect measure of actual future savings rates.

## 5. SELECTION BIAS AND ANALYSES

Here we briefly describe the analyses carried out on this rich data set and the selection issues that necessitate them. For a large number of outcomes (e.g., earnings, hours worked), we are interested in the average response as a function of the yearly prize. If prizes had been randomly assigned to the people in our sample, simple comparisons of outcomes by prize would have provided unbiased estimates of these functions. However, such simple comparisons are unlikely to lead to credible estimates of the causal effect of winning a prize in light of the pre-lottery differences documented in Tables 2 and 3.

### 5.1 SELECTION BIAS

The first reason that the simple comparison by prize won does not lead to credible estimates of the average causal effect of winning the lottery is that the randomization is over lottery tickets and not over the population of interest. Even if everybody in the population bought at most one lottery ticket, the randomization would no longer be over the population at large, but it would still correspond to randomization over a subpopulation, namely that of lottery players. However, different people buy different numbers of lottery tickets. Suppose for example that there are two types of lottery players. The first type buys a single lottery ticket every single drawing. The second type buys two tickets every drawing. If we randomly draw from those tickets that won a small prize, the second type of person would be twice as likely to be drawn. This effect is likely to explain most of the statistically highly significant differences in number of tickets bought between losers and winners, as presented in Table 2. There is also some evidence that the number of tickets bought is correlated with characteristics of the individuals. For example, the correlation between years of college and the number of tickets bought is  $-0.20$ , with a  $t$ -statistic for the null of zero correlation equal to  $-4.5$ . Within the sample of winners the correlation is smaller at  $-0.06$ , with a  $t$ -statistic of

-1.0.

This effect does not explain any differences in the number of lottery tickets bought by prize for winners. If the probability of winning is equal for all lottery tickets, and if the probability of winning two prizes that are paid out over twenty years is ignored, then conditional on winning such a prize, the amount of the prize should not be correlated with the number of tickets bought. As shown in Table 2, we do not find a significant difference between small and big winners.

However, in principle there may be more complicated ways in which ticket buying behavior biases simple comparisons. As in many lotteries, the payout rate in Megabucks varies from drawing to drawing depending on previous payouts. There is much anecdotal evidence that the prospect of an exceptionally large prize draws additional lottery players. In the end this effect may lead to a positive or negative correlation between the number of tickets bought and the payout rate and thus with the prize. It is unlikely, however, that such effects are substantial.

The second reason for biases is the difference in the population from which the losers and the population from which the winners were drawn. The former is the population of season ticket buyers, whereas the latter includes both single ticket and season ticket buyers.<sup>10</sup> This difference may explain much of the difference in background characteristics in Tables 2 and 3. For example, we find that losers are significantly older, more highly educated, and earn more money than winners. This source of bias cannot explain any of the difference between small and big winners as both come from the same population.

The final source of bias is the differential nonresponse. If the likelihood of responding to the survey depends on the amount won and on pre-lottery variables that are correlated with post-lottery outcomes, simple comparisons of post-lottery outcomes by prize would not lead to unbiased estimates of the average causal effects even in the winners sample. There are

---

<sup>10</sup>We do not know whether a winner is a season ticket buyer or a single ticket buyer. We do know that most of the winners are single ticket buyers, although we do not know the exact mix.

two pieces of evidence supporting the hypothesis that differential nonresponse is potentially an important problem. First, we do find significant differences predating the lottery by prize within the winners sample that cannot be attributed to the previously discussed biases. Second, since we do know for all individuals, respondents or non-respondents, the prize won, we can directly investigate the correlation between response and prize. Such a non-zero correlation is a necessary condition for nonresponse to lead to bias. The t-statistic for the slope coefficient in a logistic regression of response on the logarithm of the yearly prize is -3.5 (the response rate goes down with the prize), lending credence to this argument.

## 5.2 ANALYSES

The key assumption we make to account for the selection biases discussed in the previous section is the *unconfoundedness* assumption (Rubin, 1978; Rosenbaum and Rubin, 1983; Holland, 1986), related to the assumption of *selection on observables* (e.g., Heckman and Robb, 1984) that implies that, within subpopulations defined by pre-lottery variables, differences in outcomes by prize are unbiased estimates of the average causal effect of the prize for that subpopulation. The plausibility of this assumption clearly depends on the richness of the set of observed pre-lottery variables. In Appendix C we provide some evidence on the adequacy of this assumption. Given the unconfoundedness assumption, we would like to, in principle, report estimates of the entire response function, adjusting for all differences in pre-lottery variables by calculating averages within cells defined by the the pre-lottery variables. This is obviously not feasible, given the limited sample size and the number of conditioning variables. We therefore focus on two sets of readily interpretable summary measures. First, we estimate the effect of a pair of discrete changes in the prize, first the effect of a change from a yearly prize of zero to 15,000 dollars and second the effect of a change from 15,000 to 80,000 dollars. To avoid reliance on a single method of inference, and to assess robustness of the results, we estimate these two effects in a variety of ways. Second, we report the elasticities implied by these estimates.

For the effect of a change from zero to 15,000 dollars we use the sample of 259 losers and

93 small winners for a sample size of  $N = 352$ . First we estimate this effect using a dummy variable for winners using ordinary least squares regression. Since the small winners win on average 15,000 dollars a year, we interpret the estimated coefficient on the dummy variable as an estimate of the effect of winning 15,000 dollars. Second, we use propensity score methods (See, e.g., Rubin and Rosenbaum, 1983, 1985; Dehejia and Wahba, 1998) to adjust for pre-lottery differences. In Appendix B we discuss the implementation of the propensity score methodology in more detail. The propensity score estimates tend to have less reported precision than the least squares estimates, reflecting both the reliance on linearity of the ordinary least squares methods, and the imbalance between the different groups. Our third set of estimates are based on a linear regression with the yearly prize as a regressor, and the pre-lottery variables as additional regressors. We report the coefficient on the prize, multiplied by 15,000 to get the effect of a 15,000 dollar increase in prize.

To convert the estimates of the response to a discrete change from a prize of zero to 15,000 dollars into an elasticity we first divide by the difference in average prizes (15,000 for small winners versus 0 for losers) to get an estimate of the effect per dollar, and then multiply by the ratio of average prize (6,600 for the small-winners/losers sample) to average response to get an estimate of the percentage change in the response to a percentage change in the yearly prize.

The covariates for the analyses involving losers and small winners consist of the following nineteen variables defined at the time of winning the lottery: number of tickets bought (as before, with numbers larger than ten replaced with ten), age at time of winning, gender, years of high school, years of college, year of winning, labor market status at the time of winning, earnings for each of the six years preceeding the year of winning and indicators for positive earnings for each of those six years.

For estimating the effect of a change from 15,000 to 80,000 dollars, we use the winners sample (93 small winners and 144 big winners for a total sample of  $N = 237$ ). Again we estimate the effect of interest using a dummy variable regression, as well as by regressing the

response on the actual size of the prize. The distributions of pre-lottery variables are much closer for these samples, as shown in the summary statistics in Tables 2 and 3, and differences between least squares and propensity score estimates are therefore uniformly small, both in estimates and in standard errors, and we only report the propensity score estimates for the earnings variables in graphical form and the appendix. Because in this sample all prizes are positive, we also include a regression using the logarithm of the prize as the regressor. In each case we report both the estimated effect and the corresponding elasticity.

The control variables for the estimate based on the winners sample are the same as those from the small-winners/losers sample with the exception that only earnings in the one year immediately prior to winning is used, both in level and as an indicator for these earnings being positive, leaving us with nine pre-lottery variables. We dropped earnings from the five additional pre-lottery years because of the smaller sample size and because the differences between the different groups of winners are smaller than the differences between small winners and losers. Appendix C reports some evidence to support the robustness of the estimates to the inclusion of these variables.

None of these analyses take account of the specific nature of the outcome variables. Many of the outcome variables are binary, and others have very skewed distributions, often with many zeros. Analyses taking such features into account may lead to more precise inferences and we intend to pursue them in future work.

We also report elasticities for various subpopulations. These are based on linear regressions with the amount of the prize won and interactions with various variables. For example, to estimate the elasticity for men and women, we estimate a linear regression of the dependent variable with on the right hand side: the level of the prize, the same pre-lottery variables as before (which includes a dummy for men), and an interaction of the male dummy and the level of the prize. From the estimated regression function we calculate elasticities for men and women, evaluated at the same average prize and response. We base these estimates on the linear regressions with interactions rather than on separate analyses for the different

subpopulations to increase precision.

Two caveats should be borne in mind when interpreting the estimates. First, there may be differences between the elasticity estimates from the small-winner/loser and winners samples. These may be due to the fact that the two samples are subject to different biases, or because the two elasticities are estimated on very different parts of the response function. In particular it is difficult to imagine that an increase in the prize from 1,000 to 2,000 dollars per year would lead to the same proportional response as an increase from 50,000 to 100,000 dollars per year. Even if the elasticity were constant as a function of *all* unearned income, it is likely that lottery winnings are a higher proportion of unearned income for the big winners, leading to a lower estimate of the elasticity for the small-winner/loser sample. A second issue concerns the comparison of estimates of the same effect or elasticity estimated on the same sample using different estimation methods. Consider, for example, the estimates of the effect of a discrete change based on a dummy variable regression or on a linear regression with the actual prize. The results may differ because of nonlinearity of the response function, or because of the differential susceptibility to outliers. Here it is useful to recall that the distribution of yearly prizes is highly skewed. The yearly payout ranges from 1,000 to 500,000 dollars per year, with the mean and median equal to 56,000 and 32,000 dollars respectively. The few observations with very large prizes will be very influential in the linear regressions with actual prizes but less so in the dummy variable regressions.

## 6. LABOR SUPPLY AND EARNINGS

Tables 4 and 5 presents results for for the small winners/loser sample and winners sample respectively. The labor supply variables are self-reported weekly hours, and self-reported own and spousal labor market status at the time of the survey. The earnings variables are social security earnings for the year of winning and six years after winning the lottery.

### 6.1 LABOR SUPPLY

In the small winners/losers sample the estimates of the effect of winning 15,000 are an

increase in hours of 0.2 (based on the propensity score estimates), 1.6 (based on the dummy variable regression) and 0.1 hours (based on the yearly prize regression). None is statistically significantly different from zero. The implied estimates of the elasticities are between 0.00 and 0.01, again very close to zero.

In the winners sample a change from a yearly prize of \$15,000 to a yearly prize of \$80,000 is estimated to reduce hours by 9.3, 4.2, and 8.7, based on the dummy variable, yearly prize and log yearly prize regressions respectively. Implied estimates of the elasticities are -0.25, -0.11, and -0.17. All estimates are significantly different from zero with t-statistics around 3.5.

How do these estimates relate to those reported in the literature? Pencavel (1986) and Blundell and MaCurdy (1998) report estimates of the marginal propensity to earn out of nonwage income, or the total income elasticity. This elasticity is equal to the product of the elasticity of hours with respect to unearned income and the ratio of earned and unearned income. If we take the lottery income to be the sole source of unearned income, and ignore the topcoding for social security earnings, we can calculate this as the ratio of average social security earnings and average yearly prize, for the losers/small-winners sample approximately equal to 3.6, and for the winners sample approximately equal to 0.27. It may be that this is somewhat of an underestimate of this ratio, with 17% of the winners at the topcoded value. We find estimates of the elasticity with respect to unearned income, using the winners sample ranging from -0.11 to -0.25. Using the ratio of average earned income to average unearned income of 0.27, this gives an estimate of the total income elasticity of -0.03 to -0.06, with standard errors around 0.01. These estimates are at the low end of, but not out of line with, the estimates reported by Pencavel and Blundell-MaCurdy surveys.<sup>11</sup>

Next consider estimates of the effect of winning on the probability of participation. The first measure of participation is an indicator for positive hours. Based on this measure, the effect of winning \$15,000 is estimated to be between 0.04 and 0.06 respectively, with

---

<sup>11</sup>This is not particularly surprising given the wide range of estimates in the literature.

implied elasticities between 0.01 and 0.02. None of these estimates is statistically significantly different from zero. In the winners sample the effect of winning \$80,000 rather than \$15,000 per year is estimated to reduce the probability of positive hours by -0.17, -0.06, and -0.13 for the three different estimation procedures. The implied elasticities are -0.18, -0.06 and -0.11 respectively. All three estimates are highly significant. In each case the elasticities are about two thirds of the corresponding elasticities for hours, suggesting that a substantial number of people adjust their labor supply to the new source of income by reducing hours without entirely withdrawing from the labor force.<sup>12</sup>

We also investigate the effect on participation through the “what were you doing most of last week” question. For the small-winners/losers sample we find again that none of the estimates are significantly different from zero. The estimates of the effect of a 80,000 rather than a 15,000 dollar prize suggest larger reductions in the probability of working than those based on the hours question. For this sample the elasticities range from -.18 to -.48, all three significantly different from zero.

