ESSAYS ON THE EFFECTS OF TAXATION

by

Shanthi Priya Ramnath

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2010

Doctoral Committee:

Professor Joel Slemrod, Chair
Professor Mary Corcoran
Professor James R. Hines
Professor Jeffrey Smith

In memory of my father.
For his love, laughter, and endless support–
I am eternally grateful.

ACKNOWLEDGEMENTS

I would like to thank the members of my committee – Joel Slemrod, Jim Hines, Jeff Smith, and Mary Corcoran – for all of their help and guidance. I also thank John Dinardo and Miles Kimball for their advice and valuable feedback on my research. I would especially like to thank Brendan Epstein, Marta Murray-Close, Joanne Hsu, and Gabriela Contreras for their friendship and support throughout my graduate school experience. My second and third essays were co-written with Matt Rutledge and Brendan Epstein. I thank them both for their time and efforts spent on our joint work.

I greatly appreciate support from the Robert V. Roosa dissertation fellowship and the Rackham Graduate School at the University of Michigan.

Finally, I would like to thank my parents and Suja and Veena for their support and encouragement throughout my life. I am grateful to Jon, who has been my rock throughout this process and provided me with invaluable feedback on my work.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

CHAPTER I

Introduction

This dissertation is comprised of three essays that focus on under-explored impacts of taxation, including how it influences the behavior of individuals, the interaction between firms and workers, and the economy as a whole. In the three essays, I test theoretical predictions through empirical analyses from both a micro and a macro perspective, using disparate methodologies as required by the disparate problems I address.

The first essay examines the Savers Credit, which is a tax credit given to low and middle income households for contributing to a retirement savings plan. The policy is structured such that reporting one extra dollar of income could lead to large loss in credit, giving individuals incentive to not report that last dollar. This discontinuity allows for a clear analysis of the behavior of taxpayers near the notch. I assess the distortion resulting from the policys incentive structure gauged through misreported income and I test whether the policy was effective in achieving its goal of increasing retirement contributions. I find that individuals indeed responded to the policy's unintended incentive to misreport income, but failed to increase retirement contributions on the margin.

The second essay, which is co-written with Matthew Rutledge, analyzes whether changes made to marginal tax rates on personal income affect pre-tax wage rates. Past literature often assumes that pre-tax wage rates are unchanged by a tax policy change. We formally test this assumption by focusing on the Tax Reform Act of 1986, which, most notably, made large changes to the personal income tax. Using survey data from the Survey of Income and Program Participation (SIPP) to follow individuals and their employment history, we find that changes in net-of-tax rates are negatively associated with pre-tax wage rates. Our empirical analysis explores how taxes can affect the wage rates offered to workers, and fails to support the claim that pre-tax wage rates are invariant to changes in marginal tax rates.

The third essay, which is co-written with Brendan Epstein, studies the role that taxes play in determining labor hours across countries. Past studies have explained differences in labor hours per population for a broad set of OECD countries by looking at differences in effective tax rates; our study provides additional insight on this topic by also accounting for employment changes that took place over the past 40 years. In particular, we show that the standard neoclassical model with taxes is a better predictor of hours per worker due to its inability to capture hours changes on the extensive margin. We then develop a model that allows both hours per worker and employment per population to vary. We find that our model accounts for a larger fraction of aggregate data on hours per worker than the standard neoclassical model with taxes. Thus the impact that taxes have on individuals' decisions to work can be better understood at an aggregate level when the hours decision is separated into an intensive and an extensive margin.

As a whole this dissertation explores behavioral responses to assess the impact resulting from tax policy. In the first two essays I study specific tax policies to gain a better understanding of broader public finance topics, including the impact of non-linear budget sets and the incidence of a tax on personal income. The third essay lies at the intersection of public finance and macroeconomics and analyzes tax policy more generally in an international setting. This dissertation contributes to both public finance and macroeconomic literature by helping to better understand the specific impacts of taxation on micro- and macroeconomic decisions.

CHAPTER II

Taxpayers' Response to Notches: Evidence from the Saver's Credit

2.1 Introduction

When tax incentives are used to motivate a desired behavior, they often induce unintended responses in the process. The Saver's Credit, a non-refundable tax credit given to low and middle income households for making retirement contributions, is no exception. Although the credit is meant to subsidize retirement savings, its design also effectively subsidizes people to adjust their income.[1] This paper analyzes the overall impact of the Saver's Credit by examining the consequences of this policy, both intended and unintended.

The goal of the Saver's Credit is to encourage retirement savings among low and middle income households (Gale et al. (2005)), yet its structure allows for some to lose as much as $600 in credit by earning one extra dollar of income. To provide the largest benefit for those with the lowest incomes, the amount of credit falls discontinuously as adjusted gross income (AGI) increases for a given amount of savings. The resulting discontinuity, or "notch," in an individual's budget constraint fosters a strong incentive to forego that extra dollar of income, either by altering labor supply or by misreporting income. While this is similar to the incentives created by

[1] This adjustment to the income need not be illegal.

the nonlinearities in the personal income tax, the personal income tax merely creates kinks or changes to the slope of the budget line not the shifts associated with a notch. Thus, the Saver's Credit policy provides households an even stronger reason to report taxable income just below the notch, which could manifest in the income distribution as bunching, or a number of individuals grouping their incomes just below the notch.

This paper exploits the discontinuous structure of the Saver's Credit to investigate two questions: conditional on receiving the Saver's Credit, do households adjust their income in order to receive a higher credit rate? and: do households that receive a higher credit save more? Because households with higher savings have a stronger incentive to manipulate their income and bunch at the notch, bunching has implications for savings behavior. Thus, I start by examining whether bunching exists among people who filed for the Saver's Credit. If bunching is found, then people who report incomes below the notch to receive a higher credit rate may also have higher marginal propensities to save. For instance, an individual that has a strong preference for saving and thus saved the maximum amount, also has an increased incentive to bunch as she has the most to gain from a higher credit rate. This makes disentangling the policy's influence on savings contributions difficult. In particular, if the higher credit rate is associated with higher levels of savings, then determining the motivating factor for the change in savings will be difficult. If bunching is found with no increase in the level of savings contributions, then the policy is simply providing an incentive for people to report income below the notch. If no bunching is found, then an increase in savings signals the program effectively encouraged behavior for the marginal person without the unintended consequences.

