Estimating Intensive and Extensive Tax Responsiveness: Do Older Workers Respond to Income Taxes?*

Abby Alpert†
David Powell‡

October 21, 2013

Abstract

This paper studies the impact of income and payroll taxes on intensive and extensive labor supply decisions for workers ages 55-74 using the Health and Retirement Study. The literature provides little guidance about the responsiveness of this population to tax incentives, though the tax code is potentially an important mechanism that can alter retirement incentives. We model labor force participation decisions and labor earnings as functions of taxes. We estimate the intensive and extensive margins simultaneously, introducing a new method to estimate labor supply decisions more broadly. Our method accounts for selection into labor force participation with a plausibly exogenous shock to employment. We use the results of our intensive labor supply estimation to predict the after-tax labor earnings of every person in our sample, including those that do not work in the next period. This method allows us to generate consistent estimates of the impact of taxes on employment and retirement. We find insignificant compensated elasticities on the intensive margin. On the extensive margin, we find significant effects on labor force participation and retirement decisions. Our estimates suggest that an age-targeted tax reform that eliminates payroll taxes for older workers would decrease the percentage of both men and women dropping out of the labor force by almost one percentage point, a 3% decrease. We find that most of this decrease in labor force participation is associated with an increase in retirement.

Keywords: Retirement, Income Taxes, Selection Models
JEL classification: H24, J20, J26

---

*We acknowledge support from the Michigan Retirement Research Center. We would like to thank Dan Feenberg for assistance with the NBER TAXSIM program. We also thank Robert Carroll and participants at the 2012 Retirement Research Center Consortium for their helpful comments. Finally, we are grateful for comments from John Laitner and Jim Poterba.

†The Paul Merage School of Business, University of California - Irvine, aealpert@uci.edu
‡RAND, dpowell@rand.org.

1 Introduction

Economists and policymakers have long been interested in understanding the effects of economic incentives on the retirement decisions of older workers. There are large literatures investigating the labor consequences of Social Security benefits, pensions, and health insurance.[1] However, one issue that has been overlooked in the traditional retirement literature is the effect of income taxes on the decision to work. There are a number of reasons why the responsiveness to taxes may differ for older individuals and the tax code could potentially be used more actively to encourage older workers to earn more and remain in the labor force longer. Banks and Diamond (2010) suggest that age-dependent taxation is one of the most promising areas for future tax reform, meanwhile some economists have recommended eliminating the payroll tax at certain ages or after Social Security receipt (Biggs (2012)), and the elimination of income taxes for seniors earning less than $50,000 was recently proposed by Barack Obama in the 2008 presidential election.[2] Moreover, since age is an observable variable that likely proxies for different levels of productivity and attachment to the labor market,"tagging" by age may also improve redistributive taxation (Akerlof (1978); Weinzierl (2011)).

While a vast literature studies the effects of income taxes on labor supply and taxable income, these studies typically exclude older individuals from the analysis or estimate aggregate effects combining all age groups. Consequently, there is almost no empirical evidence on the extent to which older individuals respond to taxes. In this paper, we aim to fill this significant gap in the tax and retirement literatures by providing some of the first estimates of the effects of income taxes on both the intensive and extensive labor supply decisions of older workers. We use the 2000-2008 Health and Retirement Study (HRS) which

[1] See for example the review by Feldstein and Liebman (2002) on the effects of Social Security on labor supply.

[2] http://change.gov/agenda/seniors_and_social_security_agenda/ (accessed October 21, 2013)

provides the most detailed information available on earnings, employment, and retirement for individuals over the age of 50. Given that labor decisions directly affect tax incentives, we exploit legislative tax changes during this time period which had large effects on the tax parameters of some individuals while leaving others relatively unaffected. Similar strategies have been used to study aggregate behavioral effects during this same time period and have found economically meaningful responses (see Giertz (2007), Auten and Carroll (1999), Heim (2009), Giertz (2010), Singleton (2011), Saez et al. (2012)).

There are many reasons to believe that naïve application of elasticities estimated from the aggregate or working age population is problematic. Differences in health status could affect labor supply preferences. Older workers may have different levels of productivity. Moreover, due to pensions and Social Security, this group may be receiving a regular stream of unearned income, thus behaving more similarly to "secondary earners," for whom taxes have been shown to have a larger effect (Saez et al. (2012)).

The labor supply responses to taxes of older workers is complicated by the extensive margin decision of whether to work or retire. We add to the earlier literature by recognizing that tax-based incentives to participate in the labor force are also separately identified due to non-linearities in the tax schedule, aiding joint estimation of the extensive and intensive labor supply effects. Specifically, if self-selection into employment is endogenous to tax incentives, this may bias the intensive margin effects of taxes on labor earnings. While Gruber and Saez (2002) show that changes in the marginal tax rate and after-tax income are independently identified by legislative tax changes, we show that a third variable, after-tax labor income (which represents the after-tax pecuniary incentives to work), is also separately identified. This suggests a powerful new instrument that affects selection into labor force participation, but does not - conditional on the marginal tax rate and after-tax income - independently affect intensive labor supply outcomes. We use this instrument to correct for selection using both the standard Heckman method and a semi-parametric approach. The labor literature has struggled to find appropriate instruments for selection models, often using the number of young children to identify the selection term (e.g., Eissa and Hoynes (2004)). One contribution of this paper is to introduce a new instrument for labor force participation which plausibly satisfies the exclusion restriction required by selection models. This instrument should be useful more broadly.

Our method combines two important tax and labor supply literatures. One literature studies intensive decisions, typically ignoring selection issues. The other literature studies extensive decisions - relating the pecuniary incentives to work to the probability of working (see Meyer and Rosenbaum (2001); Blau and Kahn (2007); Eissa and Hoynes (2004); Eissa et al. (2008)). This literature does not observe the variable of interest for non-workers (the additional after-tax income if the person worked) and must impute this value under the assumption that workers and non-workers are similar conditional on observable characteristics (or, in some cases, that the number of pre-school children is an appropriate instrument for a selection equation). While our method accounts for selection in the intensive equation, it also provides consistent estimates of the variable of interest for the extensive margin equation. The insight is that combining the estimation of the intensive and estimation equations solves many of the problems in each of these literatures.

Weinzierl (2011) finds that age-dependent taxes provide substantial welfare gains even when assuming that tax responsiveness is uniform across age, a necessary assumption in the paper because "empirical evidence is sparse" (p. 1516). Differential elasticities could have important ramifications for such welfare calculations. Furthermore, the labor earnings of older individuals are important outcomes on their own. Labor earnings can supplement savings and Social Security benefits. Butrica et al. (2006) estimate that an additional year of working can increase annual retirement income by 9% with even larger returns for low-income individuals.

We use our estimates of the tax elasticity of labor supply for older individuals to model a policy experiment which eliminates the Social Security payroll tax for older individuals. The payroll tax is an example of a tax that could easily be altered to exclude people eligible for Social Security. Social Security taxes can imperfectly be viewed as a forced savings mechanism for the prime-age working population. However, Reznik et al. (2009) estimate that individuals ages 62-65 can expect, on average, to only receive 2.5 cents per additional dollar paid into the Social Security program. Elimination of the payroll tax at certain ages or after Social Security receipt could drastically change the incentives for individuals to remain in the workforce or retire. We follow Laitner and Silverman (2012) which models the effects of a hypothetical reform that eliminates Social Security payroll taxes after the age of 54. The Laitner and Silverman (2012) paper uses structural estimation of a life-cycle model to conclude that this policy would delay retirement by, on average, one year. However, there is still only limited evidence of the possible impact of the tax code on retirement behavior. Our paper takes a different approach than the structural modeling of Laitner and Silverman (2012) and provides some of the first "natural experiment" evidence of the impact of taxes on the labor supply decisions of older individuals.

Our results suggest that taxes have a statistically significant and economically meaningful impact on labor force participation and retirement decisions for older workers. We find little evidence that individuals ages 55-74 respond on the intensive margin to the marginal net-of-tax rate, the amount that a worker keeps for an additional $1 in earnings. Our intensive labor supply estimates have large confidence intervals and it is difficult to rule out large elasticities. On the extensive margin, we find statistically significant and economically meaningful effects. We find that tax-driven changes in after-tax labor earnings affect the decision to work. Our preferred estimates suggest that the elimination of the employee portion of payroll taxes for our population would decrease the percentage of both men and women dropping out of the labor force by almost 1 percentage point, a 3% decrease. Furthermore, a major advantage of using the HRS is that we observe information about retirement decisions. We find that taxes have a significant effect on the retirement decisions of both men and women.

In the next section, we discuss how this paper is related to previous research in the tax and labor supply literatures. Section 3 describes the data and section 4 includes the model and empirical strategy. We present our results in section 5. Section 6 concludes.

2 Related Literature

This paper intersects the literatures on the elasticity of taxable income (ETI) and labor supply. Because we are especially interested in the extensive margin for the older population, we will highlight this paper's placement in the literature on the effects of tax incentives on working and labor force participation, particularly for secondary earners. First, we discuss the ETI literature.

