The Effects of Progressive Taxation on Labor Supply when Hours and Wages Are Jointly Determined

Daniel Aaronson
Eric French

ABSTRACT

This paper extends a standard intertemporal labor supply model to account for progressive taxation as well as the joint determination of hourly wages and hours worked. We show that these two factors can have implications for both estimating labor supply elasticities as well as for using these elasticities in tax analysis. Failure to account for wage-hours ties and progressive taxation may cause the hours response to marginal tax rate changes to be understated by 5 to 30 percent for men.

I. Introduction

Many recent studies find larger income responses to specific tax changes than what would be expected by inferring this parameter from the labor supply response to wage changes, as has been traditionally done (Feldstein 1995). Some research has attributed this difference to tax avoidance and retiming and reshifting of transactions.[1] Alternatively, it could be that the standard labor supply models are misspecified,[2] causing the labor supply response to tax changes to be understated. This paper provides new results on two misspecification problems. In particular, we present a traditional life-cycle labor supply model augmented to include two features: progressive taxation and the joint determination of hours and wages.

1. See Slemrod (1998) for a useful nontechnical summary and discussion of the literature.
2. Mulligan (1999) provides reasons why the standard model might be misspecified.

Dan Aaronson and Eric French are senior researchers in economic research at the Federal Reserve Bank of Chicago. The data used in this article can be obtained beginning October 2009 through September 2012 from Dan Aaronson at daaronson@frbchi.org or Eric French at efrench@frbchi.org. The views of the authors do not necessarily reflect the Federal Reserve Bank of Chicago or the Federal Reserve System.

[July 2004; accepted November 2007]
ISSN 022 166X E ISSN 1548 8004 8 2009 by the Board of Regents of the University of Wisconsin System

Hours and wages may be jointly determined. Not only may increases in the wage lead to increases in hours worked, but a worker who cuts his work hours receives a lower offered hourly wage. This suggests that firms may not be indifferent between a single 40-hour-per-week worker and two workers each at 20 hours per week. We show that failure to account for tied wage-hours can understate the labor supply response to a change in the marginal tax rate.

To see why tied wage-hours offers amplify the labor supply effect, note that the central impact of a tax cut is to increase the posttax wage, which potentially boosts hours worked. These extra hours of work further increase the pretax wage, causing a larger posttax wage gain (and thus labor supply response) than what is caused by just the tax change in isolation.

We also augment the model to allow for progressive taxation. Progressive taxation causes less dispersion in the posttax wage (what individuals are presumably responding to) than the pretax wage. Using the pretax wage rather than the posttax wage biases the labor supply elasticity to zero. Roughly speaking, this is a variant of the errors-in-variables problem.

We quantify these effects two different ways. First, we analytically evaluate the labor supply response to a tax change using a range of relevant parameter values for the labor supply response to a wage change, the tied wage-hours relationship, and the progressivity of the labor income tax schedule. Reasonable values for the wage-hours tie results in a difference of at least 6 and possibly 25 percent for men.

Second, we estimate the intertemporal elasticity of substitution using common instrumental variables strategies. We then show how to apply the intertemporal elasticity of substitution to infer the labor supply response to a change in marginal tax rates. Using the Panel Study of Income Dynamics (PSID) and the March supplements of the Current Population Survey (CPS), we compare labor supply responses to tax changes that account for tied wage-hours and progressivity with those that do not and find the resulting bias can be at least 5 percent, and perhaps as high as 30 percent, for men. Therefore, we conclude that failure to account for tied wage-hours offers and progressive taxation can lead to understated labor supply responses to changes in marginal tax rates.

The important caveat to the analysis is that there is considerable disagreement on the effect of hours upon wages. This is in addition to the well-known disagreement about the effect of wages on hours. Nevertheless, our estimates suggest that accounting for progressive taxation and the joint determination of hours and wages may be important for understanding labor supply.

II. Tied Wage-Hours Offers

The labor supply estimates in this paper build on the finding that individuals who cut work hours receive a lower offered hourly wage. Therefore, we begin with a discussion of why wages might be tied to hours worked and what the empirical literature has found.

Firms may not be indifferent to the number of hours worked when there are fixed costs involved in hiring and retaining workers, as in Lewis (1969) and Barzel (1973).[3] This could include the cost of training and aspects of compensation unrelated

3. Other explanations for hours-wage ties are discussed in Ermisch and Wright (1993) and Hirsch (2005).

to hours worked. In a competitive labor market, the worker pays 100 percent of the fixed cost of work in the form of lower wages. As hours increase, these fixed costs can be spread over more hours of work, leading to increases in the measured hourly wage. Appendix 1 describes how a structural model of the fixed costs of work can be reasonably approximated by the simple regression that is commonly estimated in the literature:

$$(1) \quad \log w_{it} = \alpha_{it} + \theta \log h_{it}$$

where α_{it} represents an individual's underlying productivity during a specific year and θ maps log hours worked, $\log h_{it}$, into the log wage, $\log w_{it}$.

Estimates of Equation 1 usually attempt to addresses the endogeneity of hours worked using variation in hours unrelated to α_{it}. The most common such approach is to make use of the number of children, as in Rosen (1976), Moffitt (1984), Blank (1990), and Ermisch and Wright (1993). Arguably, the number of children does not directly affect the productivity parameter α_{it} but does affect the marginal utility of leisure and thus hours worked. Other studies use variation in hours worked caused by the Social Security rules (Aaronson and French 2004), family structure, nonlabor income, and disability status (Biddle and Zarkin 1989), or cross-equation VAR restrictions (Lundberg 1985).

These papers tend to find some causal evidence that a reduction in hours reduces hourly wages. However, there remains considerable disagreement of the magnitude of θ, with estimates ranging from 0.1 or less (Lundberg 1985; Hirsch 2005) to about 0.2 (Moffitt 1984; Keane and Wolpin 2001) to 0.4 (Rosen 1976; Lettau 1997; Aaronson and French 2004) to one or more (Biddle and Zarkin 1989). Our reading is that a reasonable estimate for prime age males is around 0.4, which implies that moving from 40 to 20 hours of work will reduce the offered hourly wage by 25 percent. Appendix 1 shows that this estimate is consistent with the values of θ that are generated by a structural model, assuming 28 percent of firms' labor costs are fixed, as in Malcomson (1999). However, we fully acknowledge the wide range of estimates in the literature, and because of concern that the previous empirical approaches may not capture the true wage response to hours changes caused by tax changes, we will consider a range of estimates of θ in the analysis below.

But before turning to that, we briefly consider three other potential concerns with using Equation 1 in a model of dynamic labor supply.

First, it assumes a log linear relationship between hours and wages. Barzel (1973), for example, presents a case where nonlinearities are important. In his model, workers tire above a certain workweek length, causing hourly productivity to decline. The empirical evidence for a nonlinear relationship between hours and wages is thin and mixed (for example, Moffitt 1984; Biddle and Zarkin 1989; Aaronson and French 2004). Furthermore, Appendix 1 shows that while a structural model of fixed costs imply nonlinearities in the hours-wage relationship, these nonlinearities are not economically important.

Second, the hours-wage tie may not exist within a job, but may exist when moving across jobs. This issue has received little attention in either the theoretical or empirical literature. However, Aaronson and French (2004) and Lettau (1997) find that workers who cut their hours within a given employment relationship receive wage reductions. Thus, there is some evidence of a relationship between hours and wages in the short run (that is, within a given employment relationship).

Finally, there may be concern that the estimated elasticities are not structural if a large decline in hours causes wage losses, but a small drop in hours does not. This is a difficult hypothesis to reject since the estimates in the literature are based on an average hours change or a part-time versus full-time effect. We can, however, comment on our own work. Aaronson and French (2004) take advantage of the variation in hours worked caused by the Social Security rules at the ages of 62 and 65 to identify wage declines that correspond to cuts in hours at these exact ages. We find little evidence of a large discontinuity on hours changes at those ages.[4] Moreover, when we restrict our March CPS analysis by excluding observations in which annual hours at least double or halve, the results, while clearly less precisely estimated, cannot be statistically distinguished from the full sample estimates.[5]

III. Intertemporal Labor Supply Elasticities with Tied Wage-Hours Offers and Progressive Taxation

A. Model

We begin with the canonical intertemporal labor supply model,[6] as in MaCurdy (1985), augmented to account for tied wage-hours offers and a potentially progressive labor income tax schedule. Preferences take the form:

$$(2)\quad U = E_0 \sum_{t=1}^{T} \beta^t \left(v(c_{it}) - \exp(-\varepsilon_{it}/\sigma) \times \frac{h_{it}^{1+\frac{1}{\sigma}}}{1+\frac{1}{\sigma}} \right)$$

where U is the expected discounted present value of lifetime utility, c_{it} is consumption, $v(.)$ is some increasing concave function, h_{it} is hours worked, and ε_{it} is the person and year specific preference for work.[7] The parameter σ is the intertemporal elasticity of substitution, the usual object of interest in dynamic labor supply studies.

