Economics 250a
Lecture 3
Outline

1. Compensating Wage Differentials for Fixed Hours Packages (from last lecture)
2. Three simple papers illustrating the estimation of static labor supply models:
   Ashenfelter, Doran and Schaller (2010) uncompensated elasticity for taxi drivers
   Imbens, Rubin and Sacerdote (2001) the income effect for lottery winners
   Cesarini et al (2013) "The Effect of Wealth on Household Labor Supply"
3. Estimation with kinked budget sets
   Emmanuel Saez (2010). "Do Taxpayers Bunch at Kink Points?" AEJ-Policy August 2010.

1. Compensating Wage Differentials for Fixed Hours Packages (continued)

Recall, we define

$$R(\overline{h}, u) = \min_x x \quad \text{s.t.} \quad u(x, T - \overline{h}) \geq u.$$

This is the minimum amount of consumption that in combination with $\overline{h}$ achieves utility u. R is just the vertical distance from the x-axis to the u indifference curve when $\ell = T - \overline{h}$. If a job pays the wage w and requires $\overline{h}$ hours of work then an individual would have to receive

$$R(\overline{h}, u) - w\overline{h} = \min_x x - w\overline{h} \quad \text{s.t.} \quad u(x, T - \overline{h}) \geq u$$

in additional nonlabor income to achieve utility u. Note that if $\overline{h} = h^c(w, u^0)$ then the required non-labor income is $e(w, u^0)$:

$$R(h^c(w, u^0), u^0) - wh^c(w, u^0) = e(w, u^0) \qquad (*)$$

This holds as we vary w so differentiating:

$$R_1 \frac{\partial h^c}{\partial w} - h^c - w\frac{\partial h^c}{\partial w} = \frac{\partial e}{\partial w}$$

But since $\partial e/\partial w = -h^c$, we have that

$$R_1(h^c(w, u^0), u^0) = w.$$

If you think of R as the height of the indifference curve, and recall that w is the slope of the indifference curve at $h = h^c(w, u^0)$ this is obvious. Now this relation also holds as we vary w so differentiating again

$$\begin{aligned} R_{11}\frac{\partial h^c}{\partial w} &= 1 \\ &\Rightarrow \quad R_{11}(h^c(w, u^0), u^0) = [\frac{\partial h^c(w, u^0)}{\partial w}]^{-1} \end{aligned}$$

This shows that the inverse of the slope of the compensated labor supply curve is the rate of change of the slope of the indifference curve. When $\frac{\partial h^c(w,u^0)}{\partial w}$ is "small" the indifference curve changes slope very fast (i.e., indifference curves are closer to Leontief).

Now suppose there is an unconstrained job that pays a wage w^0, and another constrained job that requires $h = \overline{h}$. We ask: what wage w would the constrained job have to pay so

an agent is indifferent between the two jobs. The difference $(w - w^0)$ is the compensating differential for the constrained choice. Using the R function we must have

$$R(\overline{h}, u^0) - w\overline{h} = e(w^0, u^0) \qquad (**)$$

Now we use a second order expansion around $R(h^c(w^0, u^0), u^0)$, where u^0 is the utility level of the reference job. Let h^0 be the (unconstrained) hours choice on that job. We have

$$\begin{aligned} R(\overline{h}, u^0) &\approx R(h^c(w^0, u^0), u^0) + (\overline{h} - h^0)R_1(h^c(w^0, u^0), u^0) + .5(\overline{h} - h^0)^2 R_{11}(h^c(w^0, u^0), u^0) \\ &= e(w^0, u^0) + w^0 h^0 + (\overline{h} - h^0)w^0 + .5(\overline{h} - h^0)^2 [\frac{\partial h^c(w^0, u^0)}{\partial w}]^{-1} \quad \text{(using (*) above)} \\ &= e(w^0, u^0) + \overline{h}w^0 + .5(\overline{h} - h^0)^2 [\frac{w^0 \partial h^c(w^0, u^0)}{h^0 \partial w}]^{-1} \frac{w^0}{h^0}. \end{aligned}$$

Now subtract $w\overline{h}$ from both sides:

$$R(\overline{h}, u^0) - w\overline{h} = e(w^0, u^0) - \overline{h}(w - w^0) + .5\frac{w^0}{h^0}(\overline{h} - h^0)^2 \frac{1}{\epsilon^c}$$

And using (**) we get

$$\frac{(w - w^0)}{w^0} = .5\frac{(\overline{h} - h^0)^2}{h^0 \overline{h}} \frac{1}{\epsilon^c}$$

For example, if

$$\frac{(\overline{h} - h^0)}{h^0} = .1$$

then the compensating differential is

$$\frac{(w - w^0)}{w^0} \approx \frac{.5 \times .1 \times .1}{\epsilon^c} = \frac{.005}{\epsilon^c}$$

For example, if $\epsilon^c = 0.1$, this formula implies you need a 5% higher wage to take a job with 10% lower hours (implying that your total earnings are 5% lower). This strikes me at least as a relatively small compensating differential. If $\epsilon^c = 0.2$, the compensating differential is only 2.5%. Note that the formula applies for low hours jobs and high-hours jobs – in fact the formula is symmetric.

