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

By GUIDO W. IMBENS, DONALD B. RUBIN, AND BRUCE I. SACERDOTE*

This paper provides empirical evidence about the effect of unearned income on earnings, consumption, and savings. Using an original survey of people playing the lottery in Massachusetts in the mid-1980's, we analyze the effects of the magnitude of lottery prizes on economic behavior. The critical assumption is that among lottery winners the magnitude of the prize is randomly assigned. We find that unearned income reduces labor earnings, with a marginal propensity to consume leisure of approximately 11 percent, with larger effects for individuals between 55 and 65 years old. After receiving about half their prize, individuals saved about 16 percent. (*JEL* C81, D12, E21, J22, J26)

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, William 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). Although welfare programs are typically a combination of lump-sum grants and tax rates, the effect of unearned income on labor supply is at least part of what is needed to evaluate such programs. 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-1980's, including both winners of large prizes and people who won small, one-time prizes.

We investigate the relationship between the magnitude of the prize 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. In the context of a

* Imbens: Department of Economics, 8256 Bunche Hall, UCLA, 405 Hilgard Avenue, Los Angeles, CA 90095 (e-mail: imbens@econ.ucla.edu), Berkeley, and NBER; Rubin: Department of Statistics, Science Center 709, Harvard University, Cambridge, MA 02138 (e-mail: rubin@hustat.harvard.edu); Sacerdote: Department of Economics, 6106 Rockenfeller Hall, Dartmouth College, Hanover, NH 03755 (e-mail: Bruce.I.Sacerdote@Dartmouth.edu). We are grateful for support by the National Science Foundation through Grant Nos. 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, Susan Athey, Gary Chamberlain, Charles Clotfelter, Paul Devereux, Kei Hirano, Dean Hyslop, Larry Katz, Alan Krueger, John Pencavel, and participants in seminars at MIT, Princeton University, Columbia University, UCLA, Brown University, Stanford University, the University of Virginia, Tilburg University, the European University Institute in Florence, UCL, the Malinvaud seminar, the University of Toulouse, and the NBER summer institute. Finally, we acknowledge superb and tireless research assistance from David Grossman, Aditi Shrikhande, and Eduardo Fajnzylber. We alone are responsible for any errors.

standard life-cycle model of labor supply with Stone-Geary preferences, we estimate the marginal propensity to earn out of unearned income to be approximately −11 percent. This does not differ significantly between men and women. It is significantly greater in magnitude for individuals close to retirement age, but not for individuals already over 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 percent, and housing, with a marginal propensity to consume around 3.7 percent. Approximately 16 percent of the prize money accumulated so far (that is on average 10 years into the 20 years of payments) goes into general savings. As predicted by the theory, individuals who won more recently are estimated to have lower savings rates compared with individuals who won the lottery longer ago but, somewhat surprisingly, the savings rates do not seem to vary by age.

These results are robust against a variety of specifications and we conclude that they can be interpreted as estimates of the causal effect of lottery prizes on labor earnings, savings and consumption. Two caveats should be kept in mind, however. First, the population of lottery players is not necessarily representative of the U.S. population. To investigate this further, we compare our sample to the New England subsample of the Current Population Survey. We find that in our lottery sample, middle-aged people are overrepresented, consistent with findings of other studies of lottery players. Conditional on age, gender, and education, their earnings are somewhat lower than those of the general population, but least-squares estimates of the returns to education are similar in both populations. A second caveat is that responses to lottery prizes need not be typical of responses to other forms of unearned income such as government-provided cash assistance—what Richard Thaler (1990) refers to as fungibility. It is likely, however, that the response to lottery prizes is informative about the response to other types of unearned income. The finding that our estimates of the marginal propensity to earn out of unearned income are in line with those of nonexperimental studies is supportive of this interpretation.

I. Literature

There is a large literature concerned with estimating the effect of unearned income on labor supply. See John Pencavel (1986) and Richard Blundell and Thomas MaCurdy (2000) for surveys of men and Mark Killingsworth and James Heckman (1986) for a survey of women. Most of the studies utilized data from large, representative 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 in such data sets. Researchers have often used capital income or spousal-labor earnings, but the assumption that these are exogenous to labor supply decisions is tenuous.

Another strand of the literature on estimation of income effects analyzed experimental data with clearly exogenous components of unearned income. In the early 1970's several negative income tax (NIT) experiments were conducted in the United States in which a selected population received randomly assigned tax schedules characterized by a guarantee level combined with a tax rate.[1] Although the NIT experiments provided valuable and relatively uncontroversial estimates due to the randomized assignment, their value is limited by 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. Additional limitations stem from the modest size of the amounts of income randomly assigned, as well as from the attrition in the sample over time.[2]

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 are Mordechai Kreinin (1961) and Michael Landsberger (1963), who looked at one-time war reparations paid to Israeli citizens by the German government; Ronald Bodkin (1959), who looked at one-time payments by the U.S. government to selected service men after World War II; and Douglas Holtz-Eakin et al. (1993), who looked at the effects of inheritances on employment.[3] Compared with these studies in this literature, the physical randomization of the lottery strengthens

[1] See, for example, Albert Rees (1974) and the references in Pencavel (1986).

[2] See, for example, Jerry Hausman and David Wise (1985).

[3] See Joshua Angrist and Alan Krueger (2000) for a general discussion of natural experiments.

our case for exogeneity, although like many other natural experiments, a limitation of our study is the potential lack of representativeness of the populations studied.

Finally, as in the current paper, H. Roy Kaplan (1985) analyzed a survey of lottery winners. Kaplan, however, only collected data on economic behavior immediately prior and immediately subsequent to the lottery winning, with limited controls. In contrast, we have six years of accurate post-lottery earnings data from the Social Security Administration and detailed background variables.

II. The Data

Our lottery data set consists of two samples, the "winners" sample and the "nonwinners" 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 20 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.[4] The "nonwinners" 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 $100 to $5,000.[5] The people in this sample are for simplicity referred to as the "nonwinners," although it should be stressed that they did actually win small, one-time prizes.

A. *The Survey*

The survey questionnaire, available in Appendix A in the working paper (Imbens et al., 1999), consists of three sets of questions, the first concerning outcomes at the time of the survey, the second concerning economic behavior and background characteristics at the time of winning, and the third concerning earnings. The first set of questions is about the current (time of survey) circumstances 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. Estimating the effect of the lottery prize on these outcomes is one of the primary goals of the current study.

Second, there are a number of questions concerning background characteristics and economic behavior of the winners at the time they won their prize in the lottery. There are three reasons for including these questions. First, we wish to investigate the heterogeneity of the income effects by individual characteristics such as gender, prelottery labor market status, and age. Second, the inclusion of control variables can improve the precision of the estimates just as in a randomized experiment. Third, and most important, the variables can be used to make the inferences more credible and provide us with checks on the validity of the inferences. In principle, the randomization should ensure that the subsamples of winners and nonwinners 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, the randomization is over tickets and individuals buy different numbers of tickets. Second, there are only season ticket holders in the nonwinners sample and an unknown mix of season and single ticket buyers in the winners sample. Third, there is nonresponse, which may be correlated with individual differences as well as the prize. Similar concerns arose in the analysis of the NIT experiments (e.g., Hausman and Wise, 1985). The covariates can be used to help adjust for such differences. Note that of these three arguments only the nonresponse argument is relevant for biases in analyses involving only winners, and for this reason we limit some of the analyses to this subsample.