A rich literature studies the effect of changes in marginal tax rates on taxable income. Auten and Carroll (1999) take advantage of the differential effects that the Tax Reform Act of 1986 had on different households as an instrument for the marginal tax rate. Gruber and Saez (2002) employed a similar strategy but used these tax schedule changes to separately identify the substitution and income effects so that the effect of the marginal tax rate can be interpreted as a compensated elasticity. Singleton (2011) exploits a marriage penalty relief provision in the Economic Growth and Tax Relief Reconciliation Act of 2001. We utilize a specification similar in spirit to Gruber and Saez (2002) and build on this specification by recognizing that tax changes can also separately identify a selection mechanism, which is especially useful for our purposes. The ETI literature typically excludes households with less than a certain threshold of taxable income, assuming that selection is random. Details concerning the ETI literature are further discussed below and are well-summarized in Auten et al. (2008) and Saez et al. (2012).

The ETI literature has never focused on older individuals. Auten and Carroll (1999) limit their sample to ages 25-55. Feldstein (1995) excludes individuals above age 65. The majority of the remaining studies mentioned above estimate aggregate responsiveness.

A related literature has studied the effects of taxes and wages on labor supply, typically measured as hours worked or labor force participation. This literature is summarized in Hausman (1985a) and Blundell and Macurdy (1999) and typically finds that women are very responsive to taxes and wages, while working-aged men are not. In general, this literature does not study older workers and frequently even eliminates them from the analysis. In fact, it is difficult to find any study that examines the effect of taxes on the labor supply of the older segment of the population. Many influential studies have selected on individuals by age, usually using a maximum cutoff of 50, 55, or 60.[3] It is unlikely that these estimates could be extrapolated to an older population.

There is also a small related literature studying the effects of the Social Security Earnings Test including Friedberg (2000), Gruber and Orszag (2003), Song and Manchester (2007), Haider and Loughran (2008), and Gelber et al. (2013). The literature finds mixed results with some evidence that Social Security recipients respond to reductions in benefits on the margin of labor earnings. Recent evidence (Gelber et al. (2013)) finds large behavioral responses. Given that the Earnings Test is not a pure tax since it returns the reduced benefits in an actuarially fair manner, it is possible that individuals' responsiveness to to a pure tax may also be large.

[3] Hausman (1985b), Ziliak and Kniesner (2005), Eissa (1995), Blundell et al. (1998), Blomquist and Selin (2010), Triest (1990), Blomquist and Hansson-Brusewitz (1990) are just a small subsample of these studies.

This project builds on the tax and labor supply literature that considers extensive margin questions, such as labor force participation, as a function of the economic return to work. For those not working, this literature imputes wages or total earnings as if they had worked. Typically, the literature predicts wages or earnings by assuming that workers and non-workers are the same conditional on covariates (see Meyer and Rosenbaum (2001) and Blau and Kahn (2007)). Eissa and Hoynes (2004) and Eissa et al. (2008) also use selection models to impute earnings for non-workers. The excluded variable identifying the selection equation is typically the number of preschool-aged children (a selection instrument used frequently in the labor literature). We build on this framework, but improve on the identification of the selection equation using a new instrumental variable for older individuals. We believe that our strategy offers credibly exogenous variation in the selection mechanism. Furthermore, we test the robustness of our results by using methods that do not impose the strong distributional assumptions of conventional selection models used in the literature.

Our empirical strategy models both the intensive and extensive margins of the labor supply decisions of older individuals. We use each equation to aid estimation of the other equation. This is helpful for two reasons. First, the extensive margin equation provides a useful exclusion restriction to identify the selection mechanism in the intensive margin equation. It is well-known that there are major difficulties finding appropriate instruments for selection into working. Shocks to selection need to impact labor force participation without separately affecting earnings. Our equations provide a natural and plausibly exogenous determinant of selection, which we find has a strong relationship with employment. We believe that this selection shock is in itself a major contribution given the literature's difficulty in finding appropriate instruments for selection.

Second, we model the decision to work as a function of the pecuniary benefits of working, which we define as after-tax labor income. This is the additional amount that the individual would make in after-tax dollars due to working. This variable is not observed for the sample that does not actually work. There are many widely understood challenges to imputing labor income (and the related tax calculations) for individuals without labor income. Our intensive labor supply equation, however, provides consistent predictions of earnings for our sample. We can use this output to generate labor income for the entire sample and calculate the associated taxes. This predicted after-tax labor income allows for estimation of the extensive margin equation.

Our methodology allows for estimation of both intensive and extensive labor supply dimensions to fully understand the impact that taxes can have on the labor supply decisions of our population. This methodology should be useful more generally as it provides a credible means of accounting for selection and estimating the relationship between taxes and labor force participation.

3 Data

We use the Health and Retirement Study (HRS) as our primary data source. The HRS is a panel data set with a rich set of variables including demographics, income variables, and labor supply. The panel nature of the data is helpful for our empirical strategy. The detailed income variables are also crucial for generating tax variables. We use NBER's TAXSIM program (Feenberg and Coutts (1993)) to derive tax rates, tax liability, and labor taxes. We use federal taxes plus one-half of FICA taxes in our calculations. We define "labor taxes" as the additional taxes paid due to the person's own labor earnings. The HRS income, asset, and demographic variables used as inputs to TAXSIM are taken from RAND generated tax data files.[4]

The HRS provides data for each person every two years. We use the 2000-2008 samples in our analysis. Our sample includes everyone ages 55 to 74 who works at some time t linked to their outcomes in period $t + 1$. We use 2 year intervals in our specifications. Individuals appear multiple times in our data so we use clustered bootstraps for inference.

We define retirement in our data by two criteria: (1) no labor earnings in the year and (2) self-declared as "retired" in the following year. We only observe self-reported retirement status in the year of the survey while the labor earning and tax variables all relate to the previous calendar year. We do not view this as a limitation, however, given that we are using the retirement variable as a more permanent indicator of leaving the labor force and observing this status in the following year is consistent with that interpretation. It is well-known that labor supply decreases with age for this population. This pattern is evident in our data and illustrated by the trajectory of working and retirement in Figure 1.

Identification of our tax variables benefits from legislative changes in the tax schedule. There were two key tax reforms during our study period: first, the Economic Growth and Tax Relief Reconciliation Act (EGTRRA) of 2001, which reduced tax rates for each bracket with especially large changes for those with relatively low incomes; second, the Jobs and Growth Tax Relief Reconciliation Act (JGTRRA) of 2003, which also reduced tax rates, primarily focusing on high income households. The tax reductions can be seen in our sample in Figure 2, which graphs (a) the simulated marginal tax rate and (b) the fraction of labor earnings that an individual keeps. The figure holds the sample constant over time for ages 55-74. These tax changes were previously studied in Auten et al. (2008).

As Figure 3 shows, this is a population that tends to experience relatively large

[4]See RAND Contributions at http://hrsonline.isr.umich.edu/ for tax files for 2000.

year-to-year changes in labor supply incentives, mainly due to own and spousal behavioral changes such as retirement. This figure emphasizes the importance of studying the impact of tax changes on older workers, given that the large changes that they face.

We present summary statistics for our data in Table 1.

4 Model and Empirical Strategy

4.1 Theoretical Framework

We introduce a basic theoretical framework to guide our empirical strategy. We examine one individual maximizing a utility function with consumption and labor as its two arguments. The budget constraints include non-labor income (assumed exogenous in this model) and tax liability which is a function of labor earnings and non-labor income. The utility function includes a parameter related to a cost of working at all and is similar in spirit to the model found in Eissa et al. (2008).

$$\max_{c,\ell} U(c,\ell) - \mathbf{1}(\ell > 0)q \quad \text{s.t.} \quad c = \ell + y^o - T[\ell + y^o]$$

where c represents consumption, ℓ is labor earnings (which we assume requires additional effort), $T[z]$ is total tax liability given income z, and q represents some fixed psychic cost of working. There are alternative ways to model participation which will generate similar results, but we include this term (which can vary by individual) in the utility function which is only relevant to those that earn any money by working.

If we assume an interior solution, then the first-order conditions are

$$\frac{U_c}{U_\ell} = -(1-\tau),$$
$$c = \ell + y^o - T[\ell + y^o].$$

The insight from these conditions is the same as the derivation found in Gruber and Saez (2002). Changes in labor earnings (conditional on working) are a function of changes in $1-\tau$ and changes in $\ell + y^o - T[\ell + y^o]$ due to changes in $T[\cdot]$.

Individuals may not work, however, and this decision may be related to the tax schedule. In the above conditions, we can solve for c^* and ℓ^*. Then, we can compare the utility from working to utility from not working. Let us look at an individual that is indifferent between working and not working:

$$U(\ell^* + y^o - T[\ell^* + y^o], \ell^*) - q = U(y^o - T[y^o], 0).$$

For fixed preferences and ℓ^*, note that changes in $\ell^* + y^o - T[\ell^* + y^o]$ and $y^o - T[y^o]$ drive labor force participation changes. These are dependent on $T[\cdot]$ and this dependence is a function of y^o.

4.2 Empirical Strategy

Our empirical strategy models and estimates the impact of taxes on both the intensive and extensive margins of labor earnings for older workers, using the insights of the above theoretical framework. We start with the intensive labor supply equation:

$$\ln L_{i,t+1} - \ln L_{it} = \alpha_t + X'_{it}\delta + \beta^I \left[\ln(1-\tau_{i,t+1}) - \ln(1-\tau_{it})\right] \qquad (1)$$
$$+\theta^I \left[\ln\left(y_{it} - T_{t+1}(y_{it})\right) - \ln\left(y_{it} - T_t(y_{it})\right)\right] + f_t(L_{it}) + (\epsilon_{i,t+1} - \epsilon_{it})$$

where L represents labor income, τ is the individual's marginal tax rate, y is pre-tax income and $T(y)$ is the tax liability for income y. $\ln\left(y_{it} - T_{t+1}(y_{it})\right) - \ln\left(y_{it} - T_t(y_{it})\right)$ is the change in the log of after-tax income and will be discussed in detail later. X is a vector of covariates. This specification models intensive labor supply decisions as a function of the marginal net-of-tax rate and shifts in the individual's budget constraint, the substitution and income effects.

