SIEPR Discussion Paper No. 10-026

Estimating Income Tax Salience

by

Dominic Coey

Stanford Institute for Economic Policy Research
Stanford University
Stanford, CA 94305
(650) 725-1874

Estimating Income Tax Salience

Dominic Coey*

March 20, 2011

Abstract

I develop a procedure for estimating how accurately people perceive their income tax schedules. Using Survey of Income and Program Participation data on hourly-paid workers, I reject the hypothesis of fully salient taxes. Younger workers seem particularly prone to misperceiving their taxes. Even for younger workers, however, the welfare losses associated with this less than full salience appear to be small.

1 Introduction

Tax schedules can be complicated. Changes in marginal federal and state income tax rates, the phasing in and out of the Earned Income Tax Credit, and the cap on earnings subject to the payroll tax are all example of policies which induce kinks in budget constraints. Given these complications, people may incorrectly judge their tax liability. Put otherwise, their taxes may not be fully salient to them. This paper describes how tax salience might be identified and estimated from data on the labor supply of hourly paid workers. The data do not support the hypothesis of full tax salience. A natural conjecture is that workers'

*Department of Economics, Stanford University. Email: dcoey@stanford.edu. I am grateful to Douglas Bernheim, Liran Einav, Caroline Hoxby, Seema Jayachandran, Jonathan Levin, Thomas MaCurdy, Luigi Pistaferri and seminar participants at Stanford. Their comments and suggestions improved this paper enormously.

misperceptions of their tax schedule decrease with age, as they become more experienced and familiar with the details of their taxes. Consistent with this conjecture, the estimation results suggest that tax salience does increase with age. The overall welfare costs of this misperception are small, and they fall with age.

This work has some policy implications. It suggests that imperfect awareness of one's taxes does not necessarily entail substantial welfare losses. Educating people about the precise nature of their tax liabilities need not be of substantial benefit to them. More generally, imperfectly informed workers might be making choices which are close to optimal for them. Imperfect information by itself does not constitute a rationale for corrective action. The costs of a policy aimed at improving people's choices should be weighed against its potential benefits, which as in this application, may be relatively minor. Nor is there evidence that welfare losses from misperception small on average but large for some particular demographic groups - even for the young workers, who are especially likely to misunderstand their tax liabilities, these losses are quite small.

If individuals correctly perceive their tax schedules, basic microeconomics suggests that they will tend to bunch at those points where the marginal tax rate jumps. Figure 1 illustrates this: on the budget constraint AA', there are a range of marginal rates of substitution between leisure and consumption which are consistent with utility maximization at leisure demand level L. L is optimal given indifference curve CC'; it is also optimal given indifference curve DD'. At all other interior points of the budget constraint, there is only one such marginal rate of substitution. If preference heterogeneity is continuously distributed in the population, the distribution of leisure demand has a point mass at L, as in Figure 2a. With classical measurement error in observed hours the point mass disappears, but there is still a mode at L, as in Figure 2b.

Labor supply data typically does not reveal as much bunching around kink points as this analysis seems to suggests.[1] As others have recognized, this may be because people

[1] MaCurdy et al. (1990) make this point, although notable exceptions of bunching do exist, as documented

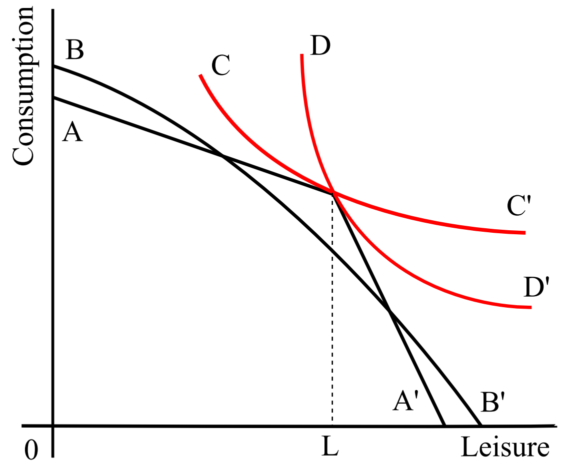


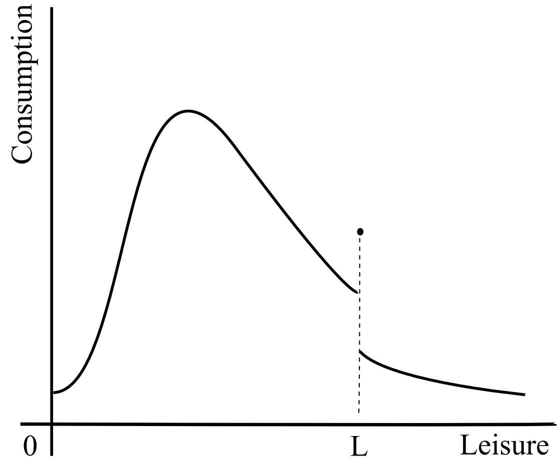
Figure 1: Kinked Budget Constraint

do not correctly perceive their taxes (Chetty et al. (2011); Liebman and Zeckhauser (2004); Saez (2010)). Suppose that when choosing labor supply, workers have in mind not their true schedule AA' but the smoothed schedule BB', also shown in Figure 1. Then, as depicted in Figure 2c, there need not be a mode at L. This suggests that it may be possible to identify tax salience from the degree of bunching at kinks. Misperception of taxes can explain labor supply choices which do not tend to accumulate around kink points. Correctly perceived taxes with classical measurement error cannot. This is essentially the variation in the data which identifies tax salience in this paper. Section 2.2 goes through the identification argument in more detail.

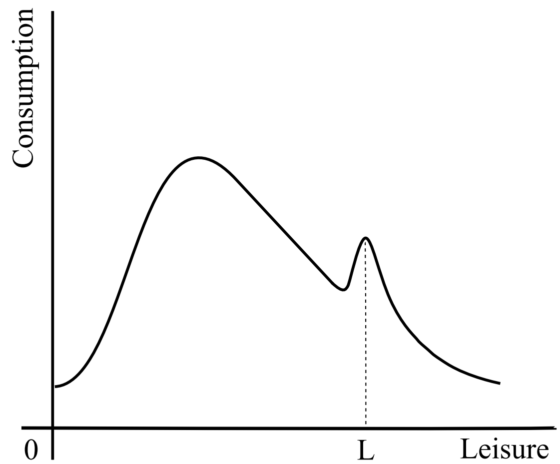
Tax misperception is not the only explanation for non-bunching at kinks. Firms may constrain the number of hours their employees work. The model accounts for some people being constrained to work roughly 40 hour weeks. I also present survey evidence supporting the claim that it is truly salience, rather than job flexibility, that is driving the results.