The individual faces the dynamic budget constraint:

$$(3)\quad A_{it+1} = (1 + r_t(1 - \tau_A))(A_{it} + w_{it}h_{it} + y_{it} - \tau_{it} - c_{it})$$

4. For example, among outgoing rotation group respondents from the March Current Population Survey, the 10th percentile of hours changes from ages 60 to 67 are -12, -13, -16, -15, -15, -17, -15, and -15. The 20th percentile of hours changes at the same ages are -7, -8, -8, -8, -8, -10, -8, and -8. At the other end of the hours change distribution, there is virtually no change between ages 60 to 67.

5. Our estimated θ goes up from 0.39 (with a standard error of 0.25) in the full sample to 0.77 (0.54) with the restricted hours change sample when using a continuous hours measure and 0.27 (0.15) to 0.38 (0.22) when using a part-time threshold based on a 35 hour, 50 week work year.

6. The key results from this section do not depend on whether the model is static or dynamic. However, the intertemporal model simplifies the analysis because it allows us to focus more on the substitution effect of a tax change. In static models and models with liquidity constraints, tax changes cause an additional change in the marginal utility of wealth. Moreover, if individuals do make forward looking decisions, many measures of nonlabor income that are used in static models are endogenous and inconsistent estimates will result.

7. Nonseparable preferences between consumption and hours worked is a strong assumption. See Browning and Meghir (1991) for evidence.

where A_{it} is time t assets, r_t is the interest rate, τ_A is the tax rate on capital income, y_{it} is spousal income, and τ_{it} is labor income taxes, which potentially depend on individual characteristics (such as the number of children):[8]

$$\tau_{it} = \tau_{it}(w_{it}h_{it} + y_{it}). \tag{4}$$

Maximization of Equation 2 subject to the dynamic budget constraint Equation 3 yields the labor supply function:

$$\log h_{it} = \sigma[\log(1-\tau'_{it}) + \log w_{it} + \log(1+\theta)] + \sigma \log \lambda_{it} + \varepsilon_{it}. \tag{5}$$

The term in square brackets is the logarithm of the opportunity cost of time, and is composed of three parts. First, $(1-\tau'_{it})$ reflects the cost of taxation that arises from additional working hours and is sometimes referred to as the log of the "net of tax price." Note that τ'_{it} is the marginal tax rate and thus $1-\tau'_{it}$ is the share of labor income that the individual keeps at the margin. The second part is the wage. The third part arises because the worker is paid a higher hourly wage when he works more hours, if hours and wages are tied. If changes in hours of work impact neither the wage (that is, $\theta = 0$) nor the amount of taxes paid (that is, $\tau'_{it} = 0$), Equation 5 becomes the standard estimating equation in intertemporal labor supply models. The term $\lambda_{it} \equiv v'(c_{it})$ represents the marginal utility of wealth.

To estimate σ, we first difference Equation 5:

$$\Delta \log h_{it} = \sigma[\Delta \log(1-\tau'_{it}) + \Delta \log w_{it}] + \sigma \Delta \log \lambda_{it} + \Delta\varepsilon_{it}. \tag{6}$$

Following MaCurdy (1985), Appendix 2 shows that the marginal utility of wealth follows a random walk with drift, allowing us to rewrite Equation 6 as:

$$\begin{aligned} \Delta \log h_{it} = {} & \sigma[\Delta \log(1-\tau'_{it}) + \Delta \log w_{it}] - \sigma \log \beta(1 + r_{t-1}(1-\tau_A)) \\ & + \sigma \frac{\beta(1 + r_{t-1}(1-\tau_A))\varepsilon_{it}}{\lambda_{it-1}} + \Delta\varepsilon_{it}. \end{aligned} \tag{7}$$

where ε_{it} is the innovation to the marginal utility of wealth. Equation 7 is our estimating equation. In that equation, failure to account for progressive taxation likely leads to inconsistent estimates of σ. On the other hand, the parameters relating to the hours-wage tie have been differenced out, meaning we do not need to explicitly account for hours-wage ties in the estimation.[9]

8. This analysis looks at anticipated changes in tax rates. If a tax change is unanticipated, we must consider both movements along and "parametric shifts" (e.g. MaCurdy 1985) in the lifecycle wage profile. Furthermore, we assume that capital income does not affect labor income tax rates, which simplifies the analysis (Blomquist 1985) but is problematic in that interest and dividends are taxed like ordinary income. Capital gains were taxed like ordinary income prior to 1997 and are still taxed that way for investments held less than one year. For long-term investments, there are currently two marginal rates. However, if capital gains are primarily concentrated among higher income households (see Burman and Ricoy (1997) for evidence), these rates could be considered significantly more proportional in practice than labor income. For tractability and due to limitations in the data, we therefore ignore these aspects of the progressive tax schedule. Moreover, since the PSID analysis stops in 1996, prior to the capital gains changes, and our CPS analysis is unaffected by limiting the data to pre-1997, we do not believe this has any practical impact on our results.

9. However, hours-wage ties imply that the wage and ε_{it} are correlated, meaning that we must instrument for the wage.

The remainder of this paper examines two general questions: how to obtain consistent estimates of σ and how to use σ to infer the labor supply response to a tax change.

B. The Labor Supply Response to Taxes

In order to capture a potentially progressive (or regressive) tax schedule, we let the marginal tax rate depend on a polynomial in $\log(w_{it}h_{it}+y_{it})$ for each household.[10]

$$(8) \quad \log(1-\tau'(w_{it}h_{it}+y_{it})) = \sum_{k=0}^{K} \gamma_{ik}[\log(w_{it}h_{it}+y_{it})]^k.$$

Next, we describe the effect of a change in γ_{0i} on labor supply. Note that a one percentage point change in γ_0 increases the after tax wage by one percentage point, holding pretax income constant. Combining Equations 1, 5, and 8 and differentiating yields:[11]

$$(9) \quad \frac{d\log h_{it}}{d\gamma_{0i}} = \sigma\left(\frac{d\log(1-\tau_{it}')}{d\gamma_{0i}} + \theta\frac{d\log h_{it}}{d\gamma_{0i}} + \frac{d\log\lambda_{it}}{d\gamma_{0i}}\right).$$

There are three pieces on the right hand side of Equation 9, reflecting different labor supply incentives arising from a tax change. The first term reflects changes in the posttax wage, holding the pretax wage fixed. A reduction in taxes causes an increase in the posttax wage, which in turn affects labor supply. This is the usual object of interest in intertemporal labor supply studies. The second term arises from the effect of hours worked upon the wage. If $\sigma > 0$, reductions in taxes cause increases in hours worked, which in turn increases the pretax wage (because of tied wage-hours offers). Because the pretax wage increases, hours worked increase further. The final term is the effect of the tax change on the marginal utility of wealth. Increases in γ_{0i} (that is, decreases in marginal tax rates) tend to increase lifetime wealth and thus decrease its marginal utility, $\frac{d\log\lambda_{it}}{d\gamma_{0i}} \leq 0$. Nevertheless, the labor supply response to tax changes, holding the marginal utility of wealth constant, is an important object since it is used to calibrate many of the key tax models (Altig *et al.* 2001) and it is a measure of the deadweight loss associated with tax changes (Ziliak and Kniesner 1999).

The elasticity that is of most interest to tax analysts can be derived using Equation 9 to solve for $\frac{d\log h_{it}}{d\log(\tau_{it}')}|_{\lambda_{it}}$ as a function of $\frac{d\log(\tau_{it}')}{d\gamma_{i0}}|_{\lambda_{it}}$, then dividing through by $\frac{d\log(\tau_{it}')}{d\gamma_{i0}}|_{\lambda_{it}}$:[12]

10. This approach follows MaCurdy *et al.* (1990) and Ziliak and Kniesner (1999). In practice, we use a third-order polynomial in log income. We also tried higher order polynomials, although this adjustment did not affect our results. A differentiable tax function makes the evaluation of the labor supply response to tax changes more straightforward, as in Equation (10).

11. Note that the budget set may not be convex if either $\theta > 0$ or if τ_{it}' declines with income (which happens when the Social Security payroll tax is phased out, for example). However, Equation (9) still represents an equilibrium condition given our chosen values of σ, θ, and γ_{ik} parameters.