To analyze how households respond to the Saver's Credit, I use the IRS Statistics of Income (SOI) Individual Public Use Tax Files spanning 2002 through 2004. The data contain information obtained directly from individual tax returns, which I use to estimate the effects of the Saver's Credit. I conduct a formal test for bunching by adapting a technique developed by McCrary (2008) and find evidence that bunching exists. Further inspection of the results reveals that, although a significant break exists in the pooled sample, the 2003 data appear to be driving the results. This result is puzzling, as the credit's effect on bunching appears to lessen over time, whereas intuitively one might expect the bunching to increase over time as people learn about and adapt behavior in response to the notch.

It may be the case that competing programs also influence the behavior of the targeted population. Thus, I consider the confounding effects of additional federal programs, each with its own set of incentives, aimed at the same demographic. The largest anti-poverty program targeted at potential recipients of the Saver's Credit is the Earned Income Tax Credit (EITC); those accepting the EITC credit may be reacting to an alternative set of incentives, thereby confounding the bunching results. After excluding EITC recipients from the sample, the magnitude and significance of the break increases over time, consistent with people learning.

The nature of the program makes the regression discontinuity research design seem ideal for studying the effect of the credit rate changes on savings contribution levels. However, the bunching complicates these estimates by potentially violating the identification assumption necessary for estimation. Although bunching is found in the data, I place bounds on the estimated treatment effect, following Lee (2002), which account for the potential resulting bias. Conditional on taking the Saver's Credit, I find no significant evidence that receiving a higher credit rate increased individual savings contributions for the marginal person. The overall impact of the Saver's Credit appears to be that taxpayers taking the Saver's Credit understand

and respond to the incentive to bunch at the notch but their savings contributions are unresponsive to the change in its price.

2.2 The Saver's Credit

The Saver's Credit targets households who earn below a threshold income level, where the income level is determined by filing status.[2] Individuals may receive a non-refundable tax credit on retirement contributions of up to $2,000 made to both traditional and Roth IRA plans as well as elective deferrals plans such as 401(k) and 403(b) plans.[3] Because the credit is non-refundable, individuals must have positive tax liability to receive a Saver's Credit. Details regarding the credit rates for the Saver's Credit are presented in Table 2.1. The last row of Table 2.1 calculates the equivalent match rate by interpreting the Saver's Credit like an employer match on elective deferrals.[4] For example, a taxpayer who contributes $1 earns a $0.50 credit that immediately offsets tax liability and puts $0.50 back in that taxpayer's pocket. That 50% credit rate has an economically equivalent match rate of 100%, since the taxpayer and the government effectively each contribute $0.50. This calculation allows for a comparison to studies on the success of employer matching as a savings incentive, which I will draw from when discussing the impact of the Saver's Credit on retirement contributions.

Because the Saver's Credit rate changes discontinuously with AGI (for a given amount of savings), taxpayers face a discrete jump in their after-tax income at the program's income cutoffs. Suppose a single filer with a positive tax liability and an AGI of $15,000 contributes $2,000 to a retirement plan. She will receive a tax credit

[2] After becoming permanent under the Pension Protection Act of 2006, the Saver's Credit was indexed for inflation causing the threshold income levels to rise from 2007 on.

[3] Couples that are married filing jointly can earn a credit on contributions up to $4,000.

[4] This follows Duflo et al. (2006) and Gale et al. (2005) who compare the Saver's Credit to employer matching on contributions to a company retirement plan where the credit rate, s, is equivalent to an employer match rate of $\frac{s}{1-s}$.

of \$1,000, or 50% of her savings contribution to offset her tax liability. However, if she makes one extra dollar of income, her credit rate falls from 50% to 20%. Since each dollar of her contribution now earns the lower credit rate, her total credit will then fall from \$1,000 to \$400, forming a \$600 notch in her budget constraint. Figure 2.1 illustrates the specific notches that result from the Saver's Credit seen in the before- and after-tax budget constraint for a married couple filing jointly. The couple's budget constraint with the Saver's Credit policy maintains the same slope, but at the income cutoffs, the couple faces downward jumps in their after-tax income for a fixed retirement contribution. Thus, moving from an income of \$30,000 to \$30,001 strictly lowers utility.

Given the complexity of the Saver's Credit, there is evidence that the credit has a low take-up rate. Between 2002 and 2006, roughly 5.3 million credits were filed each year and the average credit payment was around \$190. Following the first year the credit was offered, Koenig and Harvey (2005) found that 34% of eligible taxpayers failed to claim up to \$496 million dollars in credits, and 43% of claimed credits were limited by tax liability. This low participation rate is even more staggering in light of Gale et al. (2005) who note that the Saver's Credit complements employer matching, making the effective match rate as high as 200% for a 50% employer match rate. Table 2.2 gives a detailed summary of aggregate participation rates by AGI for the Saver's Credit, as well as the average credit amounts.

There is a small but growing literature that looks specifically at the Saver's Credit. Koenig and Harvey (2005) study the Saver's Credit following its first year in existence and conclude that the credit's non-refundability is a limiting factor for eligibility. Also, the lack of knowledge for the credit substantially decreased the number of credits claimed. Gale et al. (2005) provide a general discussion of the Saver's Credit and suggest possible ways to improve the credit as a policy tool for encouraging retirement savings among low and middle income households. These studies offer descriptive analyses of the Saver's Credit, which I expand upon by examining not only the incentive to save but the incentive to alter income in order to avoid a credit loss at the notch.