This paper builds on both the established literature of structural labor supply models with nonlinear budget constraints, and the burgeoning work on tax salience. Burtless and Hausman (1978) is an early contribution in the former tradition, and MaCurdy et al. (1990) provide

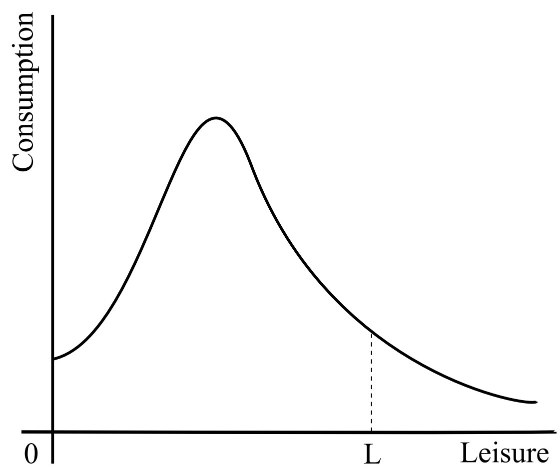
by Blundell et al. (2001); Chetty et al. (2011); Friedberg (2000) and Saez (2010), among others.



(a) Full Salience, No Measurement Error



(b) Full Salience, Measurement Error



(c) Partial Salience, No Measurement Error

Figure 2: Distribution of Leisure Demand

a detailed overview of these models. MaCurdy et al. (1990) also “smooth” the budget constraint as in this paper, but in their analysis the purpose of smoothing is to simplify the likelihood. Their smoothing parameter is chosen, not estimated. Friedberg (2000) uses jumps in marginal tax rates induced by the Social Security earnings test to identify income and substitution effects. The earnings test generates particularly high marginal tax rates and Friedberg documents significant bunching amongst the elderly at the corresponding kink. In common with these papers, I estimate preference parameters by maximum likelihood in a labor supply model with nonlinear budget constraints. In contrast to them, I also estimate parameters governing tax salience. Rosen (1976) estimates a “coefficient of tax perception” for married women, but abstracts from the the tax-induced nonlinearities.

A considerable body of evidence suggests that people do not always judge their tax liabilities accurately (Bartolome (1995); Cabral and Hoxby (2010); Chetty et al. (2009); Finkelstein (2009); Fujii and Hawley (1988); Slemrod (2006))[2]. Using tax return data, Saez (2010) documents bunching around the first kink point of the Earned Income Tax Credit and also where income tax liability starts, but nowhere else. As with Liebman and Zeckhauser (2004), I take the approach of assuming agents optimize subject to a misperceived tax schedule. They too note that this can explain non-bunching at kinks, and go on to derive the welfare implications of two particular kinds of misperceptions.

Using Danish tax records, Chetty et al. (2011) find substantial bunching at sharp kink points but negligible bunching for those kink points where net wages change by less than 10%. They explain bunching patterns with reference to the adjustment costs involved in searching for a new job, and the hours constraints firms impose. One interpretation of these adjustment costs is the cost of becoming well-informed about one’s tax liability. Chetty (2010) proposes a method of bounding structural parameters without explicitly specifying how optimization frictions such as inattentiveness affect labor supply. Interestingly, despite

[2]Fujii and Hawley find that on average people’s perceptions of their marginal income tax rate are only slightly below the truth, but the standard deviation of the difference between the two is large.

a very different modelling approach, he also concludes that the welfare losses relative to the no friction case are likely to be small - less than 1% of consumption.

There are some caveats to this study. First, I assume that peoples' misperceptions of the tax schedule take the form of the "smoothed budget constraint". It may instead be the case that people systematically over or underestimate their tax liability. In this case the welfare cost of less than full salience could be much larger. The point remains, however, that non-bunching due to tax misperception does not necessarily imply substantial welfare losses. Second, this model cannot accommodate all forms of hours constraints employers might impose upon employees. Third, the model estimates may be misleading if measurement error is incorrectly specified. In particular, I rule out a measurement error variance changing with age.

2 Model

2.1 Description

The model I propose differs from standard models of labor supply subject to nonlinear budget constraints, as the budget constraint itself is parameterized. Figure 3 illustrates this. $\gamma \in [0, \infty)$ governs how smooth the parameterized budget constraint is. As $\gamma \to \infty$, it converges to the least squares line of best fit through the real budget constraint. As $\gamma \to 0$, it converges to the real budget constraint. The details of this construction are in Appendix A. I assume that agents supply labor as if they were facing the parameterized, or perceived, constraint and estimate the γ which best explains the patterns of bunching or non-bunching in the data. $\gamma = 0$ corresponds to the standard model of utility maximization under full information. Under the maintained assumptions of the model the hypothesis of fully salient taxes ($\gamma = 0$) is testable.

There are other ways to model less than full salience. Liebman and Zeckhauser (2004)

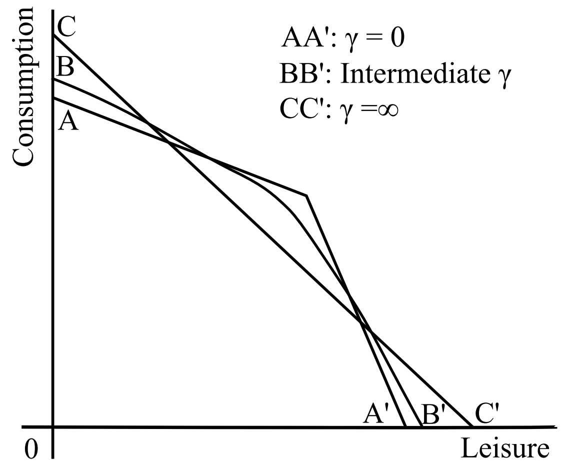


Figure 3: Parameterized Budget Constraints

suggest two: ironing, in which people believe their marginal tax rate to be their average tax rate, and spotlighting, in which they imagine their current marginal rate applies to all income levels. The parameterization I select has much in common with ironing and spotlighting. It allows for the possibility that workers are unaware of where the kinks in their constraints are. The perceived constraint is not itself explicitly derived from an economic model of decision making with optimization frictions such as costs to becoming informed. It is however parsimonious - the scalar γ entirely determines the constraint's shape. However, unlike Liebman and Zeckhauser's alternatives, it has the feature that workers' perceptions of the net income they would earn from various level of labor supply are roughly right on average, and not consistently above or below the truth.

Agent i has constant elasticity of substitution preferences over perceived yearly income Y and leisure L:

$$U_i(Y, L) = (\alpha_i Y^\rho + (1 - \alpha_i)L^\rho)^{1/\rho} \tag{1}$$

with $\alpha_i \in [0, 1]$ and $\rho \in (-\infty, 1]$. Perceived income depends on leisure demand and tax salience: $Y = Y_i(L, \gamma_i)$, where $\gamma_i \in [0, \infty)$. The function Y_i varies with i because wages and tax circumstances (such as state of residence) vary by individual. The true budget constraint is the function $Y_i(\cdot, 0)$. Testing full tax salience is testing the hypothesis that γ_i is constant across individuals and equal to 0.

In general α_i and γ_i are allowed to vary by individual: $\alpha_i = F(\beta_0 + \beta_1 X_i^w + \varepsilon_i^w)$ with $F(x) = \exp(x)/(1 + \exp(x))$, and $\gamma_i = \exp(\delta_0 + \delta_1 X_i^s)$, where X^w and X^s are variables affecting taste for work and tax salience.[3] The unobserved heterogeneity term ε^w is assumed independent of X^w and X^s, and has distribution $N(0, \sigma^2_{\varepsilon^w})$.

