Estimating the Effect of Unearned Income on Labor Earnings, Savings and Consumption: Evidence from a Survey of Lottery Players[1]

Guido W. Imbens – UCLA[2] and NBER

Donald B. Rubin – Harvard University[3]

Bruce Sacerdote – Dartmouth College[4]

July 2000
July 17, 2000

[1]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 commision 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.

[2]Department of Economics, 8256 Bunche Hall, UCLA, 405 Hilgard Avenue, Los Angeles, CA 90095.

[3]Department of Statistics, Science Center 709, Harvard University, Cambridge, MA 02138.

[4]Department of Economics, 6106 Rockenfeller Hall, Dartmouth College, Hanover, NH 03755.

Estimating the Effect of Unearned Income on Labor Supply, Earnings, Savings and Consumption: Evidence from a Survey of Lottery Players

Imbens, Rubin and Sacerdote

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 often complicated by the fact that exogenous changes in unearned 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, consumption and savings.

We find that unearned income reduces labor earnings, with a marginal propensity to consume leisure of approximately 0.10, for men as well as women. Individuals between 55 and 65 years old when winning the lottery reduce their labor earnings more sharply than older or younger individuals. Individuals with higher earnings before winning the lottery also show greater reductions in labor earnings.

Some of the unearned income is spent on cars and housing. Although the winnings do not appear to affect retirement savings much, other savings go up by about 15 to 20 percent of the winnings to date. Those who won longer ago are saving at a higher rate than recent winners.

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 concerning 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 the effects of unearned income.

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, consumption and savings, and report estimates of the marginal propensity to allocate the unearned income to various categories. Based on a static model of labor supply with Stone-Geary preferences we estimate the marginal propensity to consume leisure to be around 10%. This does not differ significantly between men and women. It is significantly higher for individuals close to the retirement age, but not for individuals already over the retirement age at the time of winning. Some of the lottery winnings are spent on cars, with a marginal propensity to consume of 1.4%, and housing, with a marginal propensity to consume around 3.7%. A significant portion goes into general

(i.e., non-retirement) savings, with a marginal propensity to save around 16%. Individuals who won more recently are estimated to have lower savings rates compared to individuals who won the lottery longer ago, but the savings rates do not appear to vary by age.

Based on our evidence we interpret the results as estimates of the causal effect of lottery prizes on labor earnings, savings and consumption. Two caveats should be kept in mind, however. First, 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. Second, 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 income on labor supply. See Pencavel (1986), and Blundell and MaCurdy (2000) 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 Rees (1974) and the references in Pencavel (1986). 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 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 unanticipated infusion of income.

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. The total prizes range from $22,000 to $9,696,000, with the sample mean and median equal to $1,104,000 and $635,000 respectively.[2] 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.[3] 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, available in Appendix A in Imbens, Rubin and Sacerdote (1999), 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 car values. Second, there are a number of questions regarding the situation at the time they won their prize in the lottery . 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 and provide us with checks on the validity of the inferences. In principle the randomization should ensure that the different groups of winners and losers are comparable at the time of playing the lottery. In practice there are three reasons why this need not be true in our sample. First, individuals buy different numbers of tickets, and the randomization is

[2] All dollar amounts are in 1986 dollars.

[3] 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.

over tickets rather than individuals. Second there are only season ticket holders in the losers sample, and an unknown mix of season ticket holders and single ticket buyers in the winners sample. Third, there is nonresponse which is potentially correlated with individual differences as well as the amount of the prize. Similar concerns, in that case primarily due to nonresponse, arose in the analysis of the NIT experiments (Hausman and Wise, 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 individual characteristics such as 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 preceeding 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

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 and 237 winners. For analyses here involving additional variables (e.g., savings or consumption) 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).

Table 2 presents summary statistics for the variables used in the analyses. For each variable the mean and standard deviation for the entire sample are given in the first two columns. We also present averages separately for the losers and winners, as well as t–statistics for the null hypotheses that the averages for the loser and winner subpopulations are identical. Finally we present averages for the sample of 46 "big" winners, who win more than $100,000 per year (more than $2,000,000 total), and t-statistics for the null hypothesis that the averages for the big winners are different from those for the "small' winners (winners of prizes less than 2,000,000) total. We consider this group separately because in some of the regressions below we exclude the big winners to investigate the sensitivity to their presence.

On average the individuals in our basic sample won yearly prizes of $26,000. Typically they won ten years ago, implying they are on average half way through their twenty years of lottery payments. We asked all individuals how many tickets they bought in a typical week in the year they won the lottery.. 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). 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. 35% of the sample was over 55 years old at that time, and 15% was over 63% of the sample was male. The average level of schooling, calculated as years of high school plus years of college plus 8, is equal to 13.7. 64% claimed at least one year of college. In principle these characteristics should not differ between losers and winners. However, losers are significantly more educated than winners, and they are also older. This is likely to reflect the differences between season ticket holders and single ticket buyers. Note that the differences between all winners and the big winners tend to be smaller,

although there are some large differences in the sex ratio and age.

