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 $\ell = T - \overline{h}$. A key observation is that

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

(which is true by definition of $e(w,u^0)$). 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

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

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

$$\frac{(\overline{h} - h^0)}{h^0} = .2$$

then the compensating differential is

$$\frac{(w - w^0)}{w^0} \approx \frac{.5 \times .2 \times .2}{\epsilon^c} = \frac{.02}{\epsilon^c}$$

For example, if $\epsilon^c = 0.2$, this formula implies you need a ***10% higher average wage*** to take a job with ***20% longer hours*** than the reference (unrestricted) job. Note that if there are different groups of workers out there, with different preferences, the group with a **smaller** ϵ^c needs a **larger** compensating differential for long hours. The "low ϵ^c" workers will sort away from the long hour jobs.

The compensating differential formula for **low hours** will be different if workers receive unemployment compensation for their lower hours. For an application of this, see:

Emilia Del Bono and Andrea Weber. "Do wages compensate for anticipated working time restrictions? Evidence from seasonal employment in Austria." JOLE, 2008

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. The implied IV estimate is -0.23 with fixed effects. Notice that (ignoring "GE" effects) b is interpretable as an estimate of the uncompensated labor supply elasticity.

For discussion:

a) why cab drivers?

b) what if there are demand side effects?

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) scatch-ticket lotteries, known as Triss. 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 Figures 1-2, and Tables 1-4 (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 look at Figure 2you 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.

They also use a simulation technique to estimate the parameters in a very simple Stone Geary model:

$$U = \sum_{t=0}^{T-1} \frac{1}{(1+\delta)^t}[\beta log(c_t - \gamma_c) + (1-\beta) log(\gamma_h - h_t)$$

where the budget contraint is:

$$\begin{aligned} A_{t+1} &= (1+r)(A_t - c_t + w_t h_t + a_t) \\ h_t &= 0 \quad if\, t \geq R \end{aligned}$$

To estimate they assume people die at 80, retire at 65, set $r = 0.02$, $\gamma_c = 20000$, and assume $a_t = 0$ for all t¡65 and $a_t = 70\%$ of annual after-tax earnings in retirement.

Table 7 is very interesting because it shows that there is some additional negative effect on spouses, around 50% 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 partially 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:

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

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.

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.

Table 2a
Simple Estimates of Labour Supply Using All Observations and Controlling for Driver Heterogeneity with Fixed Effects

	Change in revenue per mile	Change in miles driven
Simple difference table: medallion fixed effects, no other controls		
1996 fare increase	+ \$0.14 (+ 17%)	– 477 miles (– 3.2%)
2004 fare increase	+ \$0.15 (+ 19%)	– 824 miles (– 5.6%)
Difference table: medallion fixed effects, controls for month and days since last inspection		
1996 fare increase	+ \$0.14 (+ 17%)	– 399 miles (– 2.7%)
2004 fare increase	+ \$0.15 (+ 19%)	– 818 miles (– 5.6%)

Notes
All changes are computed as the coefficient of a dummy variable indicating the year noted and are significant at the 0.1% 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 (4 months), with a standard deviation of 4 days. Since the panel is not fully balanced, these results are computed from a regression that includes medallion fixed effects in order to use all the data.
The regressions in the last two rows also contain a variable measuring the number of days since the taxi was last inspected.

Table 2b
Simple Estimates of Labour Supply Using Only Observations with Continuous Panel Data

	Change in revenue per mile	Change in miles driven
Simple difference table (balanced panel): no other controls		
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		
1996 fare increase	+ \$0.15*** (+ 19.0%)	– 758 miles* (– 5.2%)
2004 fare increase	+ \$0.15*** (+ 20.9%)	– 758 miles** (– 5.1%)

Notes
***,**,*Indicate significant at the 0.1%, 1%, 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 days in 2004.

sample of data these estimates are all roughly $-5\%/20\% = -0.25$. As we shall see below, this is close to the estimate obtained from a more complete econometric analysis.