We are also interested in understanding extensive margin decisions for both the decision to work and retire. Let y^o be total pre-tax income other than own labor earnings, such as spousal labor earnings. The decision to work at all should be a function of total after-tax income if the person works minus total after-tax income if the person does not work. We model the extensive margin in the following manner:

$$P\left(\text{Work}_{i,t+1} = 1\right) = F\Bigg(\phi_t + X'_{it}\gamma + \beta^E\Big\{\ln\left[L + y^o - T(L+y^o)\right] - \ln\left[y^o - T(y^o)\right]\Big\}_{i,t+1} \qquad (2)$$
$$+\theta^E\Big\{\ln\left(y_{it} - T_{t+1}(y_{it})\right) - \ln\left(y_{it} - T_t(y_{it})\right)\Big\} + f_t(L_{it}) + \nu_{it}\Bigg)$$

We study the decision to work (or retire) in period $t+1$ for all individuals working in period t. $T(L+y^o)$ represents the tax liability if the individual works and earns labor income L. $T(y^o)$ is the individual's tax liability if the individual does not work. The implication of equation (2) is that we model the extensive margin as a function of the additional money that the individual receives if s/he works. We use the change in log of after-tax income as our measure of the pecuniary return to working. A person that works has log of annual after-tax income $\ln[L + y^o - T(L + y^o)]$ while not working results in $\ln[y^o - T(y^o)]$. These are the variables suggested by our theoretical framework. We will refer to this variable as "after-tax labor income." A similar variable is referred to as "effective net-of-tax share" in Gelber and Mitchell (2011). We also model the probability of retiring in the same manner as equation (2). Note that L and, consequently, $T(L + y^o)$ are not observed for non-workers.

Equations (1) and (2) pose a few identification challenges. First, OLS will not provide consistent estimates. Changes in tax rates and after-tax income are functions of changes in labor earnings. In equation (2), individuals with higher L may be more likely to work for reasons unrelated to after-tax earnings. Second, we do not observe $L_{i,t+1}$ for individuals that do not work in period $t + 1$. Consequently, we can only estimate equation (1) for a selected sample of individuals. And, we do not observe $[\ln(L + y^o - T(L + y^o)) - \ln(y^o - T(y^o))]_{i,t+1}$ in equation (2).

To address the issue of endogeneity, we create a set of instruments that exploit plausibly exogenous shocks to the tax variables. These instruments are discussed in detail in Section 4.2.1 below. To adjust for selection, we follow standard methods in the labor supply literature with three important differences that improve the consistency of the estimates. First, we take advantage of the panel nature of the HRS data to incorporate previous income in our earnings predictions. The literature typically predicts earnings on the basis of only cross-sectional variables. An assumption that selection is random conditional on observables is more plausible in our data set compared to the same assumption in cross-sectional data. Second, we use a new instrument for selection that plausibly satisfies the exclusion restriction. Conventional selection instruments used in the labor supply literature (such as number of children) have numerous identification concerns and are not relevant in the context of older workers. Third, we also consider semi-parametric alternatives to Heckman selection methods commonly employed in the literature so that we do not identify off of strong distributional assumptions.

We utilize both equations (1) and (2) and allow each equation to inform the other. Insights from the extensive margin equation can provide information on selection in the intensive equation by suggesting a natural excluded variable. Individuals who do not work are not making intensive labor decisions and are unlikely to be responding to the marginal net-of-tax rate at the first dollar of labor earnings. These individuals are excluded from the estimation of the intensive labor supply equation, which may create issues related to a selected sample. The extensive labor supply equation allows us to address this selection concern and obtain consistent estimates. The extensive equation provides a necessary exclusion restriction to properly address selection.

If the tax schedule becomes more generous, individuals with less proneness to labor supply (relative to their previous labor supply) will enter the labor force. These people will likely work less and, consequently, we will associate generous tax schedules (higher marginal net-of-tax rates) with lower intensive labor supply decisions. Accounting for selection, then, should increase our estimates of β^I.

4.2.1 Instruments

Our two equations of interest include three different tax-related variables: the marginal net-of-tax rate, after-tax income, and after-tax labor income. These are all potentially endogenous as taxes are a function of labor income. We implement an instrumental variables strategy with shocks to all three of these tax-related variables. The last instrument for after-tax labor income will also be used to identify the selection equation.

The tax schedule underwent two major reforms (EGTRRA and JGTRRA) during the study period, as shown previously. We use these reforms for exogenous variation. We employ an instrumental variables strategy similar to the one found in Gruber and Saez (2002). We take this procedure one step further, however, in attempt to reduce concerns related to wage trends and mean reversion. Gruber and Saez (2002) show that their results are sensitive to controls for initial taxable income. Our implicit experiment is to compare people with the same initial labor income but different other initial characteristics, such as spousal labor income. When the tax schedule changes, people with the exact same labor income may be differentially affected. Our variation originates from the interaction of tax schedule changes and initial non-labor income. We also control for initial labor earnings to account for possible trends and mean reversion.

We implement this strategy by creating the Gruber and Saez (2002) instruments for the tax variables. We calculate tax liability, the marginal tax rate, and labor taxes for each individual in the initial year t. We also calculate their predicted tax variables in year $t + 2$. Consequently, variation in tax changes originates solely from legislative tax schedule changes, not individual-level behavioral changes. Gruber and Saez (2002) calculate predicted tax variables by adjusting all income measures for inflation. Our population experiences relatively large year-to-year changes in labor income that are not accurately reflected, even on average, by an inflation adjustment. Instead, we actually predict each person's earnings in the next year using covariates and initial labor income. Note that we include these exact same variables in all of our specifications to ensure that no identification originates from this prediction. We find that the first stage predictions are much better when we use a more accurate value for next year's labor earnings. There is little cost in obtaining a more accurate prediction of earnings growth in our sample.

The Gruber and Saez (2002) method uses their predicted changes as the instruments and then controls flexibly for initial income so that identification originates from differential changes not captured by the flexible controls. We take this approach one step further. We regress each of the predicted tax change variables on initial income measures and use the predictions from these regressions as instruments. The benefit of this approach is that we can know the exact functional form in which the instruments vary. We can then control for initial labor income using the same function. We estimate

$$\begin{aligned} \text{Gruber-Saez Instrument}_{it} &= \tilde{\alpha}_t + X_{it}'\tilde{\delta} + f_t(L_{it}) \\ &+ g_t(\text{Initial Other Income Measures}_{it}) + \tilde{\epsilon} \end{aligned} \tag{3}$$

where X represents a set of covariates based on age and education. In all of our specifications, we will control for $f_t(L_{it})$ and X. Note that we have allowed the impact of initial labor income to have a different effect in each year. The advantage of this method is that it reduces concerns that mean reversion and wage trends are potentially biasing our results.

It is important to note that our three separate shocks vary independently. A tax change may shift the marginal tax rates of people in different tax brackets differentially. Given different distances from the nearest kink point, the tax change will have different impacts on people's after-tax income. This point was made by Gruber and Saez (2002) and can be see in Figure 4. This figure shows the nonlinear budget set created by the tax schedule with two brackets. In period $t = 1$, the tax rate for the top bracket is reduced. Person A and Person B experience different changes in their marginal tax rates. Person B and Person C face the same marginal tax rate change but difference changes in after-tax income, even proportional to initial income.

Our addition to this model can be seen in Figure 5. In this figure, we focus on point C in Figure 4. We study two different people with the same total taxable income. C_1^{NW}

represents the non-wage income of C_1. C_2^{NW} represents the non-wage income of C_2. The top tax rate decreases. Note that the pecuniary incentives to work (labeled P_1) have increased much more for C_1 than for C_2 (equal to P_2). This result occurs because C_2 benefits from the tax decrease regardless of labor force participation while C_1 benefits only through labor force participation. Consequently, we can find two people that look identical in the Gruber and Saez (2002) framework but experience different shocks due to legislative tax schedule changes in the pecuniary incentives to work. We exploit this additional source of variation.

Finally, people with different non-labor income will experience differential changes in after-tax labor income.

In the end, we have three instruments:

1. Predicted change in log of the marginal net-of-tax rate: $\hat{\Delta}\text{MTR}_{it}$
2. Predicted change in log of after-tax income: $\hat{\Delta}\text{ATI}_{it}$
3. Predicted change in log of after-tax labor income: $\hat{\Delta}\text{ATLI}_{it}$

4.2.2 Selection

Both of our equations of interest are potentially affected by selection. Even with randomly-assigned tax schedules, selection could bias equation (1) since we do not observe $L_{i,t+1}$ for people who do not work in period $t+1$. Our primary concern is that these non-workers may have systematically different $\epsilon_{i,t+1}$, implying that $E\left[(\epsilon_{i,t+1} - \epsilon_{it})|\hat{\Delta}\text{MTR}_{it}, \hat{\Delta}\text{ATI}_{it}, \tilde{X}_{it}, \text{Work}_{i,t+1} = 1\right] \neq 0$ where $\tilde{X}$ is our set of covariates, including all fixed effects, in equation (1). Imagine if these non-workers were forced to work. We might guess that they would work less on average, given observable characteristics, than the population that actually does work. This in itself is not problematic if the decision to work or not work is independent of our shocks to $\ln(1-\tau)$ and $y - T(y)$. However, we may think that changes in incentives due to taxes influence the decision to earn labor income. While the tax variables explicitly included in equation (1) may not affect this decision, they are likely correlated with the variables that do. For example, an increase in taxes may lead to lower after-tax labor income and a higher marginal net-of-tax rate. The lower after-tax labor income will induce people to drop out of the labor market in period $t+1$, excluding them from the sample for estimation of equation (1). Since this is correlated with higher taxes and, specifically, higher marginal net-of-tax rates, we may have a selected sample.