Suppose agent i is unconstrained by his employer in his choice of hours worked. He demands leisure $L^*_{i,uncon}$ satisfying:

$$L^*_{i,uncon} \in \underset{L}{\operatorname{argmax}}\, U(Y_i(L, \gamma_i), L) \tag{2}$$

The bundle of consumption goods and leisure the worker ends up with is determined by his actual budget constraint: misperception does not change the worker's choice set. I nevertheless assume the consumer demands leisure *as if* he were maximizing utility subject to the perceived constraint $Y_i^{perc}(L, \gamma_i)$. Perhaps after repeated mistakes - that is, after repeatedly finding anticipated and actual post-tax income differ - people gradually correct their perceptions of their true tax liability. This is precisely why salience might be expected to increase with age. The results in Section 4.1 support this: as age increases, estimated perceived budget constraints more closely approximate the truth.

It is likely that some workers are constrained in their choice of labor supply. In particular, one might expect many employees to be required to work roughly 40 hour weeks, or 2000 hour years. I treat each worker as being so constrained with some probability and unconstrained otherwise. Specifically, each worker is restricted to working between 2025 and 2075 hours with probability ϕ_1, and between 2075 and 2125 hours with probability ϕ_2 (or since labor

[3] γ_i could also include an unobserved random effect, as in $\gamma_i = \exp(\delta_0 + \delta_1 X_i^s + \varepsilon_i^s)$, with $\varepsilon^s \sim N(0, \sigma^2_{\varepsilon^s})$. In practice, the maximum likelihood estimate of $\sigma^2_{\varepsilon^s}$ is 0. This is because some people's behaviour is very poorly explained by salient budget constraints. Incorporating a salience error term will greatly decrease these individuals' contributions to the likelihood function. It is likely that there is unobserved heterogeneity in salience, but it is not well captured by a simple homoskedastic random effect. Since unobserved salience heterogeneity is not the main question here, I avoid complicating the model further by incorporating a more realistic distribution for ε^s.

supply is discretized, working 2050 or 2100 hours). These intervals are natural choices given the observed data - as discussed in Section 3, they are the most frequently chosen labor supply levels in the sample. With probability $1 - \phi_1 - \phi_2$ actual hours worked, denoted L_i^*, is unconstrained and satisfies $L_i^* = L_{i,uncon}^*$.

I allow for classical measurement error on hours worked. Observed hours worked for i is $L_i = L_i^* + \varepsilon_i^m$, where $\varepsilon^m \sim N(0, \sigma_{\varepsilon^m}^2)$ is independent of ε^w and the covariates X^w, X^s. Another interpretation of the error is imperfect control over hours worked: although i would ideally consume L_i^* hours of leisure, he actually ends up consuming some perturbation of his optimal choice.[4]

2.2 Identification

The basic intuition behind the identification of the salience parameter γ was mentioned earlier - it explains why we do not observe clustering around kink points. This section develops this intuition in more detail.

One question is how to separately identify measurement error and tax salience. With measurement error, modes in the distribution of hours worked ought to be observed at kink-points, as in Figure 2b. But for large enough measurement error variance, these modes will be less pronounced, to the point of being difficult to discern in the data. In principle it might be large measurement error variance rather than low tax salience which explains the lack of bunching at kinks. However, high measurement error variance can only predict minimal clustering at the expense of predicting very large variance in the distribution of hours worked conditional on observables. This is not true of low tax salience, which can explain non-bunching without requiring such dispersion in hours worked. To the extent that non-bunching is not accompanied by such a large conditional variance in hours, this must be because of salience, not measurement error.

[4] Implicit in this interpretation of ε^M is some myopia on the part of the worker. A forward-looking worker would anticipate the perturbation and this would affect his choice of L_i^*.

The degree to which clustering at kink points changes with age identifies how the salience parameter γ changes with age. But perhaps it is incorrect to interpret this variation, and therefore the estimate of the salience parameter, as truly speaking to tax salience. It could be that older workers cluster around kinks to a greater degree than younger workers simply because they have more flexibility in choosing their hours. I present evidence in Section 4.2 which suggests that, at least for this sample, this is not the case.

Incorporating heterogeneity in tastes for work as well as measurement error (ε^w and ε^m) is standard in the literature on estimating labor supply behavior subject to nonlinear budget constraints (see MaCurdy et al. (1990) for more). The variability in observed hours worked induced by measurement error is constant with respect to wage changes, whereas that induced by heterogeneity in tastes for work is not. $\sigma^2_{\varepsilon^w}$ and $\sigma^2_{\varepsilon^m}$ are therefore identified by how cross-sectional dispersion in hours worked changes with wages.

2.3 Estimation

Estimating this model by full maximum likelihood would require obtaining the density of hours worked for each individual in the sample. These densities may be complicated - depending on the kinks and nonlinearities in the budget set, they may be a discrete-continuous mixture supported on not on a single interval but several intervals. Unlike other maximum likelihood models of labor supply with nonlinear budget constraints, this exercise would need to be carried out for multiple values of the smoothing parameter in the search for an optimum. As Blundell et al. (2007b) emphasize, discretizing labor supply avoids the technical difficulties involved in dealing with these densities. I therefore discretize the hours grid to make this problem tractable, and consider the 75 point annual hourly labor supply grid $\{300, 350, ..., 3950, 4000\}$, corresponding to the leisure demand grid $\{l_1 = 4760, 4810, ..., 8410, l_{75} = 8460\}$.

Given values of the parameters and covariates, the model predicts that i will be observed to choose each of $l_1, l_2, ...l_{75}$ with probabilities denoted $p^i_1, p^i_2, ..., p^i_{75}$. Letting $\Theta = (\rho, \beta_0, \beta_1, \delta_0, \delta_1)$, $X_i = X^w_i \cup X^s_i$ and $\varepsilon = (\varepsilon^w, \varepsilon^m)$, these probabilities are defined by:

$$p^i_k = \int_\varepsilon \mathbb{I}(i \text{ is observed to choose } l_k \mid \Theta, X_i, \varepsilon) dF(\varepsilon) \tag{3}$$

Denoting the closest grid point to actual observed leisure demand L_i by l_{k_i}, agent i's contribution to the likelihood function is $p^i_{k_i}$, and the log-likelihood function is $\sum_{i=1}^n \ln p^i_{k_i}$. The calculation of the integral in (3) is described in Appendix B.

2.4 Discussion

An alternative to a fully specified structural model is a regression analysis. A natural regression to run in this context is

$$Dist_i = \zeta_0 + \zeta_1 X_i + \epsilon_i \tag{4}$$

where $\mathbb{E}(\epsilon_i X) = 0$ and $Dist_i$ is i's distance in hours from the nearest kink in his budget constraint. If $Dist_i$ decreases with age this might appear to suggest that tax salience increases with age. There are unfortunately several problems with this approach. First, age is likely to be correlated with the unobservables ϵ_i. There tend to be more kinks in the part of the budget constraint associated with lower incomes, due to the Earned Income Tax Credit and changes in federal income tax rates. $Dist_i$ is therefore correlated with taste for work, which in turn changes with age. The above regression conflates changes in $Dist_i$ due to changes in work preferences with those due to changes in salience. In the structural model this is not the case: taste shifters change the relative weights on hours and consumption in the utility function, whereas the tax salience parameter governs the smoothness of the perceived budget constraint.