2.3 Related Literature

The bunching incentives of the Saver's Credit are not unique; the US tax code creates similar incentives by imposing kinks and notches within a household's budget constraint. Whereas a notch creates a discontinuous jump within a budget constraint, a kink creates a slope change. Although theory predicts bunching in both cases, the incentive to bunch is stronger in the case of a notch. Past literature has looked at whether people respond to kinks within a budget constraint by bunching reported income. Saez (2009) finds very little bunching of AGI at the kinks created by the personal income tax, but finds that bunching exists at the first kink of the Earned Income Tax Credit (EITC) among self-employed individuals. This bunching disappears for those who are not self-employed, which may indicate that people who have more control over their income and/or reporting of income are more likely to bunch. Chetty (2009a) posits that the lack of consistent bunching in the data may be the result of optimization error. If individuals face some cost to adjusting their income to bunch, then depending on the size of the economic incentive, this cost may not be recouped by the benefit of reoptimizing. This also suggests that for sizeable economic incentives, such as those created by large kinks or notches, bunching may be found in the data more regularly. Chetty et al. (2009) incorporate potential frictions in optimizing and find bunching of income at large and salient kinks in the

Danish tax code, but find very little evidence for bunching at the smaller, less salient kinks. The Saver's Credit introduces a notch, which creates a large economic incentive to bunch. However, this bunching could be mitigated through issues of salience and income control, which will ultimately impact whether and how much bunching appears in the data.

Of particular interest to the questions explored in this paper are the papers by Duflo et al. (2006) and Duflo et al. (2007). Duflo et al. (2006) conduct an experimental program with incentives similar to the Saver's Credit to analyze the impact of offering a match for retirement contributions on participating in retirement savings plans. The study focuses particularly on participation in Express IRA (X-IRA) plans by H&R Block clients. X-IRAs are IRAs that can be opened at the time of filing using the tax refund earned on that filing. In the experiment, match rates of 0%, 20%, and 50% for IRA contributions are randomly assigned to taxpayers filing at H&R Block. The authors estimate when the match rate is increased from 20% to 50%, participation in X-IRA plans increases by 6.4%, while retirement contributions increase by $310, conditional on take-up. These experimental results are then compared to quasi-experimental results obtained on the Saver's Credit. The Saver's Credit effectively offers match rates of 0%, 11%, 25%, and 100%, though these rates are not randomly assigned. Using a difference-in-difference approach, where those who are ineligible for the Saver's Credit act as a comparison group, the authors estimate that increasing the effective match rate increases participation in X-IRA plans by 1.3% and, conditional on take-up, increases retirement contributions by $81. Duflo et al. (2006) find that the experimental results are more pronounced than the Saver's Credit in terms of participation and savings contributions. In a separate study, Duflo et al. (2007) use data from H&R Block to study these differences in the household response.

Both Duflo et al. (2006) and Duflo et al. (2007) observe a spike in the histogram of X-IRA participation at the location of the first notch in the Saver's Credit. Given that the benefit to opening an account is constant in the range of the 50% credit rate, there is no reason to expect a spike at the first notch in participation unless people are bunching to avoid the credit loss. This paper will provide a direct test for whether this bunching exists in the density of AGI along with an estimate for the amount of bunching that took place. By using representative data, this paper will also generate more generalizable estimates of the Saver's Credit's impacts on the individuals it seeks to affect.

In my empirical study, I focus my analysis on individuals who take the Saver's Credit to determine how the structure of the credit influences their behavior. Those that file for the credit are arguably more informed about the Saver's Credit's structure and incentives than those that do not file for the credit. This generates a number of questions. Are people who take the credit able to fully optimize their credit rate by altering their reported income? If differing credit rates are known ahead of time, do they have any impact on savings contributions? In order to motivate the empirical estimation, I start with a theoretical model that generates specific behavioral predictions arising from the Saver's Credit.

2.4 Theory

2.4.1 Exogenous Income

The incentive structure of the Saver's Credit can be modeled in a two-period framework. I start with a standard intertemporal budget constraint where the agent lives for two periods and maximizes utility over consumption, given by $U(c_1, c_2)$. In the first period the agent inelastically supplies labor and thus earns an exogenous income, y. The agent chooses how much to save, a, and how much to consume, c_1, in period 1. For simplicity, I assume the only means of savings is through a retirement plan and thus I use the terms savings and retirement contributions interchangeably.[5] The policy dictates for a given amount of savings less than $\bar{A}$, the government will provide a credit equal to a proportion, s, of the agent's savings, where s depends on income. For any savings contribution above $\bar{A}$, the agent simply receives $s\bar{A}$ and the amount of savings no longer impacts the amount of the total credit. For simplicity, the marginal tax rate on income, $\tau > 0$, is assumed to be constant. The first period budget constraint is given by: for $a < \bar{A}$,

$$y(1-\tau) + s(y)a = c_1 + a,$$

and for $a \geq \bar{A}$,

$$y(1-\tau) + s(y)\bar{A} = c_1 + a,$$

$$\text{where } s(y) = \begin{cases} .5 & \text{if } 0 < y \leq Y_a \\ .2 & \text{if } Y_a < y \leq Y_b \\ .1 & \text{if } Y_b < y \leq Y_c \\ 0 & \text{if } y > Y_c. \end{cases}$$

In the second period, the agent consumes her savings plus the interest earned from the first period. The second period budget constraint is given by,

$$c_2 = a(1+r).$$

Substituting for savings, the intertemporal budget constraint is: for $a < \bar{A}$

$$y(1-\tau) = c_1 + (1-s(y))\frac{c_2}{1+r},$$

and for $a \geq \bar{A}$

$$y(1-\tau) + s(y)\bar{A} = c_1 + \frac{c_2}{1+r}.$$

[5]For simplicity this set up does not allow for retirement contributions to be tax deductible, which would lower taxable income and tax liability.

For $a < \bar{A}$, the price of consumption in period 2 falls to $\frac{1-s(y)}{1+r}$, while for $a \geq \bar{A}$, the price of consumption in period 2 remains $\frac{1}{1+r}$ – the same price as in the case of no credit. Thus, for those who save $a \geq \bar{A}$, the credit creates an income effect but no substitution effect. When labor is exogenous or inflexible over the range $0 \leq y \leq Y_3$, the credit rate is also exogenous since the agent has no control over y. However, the credit changes the price of saving, and increases overall income.