Finally, consider the effect on spousal labor supply. We find no evidence of any effect of lottery winnings on the labor supply of the spouse. Neither the estimate of the effect of a 15,000 dollar increase, nor the estimate of the effect of an increase from 15,000 to 80,000 dollars is significantly different from zero. The estimated elasticities range from -0.09 to 0.05. The benefits of lottery prizes in terms of increased leisure appear to accrue mainly to the winner of the prize, rather than be shared among members of the household. This is somewhat surprising, given Massachusetts law that would regard lottery prizes as community property and the common presumption that family behavior can be modeled using a single family utility function and budget constraint (e.g., Ashenfelter and Heckman, 1974)

## 6.2 EARNINGS

Using social security earnings records we can trace the effect of lottery winnings from

---

<sup>12</sup>Holtz-Eakin, Joulfaian and Rosen similarly find that inheritances lead to reduction in labor supply for those who do not withdraw entirely from the labor market.

the year in which the prize was won to the sixth year following the prize. In Tables 4 and 5 estimates of the effect on average earnings and on the probability of positive earnings for the two samples are reported. Estimates of the effect of winning 15,000 rather than zero dollars per year on social security earnings vary across estimation methods but are not significantly different from zero. Winning 80,000 rather than 15,000 dollars per year reduces earnings by an amount that is both economically and statistically significant. This reflects both lower participation and lower earnings per worker with positive earnings. For example, the estimated earnings elasticity six years after winning is -0.41 for the estimate based on the dummy variable regression, -0.20 for the regression based on the actual prize and -0.29 for the regression based on the logarithm of the prize. In the sixth year following the prize, the estimated elasticities of participation as measured by positive earnings are -0.30 for the dummy variable regression, -0.16 for the prize regression, and -0.23 for the log prize regression. In each case these elasticities are approximately 75% of the corresponding earnings elasticities.

The elasticities, both for earnings and for participation markedly increase in absolute value as time since winning passes. In the first full year after winning, the estimated elasticities of earnings are -0.22, -0.18, and -0.23, rising (in absolute value) in the fourth year to -0.49, -0.24, and -0.38, suggesting it takes a while for individuals to adjust fully to their new economic circumstances (e.g., Lemieux and MacLeod, 1998).

The estimates of the effect of changes from 0 to \$15,000 and from \$15,000 to \$80,000 are also reported in Figures 4-7. The solid lines present 95% confidence intervals based on propensity score estimates, and the dashed lines present 95% confidence intervals based on least squares estimates. These figures demonstrate the substantial differences in reported precision between the more robust propensity score and more model-dependent least squares estimates for the loser/small-winner sample, and the broad agreement between the two estimates for the winners sample.

Overall it appears that modest changes in unearned income do not affect labor supply

very much, whether directly measured by hours, or indirectly measured by social security earnings. It is only with changes in unearned income larger than most government programs that we see clear and sizeable reductions in hours, participation rates and social security earnings.

## 7. CONSUMPTION AND SAVINGS

Tables 4 and 5 also report results for the consumption and savings variables, for the small-winners/losers and big/small-winners samples respectively.

### 7.1 CARS, CAR DEBTS, HOUSING AND MORTGAGES

Winning \$15,000 rather than \$0 appears to have little effect on expenditures on cars or housing. There is some evidence that debt for cars and houses goes up somewhat, but given that the value of cars and houses does not go up, the implication that net value of cars and houses would actually go down appears unlikely.

Winning \$80,000 rather than \$15,000 affects housing and car values considerably. The value of cars rises sharply, on average by between \$5,500 and \$8,000 dollars, roughly 30% of average car values, depending on the estimator used, with an estimated elasticity of around 0.30. Car debts do not seem to be affected, leaving the net car value to rise by almost the same amount, with again an elasticity of around 0.30. For housing the story is different. The value of houses also goes up, by about \$30,000 on average, out of average house prices around \$166,000, but mortgages also go up, by about \$20,000. Both these changes are highly significant, but the difference, the effect on net housing values, is not significantly different from zero. The estimated elasticities for gross housing values are around 0.16, those for debt around 0.27, and those for net housing values around 0.08.

### 7.2 WEALTH AND SAVINGS RATES

Reported values of savings in various accounts are significantly affected by winning the lottery. Winning \$15,000 rather than \$0 considerably reduces the amount held in retirement accounts, in bonds and mutual funds, and those in general savings. Estimated elasticities

for the total of these three categories are estimated to be between -0.12 and -0.31 depending on the estimator. These reductions are supportive of models that attach importance to the precautionary savings motive. During the twenty years that the respondents are receiving lottery payments, there would appear to be little reason for precautionary savings. Given the average age of the respondents, around fifty years at the time of winning, however, it is not necessarily inconsistent with life-cycle models to find that savings reduce sharply in response to the windfall income. In all cases these reductions in savings are highly significant. In contrast to accumulated savings, reported savings rates are not affected by winning \$15,000 rather than \$0. One possibility is that people do not accurately report savings rates. This is in line with the fact that many report planned savings rates for next year that exceed savings rates for the current year. These planned savings rates increases appear unlikely to be accurate predictions given that not a single individual reported planning a reduction in their savings rate. The change in savings between this year and next year is highly correlated with winning the lottery. If one interprets the planned savings rate as the rate that respondents feel they ought to be saving, this may reflect the respondents' view that they are not saving enough of their lottery windfall.

A very different picture emerges for the comparison between small and big winners. Savings in retirement type accounts still appear to have dropped slightly, although this effect is not significant at conventional levels. However, savings in bonds and mutual funds type accounts, as well as general savings, have gone up markedly. The estimated elasticities for overall savings are between 0.33 and 0.51 depending on the estimator, and highly significant in all cases. Winning \$80,000 rather than \$15,000 also increases reported savings rates, by around 3 percentage points (on an average reported savings rate of around 10%).

## 8. INTERACTIONS WITH GENDER, EMPLOYMENT STATUS AND AGE

In Tables 6-8 we reported elasticities for subpopulations by gender, employment status prior to winning the lottery and age. In each case we estimate the regression with the actual

amount of the prize and the interaction of the actual amount of the prize with gender, employment status, or age, and calculate from those estimates the implied elasticities for the various subpopulations. In addition to the elasticities and their standard errors, we report t-statistics for the null hypothesis that the elasticities for two groups are identical, or, equivalently, the null hypothesis that the coefficient on the interaction term is zero.

### 8.1 INTERACTIONS WITH GENDER

First consider the estimates based on the loser/small-winners sample, reported in Table 7. Most of the differences between elasticities for men and women are statistically insignificant. The exception is in the participation elasticities which are less negative for women than for men. In fact, the estimated elasticities for men are negative in six out of the seven years, whereas for women, the estimated elasticities are positive in six out of the seven years and significantly so in three of the seven years. In contrast, social security earnings do not appear to go up for women as a result of winning \$15,000. One interpretation is that women do not necessarily withdraw from the labor force, but rather find more attractive jobs with less focus on wages.

The reduction in retirement account savings reported in Table 4 for the combined men and women sample appears largely due to reduced savings for women, although the difference with men is not statistically significant. Both men and women intend to increase their savings rate for next year, but the increase is much bigger for men than for women.

Next, consider the estimates based on the winners sample. Both men and women reduce their hours significantly as a result of winning large amounts of money, but the reduction is much larger for women than for men (elasticities of -0.31 and -0.10 respectively), although the difference is not statistically significant. It is interesting in this respect that the elasticities for spousal labor force participation is not significantly different from zero for either men or women. For this sample, elasticities of earnings are larger in absolute value for women than for men, but again the differences are never statistically significant.

## 8.2 INTERACTIONS WITH EMPLOYMENT STATUS

In Table 8 results for the interaction with employment status prior to winning are reported. Employment status is measured here as having positive social security earnings in the year prior to winning.

First consider the losers/small-winners sample. Earnings go down more for those who were employed prior to winning, with the estimated elasticities for those with zero earnings prior to winning essentially equal to zero. These differences are not statistically significant. Participation elasticities, however, differ significantly between those previously employed and those previously not employed. Estimated elasticities for those with zero earnings prior to winning the lottery are positive in every year and significantly different from zero in three of the seven years. For those with positive earnings prior to winning the estimated elasticities are negative in all but one year, and significantly different from zero in one year.

The reductions in savings found in Table 4 appear largely due to reductions in savings by those employed prior to winning the lottery. Reductions in savings are not significantly different from zero in any of the categories for those with prior earnings equal to zero. In fact, the elasticity of the current savings rate with respect to unearned income is positive and significantly different from zero for this group.

Next, consider the winners sample. Those with positive prior earnings reduce their subsequent earnings much more than those with zero prior earnings. This is not surprising because for the latter group there is little to reduce as far as earnings are concerned. Elasticities for savings rates and amounts are generally similar for the two groups, with only savings in the general savings category substantially higher for those with positive prior earnings.

## 8.3 INTERACTIONS WITH AGE

In Table 8 we report elasticities by age. We estimate these using a regression function with an interaction of the amount of the prize and an interaction with age at the time of winning. We then calculate the implied elasticities at two age levels, 40 and 60, motivated by the fact that the average age at the time of winning was approximately 50.

First consider the loser/small-winner sample. None of the estimated elasticities are significantly correlated with age for this sample. The older winners do not appear more likely to drop out of the labor force, or less likely to save, as might have been expected.

Next, consider the estimates based on the winners sample. Older winners are estimated to reduce their earnings and participation less than younger winners, but as in the previous table, this is partly a mechanical result coming from the fact that their earnings are less to begin with. Somewhat more surprising is that younger and older winners are likely to have similarly increased wealth. In fact the elasticities for total wealth are 0.25 at age 40 and 0.37 at age 60. In addition, the elasticities for the savings rate this year are 0.07 at age 40 and 0.24 at age 60, statistically significant at the 10% level.

## 9. CONCLUSION

In this paper we report estimates of the effect of winning the lottery on economic behavior, including labor supply, earnings, consumption and savings. These estimates are based on a new survey of lottery players. We provide some evidence to support our view that these estimates can be interpreted causally. We find that if the amount of winnings per year is around \$15,000, there is little or no evidence that labor supply is reduced as a result. If anything, it appears that those with zero earnings prior to winning the lottery increase participation as a result. Although there is little effect on earnings and labor supply, it does appear that savings are sharply reduced as a consequence of winning around \$15,000, consistent with an important role for the precautionary motive for savings.

Winning a large amount of money does lead to significant reductions in labor supply. In addition, some of the extra income can be traced to increased expenditures on cars and housing, as well as to increased savings. These effects appear to differ little between men and women, or by age.

A question not addressed in this paper is how representative these results are for other populations, or how similar the response to lottery winnings is to the response to other forms

of unearned income. The former question can be addressed by comparing the losers sample to more representative data sets such as the PSID or CPS. It is more difficult to see how the second concern can be addressed with similar amounts of unearned income.

Finally, on a methodological note, our results suggest that even in cases where the treatment of interest is randomly assigned, non-response and other complications can lead to biases in simple comparisons. Even in many of the natural experiments studies, it may therefore still be important to use additional methods to control for differences predating the assignment of treatment.