III. A Simple Model of Behaviour for Taxi Driver Labour Supply

What is apparent from the previous discussion is that drivers do not face explicit wage rates, but instead face a taxi fare function that relates their income to hours worked through the miles they travel. A simple model of this behaviour starts with the standard assumption that a driver has utility function

(1) $$u = u(h, y),$$

TABLE 5
(LOG) MILES DRIVEN AS A FUNCTION OF (LOG) REVENUE PER MILE

	(1) OLS	(2) OLS IV	(3) Fixed effects	(4) Fixed effects IV
ln(real revenue per mile)	− 0.42	− 0.13	− 0.40	− 0.23
	(0.01)	(0.03)	(0.01)	(0.02)
ln(days since inspection)	0.63	0.63	0.72	0.79
	(0.06)	(0.10)	(0.04)	(0.06)
February	0.00	− 0.00	0.02	0.04
	(0.01)	(0.02)	(0.02)	(0.09)
March	− 0.03	− 0.03	− 0.00	− 0.01
	(0.01)	(0.02)	(0.02)	(0.09)
April	− 0.01	− 0.01	− 0.04	− 0.06
	(0.01)	(0.02)	(0.02)	(0.09)
May	0.02	0.02	0.02	0.02
	(0.01)	(0.02)	(0.01)	(0.01)
June	0.03	0.04	0.04	0.06
	(0.01)	(0.02)	(0.02)	(0.09)
July	− 0.01	− 0.02	0.02	0.01
	(0.01)	(0.02)	(0.02)	(0.09)
August	− 0.02	− 0.01	− 0.05	− 0.06
	(0.01)	(0.02)	(0.02)	(0.09)
September	− 0.02	− 0.03	− 0.02	− 0.02
	(0.01)	(0.02)	(0.01)	(0.01)
October	− 0.01	− 0.02	− 0.01	0.03
	(0.01)	(0.02)	(0.02)	(0.09)
November	− 0.04	− 0.03	− 0.02	− 0.01
	(0.01)	(0.02)	(0.02)	(0.09)
December	− 0.03	− 0.00	− 0.05	− 0.06
	(0.01)	(0.02)	(0.02)	(0.09)
Constant	6.38	6.46	5.95	5.67
	(0.31)	(0.48)	(0.18)	(0.31)
Observations	33,962	12,281	33,962	12,244
R-squared	0.06	0.03	0.07	
Number of medallions		2645		2514

Notes
Standard errors in parentheses. Unit of observation: one driver during a four-month period. Instrument: 1996 fare increase and 2004 fare increase. Fixed effects: medallion level.

lease rate by only 8% in order that most of the fare increase 'would end up in drivers' pockets'. (See the timeline in the Appendix for more details.) Although we cannot document all the details, there is a potential for large changes in the lease rate around the time of the fare changes that might affect the labour supply of drivers who own medallions, despite the fact that these lease rates do not affect them directly, because of the potential incentives that a change in the lease rate might give an owner–driver.

For example, this change in the incentive to lease to others could potentially cause at least one serious bias: it could selectively remove people from our sample after the fare increase, because they then start renting their evening shift to others.

Since our dataset contains the universe of drivers inspected by the TLC between 1990 and 2005, we may examine the number of medallions that switched from being associated

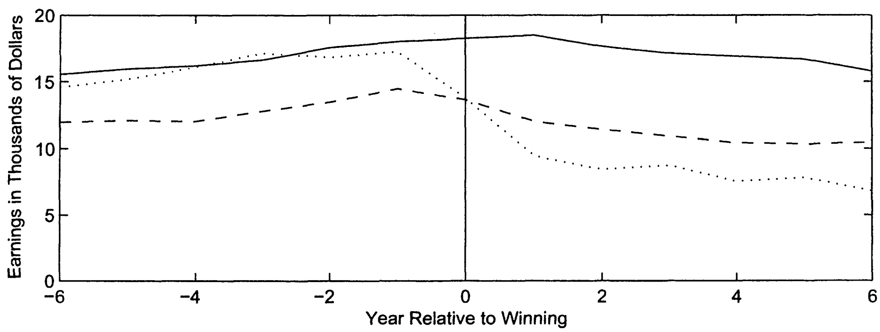


FIGURE 1. AVERAGE EARNINGS FOR NONWINNERS, WINNERS, AND BIG WINNERS

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.