If consumption is a normal good, then the income effect from the credit should lead to increased consumption in period 1, which decreases savings. However, because the price of second period consumption falls from the credit, the substitution effect would lower period 1 consumption, thereby increasing savings. Thus, by lowering the price of consumption in the second period, the credit may increase savings depending on how the income and substitution effects interact. In the empirical section of this paper, I will test whether retirement contributions were indeed impacted by the Saver's Credit and in what direction.

2.4.2 Endogenous Income and Income Reports

Next, I relax the exogenous labor income assumption, and incorporate the agent's choice between labor and leisure, where labor hours is denoted as l. The agent must choose the number of hours to work in period 1, for a given wage rate, w. Income, y, is calculated as

$$y = wl.$$

Because the proportion of savings returned to the agent as a credit depends on income, the incentive to earn an extra dollar is distorted at the income cutoffs for differing credit rates. As a result, there is an incentive to either forego the extra dollar of earned income or, if possible, to misreport income. I therefore extend the model to include the possibility that individuals have an incentive to alter their income report through a choice variable, x, where x is unreported income.[6] The agent's reported income, y^R, now differs from her earned income, y, so that

$$y^R = y - x.$$

The savings credit and taxes will now both depend on y^R rather than y and the new intertemporal budget constraint is, For $a < \bar{A}$

$$wl(1-\tau) + \tau x + s(wl - x)\frac{c_2}{1+r} = c_1 + \frac{c_2}{1+r},$$

and for $a \geq \bar{A}$

$$wl(1-\tau) + \tau x + s(wl - x)\bar{A} = c_1 + \frac{c_2}{1+r}.$$

The optimal reported income, if there were no cost to misreporting income would be trivial and anyone with savings would report $y \leq y_1$. However, misreporting carries a risk. Following Slemrod (2001), I include a cost to misreporting income denoted by $\phi(x, wl)$, where $\phi_x(x, wl) > 0$ and $\phi_l(x, wl) < 0$. An individual-specific parameter, γ, where $0 < \gamma \leq 1$, multiplies this cost function to represent idiosyncratic costs of misreporting. For example, a low value of γ indicates the individual has a low cost for misreporting income as in the case of self-employment. On the other hand, a high value of γ indicates income that is difficult to misreport such as when income is predominately earned through wages and salary, which are also reported by employers.

Agents are endowed with a fixed amount of time, $\bar{L}$, that is allocated to either labor hours or to leisure. The utility function is expanded so that the agent derives utility from leisure. The agent's problem is to choose c_1, c_2, l and x to maximize

[6] I do not distinguish between legal misreporting through tax avoidance and illegal misreporting through tax evasion.

utility,

$$\max_{c_1,c_2,x,l} U(c_1, c_2, \bar{L} - l)$$

subject to their intertemporal budget constraint: for $a < \bar{A}$

$$wl(1-\tau) + \tau x + s(y^R)\frac{c_2}{1+r} - \gamma\phi(x, wl) = c_1 + \frac{c_2}{1+r}$$

and for $a \geq \bar{A}$

$$wl(1-\tau) + \tau x + s(y^R)\bar{A} - \gamma\phi(x, wl) = c_1 + \frac{c_2}{1+r}.$$

Because the schedule of rates for the Saver's Credit is discontinuous, notches are formed in the budget constraint. Thus, the maximization problem is solved for each income range and the agent chooses the bundle that gives the most overall utility. Given standard assumptions for the utility function, bunching will occur by some individuals at the notches. An individual will bunch as long as the cost of lowering their income report or their labor hours will be regained by the benefit from receiving a higher credit rate on their savings. Let x^* denote the unique amount of misreported income that positions an individual's income report at a notch. Each person faces a benefit and a cost to misreporting income where the marginal cost of misreporting is equal to $\gamma\phi_x(.)$, and differs across individuals. In some situations, a person will not misreport to x^* because the amount of optimal misreported income is either greater than the amount it takes to reach the notch, or the benefit from the credit does not cover the additional cost of the extra misreported income. This group will report $y^R > y - x^*$. For some, however, the extra benefit that comes from misreporting x^* at the notch, may raise them beyond the cost of misreporting. Accordingly, this group will report $y^R \leq y - x^*$, which will include people that optimally bunch at the notch.

An individual can bunch either by changing their labor supply or by changing their income report and both actions are associated with costs and benefits. For the purpose of intuition only, I assume that $s(y^R)$ is a continuously differentiable function and $s'(y^R) < 0$. In this case, the first order conditions for the variables are as follows,

$$c_1: \quad \frac{\partial U}{\partial c_1} - \lambda = 0$$

$$c_2: \quad \frac{\partial U}{\partial c_2} - \lambda \frac{1 - s(wl - x)}{1 + r} = 0$$

$$l: \quad -\frac{\partial U}{\partial l} - \lambda w[1 - \tau + s'(y^R)\psi(a) - \gamma\phi_l(.)] = 0$$

$$x: \quad \lambda[\tau - s'(y^R)\psi(a) - \gamma\phi_x(.)] = 0,$$

where $\psi(a) = \begin{cases} \frac{c_2}{1+r} & \text{if } a < \bar{A} \\ \bar{A} & \text{if } a \geq \bar{A}. \end{cases}$

Solving for the marginal rate of substitution between labor and first period consumption (MRS_{lc}) yields

$$(\frac{\partial U}{\partial l})/(\frac{\partial U}{\partial c_1}) = w[1 - \tau + s'(y^R)\psi(a) - \gamma\phi_l(.)]. \tag{2.1}$$

The expressions within the brackets can be broken down into two parts. Recall that $s'(y^R) < 0$, and $\phi_l < 0$. Then $s'(y^R)$ is interpreted as the decrease in benefit from working an additional hour due to the decrease in credit rate. On the other hand, ϕ_l can be interpreted as the increase in the benefit of working an additional hour that comes from lowering the cost of misreporting income. Thus the marginal rate of substitution between consumption and labor is equal to the net benefit from working an additional hour, taking into account the reduced credit rate and reduced cost to misreporting.