We observe for each individual in the basic sample social security earnings for six years preceeding the time of winning the lottery, for the year they won (year zero), and for six years 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. Participation rates, as measured by positive social security earnings, gradually decline over the thirteen years, starting at 69% and reaching as high as 71% before going down to 56%. Figures 1 and 2 present graphs for average earnings and the proportion of individuals with positive earnings for the three groups, losers, all winners and big winners. One can see a modest decline in earnings and proportion of individuals with positive earnings for the winner sample compared to the losers, and a sharp and much larger decline for big winners. The t-statistics in Table 2 test the equality of the averages in these graphs.

On average the value of all cars was \$18,200, with an average debt of \$3,000, leading to an average net car value of \$15,2000. For housing the average value was \$166,300, with an average mortgage of \$44,200. 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 aggregated the responses to financial wealth into two 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 and general savings.[4] We construct an additional variable adding up the two savings categories as "total financial wealth".[5] For the losers, wealth in the various savings accounts is considerably higher than

[4]See the appendix in Imbens, Rubin and Sacerdote (1999) for the questionnaire with the exact formulation of the questions.

[5]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.

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.

5. Conceptual Framework and Specification

As a framework for analyzing the data from our survey we use a static model for labor supply where an individual allocates his available time to work and leisure, and the total of earned and unearned income to a number of categories, including consumption as well as savings. To reflect the fact that we do not have wages and prices, we use an augmented Stone-Geary utility function, with the specification:

$$U(l, x_1, \ldots, x_K) = (l - \gamma_0)_0^{\beta} \cdot \prod_{k=1}^{K} (x_k - \gamma_k)_k^{\beta},$$

where l is the amount of leisure, and $x_1, \ldots, X_K$ are the amounts of various commodities, including savings as one of the commodities, and the β_k and γ_k are the preference parameters.

The agent is assumed to maximize this utility function each period subject to the budget constraint

$$wl + \sum_{k=1}^{K} p_k x_k = wT + LP + Y,$$

where w is the wage rate, p_k the price of commodity k, T the maximum number of hours worked, and $Y + LP$ is total unearned income, equal to the lottery prize LP plus other unearned income Y. The expenditure function for commodity k based on this utility function is

$$E_k = p_k x_k = \gamma_k p_k + \beta_k \cdot (wT + LP + Y - \sum_{m=1}^{K} \gamma_k p_k - \gamma_k w), \tag{1}$$

That is, if someone reports positive savings in the category "retirement accounts", 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 30 individuals for mutual funds and general savings. As a result, the average of the two 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 two savings categories.

and the labor earnings function

$$y = wl = \gamma_0 w + \beta_0 \cdot (wT + LP + Y - \sum_{m=1}^{K} \gamma_k p_k - \gamma_k w). \tag{2}$$

Our key assumption is that, at least conditional on some individual characteristics, the amount of lottery winnings is independent of the wages, other unearned income and other determinants of expenditures and labor supply. Even without direct measures of prices and wages we can in that case estimate the β_k, the marginal propensities to consume.

The first regression function we estimate is

$$y_{it} = \alpha_t + \beta_{0t} YP_i + \delta_t' X_i + \varepsilon_{it}, \tag{3}$$

where y_{it} is social security earnings for individualo i in year t after winning the lottery, YP_i is the yearly amount paid by the lottery (equal to the total lottery prize divided by twenty), and X_i is a set of control variables. We estimate this regression function for the year of winning the lottery as well as for the six subsequent year ($t = 0, 1, \ldots, 6$), and for the average of the six post-lottery years, $\sum_{t=1}^{6} y_{it}/6$, in seven different specifications.

First we estimate this equation without any control variables using the entire sample of 496 observations. Second, we estimate the regression function on the same sample but including a small set of individual characteristics typical of the type of individual characteristics one might be able to control for using a typical large survey data set: a dummy for men, years of education, a dummy for some college, age, and dummies for being over 55 and over 65. In the third specification we subtract from the outcome variable earnings in the last pre-lottery year:

$$y_{i,t} - y_{i,-1} = \alpha_t + \beta_{0t} YP_i + \delta_t' X_i + \varepsilon_{it}, \tag{4}$$

still allowing for individual differences, which now could affect the growth rates. In the fourth specification we expand the set of control variables to include lagged earnings and indicators for positive lagged earnings, for all six pre-lottery years, as well as an indicator for pre-lottery

participation in the labor force, number of tickets bought, and time since winning. Given that lagged earnings are included it is immaterial whether the outcomes are differenced or not. In the fifth specification we include a quadratic term in the lottery prize. Here we used the expanded set of control variables and the same basic sample of 496 observations. Instead of reporting the estimate of the quadratic term, we report the derivative of the response function at two values for the lottery prize, namely at zero and at the median of the yearl prize, $32,000.

The last three specifications use the large set of control variables but differ in the sample used. First, in the sixth specification we restrict the sample to winners only. This eliminates any biases that might arise from differences between season ticket buyers and single ticket buyers. In the seventh specification we again include the losers but now use only winners with yearly prizes less than $100,00, thus excluding the 46 big winners. Because some of their prizes are so large, these relatively few big winners may exert a disproportionate effect on the estimates. Finally we exclude both losers and big winners, leaving us with a sample of 191 winners with a prize less than $100,000 per year.