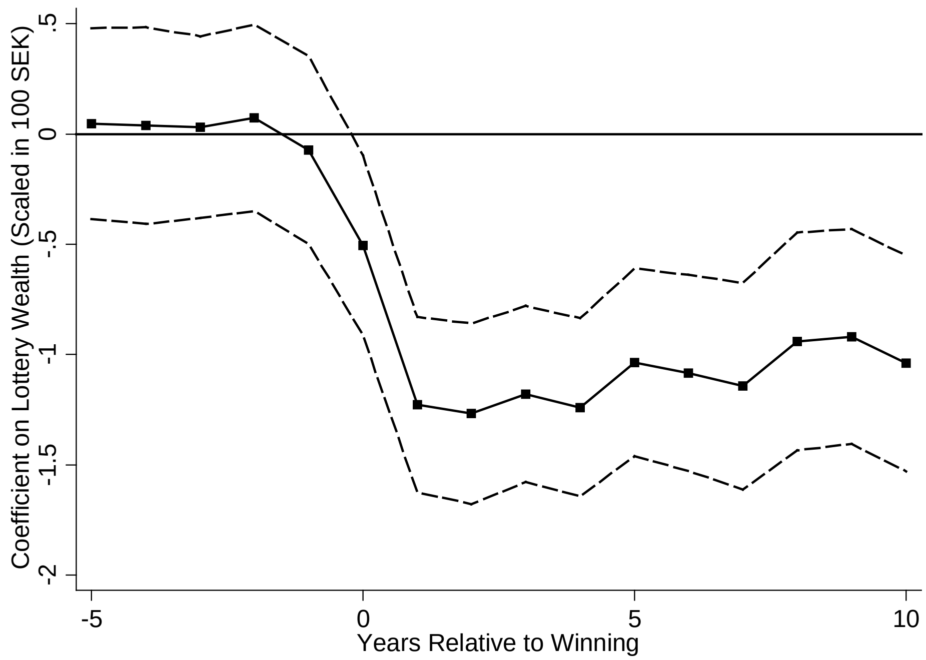
In the fifth specification we add a quadratic term in the prize. Rather than report the coefficient on the quadratic term, we report the derivative of the expected earnings as a function of the prize at two values of the prize, zero and the median prize (\$32,000 per year). The estimates of the MPE based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