2. Simple Static Labor Supply Estimation and Findings

a) Ashenfelter, Doran and Schaller (2010) – uncompensated elasticity of labor supply for taxi drivers

ADS study the effects of two major fare increases institututed for NYC cabs in March 1996 and May 2004. Their data consist of information collected each time a cab is inspected – roughly every 4 months (the mean time between inspections is 122 days with std dev = 4 days). Their measure of labor supply is m = miles driven (in the 4 months prior to the inspection). Their measure of the "wage", which they call θ, is revenue per mile (averaged over the 4 months prior to the inspection), which they estimate from R = revenues (over the 4 months): $\theta = R/m$. Note that with given levels of congestion, weather, etc, the rate of earnings per hour is just a multiple of θ. The will assume that labor supply depends on $\log \theta$, so the factor of proportionality drops out.

As in the case where we divide earnings by hours, there is a mechanical negative correlation between θ and m. The idea is to isolate the two major fare increase episodes, and examine the changes in miles and revenues/mile that occur at these events. Thus, their data are restricted to inspections in 4 periods:

March 1 1995	$\rightarrow$	February 9 1996 (pre-data for 1st increase)
July 1 1996	$\rightarrow$	July 1 1997 (post-data for 1st increase)
May 12 2003	$\rightarrow$	May 3 2004 (pre-data for 2nd increase)
September 7 2004	$\rightarrow$	September 7 2005 (post-data for 2nd increase)

For some of their analysis, they use a "balanced" sample that includes cabs that have complete data from the pre- and post-period for each increase event. Figure 1 of their paper plots mean miles and mean revenues per mile for inspections occurring in these intervals. You can see very clearly that (a) average revenues per mile went up sharply (b) average miles driven is either flat or falls off slightly, and definitely did not increase!

ADS use a log-linear labor supply model:

$$\log m_{it} = x_{it}a + b \log \theta_{it} + e_{it}$$

where m_{it} = miles driven by cab i in the 4-month period before the inspection at time t, $x_{it}a$ includes fixed effects for the month of the inspection, a control for the length of the interval covered by the retrospective period, and in some models fixed effects for each "medallion". Their sample is constructed to try to ensure that a medallion corresponds to a single owner-driver. Thus the fixed effects models control for preference variation, and also for any permanent differences in non-labor income (e.g., differences in spousal earnings). Transitory changes in non-labor income are not controlled – it is presumed that these average to 0. Likewise, factors that affect the relationship between hours and miles (traffic, weather, presence of conventioneers, etc) are assumed to average to

Their preferred estimation strategy is to fit the model by IV, using as an instrument a dummy =1 if period t is a post-increase period. This is not explained as clearly as it could be, though it should be clear from the graphs what is going on. The "first stage" models (shown in their table 3) show that revenue per mile increased by 19 percent after the fare increases. The "reduced form" models (shown in table 4) show that log miles driven fell by 2.39% (without fixed effects) or -4.23% (with fixed effects). The implied IV estimates are $-0.13 = -0.0239/0.19$ without fixed effects and $-0.22 = -0.0423/0.19$ with fixed effects. Notice that b is directly interpretable as an estimate of the uncompensated labor supply elasticity. ADS's estimates are between -0.13 and -0.22, which is pretty consistent with other long run evidence, though perhaps on the negative side.

For discussion:

a) why cab drivers?

b) can we generalize to other types of workers?

b) Imbens, Rubin and Sacerdote (2001) the income effect for lottery winners

IRS survey winners and what they call "non-winners" (who are in fact "very small prize winners") who purchased tickets to the Massachusetts "Megabucks" lottery in the 1984-88 period. The winners got big prizes – the median is $635,000 – which were paid out over 20 years. They asked people to allow them to use their SSA earnings records. They managed to get response rates of around 45%, yielding a sample of about 500. The sample respondents

are slightly older than average adults (mean age ~50), 63% men, with 13.7 years of education (about average for the cohort). The SSA records include earnings data for 6 years per-win and 6 years post-win: Figure 1 shows the 'event' study graph which suggests a modest decline in earnings after the win.