The third set of questions concerns labor earnings. We asked respondents to authorize the release of their Social Security earnings records to us. 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.[6]

The survey was conducted in three stages. In July 1995 we sent out by regular mail pilot surveys to 50 winners and 50 nonwinners 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

[4] All dollar amounts are converted to 1986 dollars.

[5] 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.

[6] 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 response rates.

TABLE 1—RESPONSE RATES BY MAILING

Mailing	Date	Sent		Responses		Response rates		
		Winners	Nonwinners	Winners	Nonwinners	Winners	Nonwinners	Total
Pilot	July '95	50	50	17	25	0.34	0.50	0.42
Main	July '96	752	637	272	262	0.36	0.41	0.38
Follow-up ($50 check)	Sept. '96	248	248	39	40	0.16	0.16	0.16
Follow-up ($10 cash, $40 check)	Sept. '96	49	49	11	12	0.22	0.24	0.23
Total		802	687	339	339	0.42	0.49	0.46

Notes: For each of the mailings, we report the number of surveys mailed out and received separately for the nonwinners and winners, as well as the response rates. In the follow-up, we offered some people a $50 check and for others included a $10 bill in the mailing, combined with an offer for a $40 check in case of response. The follow-up was done using Federal Express.

nonwinners. Finally, in September 1996 we sent out reminders to 297 nonresponding winners and 297 nonresponding nonwinners. 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 nonwinners were sent ten dollars in cash and were offered a check for an additional 40 dollars in exchange for returning the survey.[7] The other 248 winners and 248 nonwinners approached in the follow-up were offered a check for 50 dollars for returning the survey.

Table 1 summarizes the response rates for the different mailings. The overall response rate is approximately 46 percent, somewhat higher for nonwinners at 49 percent than for winners at 42 percent.[8] 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 $[0.38 + (1 - 0.38) \times 0.23] \times 100$ percent = 53 percent, rather than the actual 46 percent response rate.

B. *Summary Statistics*

Our basic sample for the analyses presented below consists of individuals with complete answers to the questions on selected pre-lottery conditions (i.e., 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 nonwinners and 237 winners. For analyses 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. 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 investigate alternative approaches to missing data involving models for nonresponse and multiple imputation.[9]

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

[7] Incentive schemes where potential respondents are paid prior to responding were previously implemented in Thomas Philipson (1997), who discusses the merits of such schemes in detail.

[8] One might have expected a lower response rate for winners because, with the incentives equal in absolute terms for winners and nonwinners, the relative incentives are much lower for winners. On the other hand, for the winners the addresses are almost guaranteed to be up to date, whereas it is likely that some of the addresses for nonwinners are out of date. 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 percent). The 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.

[9] See, for example, Roderick Little and Rubin (1987) and Rubin (1987).

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

	All (N = 496)		Nonwinners (N = 259)	Winners (N = 237)		Big winners (N = 43)	
Variable	Mean	(SD)	Mean	Mean	[t-stat]	Mean	[t-stat]
Yearly prize	26.4	(50.8)	0	55.2	[14.4]	160.0	[20.4]
Year won	1986.2	(1.2)	1986.4	1986.1	[−3.0]	1985.9	[−1.1]
Tickets bought	3.3	(2.9)	2.2	4.6	[10.2]	5.0	[0.9]
Age	50.2	(13.7)	53.2	46.9	[−5.2]	50.3	[1.8]
Age > 55	0.35	(0.48)	0.43	0.27	[−3.9]	0.40	[2.1]
Age > 65	0.15	(0.36)	0.19	0.10	[−2.9]	0.21	[2.6]
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	13.8	(13.4)	15.6	12.0	[−3.0]	14.6	[1.6]
Earnings year −5	14.1	(13.8)	16.0	12.1	[−3.1]	15.2	[1.9]
Earnings year −4	14.2	(14.1)	16.2	12.0	[−3.3]	16.1	[2.5]
Earnings year −3	14.8	(14.8)	16.6	12.8	[−2.9]	17.1	[2.5]
Earnings year −2	15.6	(15.3)	17.6	13.5	[−3.0]	16.8	[1.9]
Earnings year −1	16.3	(15.7)	18.0	14.5	[−2.5]	17.3	[1.5]
Earnings year 0	16.1	(15.8)	18.2	13.7	[−3.3]	13.8	[0.1]
Earnings year 1	15.4	(16.2)	18.5	12.0	[−4.5]	9.5	[−1.4]
Earnings year 2	14.7	(16.3)	17.7	11.4	[−4.3]	8.4	[−1.6]
Earnings year 3	14.2	(16.3)	17.1	10.9	[−4.3]	8.7	[−1.2]
Earnings year 4	13.8	(16.3)	16.9	10.4	[−4.5]	7.5	[−1.6]
Earnings year 5	13.6	(16.3)	16.7	10.3	[−4.4]	7.8	[−1.4]
Earnings year 6	13.2	(16.4)	15.8	10.5	[−3.6]	6.8	[−2.0]
Positive earnings year −6	0.69	(0.46)	0.69	0.70	[0.3]	0.70	[−0.0]
Positive earnings year −5	0.71	(0.45)	0.68	0.74	[1.5]	0.65	[−1.5]
Positive earnings year −4	0.71	(0.45)	0.69	0.73	[1.1]	0.72	[−0.2]
Positive earnings year −3	0.70	(0.46)	0.68	0.73	[1.4]	0.74	[0.2]
Positive earnings year −2	0.71	(0.46)	0.68	0.74	[1.6]	0.70	[−0.7]
Positive earnings year −1	0.71	(0.45)	0.69	0.74	[1.2]	0.70	[−0.7]
Positive earnings year 0	0.71	(0.45)	0.69	0.73	[1.1]	0.70	[−0.6]
Positive earnings year 1	0.68	(0.47)	0.68	0.68	[−0.0]	0.49	[−3.0]
Positive earnings year 2	0.63	(0.48)	0.64	0.62	[−0.5]	0.42	[−3.0]
Positive earnings year 3	0.60	(0.49)	0.62	0.58	[−0.8]	0.40	[−2.8]
Positive earnings year 4	0.58	(0.49)	0.61	0.55	[−1.3]	0.33	[−3.4]
Positive earnings year 5	0.59	(0.49)	0.59	0.58	[−0.4]	0.35	[−3.4]
Positive earnings year 6	0.56	(0.50)	0.57	0.55	[−0.4]	0.33	[−3.4]
Car value	18.2	(17.8)	16.7	20.0	[2.0]	29.6	[3.5]
Net car value	15.5	(14.9)	15.3	15.7	[0.3]	25.7	[4.0]
Housing value	166.3	(111.6)	174.9	156.9	[−1.8]	218.1	[4.4]
Net housing value	122.1	(95.5)	144.6	97.6	[−5.4]	112.3	[1.4]
Retirement accounts	64.7	(102.8)	92.6	34.4	[−6.1]	34.6	[0.0]
Other financial assets	84.3	(151.9)	91.8	76.1	[−1.1]	127.1	[2.0]
Total financial assets	133.4	(192.5)	164.5	99.4	[−3.8]	150.9	[2.0]

Notes: The first two columns report the sample average and standard deviation for the basic sample of 496. For the consumption and savings variables the sample size is slightly smaller due to item nonresponse. The third and fourth columns report sample averages for the nonwinners and winners respectively, with the fifth column the t-statistic for the null hypothesis that the averages for winners and nonwinners are identical. The sixth column reports the sample average for the 43 big winners (winners with a yearly prize at least \$100,000), and the seventh column reports the t-statistic for the null hypothesis that the average for the big winners is the same as the average for the other winners.

averages separately for the nonwinners and winners, as well as t-statistics for the null hypotheses that the averages for the nonwinner and winner subpopulations are identical. Finally, we present averages for the subsample of 43 "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.[10]

[10] We consider this group separately because in some of the regressions below we exclude the big winners to investigate the sensitivity of the estimates to their presence. This small group of big winners have average yearly prizes of \$160,000. Excluding these 43 big winners, the average yearly prize for the winners drops from \$55,200 to \$31,500.