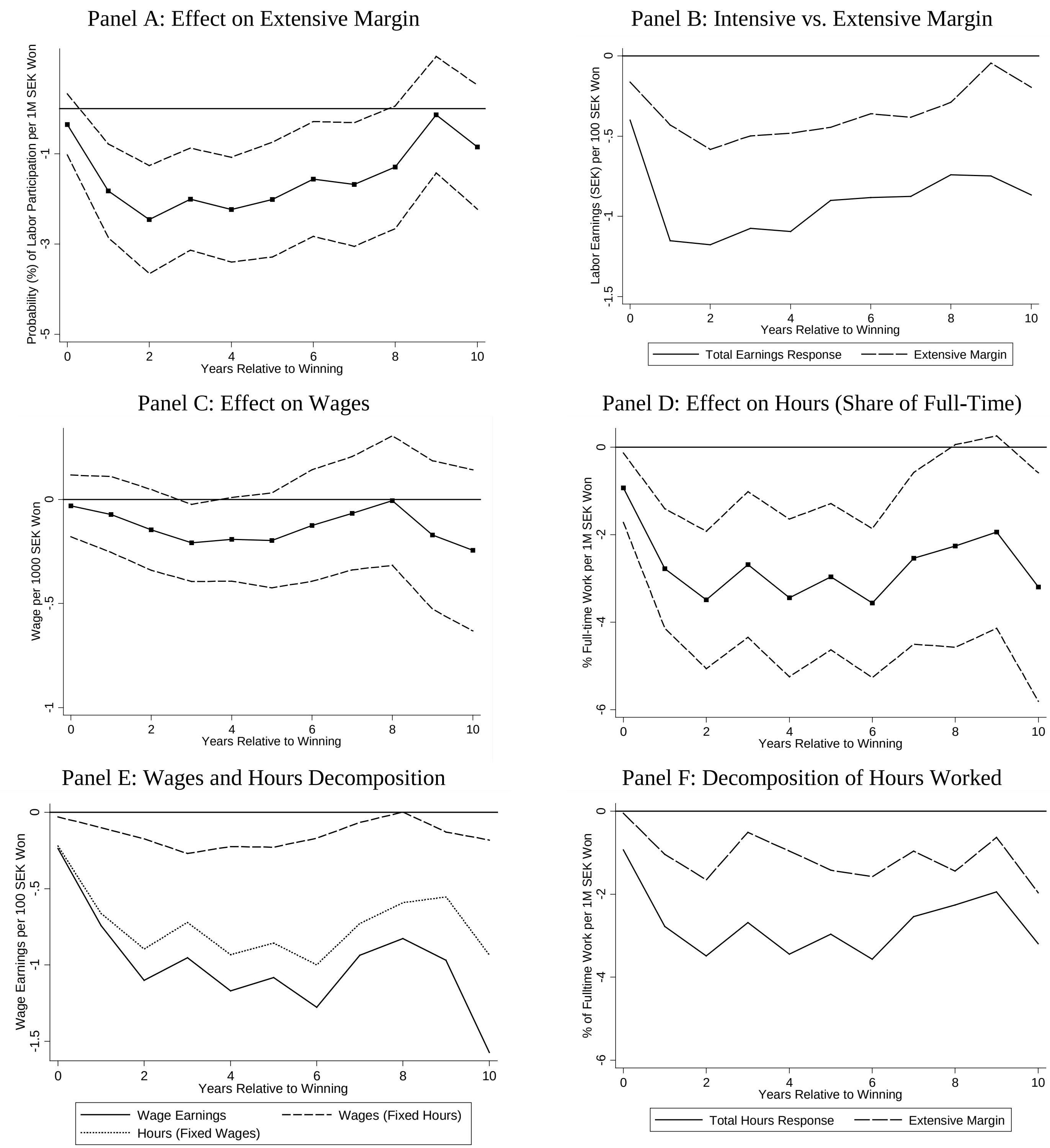
[22] Although more than half the original sample is dropped in this specification, the precision is not significantly affected because most of the variation in the lottery prize is among the winners.

Figure 1: Effect of Wealth on Individual Gross Labor Earnings



Notes: This figure reports estimates obtained from equation (2) estimated in the pooled lottery sample with gross labor earnings as the dependent variable. A coefficient of 1.00 corresponds to an increase in annual labor earnings of 1 SEK for each 100 SEK won. Each year corresponds to a separate regression and the dashed lines show 95% confidence intervals.

Figure 2: Margins of Adjustment



Notes: This figure reports estimates obtained from equation (2) estimated for the different outcomes discussed in section III.A. Each year corresponds to a separate regression. The dashed lines in Panel A, C, and D display 95% confidence intervals. Panels A and B are estimated in the full sample, whereas Panels C to F are estimated in the subsample with observable wages.

Table 1. Distribution of Prizes

	Pooled Sample		Individual Lottery Samples							
			PLS		Kombi		Triss-Lumpsum		Triss-Monthly	
	Count	Share	Count	Share	Count	Share	Count	Share	Count	Share
0 to 1K SEK	25,172	10.0%	0	0.0%	25,172	99.0%	0	0.0%	0	0.0%
1K to 10K SEK	204,626	81.3%	204,626	92.0%	0	0.0%	0	0.0%	0	0.0%
10K to 100K SEK	16,429	6.5%	15,520	7.0%	0	0.0%	909	27.8%	0	0.0%
100K to 500K SEK	3,685	1.5%	1,654	0.7%	0	0.0%	2,031	62.1%	0	0.0%
500K to 1M SEK	355	0.1%	195	0.1%	0	0.0%	160	4.9%	0	0.0%
≥1M SEK	1,481	0.6%	481	0.2%	263	1.0%	168	5.1%	569	100.0%
TOTAL	251,748		222,476		25,435		3,268		569	

Notes: This table reports the distribution of lottery prizes for the pooled sample and the four lottery subsamples.