The first order condition for x is given by:

$$\lambda[\tau - s'(y^R)\psi(a) - \gamma\phi_x(.)] = 0. \tag{2.2}$$

This condition shows that the agent has an incentive to misreport income since increasing x by \$1 will reduce tax liability by τ. Also, x increases the credit rate by lowering y^R. The optimal amount of misreported income is that which sets the marginal cost of misreporting income, measured as ϕ_x, equal to the marginal benefit of increasing x, measured as $\tau - s'(y^R)\psi(a)$, where the change in credit rate is scaled by the amount of savings.

A certain amount of time and learning must be invested in order to claim the credit. However, beyond just knowing the credit exists an additional investment must be made to understand how it works. This information cost is excluded from the model but could prove to be an important factor for my empirical results. By ignoring the complexity of the program, I may find that agents base their actions on a slightly different problem which comes from this lack of understanding. This would impact whether bunching and savings behavior respond to the credit as predicted and thus poses a concern for the empirical estimation.

Two predictions emerge from the theory presented above: (1) people will bunch at the notch and (2) savings contributions are influenced by the offer of a credit, though the direction is ambiguous because of income and substitution effects. Whether these predictions appear in the data will be influenced by other factors. Starting with the former, the amount of bunching is affected by a number of additional factors, including the distribution of cost functions for misreported income among the population. The latter hypothesis is impacted by the distribution of preferences for saving along with differences in people's ability and propensity to save. These predictions on the potential behavioral responses of the Saver's Credit form the basis for the empirical section that follows.

2.5 Data Description and Summary Statistics

The Individual Public Use Tax Files are an annual cross-section of tax returns spanning 1960 to 2004, available at the Statistics Division of the Internal Revenue Service. The 2002 through 2004 data contain dollar amounts for all Saver's Credits filed during that period. After dropping observations pertaining to previous tax years, each sample represents roughly 127 million tax returns. Data are obtained through stratified probability sampling where each stratum is defined by a combination of AGI and the presence of particular tax forms. Sampling rates within each stratum range from 0.05% to 100%.[7] The Public Use Tax File over-samples wealthy individuals to achieve a broad range of tax rates in the data. Unfortunately, this limits the number of returns in the sample from low and middle income households. In particular, this greatly reduces returns that are both eligible for, and filed for the Saver's Credit. There are 7,718 returns in the combined sample of taxpayers claiming the Saver's Credit between 2002 and 2004 representing roughly 5.1 million claims in the population.

I calculate savings contributions to retirement plans using the amount of credit claimed on a tax return and dividing by the eligible credit rate. Because the credit is non-refundable, this calculation is bounded by total tax liability less additional credits that include the foreign tax credit, child care credit, elderly credit, and education credit. Those who fall in the 50% credit rate are more likely to reach this bound as they typically have the lowest tax liability. In the sample, roughly 52% of people receiving the 50% credit rate are at their credit limit. Overall, those with a Saver's Credit equivalent to the tax limit account for 15% of all Saver's Credit filers.

Similar to Duflo et al. (2006), I combine all taxpayers by normalizing AGI to

[7] A more complete description of the data can be found at http://www.nber.org/ taxsim/gdb/.

align the notch for each filing status to match married couples filing jointly. This entails multiplying single filers' AGI by 2, and head of households' AGI by 4/3. In addition, the 20% credit falls within a narrow income band, between $1,250 to $2,500, depending on income status; thus, I group those receiving the 10% credit rate and 20% credit rates. This creates one "Saver's Credit notch", marking the jump from receiving a "high" credit rate of 50% to receiving a "low" credit rate of 10% or 20%.

Because the Saver's Credit is targeted at low and middle income households, some individuals may be credit constrained, raising a question as to their ability to save for retirement. Table 2.3 shows participation in retirement plans by taxpayers that have a positive AGI below $25,000 and a positive AGI below $50,000. These aggregate data are unable to control for filing status or tax liability, but nonetheless show the existence of low and middle income household savers. Another concern arising from the lower end of the income distribution is the ability to control income. Underlying the theoretical prediction of bunching is the assumption that people have some control over reported income through labor supply or misreporting. The extent to which households have control over their income is therefore an important factor that will impact the results on bunching.[8] As a proxy for income control, Table 2.4 provides additional summary statistics on how Saver's Credit filers compare to eligible (based on their AGI) taxpayers that did not claim the Saver's Credit in terms of the types of income they report. The "eligible" group contains people that fall below the appropriate income limits and have positive tax liability. Because savings data are unavailable for people who did not take the credit, the eligible group will overstate the actual number of people eligible for the Saver's Credit by including those who did not contribute to a qualified retirement plan; however, this

[8]Saez (2009),Chetty et al. (2009) and others have estimated behavioral elasticities based on the amount of bunching induced by non-linearities within a budget constraint.

group can still serve as a useful comparison for Saver's Credit filers. Table 2.4 shows that Saver's Credit takers typically have more schedule income, which potentially indicates a lower cost to manipulating income. Additional summary statistics show that mean tax liability is greater for people who filed for the credit, consistent with a binding nonrefundability constraint.

In Table 2.5, I present probit results looking at the factors affecting the probability of taking the credit conditional on eligibility as defined in the paragraph above. Factors including e-filing and having a paid preparer increase the likelihood of take-up. The results show that having a higher credit is not associated with higher take-up, which is expected given that the credit rate is not randomly assigned. However, past studies, including the experimental results from Duflo et al. (2006), show that offering higher matches increases participation. For the empirical portion of this paper, I focus my analysis on those who take the Saver's Credit, and draw inference based on comparisons between the groups receiving different credit treatment. I start by analyzing whether people bunch their incomes at the notch, and then move on to look at the relative savings contributions within the groups.