The second set of regression functions we estimate is

$$E_k = \alpha_0 + \beta_k TP + \alpha' X + \varepsilon, \tag{5}$$

where E_k is the amount spent in various categories. Here we use current value of cars, both total and net of loans, and current value of housing, again both total and net of mortgages. We also estimate this regression function with current values of various savings categories on the left-hand side. The three savings categories are retirement savings, other savings, and total savings. For all these outcomes all we observe is the value of these items at the time of the survey. Rather than attribute a fixed portion of that to consumption in that year, we use the total value on the left-hand side and use the total amount of the lottery prize paid up to that time on the right-hand side (TP). Since on average people have been receiving lottery payments for ten years, the amount received is around 50% of the total prize. To

investigate the sensitivity to this specification we also estimate one regression where the total (twenty-year) prize is the right-hand side variable. Withe the amount paid out up to now as the explanatory variable, the estimates can be interpreted as the fraction of the money the lottery has paid out up to that point that is allocated to cars, housing and savings.

The third and fourth set of regressions correspond to the first and second set but allow for interactions of the lottery prize with six background characteristics. We use specification VIII, based on the sample of 191 winners with prizes less than $100,000 only, and include the large set of controls. We interact the lottery prize with an indicator for positive earnings in the year prior to winning the lottery, an indicator for men, an indicator for being between 55 and 65 years old at the time of winning the lottery, an indicator for some college, and the number of years since winning the lottery. All interactions are included in the same regression.

6. Results

6.1 Marginal Propensity to Earn out of Unearned Income

In Table 3 we present the results for the marginal propensity to earn (mpe) out of unearned income, or minus the marginal propensity to consume leisure. As discussed before, the explanatory variable is the yearly lottery prize. In the first row the outcome is the average social security earnings over the six post-lottery years, arguably the most reliable measure of the longterm income effect. The first column gives the result with no control variables, an estimate of -0.051 with a standard error of 0.014. The estimate changes very little when we include the small or large set of regressors, or first difference the earnings variable. This robustness may come as somewhat of a surprise given that Table 2 shows that average characteristics including lagged earnings differ considerably by prize. However, these significant difference do not imply that these characteristics actually explain much of the variation in prizes. The R-squared in a regression of prizes on all individual characteristics and lagged earnings is only 0.17, and if the number of tickets bought is excluded the R-

squared drops to 0.08. Since the number of tickets bought itself is not significantly correlated with earnings, it may be less surprising that the control variables do not affect the results very much. Because some of the control variables, notably lagged earnings, are highly correlated with the outcome here, the standard errors in the specifications III and IV are much lower than those in the first specification without any control variables.

In the fifth specification we add a quadratic term in the prize. Rather than report the coefficient on the quadratic term we report the derivative of the expected earnings as a function of the prize at two values of the prize, zero and the median prize ($32,000 per year). The estimates of the marginal propensity to earn based on this specification are much larger, equal to -0.114 (0.015) at a prize equal to zero, and -0.097 (0.012) at a prize equal to $32,000. Although these estimates appear very close, the quadratic term is in fact highly significant, with a t-statistic equal to 4.8. Because the distribution of prizes is so skewed, with a median yearly prize equal to $32,000 and a maximum equal to $500,000, the few very large observations disproportionally affect the linear regression estimates.

The next specification excludes the 259 losers, more than half the sample. This specification avoids potential biases from the differences between season ticket holders and single ticket buyers, and thus stays closer to the ideal experiment. The results for this specification are very similar to those from specification IV with the same set of control variables. Excluding the big winners (winners with a yearly prize larger than $100,000) does change the results considerably and gets them much closer to the quadratic specification, with an estimate for the marginal propensity to earn of -0.122 (0.020). Finally we exclude both loser and big winners. This again leads to a much larger estimate than the simple linear specification for the entire sample. From the full set of estimates it appears that the linear specification has trouble fitting the response function for the big winners. This can be fixed by either including a quadratic term in the prize, or by excluding the big winners. Once we exclude the big winners including a quadratic term in the prize leads to a t-statistic is 0.6, suggesting that the linear specification fits fairly well.

The next seven rows present the results for the year of winning the the six subsequent years. We mainly consider specification VIII. Since we do not have information on division of earnings in the year of winning between earnings prior to and after winning, one might expect the marginal propensity to earn to be closer to zero for this year than for subequent years. This is confirmed by the data with an estimated mpe of 0.004. However, even during the first full year after winning the lottery the estimated mpe is much lower than in subsequent years, at -0.056 (0.25). It appears to take individuals some time to adjust their labor supply to the desired level. After the first post-lottery year the mpe stabilizes around -0.10, with a typical standard error of around 0.03.