IRS use a Stone-Geary model, which gives rise to a "linear expenditure" model for earings in the post-win period

$$y_{it} = \alpha + \beta\overline{\lambda}\frac{L_i}{20} + e_{it}$$

where y_{it} = earnings of i in year t, L_i = lottery amount won by person i (=0 for the very small prize winners, and a number like 650,000 for the winners), $\overline{\lambda}$ is an average annuitizing factor, which adjusts for the fact that the lottery only lasts for 20 years and that people's rate of time preference may be different than the interest rate, and α and β are parameters from the SG utility function. This specification is predicated on the idea that L_i is randomly assigned – the very small winners got 0, the winners got a big prize, so preference differences can be rolled into the error and should not be correlated with the winning amount. Note that (apart from $\overline{\lambda}$, which should be on the order of .9 or so), we get an estimate of the *mpe* – the marginal propensity to reduce earnings per dollar of non-labor income.

In fact they actually estimate a model of the form:

$$y_{it} = \alpha + b_1\frac{L_i}{20} + b_2\left(\frac{L_i}{20}\right)^2 + e_{it}$$

since they find that the dependent variable is very skewed and the response seems to be affected by a few very large prizes. (It would have been nice to see a graph). Their "preferred" model gives an estimate of the "average" *mpe* ≈ -0.12 in an average year after the win, which accords very well with literature.

For discussion:

a) can we generalize to other types of people?

b) can we think of other ways to identify the *mpe* credibly?

b) Cesarini, Lindqvist, Notwidigdo, and Ostling, 2013.

CLNO study a very large sample of lottery winners using data from Sweden. Unlike IRS they have access to data on earnings for the entire country, and they know who "played" a lottery (including how many tickets they purchased) and who won, so they can implement very clean models in which the control group includes everyone who bought a ticket in the same lottery. They can also look at the spouses of lottery winners to see if it matters who wins.

CLNO study three types of lotteries: (1) "prize-linked savings" PLS lotteries, which gave awards to holds of certain savings accounts; (2) Kombi lottery, a montly subscription lottery; and (3) two types of scatch-ticket lotteries, known as TV-Triss and Clover. They estimate models of the form:

$$y_{it} = \beta_t L_{i0} + Z_{it}\gamma_t + X_i\delta_t + \epsilon_{it}$$

where y_{it} is individual i's income in time t, where $t = 0$ is the year of winning, L_{i0} is lottery winnings (measured in present value terms), Z_{it} are pre-determined controls (like earnings in earlier years), and X_i is a set of lottery fixed effects that ensure random assignment. Notice that they do not attempt to "annuitize" lottery winnings – so you should expect the estimate of β_t to be (approximately) 10-20 times smaller than the estimated effect in IRS.

Some of their main results are shown in Figure 1, Table 6, and Table 11 (at the end of the lecture). As shown in Figure 1, they get a wealth effect of about -0.01 per year (i.e., each 100 kronar of winnings causes a reduction of about 1 kronar in earnings) that is effective immediately and persists for 10 years. There does not seem to be a "cumulative" effect, or evidence of slow adjustment. If you read the paper carefully you will notice that they also find about the same effect for older and younger workers, and for male vs. female winners. The magnitude of the effect is in the range of IRS's finding of an *mpe* of -0.15 or so.[1]

Table 11 is very interesting because it shows that there is some additional negative effect on spouses, around 10-20% as big as the effect on the winner him/her self. Thus: (1) the effect on family earnings is a little larger; and (2) it looks like who wins the money largely determines who gets to "slack off" in the family! The latter finding is an important addition to the large but relatively low quality empirical literature on family labor supply, where many studies lack credible identification, and others try to use randomized experiments like Progressa but are often severely under-powered (and confounded by multiple channels).

3. Estimation with kinked budget sets - a brief introduction

Let's consider an agent who faces an non-linear tax: the tax rate is 0 for earnings less than E_1, then rises to $t > 0$. If the agent has a wage rate w and nonlabor income y (which is included in the tax base) then the agent pays no tax until

$$\begin{aligned} y + wh &= E_1 \\ \Rightarrow \quad h = h^* &= \frac{E_1 - y}{w}. \end{aligned}$$

For additional hours she pays a marginal tax of t. This is usually illustrated as in Figure 3.1

Note that the "linearization" of the flatter budget segment hits the $h = 0$ line at the level of income

$$y' = E_1 - w(1-t)h^* = tE_1 + (1-t)y > y \quad \text{if } E_1 > y.$$

Lets suppose agents have a labor supply function $h(w, y; \theta)$, where θ represents an unobserved heterogeneity component, such that $h(w, y, \theta') > h(w, y, \theta)$ whenever $\theta' > \theta$. Then looking at the graph we can see there 3 possible regimes:

$$\begin{aligned} I &: \quad h = h(w, y, \theta) \quad \text{if} \quad h(w, y, \theta) < h^* \\ II &: \quad h = h(w(1-t), y', \theta) \quad \text{if} \quad h(w(1-t), y', \theta) > h^* \\ III &: \quad h = h^* \quad \text{if} \quad h(w(1-t), y', \theta) \leq h^* \leq h(w, y, \theta). \end{aligned}$$

The fraction of the population who fall into range III – and who therefore have earnings exactly equal to the kink-point level E_1 – depends on the curvature of indifference curves. If people have Leontief preferences there is no one in range III. If preferences are very flat, however, there will be a lot of people who "bunch" at the kink point.