Second, it is unclear whether the distance from the nearest kink ought to be measured in terms of hours of labor supply or dollars of consumption. The consumer has preferences defined over both. As it stands, it seems difficult to interpret a coefficient from this regression in economic terms.

Third, implicit in $Dist_i$ is some definition of what a kink-point is. One possibility is to define any jump in the marginal income tax rate, no matter how small, as a kink. This regression would effectively weight all kinks equally, which is peculiar. We expect - and the structural model predicts - more clustering around sharper kinks than moderate ones, and that large jumps in tax rates are more informative about the true degree of tax salience than small jumps.

Fourth, specifying and estimating preferences allows the calculation of the welfare costs from less than complete tax salience. This would be hard to do using only the estimates from (4).

Although preferences and tax salience are allowed to vary with age, the structural model is still static in that it does not model saving. Regardless of how much of post-tax income is saved or consumed, only total post-tax income enters into the utility function. Incorporating dynamics here would likely complicate the model without delivering much additional insight into the issue of tax salience, so I omit this additional structure.

3 Data

I use the Survey of Income and Program Participation (SIPP), which has advantages over alternative datasets. Unlike individual tax return data, the SIPP allows for the calculation of total hourly labor supply, which is crucial to the analysis. It also yields a larger sample size than the Panel Study of Income Dynamics, and more accurate tax data than the Current Population Survey.

I draw data from three separate waves for the years 1997, 2002 and 2005. The data is a repeated cross-section. In addition to the core survey, the Tax Topical Module is available for these years. Together, this is enough information to construct each worker's budget constraint, that is, to specify his post-tax income for any given level of labor supply. I use the NBER's "TAXSIM" tax simulator for this calculation.[5] TAXSIM calculates federal and state income tax liabilities from survey data, given relevant information such as wage income, social security benefits, number of dependent exemptions, and itemized deductions. For each individual, I use TAXSIM to compute the total tax liability they would face given labor supply equal to each of the 75 values on the grid. This pins down 75 points on each individual's budget constraint.

To increase the credibility of the estimates, I restrict attention to a subsample which is relatively well-suited for this application. From the SIPP, I select hourly paid workers with exactly one job who are single for tax purposes, and who have a copy of their tax form to refer to when answering survey questions.[6] This leaves 1094 people. Salaried workers are omitted because their earned income is not directly linked to their hours worked, as the model I propose assumes. The labor supply of married workers is unlikely to be well described by considering each spouse's decision in isolation, so they are excluded from the sample.[7] Selecting only those who have a copy of their tax form at hand when answering questions reduces reporting error in responses to questions about tax exemptions and deductions. Also of note is that the SIPP only requires respondents to recall earnings over the last 4 months, unlike the biennial Panel Study of Income Dynamics. Any reporting error which remains is explicitly modelled as the random variable ε^m.

Focusing on this subsample is not meant to prejudge the importance of tax salience in other groups. The preferences and tax perceptions of this group may not be representative of the general population, and these results should be extrapolated only with caution. For example, the decision to participate in the workforce is left unmodelled here, and so estimated

[5]http://www.nber.org/~taxsim/

[6]There are also 2 people who sustain such large losses on their property income that their implied consumption is negative for some grid points of labor supply. Since the utility function is not defined for negative consumption, I drop them from the sample. Including them and rescaling consumption so that utility is defined does not change the main results.

[7]Blundell et al. (2005, 2007a) present results on the identification and estimation of collective labor supply models.

preferences and perceptions should not be taken as indicative of those of the unemployed. Hourly workers may have more job flexibility than salaried workers, and perhaps those with a copy of their tax forms at hand are more aware of their taxes.

The SIPP includes data on hourly wages, monthly earned income and hours worked. The survey validation literature indicates that reported hours worked are subject to significant measurement error, but that annual earnings are more accurately reported (Bound et al. (2001); Angrist and Krueger (1999); Rodgers et al. (1993)). It also seems reasonable that hourly paid workers have a good idea of how much their hourly wages are. I therefore calculate annual observed labor supply as annual earned income divided by wages, instead of using reported hours worked.[8] No simple specification of measurement error for income and wages gives a normal measurement error for their quotient, hours. The assumption of normal errors on hours is only intended as a parsimonious approximation to the truth. As noted in Section 4.1, normality of errors appears not to be crucial.

Table 1 shows summary statistics for the sample, and describes some of the covariates. Incomes are inflated to 2005 prices using the CPI. Hours worked are winsorized above and below at 4000 and 300 hours (this affects 9 and 16 observations). Figure 4 depicts the histogram of labor supply, each bar of which corresponds to a point on the grid used for estimation. 127 and 98 workers are located at the two highest bars, centered at 2050 and 2100 hours. It is at these labor supply levels that workers are assumed to be constrained to work with probabilities ϕ_1 and ϕ_2.

One possible concern is that although individuals have a clear idea of their tax schedules, unpredictable asset income hampers them from locating precisely at the kink-points on their budget constraints. However, the SIPP oversamples the relatively poor. This is reflected in

[8] An adjustment for overtime pay is in order here: I assume that hours worked over 2000 hours per year earn the federally mandated overtime rate for non-exempt workers of 1.5 times the normal wage. Different assumptions on overtime pay will affect calculated labor supply, but not earned income or taxes paid, or how close the worker is in dollar terms to a kink-point. The main results are robust with respect to these assumptions - ignoring overtime pay entirely, for example, I still find significant misperception of taxes which decreases with age.

	All Years, N = 1094		1997, N = 295		2002, N = 326		2005, N = 473	
	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Annual Wage Income	25,778	17,062	26,244	18,029	25,152	15,609	25,918	17,420
Hourly Wage	11.95	5.73	10.45	5.32	11.97	5.57	12.88	5.90
Annual Hours Worked	1833.14	638.76	1869.29	648.67	1798.81	612.49	1834.24	650.22
Age	40.38	15.56	38.80	15.21	39.70	15.15	41.83	15.96
Female	0.54	0.50	0.53	0.50	0.55	0.50	0.54	0.50
Disabled	0.06	0.25	0.06	0.23	0.06	0.25	0.07	0.26
High School	0.89	0.31	0.87	0.33	0.89	0.32	0.91	0.29
College	0.15	0.35	0.16	0.36	0.16	0.36	0.14	0.34
Itemize	0.21	0.41	0.20	0.40	0.21	0.41	0.21	0.41

Table 1: Summary Statistics, All Years

"Disabled" is a dummy variable for the respondent's claiming to have a work-limiting physical or mental condition. "High School" is a dummy variable for completion of high school. "College" is defined analogously. "Itemize" is a dummy variable for itemization of tax deductions.