12. Recall that $\frac{d\log h_{it}}{d\log(\tau_{it}')}|_{\lambda_{it}}$ is somewhat difficult to interpret because the marginal tax rate is a function of hours worked. Because of progressive taxation, longer work schedules also force some households into a higher tax bracket. However, we can interpret this function as $\frac{\frac{d\log h_{it}}{d\gamma_{i0}}|_{\lambda_{it}}}{\frac{d\log(\tau_{it}')}{d\gamma_{i0}}|_{\lambda_{it}}}$.

$$(10)\quad \frac{d\log h_{it}}{d\log \tau'_{it}}\Big|_{\lambda_{it}} = \left(\frac{\tau'_{it}}{\tau'_{it}-1}\right)\left(\frac{\sigma}{1-\sigma\theta}\right).$$

If $\theta = 0$, Equation 10 reduces to something familiar. However, the term $1-\sigma\theta$ measures the extent to which a tax cut raises hours, which then raises the wage, which in turn further increases hours. This feedback increases the responsiveness of hours worked.

C. Inferring the Labor Supply Response to Tax Changes using the Labor Supply Response to Wage Changes: Analytic Results

This section shows how failure to account for progressive taxation has potentially led to inconsistent estimates of σ in previous work. As such, our discussion is intended to provide intuition for the empirical and calibration results to follow.

Define $\sigma_{IV}^{notaxes} \equiv$ the probability limit of the IV estimator of Equation 7 when the econometrician incorrectly assumes $\Delta\log(1-\tau'_{it}) = 0$. We are interested in how this differs from σ (that is, we are interested in how ignoring progressive taxation affects the estimated labor supply elasticity). In order to derive some simple analytic results, we assume that $\log(1-\tau'_{it})$ is linear in log labor income, and that taxes depend only on labor income (that is, $\gamma_{ik} = \gamma_k$ and $\gamma_k = 0$ for $k > 1$), so that $\log(1-\tau'_{it}) = \gamma_0 + \gamma_1 \log(w_{it}h_{it})$. Appendix 3 shows that

$$(11)\quad \sigma_{IV}^{notaxes} = \frac{\sigma(1+\gamma_1)}{1-\sigma\gamma_1}.$$

Thus, $\sigma_{IV}^{notaxes} < \sigma$ under progressive taxation (that is, when $\gamma_1 < 0$). The difference between σ and $\sigma_{IV}^{notaxes}$ arises because σ measures the labor supply response to the posttax wage, whereas $\sigma_{IV}^{notaxes}$ measures the labor supply response to the pretax wage. Using the pretax wage as a proxy for the posttax wage leads the econometrician to overstate the amount of variability in the posttax wage, biasing the σ coefficient downwards.

Finally, the relationship between $\sigma_{IV}^{notaxes}$ and $\frac{d\log h_{it}}{d\gamma_0}\big|_{\lambda_{it}}$ can be derived analytically using Equations 10 and 11:

$$(12)\quad \frac{d\log h_{it}}{d\log \tau'_{it}}\Big|_{\lambda_{it}} = \left(\frac{\tau'_{it}}{\tau'_{it}-1}\right)\left(\frac{\sigma_{IV}^{notaxes}}{1-\sigma_{IV}^{notaxes}\,\theta+\gamma_1(1+\sigma_{IV}^{notaxes})}\right).$$

Comparing Equation 12 to 10, note that Equation 12 has the extra term $\gamma_1(1+\sigma_{IV}^{notaxes})$ which is negative if taxation is progressive.

Note that if there is no progressive taxation (captured in the term $\gamma_1(1+\sigma_{IV}^{notaxes})$) or tied wage-hours offers (captured in the term $(\sigma_{IV}^{notaxes}\theta)$), Equation 12 reduces to $(\frac{\tau'_{it}}{\tau'_{it}-1})\sigma$. Thus both progressive taxation and tied wage-hours may lead to larger labor supply responses than in an analysis where both of these effects are ignored.

D. Calibration

To get a better sense of whether progressive taxation and tied wage hours offers are important, Table 1 describes calibrations of the key tax derivative, $\frac{d\log h_{it}}{d\log\tau'_{it}}\big|_{\lambda_{it}}$, expressed in Equation 12, using plausible ranges of the underlying parameters, θ, $\sigma_{IV}^{notaxes}$, and γ_1. For θ, we allow the wage-hours relationship to vary from 0 to 0.60, which seems to cover the range of estimates in the literature. Most studies measure $\sigma_{IV}^{notaxes}$ to be between 0 and 0.5 for continuously employed men but this parameter is often estimated to be greater than one for women (for example, Heckman

Table 1

Value of $\frac{d \log h_{it}}{d \log \tau'_{it}}|_{\lambda_{it}}$

A $\gamma_1 = 0$

	θ				
$\sigma_{IV}^{notaxes}$	0	0.1	0.2	0.4	0.6
0	0	0	0	0	0
0.5	-0.33	-0.35	-0.37	-0.42	-0.48
1	-0.67	-0.74	-0.83	-1.11	-1.67

B $\gamma_1 = -0.10$

	θ				
$\sigma_{IV}^{notaxes}$	0	0.1	0.2	0.4	0.6
0	0	0	0	0	0
0.5	-0.39	-0.42	-0.44	-0.51	-0.61
1	-0.83	-0.95	-1.11	-1.67	-3.33

and MaCurdy 1980). Therefore, we allow $\sigma_{IV}^{notaxes}$ to vary between zero and one to account for the vast majority of estimates in the literature.

Lastly, we allow γ_1 to take on two values: 0 and -0.10. Zero represents a proportional tax schedule. Larger negative values of γ_1 characterize more progressive tax systems. In the United states we estimate γ_1 to be, on average, -0.11 for the 1977–96 period using OLS.[13] We assume that marginal tax rates are at the sample mean of $\tau'_{it} = 0.4$ throughout.

Panel A of Table 1 displays the proportional tax case. When $\sigma_{IV}^{notaxes} = 0.5$ and $\theta = 0$, $\frac{d \log h_{it}}{d \log \tau'_{it}}|_{\lambda_{it}}$ = -0.33. However, when $\theta = 0.4$, the value of $\frac{d \log h_{it}}{d \log \tau'_{it}}|_{\lambda_{it}}$, rises to -0.42. Thus, failure to account for tied wage-hours offers of this magniture would lead an analyst to understate the importance of taxation by 25 percent (-0.42 versus -0.33). With $\sigma_{IV}^{notaxes} = 1$, a relevant case for women, the difference between $\theta = 0.0$ and $\theta = 0.4$ is 66 percent (-0.67 versus -1.11). However, inelastic labor supply or a more modest assumption about the magnitude of the wage-hours tie, including plausible estimates such as 0.1 or 0.2, results in smaller differences.

Panel B introduces progressive taxes at the level observed for the United States ($\gamma_1 = -0.10$). Given $\theta = 0.4$ and $\sigma_{IV}^{notaxes} = 0.5$, $\frac{d \log h_{it}}{d \log \tau'_{it}}|_{\lambda_{it}}$ becomes -0.51, which is 31 percent larger than the case where $\theta = 0.0$. For $\sigma_{IV}^{notaxes} = 1$, $\frac{d \log h_{it}}{d \log \tau'_{it}}|_{\lambda_{it}}$ becomes -1.67, which doubles the case where $\theta = 0.0$.

13. This is based on a regression of the PSID respondents' effective marginal tax rate on log income. Adding a more complicated log income polynomial has only a marginal impact on the progressivity parameters as well as the general fit of the regression. Instrumenting for income using age or education results in estimates of γ_1 that are closer to 0.

Finally, these calibrations suggest that progressive taxation can be important. In particular, if we set $\theta = 0$ and $\sigma_{IV}^{notaxes} = 0.5$, $\frac{d \log h_{it}}{d \log \tau_{it}^{'}}|_{\lambda_{it}}$ changes from -0.33 to -0.39 when γ_1 changes from 0 to -0.10. That is, going from the case of proportional taxation to a progressive tax system similar in size to some estimates of the United States, increases the magnitude of the marginal tax rate elasticity by 18 percent.

IV. Estimation Strategy

This section briefly describes our approach for estimating Equation 7, and in particular how we measure and identify the key right hand side covariates. The first, changes in the marginal tax rate, is generated for each individual using the NBER's TAXSIM program. We describe this procedure in more detail when we describe the data in Section V. The second term is the wage. The third term, $\log \beta(1 + r_{t-1}(1 - \tau_A))$, measures changes in interest rates over time, and is accounted for in the regressions by year dummies. The fourth term, the marginal utility of wealth, is described below. Health status change regressors capture the observed component of preference shifters, the fifth term. The remaining portion of that term is potentially correlated with wages, and will be instrumented for, as described below.