Selection bias in equation (1) creates further problems for our empirical strategy since we use this equation to predict $L_{i,t+1}$ to use for equation (2). This procedure is similar in spirit to the methodology found throughout the labor supply literature studying extensive margin questions. Meyer and Rosenbaum (2001), Blau and Kahn (2007), Gelber and Mitchell (2011) impute earnings for non-workers based on covariates, assuming that workers and non-workers are the same after conditioning on observable characteristics. These studies predict earnings by estimating equations such as:

$$\ln(L_i) = \alpha + X_i\beta + \epsilon_i. \tag{4}$$

It is also common to account for selection bias in these equations using a Heckman (1979) correction term and adding the inverse Mills ratio to equation (4). This method is used in Eissa and Hoynes (2004); Eissa et al. (2008); and - as a robustness check - Gelber and Mitchell (2011). These studies include an additional instrument for selection such as number of children or preschool-aged children.

We improve on this method in several ways for our context. First, we employ panel data, providing us with data on previous earnings. Instead of assuming that people who do not work have the same earning power and labor supply preferences as non-workers, we observe previous earnings in our sample. We can assign people the same earnings as the previous year, making adjustments to the new earnings based on observable characteristics. This method reduces concerns of selection issues by accounting for this individual fixed effect and not assuming that workers and non-workers are the same. Furthermore, note that equation (4) does not use information about taxes to predict earnings. If taxes (or, in other contexts, the policy variables of interest) affect earnings, then is potentially problematic to predict earnings without these policy variables. We model changes in earnings as a function of covariates and changes in the tax variables.

Second, we have a plausibly exogenous instrument for selection. Even panel data may not be adequate to impute earnings since people who stop working may have experienced different shocks to productivity or labor supply preferences than those who continue working. Our framework, however, provides a natural instrument for selection. The literature has used family characteristics such as the number of preschool-aged children as a cross-sectional shock to participation in the labor force. This instrument poses problems because the selection method requires an instrument that affects participation without separately affecting labor earnings. However, people with children are likely different than people without children on dimensions such as productivity. Furthermore, children may also change intensive labor supply preferences or occupational choice, affecting wages or earnings independent of selection.

Our intensive and extensive labor supply equations offer a useful exclusion restriction. The extensive equation includes a variable (after-tax labor income) that is not included in the intensive equation, suggesting that the instrument for this variable is an appropriate instrument for selection into equation (1). Thus, we use predicted after-tax labor income as an exogenous shock to employment. To illustrate, take two people currently in the labor force with the same pre-tax labor income. One person's labor taxes are predicted to decrease by $1000, while another's are predicted to stay the same. We predict that the first person is more incentivized to remain in the labor force.

Of course, changes in after-tax labor income are correlated with changes in after-tax income. Changes in taxes shift the budget constraint and could have important income effects, including impacts on intensive labor supply preferences. Consequently, our selection instrument is only orthogonal to the disturbance term in our intensive labor supply equation *conditional on changes in after-tax income*. It is important to highlight that these two terms - changes in labor taxes and changes in total taxes - are independently identified due to nonlinearities in the tax schedule. Thus, we use our shocks to labor taxes as an instrument for selection while conditioning on the change in total taxes (and the change in the marginal tax rate). Since these are not collinear, this separates the selection shock from an income effect.

Our strategy depends on people responding to our selection instrument. Predicted changes in after-tax labor income must be correlated with employment probabilities. This is an empirical question and we will show that there is a relationship. A main insight of our empirical strategy is that, conditional on shocks to after-tax income, shocks to after-tax labor income are a useful instrument for selection. This strategy should be useful in contexts beyond this paper. It is commonly argued that selection models are difficult to utilize because of the standards required for the selection mechanism. We suggest a credibly exogenous instrument for selection, which can be used for broader research questions.

Finally, we also account for selection using a method that does not rely on strong distributional assumptions found in the Heckman (1979) method. While we present results that condition on an inverse Mills ratio, we also report results using semi-parametric methods. Systematic selection into employment implies that

$$E\left[(\epsilon_{i,t+1} - \epsilon_{it})|\hat{\Delta}\text{MTR}_{it}, \hat{\Delta}\text{ATI}_{it}, \tilde{X}_{it}, \text{Work}_{i,t+1} = 1\right] = \lambda(\mathbf{W}_i'\zeta) \qquad (5)$$

where W includes our instruments for the intensive labor supply equation, the selection instrument, and all exogenous variables in equation (1). We do not assume any functional form for $\lambda(\cdot)$ and instead use a series approximation, as suggested in Newey (2009). We estimate the selection equation using a monotone rank approach introduced in Cavanagh and Sherman (1998), which requires no distributional assumptions to obtain consistent estimates (up to scale).

4.2.3 Income Effect

Before discussing the practical implementation of our empirical strategy, we will address the after-tax income term in our two equations. Gruber and Saez (2002) discuss separate identification of the substitution and income effects, allowing for interpretation of the coefficient on the change in log of marginal net-of-tax rate variable as a compensated elasticity. This was an important innovation. The Gruber-Saez after-tax income variable is $\ln(y_{i,t+1} - T_{t+1}(y_{i,t+1})) - \ln(y_{it} - T_t(y_{it}))$ and the paper models changes in taxable income as a function of this variable.

We note that this structural relationship may be problematic. Households respond to shifts in their budget constraints, but this variable assumes that they are responding to their final after-tax income, which includes the response to this budget constraint shift and changes in the marginal tax rate. This point is discussed briefly in Powell and Shan (2012). An example can help make this point clearer. Imagine an individual with income $\tilde{y}$ in period 1 and no tax liability. The tax code changes such that this person is given a lump sum equal to $\tilde{y}$ in period 2. In response, the person changes behavior and earns income equal to 0. This is a strong income effect (i.e., income targeting) resulting from the outward budget constraint shift. However, the Gruber-Saez specification models the change in earnings $(-\tilde{y})$ as a function of the final change in after-tax income, which is 0. This is because using actual after-tax income changes includes the response to the tax change as well. The variable capturing the income effect should not include this endogenous response, only the budget constraint shift.

We include an after-tax income term because our shocks to the marginal net-of-tax rate also shock the individual's budget constraint. We want to instrument the marginal net-of-tax rate with an instrument that is exogenous to changes in the budget constraint so we explicitly include this term. Our after-tax income term is the expected change in after-tax income given legislative changes in the tax schedule and predicted changes in labor earnings. We hold everything constant and estimate the budget constraint shift due only to these changes. This separates the substitution and income effects as before, and we interpret our estimates as compensated elasticities.

The important consideration is that our other instruments are orthogonal to income effect considerations conditional on the after-tax income variable or, alternatively, the instrument for after-tax income. We expect the coefficient on our income variable to be negative in the intensive labor supply equation.

4.3 Implementation

We implement our empirical strategy in three steps. Given that the labor supply literature has consistently found that men and women respond to labor market incentives in different ways, we perform all estimations separately by gender. We also include several covariates in our estimation. We create cells based on age and education of individual i and the age of individual i's spouse. There are 4 education categories (less than high school, high school graduate, some college, college graduate) and 5 age group categories (55-60, 61-64, 65-67, 68-70, 71+) for a total of 20 cells.[5] We include an indicator variable for each cell. We

[5] The results are not meaningfully affected by the use of different age categorization.

also include indicator variables based on spouse's age (under 55,[6] 55-60, 61-64, 65-67, 68-70, 71+). X also includes functions of initial labor income (L_{it}), which are allowed to have different effects in each year. Note, again, that the initial labor income function is the same used to predict the tax instruments. We also include year fixed effects which are allowed to have different effects based on initial marital status.

4.3.1 Step 1:

In the first step, we model the selection mechanism. When we report estimates that do not account for selection, this step is skipped. It is only appropriate to include exogenous variables in the selection equation so we estimate a reduced form version of equation (2). Furthermore, we must include all of the instruments used in the intensive labor supply equation. In the end, we estimate

$$P\left(\text{Work}_{i,t+1}=1\right)=F\Big(\phi_t+X_{it}'\gamma+\beta_1\hat{\Delta}\text{MTR}_{it}+\beta_2\hat{\Delta}\text{ATI}_{it}+\beta_3\hat{\Delta}\text{ATLI}_{it}+f_t(L_{it}),\eta\Big) \quad (6)$$

The predictions provided by equation (6) are used as selection adjustments for the intensive equation. We do this in two different ways. First, we assume that $F(\cdot)=\Phi(\cdot)$, estimate equation (6) using a probit regression, and form an inverse Mills ratio as discussed in Heckman (1979). This method is frequently used in the literature.