Table 2. Effect of Wealth on Individual Gross Labor Earnings

	t = 1	t = 2	3-year total	5-year total	10-year total	Event study estimate t = 1-5
	(1)	(2)	(3)	(4)	(5)	(6)
Prize Amount (SEK/100)	-1.152	-1.177	-3.219	-4.681	-8.033	-1.068
SE	(0.153)	(0.191)	(0.517)	(0.917)	(1.961)	(0.149)
p	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
N	199,168	211,555	193,312	186,819	173,129	249,278

Notes: This table reports results of estimating equation (2) in the pooled lottery sample with gross labor earnings as the dependent variable. The prize amount is scaled so that a coefficient of 1.00 implies a 1 SEK increase in earnings per 100 SEK won.

Table 3. Effect of Wealth on Different Measures of Individual Earnings

	Labor Earnings		Wage Earnings	Self-employment Income	Unemployment Benefits	Pensions	Taxable earnings		
Taxes	Pre-tax	Pre-tax incl. SSC	Pre-tax	Pre-tax	Pre-tax	Pre-tax	Pre-tax	Post-tax	Post-tax incl. SSC benefits
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Prize Amount (SEK/100)	-1.068	-1.412	-0.967	-0.142	0.035	0.157	-0.900	-0.580	-0.624
SE	(0.149)	(0.199)	(0.151)	(0.036)	-0.026	-0.085	-0.131	-0.081	-0.084
p	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[0.177]	[0.064]	[<0.001]	[<0.001]	[<0.001]
N	249,278	247,847	247,915	248,058	248,058	248,058	249,278	247,847	247,847

Notes: This table reports event-study estimates obtained by estimating equation (2) in the pooled lottery sample with different earnings measures as the dependent variable. The earnings measure in column (2) includes SSC paid by the employer and the column (9) earnings measure includes the implicit employee benefit of SSC. Labor earnings in column (1) and (2) includes wage earnings and self-employment income used in columns (3) and (4). Taxable earnings in columns (7) to (9) includes labor earnings (column 1), unemployment benefits (column 5) and pension income (column 6). The variables are scaled so that a coefficient of 1.00 implies a 1 SEK increase in earnings per 100 SEK won.

Table 4. Margins of Adjustment

	Extensive Margin (> 25K SEK)					
	Labor Earnings	Wage Earnings	Self-employment	Pension Income (≥ Age 50)	Hours (Percent of Full-time)	Pre-tax Monthly Wages
	(1)	(2)	(3)	(4)	(5)	(6)
Prize Amount	-2.067	-2.244	-0.623	0.458	-3.109	-0.158
SE	(0.449)	(0.475)	(0.253)	(0.507)	(0.616)	(0.085)
p	[<0.001]	[<0.001]	[0.014]	[0.366]	[<0.001]	[0.063]
Proportion/mean	77.3%	71.0%	5.4%	36.4%	81.6%	22,973
N	249,278	247,915	248,058	130,848	110,080	110,080

Notes: This table reports event-study estimates obtained by estimating equation (2) in the pooled lottery sample. The variables in columns (1) to (5) are scaled so that a coefficient of 1.00 implies a 1 percentage point increase in participation or fraction of full-time worked per million SEK won, whereas the prize amount is scaled by 1000 SEK in column (6).