The first and second terms of Equation 7 are endogenous. The marginal tax rate is endogenous because hours choices affect it. The wage change is potentially correlated with the innovation to the marginal utility of wealth (that is, the fourth term in Equation 7) if the wage change is unanticipated. Therefore, we need anticipated sources of posttax wage variation that are uncorrelated with unobserved preferences and the marginal utility of wealth to identify σ.

We use two different approaches to measure anticipated wage growth. First, we exploit the life cycle wage profile and assume that workers are able to anticipate future posttax wage growth based on their age, as in Browning *et al.* (1985), among many others. The age profile will give consistent estimates of σ so long as age-specific variation in preferences is fully accounted for using health status.[14] However, preferences for leisure potentially change over the life cycle, even controlling for health status. Therefore, we also exploit education-specific differences in wage growth. In our second IV approach we use education and education interacted with age polynomials as instruments for wage growth, while including an age polynomial in the labor supply function to control for life cycle changes in preferences. Appendix 3 shows that using age-specific variation in productivity will yield consistent estimates of σ.[15]

14. Note that the estimating Equation 7 allows for an age trend for preferences: that is, $\varepsilon_{it} = \tilde{\varepsilon}_{it} + \omega \times \text{age}_{it}$. In first differences, $\Delta\varepsilon_{it} = \Delta\tilde{\varepsilon}_{it} + \omega$, so ω just enters the constant term.

15. An alternative strategy is to assume workers can anticipate future wage growth based on their current wage and thus use lagged wages or wage changes as instruments, as in Altonji (1986), Holtz-Eakin *et al.* (1988), and Ziliak and Kniesner (1999), among others. However, in the presence of tied wage-hours offers, changes in hours worked caused by changes in preferences will impact the wage. This violates the orthogonality assumptions of the life cycle labor supply model. Because lagged wages depend on lagged hours, lagged wages will only be a valid instrument for the current wage if $E[\Delta\varepsilon_{it}\varepsilon_{it-k}] = 0$ for wages lagged k periods. It is possible to show that a slightly modified version of the lagged wage instrument that adjusts lagged wages by $\theta \log h_{it}$ can potentially eliminate this feedback effect. Results are available upon request. But it appears to us that the age profile is clearly a cleaner instrument in a setting with tied wage-hours offers.

In short, the estimating equation is:

$$(13) \quad \Delta \log h_{it} = \sigma[\Delta \log(1-\tau_{it}^{'}) + \Delta \log w_{it}] + X_{it}^{'}\Gamma + \Delta u_{it}$$

where the residual $\Delta u_{it} = \sigma \frac{\beta(1+r_{t-1}(1-\tau_A))\varepsilon_{it}}{\lambda_{it-1}} + \Delta\varepsilon_{it}$ and X_{it} includes changes in health status and time dummies. Depending on the specification, X_{it} also includes either an age polynomial or education.

V. Data

We use the PSID and March supplements to the CPS to estimate σ. Besides allowing for some corroboration across independent data sets, each data set offers specific advantages. In particular, the PSID includes key covariates not available in the CPS, such as health status and detailed income and deduction information used to compute marginal tax rates. The CPS provides larger sample sizes. Although the questions are more limited in the CPS, we can still roughly recreate the PSID specification.

In both data sets, we restrict the sample to male household heads aged 25 to 60 that are not self-employed. The latter restriction acknowledges the difficulty of distinguishing capital and labor income among the self-employed. We also drop workers with fewer than 300 or more than 4,500 hours and those who earn less than three dollars or more than $200 per hour in 1996 dollars.[16] Our selection criterion leads to a sample of 5,618 working men encompassing 43,348 person-year observations in the PSID between 1977 and 1996. The matched CPS sample consists of 164,147 working men between 1979 and 2003.[17] Descriptive statistics are available in Appendix 4.

Two variables require further elaboration. First, we use a common measure of the hourly wage, annual earnings divided by annual hours. However, such a measure introduces a nonstandard measurement error problem called "division bias" by allowing measurement error in hours to enter both the left hand and right hand side of the estimating Equation 13.[18] This can drive estimates of the wage elasticity to negative values given finite sample size. Consequently, we also provide estimates that use log annual earnings rather than log wages.

16. Because of the topcoding of annual income in the CPS, especially prior to 1988, we restrict that sample to those with a nominal wage at or below $100. We also restrict the CPS sample by eliminating anyone with annual hours or an hourly wage that changes by more than a factor of three in either direction. This is to alleviate concerns about multiple job holding and measurement error.

17. CPS respondents are observed in only two consecutive Marches. Therefore, there is only one differenced observation per person. The match itself is based on household ID, line number, year, month in sample, gender, race, age, and education. There is no difference in the results if we exclude the main time varying variable, education. Restricting the sample to 1979 to 1996 has little impact on our results as well.

18. Some papers, such as Altonji (1986), Kimmel and Kniesner (1998), Ziliak and Kniesner (1999), and French (2004) overcome this problem using the reported current hourly wage of hourly workers. See Altonji (1986) and French (2004) for reviews of the measurement issues with this measure.

Second, effective marginal rates are computed for each household using the NBER's TAXSIM program, which calculates state and federal income taxes, as well as payroll taxes. In the PSID, we allow these taxes to depend on the year, state of residence, marital status, number of dependents (as well as the number younger than age 17), labor income of husband and spouse, dividend, rental, and interest income, Social Security income, pensions, government transfers, property taxes, mortgage interest,[19] and unemployment benefits. Because of data limitations, we use only year, state of residence, marital status, number of dependents (as well as the number younger than age 17), number of taxpayers older than age 65, and labor income of the husband and spouse with the CPS respondents. We augment the federal and state rates computed in TAXSIM with payroll tax schedules obtained from the Tax Policy Center, which is a joint project of the Urban Institute and Brookings Institution.

Figure 1 displays marginal tax rates for individuals in our PSID sample.[20] What stands out is the considerable heterogeneity in taxes, above and beyond what is caused by differences in income. A regression of $\log(1-\tau_{it}^{'})$ on log income (that is, a simplified version of Equation 8) results in a relatively low R^2 of 0.23, with an estimate of $\gamma_0 = 0.67$ and $\gamma_1 = -0.11$.[21] On the other hand, we get mixed results for γ_1 when instrumenting for income using our instrument sets. For example, the age instruments result in $\gamma_1 = -0.02$. The education instruments flip sign, with $\gamma_1 = 0.07$.

In order to account for heterogeneity in taxes, we generate $\log(1-\tau_{it}^{'})$ for each person in the data. We follow the general approach of MaCurdy *et al.* (1990); Ziliak and Kneisner (1999); and advocated by Blundell and MaCurdy (2002) by smoothing the tax function. However, we extend previous approaches by allowing more individual heterogeneity in marginal tax rate schedules.[22] Our procedure accounts for the fact that a large amount of individual heterogeneity in marginal tax rates comes from heterogeneity caused by mortgage interest, for example.

VI. Results

Table 2 reports the PSID results. There are four sets of estimates, depending on whether the dependent variable is the hourly wage or annual earnings

19. Mortgage payments are measured in the PSID, although mortgage interest is not. We impute mortgage interest using mortgage payments × 0.8. For an interest rate of 7 percent, mortgage interest is between 74 and 87 percent of the mortgage payment for the first ten years of a 30 year fixed rate mortgage. Given that 60 percent of all mortgages are less than six years old (Bucks and Pence, 2006), assuming 80 percent of mortgage payments go to interest is an accurate approximation for the majority of all mortgages.

20. To account for substantial changes in the tax code introduced by the 1986 law changes, we show the rates separately pre- and postreform. It is also important to note that there are few households facing negative marginal tax rates because we include payroll taxes and limit the sample to those households headed by men with at least $5,000 in annual income. However, the EITC is accounted for in the calculations.

21. Adding in a fifth order polynomial in log income only increases the R^2 to 0.25.

22. Because this approach requires a smooth approximation of the budget constraint for every individual (which is computationally intensive), and because it is difficult to recover the exact tax function using TAXSIM, we evaluate each individual's labor income at nine different income levels (income × 0.8, income ×0.85, …, income × 1.2). We fit the nine data points using a fourth order polynomial, although our results were largely unaffected by choosing other polynomial orders.