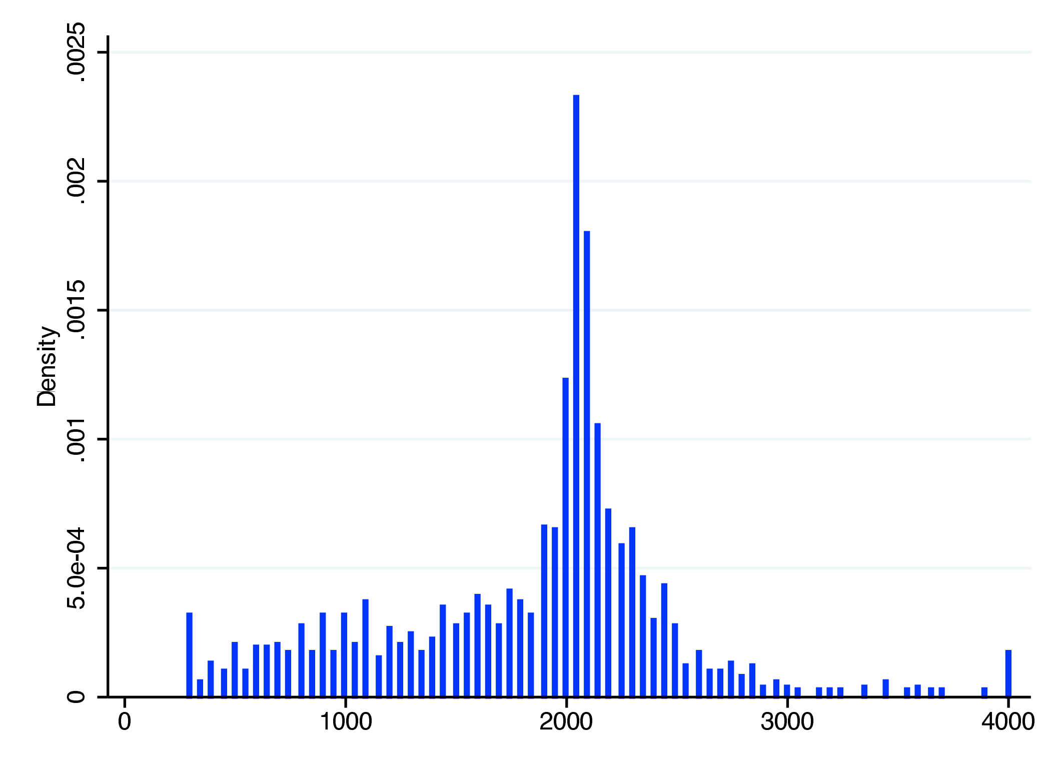


Figure 4: Histogram: Imputed Annual Hourly Labor Supply

below-average incomes in Table 1. Only 10.8% of the sample report owning any stocks. The fraction which realize taxable capital gains or losses in a given year is smaller still: only 2.4% respondents report non-zero capital gains in the Tax Topical Module. It seems unlikely that uncertainty over asset returns is driving the findings here.

4 Results

4.1 Model Estimates

Table 2 displays the estimates of tax perception and work preference parameters. The first column of estimates restricts the tax salience parameter γ_i to be constant across people and equal to 0. This is the special case of agents optimizing subject to perfectly known budget constraints. The second column restricts γ_i to be constant across people and possibly positive. The third column introduces heterogeneity: γ_i is modelled as $\exp(\delta_0 + \delta_1 X_i^S)$, so that it varies across people with the regressors X^s.

The salience shifters X^s include age (in logarithms so that γ_i moves linearly with age) and education, as well as dummy variables for sex and whether or not the respondent itemizes. The taste for work shifters X^w include year dummy variables, wages, age, age squared, and dummy variables for sex and disability. Including wages in X^w controls for the fact that higher wages may be correlated with higher unobserved tastes for work.

In all specifications the estimates of the coefficients on the taste for work shifters are of the expected sign. Positive values of $\hat{\mu}_{wages}$ indicate that those with higher wages also enjoy working more. Corresponding to the hump-shaped pattern of labor supply over the life-cycle, $\hat{\mu}_{age}$ and $\hat{\mu}_{agesq}$ are positive and negative. Women and those who report having a disability restricting their labor supply work fewer hours, reflected in the negative values of $\hat{\mu}_{fem}$ and $\hat{\mu}_{dis}$.

As tax salience specification becomes less restricted, the fit of the model improves appre-

N=1096	Full Salience	Constant Salience	Heterogeneous Salience
		Salience	
δ_{cons}	-	6.26	26.15
		(0.20)	(6.55)
$\delta_{ln(age)}$	-	-	-4.57
			(1.48)
δ_{fem}	-	-	-3.28
			(2.45)
$\delta_{HighSch}$	-	-	1.33
			(1.52)
$\delta_{College}$	-	-	-0.65
			(0.93)
$\delta_{Itemize}$	-	-	-0.84
			(1.53)
		Work Preferences	
ρ	-2.27	-0.38	-0.34
	(0.66)	(0.19)	(0.16)
μ_{1997}	-5.68	-2.92	-2.79
	(1.01)	(0.35)	(0.29)
μ_{2002}	-5.90	-3.03	-2.90
	(1.06)	(0.36)	(0.30)
μ_{2005}	-5.81	-2.98	-2.85
	(1.00)	(0.34)	(0.28)
$100 \cdot \mu_{wages}$	13.47	2.27	1.94
	(4.00)	(1.29)	(1.07)
μ_{age}	0.21	0.08	0.07
	(0.05)	(0.02)	(0.01)
$100 \cdot \mu_{agesq}$	-0.17	-0.06	-0.06
	(0.05)	(0.02)	(0.01)
μ_{fem}	-0.66	-0.25	-0.23
	(0.20)	(0.07)	(0.06)
μ_{dis}	-0.55	-0.18	-0.17
	(0.28)	(0.10)	(0.09)
		Error Variances and Hours Constraints	
σ_{ε^w}	1.58	0.67	0.61
	(0.30)	(0.09)	(0.07)
σ_{ε^m}	137.29	110.89	110.30
	(6.74)	(6.12)	(6.00)
ϕ_1	0.09	0.12	0.13
	(0.08)	(0.06)	(0.06)
ϕ_2	0.35	0.27	0.27
	(0.08)	(0.06)	(0.06)
Likelihood	-4082.84	-4076.32	-4068.33

Table 2: Estimation Results

ciably. A likelihood ratio test of full salience against constant salience ($\gamma = 0$ against $\gamma > 0$) rejects the null at the the 1% significance level.[9,10] The improved model fit is also reflected in the decrease in the standard errors of the unobserved error terms. $\hat{\sigma}_{\varepsilon^m}$ falls by 19% from the full to the heterogeneous salience cases. $\hat{\sigma}_{\varepsilon^w}$ falls by over 60%. The estimate of the constrained fraction of the population $\hat{\phi}_1 + \hat{\phi}_2$ also falls, although even in the heterogeneous salience case it remains quite large, at 40%. [11]

Chetty et al. (2009) emphasize that price elasticities depend on the salience of price changes. This is borne out in these data. The elasticity of substitution is $1/1-\rho$. Imposing the assumption of perfect knowledge of tax schedules results in underestimating the elasticity of substitution between labor and consumption. Intuitively this is because we do not observe pronounced behavioral responses when the marginal benefit to working jumps discretely, as at kinks in the budget constraint. The implied elasticities when the restriction of perfect knowledge is lifted are over twice as large: 0.72 and 0.75 with constant and heterogeneous salience, compared with 0.31 in the full salience case. Elasticities with respect to perceived price changes are substantially different from those with respect to actual price changes.