The estimates for the marginal propensity to earn of around -0.10 are not out of line with those in the literature, which is not surprising given the wide range of estimates reported there. Pencavel reports on thirteen studies using non-experimental data for US men, ranging from -0.70 to 0.08. Estimates based on the negative income tax experiments range from -0.29 to 0.02 in his survey. Blundell and MaCurdy find estimates ranging from -0.95 to 0.002 for men, and from -0.33 to 0.27 for women.

In Table 4 we present the results for the mpe with interactions for positive prior earnings, men, age at winning between 55 and 65, age at winning greater than 65, some college and years since winning. For the 71% of the sample who have positive earnings in the year prior to winning the lottery the mpe based on the average post-lottery earnings is -0.209 (0.084) lower thanm for the remainder of the sample who had zero earnings. Similar results are found for the seven post-lottery years separately. More surprisingly we find no significant differences between men and women in terms of the marginal propensity to earn out of unearned income. All estimates are extremely close to zero, with some negative and some positive, and none are significant for the average of post-lottery years or for any of the post-lottery years separately. Men and women do have very different labor market experiences, with in the last pre-lottery year 75% of the men working, with average earnings for those with positive earnings equal to $26,700, and 65% of the women working, with their average

earnings equal to$15,400. These differences do not lead to different estimates of the marginal propensity to earn. We do find differences by age. We experimented with just age interacted with the prize, and with dummies for age between 55 and 65 and age over 65. The latter specification is the one reported here. Individuals between 55 and 65 at the time of winning the lottery reduce their labor earnings significantly more than younger workers. Looking at the average post-lottery earnings, their mpe is lower by -0.167 (0.070). Individuals even older, that is older than 65 at the time of winning do not reduce their earnings as much. The effect of the lottery winnings by age suggests that some individuals reduce their labor supply earlier than they might otherwise have. Regressions using only a simple interaction of age and prize do not show any evidence of age differences in the marginal propensity to consume, possibly missing the negative effect for the middle age group. Having some college education does not affect the marginal propensity to earn out of unearned income, and neither does the timing of the prize. In both cases the interactions are small and insignificant in all seven years as well as in the average earnings measure.

6.2 Consumption and Savings

In Table 5 we report the results for expenditures on cars and housing and savings. Recall that here the explanatory variable is the cumulative lottery prize up to the time of the survey. For cars we find that there is a small but highly signicant effect of lottery prize. As long as the sample includes the big winners the marginal propensity to consume is very precisely estimated to be 0.009 (0.002), meaning that out of the total amount won so far in the the lottery 0.9% is spent on cars (or 0.7% on car values net of loans).[6] If we exclude the big winners the marginal propensity to consume goes up to about 1.4% (0.7%). For housing there is a similar story. As long as the sample includes the big winners the value of housing is significantly affected by the lottery prize, with a marginal propensity to consume of around 3% (1%). If we exclude the big winners the effect is still of similar size, but it is no longer

[6]This is in fact not total expenditure on cars, as we only observe the current value as reported by the individual.

significant. Note that if we use housing values net of mortgages, the effect disappears in all specifications. After winning the lottery people appear to be buying more expensive houses, but they do finance them through correspondingly larger mortgages.

Next, consider the savings outcomes. First we look at the retirement accounts. The basic specifications suggest a large and significant negative effect of unearned income on retirement savings. However, this effect disappears entirely if we drop the losers from the sample. It appears that the loser sample saves considerably more in retirement accounts than the winners. Evidence of this can already be seen in the summary statistics in Table 2 where the losers on average have $92,000 in retirement savings, compared to $34,000 for the winners. Part of this is obviously due to the six year difference in average age between the two groups. However, there is also a three year age gap between the big winners and the other winners, but this does not show up in a difference in average retirement savings. Part of this savings differential may therefore be real, but some of it may also be due to differences in the population of seasonticket holders and single ticket buyers that we do not adequately control for. Looking at other financial savings we do find a positive relation with unearned income. For the same reason we are skeptical of the results for retirement savings that rely on the loser sample, we prefer the estimates based on winners only. Including the big winners we find a marginal propensity to save out of unearned income of 0.039 (0.019). Excluding the big winners the estimate is much higher, at 0.183 (0.061). Given concerns about the reliability of the reports on savings for the biggest savers, which are likely to also be the biggest winners, we view the estimates on the sample excluding the big winners as the most reliable. Adding up the two savings measures leads to similar results, with our preferred estimate of the marginal propensity to save out of unearned income based on the sample without losers and big winners equal to 15.8% (5.6%).

In Table 6 we report estimates for the same outcomes based on interacting the lottery prize with the six individual background variables (prior earnings positive, men, age between 55 and 65, age over 65, college, and years since winning). Having positive prior earnings

reduces the effect of unearned income on car values. It may be that individuals with positive earnings prior to winning the lottery were less constrained in their consumption behavior. Men appear to be saving more from unearned income, with their marginal propensity to save (other than in retirement accounts) larger by 18% (6.7%). There is no evidence that older people save less out of unearned income. For the 55-65 year old group the estimates are negative and close to significant, but for the older group they are in fact positive, although far from significant. Having some college does not affect the marginal propensity to spend on cars or housing, or the marginal propensity to consume. The longer ago someone has won the lottery, and thus the closer to the end of the twenty years of lottery payments, the larger the marginal propensity to save. This is consistent with some consumption smoothing where large items are purchased early on during the period of lottery payments, followed by a period with higher savings.