Second, we use a monotone rank estimator introduced in Cavanagh and Sherman (1998). This estimator does not estimate $F(\cdot)$ but provides $\sqrt{n}$-consistent estimates up to scale of the coefficients in the argument of the function. We then predict the index function, which we denote as $W_{it}'\hat{\zeta}$. The selection correction term is a function of this index and we follow the method of Newey (2009) by approximating this term with a spline using $W_{it}'\hat{\zeta}$.[7] The

[6] While our sample includes people 55+, a respondent's spouse may be younger.

[7] Newey (2009) recommends the use of a spline over a power series.

advantage of this approach is that the maximum rank estimator requires no distributional assumptions to obtain consistent $\hat{\zeta}$. We will refer to this as the semi-parametric selection correction.

We coded the monotone rank estimator, generated initial values through a probit regression, and maximized the objective function using an MCMC optimization algorithm (see Chernozhukov and Hong (2003)).[8] Standard errors are generated using a clustered bootstrap.[9]

4.3.2 Step 2:

The second step estimates the intensive labor supply equation. Because of selection, we estimate the following:

$$\ln L_{i,t+1} - \ln L_{it} = \alpha_t + X_{it}'\delta + \beta^I \left[\ln(1-\tau_{i,t+1}) - \ln(1-\tau_{it})\right] \quad (7)$$
$$+\theta^I \left[\ln\left(y_{it} - T_{t+1}(y_{it})\right) - \ln\left(y_{it} - T_t(y_{it})\right)\right] + \lambda(W_{it}'\hat{\zeta}) + f_t(L_{it}) + (\mu_{i,t+1} - \mu_{it})$$

In practice, we use an inverse Mills ratio or a 5-piece spline in $W_{it}'\hat{\zeta}$. We use $\hat{\Delta}\text{MTR}_{it}$ and $\hat{\Delta}\text{ATI}_{it}$ as instruments.

Because of estimation of the selection terms, we bootstrap Steps 1 and 2 to account for the inclusion of an estimated term in equation (7). Since each individual may be included multiple times in our data, we use a clustered bootstrap.

We should highlight that 2SLS includes the selection adjustment terms in the first stage as well. This has practical importance in our strategy. Notice that for individuals not working, we do not observe the change in their marginal net-of-tax rate if they had actually

[8]Similar results are found using a Nelder-Mead optimization algorithm.
[9]Subbotin (2007) discusses properties of bootstrapping rank regression estimates.

worked.[10] We predict this variable from the first-stage regression in the same way that we will predict labor earnings for period $t+1$.

We use 2SLS to obtain consistent estimates. Once we have consistent estimates for equation (7), we can predict $\ln L_{i,t+1} - \ln L_{it}$ for our entire sample. This includes people who did not work in period $t+1$. However, it is important to note that when using the Newey (2009) method, the constant term is not separately identified from the selection correction term. A method to estimate the constant term was introduced in Heckman (1990). Schafgans and Zinde-Walsh (2002) discuss consistency and asymptotically normality of this estimator. We implement this estimator to derive the constant term.

We predict the change in earnings for the entire sample using the estimated coefficients in equation (7), the estimated constant term, and the imputed change in $\ln(1-\tau_{i,t+1}) - \ln(1-\tau_{it})$. In other words, we have consistent predictions of the tax variables and the coefficients to predict $\ln L_{i,t+1} - \ln L_{it}$ for everyone in our sample. We use this to predict Ł$_{i,t+1}$ using

$$\widehat{L}_{i,t+1} = \exp(\ln L_{it} + \overbrace{\ln L_{i,t+1} - \ln L_{it}}^{\text{Predicted}}). \tag{8}$$

4.3.3 Step 3:

Once we have $\widehat{L}_{i,t+1}$, we can calculate $T(\widehat{y^o + \hat{L}})$ using NBER's TAXSIM program. Then, we estimate

$$P\left(\text{Work}_{i,t+1} = 1\right) = F\Big(\phi_t + X_{it}'\gamma + \beta^E \Big[\ln\Big(y^o + \widehat{L} - T(y^o + \widehat{L})\Big) - \ln\Big(\widehat{L} - T(\widehat{L})\Big)\Big]_{i,t+1} \tag{9}$$
$$+\theta^E \left[\ln\left(y_{it} - T_{t+1}(y_{it})\right) - \ln\left(y_{it} - T_t(y_{it})\right)\right] + f_t(L_{it}) + \nu_{i,t+1}\Big).$$

[10] Since our endogenous after-tax income measure is the shift in the budget constraint, we do observe this variable.

We estimate this equation using the IV-Probit technique introduced in Newey (1987) with instruments

$$\hat{\Delta}\text{ATLI}_{it} \quad \text{and} \quad \hat{\Delta}\text{ATI}_{it}$$

We use a clustered bootstrap for all 3 steps to generate all confidence intervals in this paper. We present 95% confidence intervals.

5 Results

5.1 Estimates

Our empirical strategy requires three different steps. First, we estimate the selection mechanism and examine the impact of our labor taxes instrument on employment. This equation is the reduced form equivalent of the extensive labor supply equation. Second, we estimate the intensive labor supply equation to measure the effect of the marginal tax rate on labor earnings. Finally, we estimate the extensive labor supply dimension to obtain the structural estimates of the impact of after-tax labor earnings on labor force participation. For these last two steps, we include estimates with (1) no selection adjustment; (2) a Heckman (1979) adjustment method; (3) a semi-parametric adjustment method. These adjustments are marked in the tables. We should highlight that no selection term is actually included in the extensive margin estimation. Instead, the different selection adjustments refer to the method used to generate the earnings (and the corresponding tax variables) used in the extensive equation.

In Table 2, we present results for the reduced form selection mechanism. Note that all exogenous variables are included in these regressions. Adjusting for selection in the intensive labor supply equation requires an excluded variable that is correlated with selection (i.e., participation in the labor force). Technically, it is possible to identify purely off of distributional assumptions using Heckman (1979) and a probit regression. Even with an excluded variable, some of the identification will originate from distributional assumptions. The semi-parametric method that we employ, however, does require an excluded variable that predicts labor force participation. All identification of the selection term originates from our selection instrument.

Columns (1) and (3) present the results from probit regressions and show that there is a statistically significant relationship between our selection instrument and employment for women and men. Columns (2) and (4) are the results of a maximum rank estimator introduced in Cavanagh and Sherman (1998). The coefficients are only identified up to scale. We normalize all coefficients so that the norm is equal to 1. It appears that our selection equation is well-identified. The results also suggest that labor taxes impact the probability of working. Estimating the intensive labor supply equation and generating the endogenous "after-tax labor income" is necessary for parameterization of this relationship, but we can surmise from Table 2 that there is likely a positive relationship.

Table 3 presents the first stage results for the intensive labor supply equation. We present Partial F-statistics, which measure the strength of the instrument on the endogenous variable. We find strong relationships.

Table 4 includes the results from IV estimation of the intensive labor supply equations. We can interpret the coefficients on the marginal net-of-tax rate as compensated elasticities. The confidence intervals in this table are rather large, likely due to the relatively high variance of labor earnings for this population. Our estimates do not allow us to rule out that older workers do not respond to changes in taxes on the intensive margin. We are also unable to rule out rather large elasticities. We find important income effects on the intensive margin for men.

Finally, we use the results found in Table 4 to generate predicted labor earnings. As detailed earlier, we then calculate the additional taxes that the individual would pay had s/he earned this income. We use this variable in our extensive margin equation and estimate with instrumental variables.

Table 5 shows significant effects on the extensive margin dimension. We present IV-Probit results using a control function approach found in Newey (1987). We find large effects for both men and women and these effects are statistically significant. The method used to predict labor earnings does not appear to matter for the magnitude of the estimates. The Appendix includes results using an IV linear probability model (see Table A.1). These also find significant effects. The Table 5 estimates are difficult to interpret for several reasons and we will discuss interpretation in section 5.2.

When we focus on retirement in Table 6 (linear probability model results found in Table A.2), we also find strong effects, suggesting that an increase in after-tax labor income delays retirement. Note that, for women, it appears that the entire effect on increasing the probability of working in Table 5 appears to result from a reduction in retirement. For men, only half of the effect an be attributed to retirement.

5.2 Discussion

We find little evidence that older workers are responsive on the intensive margin. Due to the imprecision of the estimates we cannot rule out more elastic relationships either.

The key finding in our analysis is the level of responsiveness of older workers to the additional taxes that they would have to pay if they worked. The parameter estimated in Table 5 is hard to interpret, however, so we consider the following policy experiment, inspired by Laitner and Silverman (2012). We study how extensive margin work decisions would respond to the elimination of the employee portion of the FICA taxes for our sample. We assume that an equivalent lump sum tax is levied on each person in a way that that does not distort labor decisions. Table 7 shows the results of this simulation.

In our sample, 25% of working women and 26.2% of working men will not earn positive wages 2 years later. Elimination of the employee FICA taxes will reduce this percentage by 0.7 percentage points for women and 0.8 percentage points for me. This is a 2.8% decrease for women and an 3.1% decrease for men. We estimate that there would be a 0.8 percentage point drop in retirement for women and a corresponding 0.4 decrease in retirement for men, a 5.7% and 2.4% decrease, respectively.

The retirement changes are especially interesting because they suggest that tax incentives can have dynamic effects on labor supply by causing older workers to drop out of the labor force in a more permanent manner. Understanding the permanence of these behavioral responses will be studied in future research.