Figure 1: Histogram of Yearly Prize Won

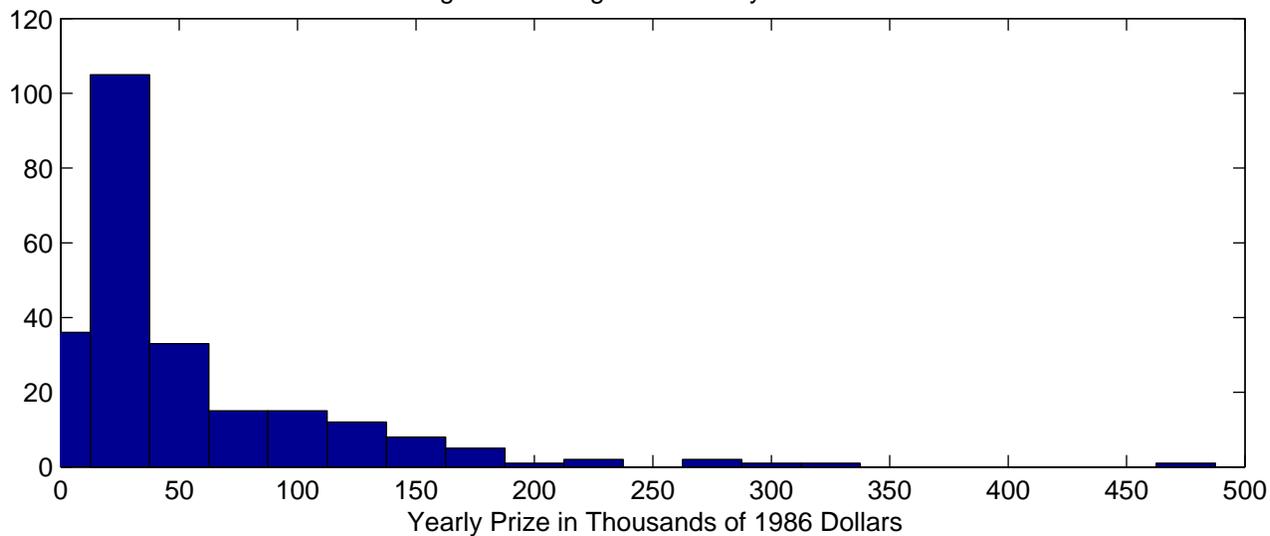


Figure 2: Histogram of Yearly Prize Won for Small Winners

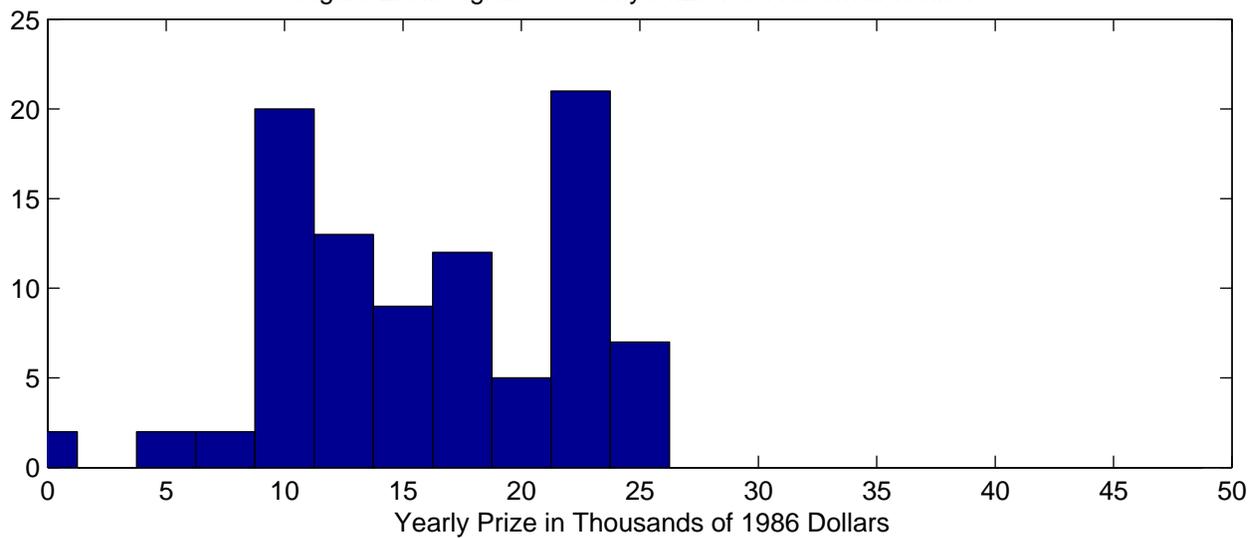


Figure 3: Histogram of Earnings Year -1 for those Positive Earnings

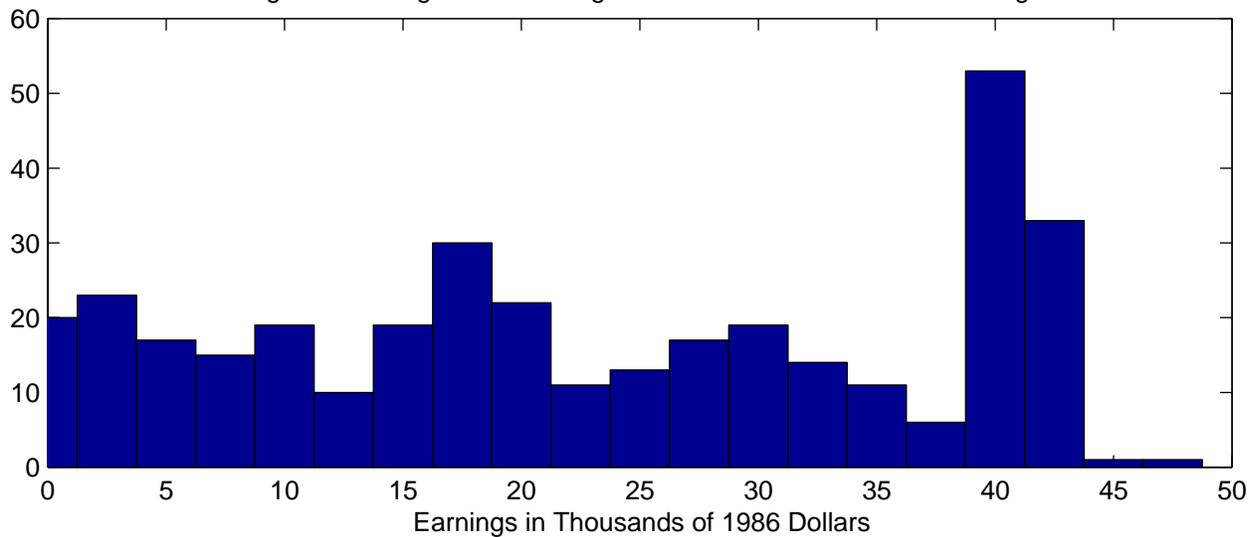


Figure 4: Confidence Intervals for the Effect on Earnings, Small Winners Versus Losers

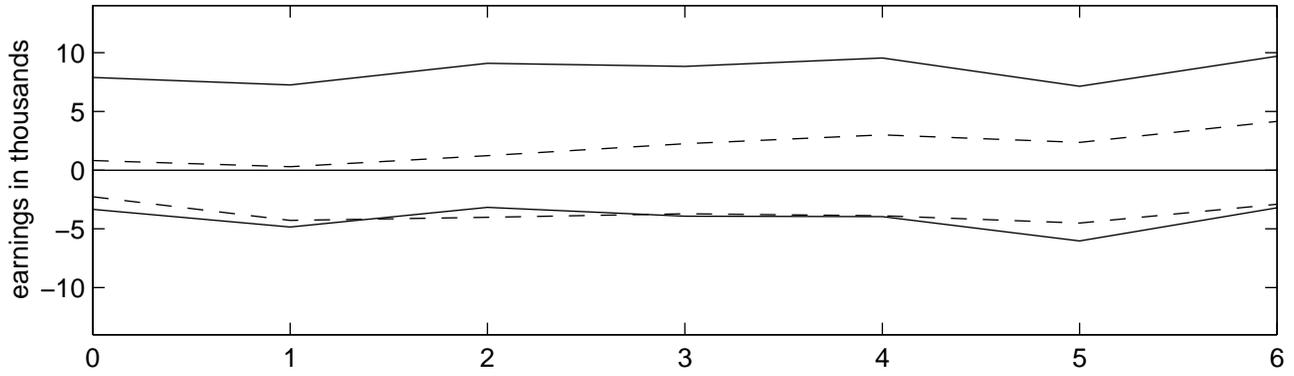


Figure 5: Confidence Intervals for the Effect on Indicator for Positive Earnings, Small Winners Versus Losers

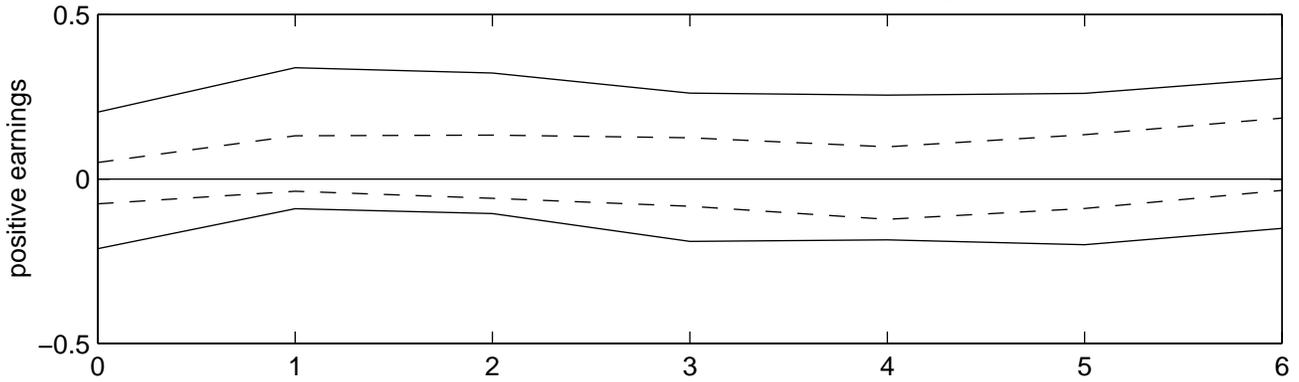


Figure 6: Confidence Intervals for the Effect on Earnings, Big Winners Versus Small Winners

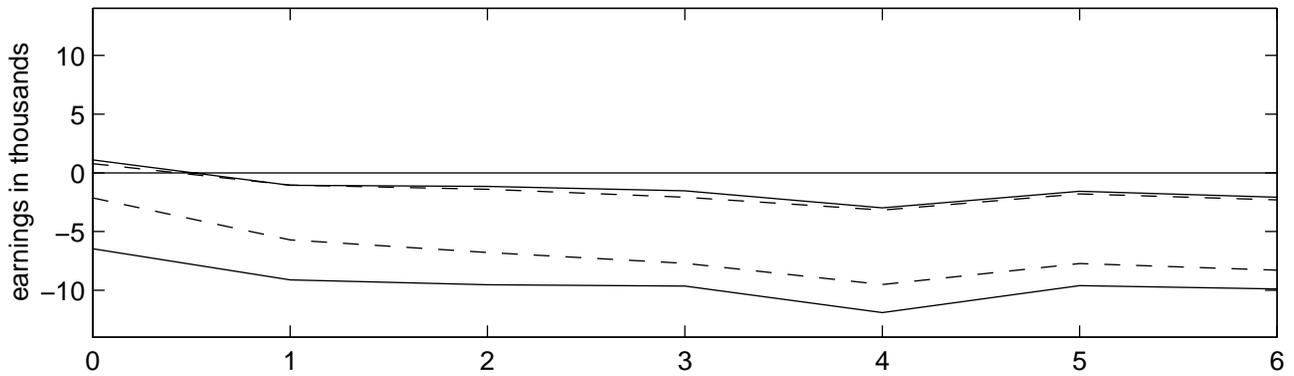


Figure 7: Confidence Intervals for the Effect on Indicator for Positive Earnings, Big Winners Versus Small Winners

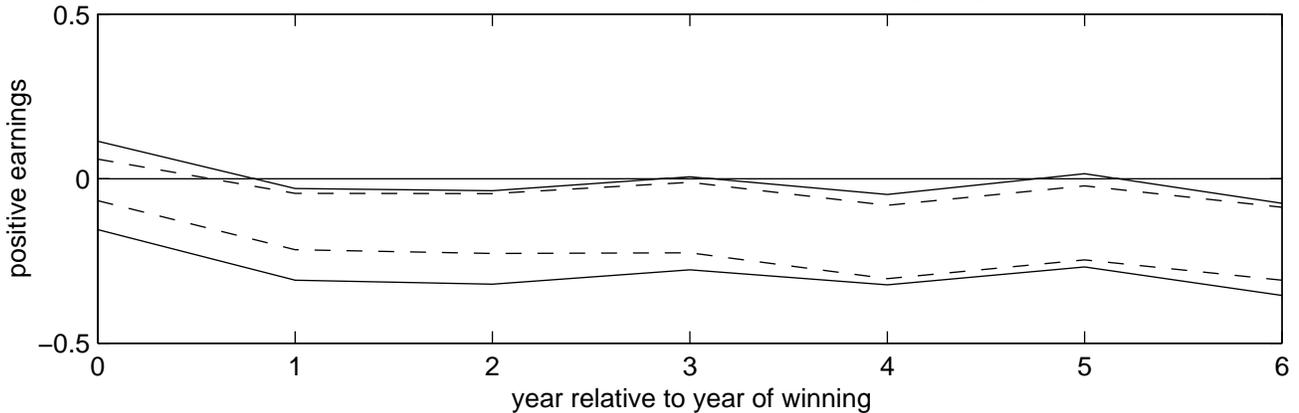


Table 1: RESPONSE RATES BY MAILING

mailing	Sent		Responses		Response Rates		
	winners	losers	winners	losers	winners	losers	total
pilot	50	50	17	25	0.34	0.50	0.42
main	752	637	272	262	0.36	0.41	0.38
follow-up (\$50 check)	248	248	39	40	0.16	0.16	0.16
follow-up (\$10 cash, \$40 check)	49	49	11	12	0.22	0.24	0.23
total	802	687	339	339	0.42	0.49	0.46