An increase in a variable associated with a negative $\hat{\delta}$ coefficient decreases the estimated smoothing parameter $\hat{\gamma} = \exp(\hat{\delta}_0 + \hat{\delta}_1 X^s)$ and increases salience. Table 2 shows that tax salience increases with age, and that the increase is statistically significant.[12] The welfare calculations in Section 4.3 suggest that this effect is small but not negligible. The other variables in X^s do not affect salience in a statistically significant way. It may seem surprising

[9] Since the null hypothesis is on the boundary of the parameter space the LR test statistic has a limiting distribution which is 0 with probability .5 and χ^2_1 with probability .5. Self and Liang (1987) derive this result.

[10] Following the construction of the perceived budget constraint described in Appendix A, the estimated constant salience parameter of 6.26 corresponds to a bandwidth of $\exp(6.26) = 523.22$ hours, indicating substantial smoothing.

[11] The estimate $\hat{\phi}_1$ is less than $\hat{\phi}_2$, although there are more people observed at a labor supply of 2050 hours than 2100 hours. Because of measurement error, the estimates $\hat{\phi}_1, \hat{\phi}_2$ are driven not just by how many people are observed to located exactly at the corresponding grid point, but also by how many are observed to locate nearby. These estimates indicate that measurement error may have a larger role to play in explaining observations above the mode than below it.

[12] 46 coefficients are estimated in Table 2. To account for the multiple comparisons problem I use the Bonferroni correction of testing hypotheses at the 5%/46 level. For each age and age squared estimate, the null hypothesis of no effect can be rejected.

that a college education is not associated with significantly more accurate tax perceptions. However this sample consists only of hourly-paid workers, and the college-educated who work in such jobs are unlikely to be representative of the college-educated at large.

These results are robust to different distributional assumptions on the error terms. Allowing for longer tails by treating ε^m and ε^w as logistically distributed, or incorporating asymmetry around the mean as in a Gumbel distribution do not change the main finding that tax salience is less than perfect, but increases with age.

This is not panel data, and a natural concern is that the results may be attributable to selection. Young and single, and old and single hourly-paid workers may differ in some unobservable respects. But is rather unclear what kind of selection could plausibly be responsible for the finding of salience increasing with age. If high skilled young workers initially work hourly-paid jobs before later transitioning to salaried positions, and if these workers are more aware of their taxes than those who remain in hourly-paid jobs, this would attenuate the increase in salience with age, and imply that the estimates presented understate the true effect. If selection were to explain these results, it would have to be part of a more complicated story.

It is instructive to compare these results to Saez (2010). Saez finds evidence of bunching only around the first kink of the Earned Income Tax Credit and where income tax liability starts. The increased tendency of the elderly to bunch may be less apparent when pooling data over all ages. Moreover, his data are a stratified random sample of the entire population of U.S. taxpayers, and so include groups omitted from this analysis. In particular, hourly workers may have more flexibility in labor supply decisions than salaried workers, so aggregating over both categories could reduce observed bunching relative to hourly workers alone.

The assumption that measurement error is homoskedastic is important. If the variance of measurement error changes with age, for example, the model estimates may falsely attribute this change in the measurement error distribution to a change in tax salience. Accommodating more general forms of measurement error would likely require a somewhat different identification strategy, or richer data.

4.2 Job Flexibility and Age

As mentioned in the discussion of identification, older workers might have more latitude in choosing their hours than younger workers. The tendency of the old to locate closer to kinks might be attributable not to salience changing with age, but to job flexibility changing with age. This appears not to be the case in this sample. The SIPP Work Schedule Topical Module contains information on the kind of shift worked, e.g. "Regular daytime schedule", "Split shift" and "Irregular schedule". It seems reasonable that those who work regular shifts are likely to be those with the least flexibility over their hours. Working an irregular shift is thus a rough proxy for job flexibility. I regress a dummy variable for working an irregular shift on age dummy variables and other controls. Table 3 shows the results. Workers in their teens (the omitted age category) and twenties are the most likely to have irregular shifts; those in their fifties are the least likely. This suggests that it is indeed tax salience which explains the behavior of older workers, rather than greater job flexibility.

The SIPP also asks why respondents work the shifts they work. Responses to this latter variable fall into two broad categories: voluntary reasons, such as better pay, child care arrangements or allowing time for school, and involuntary reasons, such as the shift being a requirement of the job, or the respondent's not being able to get any other job. It seems plausible that those who work the shifts they do for involuntary reasons would have less job flexibility on average than other workers. If job flexibility increases with age, older workers should be less likely to have their work schedules determined by involuntary reasons. Again, this is not supported by the data. Table 3 shows that it is workers in their fifties who are most likely to work the hours they do for involuntary reasons.

The two proxies for job flexibility used above are admittedly imperfect. They measure

N=1094	100·(Irregular Shift)	100·(Involuntary Reasons)
	Year	
1997	31.49	37.63
	(4.28)	(5.68)
2002	34.19	44.95
	(4.32)	(5.73)
2005	33.62	45.57
	(4.24)	(5.62)
	Age	
20 - 29	−16.02	39.77
	(4.43)	(5.88)
30 - 39	−24.64	49.02
	(4.66)	(6.18)
40 - 49	−23.63	49.44
	(4.60)	(6.11)
50 - 59	−26.01	49.94
	(4.59)	(6.09)
60 - 69	−23.77	42.89
	(5.05)	(6.70)
Over 70	−19.92	46.73
	(6.65)	(8.82)
	Other	
Female	2.84	−3.38
	(2.00)	(2.65)
Disabled	3.12	−2.83
	(3.99)	(5.29)
High School	−2.72	−11.31
	(3.49)	(4.63)
College	2.92	3.56
	(2.81)	(3.73)
Itemize	−1.04	−4.43
	(2.51)	(3.33)
R^2	0.17	0.76

Table 3: Job Flexibility

job flexibility only with some error. However, to overturn these regression results, this error would have to be robustly correlated with age. It is not clear why such a correlation would obtain, and so it seems that job flexibility is not a major confounding factor in interpreting the estimation results.

4.3 Importance of Tax Misperceptions

One way to assess the importance of imperfect salience is to calculate how much labor supply would change if the budget constraint were perfectly known. I compute predicted labor supply given full tax salience and given heterogeneous tax salience (the first and third columns of Table 2). In this exercise I hold the unobserved tastes for work error ε^w constant and equal to 0, and initially assume no constraints in the choice of hours. The difference between the predicted labor supply in these two cases is relatively modest, as Table 4 shows. Estimated salience labor supply exceeds full salience labor supply by only 12 hours a year. If hours constraints are taken into account the costs are smaller still, since in the presence of hours constraints labor supply will not vary with tax salience. Many people would choose the same labor supply level under full and estimated salience: even if they were free to choose the hours they wished, 37% of the sample would not change labor supply at all (to the nearest grid point).