The magnitudes of these effects are large. The Song and Manchester (2007) study finds little evidence of any effect of the elimination of the Earnings Test in 2000 for ages 65-69 on labor force participation. Our estimates are larger than their largest estimates on labor force participation.

6 Conclusion

This paper models both the intensive and extensive margins of labor supply, using each dimension to provide more accurate and consistent estimation of the other. We model the intensive labor supply as a function of the marginal net-of-tax rate and shifts in the budget constraint. Extensive labor supply is a function of the monetary benefit of working as measured by after-tax labor earnings. Both of these equations are difficult to estimate even with proper instrumental variables. The extensive labor supply equation, however, provides a natural exclusion restriction to account for selection in the intensive labor supply. This instrument is, to our knowledge, new to the labor supply literature, which has consistently suggested that selection instruments meeting the required conditions are difficult to find. The intensive labor supply equation provides a crucial variable to the extensive equation, allowing us to generate consistent estimates for that equation as well.

We find statistically significant and economically meaningful effects of taxes on the labor force participation behavior of older workers. We find evidence that women and men retire in response to high taxes. The labor supply and tax literatures rarely study the older segment of the population, frequently excluding them from the analysis. This paper fills a large gap in these literatures and provides important estimates about the incentivizing potential of the tax code. Age-specific tax reductions could cause this population to remain in the labor force and delay retirement.