7. Conclusion

References

Angrist, J., and A. Krueger, (1998), "Empirical Strategies in Labor Economics", forthcoming in *Handbook of Labor Economics*, Ashenfelter and Card, eds.

Ashenfelter, O., and Heckman, (1974), "The Estimation of Income and Substitution Effects in a Model of Family Labor Supply" *Econometrica*, Vol 42, 73–85.

Blundell, R., and T. MaCurdy, (2000), "Labor Supply", *Handbook of Labor Economics*, Ashenfelter and Card, eds.

Bodkin, R., (1959), "Windfall Income and Consumption", *American Economic Review*, Vol 49, 602–614.

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.

Imbens, G., D. Rubin and B. Sacerdote, (1999), "Estimating the Effect of Unearned Income on Labor Earnings, Savings and Consumption: Evidence from a Survey of Lottery Players", NBER Working Paper 7001.

Kaplan,

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.

Landsberger, M., (1963), "Windfall Income and Consumption", *American Economic Review*, Vol 53, 534–540.

Little, R. J. A., and D. B. Rubin, (1987), *Statistical Analysis with Missing Data*, Wiley: New York.

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.

RUBIN, D. B., (1977), "Assignment to a Treatment Group on the Basis of a Covariate", *Journal of Educational Statistics*, 2, 1-26.

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.

THALER, R., (1990), "Savings, Fungibility, and Mental Accounts", *Journal of Economics Perspectives*, Vol. 4, No. 1, 193-205.

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

Variable	All (N=496)		Losers (N=259)	Winners (N=237)		Big Winners (N=46)	
	mean	(s.d.)	mean	mean	t-stat	mean	t-stat
Yearly Prize	26.4	(50.8)	0	55.2	[14.4]	160.0	[20.4]
Year Won	9.8	(1.2)	9.6	9.9	[3.0]	10.1	[1.1]
Tickets Bought	3.3	(2.9)	2.2	4.6	[10.2]	5.0	[0.9]
Age	50.2157	(13.6845)	53.2085	46.9451	[-5.2252]	50.3256	[1.7840]
Age>55	0.3528	(0.4783)	0.4324	0.2658	[-3.9310]	0.3953	[2.1365]
Age>65	0.1492	(0.3566)	0.1931	0.1013	[-2.8841]	0.2093	[2.6221]
Male	0.63	(0.48)	0.67	0.58	[-2.1]	0.84	[3.9]
Years of Schooling	13.7	(2.2)	14.4	13.0	[-7.8]	12.8	[-0.6]
College	0.65	(0.48)	0.78	0.51	[-6.6]	0.53	[0.4]
Working Then	0.78	(0.41)	0.77	0.80	[0.9]	0.86	[1.1]
Earnings Year -6	13.8426	(13.3583)	15.5606	11.9651	[-3.0187]	14.5792	[1.6124]
Earnings Year -5	14.1206	(13.7580)	15.9557	12.1153	[-3.1328]	15.1754	[1.8591]
Earnings Year -4	14.2089	(14.0553)	16.1958	12.0374	[-3.3246]	16.0929	[2.4588]
Earnings Year -3	14.8037	(14.7662)	16.6192	12.8196	[-2.8836]	17.1248	[2.4930]
Earnings Year -2	15.6196	(15.2683)	17.5787	13.4787	[-3.0115]	16.8347	[1.8863]
Earnings Year -1	16.3112	(15.7018)	17.9982	14.4676	[-2.5149]	17.2551	[1.4868]
Earnings Year 0	16.0524	(15.7999)	18.2495	13.6515	[-3.2689]	13.7787	[0.0691]
Earnings Year 1	15.4083	(16.2157)	18.4949	12.0352	[-4.5177]	9.4545	[-1.4154]
Earnings Year 2	14.6859	(16.3457)	17.6500	11.4465	[-4.2958]	8.4394	[-1.6174]
Earnings Year 3	14.1814	(16.2855)	17.1406	10.9474	[-4.3048]	8.7371	[-1.1963]
Earnings Year 4	13.7961	(16.2641)	16.8904	10.4146	[-4.5154]	7.5318	[-1.6097]
Earnings Year 5	13.6247	(16.3155)	16.6504	10.3181	[-4.3968]	7.7901	[-1.3948]
Earnings Year 6	13.2302	(16.4209)	15.7680	10.4568	[-3.6425]	6.8063	[-1.9714]
Pos Earnings Year -6	0.6935	(0.4615)	0.6873	0.7004	[0.3170]	0.6977	[-0.0433]
Pos Earnings Year -5	0.7097	(0.4544)	0.6795	0.7426	[1.5466]	0.6512	[-1.5171]
Pos Earnings Year -4	0.7117	(0.4534)	0.6911	0.7342	[1.0565]	0.7209	[-0.2164]
Pos Earnings Year -3	0.7036	(0.4571)	0.6757	0.7342	[1.4252]	0.7442	[0.1635]
Pos Earnings Year -2	0.7077	(0.4553)	0.6757	0.7426	[1.6384]	0.6977	[-0.7428]
Pos Earnings Year -1	0.7137	(0.4525)	0.6911	0.7384	[1.1627]	0.6977	[-0.6693]
Pos Earnings Year 0	0.7097	(0.4544)	0.6873	0.7342	[1.1491]	0.6977	[-0.5968]
Pos Earnings Year 1	0.6794	(0.4672)	0.6795	0.6793	[-0.0050]	0.4884	[-3.0091]
Pos Earnings Year 2	0.6310	(0.4830)	0.6409	0.6203	[-0.4758]	0.4186	[-3.0578]
Pos Earnings Year 3	0.6008	(0.4902)	0.6178	0.5823	[-0.8049]	0.3953	[-2.7801]
Pos Earnings Year 4	0.5827	(0.4936)	0.6100	0.5527	[-1.2921]	0.3256	[-3.3763]
Pos Earnings Year 5	0.5867	(0.4929)	0.5946	0.5781	[-0.3729]	0.3488	[-3.4327]
Pos Earnings Year 6	0.5625	(0.4966)	0.5714	0.5527	[-0.4183]	0.3256	[-3.3763]
Car Value	18.2305	(17.7915)	16.6527	19.9549	[2.0229]	29.6250	[3.4864]
Net Car Value	15.5190	(14.8528)	15.3320	15.7146	[0.2595]	25.7203	[4.0333]
Housing Value	166.2986	(111.5519)	174.9142	156.8623	[-1.7822]	218.1190	[4.4361]
Net Housing Value	122.1247	(95.4538)	144.6135	97.5914	[-5.4396]	112.2500	[1.3584]
Retirement Accounts	64.6640	(102.7855)	92.6247	34.4290	[-6.1446]	34.5609	[0.0166]
Other Financial Ass.	84.3045	(151.9460)	91.7815	76.0972	[-1.0951]	127.0805	[2.0073]
Total Financial Ass.	133.3673	(192.4901)	164.4541	99.3948	[-3.8110]	150.9082	[2.0481]