Table 2: SUMMARY STATISTICS BASIC SAMPLE: PRE-LOTTERY CHARACTERISTICS AND POST-LOTTERY OUTCOMES

Variable	Sample Size	All (N=496)		Losers (N=259)	Small Winners (N=93)	Big Winners (N=144)	t-statistics	
		mean	(s.d.)	mean	mean	mean	L/SW	SM/BW
<u>Pre-lottery Characteristics</u>								
Year Won	496	1986.2	(1.2)	1986.4	1986.0	1986.1	[-3.2]	[0.9]
Tickets Bought	496	3.3	(2.9)	2.2	4.4	4.7	[8.1]	[0.7]
Years High School	496	3.7	(0.9)	3.9	3.5	3.7	[-3.7]	[1.5]
Years College	496	2.0	(1.8)	2.6	1.2	1.5	[-6.8]	[1.0]
Work Then	496	0.78	(0.41)	0.77	0.78	0.81	[0.3]	[0.5]
Male	496	0.63	(0.48)	0.67	0.46	0.65	[-3.5]	[2.9]
Age Year Won	496	50.2	(13.7)	53.2	44.6	48.5	[-5.5]	[2.1]
Yearly Prize	496	26.4	(50.8)	0.0	15.8	80.6	[42.4]	[9.2]
<u>Post-lottery Outcomes</u>								
<u>Labor Supply</u>								
Hours Worked	414	30.6	(21.5)	30.1	37.5	26.9	[2.7]	[-3.5]
Positive Hours Worked	414	0.76	(0.43)	0.74	0.90	0.70	[3.1]	[-3.4]
Working Last Week	494	0.53	(0.50)	0.54	0.71	0.39	[2.9]	[-5.0]
Working Spouse	448	0.40	(0.49)	0.37	0.45	0.41	[1.3]	[-0.5]
<u>Consumption</u>								
Value Cars	473	18.2	(17.8)	16.7	14.3	23.6	[-1.3]	[3.6]
Debt Cars	413	3.0	(6.5)	1.9	3.7	4.4	[2.5]	[0.7]
Net Value Cars	407	15.5	(14.9)	15.3	11.4	18.75	[-2.2]	[3.3]
Value Home	484	166.3	111.6	(174.9)	133.1	171.8	[-3.1]	[2.8]
Debt Home	461	48.8	65.5	(33.8)	54.8	71.7	[2.8]	[1.8]
Net Value Home	460	122.1	(95.5)	144.6	83.0	106.8	[-5.0]	[2.3]
<u>Savings Accounts</u>								
Retirement Accounts	435	64.7	(102.8)	92.6	32.4	35.7	[-4.2]	[0.5]
Bonds/Mutual Funds	414	53.0	(126.9)	60.0	16.0	63.3	[-3.6]	[2.2]
General Savings	435	37.0	(67.5)	37.58	23.0	45.0	[-2.3]	[1.9]
Total Savings	462	143.2	(195.9)	176.7	66.6	131.6	[-4.9]	[2.5]
Log Retirement Accounts	435	2.85	(1.96)	3.38	2.19	2.33	[-4.8]	[0.5]
Log Bonds/Mutual Funds	414	2.27	(2.01)	2.67	1.08	2.31	[-6.4]	[4.4]
Log General Savings	435	2.68	(1.50)	2.81	1.93	2.90	[-4.6]	[4.5]
Log Total Savings	462	4.11	(1.58)	4.49	3.11	4.06	[-7.4]	[4.3]
<u>Savings Rates</u>								
Savings This Year	472	0.12	(0.13)	0.10	0.10	0.15	[-0.1]	[2.3]
Savings Next Year	460	0.13	(0.14)	0.11	0.13	0.17	[1.2]	[1.9]
Change in Savings Rate	460	0.01	(0.05)	0.01	0.02	0.02	[2.5]	[-0.1]

Table 3: SUMMARY STATISTICS BASIC SAMPLE: SOCIAL SECURITY EARNINGS (IN THOUSANDS OF '86 DOLLARS)

Variable	All (N=496)		Losers (N=259)	Small Winners (N=93)	Big Winners (N=144)	t-statistics	
	mean	(s.d.)	mean	mean	mean	L/SW	SM/BW
<u>Earnings</u>							
Earnings Year-6	13.9	(13.6)	15.6	9.8	13.7	[-3.6]	[2.4]
Earnings Year-5	14.3	(14.2)	16.2	9.8	13.8	[-3.7]	[2.5]
Earnings Year-4	14.3	(14.3)	16.4	9.2	13.9	[-4.1]	[2.9]
Earnings Year-3	14.8	(14.8)	16.7	10.0	14.7	[-3.7]	[2.9]
Earnings Year-2	15.6	(15.3)	17.6	10.9	15.1	[-3.6]	[2.5]
Earnings Year-1	16.3	(15.7)	18.0	12.3	15.9	[-3.0]	[2.0]
Earnings Year 0	16.3	(16.3)	18.5	12.3	14.8	[-3.1]	[1.4]
Earnings Year+1	15.6	(16.6)	18.6	12.7	11.9	[-2.9]	[-0.4]
Earnings Year+2	14.8	(16.6)	17.7	12.8	10.9	[-2.4]	[-1.0]
Earnings Year+3	14.4	(16.9)	17.5	12.9	9.9	[-2.1]	[-1.6]
Earnings Year+4	14.1	(17.2)	17.1	13.6	8.9	[-1.6]	[-2.4]
Earnings Year+5	13.8	(16.8)	16.9	12.4	9.2	[-2.1]	[-1.8]
Earnings Year+6	13.3	(16.6)	15.9	12.9	8.9	[-1.4]	[-2.3]
<u>Indicator for Pos. Earn.</u>							
Pos. Earn. Year-6	0.69	(0.46)	0.69	0.69	0.71	[0.0]	[0.3]
Pos. Earn. Year-5	0.71	(0.45)	0.68	0.74	0.74	[1.1]	[0.0]
Pos. Earn. Year-4	0.71	(0.45)	0.69	0.71	0.75	[0.3]	[0.7]
Pos. Earn. Year-3	0.70	(0.46)	0.68	0.72	0.74	[0.8]	[0.4]
Pos. Earn. Year-2	0.71	(0.46)	0.68	0.74	0.74	[1.2]	[0.0]
Pos. Earn. Year-1	0.71	(0.45)	0.69	0.74	0.74	[0.9]	[-0.1]
Pos. Earn. Year 0	0.71	(0.45)	0.69	0.73	0.74	[0.8]	[0.1]
Pos. Earn. Year+1	0.68	(0.47)	0.68	0.75	0.63	[1.3]	[-2.0]
Pos. Earn. Year+2	0.63	(0.48)	0.64	0.71	0.56	[1.2]	[-2.3]
Pos. Earn. Year+3	0.60	(0.49)	0.62	0.67	0.53	[0.8]	[-2.1]
Pos. Earn. Year+4	0.58	(0.49)	0.61	0.68	0.47	[1.2]	[-3.2]
Pos. Earn. Year+5	0.59	(0.49)	0.59	0.67	0.52	[1.2]	[-2.2]
Pos. Earn. Year+6	0.56	(0.50)	0.57	0.68	0.47	[1.8]	[-3.2]

Table 4: SMALL WINNERS AND LOSER SAMPLE (N=352), EFFECTS OF LOTTERY PRIZES ON LABOR SUPPLY AND SOCIAL SECURITY EARNINGS (IN THOUSANDS OF 1986 DOLLARS)

	Winning \$15,000 versus 0						Elasticity					
	Prop Score		OLS(dummy)		OLS(level)		Prop Score		OLS(dummy)		OLS(level)	
	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.
Hours	0.2	(5.8)	1.6	(2.4)	0.1	(2.1)	0.00	(0.05)	0.01	(0.02)	0.00	(0.02)
Positive Hours	0.04	(0.09)	0.06	(0.05)	0.06	(0.05)	0.01	(0.03)	0.02	(0.02)	0.02	(0.02)
Working	0.08	(0.11)	0.06	(0.06)	0.02	(0.05)	0.03	(0.05)	0.03	(0.03)	0.01	(0.02)
Working Spouse	0.06	(0.10)	-0.05	(0.07)	-0.01	(0.06)	0.04	(0.07)	-0.03	0.05	-0.00	(0.05)
Earnings 0	2.3	(2.9)	-0.7	(0.8)	-0.8	(0.7)	0.04	(0.04)	-0.01	(0.01)	-0.01	(0.01)
Earnings 1	1.2	(3.1)	-2.0	(1.2)	-2.1	(1.0)	0.02	(0.05)	-0.03	(0.02)	-0.03	(0.02)
Earnings 2	3.0	(3.1)	-1.4	(1.3)	-1.7	(1.2)	0.05	(0.05)	-0.02	(0.02)	-0.03	(0.02)
Earnings 3	2.5	(3.3)	-0.7	(1.5)	-1.4	(1.3)	0.04	(0.05)	-0.01	(0.02)	-0.02	(0.02)
Earnings 4	2.8	(3.5)	-0.5	(1.8)	-1.5	(1.5)	0.05	(0.06)	-0.01	(0.03)	-0.02	(0.02)
Earnings 5	0.6	(3.4)	-1.1	(1.8)	-1.4	(1.5)	0.01	(0.06)	-0.02	(0.03)	-0.02	(0.03)
Earnings 6	3.3	(3.3)	0.6	(1.8)	-0.0	(1.6)	0.06	(0.06)	0.01	(0.03)	-0.00	(0.03)
Pos Earn 0	-0.00	(0.11)	-0.01	(0.03)	-0.02	(0.03)	-0.00	(0.04)	-0.00	(0.01)	-0.01	(0.01)
Pos Earn 1	0.12	(0.11)	0.05	(0.04)	0.04	(0.04)	0.05	(0.04)	0.02	(0.02)	0.01	(0.01)
Pos Earn 2	0.11	(0.11)	0.04	(0.05)	0.05	(0.04)	0.04	(0.04)	0.01	(0.02)	0.02	(0.02)
Pos Earn 3	0.04	(0.12)	0.02	(0.05)	-0.01	(0.05)	0.01	(0.05)	0.01	(0.02)	-0.00	(0.02)
Pos Earn 4	0.03	(0.11)	-0.01	(0.06)	-0.03	(0.05)	0.01	(0.05)	-0.01	(0.02)	-0.01	(0.02)
Pos Earn 5	0.03	(0.12)	0.02	(0.06)	0.01	(0.05)	0.01	(0.05)	0.01	(0.02)	0.00	(0.02)
Pos Earn 6	0.08	(0.12)	0.08	(0.06)	0.03	(0.05)	0.03	(0.05)	0.03	(0.02)	0.01	(0.02)
Value Cars	2.1	(4.9)	-0.7	(2.3)	-0.3	(2.0)	-0.03	(0.08)	-0.01	(0.04)	-0.01	(0.03)
Debt Cars	3.2	(1.4)	1.4	(0.9)	1.5	(0.8)	0.37	(0.17)	0.17	(0.11)	0.18	(0.09)
Net Value Cars	-2.0	(3.4)	-2.7	(2.3)	-2.0	(2.0)	-0.04	(0.07)	-0.05	(0.04)	-0.04	(0.04)
Value Home	8.4	(17.5)	-9.1	(17.3)	-10.2	(14.9)	0.01	(0.03)	-0.01	(0.03)	-0.02	(0.02)
Debt Home	19.2	(13.2)	17.5	(9.3)	11.9	(8.0)	0.13	(0.09)	0.12	(0.06)	0.08	(0.05)
Net Value Home	-8.8	(15.1)	-26.0	(15.8)	-22.5	(13.6)	-0.02	(0.03)	-0.05	(0.03)	-0.05	(0.03)
Log Ret. Accts	-0.69	(0.44)	-0.45	(0.28)	-0.25	(0.24)	-0.18	(0.12)	-0.12	(0.07)	-0.06	(0.06)
Log B/MF.	-1.87	(0.47)	-0.90	(0.33)	-0.67	(0.29)	-0.48	(0.12)	-0.23	(0.08)	-0.17	(0.07)
Log Gen Sav	-0.72	(0.37)	-0.44	(0.25)	-0.14	(0.22)	-0.19	(0.10)	-0.12	(0.06)	-0.04	(0.06)
Log Tot Fin Assts	-1.18	(0.33)	-0.71	(0.22)	-0.47	(0.20)	-0.31	(0.09)	-0.18	(0.06)	-0.12	(0.05)
Sav. This Year	0.01	(0.03)	0.02	(0.02)	0.02	(0.01)	0.03	(0.07)	0.04	(0.04)	0.06	(0.04)
Sav. Next Year	0.01	(0.02)	0.04	(0.02)	0.05	(0.02)	0.03	(0.05)	0.10	(0.04)	0.12	(0.04)
Chang in Sav. Rate	0.02	(0.01)	0.03	(0.01)	0.03	(0.01)	0.67	(0.22)	0.72	(0.23)	0.74	(0.20)

Table 5: WINNERS SAMPLE (N=237), EFFECTS OF LOTTERY PRIZES ON LABOR SUPPLY AND SOCIAL SECURITY EARNINGS (IN THOUSANDS OF 1986 DOLLARS)