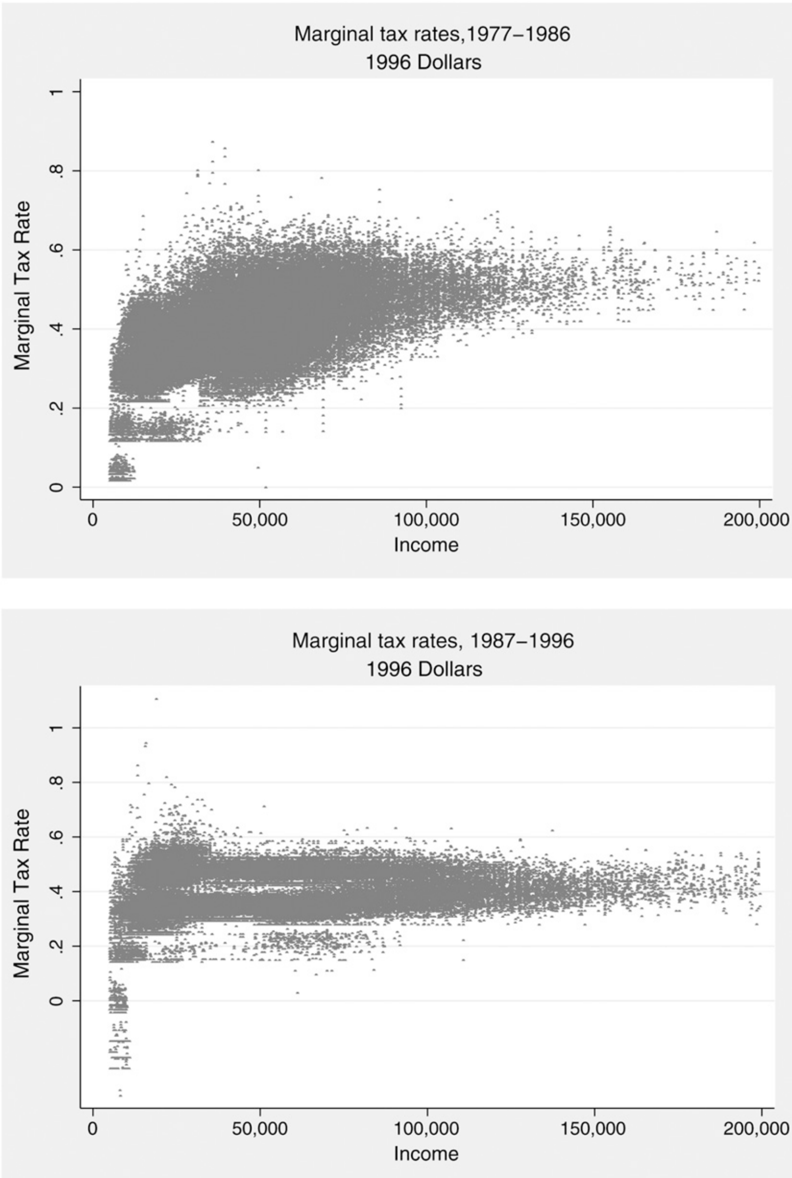

Figure 1
Marginal Tax Rates

Table 2
Estimated Labor Supply Elasticities, PSID 1977-96

Column	(1)	(2)	(3)	(4)
Dependent variable	Hourly wage	Hourly wage	Annual earnings	Annual earnings
Instrument set	Age	Education	Age	Education
First-stage estimates, dependent variable is		$\Delta \log w_{it}$		
Number of instruments	3	16	3	16
Overidentification statistic	12.40	4.65	11.10	4.94
P-value	0.006	0.996	0.011	0.995
F-statistic	19.0	5.0	37.0	6.5
P-value	0.000	0.000	0.000	0.000
Partial R^2	0.0013	0.0018	0.0013	0.0018
N	43,348	43,348	43,348	43,348
Second-stage estimates, dependent variable is		$\Delta \log h_{it}$		
$\sigma_{IV}^{notaxes}$	0.343	0.161	0.504	0.219
	(0.089)	(0.065)	(0.084)	(0.070)
σ	0.344	0.140	0.514	0.196
	(0.092)	(0.060)	(0.087)	(0.065)
Marginal tax rate elasticity				
$\theta = 0.0, \sigma = \hat{\sigma}_{IV}^{notaxes}$	−0.237	−0.112	−0.348	−0.151
	(0.061)	(0.040)	(0.057)	(0.048)
$\theta = 0.0, \sigma = \hat{\sigma}$	−0.238	−0.097	−0.356	−0.135
	(0.063)	(0.042)	(0.060)	(0.045)
$\theta = 0.1, \sigma = \hat{\sigma}$	−0.246	−0.098	−0.374	−0.138
	(0.068)	(0.043)	(0.066)	(0.047)
$\theta = 0.4, \sigma = \hat{\sigma}$	−0.275	−0.103	−0.447	−0.147
	(0.085)	(0.047)	(0.095)	(0.053)

Notes: The age instrument set is a third-order age polynomial. In this case, the education dummies are included in the labor supply equation. The education instrument set is five dummy variables interacted with a third-order age polynomial. In this case, the age polynomial is included in the labor supply equation. Other variables included in the labor supply equation are year dummies and health status change.

and whether the instrument set is a third-order age polynomial or education and education interacted with the age polynomial. To demonstate the power of these instruments, the top panel displays the F-statistic, partial R^2 and the overidentification test statistic from the first-stage regression. The middle panel reports the estimated intertemporal elasticity of substitution, σ, as well as $\sigma_{IV}^{notaxes}$, when taxation is ignored. Finally, the bottom panel shows the elasticity of hours with respect to the marginal tax rate ($\frac{d \log h_{it}}{d \log \tau'_{it}}|_{\lambda_{it}}$), given different assumptions on σ and θ.

With regard to the age polynomial instrument set (Columns 1 and 3), we find that the partial R^2 is low, but the F-statistic easily exceeds standard thresholds of significance. On the other hand, the overidentification test statistics reject the hypothesis that age is a valid instrument. When using the education and education-age polynomial interaction (Columns 2 and 4), we find that the partial R^2 and the F-statistic drop significantly and the F-statistic is on the border of being acceptable. On the other hand, the overidentification test statistic does not reject the hypothesis of proper model specification.

The middle panel displays the estimated intertemporal elasticity of substitution parameter σ and $\sigma_{IV}^{notaxes}$ for the four combinations of dependent variables and instrument sets. One common theme is that $\sigma_{IV}^{notaxes}$ and σ are very similar in magnitude. For example, in Column 1, we estimate that $\sigma_{IV}^{notaxes}$ is 0.343 (with a standard error of 0.089)[23] and σ is 0.344 (0.092).[24] As mentioned above, and formally shown in the appendixes, failure to account for progressive taxation likely leads to a downward biased estimate of σ (that is, 0.344 versus 0.343). But in this example, and the remaining columns as well, this effect is small. As we pointed out in Section V, instrumenting for income using age or education suggests that taxation is not very progressive.[25]

The bottom panel shows the elasticity of hours with respect to the marginal tax rate ($\frac{d\log h_{it}}{d\log \tau'_{it}}|_{\lambda_{it}}$), given different assumptions on σ and θ. In Column 1, this elasticity is -0.237 when using using our estimate of $\sigma_{IV}^{notaxes}$ as a proxy for σ, and setting $\theta = 0$. As we have argued, that is essentially the approach that many previous scholars have used. When our estimate of σ that accounts for taxation is used (but $\theta = 0$), $\frac{d\log h_{it}}{d\log \tau'_{it}}|_{\lambda_{it}}$ changes only slightly, to -0.238. But introducing hours-wage ties causes $\frac{d\log h_{it}}{d\log \tau'_{it}}|_{\lambda_{it}}$ to jump, by 4 percent if θ is 0.1 but by 16 percent to -0.275 if $\theta = 0.4$.

To provide intuition for the last estimate, consider the effect of a 1 percent increase in τ'_{it}. This marginal tax rate increase reduces the after tax wage $\frac{\tau'_{it}}{1-\tau'_{it}}$ percent (at sample means, τ'_{it} is 0.40 and $\frac{\tau'_{it}}{1-\tau'_{it}}$ is 0.69). Given an estimate of $\sigma = 0.34$, hours worked drop $0.69 \times 0.34 = 0.238$ percent. But because hours drop, the wage drops further, causing hours to decline 0.275 percent in total.

Column 2, which uses the education instrument set produces much smaller estimates of $\sigma_{IV}^{notaxes}$ and σ and consequently any impact of hours wage ties are muted. In particular, the labor supply elasticity with respect to the marginal tax rate is -0.097 when $\theta = 0$ and -0.103 when $\theta = 0.4$, a difference of only 6 percent.

The final two columns of Table 2 attempt to minimize division bias by respecifying the labor supply function in terms of log earnings rather than log wages. The estimating Equation 13 becomes

23. Standard errors are computed using the multivariate delta method and correct for arbitrary forms of heteroskedasticity and serial correlation.