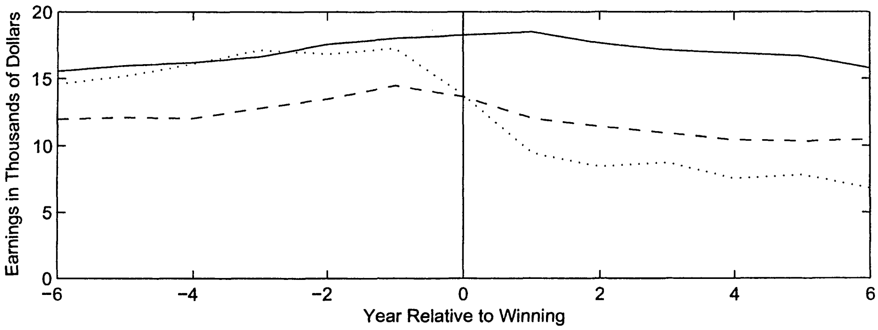

FIGURE 1. AVERAGE EARNINGS FOR NONWINNERS, WINNERS, AND BIG WINNERS

Note: Solid line = nonwinners; dashed line = winners; dotted line = big winners.

On average the individuals in our basic sample won yearly prizes of $26,000 (averaged over the $55,000 for winners and zero for nonwinners). Typically they won 10 years prior to completing our survey in 1996, implying they are on average halfway through their 20 years of lottery payments when they responded in 1996. We asked all individuals how many tickets they bought in a typical week in the year they won the lottery.[11] As expected, the number of tickets bought is considerably higher for winners than for nonwinners. 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 percent of the sample was over 55 and 15 percent was over 65 years old at the time of winning; 63 percent of the sample was male. The average number of years of schooling, calculated as years of high school plus years of college plus 8, is equal to 13.7; 64 percent claimed at least one year of college.

We observe, for each individual in the basic sample, Social Security earnings for six years preceding 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 then decline back to $13,290 over the post-winning period. For those with positive Social Security earnings, average earnings rise over the entire 13-year period from $20,180 to $24,300. Participation rates, as measured by positive Social Security earnings, gradually decline over the 13 years, starting at around 70 percent before going down to 56 percent. Figures 1 and 2 present graphs for average earnings and the proportion of individuals with positive earnings for the three groups, nonwinners, winners, and big winners. One can see a modest decline in earnings and proportion of individuals with positive earnings for the full winner sample compared to the nonwinners after winning the lottery, and a sharp and much larger decline for big winners at the time of winning. A simple difference-in-differences type estimate of the marginal propensity to earn out of unearned income (mpe) can be based on the ratio of the difference in the average change in earnings before and after winning the lottery for two groups and the difference in the average prize for the same two groups. For the winners, the difference in average earnings over the six post-lottery years and the six pre-lottery years is −$1,877 and for the nonwinners the average change is $448. Given a difference in average prize of $55,000 for the winner/nonwinners comparison, the estimated mpe is $(-1{,}877 - 448)/(55{,}000 - 0) = -0.042$ (SE 0.016). For the big-winners/small-winners comparison, this estimate is −0.059 (SE 0.018). In Section IV we report estimates for this quantity using more sophisticated analyses.

On average the value of all cars was $18,200. For housing the average value was $166,300, with an average mortgage of $44,200.[12] We aggregated the responses to financial wealth into two categories. The first concerns retirement

[11] Because there were some extremely large numbers (up to 200 tickets per week), we transformed this variable somewhat arbitrarily by taking the minimum of the number reported and ten. The results were not sensitive to this transformation.

[12] Note that this is averaged over the entire sample, with zeros included for the 7 percent of respondents who reported not owning their homes.

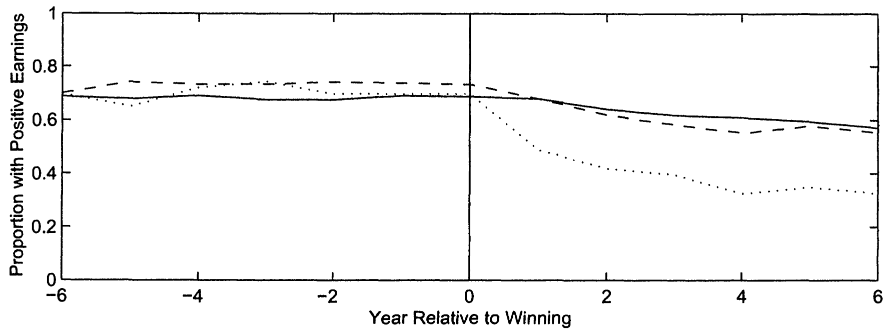

FIGURE 2. PROPORTION WITH POSITIVE EARNINGS FOR NONWINNERS, WINNERS, AND BIG WINNERS

Note: Solid line = nonwinners; dashed line = winners; dotted line = big winners.

type accounts, including IRA's, 401(k) plans, and other retirement-related savings. The second consists of stocks, bonds, and mutual funds and general savings.[13] We construct an additional variable "total financial wealth," adding up the two savings categories.[14] Wealth in the various savings accounts is somewhat higher than net wealth in housing, $133,000 versus $122,000. 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 percent zeros.

The critical assumption underlying our analysis is that the magnitude of the lottery prize is random. Given this assumption the background characteristics and pre-lottery earnings should not differ significantly between nonwinners and winners. However, the *t*-statistics in Table 1 show that nonwinners are significantly more educated than winners, and they are also older. This likely reflects the differences between season ticket holders and single ticket buyers as the differences between all winners and the big winners tend to be smaller.[15] To investigate further whether the assumption of random assignment of lottery prizes is more plausible within the more narrowly defined subsamples, we regressed the lottery prize on a set of 21 pre-lottery variables (years of education, age, number of tickets bought, year of winning, earnings in six years prior to winning, dummies for sex, college, age over 55, age over 65, for working at the time of winning, and dummies for positive earnings in six years prior to winning). Testing for the joint significance of all 21 covariates in the full sample of 496 observations led to a chi-squared statistic of 99.9 (dof 21), highly significant ($p < 0.001$). In the sample of 237 winners, the chi-squared statistic was 64.5, again highly significant ($p < 0.001$). In the sample of 193 small winners, the chi-squared statistic was 28.6, not significant at the 10-percent level. This provides some support for assumption of random assignment of the lottery prizes, at least within the subsample of small winners.

[13] See the Appendix in Imbens et al. (1999) for the questionnaire with the exact formulation of the questions.

[14] 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 "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 and 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.