	Winning \$80,000 versus \$15,000						Elasticity					
	OLS(dummy)		OLS(level)		OLS(logs)		OLS(dummy)		OLS(level)		OLS(logs)	
	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.
Hours	-9.3	(2.6)	-4.1	(1.4)	-8.7	(2.2)	-0.25	(0.07)	-0.11	(0.04)	-0.17	(0.04)
Positive Hours	-0.17	(0.05)	-0.06	(0.03)	-0.13	(0.04)	-0.18	(0.05)	-0.06	(0.03)	-0.11	(0.04)
Working	-0.29	(0.06)	-0.11	(0.03)	-0.24	(0.05)	-0.48	(0.09)	-0.18	(0.05)	-0.28	(0.06)
Working Spouse	0.02	(0.06)	-0.04	(0.03)	-0.01	(0.05)	0.05	(0.13)	-0.09	(0.07)	-0.02	(0.08)
Earnings 0	-0.6	(0.8)	-1.1	(0.4)	-1.6	(0.6)	-0.04	(0.05)	-0.06	(0.02)	-0.07	(0.03)
Earnings 1	-3.2	(1.2)	-2.6	(0.6)	-4.5	(1.0)	-0.22	(0.08)	-0.18	(0.04)	-0.23	(0.05)
Earnings 2	-3.9	(1.4)	-2.9	(0.7)	-5.3	(1.2)	-0.28	(0.10)	-0.21	(0.05)	-0.28	(0.06)
Earnings 3	-4.6	(1.5)	-2.6	(0.8)	-5.5	(1.2)	-0.36	(0.11)	-0.20	(0.06)	-0.30	(0.07)
Earnings 4	-6.2	(1.6)	-3.1	(0.9)	-6.7	(1.3)	-0.49	(0.13)	-0.24	(0.07)	-0.38	(0.08)
Earnings 5	-4.5	(1.5)	-2.3	(0.8)	-5.0	(1.3)	-0.36	(0.12)	-0.19	(0.07)	-0.29	(0.08)
Earnings 6	-5.0	(1.6)	-2.4	(0.8)	-5.0	(1.3)	-0.41	(0.13)	-0.20	(0.07)	-0.29	(0.08)
Pos. Earn 0	-0.00	(0.03)	-0.02	(0.02)	-0.01	(0.03)	-0.00	(0.04)	-0.03	(0.02)	-0.01	(0.02)
Pos. Earn 1	-0.13	(0.05)	-0.09	(0.02)	-0.15	(0.04)	-0.16	(0.06)	-0.12	(0.03)	-0.13	(0.03)
Pos. Earn 2	-0.14	(0.05)	-0.09	(0.02)	-0.15	(0.04)	-0.19	(0.07)	-0.12	(0.03)	-0.14	(0.04)
Pos. Earn 3	-0.11	(0.06)	-0.08	(0.03)	-0.15	(0.05)	-0.17	(0.08)	-0.12	(0.04)	-0.16	(0.05)
Pos. Earn 4	-0.19	(0.06)	-0.10	(0.03)	-0.21	(0.05)	-0.30	(0.09)	-0.16	(0.05)	-0.23	(0.05)
Pos. Earn 5	-0.13	(0.06)	-0.10	(0.03)	-0.18	(0.05)	-0.19	(0.08)	-0.14	(0.04)	-0.19	(0.05)
Pos. Earn 6	-0.19	(0.06)	-0.10	(0.03)	-0.21	(0.05)	-0.30	(0.09)	-0.16	(0.05)	-0.23	(0.05)
Value Cars	7.9	(2.7)	5.3	(1.4)	8.9	(2.3)	0.34	(0.12)	0.23	(0.06)	0.28	(0.07)
Debt Cars	0.4	(1.1)	0.4	(0.5)	0.7	(0.9)	0.08	(0.23)	0.09	(0.11)	0.11	(0.14)
Net Value Cars	6.1	(2.3)	4.8	(1.1)	7.8	(1.9)	0.33	(0.12)	0.26	(0.06)	0.31	(0.07)
Value Home	29.1	(13.9)	31.3	(7.0)	41.5	(11.6)	0.16	(0.08)	0.17	(0.04)	0.16	(0.04)
Debt Home	20.9	(9.1)	23.6	(4.5)	31.9	(7.5)	0.27	(0.12)	0.31	(0.06)	0.30	(0.07)
Net Value Home	9.4	(10.5)	7.3	(5.4)	9.7	(8.8)	0.08	(0.09)	0.06	(0.05)	0.06	(0.05)
Log Ret. Accts	-0.23	(0.25)	0.01	(0.13)	0.01	(0.22)	-0.19	(0.21)	0.01	(0.11)	0.00	(0.13)
Log B/MF.	0.84	(0.29)	0.58	(0.14)	1.13	(0.25)	0.71	(0.25)	0.49	(0.12)	0.67	(0.15)
Log Gen Sav.	0.73	(0.21)	0.37	(0.11)	0.82	(0.18)	0.62	(0.18)	0.32	(0.09)	0.50	(0.11)
Log Tot Fin Assts	0.60	(0.22)	0.38	(0.11)	0.79	(0.19)	0.51	(0.19)	0.33	(0.09)	0.48	(0.11)
Sav. This Year	0.03	(0.02)	0.03	(0.01)	0.06	(0.02)	0.23	(0.13)	0.17	(0.07)	0.27	(0.08)
Sav. Next Year	0.03	(0.02)	0.03	(0.01)	0.06	(0.02)	0.16	(0.13)	0.17	(0.06)	0.26	(0.08)
Change in Sav. Rate	-0.00	(0.01)	0.00	(0.01)	0.01	(0.01)	-0.09	(0.44)	0.17	(0.22)	0.23	(0.27)

Table 6: ELASTICITIES BY GENDER

	Losers/Small-Winners (N=352)					Winners (N=237)				
	Men		Women		t-stat	Men		Women		t-stat
	elast.	s.e.	elast.	s.e.	dif	elast.	s.e.	elast.	s.e.	dif
Hours	-0.01	(0.02)	0.01	(0.02)	[-0.6]	-0.10	(0.04)	-0.31	(0.14)	[1.5]
Positive Hours	0.01	(0.02)	0.03	(0.02)	[-0.4]	-0.06	(0.03)	-0.17	(0.11)	[1.0]
Working	0.02	(0.03)	0.00	(0.03)	[ 0.4]	-0.17	(0.05)	-0.25	(0.13)	[0.6]
Working Spouse	-0.03	(0.06)	0.02	(0.06)	[-0.7]	-0.11	(0.07)	0.08	(0.19)	[-0.9]
Earnings 0	-0.02	(0.01)	-0.01	(0.01)	[-0.8]	-0.06	(0.03)	-0.07	(0.07)	[ 0.0]
Earnings 1	-0.04	(0.02)	-0.02	(0.02)	[-0.7]	-0.18	(0.05)	-0.14	(0.12)	[-0.4]
Earnings 2	-0.02	(0.02)	-0.03	(0.02)	[ 0.3]	-0.22	(0.06)	-0.19	(0.14)	[-0.2]
Earnings 3	-0.03	(0.03)	-0.02	(0.03)	[-0.3]	-0.19	(0.06)	-0.29	(0.16)	[ 0.6]
Earnings 4	-0.02	(0.03)	-0.03	(0.03)	[ 0.0]	-0.23	(0.07)	-0.28	(0.18)	[ 0.3]
Earnings 5	-0.05	(0.03)	0.00	(0.03)	[-1.2]	-0.18	(0.07)	-0.23	(0.18)	[ 0.3]
Earnings 6	-0.02	(0.04)	0.02	(0.04)	[-0.7]	-0.19	(0.07)	-0.26	(0.18)	[ 0.4]
Pos. Earn. 0	-0.03	(0.01)	0.01	(0.01)	[-2.3]	-0.04	(0.02)	0.05	(0.05)	[-1.6]
Pos. Earn. 1	-0.02	(0.02)	0.04	(0.02)	[-2.6]	-0.11	(0.03)	-0.20	(0.08)	[ 1.1]
Pos. Earn. 2	0.01	(0.02)	0.03	(0.02)	[-0.4]	-0.11	(0.04)	-0.18	(0.09)	[ 0.8]
Pos. Earn. 3	-0.02	(0.02)	0.02	(0.03)	[-1.2]	-0.11	(0.04)	-0.21	(0.11)	[ 0.9]
Pos. Earn. 4	-0.01	(0.03)	-0.01	(0.03)	[-0.1]	-0.13	(0.05)	-0.32	(0.12)	[ 1.4]
Pos. Earn. 5	-0.04	(0.03)	0.05	(0.03)	[-2.3]	-0.12	(0.05)	-0.27	(0.12)	[ 1.2]
Pos. Earn. 6	-0.01	(0.03)	0.03	(0.03)	[-1.2]	-0.14	(0.05)	-0.26	(0.13)	[ 0.9]
Value Cars	0.00	(0.04)	-0.01	(0.04)	[ 0.3]	0.22	(0.06)	0.29	(0.16)	[-0.4]
Debt Cars	0.23	(0.12)	0.12	(0.12)	[ 0.8]	0.13	(0.12)	-0.22	(0.32)	[ 1.1]
Net Value Cars	-0.05	(0.05)	-0.03	(0.05)	[-0.2]	0.26	(0.06)	0.26	(0.16)	[-0.0]
Value Home	-0.04	(0.03)	0.01	(0.03)	[-1.3]	0.18	(0.04)	0.11	(0.10)	[ 0.7]
Debt Home	0.03	(0.07)	0.12	(0.07)	[-1.0]	0.32	(0.06)	0.23	(0.16)	[ 0.5]
Net Value Home	-0.06	(0.04)	-0.03	(0.04)	[-0.8]	0.07	(0.05)	-0.01	(0.13)	[ 0.6]
Log Ret. Accts	0.01	(0.09)	-0.13	(0.08)	[ 1.3]	-0.03	(0.05)	0.24	(0.13)	[-2.0]
Log B/MF	-0.24	(0.10)	-0.12	(0.09)	[-1.1]	0.24	(0.07)	0.45	(0.19)	[-1.0]
Log Gen Sav	0.00	(0.08)	-0.06	(0.07)	[ 0.7]	0.11	(0.04)	0.23	(0.10)	[-1.2]
Log Tot Fin Assts	-0.11	(0.07)	-0.14	(0.06)	[ 0.4]	0.07	(0.03)	0.21	(0.07)	[-1.9]
Sav. This Year	0.05	(0.05)	0.06	(0.05)	[-0.1]	0.17	(0.07)	0.21	(0.19)	[-0.2]
Sav. Next Year	0.14	(0.05)	0.09	(0.05)	[ 0.9]	0.17	(0.07)	0.22	(0.19)	[-0.3]
Change in Sav. Rate	1.19	(0.25)	0.25	(0.26)	[ 2.8]	0.15	(0.24)	0.29	(0.67)	[-0.2]

Table 7: ELASTICITIES BY EARNINGS STATUS PRIOR TO WINNING THE LOTTERY

	Losers/Small-Winners (N=352)					Winners (N=237)				
	Pos Earn elast.	s.e.	Zero Earn elast.	s.e.	t-stat dif	Pos Earn elast.	s.e.	Zero Earn elast.	s.e.	t-stat dif
Hours	0.00	(0.02)	-0.01	(0.03)	[0.3]	-0.13	(0.05)	-0.09	(0.05)	[-0.6]
Positive Hours Working	0.03	(0.02)	-0.00	(0.03)	[1.0]	-0.05	(0.04)	-0.08	(0.04)	[ 0.5]
Working Spouse	0.03	(0.03)	-0.03	(0.04)	[1.3]	-0.22	(0.06)	-0.12	(0.08)	[-1.1]
	0.03	(0.05)	-0.06	(0.07)	[1.0]	-0.12	(0.08)	-0.04	(0.10)	[-0.2]
Earnings 0	-0.02	(0.01)	0.00	(0.02)	[-1.0]	-0.11	(0.03)	-0.00	(0.04)	[-2.2]
Earnings 1	-0.04	(0.02)	-0.02	(0.03)	[-0.8]	-0.29	(0.05)	-0.02	(0.06)	[-3.2]
Earnings 2	-0.04	(0.02)	-0.01	(0.03)	[-0.8]	-0.35	(0.07)	-0.02	(0.08)	[-3.2]
Earnings 3	-0.03	(0.02)	0.00	(0.04)	[-0.8]	-0.33	(0.07)	-0.02	(0.09)	[-2.7]
Earnings 4	-0.04	(0.03)	0.01	(0.04)	[-1.1]	-0.37	(0.09)	-0.05	(0.10)	[-2.5]
Earnings 5	-0.04	(0.03)	0.00	(0.04)	[-0.8]	-0.29	(0.08)	-0.03	(0.10)	[-2.1]
Earnings 6	-0.01	(0.03)	0.03	(0.05)	[-0.7]	-0.32	(0.08)	-0.03	(0.10)	[-2.2]
Pos Earn 0	-0.02	(0.01)	0.02	(0.02)	[-2.4]	-0.05	(0.02)	-0.00	(0.03)	[-1.3]
Pos Earn 1	-0.00	(0.02)	0.05	(0.02)	[-1.9]	-0.15	(0.04)	-0.06	(0.04)	[-1.6]
Pos Earn 2	0.02	(0.02)	0.02	(0.03)	[-0.2]	-0.17	(0.04)	-0.04	(0.05)	[-2.0]
Pos Earn 3	-0.02	(0.02)	0.04	(0.03)	[-1.7]	-0.15	(0.05)	-0.07	(0.06)	[-0.9]
Pos Earn 4	-0.04	(0.02)	0.04	(0.03)	[-2.1]	-0.19	(0.06)	-0.10	(0.07)	[-1.0]
Pos Earn 5	-0.03	(0.02)	0.08	(0.04)	[-2.8]	-0.18	(0.06)	-0.09	(0.07)	[-1.1]
Pos Earn 6	-0.00	(0.02)	0.05	(0.04)	[-1.4]	-0.21	(0.06)	-0.08	(0.07)	[-1.4]
Value Cars	0.00	(0.04)	-0.02	(0.05)	[ 0.3]	0.24	(0.07)	0.22	(0.10)	[ 0.1]
Debt Cars	0.25	(0.11)	0.01	(0.15)	[ 1.4]	0.02	(0.15)	0.18	(0.17)	[-0.7]
Net Value Cars	-0.05	(0.04)	-0.02	(0.06)	[-0.5]	0.30	(0.07)	0.19	(0.09)	[ 0.9]
Value Home	-0.00	(0.03)	-0.04	(0.04)	[ 0.8]	0.21	(0.05)	0.11	(0.06)	[ 1.3]
Debt Home	0.14	(0.06)	-0.05	(0.09)	[ 1.9]	0.32	(0.07)	0.29	(0.09)	[ 0.3]
Net Value Home	-0.05	(0.03)	-0.04	(0.04)	[-0.2]	0.11	(0.06)	0.00	(0.07)	[ 1.2]
Log Ret. Accts	-0.10	(0.08)	0.01	(0.10)	[-1.0]	0.09	(0.14)	-0.10	(0.16)	[0.9]
Log B/MF.	-0.19	(0.09)	-0.15	(0.11)	[-0.3]	0.53	(0.16)	0.43	(0.19)	[0.4]
Log Gen Sav	-0.09	(0.07)	0.07	(0.09)	[-1.6]	0.47	(0.12)	0.13	(0.14)	[1.9]
Log Tot Fin Assts	-0.16	(0.06)	-0.04	(0.08)	[-1.3]	0.41	(0.12)	0.20	(0.14)	[1.2]
Sav. This Year	0.02	(0.04)	0.14	(0.06)	[-1.7]	0.19	(0.09)	0.14	(0.10)	[0.4]
Sav. Next Year	0.10	(0.04)	0.14	(0.06)	[-0.6]	0.19	(0.08)	0.14	(0.10)	[0.4]
Change in Sav. Rate	0.97	(0.23)	0.26	(0.32)	[ 1.9]	0.20	(0.29)	0.12	(0.34)	[0.2]