Table 3: Estimates of Marginal Propensity to Earn out of Unearned Income: Yearly Lottery Payments as Right-hand Side Variable

	I	II	III	IV	V	VI	VII	VIII
Ave Post Lottery Earn	-0.051	-0.052	-0.048	-0.051	-0.114	-0.097	-0.043	-0.122	-0.101
	(0.014)	(0.013)	(0.009)	(0.008)	(0.015)	(0.012)	(0.010)	(0.020)	(0.029)
Year 0 Earn	-0.019	-0.022	-0.017	-0.020	-0.038	-0.033	-0.015	-0.024	0.004
	(0.014)	(0.013)	(0.004)	(0.004)	(0.008)	(0.007)	(0.006)	(0.010)	(0.015)
Year 1 Earn	-0.048	-0.049	-0.045	-0.050	-0.103	-0.089	-0.038	-0.094	-0.056
	(0.014)	(0.014)	(0.007)	(0.007)	(0.014)	(0.011)	(0.009)	(0.017)	(0.025)
Year 2 Earn	-0.052	-0.054	-0.050	-0.054	-0.114	-0.098	-0.045	-0.117	-0.092
	(0.014)	(0.014)	(0.009)	(0.009)	(0.016)	(0.013)	(0.011)	(0.021)	(0.031)
Year 3 Earn	-0.051	-0.053	-0.048	-0.053	-0.118	-0.100	-0.043	-0.134	-0.117
	(0.014)	(0.014)	(0.010)	(0.009)	(0.017)	(0.014)	(0.012)	(0.023)	(0.033)
Year 4 Earn	-0.056	-0.057	-0.052	-0.055	-0.127	-0.107	-0.044	-0.151	-0.133
	(0.014)	(0.013)	(0.010)	(0.010)	(0.019)	(0.015)	(0.012)	(0.024)	(0.034)
Year 5 Earn	-0.052	-0.050	-0.046	-0.050	-0.117	-0.099	-0.041	-0.137	-0.116
	(0.014)	(0.013)	(0.011)	(0.011)	(0.020)	(0.016)	(0.013)	(0.026)	(0.036)
Year 6 Earn	-0.050	-0.049	-0.045	-0.046	-0.106	-0.090	-0.047	-0.101	-0.094
	(0.014)	(0.013)	(0.012)	(0.011)	(0.021)	(0.017)	(0.013)	(0.027)	(0.037)
Individual Controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lagged Earnings	No	No	No	Yes	Yes	Yes	Yes	Yes
Losers	Yes	Yes	Yes	Yes	Yes	No	Yes	No
Yearly Prize >100	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Quadratic in Prize	No	No	No	No	Yes	No	No	No
Difference Outcomes	No	No	Yes	Yes	Yes	Yes	Yes	Yes

Table 4: Estimates of Marginal Propensity to Consume Out of Unearned Income: Interactions with Prior Labor Market History, Sex, Age, Education and Time Since Winning