[15] Although the differences between small and big winners are smaller than those between winners and losers, some of them are still significant. The most likely cause is the differential nonresponse by lottery prize. Because we do know for all individuals, respondents or nonrespondents, the magnitude of the prize, 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.

TABLE 3—COMPARISONS OF LOTTERY SAMPLE AND CPS

	Lottery		CPS-raw		CPS-I		CPS-II	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
Male	0.63	(0.48)	0.47	(0.50)	0.63	—	0.63	—
Age	47.0	(13.6)	43.2	(17.0)	47.0	—	47.0	—
Education	13.7	(2.2)	12.7	(2.6)	12.7	(2.6)	13.7	—
Earnings 1983	14.8	(14.6)	11.6	(12.1)	15.0	(12.5)	17.1	(12.7)
Positive earnings 1983	0.71	(0.46)	0.71	(0.45)	0.76	(0.43)	0.80	(0.40)

Notes: Lottery: Averages and standard deviations for lottery sample of size 496. CPS-raw: Averages and standard deviations for 3,778 CPS observations from New England states in 1983 with age between 19 and 81, unweighted. CPS-I: Weighted averages of CPS-raw sample with weights to match age and sex distribution with lottery sample. CPS-II: Weighted average of CPS-raw sample to match age, sex, and education distribution with lottery sample.

C. *Comparison with the Current Population Survey*

To provide the appropriate context for the results discussed below we compare our sample to a more representative sample of the general population drawn from the Current Population Survey (CPS). We focus on a comparison of the joint distribution of gender, age, education, and earnings in 1983. We use 1983 as the comparison year because it is the last year prior to anyone in the lottery sample winning so that earnings for the lottery sample cannot yet be affected by the prize. As the measure of education in the lottery sample we add 8 to the sum of the years of college and high school reported. For the CPS education variable we take the highest grade attended and subtract 1 if the highest grade attended was not completed. To make the CPS measure comparable to our measure we set the minimum at 8 years of education and the maximum at 16. For the earnings variable in the CPS we take total yearly earnings, not including unearned income, top-coded at the Social Security maximum for 1983.

Table 3 presents the summary statistics for the lottery sample of 496 and the subsample of size 3,778 from the CPS drawn from the New England states. We discard CPS observations with ages outside the range of the lottery sample, 19 and 81 in 1983. The raw CPS sample (CPS-raw in Table 3) is more balanced in its gender composition, with relatively more women than the lottery sample. This agrees with Charles Clotfelter and Philip Cook's (1989) findings that men are more likely to play the lottery than women. The CPS sample is on average slightly younger, with a much more dispersed age distribution. This again agrees with Clotfelter and Cook, and with Lisa Farrell and Ian Walker (1999), who find that lottery participation by age follows an inverted U-shape, with lottery players more likely to come from the middle of the age distribution. Compared to the CPS the lottery sample has higher average education and earnings. This result is somewhat surprising as researchers have often found that lottery players are relatively low educated (e.g., Clotfelter and Cook, 1989), although gambling in general is often found to increase with education. To further investigate this result we match the CPS sample on age and gender to the lottery sample. The adjusted averages are reported in the CPS-I columns in Table 3. We find that after this adjustment, the lottery sample is still more highly educated than the CPS sample, although their earnings are now similar. If we also adjust for years of education (CPS-II), we find that the earnings in the lottery sample are considerably lower than those for the CPS sample, consistent with other research.

The most surprising difference between the lottery sample and the CPS is in the education distribution. In the CPS as well as in the lottery sample the education distribution has modes at 8, 12, and 16 years of education, but the relative weight of the modes is quite different. Taking the raw CPS sample, the proportion of people with less than or equal to 8 years of education is 9 percent, the proportion with exactly 12 years (high school) is 38 percent, and the percentage with 16 years of education (college) is 23 percent. For the lottery sample these proportions are 3 percent (8 years), 26 percent (12 years), and 37 percent (16 years). Although some of these differences may be due to nonresponse, with a 50-percent response rate it is difficult to attribute them entirely to nonresponse. An alternative explanation is measurement error, possibly deliberate exaggeration of education levels. To check this we ran a standard regression of log earnings on years of education, experience (calculated as age minus education minus 6), experience-squared, and a dummy for men, using only observations with positive earnings. The returns to education in the lottery sample in that regression are 7.7 percent (standard error [SE] 2.7 percent), very similar to those found in the CPS sample of 8.2 percent (SE 0.9 percent).[16] If education levels were generally misreported in the lottery sample, one might have expected an estimate for the returns to education closer to zero. Thus, measurement error is an unlikely reason for the differences in education levels between the two samples.

III. Conceptual Framework and Specification

Here we outline the basic framework for analyzing the data. Suppose individual i lives for T_i periods. Each individual chooses leisure l_{it} and consumption c_{ijt} in categories $j = 1, \ldots, J$, in all periods, $t = 1, 2, \ldots, T_i$, to maximize lifetime discounted utility. We assume the utility function for individual i is Stone-Geary[17]:

$$(1)\quad U(c_{i11}, \ldots, c_{i1T_i}, \ldots, c_{iJ1}, \ldots, c_{iJT_i}, l_{i1}, \ldots, l_{iT_i}) = \sum_{t=1}^{T_i} \frac{1}{(1+\delta)^t} \left[\beta_l \ln(l_{it} - \gamma_l) + \sum_{j=1}^{J} \beta_{cj} \ln(c_{ijt} - \gamma_{cj}) \right],$$

where δ, γ_{cj}, γ_l, β_l, and β_{cj} are the preference parameters. We normalize β_{cj} and β_l so that $\beta_l + \sum_{j=1}^{J} \beta_{cj} = 1$.

Utility is maximized over consumption and leisure, subject to an intertemporal budget constraint,

$$(2)\quad \sum_{t=1}^{T_i} \frac{1}{(1+r)^t} \times \left(w_{it}(H - l_{it}) + Y_{it}^N - \sum_{j=1}^{J} p_{tj} c_{ijt} \right) = 0,$$

where r is the interest rate, w_{it} is the wage rate for individual i in period t, H is the maximum number of hours worked, p_{tj} is the price of good j in period t, and Y_{it}^N is unearned income in period t. In this formulation Y_{i1}^N is individual i's wealth at the time of winning the lottery. Lifetime discounted discretionary income is the discounted sum of total potential labor earnings and unearned income minus essential expenditures on leisure and consumption:

$$(3)\quad F_i = \sum_{t=0}^{T} \frac{1}{(1+r)^t} \times \left(w_{it} \cdot H + Y_{it}^N - w_{it} \cdot \gamma_l - \sum_{j=1}^{J} p_{sj} \cdot \gamma_{cj} \right),$$

where the discount factor is the interest rate r.

We are interested in the effect of winning the lottery on the optimal path of labor earnings, consumption, and savings. Winning a total lottery prize L_i corresponds to a change in the path of unearned income from Y_{it}^N to $Y_{it}^{N'}$, where

$$Y_{it}^{N'} = \begin{cases} Y_{it}^N + L_i/20 & t = 1, \ldots, 20, \\ Y_{it}^N & t = 21, \ldots, T_i. \end{cases}$$

For simplicity we suppose that the life span extends beyond the 20 years of the lottery payments. The only way winning the lottery changes the optimal labor supply, consumption, and savings decision is through its effect on lifetime discounted discretionary income. Winning the lottery increases this by $\sum_{t=1}^{20} [1/(1+r)^t](L_i/20)$.