Table 5. Simulation-based Estimates of Model Parameters

Panel A: Minimum-distance Estimates	Minimum-distance estimates	Panel B: Reduced-form and Model-based Moments	Reduced-form	Model-based
β / SE	0.855 / (0.010)	100K SEK Prize	-3.12	-2.95
δ / SE	0.010 / (0.005)	1M SEK Prize	-2.99	-2.94
γ_h / SE	1852 / (39.7)	Below median earnings	-2.36	-2.99
		Above median earnings	-2.66	-2.91
Goodness-of-fit, $\chi^2(8)$	3.428	Age 21-34	-3.06	-1.71
p-value	[0.095]	Age 35-54	-3.13	-3.41
		Age 55-64	-1.06	-2.66

Notes: This table reports results of estimating the dynamic model via indirect inference, with asymptotic standard errors in parentheses. The goodness-of-fit test uses the minimized value of weighted minimum distance procedure, based on 11 moments and 3 parameters. The reduced form and model-based moments in Panel B shows the effect on total five-year after-tax earnings scaled in units of 100 SEK. The non-linear moments in the first two rows compares a lottery win of 100K and 1M using the reduced-form results from a quadratic model.

Table 6. Implied Labor Supply Elasticities from Simulated Model

Panel A: Implied Lifetime Wealth Effects at Various Ages

Age at Win	Implied Lifetime Wealth Effect (1)	Cumulative Wealth Effect Over First 10 Years (2)	Effect Over First 10 Years as Share of Lifetime Effect (3)
20	-0.186	-0.031	16.7%
25	-0.172	-0.032	18.7%
30	-0.159	-0.035	22.2%
35	-0.145	-0.041	28.4%
40	-0.129	-0.049	38.3%
45	-0.114	-0.056	49.0%
50	-0.107	-0.078	72.9%
55	-0.074	-0.074	100.0%
60	-0.046	-0.046	100.0%

Panel B: Implied Labor Supply Elasticities

(Win at Age 50, Retire at Age 65, Die at Age 80)

(1)	Effect of lottery prize on total labor earnings over remaining working life (Implied Lifetime Wealth Effect)	-0.107
(2)	Effect of permanent change in wages on total hours worked (Uncompensated (Marshallian) Elasticity)	0.001
(3)	Effect of transitory change in wages on hours worked (Intertemporal Frisch Elasticity)	0.201
(4)	Implied Compensated (Hicksian) Labor Supply Elasticity (from (1) and (2) through Slutsky equation)	0.108

Notes: This table reports key labor supply elasticities implied from the model using the parameters reported in Table 5. Panel A reports elasticities at different ages. Each row computes the lifetime wealth effect and the wealth effect over the first 10 years. Row (1) in Panel B reports the effect of a lottery prize on total labor earnings (i.e., sum of dy/dL across all remaining working years, as implied by model), row (2) reports the implied effect of a permanent increase of wages on total hours worked (summed up across all remaining working years), row (3) reports the Frisch elasticity (i.e., effect of a transitory change in wages on hours worked), and row (4) shows the implied Hicksian elasticity from the Slutksy equation.

Table 7. Effect of Wealth on Household Labor Earnings

	Panel A: Married Winners			
	Household	Winner	Spouse	Difference
	(1)	(2)	(3)	(4)
Prize Amount (SEK/100)	-1.439	-0.981	-0.458	-0.522
SE	(0.298)	(0.200)	(0.206)	(0.276)
p	[<0.001]	[<0.001]	[0.026]	[0.059]
N	144,979	144,979	144,979	144,979

	Panel B: Unmarried Winners		Panel C: Total Sample	
	Household	Winner	Household	Winner
	(5)	(6)	(7)	(8)
Prize Amount (SEK/100)	-1.259	-1.259	-1.324	-1.068
SE	(0.229)	(0.229)	(0.193)	(0.149)
p	[<0.001]	[<0.001]	[<0.001]	[<0.001]
N	101,473	101,473	249,278	249,278

Notes: This table reports event-study estimates obtained by estimating equation (2) on winners, winners' spouses, and at the household level for different subsamples. Panel A includes all winners that were married the year before the lottery event, Panel B includes those that were unmarried, and Panel C includes both married and unmarried winners. The prize amount is scaled so that a coefficient of 1.00 implies a 1 SEK increase in earnings per 100 SEK won. The estimates in Panel A includes baseline controls for the winner's spouse.