24. These estimates are at the high end of the literature for men, although slightly lower than Lee (2001), who estimates $\sigma_{IV}^{notaxes}$ to be 0.5 using a similar sample and instrument set. Lee finds that using unbalanced data and a parsimonious instrument set overcomes small sample bias and thus leads to higher estimates of the intertemporal elasticity of substitution. Using lagged wages as instruments typically yields a smaller estimate of the intertemporal elasticity of substitution (e.g., Altonji (1986), French (2004), Holtz-Eakin *et al.* (1988)).

25. Small differences do not necessarily refute the hypothesis that progressive taxation is important in general - only that progressive taxation is unimportant given our instruments. For example, there is very little life cycle variation in marginal tax rates. Moreover, there may be offsetting influences on life cycle taxable income across education groups due to the propensity to use deductible items, such as the home mortgage deduction.

$$(14)\quad \Delta \log h_{it} = \tilde{\sigma}[\Delta \log(1-\tau_{it}^{'}) + \Delta \log E_{it}] + \frac{1}{1+\sigma} X_{it}^{'}\Gamma + \frac{1}{1+\sigma}\Delta u_{it}$$

where E_{it} is earnings and $\sigma = \frac{\tilde{\sigma}}{1-\tilde{\sigma}}$. It can be easily shown that this modification results in σ being biased to zero rather than negative one because of measurement error. However, Ghez and Becker (1975) point out that omitted variables potentially lead to an upward bias using this specification. As expected this alteration increases the size of σ and the marginal tax rate elasticities. Using the age polynomial instruments (Column 3), σ is now 0.514 and $\frac{d \log h_{it}}{d \log \tau_{it}^{'}}|_{\lambda_{it}}$ jumps to -0.356 when $\theta = 0$. Because the estimate of σ increases, the importance of the feedback from tied wage hours rises as well. Specifically, $\frac{d \log h_{it}}{d \log \tau_{it}^{'}}|_{\lambda_{it}}$ increases to -0.374 and -0.447 when $\theta = 0.1$ and $\theta = 0.4$, a 7.5 percent and 28 percent larger estimate than the case when $\theta = 0$. The wage-hours tie increases $\frac{d \log h_{it}}{d \log \tau_{it}^{'}}|_{\lambda_{it}}$ by about 9 percent when the education instruments are used in Column 4.

Table 3 provides comparable estimates for the March CPS sample. First note that the larger sample sizes are helpful. The statistical significance of the instrument sets is across-the-board higher in the larger CPS samples. On the other hand, the overidentification tests reject the hypothesis of correct model specification. We also find the estimated intertemporal elasticities of substitution, as well as the elasticities of hours with respect to the marginal tax rate, are larger. That said, the CPS results are quite similar to the PSID. In particular, when we introduce hours-wage ties at $\theta = 0.4$, the resulting change in $\frac{d \log h_{it}}{d \log \tau_{it}^{'}}|_{\lambda_{it}}$ jumps by 10 to 30 percent, depending on the dependent variable and instrument set employed, just slightly higher than the 5 to 28 percent range indicated by the PSID.

VII. Conclusions

There are two important caveats to our analysis. First, we consider the decision of how many hours to work (the "intensive margin"), not the decision of whether to work (the "extensive margin").[26] Heckman (1993) contends that most of the variability in labor supply is at the extensive margin. French (2005) and Rogerson and Wallenius (2007) argue that a part-time wage penalty is necessary to fit the distribution of hours worked over the life cycle. Furthermore, these papers point out that a part-time wage penalty increases labor supply elasticities on the extensive margin. A contribution of this paper is to show that allowing for fixed costs increases labor supply elasticities on the intensive margin also.

The second concern is that we focus only on the substitution effect associated with tax wage changes. Understanding the substitution effects is arguably sufficient for understanding the labor supply response to short-term tax adjustments. However,

26. See Kimmel and Kniesner (1998) for a decomposition of labor supply elasticities into the intensive and extensive margins).

Table 3
Estimated Labor Supply Elasticities, CPS 1979-2003

Column	(1)	(2)	(3)	(4)
Dependent variable	Hourly wage	Hourly wage	Annual earnings	Annual earnings
Instrument set	Age	Education	Age	Education
First stage estimates, dependent variable is		$\Delta \log w_{it}$		
Number of instruments	3	16	3	16
Overidentification statistic	27.0	46.7	24.9	34.9
P-value	0.000	0.000	0.000	0.004
F-statistic	57.3	7.4	125.6	13.2
P-value	0.000	0.000	0.000	0.000
Partial R^3	0.0010	0.0007	0.0011	0.0007
N	164,147	164,147	164,147	164,147
Second stage estimates, dependent variable is		$\Delta \log h_{it}$		
$\sigma_{IV}^{notaxes}$	0.459	0.365	0.512	0.608
	(0.063)	(0.090)	(0.053)	(0.090)
σ	0.491	0.301	0.583	0.532
	(0.074)	(0.078)	(0.063)	(0.082)
Marginal tax rate elasticity , given,				
$\theta = 0.0, \sigma = \hat{\sigma}_{IV}^{notaxes}$	−0.350	−0.278	−0.391	−0.464
	(0.048)	(0.069)	(0.040)	(0.068)
$\theta = 0.0, \sigma = \hat{\sigma}$	−0.375	−0.230	−0.445	−0.406
	(0.057)	(0.059)	(0.048)	(0.062)
$\theta = 0.1, \sigma = \hat{\sigma}$	−0.394	−0.237	−0.473	−0.428
	(0.063)	(0.063)	(0.054)	(0.069)
$\theta = 0.4, \sigma = \hat{\sigma}$	−0.467	−0.261	−0.581	−0.515
	(0.088)	(0.076)	(0.082)	(0.100)

Notes: The age instrument set is a third-order age polynomial. In this case, the education dummies are included in the labor supply equation. The education instrument set is five dummy variables interacted with a third-order age polynomial. In this case, the age polynomial is included in the labor supply equation. Other variables included in the labor supply equation are year dummies.

to understand the importance of fundamental tax reform, it is necessary to recognize the wealth effects associated with tax changes.

Nevertheless, we believe that we have shown, both qualitatively and quantitatively, that augmenting a standard intertemporal labor supply model to account for tied wage-hours offers and progressive taxation can affect estimates of the intertemporal elasticity of substitution and the labor supply response to tax changes. Using common methods to estimate men's labor supply functions and using our interpretation of the size of the wage-hours tie, we find that the hours response to a change in marginal tax rates may be biased by at least 5 percent, and as much as 30 percent, relative to many of the estimates in the literature that do not account for tied wage-hours offers and progressive taxation. Therefore, tax analysts inferring the extent of behavioral responses to tax changes should consider the source of variation used for identification.

Appendix 1

The Specification of Tied Wage-Hours Offers

To formally capture the link between hours worked and the offered wage, we first note that, in equilibrium, perfectly competitive firms cover their fixed costs so that total output equals the wage bill plus the fixed cost of work:

$$(15) \quad p_{it}h_{it} = w_{it}h_{it} + \phi$$

where ϕ is the fixed cost per employee, p_{it} is productivity of worker i at time t, h_{it} is hours worked, and w_{it} is the offered hourly wage. By rewriting Equation 15 as

$$(16) \quad w_{it} = p_{it} - \frac{\phi}{h_{it}},$$

it is obvious that the offered hourly wage is rising in hours worked. This relationship implies that at points in the life cycle or tax cycle that hours worked are high, the offered wage should also be high.

In order to understand whether the model presented in Equation 1 is consistent with a reasonably calibrated version of Equation 16, we choose values of p_{it} and ϕ to fit Equation 1. Specifically, we use $\theta = 0.4$, and pick α_{it} to match the average work year length (1,941 hours) and wage ($17.26, in 1996 dollars) from the sample of older PSID (age 50 to 70) males for Equation 1. Next, we pick p_{it} and ϕ to match the average wage and an elasticity of 0.4 at 1,941 hours of work for our fitted Equation 16.

The left panel in Figure 2 plots the estimated relationship between hours worked and the offered hourly wage, using Equation 1. The right hand panel plots the elasticity of the wage with respect to hours worked implied by Equations 16 and 1. Between 1,700 and 2,500 hours, encompassing 68 percent of our sample, the implied elasticity from Equation 16 is between 0.48 and 0.28, as compared to the constant elasticity implied by Equation 1. Therefore, we conclude the linearized relationship in Equation 1 provides a good approximation to the structural Equation 16.