Figures

Figure 1: Working and Retirement by Age

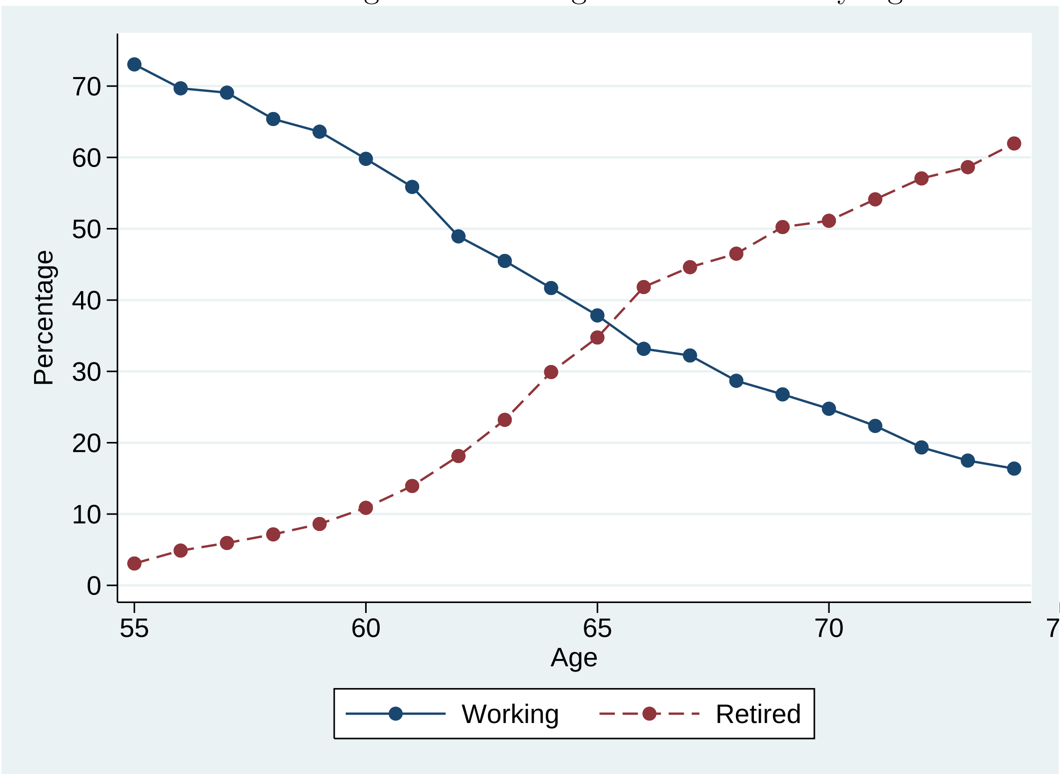


Figure 2: Simulated Federal Tax Rates in HRS

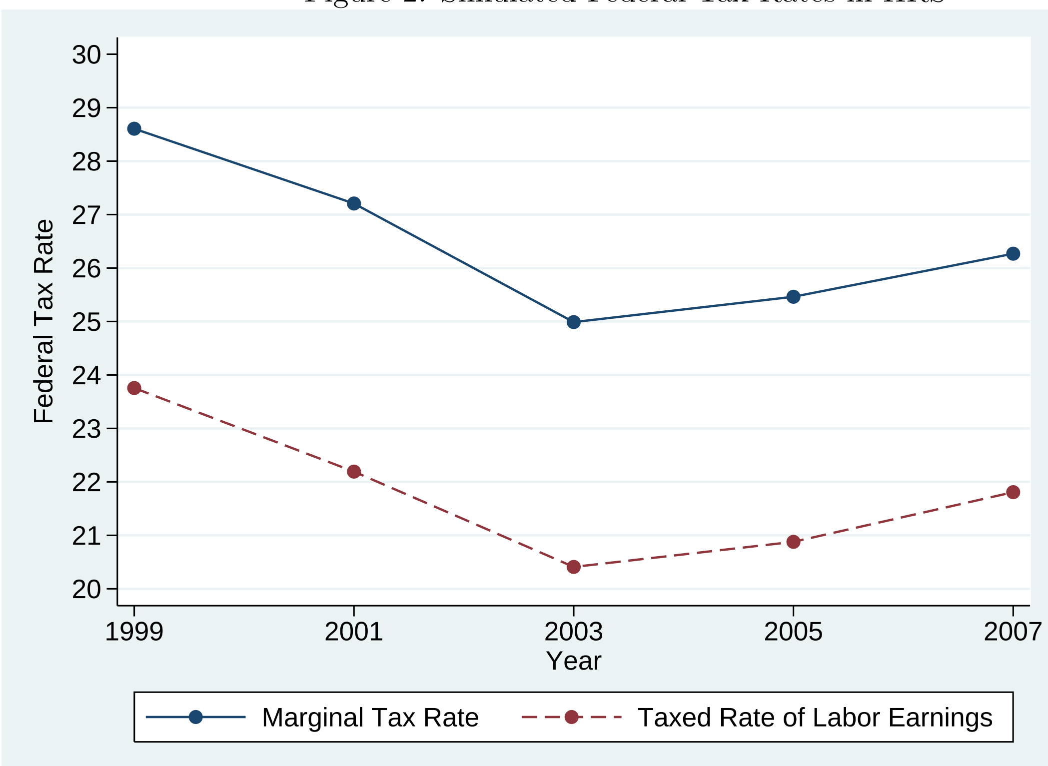


Figure 3: Change in the Marginal Tax Rate Between Year t and $t+2$ by Age

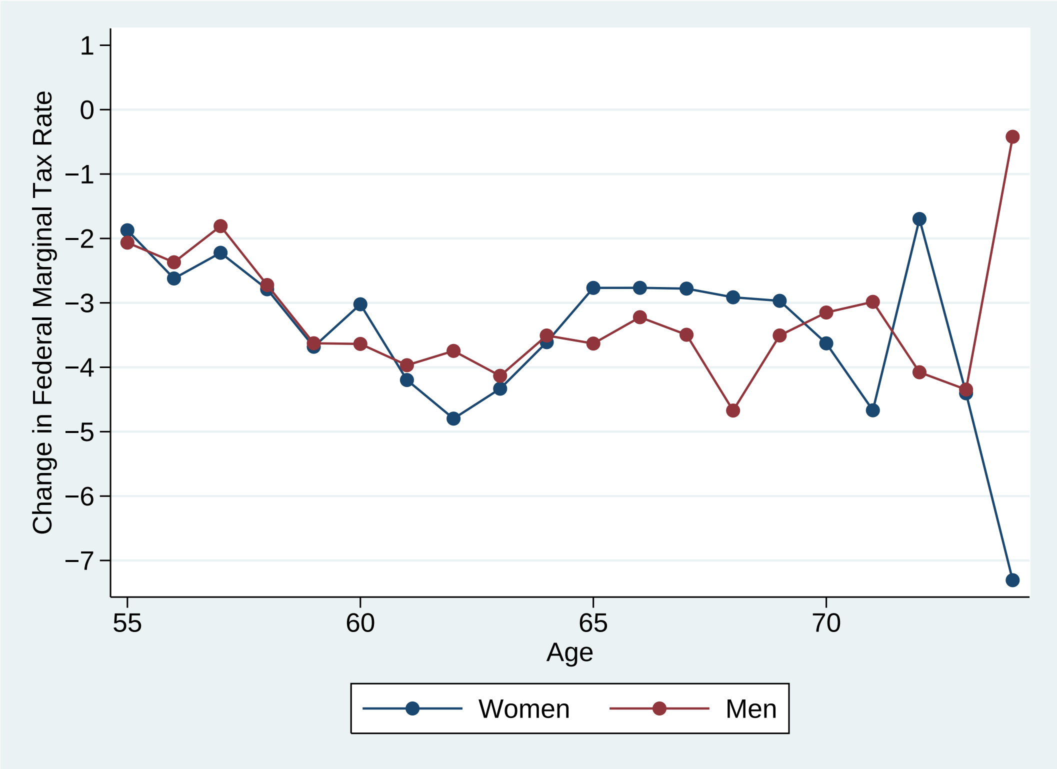


Figure 4: Marginal Tax Rates and After-Tax Income

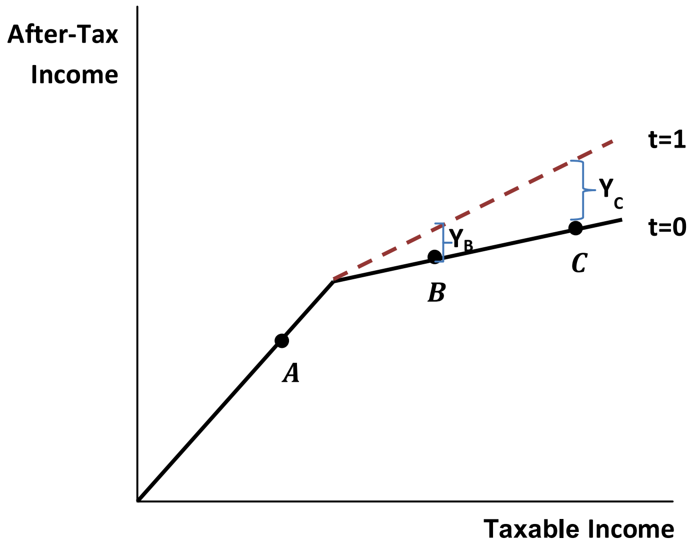
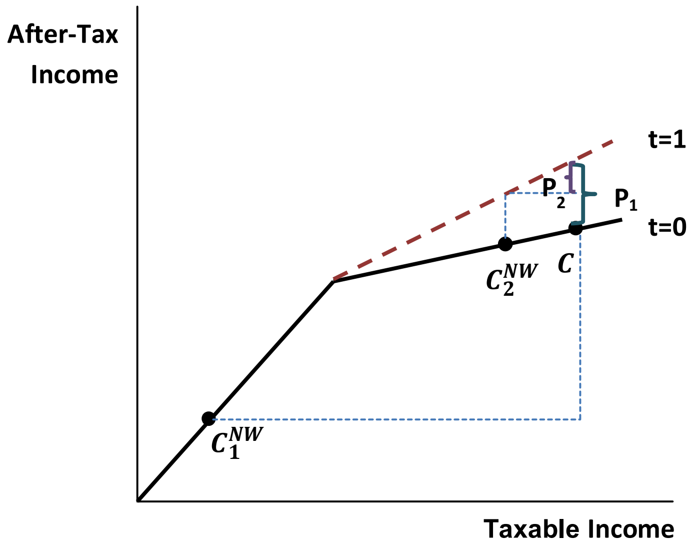


Figure 5: After-Tax Labor Income



Tables

Table 1: Descriptive Statistics

Sample:	Overall Sample
Demographics	
Age	62.1
Less than HS	0.145
HS Grad	0.359
Some College	0.236
College Grad	0.260
Male	0.475
Labor Outcomes	
Labor Earnings	35,213.59
Spouse's Labor Earnings	14,514.87
Total Income	81,700.47
Employed in Next Period	0.744
Wages in Next Period (if Employed)	36,809.07
Retired in Next Period	0.152
Tax Variables	
Marginal Tax Rate	27.3
Tax Liability	17186.16
Change in MTR (t and $t+2$)	-3.3
Change in Tax Liability (t and $t+2$)	-2,835.23
Observations	16,559

Table 2: Selection Equation, Reduced Form

Dependent Variable:	I(Employed)			
	Women		Men	
	(1)	(2)	(3)	(4)
Predicted Δ in ln($1-$ Marginal Tax Rate)	0.145	-0.089	0.371	0.271
	[-0.196, 0.487]	[-0.406, 0.228]	[0.096, 0.647]	[0.369, 0.824]
Predicted Δ in ln(After-Tax Income)	-0.373	-0.314	-0.453	-0.455
	[-0.555, -0.191]	[-0.505, -0.124]	[-0.668, -0.238]	[-0.675, -0.236]
Predicted Δ in ln(After-Tax Labor Income)	0.595	0.526	0.554	0.596
	[0.392, 0.798]	[0.316, 0.735]	[0.326, 0.783]	[0.369, 0.824]
Probit	X		X	
Monotone Rank		X		X
Observations	8,687	8,687	7,872	7,872

95% Confidence Intervals in brackets estimated using clustered (by individual) bootstrap. Coefficients are scaled so that the norm of all coefficients is equal to 1. Other variables included: year dummies; interactions for age group × education; spousal age group fixed effects. Initial income controls include splines in initial labor income. Income controls allowed to have different effects in each year.

Table 3: Intensive Labor Supply Equation, First Stage

Dependent Variable:	Δ ln(1-MTR)	Δ ln(After-Tax Income)	Δ ln(1-MTR)	Δ ln(After-Tax Income)
Predicted Δ ln(1-MTR)	0.807	0.167	0.798	0.149
	[0.671, 0.942]	[0.087, 0.246]	[0.691, 0.904]	[0.071, 0.227]
Predicted Δ ln(After-Tax Income)	0.020	1.092	0.082	1.000
	[-0.021, 0.061]	[1.068, 1.116]	[0.042, 0.123]	[0.970, 1.029]
Partial F-Statistic	150.61	2463.98	206.02	413.85
	Women	Women	Men	Men
	6511	6511	5813	5813

95% Confidence Intervals in brackets adjusted for clustering at individual-level. Other variables included: year dummies; interactions for age group × education; spousal age group fixed effects. Initial income controls include splines in initial labor income. Income controls allowed to have different effects in each year.

Table 4: Intensive Labor Supply Equation, 2SLS

Dependent Variable:	ln(Labor Income)					
	Women			**Men**		
Selection Adjustment:	None	Heckman	Semi-Parametric	None	Heckman	Semi-Parametric
	(1)	(2)	(3)	(4)	(5)	(6)
Δ ln(1-MTR)	0.341	0.671	0.239	0.091	0.644	0.577
	[-0.471, 1.153]	[-0.275, 1.617]	[-0.537, 1.817]	[-0.572, 0.754]	[-0.228, 1.516]	[-0.228, 1.382]
Δ ln(After-Tax Income)	-0.022	0.226	0.205	-1.328	-1.274	-0.746
	[-0.219, 0.175]	[-0.032, 0.484]	[0.004, 0.407]	[-1.657, -0.999]	[-1.627, -0.922]	[-1.595, -0.834]
Observations	6,511	6,511	6,511	5,813	5,813	5,813

95% Confidence Intervals in brackets estimated using clustered (by individual) bootstrap. Other variables included: year dummies; interactions for age group × education; spousal age group fixed effects. Initial income controls include splines in initial labor income. Income controls allowed to have different effects in each year.

Table 5: Extensive Labor Supply Equation (Employment), IV-Probit

Dependent Variable:	I(Employed)					
	Women			**Men**		
Selection Adjustment:	None	Heckman	Semi-Parametric	None	Heckman	Semi-Parametric
ln(After-Tax Labor Income)	0.154	0.165	0.154	0.166	0.176	0.173
	[0.117, 0.190]	[0.129, 0.201]	[0.118, 0.190]	[0.113, 0.219]	[0.131, 0.221]	[0.126, 0.220]
Observations	8,687	8,687	8,687	7,872	7,872	7,872

95% Confidence Intervals in brackets estimated using clustered (by individual) bootstrap. Other variables included: year dummies; interactions for age group × education; spousal age group fixed effects. Initial income controls include splines in initial labor income. Income controls allowed to have different effects in each year. Change in ln(After-Tax Income) also included and instrumented.

Table 6: Extensive Labor Supply Equation (Retirement), IV-Probit

Dependent Variable:	I(Retirement)					
	Women			**Men**		
Selection Adjustment:	None	Heckman	Semi-Parametric	None	Heckman	Semi-Parametric
ln(After-Tax Labor Income)	-0.157	-0.171	-0.158	-0.076	-0.085	-0.083
	[-0.203, -0.111]	[-0.219, -0.123]	[-0.205, -0.111]	[-0.132, -0.020]	[-0.148, -0.022]	[-0.142, -0.024]
Observations	8,687	8,687	8,687	7,872	7,872	7,872

95% Confidence Intervals in brackets estimated using clustered (by individual) bootstrap. Other variables included: year dummies; interactions for age group × education; spousal age group fixed effects. Initial income controls include splines in initial labor income. Income controls allowed to have different effects in each year. Change in ln(After-Tax Income) also included and instrumented.

Table 7: Effect of Eliminating Employee Portion of Payroll Tax

Outcome:	Not Working	Not Working	Retired	Retired
Effect of Age-Specific Payroll Tax	-0.007	-0.008	-0.008	-0.004
	[-0.010, -0.005]	[-0.012, -0.005]	[-0.011, -0.004]	[-0.008, -0.000]
Baseline	0.250	0.262	0.139	0.166
	Women	Men	Women	Men

Uses results from Tables 5 and 6 to simulate effects of eliminating the employee portion of the payroll tax. We calculate after-tax labor income with and without the payroll tax, comparing the probabilities of not working and retiring. 95% Confidence Intervals in brackets estimated using clustered (by individual) bootstrap.