2.6 Bunching

Since the credit's eligibility rules are known ahead of time and AGI is self-reported, taxpayers may report AGI just below the notch so as to benefit from the higher credit rate. As shown in Section 3.3, taxpayers can decrease labor hours so their income falls below the notch, or they can alter their income.[9] For the purpose of this paper, I will not distinguish between the two behaviors. Instead, I focus only on whether there exists bunching in the estimated density of AGI, as its existence is instructive

[9]According to Feldstein (1999), the welfare cost from imposing a tax can be measured simply by knowing the response in taxable income, not the mechanism of response. However, Chetty (2009b) disputes this. See Saez et al. (2009) for a critical survey.

for the welfare implications of the Saver's Credit. However, I will provide evidence that suggests income manipulation may be the mechanism filers use to bunch.

If people bunch, then a spike would appear in the density of AGI just below the notch. Figure 2.2 shows the kernel density estimate of AGI for 2002-2004, using Silverman's plug-in described in Cameron and Trivedi (2005) as a guide for choosing a bandwidth.[10] The histogram of the data shows a small spike in AGI at the notch, though the spike disappears in the kernel density plot. Figure 2.3 imposes a smaller bandwidth as a robustness check, but yields no substantial evidence that bunching exists in the kernel density estimated AGI density. Although the kernel density graphs provide a helpful first pass of the data, they are of limited use in identifying bunching since point estimates at the notch are obtained using observations on both sides of the notch. If bunching exists, then the density to right of the notch is inherently lower as people shift to the left. The kernel density estimate at the notch will mask all but the most extreme signs of bunching as the density will be averaged downward by observations from the right. Additionally, the histogram appears to provide evidence of bunching that the kernel density smooths over. This warrants a more formal test for a break in the density.

McCrary (2008) develops a test for detecting manipulation of a running variable– the variable a policy rule is based on–in the context of regression discontinuity (RD) estimation. For example, receiving a scholarship might be contingent on applicants scoring above a certain threshold on their SAT, making SAT score the running variable. When people have the power to affect the running variable, say through self-reporting, and the policy rule is known ahead of time, they may manipulate the running variable to ensure treatment. This is exactly the case for the Saver's Credit

[10]Silverman's plug-in is optimal for the normal kernel. I use the Epanechenikov kernel; however, the results do not change by kernel choice.

where AGI is the running variable. Assuming the distribution of AGI would be continuous absent the policy, a break in the estimated density at the notch would indicate manipulation of AGI took place. McCrary (2008) develops a formal test for such a break by essentially estimating whether a break exists in the estimated density of the running variable at the discontinuity from the policy. Intuitively, the test estimates the treatment effect of the policy on the density of the running variable. A point estimate is obtained for the behavioral response, which gives a sense of its magnitude. For the application of an RD design, bunching in the running variable has the potential to be problematic as it may lead to biased estimates of the treatment effect. However, in this paper, bunching is an object of interest, serving as evidence for a behavioral response to the policy.

After combining taxpayers so the notch is the same for each filing status, I define the running variable as the newly aligned (or "normalized") AGI. The test for bunching in the density of AGI proceeds as follows: first, an undersmoothed histogram is created where no one bin contains points both to the left and to the right of the break; second, local linear regression is used to smooth the histogram and provide an estimate of the density of AGI.[11] These steps are illustrated in Figures 2.4 and 2.5. Once each point is estimated, the estimated density graph provides visual evidence for whether a break exists in the data. The test statistic of the break is derived by taking the log difference in density of AGI at the notch, given by

$$\hat{\theta} = \ln \hat{f}^{+} - \ln \hat{f}^{-}. \tag{2.3}$$

This measures the difference in the density at the notch when the density is estimated separately with points to the left and points to the right of the notch. The null

[11] The binsize for the histogram is a function of the standard deviation of AGI. The estimated density is derived using triangle kernel weights for the local linear smoothing, however, the results are robust to different kernel choices. See McCrary (2008) for more details.

hypothesis is that $\hat{\theta}$ is zero at the notch, which indicates no bunching occurred. Therefore, a significant negative estimate for $\hat{\theta}$ implies that a large enough number of taxpayers are at or closely below the notch to suggest that some manipulated their AGI in order to receive the higher credit.

Figure 2.6 shows the graphical result of the test, effectively a pdf of AGI with 95% confidence intervals derived using points to the left of the break and points to the right of the break separately. Table 2.6 gives the numeric results from the break test and indicates a significant break in the distribution of AGI exists.[12] The coefficient for $\hat{\theta}$ shows that the log difference in density to the left of the notch is 27% higher than the estimated density to the right of the notch.

It is important to note the performance of $\hat{\theta}$ as an estimator is sensitive to the size of the bandwidth. To choose a bandwidth, I use the automated bandwidth selection process suggested by McCrary (2008). The procedure involves first binning the data into a histogram, then fitting a global $4th$ order polynomial on each side of the break, where the independent variable is defined by the midpoint of the bins from the histogram. The estimated second derivative is then used to calculate the rule-of-thumb bandwidth selector from Fan and Gijbels (1996) on each side of the break. These two bandwidth choices are then averaged to obtain one bandwidth to be used in the estimation process. The intuition for the selection process is that the size of the bandwidth should be inversely proportional to the curvature of the estimated density.[13] The result in Table 2.6 uses the automatic bandwidth described above. Although this is a relatively large bandwidth, I perform the break test using additional bandwidth values as a robustness check. The break remains significantly different than zero even when substantially larger bandwidth choices are employed,

[12] I thank Brian Kovak for providing me with code to run this test.

[13] A more complete discussion for the bandwidth selection can be found in McCrary (2008) and Fan and Gijbels (1996)

shown in Table 2.7. As expected, a smaller bandwidth of 3,500 reveals a larger break. As an additional robustness check, I test for bunching at breaks other than $30,000. Figures 2.7 and 2.8 show that the breaks at values other than $30,000 are either not statistically significant or are significant in the wrong direction.