Moreover, the estimated value of θ seems to provide a plausible estimate of the fixed cost of work. We find $\phi = \$13{,}450$ and $p_{it} = \$23.30$, implying that 28 percent of firms' labor costs $\frac{13{,}450}{13{,}450 + 17.26 * 1{,}941}$ are fixed. This accords reasonably well with the studies on recruitment and training costs cited in Malcomson (1999).

Appendix 2

Controlling for Changes in the Marginal Utility of Wealth

This appendix describes our approach for dealing with changes in the marginal utility of wealth in order to derive Equation 7 from the first differenced labor supply function illustrated in Equation 6. The discussion follows MaCurdy (1985), in which the marginal utility of wealth and, in approximation, the log of the marginal utility of wealth are shown to follow a random walk with drift. This result falls out of the Euler

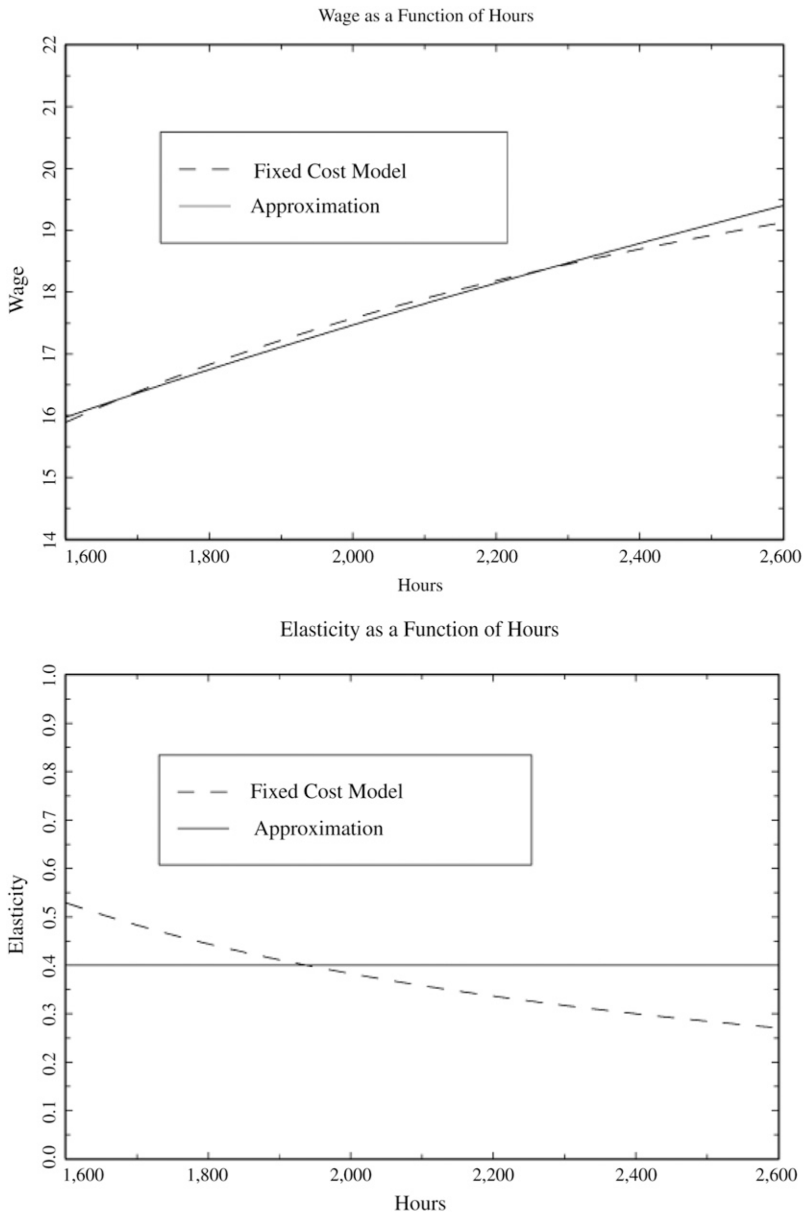

Figure 2
Offered Hourly Wage as a Function of Hours

equation of the model described in Section IIIA. In particular, the Euler equation indicates that individuals equate expected marginal utility across time according to:

$$(17) \quad \lambda_{it-1} = \beta(1 + r_{t-1}(1-\tau_A))E_{t-1}\lambda_{it}$$

where rational expectations[27] implies that innovations to the marginal utility of wealth, denoted ε_{it}, should be uncorrelated with lagged values of the marginal utility of wealth:

$$(18) \quad \lambda_{it} = E_{t-1}\lambda_{it} + \varepsilon_{it}.$$

Equations 17 and 18 can be rewritten as

$$(19) \quad \frac{\beta(1 + r_{t-1}(1-\tau_A))\lambda_{it}}{\lambda_{it-1}} = \left(1+\frac{\beta(1+r_{t-1}(1-\tau_A))\varepsilon_{it}}{\lambda_{it-1}}\right).$$

Taking logarithms of both sides of Equation 19 and approximating $\log(1+\frac{\beta(1+r_{t-1}(1-\tau_A))\varepsilon_{it}}{\lambda_{it-1}})$ yields

$$(20) \quad \log\lambda_{it} - \log\lambda_{it-1} + \log\beta(1+r_{t-1}(1-\tau_A)) = \log\left(1+\frac{\beta(1+r_{t-1}(1-\tau_A))\varepsilon_{it}}{\lambda_{it-1}}\right) \approx \frac{\beta(1+r_{t-1}(1-\tau_A))\varepsilon_{it}}{\lambda_{it-1}}.$$

We assume that the approximation in Equation 20 holds with equality, a valid assumption as innovations in the marginal utility of wealth become arbitrarily small. Combining Equations 20 and 6 results in

$$(21) \quad \Delta\log h_{it} = \sigma[\Delta\log(1-\tau'(.)) + \Delta\log w_{it}] - \sigma\log\beta(1 + r_{t-1}(1-\tau_A)) + \sigma\frac{\beta(1 + r_{t-1}(1-\tau_A))\varepsilon_{it}}{\lambda_{it-1}} + \Delta\varepsilon_{it}.$$

Because the innovation to the marginal utility of wealth is potentially correlated with wage changes if the wage change is unanticipated, the wage must be instrumented. See Section IV for a discussion on instrument selection.

Appendix 3

Estimates of σ that do not Account for Progressive Taxation

In this appendix, we show that overlooking progressive taxation potentially leads to inconsistent estimates of σ. In order to simplify the analysis, consider the case where $\log(1-\tau'_{it})$ is linear in the log of labor income, and the marginal tax rate is unaffected by spousal income:

27. If workers have rational expectations then at time t they know their state variables $\alpha_{it}, \theta, r_t, \varepsilon_{it}, \tau_{it}$, the Markov process that determines the evolution of the state variables, and optimize accordingly.

(22) $\log(1-\tau_{it}^{'}(w_{it}h_{it}+y_{it})) = \gamma_0+\gamma_1[\log(w_{it})+\log(h_{it})]$.

Further, ignore the importance of variable interest rates and observable preference shifters.[28] Therefore, Equation 13 can be rewritten as:

(23) $\Delta\log h_{it} = \sigma[\Delta\log(1-\tau^{'}(.)) + \Delta\log w_{it}] + \Delta u_{it}$

where $\Delta u_{it} = \sigma\frac{\beta(1+r_{t-1}(1-\tau_A))\varepsilon_{it}}{\lambda_{it-1}} + \Delta\varepsilon_{it}$. Combining Equations 1, 7, and 22 yields the reduced form equations of the system:

(24) $$\Delta\log h_{it} = \frac{\sigma[(1+\gamma_1)\Delta\alpha_{it}] + \Delta u_{it}}{1-\sigma(\gamma_1(1+\theta)+\theta)}$$

(25) $$\Delta\log w_{it} = \frac{[(1-\sigma\gamma_1)\Delta\alpha_{it} + \theta\Delta u_{it}]+\Delta u_{it}}{1-\sigma(\gamma_1(1+\theta)+\theta)}.$$

Typically, instrumental variables procedures are used to estimate $^{\sigma}$ within the misspecified model

(26) $\Delta\log h_{it} = \sigma_{IV}^{notaxes}[\Delta\log w_{it}] + \Delta u_{it}$

where $\sigma_{IV}^{notaxes}$ is the probability limit of the wage coefficient on the misspecified model.