A Appendix

Table A.1: Extensive Labor Supply Equation (Employment), 2SLS

Dependent Variable:	I(Employed)					
	Women			**Men**		
Selection Adjustment:	None	Heckman	Semi-Parametric	None	Heckman	Semi-Parametric
ln(After-Tax Labor Income)	0.041	0.044	0.041	0.046	0.048	0.047
	[0.031, 0.051]	[0.034, 0.054]	[0.030, 0.052]	[0.030, 0.062]	[0.032, 0.064]	[0.032, 0.062]
Observations	8,687	8,687	8,687	7,872	7,872	7,872

95% Confidence Intervals in brackets estimated using clustered (by individual) bootstrap. Other variables included: year dummies; interactions for age group × education; spousal age group fixed effects. Initial income controls include splines in initial labor income. Income controls allowed to have different effects in each year. Change in ln(After-Tax Income) also included and instrumented.

Table A.2: Extensive Labor Supply Equation (Retirement), 2SLS

Dependent Variable:	I(Retirement)					
	Women			**Men**		
Selection Adjustment:	None	Heckman	Semi-Parametric	None	Heckman	Semi-Parametric
ln(After-Tax Labor Income)	-0.026	-0.028	-0.026	-0.010	-0.010	-0.010
	[-0.035, -0.017]	[-0.037, -0.019]	[-0.035, 0.017]	[-0.022, 0.002]	[-0.022, 0.002]	[-0.022, 0.002]
Observations	8,687	8,687	8,687	7,872	7,872	7,872

95% Confidence Intervals in brackets estimated using clustered (by individual) bootstrap. Other variables included: year dummies; interactions for age group × education; spousal age group fixed effects. Initial income controls include splines in initial labor income. Income controls allowed to have different effects in each year. Change in ln(After-Tax Income) also included and instrumented.

References

Akerlof, George A, "The economics of "tagging" as applied to the optimal income tax, welfare programs, and manpower planning," *The American Economic Review*, 1978, *68* (1), 8–19.

Auten, Gerald and Robert Carroll, "The Effect Of Income Taxes On Household Income," *The Review of Economics and Statistics*, November 1999, *81* (4), 681–693.

_ , _ , and Geoffrey Gee, "The 2001 and 2003 Tax Rebate Reductions: An Overview and Estimate of the Taxable Income Response," *National Tax Journal*, September 2008, *61* (3), 345–364.

Banks, James and Peter A. Diamond, "The base for direct taxation," in James A. Mirrlees et al., ed., *The Mirrlees Review*, Oxford: Oxford University Press, 2010, pp. 548–648.

Biggs, Andrew, "A Payroll Tax Cut Could Help Social Security," *Wall Street Journal*, April 2012. April 24.

Blau, Francine D. and Lawrence M. Kahn, "Changes in the Labor Supply Behavior of Married Women: 1980 - 2000," *Journal of Labor Economics*, 2007, *25*, 393–438.

Blomquist, N Sören and Urban Hansson-Brusewitz, "The effect of taxes on male and female labor supply in Sweden," *Journal of Human Resources*, 1990, pp. 317–357.

Blomquist, Sören and Håkan Selin, "Hourly wage rate and taxable labor income responsiveness to changes in marginal tax rates," *Journal of Public Economics*, 2010, *94* (11), 878–889.

Blundell, Richard, Alan Duncan, and Costas Meghir, "Estimating Labor Supply Responses Using Tax Reforms," *Econometrica*, 1998, *66* (4), 827–861.

_ and Thomas Macurdy, "Labor supply: A review of alternative approaches," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Vol. 3 of *Handbook of Labor Economics*, Elsevier, 1999, chapter 27, pp. 1559–1695.

Butrica, Barbara, Karen E Smith, and C Eugene Steuerle, "Working for a good retirement," *The Levy Economics Institute Working Paper Series*, 2006, (463).

Cavanagh, Christopher and Robert P. Sherman, "Rank estimators for monotonic index models," *Journal of Econometrics*, 1998, *84* (2), 351 – 381.

Chernozhukov, Victor and Han Hong, "An MCMC approach to classical estimation," *Journal of Econometrics*, 2003, *115* (2), 293–346.

Eissa, Nada, "Taxation and labor supply of married women: the Tax Reform Act of 1986 as a natural experiment," Technical Report, National Bureau of Economic Research 1995.

— **and Hilary Williamson Hoynes**, "Taxes and the labor market participation of married couples: the earned income tax credit," *Journal of Public Economics*, August 2004, *88* (9-10), 1931–1958.

— **, Henrik Jacobsen Kleven, and Claus Thustrup Kreiner**, "Evaluation of four tax reforms in the United States: Labor supply and welfare effects for single mothers," *Journal of Public Economics*, April 2008, *92* (3-4), 795–816.

Feenberg, Daniel and Elisabeth Coutts, "An introduction to the TAXSIM model," *Journal of Policy Analysis and Management*, 1993, *12* (1), 189–194.

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

— **and Jeffrey B Liebman**, "Social security," *Handbook of public economics*, 2002, *4*, 2245–2324.

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

Gelber, Alexander M. and Joshua W. Mitchell, "Taxes and Time Allocation: Evidence from Single Women and Men," *The Review of Economic Studies*, 2011.

Gelber, Alexander M, Damon Jones, and Daniel W Sacks, "Earnings Adjustment Frictions: Evidence from the Social Security Earnings Test," Technical Report, National Bureau of Economic Research 2013.

Giertz, Seth, "The Elasticity of Taxable Income over the 1980s and 1990s," *National Tax Journal*, 2007, *60* (4).

Giertz, Seth H, "The elasticity of taxable income during the 1990s: new estimates and sensitivity analyses," *Southern Economic Journal*, 2010, *77* (2), 406–433.

Gruber, Jon and Emmanuel Saez, "The elasticity of taxable income: evidence and implications," *Journal of Public Economics*, April 2002, *84* (1), 1–32.

Gruber, Jonathan and Peter Orszag, "Does The Social Security Earnings Test Affect Labor Supply And Benefits Receipt?," *National Tax Journal*, 2003, *56*, 755–773.

Haider, Steven J. and David S. Loughran, "The Effect of the Social Security Earnings Test on Male Labor Supply: New Evidence from Survey and Administrative Data," *Journal of Human Resources*, 2008, *43* (1).

Hausman, Jerry, "Taxes and Labor Supply," in Alan J. Auerbach and Martin Feldstein, eds., *Handbook of Public Economics*, New York: Elsevier, 1985, pp. 213–263.

Hausman, Jerry A, "Taxes and labor supply," *Handbook of Public Economics*, 1985, *1*, 213–263.

Heckman, James J, "Sample Selection Bias as a Specification Error," *Econometrica*, January 1979, *47* (1), 153–61.

_ , "Varieties of Selection Bias," *American Economic Review*, May 1990, *80* (2), 313–18.

Heim, Bradley T, "The effect of recent tax changes on taxable income: Evidence from a new panel of tax returns," *Journal of Policy Analysis and Management*, 2009, *28* (1), 147–163.

Laitner, John and Dan Silverman, "Consumption, retirement and social security: Evaluating the efficiency of reform that encourages longer careers," *Journal of Public Economics*, 2012, *96* (78), 615 – 634.

Meyer, Bruce D. and Dan T. Rosenbaum, "Welfare, The Earned Income Tax Credit, And The Labor Supply Of Single Mothers," *The Quarterly Journal of Economics*, August 2001, *116* (3), 1063–1114.

Newey, Whitney K., "Efficient estimation of limited dependent variable models with endogenous explanatory variables," *Journal of Econometrics*, November 1987, *36* (3), 231–250.

_ , "Two-step series estimation of sample selection models," *The Econometrics Journal*, January 2009, *12* (Supplement s1), S217–S229.

Powell, David and Hui Shan, "Income Taxes, Compensating Differentials, and Occupational Choice: How Taxes Distort the Wage-Amenity Decision," *American Economic Journal: Economic Policy*, February 2012, *4* (1), 224–47.

Reznik, Gayle, David Weaver, and Andrew Biggs, "Social Security and Marginal Returns to Work Near Retirement," *Social Security Issue Paper*, 2009, (2009-02).

Saez, Emmanuel, "The effect of marginal tax rates on income: a panel study of 'bracket creep'," *Journal of Public Economics*, May 2003, *87* (5-6), 1231–1258.

_ , Joel Slemrod, and Seth Giertz, "The elasticity of taxable income with respect to marginal tax rates: a critical review," *Journal of Economic Literature*, March 2012, *50* (1), 3–50.

Schafgans, Marcia M.A. and Victoria Zinde-Walsh, "On Intercept Estimation In The Sample Selection Model," *Econometric Theory*, February 2002, *18* (01), 40–50.

Singleton, Perry, "The Effect of Taxes on Taxable Earnings: Evidence from the 2001 and Related U.S. Federal Taxes," *National Tax Journal*, 2011, *64*, 323–352.

Song, Jae G. and Joyce Manchester, "New evidence on earnings and benefit claims following changes in the retirement earnings test in 2000," *Journal of Public Economics*, April 2007, *91* (3-4), 669–700.

Subbotin, Viktor, "Asymptotic and bootstrap properties of rank regressions," MPRA Paper 9030, University Library of Munich, Germany November 2007.

Triest, Robert K, "The effect of income taxation on labor supply in the United States," *Journal of Human Resources*, 1990, pp. 491–516.

Weinzierl, Matthew, "The surprising power of age-dependent taxes," *The Review of Economic Studies*, 2011, *78* (4), 1490–1518.

Ziliak, James P and Thomas J Kniesner, "The effect of income taxation on consumption and labor supply," *Journal of Labor Economics*, 2005, *23* (4), 769–796.