I also calculate a compensating variation style measure CV_i of the welfare change for individual i.[13] Let $(Y_{Subopt,i}, L_{Subopt,i})$ be i's actual income and labor supply bundle, and let $(Y_{Opt,i}, L_{Opt,i})$ be his optimal income and labor supply bundle, when he is maximizing with respect to his true budget constraint. Then define CV_i implicitly as:

$$U_i(Y_{Subopt,i} + CV_i, L_{Subopt,i}) = U(Y_{Opt,i}, L_{Opt,i})$$

[13] This welfare analysis can be viewed as a special case of Bernheim and Rangel (2009), in which choices are frame-dependent. The welfare-relevant domain in this context consists of those choices made when attention is explicitly drawn to the shape of the tax function, and it excludes choices made when attention is not directed in that manner.

	Mean	Std. Dev.
Labor Supply Difference	-7.18	62.26
Labor Supply Difference, No Constraints	-11.97	103.76
Labor Supply Difference, Absolute Value	39.65	48.52
Labor Supply Difference, No Constraints, Absolute Value	66.09	80.86
Welfare Loss	29.62	82.92
Welfare Loss, No Constraints	49.36	138.20

Table 4: Labor Supply Changes and Welfare Losses

"Labor Supply Difference" is predicted labor supply given full tax salience minus predicted labor supply given estimated tax salience. "Welfare loss" is the compensating variation in dollars, as described above. "No Constraints" disregards the fact that each worker is constrained to work a set amount of hours with the probabilities ϕ_1, ϕ_2, and instead treats him as free to choose any labor supply level.

that is, CV_i is the amount that actual consumption would need to be increased in order to make i as well off as he would be if he optimized with respect to the correct budget constraint.

Like the change in hours, the average welfare loss implied by tax misperceptions is quite small. Table 4 displays the results. Even ignoring the estimated hours constraints, it is only $49 a year. The standard deviation of the welfare losses and the change in hours are quite large relative to their means. Large welfare losses and hours changes correspond to those people who under full salience should locate on kink-points of their budget constraints far away from the labor supply they actually choose. Overall there are few of these cases, and even the largest welfare loss in the sample is under $2000 dollars.

The estimates from the structural model suggest that the welfare costs of misperception will decrease with age. To confirm this I regress the welfare loss accounting for hours constraints on age dummy variables. Table 5 shows the results. The overall trend is of welfare costs decreasing with age, and even for younger workers they are not too substantial.[14]

These welfare costs calculations depend on how inaccurate beliefs about taxes are modelled. In this model the perceived budget constraint approximates the real budget constraint at least as well as the best fitting line through the real budget constraint. If people greatly or systematically misjudge their tax liability, the costs could be much larger. Nevertheless

[14]Other regression specifications support this result. For example, including linear and quadratic controls for age instead of decade indicators, and controlling for other demographics, gives similar results.

N=1094	Welfare Loss	Number of Observations
Under 20	18.32 (9.18)	81
20 - 29	36.78 (4.98)	275
30 - 39	43.52 (6.18)	179
40 - 49	29.26 (5.74)	207
50 - 59	21.64 (5.66)	213
60 - 69	15.60 (8.14)	103
Over 70	20.56 (13.77)	36
$R^2 = 0.12$		

Table 5: Welfare Costs and Age

these estimates show that non-bunching due to tax misperception does not necessarily imply substantial welfare losses. Chetty (2010) reaches similar conclusions with an altogether different approach: he specifies a dynamic lifecycle model with optimization frictions and finds that welfare costs of under 1% of consumption can explain labor supply elasticity estimates from 20 studies.

5 Conclusion

Specifying and estimating a model of tax salience, I find evidence that in general people do not correctly perceive their taxes, but that the accuracy of tax perceptions increases significantly with age. The welfare costs of this misperception seem to be quite small even for younger workers, and decrease further with age. It appears that imperfect awareness of tax liabilities may explain the surprisingly little bunching in labor supply data without requiring that workers suffer large welfare losses relative to the full salience case. Further, compared to imposing full tax salience, allowing for tax misperceptions yields substantially larger estimates of the elasticity of substitution between labor and consumption.

Data limitations necessitate some pragmatic modelling compromises. It is unknown exactly how flexible each worker's job is, so I assume some fraction must work standard 40 hour weeks plus a small amount of overtime. The data would not identify arbitrary forms of tax misperception, so I assume tax misperceptions have a particular, simple form. Calculating welfare costs and elasticities under different assumptions about the nature of tax misperceptions is left as a topic for future research.

Appendix

Appendix A: The Perceived Budget Constraint

Let individual i's true budget constraint be $Y = Y_i(H)$, where Y is income and H is hours worked. Let $\hat{Y}_i$ be the least squares approximation to the function Y_i. Denote by $\ker(Y(H) - \hat{Y}(H), \gamma)$ the fitted value at H from a Nadaraya-Watson kernel regression of $Y - \hat{Y}$ on hours with a standard normal kernel and bandwidth γ. Define the parameterized, or perceived, budget constraint $Y_i^{perc} : [0, \bar{H}] \times (0, \infty) \to \mathbb{R}$ by

$$Y_i^{perc}(H, \gamma) = \hat{Y}_i(H) + \ker(Y(H) - \hat{Y}(H), \gamma)$$

where $\bar{H}$ is the maximum possible hourly annual labor supply (taken to be 4000 in the model). Since $\lim_{\gamma \to 0} \ker(Y(H) - \hat{Y}(H), \gamma) = 0$ pointwise, $\lim_{\gamma \to 0} Y_i^{perc}(H, \gamma) = Y_i(H)$ pointwise. Similarly, since $\lim_{\gamma \to \infty} \ker(Y(H) - \hat{Y}(H), \gamma) = Y(H) - \hat{Y}(H)$ pointwise, $\lim_{\gamma \to \infty} Y_i^{perc}(H, \gamma) = Y_i(H)$ pointwise. Thus, when the smoothing parameter γ is small, the smoothed budget constraint is close to the real budget constraint. When γ is large, it is close to the least squares line through the real constraint.