We assume that the individual faces no uncertainty concerning prices, wages, unearned income, and her life span. The optimal value of labor earnings in period t is

[16] The same is true for the coefficients on experience, experience-squared, and gender.

[17] See, for example, Orley Ashenfelter and John Ham (1979), MaCurdy (1981), Terry Johnson and Pencavel (1984), and Dean Hyslop (2001) for empirical applications of Stone-Geary utility functions in life-cycle settings.

$$(4) \qquad y_{it} = w_{it}(H - l_{it}) = (H - \gamma_l)w_{it} - \beta_l\left(\frac{1+r}{1+\delta}\right)^t \times \left(\sum_{s=1}^{T_i}(1+\delta)^{-s}\right)^{-1} F_i - \beta_l\lambda_{it}(L_i/20),$$

where

$$\lambda_{it} = \left(\frac{1+r}{1+\delta}\right)^t \frac{\sum_{s=1}^{20}(1+r)^{-s}}{\sum_{s=1}^{T_i}(1+\delta)^{-s}}.$$

Optimal expenditures on consumption of good j in this period are given by

$$(5) \qquad p_{tj}c_{ijt} = p_{tj}\gamma_{cj} + \beta_{cj}\left(\frac{1+r}{1+\delta}\right)^t \times \left(\sum_{s=1}^{T_i}(1+\delta)^{-s}\right)^{-1} F_i + \beta_c\lambda_{it}(L_i/20).$$

Wealth in period t, for $t \leq 20$, that is, during the time of winning, is

$$(6) \quad S_{it} = \sum_{s=0}^{t}(1+r)^{t-s}\Bigg[w_{is}H - w_{is}\gamma_s - \gamma_c p_s - \left(\frac{1+r}{1+\delta}\right)^t\left(\sum_{s=0}^{T_i}(1+\delta)^{-s}\right)^{-1} F_i + Y_{it}^N + (1-\lambda_{it})\frac{L_i}{20}\Bigg].$$

Equations (4) to (6) form the basis of our empirical analyses. Given a population of individuals acting according to this model we can write the expression for labor earnings as a linear function of the yearly lottery prize

$$y_{it} = \alpha_l - \beta_l\bar{\lambda}(L_i/20) + \varepsilon_{it},$$

where $\bar{\lambda}$ is the population average of λ_{it} and the residual ε_{it} captures variation in wages, other unearned income, and life span.

As discussed in Section II, subsection B, the critical assumption we make is that of exogeneity of the lottery prize. In the current setup this requires that within the population of respondents the magnitude of the lottery prize, L_i, is independent of all the other inputs into the decision process, that is, wages, other unearned income, and life span, and thus of the lifetime discounted discretionary income in the absence of the lottery prize, F_i. Evidence for this assumption is the lack of correlation between the magnitude of the lottery prize on the one hand and individual characteristics and prior earnings on the other hand, at least within the sample of 193 small winners. If this assumption holds, then ε_{it} is uncorrelated with L_i, and thus least-squares regression gives an unbiased estimator for $-\beta_l \times \bar{\lambda}$.

In addition to this simple linear regression we estimate regression functions with additional covariates:

$$y_{it} = \alpha_l - \beta_l\bar{\lambda}(L_i/20) + \theta_l{}' X_i + \varepsilon_{it}.$$

If these covariates are independent of the lottery prize, their inclusion cannot lead to bias, irrespective of their relation to the left-hand-side variable, although it may affect precision. The main reason for including them is to correct for nonresponse bias. The critical assumption then requires only that among respondents the lottery prize is independent of ε_i once we condition on these covariates.

In the empirical section, we report estimates of this coefficient without directly attempting to separate the coefficient of primary interest, $-\beta_l$, from $\bar{\lambda}$. This coefficient $\bar{\lambda}$ reflects the fact that the lottery payments are for 20 years only rather than an annuity for the remaining lifetime, as well as differences between the discount factor δ and the interest rate r. To adjust for this limited duration would involve estimating the discount factor and average life span, for which the current data set does not appear to be well suited. To provide some indication of the relation between the estimated coefficients and coefficient of primary interest, $-\beta_l$, we calculated $\bar{\lambda}$ for different values of the discount factor and life span. Suppose the discount factor and interest rate are both equal to 0.10 and the life span is an additional 30 years after winning the lottery.[18] In that case $\lambda = 0.90$.[19]

We also estimate regression functions where we interact the lottery prize with individual characteristics, specifically labor market status prior to winning, gender, age, education, and the number of years since winning the lottery. We might expect some of these factors to be associated with variation in the preference parameters. For example, the weight of leisure in the utility function may vary with gender or education level. Age is likely to be correlated with the remaining life span, with older individuals having lower values of the remaining life span T_i, and thus higher values of λ_{it}, implying that the interaction of age and lottery prize should be negative. According to the model, the number of years since winning the lottery should not be associated with variation in the effect of the lottery prize on labor earnings, although clearly alternative functional forms for the utility function allow for such variation.

We carry out the same analysis for the consumption expenditures. Because we only observe total expenditures over the entire period, on the right-hand side we use the accumulated lottery winnings, equal to $t(L_i/20)$. The coefficient should be the relative weight in the utility function for that consumption category β_{cj}, multiplied by $\bar{\lambda}$, approximately equal to 0.90.

For the savings the coefficient on the accumulated lottery prize $\sum_{s=1}^{t} L_i/20 = t(L_i/20)$ should be equal to $[(1 - \bar{\lambda})/t] \sum_{s=1}^{t} (1 + r)^{t-s}$, a function of the discount rate and interest rate, the remaining life span and the number of years since winning. Using the same numbers as before, $\delta = r = 0.10$, $T = 30$, and assuming all individuals are 10 years into their payout period, this equals 0.15. Increasing the remaining life span to 40 years increases this to 0.21, and lowering the interest and discount rate to 0.05 increases the predicted coefficient on savings to 0.23. Interacting the lottery prize with age in this regression should lead to a negative coefficient because age is negatively correlated with the remaining life span and thus also negatively correlated with $1 - \bar{\lambda}$. Interacting it with the number of years since winning the lottery should in this model lead to a positive coefficient of approximately 0.01 given the previoulsy hypothesized values for δ, r, and T.

IV. Results

A. *Marginal Propensity to Earn Out of Unearned Income*

In Table 4 we present the results for the marginal propensity to earn (MPE) out of unearned income, or the income elasticity, based on regressions of labor earnings on the yearly lottery prize.[20] 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 long-run 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 set of regressors (years of education, age, dummies for sex, college, age over 55, age over 65) in specification II, or when we first difference the earnings variable by subtracting the earnings in the last pre-lottery year (specification III). Even with the large set of covariates (small set of covariates plus number of tickets bought, year of winning, earnings in six years prior to winning, dummies for positive earnings in six years prior to winning, dummy for working at the time of winning) in specification IV, the estimates do not change much, although they are more precisely estimated.[21]

[18] This implies 10 years beyond the 20 years of lottery payment. Given that individuals in our sample are on average 50 years old when they win the lottery, this corresponds to a total age of 80.

[19] If the remaining life span is 40 years, $\lambda = 0.87$. If the discount rate and interest rate are 0.05 (and the life span is again 30 years), $\lambda = 0.81$. Obviously if the life span is equal to the 20 years of lottery payments, $\lambda = 1$.