Outcome	Prior Earnings Positive	Male	55<Age<=65	Age>65	College	Years Since Winning
Ave Post Lottery Earn	-0.209	-0.002	-0.167	-0.001	0.037	-0.010
	(0.084)	(0.057)	(0.070)	(0.090)	(0.061)	(0.022)
Year 0 Earn	-0.014	0.015	-0.094	-0.004	-0.027	0.006
	(0.045)	(0.031)	(0.038)	(0.049)	(0.033)	(0.012)
Year 1 Earn	-0.108	-0.057	-0.204	-0.045	0.043	0.001
	(0.073)	(0.050)	(0.061)	(0.079)	(0.053)	(0.019)
Year 2 Earn	-0.175	-0.020	-0.215	-0.039	0.086	-0.025
	(0.088)	(0.060)	(0.073)	(0.095)	(0.064)	(0.024)
Year 3 Earn	-0.225	0.058	-0.178	0.003	0.040	-0.004
	(0.097)	(0.066)	(0.081)	(0.104)	(0.070)	(0.026)
Year 4 Earn	-0.158	-0.005	-0.100	0.099	0.009	-0.024
	(0.103)	(0.070)	(0.085)	(0.110)	(0.074)	(0.027)
Year 5 Earn	-0.235	0.000	-0.127	0.032	-0.001	-0.002
	(0.108)	(0.074)	(0.090)	(0.116)	(0.078)	(0.029)
Year 6 Earn	-0.355	0.009	-0.177	-0.057	0.045	-0.009
	(0.110)	(0.075)	(0.091)	(0.118)	(0.079)	(0.029)

Table 5: Estimates of Marginal Propensity to Consume out of Unearned Income: Accumulated Lottery Payments as Right-hand Side Variable

	I	II	III	IV	V	VI	VII	VIII
Value Cars	0.008	0.008	0.008	0.009	0.009	0.009	0.011	0.014
	(0.001)	(0.001)	(0.002)	(0.003)	(0.003)	(0.002)	(0.004)	(0.007)
Net Value Cars	0.006	0.006	0.006	0.006	0.006	0.007	0.007	0.013
	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.002)	(0.004)	(0.006)
Value House	0.027	0.034	0.032	0.012	0.013	0.041	0.011	0.037
	(0.009)	(0.009)	(0.009)	(0.019)	(0.018)	(0.010)	(0.023)	(0.027)
Net Value House	-0.009	-0.002	-0.005	-0.025	-0.025	0.010	-0.019	0.024
	(0.008)	(0.008)	(0.008)	(0.016)	(0.016)	(0.008)	(0.021)	(0.024)
Retirement Accounts	-0.026	-0.015	-0.015	-0.047	-0.046	-0.003	-0.049	-0.002
	(0.009)	(0.009)	(0.008)	(0.017)	(0.017)	(0.006)	(0.022)	(0.019)
Other Financial Ass.	0.031	0.041	0.039	0.065	0.064	0.039	0.094	0.183
	(0.012)	(0.012)	(0.013)	(0.027)	(0.026)	(0.019)	(0.034)	(0.061)
Total Financial Ass.	0.008	0.027	0.028	0.023	0.023	0.042	0.042	0.158
	(0.016)	(0.015)	(0.016)	(0.032)	(0.031)	(0.019)	(0.040)	(0.056)
Individual Controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lagged Earnings	No	No	No	Yes	Yes	Yes	Yes	Yes
Losers	Yes	Yes	Yes	Yes	Yes	No	Yes	No
Yearly Prize >100	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Quadratic in Prize	No	No	No	No	Yes	No	No	No
Difference Outcomes	No	No	Yes	Yes	Yes	Yes	Yes	Yes

Table 6: Estimates of Marginal Propensity to Consume Out of Unearned Income: Interactions with Prior Labor Market History, Sex, Age, Education and Time Since Winning

Outcome	Prior Earnings Positive	Male	55<Age<=65	Age>65	College	Years Since Winning
Value Cars	-0.026	-0.001	0.013	-0.012	-0.008	-0.000
	(0.012)	(0.008)	(0.010)	(0.013)	(0.008)	(0.003)
Net Value Cars	-0.001	0.012	-0.004	0.012	-0.005	0.005
	(0.010)	(0.008)	(0.009)	(0.012)	(0.007)	(0.003)
Value House	0.080	-0.018	0.002	-0.012	0.006	-0.013
	(0.063)	(0.044)	(0.050)	(0.078)	(0.048)	(0.018)
Net Value House	0.024	-0.016	0.033	-0.061	-0.005	-0.001
	(0.058)	(0.040)	(0.045)	(0.071)	(0.044)	(0.016)
Retirement Accounts	-0.015	-0.030	0.013	-0.032	-0.065	-0.020
	(0.064)	(0.046)	(0.055)	(0.085)	(0.047)	(0.017)
Other Financial Ass.	-0.080	0.182	-0.132	0.028	0.005	0.105
	(0.092)	(0.067)	(0.079)	(0.120)	(0.069)	(0.026)
Total Financial Ass.	-0.049	0.132	-0.092	-0.070	-0.071	0.064
	(0.110)	(0.077)	(0.087)	(0.136)	(0.083)	(0.030)