[1] Some of the analyses in CLNO is designed to address some analyses in a paper by Kimball and Shapiro which studies survey responses to hypothetical questions about how people would respond to winning a lottery. KS try to argue that the wealth effect in labor supply "could be" relatively large, and criticize various aspects of Imbens et al.

To make progress it is nicer to work with earnings ($g \equiv wh$) and the earnings function

$$g(w, y, \theta) = wh(w, y, \theta)$$

The reason is that the kink point is expressed in terms of earnings, not hours. Notice that the derivatives of the earnings function are closely related to the derivatives of the labor supply function:

$$\frac{\partial g}{\partial y} = w\frac{\partial h}{\partial y} \in [-1, 0]$$

and

$$\frac{w}{g}\frac{\partial g}{\partial w} = 1 + \frac{w}{h}\frac{\partial h}{\partial w} = 1 + \epsilon = 1 + \epsilon^c + w\frac{\partial h}{\partial y} \geq 0 \tag{1}$$

since $\epsilon^c \geq 0$ and $w\frac{\partial h}{\partial y} \geq -1$. Now return to Figures 3.1 and 3.2 and assume $y = 0$, so

$$y' = tE_1$$

Using the earnings function we can classify the 3 regimes as:

$$\begin{aligned}
I &: \; g = g(w, 0, \theta) \leq E_1 \\
II &: \; g = g(w(1-t), tE_1, \theta) \geq E_1 - y' = E_1(1-t) \\
III &: \; g = E_1 \;\; and \;\; g(w, 0, \theta) > E_1 \quad and \quad g(w(1-t), tE_1, \theta) < E_1(1-t)
\end{aligned}$$

Now lets go a little further and re-parameterize the earnings function as:

$$g(w, y, \theta) = k(w, y) + \theta$$

where θ is some random taste variable. The we can restate the regimes in terms of two critical cutoffs in the distribution of θ (for a given wage w) :

$$\begin{aligned}
I &: \; k(w, 0) + \theta \leq E_1 \Rightarrow \theta \leq E_1 - k(w, 0) = \theta^* \\
II &: \; k(w(1-t), tE_1) + \theta \geq E_1 - y' \Rightarrow \theta \geq E_1 - k(w(1-t), tE_1) - tE_1 = \theta^{**} \\
III &: \; \theta^* < \theta < \theta^{**}
\end{aligned}$$

Now notice that

$$\theta^{**} = \theta^* + k(w, 0) - k(w(1-t), tE_1) - tE_1$$

and taking a first order expansion

$$k(w(1-t), tE_1) = k(w, 0) - \frac{\partial g}{\partial w}tw + \frac{\partial g}{\partial y}tE_1$$

so using equation (1):

$$\theta^{**} \approx \theta^* + tE_1[\frac{w}{E_1}\frac{\partial g}{\partial w} - \frac{\partial g}{\partial y} - 1] = \theta^* + tE_1\epsilon^c$$

For a given wage w, the group of people at the kink are those with

$$\theta^* < \theta < \theta^* + tE_1\epsilon^c.$$

In the absence of the kink, these people would have earnings of $k(w, 0) + \theta$, which means that all the people with a wage w earning from E_1 to $E_1(1 + t_1\epsilon^c)$ get pushed to the kink. Now notice that this range does not depend on w. So, we can conclude that (to first order) the set of people who would have earned from the kink point E_1 to a higher level $E_1(1 + t_1\epsilon^c)$ are all pushed to the kink. If we could estimate the excess fraction at the kink, and the counterfactual density of people who would have earned amounts just above E_1 in the absence of the kink, we could potentially estimate ϵ^c, which is what Saez proposes in his AEJ-Policy paper.