Appendix B: Calculation of the Likelihood Function

Let p_k^i and p_k^{*i} denote the probabilities that i is observed to choose grid point l_k and that i actually chooses l_k (either because i desires to choose l_k, or because i is constrained to choose l_k). p_k^{*i} can be written as

$$p_k^{*i} = \int_{-\infty}^{+\infty} \mathbb{I}(i \text{ chooses } k \mid \Theta, X_i, \varepsilon^w) dF(\varepsilon^w) \tag{5}$$

Measurement error ε^m does not feature here since it only affects the observed choice, not what is optimal for i. Write i's perceived incomes associated with leisure demand levels $l_1, l_2, ..., l_{75}$ as $y_1^i, y_2^i, ..., y_{75}^i$. Since these are *perceived* incomes, they vary with salience and thus with (Θ, X_i), although this dependence is suppressed in the notation. Given (Θ, X_i), some algebra shows that agent i's desired leisure grid point will be l_k if and only if

$$\forall k' \neq k,\ (\alpha y_k^\rho + (1-\alpha) l_k^\rho)^{1/\rho} > (\alpha y_{k'}^\rho + (1-\alpha) l_{k'}^\rho)^{1/\rho}$$

$$\iff \quad \begin{cases} \varepsilon^w < \log\left(\frac{l_{k'}^\rho - l_k^\rho}{y_k^\rho - y_{k'}^\rho}\right) - \beta_0 - \beta_1 X_i^w, & \forall k' > k \\ \varepsilon^w > \log\left(\frac{l_{k'}^\rho - l_k^\rho}{y_k^\rho - y_{k'}^\rho}\right) - \beta_0 - \beta_1 X_i^w, & \forall k' < k \end{cases}$$

i.e. ε_i^w must fall within the above bounds in order for i to desire l_k. Denote these bounds as

$$\begin{aligned} \bar{\varepsilon}^w(k, \Theta, X_i, \varepsilon^s) &= \max_{k'>k} \log\left(\frac{l_{k'}^\rho - l_k^\rho}{y_k^\rho - y_{k'}^\rho}\right) - \beta_0 - \beta_1 X_i^w \\ \underline{\varepsilon}^w(k, \Theta, X_i, \varepsilon^s) &= \min_{k'<k} \log\left(\frac{l_{k'}^\rho - l_k^\rho}{y_k^\rho - y_{k'}^\rho}\right) - \beta_0 - \beta_1 X_i^w \end{aligned}$$

This implies that $p_k^{*i} = (1 - \phi_1 - \phi_2) \cdot \max\left\{0, \Phi\left(\frac{\bar{\varepsilon}^w(k,\Theta,X_i)}{\sigma_{\varepsilon^w}}\right) - \Phi\left(\frac{\underline{\varepsilon}^w(k,\Theta,X_i)}{\sigma_{\varepsilon^w}}\right)\right\} + \phi_1 \cdot \mathbb{P}(l_k = 2050) + \phi_2 \cdot \mathbb{P}(l_k = 2100)$. The first term is the probability that the employee is unconstrained and desires to work l_k. The second and third terms are the probabilities that the employee works l_k because he is constrained to do so. The p_k^i's are the p_k^{*i}'s convolved with measurement error:

$$p_k^i = \sum_{j=1}^{75} p_j^{*i} \mathbb{P}_{\sigma^2_{\varepsilon^m}} (\tilde{\varepsilon}^m = 50 \mid j - k \mid)$$

where $\tilde{\varepsilon}^m$ is distributed as the discretized analogue of $N(0, \sigma^2_{\varepsilon^m})$.

References

Angrist, Joshua D. and Alan B. Krueger, "Empirical Strategies in Labor Economics," in Orley C. Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 3, Elsevier, 1999, chapter 23, pp. 1277–1366.

Bartolome, Charles A.M. de, "Which Tax Rate Do People Use: Average or Marginal?," *Journal of Public Economics*, 1995, *56* (1), 79–96.

Bernheim, B. Douglas and Antonio Rangel, "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics," *Quarterly Journal of Economics*, 2009, *124* (1), 51–104.

Blundell, R, H Hoynes, and National Bureau of Economic Research, "Has In-Work Benefit Reform Helped the Labour Market?," in David Card, Richard Blundell, and Richard Freeman, eds., *Seeking a Premier Economy: The Economic Effects of British Economic Reforms, 1980-2000*, Vol. 8546, University of Chicago Press, 2001.

Blundell, Richard, Pierre-André Chiappori, and Costas Meghir, "Collective Labor Supply with Children," *Journal of Political Economy*, 2005, *113* (6), 1277–1306.

_ , _ , Thierry Magnac, and Costas Meghir, "Collective Labour Supply: Heterogeneity and Nonparticipation," *Review of Economic Studies*, Apr 2007, *74* (2), 417–445.

— , **Thomas MaCurdy, and Costas Meghir**, “Labor Supply Models: Unobserved Heterogeneity, Nonparticipation and Dynamics,” in James Heckman and Edward Leamer, eds., *Handbook of Econometrics*, Vol. 6, Elsevier, 2007, chapter 69, pp. 4667–4775.

Bound, John, Charles Brown, and Nancy Mathiowetz, “Measurement Error in Survey Data,” in “Handbook of Econometrics,” Vol. 5 2001, chapter 59, pp. 3705–3843.

Burtless, Gary and Jerry A. Hausman, “The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experiment,” *Journal of Political Economy*, 1978, *86* (6), 1103–1130.

Cabral, Marika and Caroline Hoxby, “The Hated Property Tax: Salience, Tax Rates, and Tax Revolts,” *Working Paper*, 2010.

Chetty, Raj, “Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply,” *Working Paper*, 2010.

— , **Adam Looney, and Kory Kroft**, “Salience and Taxation: Theory and Evidence,” *American Economic Review*, 2009, *99* (4), 1145–1177.

— , **John Friedman, Tore Olsen, and Luigi Pistaferri**, “Adjustment Costs, Firm Responses, and Labor Supply Elasticities: Evidence from Danish Tax Records,” *Quarterly Journal of Economics*, 2011.

Finkelstein, Amy, “E-ZTax: Tax Salience and Tax Rates,” *Quarterly Journal of Economics*, 2009, *124* (3), 969–1010.

Friedberg, Leora, “The Labor Supply Effects of the Social Security Earnings Test,” *Review of Economics and Statistics*, 2000, *82* (1), 48–63.

Fujii, Edwin T. and Clifford B. Hawley, “On the Accuracy of Tax Perceptions,” *Review of Economics and Statistics*, 1988, *70* (2), 344–347.

Liebman, Jeffrey B. and Richard J. Zeckhauser, "Schmeduling," *Working Paper*, 2004.

MaCurdy, Thomas, David Green, and Harry Paarsch, "Assessing Empirical Approaches for Analyzing Taxes and Labor Supply," *Journal of Human Resources*, 1990, *25* (3), 415–490.

Rodgers, Willard L., Charles Brown, and Greg J. Duncan, "Errors in Survey Reports of Earnings, Hours Worked, and Hourly Wages," *Journal of the American Statistical Association*, 1993, *88* (424), 1208–1218.

Rosen, Harvey S., "Taxes in a labor supply model with joint wage-hours determination," *Econometrica*, 1976, *44* (3), 485–507.

Saez, Emmanuel, "Do Taxpayers Bunch at Kink Points?," *American Economic Journal: Economic Policy*, 2010, *2* (3), 180–212.

Self, Stephen G. and Kung-Yee Liang, "Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests under Nonstandard Conditions," *Journal of the American Statistical Association*, 1987, *82* (398), 605–610.

Slemrod, Joel, "The Role of Misconceptions in Support for Regressive Tax Reform," *National Tax Journal*, 2006, *59* (1), 57.