The results from the test suggest that bunching took place in response to the policy. As mentioned earlier, data limitations prohibit isolating the exact mechanism for bunching. However, variation in the types of income reported can serve as a proxy for one's ability to manipulate their income. Looking at bunching behavior broken down by type of income reported may provide suggestive evidence that adjusting income reports was used to bunch. Figure 2.9 graphs the AGI of individuals who reported Schedule C income, while Figure 2.10 graphs the AGI of individuals that did not report Schedule C income. The break is indeed larger for those with Schedule C income, indicating that people with a presumably lower cost of misreporting their income comprise a larger proportion of bunchers. An additional test looks at individuals who report only wage and salaried income, illustrated in Figure 2.11. The break in the density of AGI for these individuals is no longer significantly different than zero. These graphs provide striking visual evidence, which indicates some bunching occurred most likely by misreporting income. That individuals with Schedule C income are more likely to bunch is inline with past studies on bunching. Saez (2009) finds evidence of bunching near to first kink of the Earned Income Tax Credit (EITC). After parsing the sample by individuals with income from self-employment and wage earners (those with no self-employment income), Saez finds strong bunching only for the group with self-employment income. Similar results are found in Chetty and Saez (2009), where information regarding the EITC schedule was given to tax filers. More bunching was found for filers with self-employment income, though wage earners also responded to the EITC kink by bunching.

Given the possibility for optimization error as discussed in Chetty (2009a), I run the break test separately by year to determine whether learning took place in terms of people's understanding of the notch's incentives. In particular, one might expect that over time the behavioral effect of the notch would become stronger in the sense of more bunching taking place as people become more aware of the program and the incentive to report income below the notch. Table 2.8 reveals the break found in the density for the combined years is driven solely by 2003 data. In both 2002 and 2004, the break is not significantly different than zero at a 95% confidence level, though the break in 2002 is significant at the 90% confidence level. Given that 2002 was the first year of the credit, it is not surprising that less bunching would occur compared to later years. However, the results in 2004 pose somewhat of a puzzle, as it seems unlikely that those who bunch in 2003 would stop bunching in 2004, unless the cost to bunching grew between the two years. Table 2.8 shows that a Wald test of joint equality between all three years is barely rejected at the 95% level, though 2004 is rejected as being equal to 2002 and 2003 separately. Regardless, the point estimates in all years remain negative, which still indicates a higher density to left of the notch. I will address this puzzle by considering the impact of alternative incentives that may dominate the Saver's Credit's incentives.

2.6.1 The Impact of the EITC

There are additional programs targeted at low and middle income households, also creating distortions in the household budget constraint, that may confound an attempt to identify the effect of the Saver's Credit. In particular, the Earned Income Tax Credit (EITC) is aimed at the same demographic group and creates large kinks in the marginal tax rate, but unlike the Saver's Credit is refundable. The EITC incentives are structured so that some households are encouraged to work more and others are encouraged to work less depending on their AGI and number of qualifying children.[14] Of importance to this paper is that people who claim both the EITC and the Saver's Credit may respond to the EITC incentives rather than the Saver's Credit notch. The EITC is a well-known and refundable credit, thus making its incentives potentially more salient than the Saver's Credit. This is particularly relevant as those taking the EITC make up roughly 20% of the entire sample of Saver's Credit claims. Figures 2.15-2.17 present the same analysis as above but excluding the EITC filers with the results for $\hat{\theta}$ excluding EITC filers summarized by Table 2.9.

The estimates show EITC claims indeed affected the amount of bunching in each year, by increasing the bunching found in 2002 and 2003 and muting the bunching found in 2004. In fact, after dropping individuals that take the EITC, there is significant evidence of bunching in 2004.[15] Once the EITC claims are accounted for, a statistically significant break remains in 2003 as well. The breaks in AGI are both larger in 2003 and 2004 than the break found in 2002, consistent with the hypothesis that people learned and adapted to the incentives of the notch over time.

2.6.2 Discussion

For all Saver's Credit claims and all years combined, the estimate for $\hat{\theta}$, the log difference in the estimated density of AGI at the notch, is -0.263. This difference implies 0.5% of the sample may be bunching. However, additional people may be bunching, but not precisely at the notch. Thus, I offer a simple estimate of the

[14] The IRS has rules to define a "qualified child" where the child must meet all requirements for relationship, age, and residency. More information can be found at http://www.irs.gov/publications/p596/ch02.html

[15] In June 2003, the IRS announced that it would conduct three new programs with one of the stated goals to increases compliance of claims. One of these programs, the EITC Automated Underreporter (AUR) Study, was created to address overclaims from misreported income by improving which claims the IRS flagged as high risk of error. In 2004, the AUR was added to the base of compliance programs already in place (IRS 2008). This may have made it more difficult to misreport income in 2004 compared to other years, forcing those who previously misreported to move to the right of the notch.

counter-factual density of AGI, where I leave out observations close to the notch and instead interpolate those observations, as illustrated in Figure 2.18. By taking the difference between the cumulative density including the bunchers and the cumulative counterfactual density, I can estimate the proportion of individuals that bunch due to the notch.

The estimated difference in cumulative density translates to roughly 0.75% of all Saver's Credit filers. In other words, this difference can represent up to 38,000 individuals bunched as a result of the policy. If bunchers save the maximum of $2,000, then this would translate to government paid credits of $22.8 million.[16] Given that the overall program costs roughly one billion dollars each year, the additional cost incurred by bunching may seem negligible, but this calculation does not include the lost tax revenue from individuals who would have reported more absent the program. While the lost revenue also would most likely be small relative to the overall cost of the program (since the marginal tax rate on dollars earned near the Saver's Credit notch is low), the amount lost in tax revenue depends on the amount of underreported income needed to reach the notch. So, although the Saver's Credit induces a behavioral response in bunching, the additional cost to the government of this response is relatively small.