Figure 1: How March 1996 fare change affected real revenue/mile and miles driven

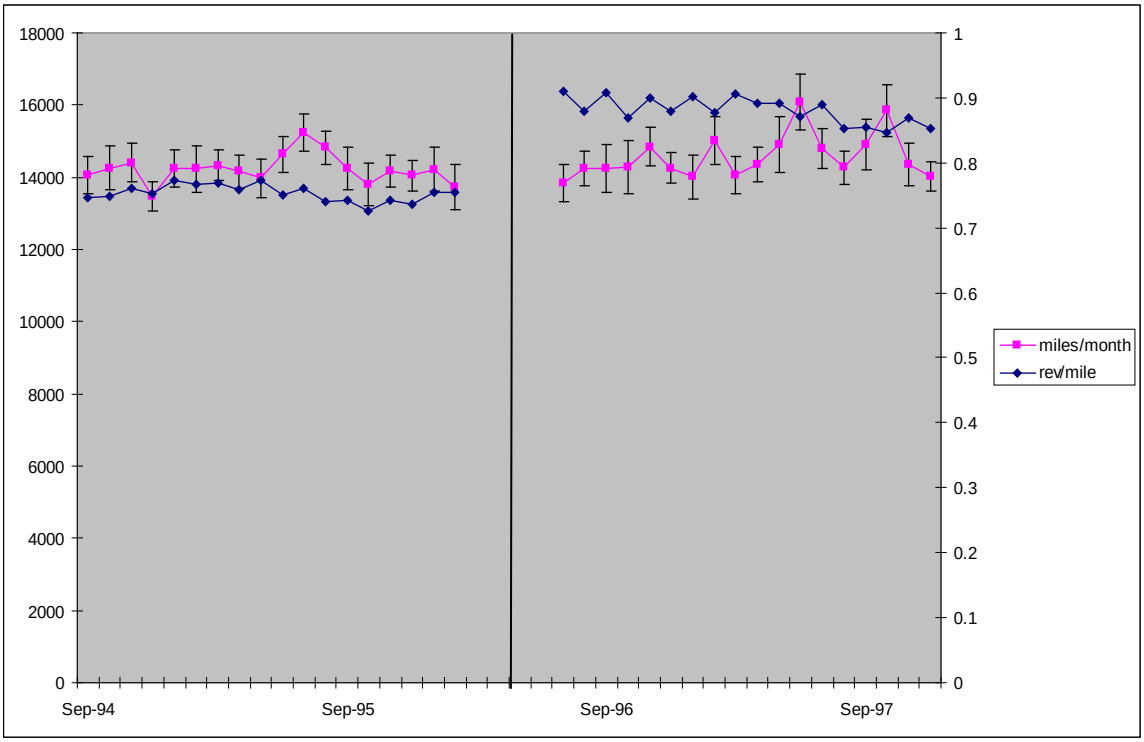
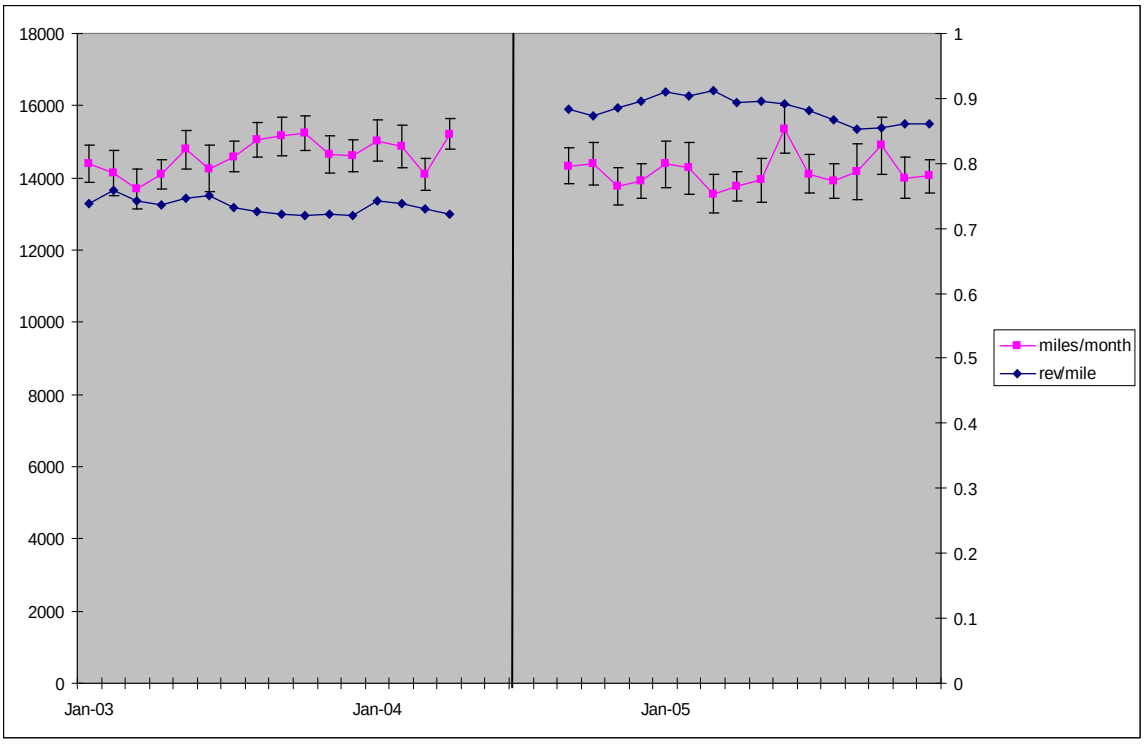