[20] More precisely, as discussed in Section III, these are estimates of $-\beta_l\bar{\lambda}$. With $\bar{\lambda}$ argued to be approximately 0.90, this means that to get estimates of the MPE directly comparable to those in the literature, one should add 10 percent to the estimates in Tables 4 and 5.

[21] This robustness may come as somewhat of a surprise, given that Table 2 shows that average characteristics, including lagged earnings, differ considerably between winners and nonwinners. However, these significant differences do not imply that these variables actually explain much of the variation in prizes. The R^2 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^2 drops to 0.08. Because the number of tickets bought itself is not significantly correlated with earnings, it may be less surprising that adjusting for differences in these variables does not greatly affect the results. Because some of the control variables, notably lagged earnings, are highly correlated with the outcome in these regressions, the standard errors in specifications III and IV are considerably lower than those in the first specification without any control variables.

TABLE 4—ESTIMATES OF MARGINAL PROPENSITY TO EARN OUT OF UNEARNED INCOME: YEARLY LOTTERY PAYMENTS AS RIGHT-HAND-SIDE VARIABLE

	Specifications								
	I	II	III	IV	V		VI	VII	VIII
Outcomes[a]	496	496	496	496	496		237	453	194
Average post-lottery earnings	−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 earnings	−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 earnings	−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 earnings	−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 earnings	−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 earnings	−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 earnings	−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 earnings	−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)

Notes: Specifications: I: No individual controls, no differencing of outcome, linear in prize; sample includes nonwinners and big winners. II: Small set of individual controls (years of education, age, dummies for sex, college, age over 55, age over 65), no differencing of outcome, linear in prize; sample includes nonwinners and big winners. III: Small set of individual controls, differenced outcomes, linear in prize; sample includes nonwinners and big winners. IV: Expanded set of individual controls (small set of controls plus number of tickets bought, year of winning, earnings in six years prior to winning, dummies for positive earnings in six years prior to winning, dummy for working at the time of winning), differenced outcomes, linear in prize; sample includes nonwinners and big winners. V: Expanded set of controls, differenced outcomes, quadratic in prize; sample includes nonwinners and big winners. Estimates reported are derivative with respect to prize at prize equal to zero and prize equal to $32,000. VI: Expanded set of individual controls, difference outcomes, linear in prize; sample includes winners only. VII: Expanded set of individual controls, difference outcomes, linear in prize; sample includes nonwinners and winners < $100,000 only. VIII: Expanded set of individual controls, difference outcomes, linear in prize; sample includes winners < $100,000 only.

[a] Outcomes: Average of Social Security earnings in years one through six after winning the lottery, and earnings in years zero to six after winning the lottery.

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 MPE based on this specification are much larger than the linear regression-based estimates, 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 two estimates are 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 minimum of zero, 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 nonwinners, 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 of randomly allocating annuities to a fixed population. The results for this specification are very similar to those from specification IV with the same set of control variables that includes the nonwinners.[22] Next, in specification VII, we exclude the big winners (winners with a yearly prize larger than $100,000). This yields results similar to those from the quadratic specification, with an estimate for the MPE of −0.122 (0.020). Finally, we exclude both nonwinner 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 specifications linear in the prize have trouble

[22] Although more than half the original sample is dropped in this specification, the precision is not significantly affected because most of the variation in the lottery prize is among the winners.

TABLE 5—ESTIMATES OF MARGINAL PROPENSITY TO EARN (MPE) OUT OF UNEARNED INCOME: INTERACTIONS WITH PRIOR LABOR MARKET HISTORY, SEX, AGE, EDUCATION, AND TIME SINCE WINNING

Outcomes	Baseline MPE[a]	Prior earnings zero[b]	Female[b]	$55 < \text{Age} \leqq 65$[b]	$\text{Age} > 65$[b]	College[b]	Years since winning[b]
Average post-lottery earnings	−0.124	0.209	0.002	−0.167	−0.001	0.037	−0.010
	(0.054)	(0.084)	(0.057)	(0.070)	(0.090)	(0.061)	(0.022)
Year 0 earnings	0.032	0.014	−0.015	−0.094	−0.004	−0.027	0.006
	(0.029)	(0.045)	(0.031)	(0.038)	(0.049)	(0.033)	(0.012)
Year 1 earnings	−0.096	0.108	0.057	−0.204	−0.045	0.043	0.001
	(0.047)	(0.073)	(0.050)	(0.061)	(0.079)	(0.053)	(0.019)
Year 2 earnings	−0.119	0.175	0.020	−0.215	−0.039	0.086	−0.025
	(0.056)	(0.088)	(0.060)	(0.073)	(0.095)	(0.064)	(0.024)
Year 3 earnings	−0.120	0.225	−0.058	−0.178	0.003	0.040	−0.004
	(0.061)	(0.097)	(0.066)	(0.081)	(0.104)	(0.070)	(0.026)
Year 4 earnings	−0.133	0.158	0.005	−0.100	0.099	0.009	−0.024
	(0.065)	(0.103)	(0.070)	(0.085)	(0.110)	(0.074)	(0.027)
Year 5 earnings	−0.138	0.235	−0.000	−0.127	0.032	−0.001	−0.002
	(0.069)	(0.108)	(0.074)	(0.090)	(0.116)	(0.078)	(0.029)
Year 6 earnings	−0.137	0.355	−0.009	−0.177	−0.057	0.045	−0.009
	(0.070)	(0.110)	(0.075)	(0.091)	(0.118)	(0.079)	(0.029)

Notes: The sample consists of the 194 winners with a yearly prize less than or equal to $100,000. All regressions include the yearly lottery prize, the lottery prize interacted with an indicator for zero earnings prior to winning, an indicator for women, an indicator for age between 55 and 65 at the time of winning, an indicator for age over 65 at the time of winning, an indicator for some college, and years since winning, as well as the large set of control variables (years of education, age, dummies for sex, college, age over 55, age over 65, small set of controls plus number of tickets bought, year of winning, earnings in six years prior to winning, dummies for positive earnings in six years prior to winning, dummy for working at the time of winning).

[a] Reports the marginal propensity to earn out of unearned income for the baseline individual, a man who won in 1986, who had positive earnings in the year prior to winning, with no college, less than 55 years old at the time of winning.

[b] The estimates are those for the coefficients corresponding to the interaction with yearly lottery prize.

fitting the response function for the big winners; this is not accounted for by including additional regressors as in specifications II, III, and IV. A quadratic specification (specification V) fits considerably better in the full sample. Once we exclude the big winners (specifications VII and VIII), including a quadratic term in the prize leads to a t-statistic of 0.6, suggesting that the linear specification fits fairly well. These three specifications (V, VII, and VIII) all lead to estimates of the MPE around −11 percent.

The next seven rows present the results for the year of winning and the six subsequent years. We focus on specification VIII because the differences between the specifications largely follow the same pattern as that for the average earnings. Because 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 subsequent years. This hypothesis 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 that in subsequent years, at −0.056 (0.25). After the first post-lottery year, the MPE stabilizes around −0.10 (0.03). It appears to take individuals some time to adjust their labor supply to the desired level.