We derive $\sigma_{IV}^{notaxes}$ when age is used as the instrumental variable. Let an individual's age-specific productivity be the sum of two orthogonal components, or $\alpha_{it} = \alpha_t + \psi_{it}$ where α_t is the age-specific component of wages and ψ_{it} is the idiosyncratic component of wages, and $E[\alpha_t\psi_{it}] = 0$. Assume $E[\Delta u_{it}\alpha_t] = 0$: that is, the life-cycle wage profile measures changes in life cycle productivity but not changes in life cycle preferences. In this case using α_t as the instrument (which is another way of saying that we use the average age-specific wage) yields

(27) $$\sigma_{IV}^{notaxes} = \frac{\sigma(1+\gamma_1)(1-\sigma\gamma_1)Cov(\Delta\alpha_{it},\Delta\alpha_t)}{(1-\sigma\gamma_1)^2 Cov(\Delta\alpha_{it},\Delta\alpha_t)} = \frac{\sigma(1+\gamma_1)}{1-\sigma\gamma_1}$$

which is Equation 11 of the paper. Finally, it is straightforward to derive the obvious result that IV estimation on Equation 23 yields a consistent estimate of σ.

28. In other words, consider a model where both the log posttax wage and posttax hours worked are the residuals from regressions of the log post tax wage and log hours worked on year dummies and observable preference shifters. Using the Frisch-Waugh-Lovell Theorem (Davidson and MacKinnon 1993), it is straightforward to show that using this approach will still yield a consistent estimate of σ.

Appendix 4
Data

PSID and March CPS Descriptive Statistics: Means (standard deviations in parentheses)

Variable	PSID	CPS
Head's hours	2,178 (525)	2,202 (470)
Head's wage	$17.9 (10.2)	$ 15.6 (10.4)
Marginal tax rate	0.397 (0.081)	0.422 (0.077)
Head's earnings	$38,651 (23,657)	$34,963 (28,416)
Wife's earnings*	$12,184 (15,259)	$ 13,510 (19,408)
Age	38.7 (8.9)	41.3 (9.2)
No high school grad	0.19 (0.39)	0.12 (0.33)
High school grad	0.33 (0.47)	0.35 (0.48)
Some college	0.23 (0.42)	0.22 (0.42)
College grad	0.14 (0.35)	0.18 (0.38)
Postgraduate	0.11 (0.31)	0.12 (0.33)
Married	0.85 (0.36)	0.79 (0.41)
Children	1.6 (1.4)	1.0 (1.2)
Bad health	0.08 (0.27)	n.a.

Note: For the CPS, "wife's earnings" are family earnings less male head earnings.

References

Aaronson, Daniel, and Eric French. 2004. "The Effect of Part-Time Work on Wages: Evidence from the Social Security Rules." *Journal of Labor Economics* 22(2):329-52.

Altig, David, Alan Auerbach, Laurence Kotlikoff, Kent Smetters, and Jan Walliser. 2001. "Simulating Fundamental Tax Reform in the United States." *American Economic Review* 91(3):574-95.

Altonji, Joseph. 1986. "Intertemporal Substitution in Labor Supply: Evidence from Microdata." *Journal of Political Economy* 94(3): S176-S215.

Barzel, Yoram. 1973. "The Determination of Daily Hours and Wages." *Quarterly Journal of Economics* 87(2):220-38.

Biddle, Jeff and Gary Zarkin. 1989. "Choice Among Wage-Hours Packages: An Empirical Investigation of Male Labor Supply." *Journal of Labor Economics* 7(4):415-37.

Blank, Rebecca. 1990. "Are Part-Time Jobs Bad Jobs?" *In A Future of Lousy Jobs? The Changing Structure of U.S. Wages*, ed Gary Burtless, 152-56. Washington, D.C.: Brookings Institution.

Blomquist, Soren. 1985. "Labour Supply in a Two-Period Model: The Effect of a Nonlinear Progressive Income Tax." Review of Economic Studies 52(3):515-24.

Blundell, Richard, and Thomas MaCurdy. 1999. "Labor Supply: A Review of Alternative Approaches." In Handbook of Labor Economics, Volume 3A, ed. O. Ashenfelter and D. Card, 1599-1695. New York: North-Holland.

Browning, Martin, Angus Deaton, and Margaret Irish. 1985. "A Profitable Approach to Labor Supply and Commodity Demands Over the Life-Cycle." *Econometrica* 53(3):503-43.

Browning, Martin and Costas Meghir. 1991. "The Effects of Male and Female Labor Supply on Commodity Demands." *Econometrica* 59(4):925-51.

Bucks, Brian, and Karen Pence. 2006. "Do Borrowers Know their Mortgage Terms?" Board of Governors. Unpublished.

Burman, Leonard, and Peter Ricoy. 1997. "Capital Gains and the People Who Realize Them." *National Tax Journal* 50(3):427-52.

Davidson, Russell, and James MacKinnon. 1993. *Estimation and Inference in Econometrics*. Oxford: Oxford University Press.

Ermisch, John, and Robert Wright. 1993. "Wage Offers and Full-Time and Part-Time Employment by British Women." *Journal of Human Resources* 28(1):111-33.

Feldstein, Martin. 1995. "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act." *Journal of Political Economy* 103(3):551-72.

French, Eric. 2004. "The Labor Supply Response to (Mismeasured but) Predictable Wage Changes." *Review of Economics and Statistics* 86(2):602-13.

———. 2005. "The Effects of Health, Wealth, and Wages on Labor Supply and Retirement Behavior." *Review of Economic Studies* 72(2):395-427.

Ghez, Gilbert, and Gary Becker. 1975. The Allocation of Time and Goods Over the Life Cycle. Cambridge: NBER.

Heckman, James. 1993. "What Has Been Learned About Labor Supply in the Past Twenty Years?" *American Economic Review* 83(2):116-21.

Heckman, James, and Thomas MaCurdy. 1980. "A Life Cycle Model of Female Labour Supply." *Review of Economic Studies* 47(1):47-74.

Hirsch, Barry. 2005. "Why Do Part-Time Workers Earn Less? The Role of Worker and Job Skills." *Industrial and Labor Relations Review* 58(4):525-51.

Holtz-Eakin, Douglas, Whitney Newey, and Harvey Rosen. 1988. "Estimating Vector Autoregressions with Panel Data." *Econometrica* 56(6):1371-95.

Keane, Micheal, and Kenneth Wolpin. 2001. "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment." *International Economic Review* 42(4):1051-1103.

Kimmel, Jean, and Thomas Kniesner. 1998. "New Evidence on Labor Supply: Employment versus Hours Elasticities by Sex and Marital Status." *Journal of Monetary Economics* 42(2):289-301.

Lee, Chul-In. 2001. "Finite Sample Bias in IV Estimation of Intertemporal Labor Supply Models: Is the Intertemporal Substitution Elasticity Really Small?" *Review of Economics and Statistics* 83(4):638-46.

Lettau, Michael. 1997. "Compensation in Part-Time Versus Full Time Jobs: What if the Job were the Same?" *Economics Letters* 56(1):101-106.

Lewis, H. G . 1969. "Employer Interests in Employee Hours of Work." Working paper, University of Chicago.

Lundberg, Shelly. 1985. "Tied Wage-Hours Offers and the Endogeneity of Wages." *Review of Economics and Statistics* 67(3):405-10.

MaCurdy, Thomas. 1985. "Interpreting Empirical Models of Labor Supply in an Intertemporal Framework with Uncertainty." In *Longitudinal Analysis of Labor Market Data*, ed J. Heckman and B. Singer, Chapter 3. Cambridge: Cambridge University Press.

MaCurdy, Thomas, David Green, and Harry Paarsch. 1990. "Assessing Empirical Approaches for Analyzing Taxes and Labor Supply." *Journal of Human Resources* 25(3):415-90.

Malcomson, James. 1999. "Individual Employment Contracts." Handbook of Labor Economics, Volume 3B, eds. O. Ashenfelter and D. Card, 2291-2372. New York: North-Holland.

Moffitt, Robert. 1984. ‘‘The Estimation of a Joint Wages-Hours Labor Supply Model.’’ In Journal of Labor Economics 2(4):550-66.
Mulligan, Casey. 1999. ‘‘Substitution over Time: Another Look at Life-Cycle Labor Supply.’’ NBER *Macroeconomics Annual* 1998, 75-134. Cambridge: NBER.
Rogerson, Richard, and Johanna Wallenius. 2007. ‘‘Micro and Macro Elasticities in a Life Cycle Model With Taxes.’’ NBER working paper 13017.
Rosen, Harvey. 1976. ‘‘Taxes in a Labor Supply Model with Joint Wages-Hours Determination.’’ *Econometrica* 44: 485-507.
Slemrod, Joel. 1998. ‘‘Methodological Issues in Measuring and Interpreting Taxable Income Elasticities.’’ *National Tax Journal* 51(4):773-88.
Ziliak, James and Thomas Kniesner. 1999. ‘‘Estimating Life Cycle Labor Supply Tax Effects.’’ Journal of Political Economy 107(2):326-59.