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

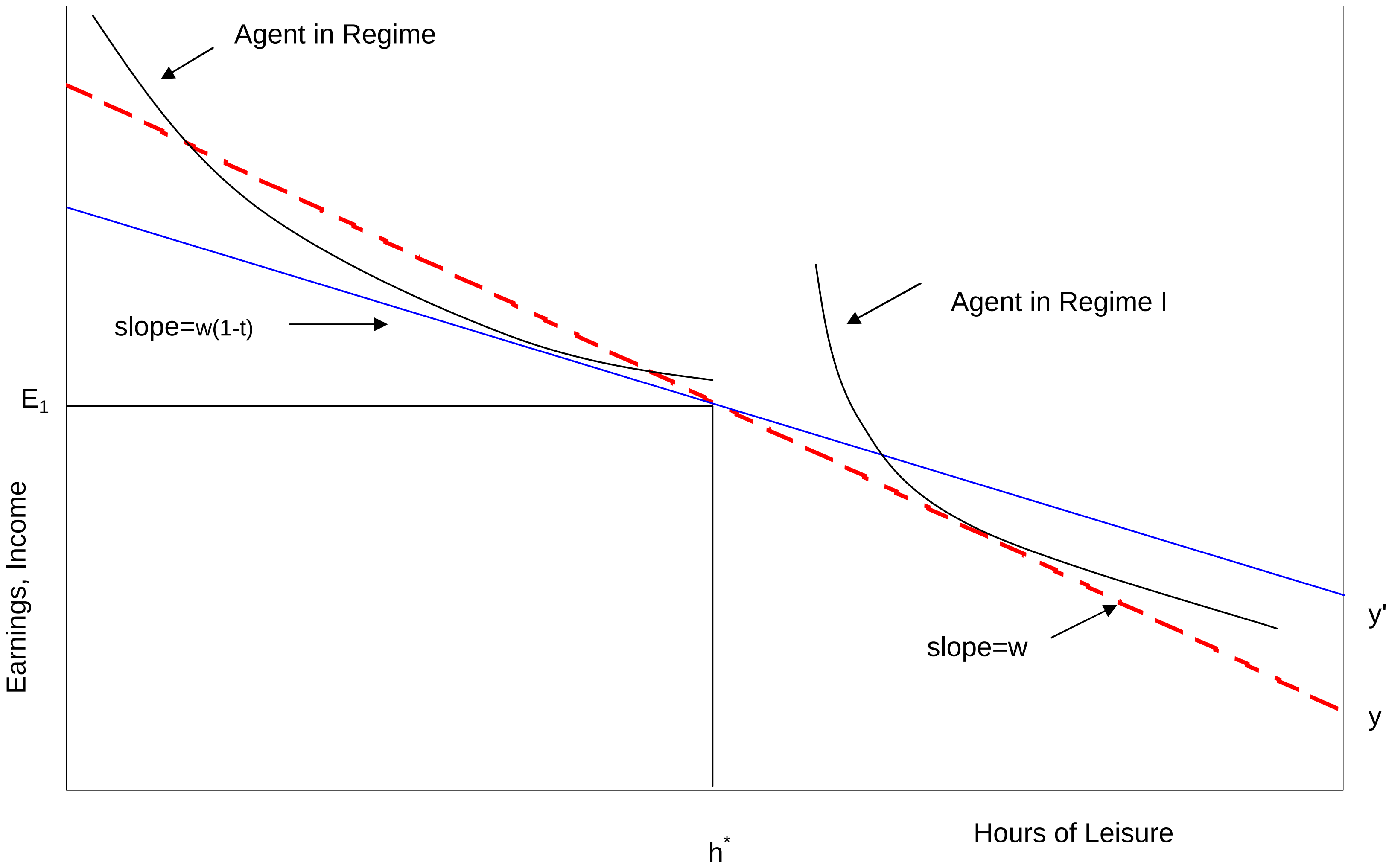


Figure 3.2: Someone in Regime III