The estimates for the marginal propensity to earn of around −0.10 (or −0.11, after adjusting for an estimate of $\bar{\lambda}$ approximately equal to 0.90 to take care of the limited duration of the lottery prize) are not out of line with those in the literature, which is not surprising given the wide range of estimates reported there. Pencavel (1986) reports in his survey on 13 studies using nonexperimental data for U.S. men, with estimates 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 (2000) find estimates ranging from −0.95 to 0.002 for men, and from −0.40 to 0.27 for women.

In Table 5 we present the results for the MPE with interactions with five variables: indicators for zero earnings in the year prior to winning, for women, for age at winning between 55 and 65, for age at winning greater than 65, for some college, and the discrete variable years since winning. This regression is based on specification VIII, with the large set of control variables and the sample limited to 194 winners with a prize less than \$100,000 per year. The first column presents the estimate for the baseline individual: a man less than 55 years old with positive earnings in the year prior to winning, with no college and who won in 1986 (10 years prior to the 1996 survey). For this baseline person, the estimated MPE is -0.124. For those with zero earnings in the year prior to winning the lottery, the MPE based on the average post-lottery earnings is 0.209 (0.084) higher, leading to an estimated MPE of $-0.124 + 0.209 = 0.085$, which is in fact positive. Although this estimate is not significantly different from zero, at least it suggests no evidence of negative effects of the lottery payments on the labor supply of low earners.

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 of the differences are extremely close to zero, with some negative and some positive estimates, and none is significant for the average of post-lottery years or for any of the post-lottery years separately. In this sample men and women have substantially different labor market experiences, with 75 percent of the men working in the last pre-lottery year, with average earnings for those with positive earnings equal to \$26,700, and 65 percent of the women working, with their average earnings equal to only \$15,400. However, these differences do not appear to lead to different estimates of the marginal propensity to earn.

We do find differences by age. We experimented with age interacted with the prize, as well as with indicators for age between 55 and 65 and age over 65 interacted with the prize. 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 that 65 at the time of winning, reduce their earnings as much as those younger than 55, though less than those between 55 and 65 years of age. The effect of the lottery winnings by age suggests that some individuals reduce their labor supply earlier than they might otherwise have, that is, take early retirement. 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. Although this negative effect was not predicted by the simple life-cycle model used in the previous section, it would be consistent with generalizations with human-capital formation that make wages a function of previous labor market experiences.

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.

B. *Consumption and Savings*

In Table 6 we report the results for expenditures on cars, housing, and savings. Recall that here the explanatory variable is the cumulative lottery prize up to the time of the survey. For car values, total or net, we find that there is a small but highly significant effect of lottery prize. As long as the sample includes the big winners (specifications I–VII) 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 lottery 0.9 percent (SE 0.2 percent) is spent on cars (or 0.7 percent on car values net of loans).[23] If we exclude the big winners, the marginal propensity to consume goes up to about 1.4 percent (0.7 percent). 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 up to 4 percent (1 percent) in specification VI. If we exclude the big winners the effect is still of similar size, but it is no longer significantly different from zero. 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, especially the big winners, but they do finance them through correspondingly larger mortgages.

Next, consider the savings outcomes. First we look at the retirement accounts. The basic no-control-variable specification suggests a large and significant negative effect of unearned income on retirement savings. Adjusting for covariates leads to slightly smaller estimates,

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

TABLE 6—ESTIMATES OF MARGINAL PROPENSITY TO CONSUME OUT OF UNEARNED INCOME: ACCUMULATED LOTTERY PAYMENTS AS RIGHT-HAND-SIDE VARIABLE

	Specifications							
	I	II	IV	V		VI	VII	VIII
Outcomes[a]	496	496	496	496		237	453	194
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 assets	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 assets	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)

Notes: Specifications: I: No individual controls, linear in prize; sample includes nonwinners and big winners. II: Small set of individual controls (years of education, age, dummies for sex, college, age over 55, age over 65), linear in prize; sample includes nonwinners and big winners. IV: Expanded set of individual controls (small set of controls plus number of tickets bought, year of winning, earnings in six years prior to winning, dummies for positive earnings in six years prior to winning, dummy for working at the time of winning), linear in prize; sample includes nonwinners and big winners. V: Expanded set of controls, quadratic in prize; sample includes nonwinners and big winners. Estimates reported are derivative with respect to prize at prize equal to zero and prize equal to $32,000. VI: Expanded set of individual controls, linear in prize; sample includes winners only. VII: Expanded set of individual controls, linear in prize; sample includes nonwinners and winners < $100,000 only. VIII: Expanded set of individual controls, linear in prize; sample includes winners < $100,000 only.

[a] Outcomes: Current value of cars, net value of cars, current value of housing, net value of housing, value of retirement accounts, value of other savings accounts, total value of financial assets.

but they are still negative and significantly different from zero. However, this effect disappears entirely if we drop the nonwinners from the sample (specifications VI and VII). It appears that the nonwinners save considerably more in retirement accounts than the winners. Evidence of this savings pattern can already be seen in the summary statistics in Table 2, which shows that nonwinners 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, but not all, because the adjustments in specifications II and IV do not eliminate this negative estimate. Part of this savings differential may also be attributed to differences in the population of season ticket holders and single ticket buyers that we do not adequately adjust for. Looking at other financial savings, we do find a positive relation with unearned income. Excluding nonwinners, we find a marginal propensity to save out of unearned income of 3.9 percent (1.9 percent). Excluding both nonwinners and big winners, the estimate is much higher: 18.3 percent (6.1 percent). 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 nonwinners and 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 nonwinners and big winners equal to 15.8 percent (5.6 percent). These numbers are consistent with the discussion of the life-cycle model in Section IV, where a discount factor and interest rate of 0.10 and a life expectancy of 80 years suggest a coefficient of 0.15 for the effect of the lottery prize on accumulated savings 10 years into winning the lottery.

In Table 7 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). Men appear to be saving more from unearned income, with their marginal propensity to save (other than in retirement accounts) greater by 24 percent (11 percent).

Surprisingly there is no evidence that older people save less out of unearned income. For the

TABLE 7—ESTIMATES OF MARGINAL PROPENSITY TO CONSUME OUT OF UNEARNED INCOME: INTERACTIONS WITH PRIOR LABOR MARKET HISTORY, SEX, AGE, EDUCATION, AND TIME SINCE WINNING

Outcomes	Baseline MPC[a]	Prior earnings zero[b]	Female[b]	$55 < \text{Age} \leqq 65$[b]	Age > 65[b]	College[b]	Years since winning[b]
Value cars	0.025	−0.012	−0.020	0.010	0.041	−0.049	−0.010
	(0.019)	(0.032)	(0.020)	(0.023)	(0.033)	(0.021)	(0.008)
Net value cars	−0.012	−0.041	−0.028	0.017	0.106	−0.002	0.010
	(0.026)	(0.042)	(0.026)	(0.031)	(0.044)	(0.028)	(0.010)
Value house	0.089	−0.087	−0.069	0.086	0.115	−0.007	−0.021
	(0.057)	(0.094)	(0.058)	(0.069)	(0.099)	(0.063)	(0.023)
Net value house	0.039	−0.095	−0.027	0.098	0.068	0.008	−0.002
	(0.053)	(0.087)	(0.054)	(0.064)	(0.092)	(0.058)	(0.022)
Retirement accounts	0.037	−0.001	−0.012	0.064	−0.113	−0.022	−0.011
	(0.040)	(0.066)	(0.041)	(0.048)	(0.070)	(0.044)	(0.016)
Other financial assets	0.280	−0.068	−0.321	0.206	−0.164	0.026	0.127
	(0.103)	(0.169)	(0.106)	(0.124)	(0.179)	(0.113)	(0.042)
Total financial assets	0.259	−0.104	−0.236	0.165	−0.155	0.023	0.104
	(0.107)	(0.176)	(0.110)	(0.129)	(0.186)	(0.118)	(0.044)

Notes: The sample consists of the 194 winners with a yearly prize less than or equal to $100,000. All regressions include the accumulated lottery prize, the lottery prize interacted with an indicator for zero earnings prior to winning, an indicator for women, an indicator for age between 55 and 65 at the time of winning, an indicator for age over 65 at the time of winning, an indicator for some college, and years since winning, as well as the large set of control variables (years of education, age, dummies for sex, college, age over 55, age over 65, small set of controls plus number of tickets bought, year of winning, earnings in six years prior to winning, dummies for positive earnings in six years prior to winning, dummy for working at the time of winning).