Table 8: ELASTICITIES BY AGE

	Losers/Small-Winners (N=352)					Winners (N=237)				
	Age 40		Age 60		t-stat	Age 40		Age 60		t-stat
	elast.	s.e.	elast.	s.e.	dif	elast.	s.e.	elast.	s.e.	dif
Hours	-0.01	(0.02)	0.04	(0.03)	[ 1.9]	-0.15	(0.05)	-0.07	(0.05)	[ 1.4]
Positive Hours	0.01	(0.02)	0.05	(0.02)	[ 1.6]	-0.07	(0.04)	-0.05	(0.04)	[ 0.4]
Working	-0.01	(0.03)	0.04	(0.03)	[ 1.3]	-0.22	(0.07)	-0.15	(0.06)	[ 0.9]
Working Spouse	0.02	(0.05)	-0.05	(0.06)	[-1.1]	-0.17	(0.09)	-0.03	(0.08)	[ 1.5]
Earnings 0	-0.01	(0.01)	-0.01	(0.01)	[-0.2]	-0.11	(0.03)	-0.04	(0.03)	[2.1]
Earnings 1	-0.04	(0.02)	-0.02	(0.02)	[ 1.1]	-0.26	(0.06)	-0.13	(0.05)	[2.0]
Earnings 2	-0.03	(0.02)	-0.02	(0.02)	[ 0.7]	-0.33	(0.07)	-0.14	(0.06)	[2.4]
Earnings 3	-0.02	(0.02)	-0.02	(0.03)	[ 0.3]	-0.33	(0.08)	-0.12	(0.07)	[2.4]
Earnings 4	-0.04	(0.03)	-0.00	(0.03)	[ 1.1]	-0.35	(0.09)	-0.17	(0.08)	[1.8]
Earnings 5	-0.04	(0.03)	-0.00	(0.03)	[ 1.0]	-0.27	(0.09)	-0.13	(0.07)	[1.5]
Earnings 6	-0.01	(0.03)	0.01	(0.04)	[ 0.6]	-0.31	(0.09)	-0.13	(0.08)	[1.8]
Pos Earn 0	-0.01	(0.01)	-0.01	(0.01)	[0.1]	-0.05	(0.03)	-0.02	(0.02)	[1.2]
Pos Earn 1	0.01	(0.02)	0.03	(0.02)	[0.9]	-0.16	(0.04)	-0.09	(0.03)	[1.7]
Pos Earn 2	0.01	(0.02)	0.03	(0.02)	[0.7]	-0.19	(0.05)	-0.08	(0.04)	[2.1]
Pos Earn 3	-0.01	(0.02)	0.00	(0.03)	[0.3]	-0.21	(0.06)	-0.06	(0.05)	[2.4]
Pos Earn 4	-0.02	(0.02)	0.01	(0.03)	[1.1]	-0.24	(0.06)	-0.11	(0.05)	[2.0]
Pos Earn 5	-0.01	(0.02)	0.03	(0.03)	[1.2]	-0.22	(0.06)	-0.10	(0.05)	[1.8]
Pos Earn 6	0.01	(0.02)	0.02	(0.03)	[0.3]	-0.23	(0.06)	-0.11	(0.05)	[1.8]
Value Cars	0.01	(0.04)	-0.03	(0.05)	[-0.9]	0.22	(0.08)	0.24	(0.07)	[ 0.2]
Debt Cars	0.18	(0.10)	0.17	(0.13)	[-0.1]	0.19	(0.16)	0.03	(0.13)	[-0.9]
Net Value Cars	-0.03	(0.04)	-0.06	(0.05)	[-0.6]	0.22	(0.08)	0.29	(0.07)	[ 0.7]
Value Home	0.00	(0.03)	-0.05	(0.03)	[-1.5]	0.22	(0.05)	0.14	(0.04)	[-1.5]
Debt Home	0.12	(0.06)	0.00	(0.07)	[-1.6]	0.34	(0.08)	0.28	(0.07)	[-0.7]
Net Value Home	-0.04	(0.03)	-0.06	(0.04)	[-0.6]	0.11	(0.06)	0.03	(0.05)	[-1.2]
Log Ret. Accts	-0.10	(0.07)	0.03	(0.09)	[ 1.5]	-0.04	(0.14)	0.05	(0.12)	[0.6]
Log B/MF.	-0.16	(0.08)	-0.20	(0.10)	[-0.5]	0.42	(0.17)	0.53	(0.14)	[0.6]
Log Gen Sav	-0.05	(0.06)	-0.01	(0.08)	[ 0.6]	0.25	(0.12)	0.36	(0.11)	[0.8]
Log Tot Fin Assts	-0.13	(0.06)	-0.11	(0.07)	[ 0.4]	0.25	(0.13)	0.37	(0.11)	[0.9]
Sav. This Year	0.04	(0.04)	0.10	(0.05)	[ 1.1]	0.07	(0.09)	0.24	(0.08)	[ 1.8]
Sav. Next Year	0.11	(0.04)	0.13	(0.05)	[ 0.4]	0.10	(0.09)	0.22	(0.07)	[ 1.3]
Change in Sav. Rate	0.86	(0.22)	0.49	(0.27)	[-1.4]	0.32	(0.30)	0.07	(0.26)	[-0.7]

Table 9: ESTIMATES OF PROPENSITY SCORE (LOGIT MODEL)

	Small-Winners/Losers (N=352)		Big/Small Winners (N=237)	
	est.	t-stat	est.	t-stat
Year Won	-0.30	[-1.8]	-0.12	[-1.1]
Number of Tickets Bought	0.40	[5.3]	-0.01	-[0.2]
Years of High School	-0.42	[-2.3]	-0.21	[-1.5]
Years of College	-0.63	[-5.2]	-0.10	[-1.0]
Working at Time of Winning	0.16	[0.3]	0.03	[0.1]
Male	-0.69	[-1.8]	-0.64	[-2.0]
Age at Time of Winning	-0.10	[-5.4]	-0.03	[-2.4]
Earnings Year-6	0.00	[ 0.1]		
Earnings Year-5	-0.00	[-0.0]		
Earnings Year-4	-0.07	[-1.2]		
Earnings Year-3	0.09	[ 1.4]		
Earnings Year-2	-0.07	[-1.2]		
Earnings Year-1	-0.01	[-0.2]	-0.01	[1.0]
Positive Earnings Year-6	0.45	[ 0.7]		
Positive Earnings Year-5	1.01	[ 1.1]		
Positive Earnings Year-4	-0.69	[-0.8]		
Positive Earnings Year-3	-0.15	[-0.2]		
Positive Earnings Year-2	0.13	[ 0.2]		
Positive Earnings Year-1	-0.05	[-0.1]	0.03	[0.1]
Test for Zero Coeff on Earn-6 to Earn-2 $\chi^2(10)$		7.4		3.9
Test for Zero Coeff on Earn-6 to Earn-1 $\chi^2(12)$		17.7		5.0
Test for Zero Coeff on All Slope Coeff. $\chi^2(19)$		177.3		24.0

Table 10: SMALL WINNERS AND LOSER SAMPLE (N=352), SENSITIVITY CHECKS ON LEAST SQUARES ESTIMATES OF EFFECT OF WINNING \$15,000 VERSUS \$0 ON SOCIAL SECURITY EARNINGS

	Last Year Used as Pre-lottery Year											
	Year-6		Year-5		Year-4		Year-3		Year-2		Year-1	
	est.	t-stat	est.	t-stat	est.	t-stat	est.	t-stat	est.	t-stat	est.	t-stat
Year-5	-0.9	[-1.4]										
Year-4	-2.6	[-2.6]	-2.0	[-2.3]								
Year-3	-2.5	[-2.1]	-1.8	[-1.6]	0.2	[0.3]						
Year-2	-3.1	[-2.3]	-2.4	[-1.9]	-0.6	[-0.6]	-0.8	[-1.2]				
Year-1	-3.5	[-2.4]	-2.7	[-2.0]	-0.9	[-0.7]	-1.0	[-1.1]	0.4	[-0.5]		
Year 0	-4.1	[-2.6]	-3.2	[-2.2]	-1.5	[-1.1]	-1.6	[-1.4]	-1.0	[-1.0]	-0.7	[-0.9]
Year+1	-4.9	[-2.9]	-4.1	[-2.6]	-2.5	[-1.7]	-2.6	[-2.0]	-2.2	[-1.8]	-2.0	[-1.7]
Year+2	-4.2	[-2.4]	-3.3	[-2.0]	-1.9	[-1.2]	-2.0	[-1.4]	-1.6	[-1.1]	-1.4	[-1.0]
Year+3	-3.3	[-1.7]	-2.4	[-1.3]	-1.2	[-0.7]	-1.3	[-0.8]	-0.9	[-0.6]	-0.7	[-0.5]
Year+4	-2.8	[-1.4]	-2.1	[-1.1]	-0.8	[-0.4]	-1.0	[-0.5]	-0.7	[-0.4]	-0.5	[-0.3]
Year+5	-3.5	[-1.8]	-2.9	[-1.5]	-1.4	[-0.7]	-1.4	[-0.8]	-1.3	[-0.7]	-1.1	[-0.6]
Year+6	-1.9	[-1.0]	-1.2	[-0.6]	0.2	[0.1]	0.2	[0.1]	0.4	[0.2]	0.6	[0.4]

Table 11: SMALL WINNERS AND LOSER SAMPLE (N=352), SENSITIVITY CHECKS ON PROPENSITY SCORE ESTIMATES OF EFFECT OF WINNING \$15,000 VERSUS \$0 ON SOCIAL SECURITY EARNINGS

	Last Year Used as Pre-lottery Year											
	Year-6		Year-5		Year-4		Year-3		Year-2		Year-1	
	est.	t-stat	est.	t-stat	est.	t-stat	est.	t-stat	est.	t-stat	est.	t-stat
Year-5	-0.1	[-0.0]										
Year-4	-1.1	[-0.5]	-2.1	[-1.0]								
Year-3	-0.3	[-0.1]	-1.6	[-0.7]	1.9	[1.0]						
Year-2	-1.3	[-0.6]	-2.0	[-0.9]	0.6	[0.3]	1.0	[0.5]				
Year-1	-1.9	[-0.7]	-2.0	[-0.8]	0.1	[0.1]	0.6	[0.3]	2.8	[1.1]		
Year 0	-2.7	[-1.0]	-1.8	[-0.7]	-0.6	[-0.3]	0.1	[0.1]	2.5	[0.9]	2.3	[0.8]
Year+1	-3.9	[-1.4]	-3.5	[-1.3]	-3.3	[-1.4]	-2.5	[-1.1]	-0.8	[-0.3]	1.2	[0.4]
Year+2	-2.5	[-0.9]	-0.9	[-0.3]	-1.1	[-0.4]	-0.5	[-0.2]	1.0	[0.3]	3.0	[1.0]
Year+3	-2.4	[-0.8]	-0.7	[-0.2]	-1.2	[-0.4]	-0.5	[-0.2]	0.6	[0.2]	2.5	[0.8]
Year+4	-2.1	[-0.7]	-0.2	[-0.1]	-0.5	[-0.2]	-0.0	[-0.0]	1.6	[0.5]	2.8	[0.8]
Year+5	-3.2	[-1.1]	-2.3	[-0.8]	-1.7	[-0.6]	-1.5	[-0.6]	0.4	[0.1]	0.6	[0.2]
Year+6	-1.2	[-0.4]	0.8	[ 0.3]	0.9	[ 0.3]	0.9	[ 0.3]	2.8	[0.9]	3.3	[1.0]

Table 12: WINNERS SAMPLE (N=237), SENSITIVITY CHECKS ON LEAST SQUARES ESTIMATES OF EFFECT OF WINNING \$80,000 VERSUS \$15,000 ON SOCIAL SECURITY EARNINGS