Figure 2: How May 2004 fare change affected real revenue/mile and miles driven



Appendix: Timeline regarding taxi decisions after 1998:

May 13th, 1998: city wide taxi drivers strike

May 28th, 1998: city wide taxi drivers strike
http://socialjustice.ccnmtl.columbia.edu/index.php/Alliance_Achievements
..\Labor Supply Project\Alliance_Achievements.htm

March 2002: New York City Taxi Workers Alliance organized forum to hear taxi driver's stories of their financial deterioration after September 11. Federal Emergency Management Agency had assisted taxi garages and brokers but not the drivers and at this hearing, FEMA officials heard the taxi drivers' stories. Soon after, FEMA opened a new Rental and Mortgage Assistance program- over 2,000 drivers participated.
http://socialjustice.ccnmtl.columbia.edu/index.php/Alliance_Achievements
..\Labor Supply Project\Alliance_Achievements.htm

Table 2b: Simple Difference Table (Balanced Panel):
(no other controls)

	Change in Revenue per Mile		Change in Miles Driven	
1996 Fare Increase	+ $0.15***	(+ 19.2 %)	- 819 miles*	(- 5.6 %)
2004 Fare Increase	+ $0.15***	(+ 20.9 %)	- 764 miles**	(- 5.1 %)

Difference Table:
(controls for month and days since last inspection)

	Change in Revenue per Mile		Change in Miles Driven	
1996 Fare Increase	+ $0.15***	(+ 19.0 %)	- 758 miles*	(- 5.2 %)
2004 Fare Increase	+ $0.15***	(+ 20.9 %)	- 758 miles**	(- 5.1 %)

All changes labeled with *** are significant at the 0.1% level; those with ** are significant at the 1% level, and those with * at the 10% level. Revenue is in December 2005 Dollars. Miles driven measures the number of miles driven since the last inspection. The average number of days between inspections is 122.6 with a standard deviation of 3.86 days in 1996, and 121.7 with a standard deviation of 2.08 in 2004.

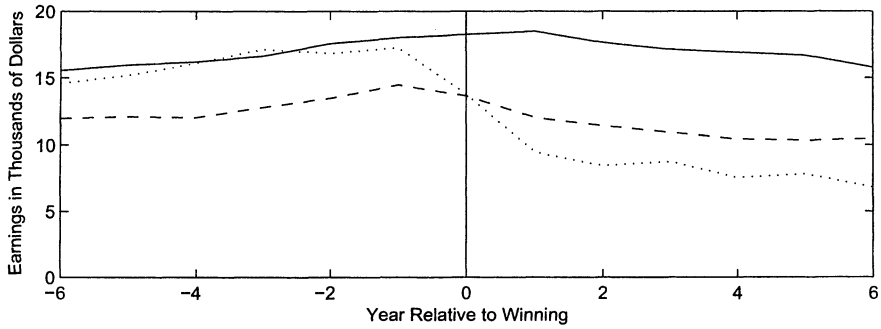


FIGURE 1. AVERAGE EARNINGS FOR NONWINNERS, WINNERS, AND BIG WINNERS

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.

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.