[a] Reports the marginal propensity to consume out of unearned income for the baseline individual, a man who won in 1986, who had positive earnings in the year prior to winning, with no college, less than 55 years old at the time of winning.

[b] The estimates are those for the coefficients corresponding to the interactions with accumulated lottery prize.

55- to 65-year-old group, the estimates are in fact positive and close to significant, although for the older group they are negative but far from significant. Having some college does not affect either the marginal propensity to spend on cars or housing, or the marginal propensity to save. The longer ago someone has won the lottery, and thus the closer to the end of the 20 years of lottery payments, the larger the marginal propensity to save. Although the direction is consistent with the predictions of the model in Section III, the magnitude of the interactions appears to be too large for this interpretation. This pattern is more consistent with consumption smoothing where large expenditures are incurred early on during the period of lottery payments, followed by a period with higher savings.

V. Conclusion

In this paper we exploit the random assignment of large sums of money in lotteries to estimate the effect of unearned income on labor earnings, consumption, and savings. We find that over the range of annual unearned income from zero to $100,000 the marginal propensity to earn out of unearned income is around −11 percent. This estimate is robust against a variety of specifications once we either allow for a nonlinear effect or exclude winners who receive more than $100,000 per year from the lottery. Surprisingly this effect does not differ much between men and women. The effect is stronger for individuals close to the standard retirement age. The savings rate for unearned income is estimated at 16 percent and increases with the proportion of the prize received.

A comparison between the lottery sample and the Current Population Survey suggests that, although the marginal distribution of economic variables varies substantially between the samples, economically meaningful parameters such as the returns to education are very similar. The estimates obtained in this study may therefore have some relevance for more general populations, although there is a caveat that we have no direct evidence concerning the difference of responses to lottery income versus other sources of unearned income.

REFERENCES

Angrist, Joshua and Krueger, Alan. "Empirical Strategies in Labor Economics," in Orley

Ashenfelter and David Card, eds., *Handbook of labor economics*. Amsterdam: North-Holland, 2000, pp. 1277–1366.

Ashenfelter, Orley and Ham, John. "Education, Unemployment and Earnings." *Journal of Political Economy*, October 1979, *87*(5), pp. S99–116.

Ashenfelter, Orley and Heckman, James. "The Estimation of Income and Substitution Effects in a Model of Family Labor Supply." *Econometrica*, January 1974, *42*(1), pp. 73–85.

Blundell, Richard and MaCurdy, Thomas. "Labor Supply," in Orley Ashenfelter and David Card, eds., *Handbook of labor economics*. Amsterdam: North-Holland, 2000, pp. 1559–1695.

Bodkin, Ronald. "Windfall Income and Consumption." *American Economic Review*, September 1959, *49*(4), pp. 602–14.

Clotfelter, Charles and Cook, Philip. *Selling hope: State lotteries in America*. Cambridge, MA: Harvard University Press, 1989.

Clotfelter, Charles and Cook, Philip. "The Demand for Lottery Products." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 2928, April 1989.

Farrel, Lisa and Walker, Ian. "The Welfare Effects of Lotto: Evidence from the UK." *Journal of Public Economics*, April 1999, *72*(1), pp. 99–120.

Hausman, Jerry and Wise, David. "Technical Problems in Social Experimentation: Cost versus Ease of Analysis," in Jerry Hausman and David Wise, eds., *Social experimentation*. Chicago: Chicago University Press, 1985, pp. 187–219.

Holtz-Eakin, Douglas; Joulfaian, David and Rosen, Harvey. "The Carnegie Conjecture: Some Empirical Evidence." *Quarterly Journal of Economics*, May 1993, *108*(2), pp. 413–35.

Hyslop, Dean. "The Covariance Structure of Intrafamily Earnings, Rising Inequality and Family Labor Supply." *American Economic Review*, September 2001, *91*(4), pp. 755–77.

Imbens, Guido; Rubin, Donald; and Sacerdote, Bruce. "Estimating the Effect of Unearned Income on Labor Earnings, Savings and Consumption: Evidence from a Survey of Lottery Players." National Bureau of Economic Research (Cambridge, MA), Working Paper No. 7001, March 1999.

Johnson, Terry and Pencavel, John. "Dynamic Hours of Work Functions for Husbands, Wives, and Single Females." *Econometrica*, March 1984, *52*(2), pp. 363–90.

Kaplan, H. Roy. "Lottery Winners and Work Commitment." *Journal of the Institute for Socioeconomic Research*, Summer 1985, *10*(2), pp. 82–94.

Killingsworth, Mark and Heckman, James. "Female Labor Supply: A Survey," in Orley Ashenfelter and Richard Layard, eds., *Handbook of labor economics*. Amsterdam: North-Holland, 1986, pp. 103–204.

Kreinin, Mordechai. "Windfall Income and Consumption—Additional Evidence." *American Economic Review*, June 1961, *51*(3), pp. 388–90.

Landsberger, Michael. "Windfall Income and Consumption." *American Economic Review*, June 1966, *56*(3), pp. 534–40.

Little, Roderick and Rubin, Donald. *Statistical analysis with missing data*. New York: Wiley, 1987.

MaCurdy, Thomas. "An Empirical Model of Labor Supply in a Life-Cycle Setting." *Journal of Political Economy*, December 1981, *89*(6), pp. 1059–85.

Morrill, William. "Introduction: Special Issue on Negative Income Tax." *Journal of Human Resources*, Spring 1974, *9*(2), pp. 156–57.

Pencavel, John. "Labor Supply of Men: A Survey," in Orley Ashenfelter and Richard Layard, eds., *Handbook of labor economics*. Amsterdam: North-Holland, 1986, pp. 3–102.

Philipson, Thomas. "Observational Agency and Supply-Side Econometrics." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 210, February 1997.

Rees, Albert. "An Overview of the Labor Supply Results." *Journal of Human Resources*, Spring 1974, *9*(2), pp. 158–80.

Rubin, Donald. *Multiple imputation*. New York: Wiley, 1987.

Sacerdote, Bruce. "The Lottery Winner Survey, Crime and Social Interactions, and Why Is There More Crime in Cities." Ph.D. dissertation, Harvard University, 1996.

Thaler, Richard. "Savings, Fungibility, and Mental Accounts." *Journal of Economic Perspectives*, Winter 1990, *4*(1), pp. 193–205.