	Last Year Used as Pre-lottery Year											
	Year-6		Year-5		Year-4		Year-3		Year-2		Year-1	
	est.	t-stat	est.	t-stat	est.	t-stat	est.	t-stat	est.	t-stat	est.	t-stat
Year-5	0.3	[0.4]										
Year-4	1.2	[1.3]	0.9	[1.4]								
Year-3	1.1	[1.1]	0.9	[1.1]	0.2	[0.2]						
Year-2	1.1	[0.9]	0.8	[0.8]	0.1	[0.1]	-0.0	[-0.1]				
Year-1	0.9	[0.7]	0.7	[0.6]	-0.0	[-0.0]	-0.1	[-0.1]	-0.01	[-0.1]		
Year 0	0.2	[0.1]	-0.1	[-0.0]	-0.6	[-0.5]	-0.8	[-0.7]	-0.8	[-0.7]	-0.7	[-0.9]
Year+1	-2.6	[-1.6]	-2.8	[-1.9]	-3.3	[-2.4]	-3.5	[-2.7]	-3.4	[-2.8]	-3.4	[-2.9]
Year+2	-3.3	[-1.9]	-3.5	[-2.3]	-4.0	[-2.7]	-4.1	[-2.9]	-4.1	[-2.9]	-4.1	[-3.0]
Year+3	-4.2	[-2.5]	-4.4	[-2.9]	-4.8	[-3.1]	-4.9	[-3.4]	-4.9	[-3.4]	-4.9	[-3.4]
Year+4	-5.8	[-3.2]	-6.0	[-3.6]	-6.3	[-3.7]	-6.4	[-3.9]	-6.4	[-3.9]	-6.3	[-3.9]
Year+5	-4.0	[-2.4]	-4.3	[-2.7]	-4.7	[-3.0]	-4.8	[-3.1]	-4.8	[-3.1]	-4.8	[-3.2]
Year+6	-4.7	[-2.8]	-4.9	[-3.1]	-5.2	[-3.3]	-5.3	[-3.5]	-5.3	[-3.5]	-5.3	[-3.5]

Table 13: WINNERS SAMPLE (N=237), SENSITIVITY CHECKS ON PROPENSITY SCORE ESTIMATES OF EFFECT OF WINNING \$80,000 VERSUS \$15,000 ON SOCIAL SECURITY EARNINGS

	Last Year Used as Pre-lottery Year											
	Year-6		Year-5		Year-4		Year-3		Year-2		Year-1	
	est.	t-stat	est.	t-stat	est.	t-stat	est.	t-stat	est.	t-stat	est.	t-stat
Year-5	0.3	[0.2]										
Year-4	1.2	[0.8]	2.0	[1.2]								
Year-3	1.0	[0.6]	1.7	[1.0]	0.2	[0.1]						
Year-2	0.8	[0.5]	1.3	[0.7]	0.1	[0.1]	0.1	[0.1]				
Year-1	0.6	[0.3]	1.2	[0.6]	-0.6	[-0.3]	0.0	[0.0]	-0.7	[-0.4]		
Year 0	0.2	[ 0.1]	0.6	[ 0.3]	-1.0	[-0.5]	-1.2	[-0.6]	-1.4	[-0.7]	-2.7	[-1.4]
Year+1	-1.8	[-0.9]	-1.5	[-0.7]	-3.2	[-1.6]	-3.5	[-1.6]	-3.7	[-1.8]	-5.1	[-2.5]
Year+2	-2.2	[-1.0]	-2.1	[-1.0]	-3.6	[-1.8]	-4.1	[-1.9]	-4.2	[-2.0]	-5.3	[-2.5]
Year+3	-2.9	[-1.4]	-3.1	[-1.4]	-4.3	[-2.2]	-4.5	[-2.1]	-4.7	[-2.3]	-5.6	[-2.7]
Year+4	-4.8	[-2.2]	-4.7	[-2.0]	-5.9	[-2.7]	-6.3	[-2.6]	-6.3	[-2.8]	-7.4	[-3.3]
Year+5	-3.0	[-1.5]	-2.8	[-1.3]	-4.1	[-2.1]	-4.7	[-2.2]	-4.7	[-2.3]	-5.6	[-2.7]
Year+6	-3.7	[-1.8]	-3.5	[-1.7]	-4.4	[-2.3]	-5.4	[-2.6]	-5.1	[-2.5]	-6.0	[-3.0]

# *Lottery Winner Survey*

Respondent Code: XXXX  
Year won: 1986  
Prize: XXXXXX

***INTRODUCTION:*** *This survey is about how winning a Lottery prize has affected your life. It will take about 25 minutes to complete. All the information you give will be used for research purposes only and your responses will be kept completely confidential.*

*Any questions that ask about your 'spouse' should be interpreted broadly to refer to any partner with whom you live whether or not you are married.*

*Please print neatly using black ink*

Date survey completed: \_\_\_\_\_

***This first group of questions is about your background:***

1. Your date of birth \_\_\_\_\_

2. Your gender (please circle one)            male            female

3. Your race (optional, circle one)

White (including Hispanic)

Black

Asian or Pacific Islander

Other (please specify) \_\_\_\_\_

*You*                      *Current Spouse*

4. Years of high school completed (circle one)    0 1 2 3 4                      0 1 2 3 4

5. Years of college completed (circle one)        0 1 2 3 4                      0 1 2 3 4

6. What were you doing most of last week? (circle one)

- working                      with a job but not at work                      looking for work
- keeping house              going to school                      unable to work                      retired
- other (please specify) \_\_\_\_\_

7. Your current occupation (please state)

8. In a typical week, how many hours per week do you currently work?

9. What was your current spouse doing most of last week? (circle one)

- working                      with a job but not at work                      looking for work
- keeping house              going to school                      unable to work                      retired
- other (please specify) \_\_\_\_\_

10. Please complete the following table showing your children's ages and how much schooling they have completed.

The table also asks you to estimate how much (if any) you spent on each child's high school, college, and graduate education.

Please give a total dollar figure and include any tuition, fees, room, and board that you paid. (Exclude money from other sources such as scholarships and grants.)

	Year of birth	Years of high school completed (0-4)	Total amount you spent on <i>private high school</i> tuition, fees & expenses	Years of college completed (0-4)	Total amount you spent on <i>college</i> tuition, fees & expenses	Total amount you spent on <i>graduate school</i> tuition, fees & expenses
child #1						
child #2						
child #3						
child #4						
child #5						
child #6						
child #7						

11. Do you own or rent your home? (circle one) own                      rent

12. If renter, approximate monthly rent \_\_\_\_\_

13. The following table asks you for some information on your assets, including any cars, real estate, homes, or businesses that you own. Your assets also include any investments that you hold such as bank accounts, stocks, bonds, mutual funds, and pension plans.

The table asks you to value each of these assets. In each case, please try to give a rough dollar figure (your best guess) of the value. In certain cases, the table asks if you have any debt or mortgages on the asset.

If you do not have any of a particular asset, please write in \$0. If there is no debt or mortgage on the asset, please write in \$0 in the appropriate box. *Do not to leave any blanks in the table. If you do not know the value the asset, please make your best guess.*

Asset	approximate dollar value	Debt or mortgage you owe on asset
primary home (and property)		
other homes or real estate (such as vacation homes)		
automobiles		
businesses (incl. building, property etc.)		
any other commercial property (such as rental real estate)		
IRAs, 401K plans, 403B plans, KEOUGH, or employer sponsored retirement plans		<i>not applicable</i>
savings accounts or CDs		<i>not applicable</i>
other stocks, bonds, and mutual funds		<i>not applicable</i>
life insurance		<i>not applicable</i>
other major assets		

14. Considering your and your spouse's total after-tax income this year, what portion will you save or invest this year and next year? (please give a percentage)

This Year                      Next Year  
 \_\_\_\_\_%                      \_\_\_\_\_%

***During 1986 you won a Lottery prize of \$100-\$5,000 playing Megabucks. The following questions concern your situation at the time you won this Lottery prize.***

**15. Around 1986, what were you doing during a typical week? (circle one)**

**working                      keeping house                      going to school                      retired**

**other (please specify) \_\_\_\_\_**

**16. Your occupation during 1986 (please state)**

**17. During the period *before* you won the Lottery prize (in 1986), how many Megabucks tickets did you typically buy per week?**

**Megabucks tickets \_\_\_\_\_**

**18. Where did you live at the time you won the lottery prize (i.e.1986)? (city name and zipcode) (example: Cambridge, MA 02138)**

**19. At the time you won the lottery prize (1986) did you own or rent your home? (circle one)**

**own                      rent**

**20. In 1986 how many years of schooling had you completed (Don't include any schooling completed after 1986):**

**high school (circle one)**

**college (circle one)**

**0 1 2 3 4**

**0 1 2 3 4**

21. Your marital status in 1986 (circle one)

married

widowed

divorced

separated

never married

22. Number of times you got divorced or separated AFTER 1986?  
(circle one)

0

1

2

3

4

23. Are you generally happy? (circle one)

yes

no

*Finally, please sign the white form from Social Security and return it with your completed survey. Thank you very much for your participation. Please return the survey, the Social Security form and the slip for your season's ticket to Megabucks in the envelope provided.*

## APPENDIX B: THE PROPENSITY SCORE

One of the procedures we use to adjust for pre-lottery differences is based on the propensity score methodology. This methodology was originally developed by Rosenbaum and Rubin (1983a, 1983b, 1985). It was recently applied by Dehejia and Wahba (1998) to the data originally analyzed by Lalonde (1985).<sup>13</sup> Dehejia and Wahba find that with the Lalonde data the propensity score approach is much more successful than any of the alternatives considered by Lalonde in replicating the experimental results in an evaluation of a job training program. In particular its success is in dealing with comparison groups that differ substantially in a number of dimensions when least squares adjustment rely heavily on linearity. In this appendix we describe briefly the application of this methodology to the problem at hand.

Using the potential outcome notation, (Rubin, 1974, Holland, 1986), let  $Y(0)$  and  $Y(1)$  be the potential outcomes under the two levels of the treatment, in our application “losing” and “winning”, respectively. Furthermore let  $T$  be a binary indicator for the treatment received,  $T \in \{0, 1\}$ , and let  $Y$  be the realized outcome,  $Y = T \cdot Y(1) + (1 - T) \cdot Y(0)$ . Finally, let  $X$  be a vector of pre-treatment variables. In our application this vector includes background characteristics such as age at the time of winning, gender, education, and earnings in years prior to winning. The key assumption is the *unconfoundedness* assumption that conditional on the pre-treatment variables assignment to the treatment is independent of the two potential outcomes:

$$T \perp Y(0), Y(1) \mid X. \tag{1}$$

In Appendix C we evaluate the plausibility of this assumption in the current application. The unconfoundedness assumption validates comparisons for units with the same value of

---

<sup>13</sup>Other applications in economics include Heckman, Ichimura, Smith and Todd (1998), and Hotz, Imbens and Mortimer (1999).

the pre-treatment variables:

$$E[Y|T = 1, X] - E[Y|T = 0, X] = E[Y(1) - Y(0)|X]. \quad (2)$$

To get the population average of the causal effect  $E[Y(1) - Y(0)]$  can then be obtained by averaging the within-cell average over the distribution of  $X$ :

$$E[Y(1) - Y(0)] = E\left[E[Y(1) - Y(0)|X]\right]. \quad (3)$$

The practical problem is that with many pre-treatment variables it is not feasible to divide the sample into subsamples with the exact value of the pre-treatment variables. With a single, continuous pre-treatment variable this is of course not possible either, but in that case one can coarsen the pre-treatment variable and ignore the variation of the pre-treatment variable within coarser cells. The difficulty in doing this with many pre-treatment variables is in finding a metric that allows comparisons of units with differences in many pre-treatment variables. This is where the propensity score enters the analysis. Let  $e(X) = Pr(T = 1|X)$  be the conditional probability of receiving the treatment, or the propensity score. Rosenbaum and Rubin (1983) show that if assignment is independent of the potential outcomes conditional on the pre-treatment variables, it is also independent conditional on the propensity score:

$$T \perp Y(0), Y(1) \mid X \longrightarrow T \perp Y(0), Y(1) \mid e(X).$$

We implement this methodology, following Rosenbaum and Rubin (1985), and Dehejia and Wahba (1998) as follows: First we estimate a logistic regression of the binary treatment indicator on a vector of functions of the pre-treatment variables. Given the estimated coefficients from this logistic regression we calculate the estimated propensity score  $\hat{e}(x)$  for each unit.

Second, we calculate the minimum value of the estimated propensity score for the treated units and discard all control units with an estimated propensity score less than this value.

Similarly we discard all treated units with an estimated propensity score larger than the maximum value of the estimated propensity score for the control units. This step was of great importance in the Dehejia-Wahba application to the Lalonde data because of the large differences in average pre-treatment variables between the treated and control units.

Third, we then divide the sample into five subsamples, based on the quintiles of the empirical distribution of the estimated propensity score of the units with  $T = 1$ . Within these subsamples we estimate the average treatment effect by ordinary least squares using three of the pre-lottery variables and an indicator for the group. The three control variables are age at the time of winning, social security earnings in the last year prior to winning, and an indicator for these earnings being positive. The choice for these variables was based on prior beliefs that these were among the ones most highly correlated with the outcome variables. The estimated within-block average treatment effect is the estimate of the coefficient on the group indicator.