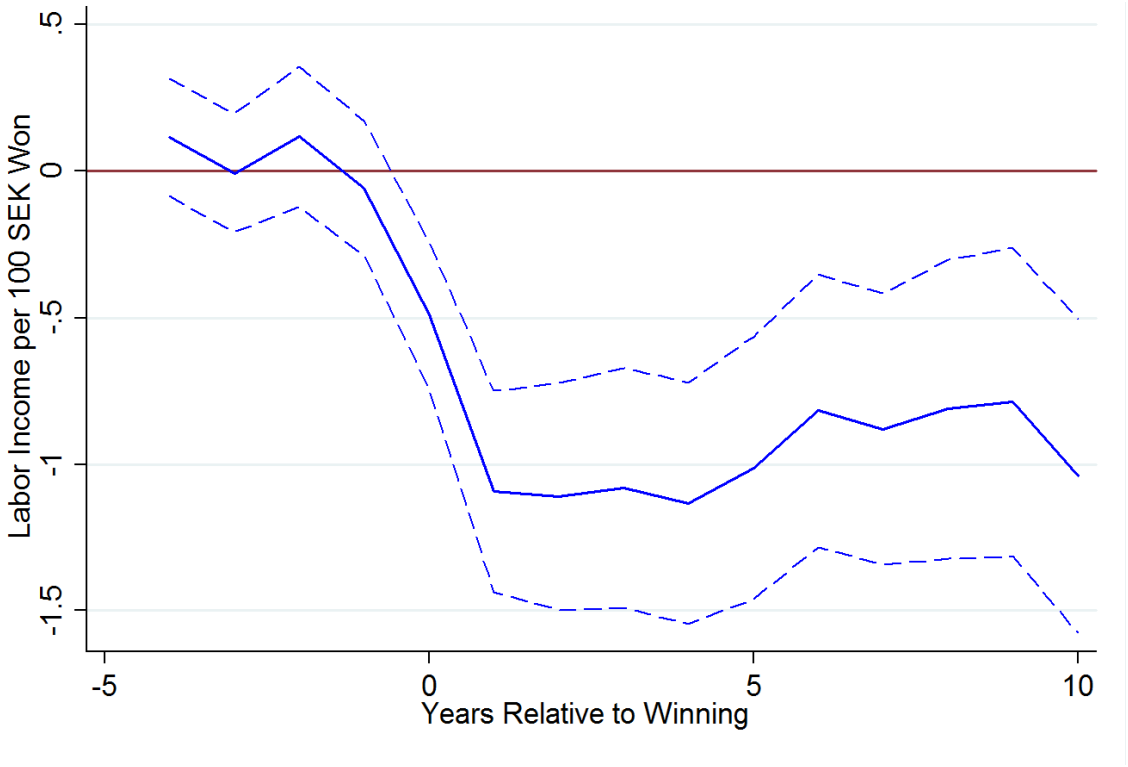
In the fifth specification we add a quadratic term in the prize. Rather than report the coefficient on the quadratic term, we report the derivative of the expected earnings as a function of the prize at two values of the prize, zero and the median prize ($32,000 per year). The estimates of the MPE based 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.

Figure 1: The Effect of Wealth on (Individual) Labor Earnings Over Time, Pooled Lottery Sample



Notes: This figure reports results from regressions using labor income as dependent variable for the pooled sample of lottery winners. The figure reports coefficients and standard errors on each year before and after winning the prize. The sample is restricted to lottery winners who won between the ages of 21 and 64.

Table 6
The Effect of Wealth on Labor Earnings

	Labor earnings (1k)		Labor earnings > 25k SEK		Longer run labor earnings (1k)		
	1 year after win	2 years after win	1 year after win	2 years after win	3-year total	5-year total	10-year total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Prize amount (in 100k SEK)	-1.036	-1.093	-0.0017	-0.0025	-2.966	-4.582	-9.436
	(0.133)	(0.169)	(0.0005)	(0.0006)	(0.449)	(0.714)	(1.523)
	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]	[0.000]	[0.000]
N	340,053	338,179	340,053	338,179	336,105	331,409	311,756
R^2	0.756	0.687	0.393	.4078489	0.737	0.714	0.679

Notes: The sample is restricted to lottery participants who were between the age of 21 and 64 at the time of winning the lottery or being assigned to the control group. All earnings and prize amounts are in 2010 Swedish Krona (SEK). The baseline controls include an indicator variable for gender, a quintic in age, educational attainment and controls for last year's labor earnings. Standard errors are clustered by individual and are reported in parentheses, and p-values are in brackets.

Table 11
The Effect of Wealth on Household Labor Earnings

	Labor earnings		Labor earnings > 25k SEK		Longer-run labor earnings		
	1 year after win	2 years after win	1 year after win	2 years after win	3-year total	5-year total	10-year total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Labor Earnings of Winners and Spouses							
Winners							
Prize amount (in 100k SEK)	-1.112	-1.088	-0.001	-0.003	-2.808	-4.485	-9.438
	(0.180)	(0.206)	(0.001)	(0.001)	(0.537)	(0.921)	(2.007)
	[0.000]	[0.000]	[0.043]	[0.001]	[0.000]	[0.000]	[0.000]
Spouses							
Prize amount (in 100k SEK)	-0.236	-0.400	0.000	-0.001	-0.607	-1.846	-1.599
	(0.166)	(0.239)	(0.001)	(0.001)	(0.604)	(1.134)	(2.748)
	[0.155]	[0.094]	[0.870]	[0.349]	[0.316]	[0.104]	[0.561]
p-value of test of equal effects	[0.001]	[0.024]	[0.103]	[0.040]	[0.007]	[0.061]	[0.018]
Panel B: Total Labor Earnings of Household							
Prize amount (in 100k SEK)	-1.407	-1.531	-0.0003	-0.0016	-3.568	-5.667	-8.407
	(0.326)	(0.362)	(0.0005)	(0.0006)	(1.053)	(1.811)	(3.759)
	[0.000]	[0.000]	[0.597]	[0.009]	[0.001]	[0.002]	[0.025]
N	221,973	220,868	221,973	220,868	219,695	216,949	205,115
Include baseline controls	y	y	y	y	y	y	y