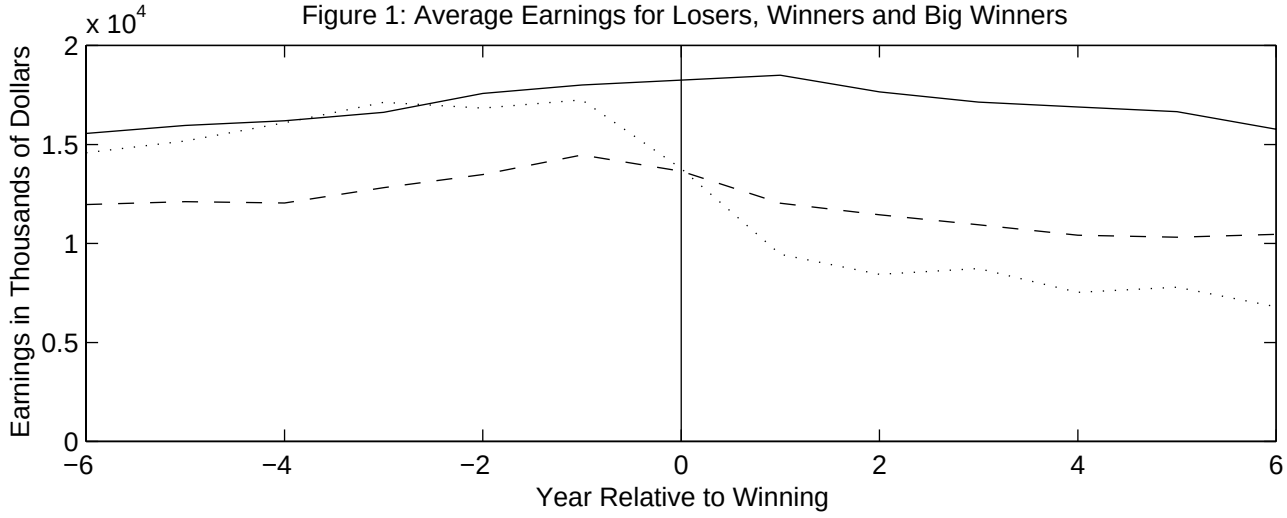
Figure 1: Average Earnings for Losers, Winners and Big Winners
x 10^4
Earnings in Thousands of Dollars
2
1.5
1
0.5
0
−6
−4
−2
0
2
4
6
Year Relative to Winning

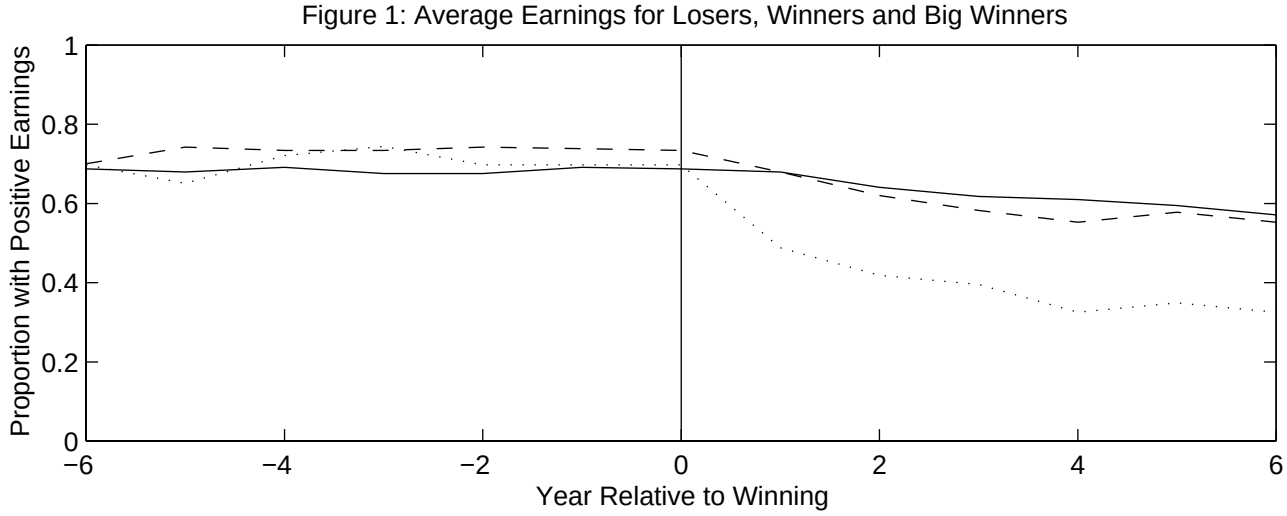
Figure 1: Average Earnings for Losers, Winners and Big Winners
Proportion with Positive Earnings
1
0.8
0.6
0.4
0.2
0
−6
−4
−2
0
2
4
6
Year Relative to Winning