Finally we weight these subsample average treatment effects by the fraction of treated units in each sample to get an estimate of the population average treatment effect.

We apply this methodology to both the small-winners/losers and the big-winners/small-winners samples. Table 9 presents the estimates of the two logistic regressions.<sup>14</sup> For the small-winners/losers sample we use as pre-treatment variables year won, number of tickets bought, age, years of high school, years of college, an indicator for working at the time of winning, an indicator for male, earnings for each of the six years preceding winning, and indicators for positive earnings in each of those six years. For the big-winners/small-winners sample we use the same pre-treatment variables with the one modification that we only use earnings and indicator for positive earnings for the one year prior to winning the lottery. The reason for using only one year of earnings data for the big-winners/small-winners sample

---

<sup>14</sup>This is for the basic samples with 259 losers, 93 small winners and 144 big winners. For the calculations involving additional variables the samples are sometimes smaller and the logistic regressions can be slightly different.

is that the distributions of pre-treatment variables are much closer for treated and control units in that sample. We carried out a number of tests to support this, also reported in Table 13. First we included all earnings and indicators for positive earnings and tested the hypothesis that all slope coefficients were equal to zero. For the small-winners/losers sample this leads to a statistic of 177.3. Under the null hypothesis this statistic should have a  $\chi^2$  distribution with 19 degrees of freedom. For the big-winners/small-winners sample the statistic is 24.0, suggesting the hypothesis should not be rejected at conventional levels. Second, we test whether the six years of pre-lottery earnings contribute to the propensity score given the other characteristics. Here the statistics, under the null hypothesis from a  $\chi^2$  distribution with 12 degrees of freedom, are 17.7 and 5.0 respectively. Finally we test whether the first five years of earnings contribute anything to the propensity score given the final pre-lottery year. Here the statistics under the null hypothesis from a  $\chi^2$  distribution with 10 degrees of freedom, are 7.4 and 3.9 respectively. On the basis of these test-statistics and other calculations we decided on the specification of the propensity scores.

Given the estimates of the propensity score for the small-winner/loser sample 11 (out of 93) small winners were discarded with estimated propensity scores larger than the largest estimated propensity score for losers. Similarly 104 (out of 259) losers were discarded with propensity scores smaller than the smallest propensity score for small winners. In the big-winners/small-winners sample only 17 big winners were discarded, again suggesting the balance is much better in this sample.

In Tables 4 and 5 one can compare the least squares and propensity score estimates for the small-winners/losers sample. The estimates are generally very close with the standard errors for the least squares estimates much smaller than those for the propensity score estimates. This can be seen very clearly in Figures 4 and 5 where the solid lines give the 95% confidence intervals based on propensity score estimates and the dashed lines the corresponding intervals based on least squares estimates. For the big-winner/small-winner sample the differences between least squares and propensity score estimates were much smaller, as evidenced by

Figures 6 and 7, and this is the reason we omitted propensity score estimates for these samples in the body of the paper.

## APPENDIX C: SENSITIVITY OF ADJUSTMENT PROCEDURES

In this appendix we discuss the sensitivity of the estimates to the choice of adjustment procedures and pre-lottery variables. Like Heckman and Hotz (1989) we exploit the presence of multiple years of pre-lottery earnings to evaluate the adequacy of short earnings histories to adjust pre-lottery difference. The importance of earnings histories in adjusting for pre-treatment differences has long been the focus of much discussion. See, among others, Ashenfelter and Card (1978), Card and Sullivan (1989), and Dehejia and Wahba (1998).

First consider the small-winners/losers sample. In the body of the paper we report results based on propensity score and least squares estimates using as pre-lottery variables earnings and indicator for positive earnings for six years prior to winning the lottery. Here we report estimates based on adjusting for only the first  $K$  years of these six years of pre-lottery earnings, both in levels and with an indicator for positive earnings. In Table 10 the first two columns reports least squares estimates based on adjusting for the same background characteristics as before (year won, age, years high school, years college, working then, male) but using only earnings six years prior to winning the lottery. We report estimates using earnings from five years prior, to six years subsequent, to winning the lottery. We are interested in these estimates for two reasons. For the five years prior to winning the lottery the true average causal effect is by definition equal to zero. We wish to see whether least squares estimates are consistent with these known true values. In addition we wish to check whether the estimates for the years subsequent to winning the lottery are sensitive to the exclusion of the other pre-lottery earnings. In the first column we see that estimates for several of the pre-lottery years are significantly different from zero, suggesting that least squares adjustment does not adequately adjust for the pre-lottery differences. The next pair of columns reports results based on adjusting for the first two years of pre-lottery earnings. The estimates for the remaining pre-lottery years are closer to zero, suggesting that with two years least squares adjustment works better. As we adjust for more pre-lottery years,

the significance of the remaining pre-lottery years goes down, and in addition the estimates for the post-lottery years stabilize.

In Table 11 we report propensity score estimates for the same samples and combinations of pre-lottery variables. Now estimates for pre-lottery years are always statistically close to zero.

In Tables 12 and 13 we report least squares and propensity score estimates for the big-winners/small-winners samples. Here the least squares estimates perform very well even adjusted for only few years of pre-lottery earnings. This is why in the body of the paper we do not report propensity score estimates for this sample.

## REFERENCES

- ANGRIST, J., AND A. KRUEGER, (1998), "Empirical Strategies in Labor Economics", forthcoming in *Handbook of Labor Economics*, Ashenfelter and Card, eds.
- ASHENFELTER, O., (1978), "The Labor Supply Response of Wage Earners", in Palmer and Pechman (eds.) *Welfare in Rural Areas: the North Carolina-Iowa Income Maintenance Experiment*, Washington, DC, Brookings, 109-138.
- ASHENFELTER, O., AND HECKMAN, (1974), "The Estimation of Income and Substitution Effects in a Model of Family Labor Supply" *Econometrica*, Vol 42, 73-85.
- BEHRMAN, J., R. POLLAK AND P. TAUBMAN, (1982), "Parental Preferences and Provisions for their Progeny", *Journal of Political Economy*, Vol 90, 52-73.
- BISHOP, J., (1980), "Jobs, Cash Transfers, and Marital Instability", *Journal of Human Resources*, 15, 301-334.
- BLUNDELL, R., AND T. MACURDY, (1998), "Labor Supply", forthcoming, *Handbook of Labor Economics*, Ashenfelter and Card, eds.
- BODKIN, R., (1959), "Windfall Income and Consumption", *American Economic Review*, Vol 49, 602-614.
- BODKIN, R., (1963), "Windfall Income and Consumption: Comment", *American Economic Review*, Vol 53, 445-447.
- CAIN, G. ET AL, (1974), "The Labor-supply Response of Married Women, Husbands Present" *Journal of Human Resources*, Vol 9, 202-222.
- CARD, D., AND SULLIVAN, (1989), "Measuring the Effect of Subsidized Training Programs on Movements into and out of Employment", *Econometrica*, Vol 56, 497-530.
- COX, D. (1958), *Planning of Experiments*, New York, Wiley.
- DEHEJIA, R., AND S. WAHBA, (1998), "Causal Effects in Non-experimental Studies: Re-evaluating the Evaluation of Training Programs" NBER working paper #6829.
- DYNAN, K., J. SKINNER, AND S. ZELDES, (1996), "Do the Rich Save More", unpublished

- manuscript, department of economics, Dartmouth College.
- FISHER, R. A., (1925), *Statistical Methods for Research Workers*, Edinborough.
- GALLUP, J., P. GLEWWE AND M. KREMER, (1997)
- HAUSMAN, J., AND D. WISE, (1977), "Social Experimentation, Truncated Distributions, and Efficient Estimation", *Econometrica*, Vol. 45, No. 4, 919-938.
- HAUSMAN, J., AND D. WISE, (1979) "Attrition Bias in Experimental and Panel Data: The Gary Income Maintenance Experiment" *Econometrica*, Vol. 47: 455-73.
- HAUSMAN, J., AND D. WISE, (1985) "Technical Problems in Social Experimentation: Cost versus Ease of Analysis", in Hausman and Wise (eds.) *Social Experimentation*, p 187-219, Chicago University Press.
- HECKMAN, J., AND R. ROBB, (1984), "Alternative Methods for Evaluating the Impact of Interventions," in Heckman and Singer (eds.), *Longitudinal Analysis of Labor Market Data*, Cambridge, Cambridge University Press.
- HECKMAN, J., AND J. HOTZ, (1989) "Alternative Methods for Evaluating the Impact of Training Programs", (with discussion), *Journal of the American Statistical Association*.
- HECKMAN, J., H. ICHIMURA, AND P. TODD, (1997), "Matching as an Econometric Evaluation Estimator: Evidence from an Evaluation of a Job Training Program", *Review of Economic Studies*, Vol 64, No 4, 605-654.
- HOLLAND, P., (1986), "Statistics and Causal Inference," (with discussion), *Journal of the American Statistical Association*, 81, 945-970.
- HOTZ, J., G. IMBENS AND J. MORTIMER, (1998), "Predicting the Efficacy of Future Training Programs Using Past Experiences", unpublished manuscript, Department of Economics, UCLA.
- KILLINGSWORTH, M., AND J. HECKMAN, (1986), "Female Labor Supply: A Survey", in Ashenfelter and Layard, (eds.), *Handbook of Labor Economics*, p 103-204.
- KREININ, M., (1961), "Windfall Income and Consumption—Additional Evidence" *American Economic Review*, Vol 51, 388-390.

- LALONDE, R. (1988), "Evaluating the Econometric Evaluations of Training Programs", *American Economic Review*.
- LANDSBERGER, M., (1963), "Windfall Income and Consumption", *American Economic Review*, Vol 53, 534–540.
- LANDSBERGER, M., (1970), *Restitution Receipts, Household Savings, and Consumption Behavior in Israel*, Jerusalem Academic Press, Jerusalem.
- LITTLE, R. J. A., AND D. B. RUBIN, (1987), *Statistical Analysis with Missing Data*, Wiley: New York.
- MASTERS, S., (1969), "The effect of Family Income on Children's Education: Some Findings on Inequality of Opportunity", *Journal of Human Resources*, Vol 4, 158–175.
- MOFFITT, R., AND K KEHRER, (1981), "The Effect of Tax and Transfer Programs on Labor Supply: The Evidence from the Income Maintenance Experiment *Research in Labor Economics*, Vol 4.
- MORRILL, W., (1974), "Introduction", Special Issue on Negative Income Tax, *Journal of Human Resources*, Vol 9, 156–157.
- PHILIPSON, T., (1997) "Observational Agency and Supply-Side Econometrics", NBER Technical Working Paper 210.
- REES, A., (1974), "Summary of Results on Negative Income Tax", *Journal of Human Resources*, Vol 9, 158–190.
- ROSENBAUM, P., AND D. RUBIN, (1983a), "The central role of the propensity score in observational studies for causal effects", *Biometrika*, 70, 1, 41–55.
- ROSENBAUM, P., AND D. RUBIN, (1983b), "Assessing Sensitivity to an Unobserved Binary Covariate in an Observational Study with Binary Outcome." *Journal of the Royal Statistical Society*, Series B, Vol 45.
- ROSENBAUM, P., AND D. RUBIN, (1985), "Reducing bias in observational studies using subclassification on the propensity score," *Journal of the American Statistical Association*, Vol 79, 516–524.

- RUBIN, D. (1974), "Estimating Causal Effects of Treatments in Randomized and Non-randomized Studies," *Journal of Educational Psychology*, 66, 688-701.
- RUBIN, D., (1976), "Inference and Missing Data", *Biometrika*, Vol 63, 581-92.
- RUBIN, D. B., (1977), "Assignment to a Treatment Group on the Basis of a Covariate", *Journal of Educational Statistics*, 2, 1-26.
- RUBIN, D., (1978), "Bayesian inference for causal effects: The Role of Randomization", *Annals of Statistics*, 6:34-58.
- RUBIN, D., (1987), *Multiple Imputation for Nonresponse in Sample Surveys*, Wiley, New York.
- SACERDOTE, B., (1996), "The Lottery Winner Survey, Crime and Social Interactions, and Why is there more Crime in Cities", PhD Thesis, Department of Economics, Harvard University.
- SCHULTZ, T. P., (1980), "Estimating Labor Supply Functions for Married Women", in: Smith (ed) *Female Labor Supply*, Princeton University Press, Princeton, NJ.
- SKIDMORE, F., (1974) "Availability of Data from the Graduated Work Incentive Experiment", *Journal of Human Resources*, Vol 9.
- STAFFORD, F., (1985), "Income-Maintenance Policy and Work Effort: Learning from Experiments and Labor-Market Studies", in Hausman and Wise (eds.), *Social Experimentation* 95-134.
- THALER, R., (1990), "Savings, Fungibility, and Mental Accounts", *Journal of Economics Perspectives*, Vol. 4, No. 1, 193-205.
- WATTS, H., R. AVERY, D. ELESH, D. HORNER, M. LEFCOWITZ, J. MAMER, D. POIRIER, S. SPILERMAN, AND S. WRIGHT, (1974) "The Labor Supply Response of Husbands", *Journal of Human Resources*, Vol 9, 181-201.