Notes: This table reports results of estimating equation (1) on the baseline sample of lottery winners. All earnings and prize amounts are in 2010 Swedish Krona (SEK). Panel A investigates nonlinear treatment effects directly by estimating a quadratic in prize amount, while Panel B allows for a linear spline with a knot at 1M SEK. The baseline controls include an indicator variable for gender, a quintic in age, and controls for last year's labor earnings and capital income. Standard errors are clustered by individual and are reported in parentheses, and p-values are in brackets.

Figure 3.1: Nonlinear Budget Set Caused by Tax on Income above E_1

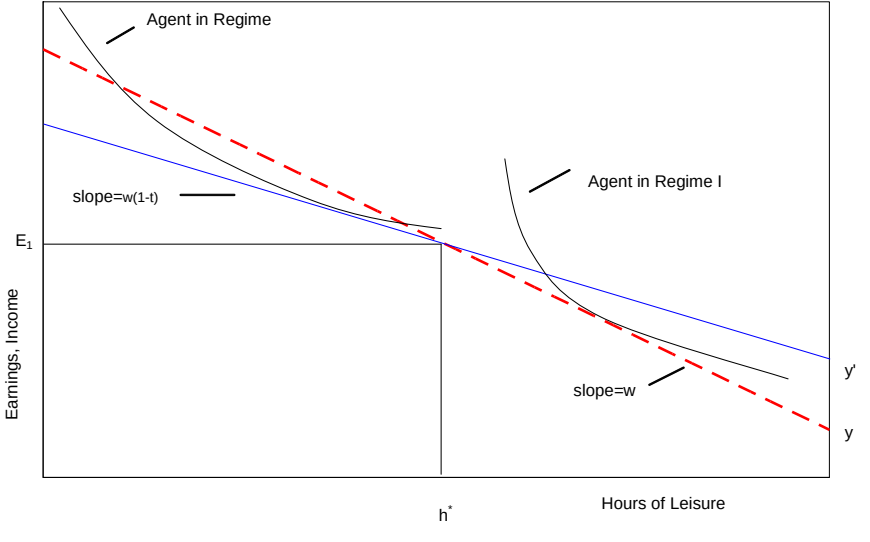


Figure 3.2: Someone in Regime III

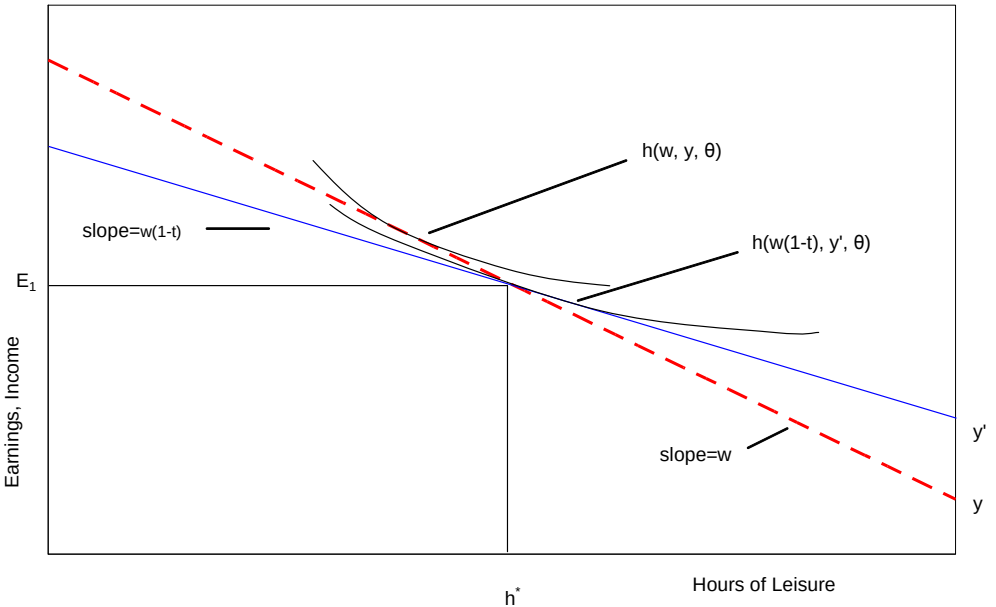


Figure 3.3: Calculating the Earnings of the Marginal "Buncher" following Introduction of tax t on Earnings in Excess of E_1