Given that bunching is found in the data, one would expect that those with higher savings contributions also have a higher incentive to bunch. This is confirmed in Section 3.3, where the benefit from misreporting that results from the increase in credit rate is scaled by the amount of savings. Thus, a spike in savings contributions should exist to the left of the notch as well. In the next section, I examine the

[16] Assuming all bunchers saved $2,000, the government would owe each of the 38,000 bunchers a $1000 credit rather than a $400 credit, thus creating an additional cost of $22.8 million. This number is an upper bound due contributions, on average, being less than $2000. But as bunchers theoretically are more likely save, this does provide a rough estimate for the cost to the government.

retirement contributions behavior around the notch further to see whether individuals appear to be making higher contributions to the left of the notch.

2.7 Retirement Contributions

The Saver's Credit is meant to stimulate retirement savings among low and middle income households. Thus, a complete assessment of households responses to the Saver's Credit includes observing any resulting changes made to retirement contributions. The ability to make contributions is a function of income, where those who earn more tend to have more disposable income with which to make contributions. Looking at the average savings contributions by the high credit and lower income filers versus the low credit and higher income filers would therefore generate biased results. Instead, I exploit the discontinuity in the program and adopt a regression discontinuity approach.

Because the notch is arbitrarily assigned by the government, I can isolate individuals to the left and to the right of the notch to control for income effects on savings. To identify the treatment effect of the higher credit rate, I must assume those to the left and to the right of the notch are similar other than receiving different credit rates. A concern with this procedure arises from the behavioral response to the notch. If the bunching found in Section 2.6 is comprised of people who save more and therefore have the most to gain from the higher credit rate, then the RD estimate of the average treatment effect would be biased towards finding a positive effect of the higher credit on savings contributions. In other words, the average treatment effect would indicate people made larger contributions if they received a higher credit rate, when in fact the positive estimate could simply be the result of selection bias. However, because it is possible to sign the bias the result can be interpreted as a upper bound for the treatment effect. In addition, I use a procedure described below to more formally place an upper and lower bound on the impact.

Following Imbens and Lemieux (2008) I estimate the following equation where the level of savings contributions, denoted by A_i, is given by

$$A_i = \alpha + \tau D_i + \beta_0(Y_i - c) + \beta_1(Y_i - c) * D_i + \epsilon_i, \tag{2.4}$$

D_i is an indicator for whether someone received the high credit rate, c is the cutoff to receive a high credit, and Y_i is AGI. D_i depends on AGI, Y_i, such that $D_i = \mathbf{1}[0 < Y_i \leq c]$ and 0 otherwise, and c is the threshold level of adjusted gross income to receive the high credit rate. The average treatment effect from receiving the high credit as opposed to the low credit is estimated as τ. In the absence of bunching, ϵ_i will be orthogonal to D_i yielding a consistent estimate for τ. As noted above, bunching has the potential to cause the orthogonality condition to be violated. For now I will proceed with the caveat that the estimate τ may be biased due to the selection; however, I will return to this issue and address the bias by placing bounds on the true value.

Figure 2.19 shows the average savings contributions within bins of \$200.[17] No visible change is apparent in savings behavior at the notch, which is confirmed by the estimates of the treatment effect. These results are summarized in Table 2.10. I fail to reject the null that the average treatment effect of receiving a high credit rate is significantly different than zero. This says that within a band around the notch, having a higher credit rate does not appear to be associated with having higher contribution levels. Thus, the higher credit rate does not appear to affect the

[17] Recall from Section 3.4, that savings are calculated based on Saver's Credit claims. This means that savings are necessarily bounded by \$2000 for both single and head of household filers and \$4000 for married couples filing jointly. Also, due to the credit's non-refundability, savings are also bounded by tax liability. The latter constraint is more relevant for the lower half of the income distribution, however, because tax liability is continuous across the notch, it should not present an issue in estimating the results.

savings contributions behavior for the marginal saver.

2.7.1 Bounds on the Results

Given that bunching occurs, I can account for the potential bias by bounding the estimated treatment effect. Lee (2002) shows that when observations are missing in a non-random way, the data can be trimmed in order to obtain bounds for the average treatment effect. In the context of this paper, the missing observations can be interpreted as observations that would have received a low credit rate if income were completely exogenous, but instead received the high credit rate. Because individuals to the left of the notch include both those that should have received the credit and those who bunched to receive the credit, Lee (2002) shows that the average treatment effect can be interpreted as a weighted average of the two groups. If people bunch monotonically, then by trimming the observations, the weighted average can be translated into bounds for the true treatment effect.[18]

In the absence of bunching the treatment effect of the Saver's Credit at the notch is given by

$$\tau = E[A_i|D = 1] - E[A_i|D = 0].$$

Suppose a fraction, ρ, of those who receive the high credit would have received the high credit without bunching, while $(1 - \rho)$ should have received the low credit but bunched in order to receive the high credit. Then the treatment effect being measured is

$$\rho E[A_i|D = 1, B = 1] + (1 - \rho)E[A_i|D = 1, B = 0] - E[A_i|D = 0, B = 0],$$

where B is an indicator that takes on a value of one if someone is a buncher. This

[18] A similar approach is used in Sallee (2009).

equation can be rewritten so that

$$E[A_i|D=1,B=0]-E[A_i|D=0,B=0]+\rho(E[A_i|D=1,B=1]-E[A_i|D=1,B=0]),$$

which includes the true treatment effect and the bias from those who bunch or,

$$\tau|B=0+\rho(E[A_i|D=1,B=1]-E[A_i|D=1,B=0]). \tag{2.5}$$

The true treatment effect can be bounded if assumptions are made that the bunching occurs only in one direction. In other words, individuals with either the highest values of A_i or the lowest values of A_i bunch.

I calculate the proportion of observations to be trimmed by taking the difference in actual and counterfactual densities for the high credit and dividing by the actual density for the high credit.[19] I calculate the amount of data to be trimmed from the high credit takers to be 2.9%. Table 2.11 presents the results from trimming the data for observations within $1,000 of the notch. The lower bound is obtained by trimming those with the highest values for A_i from the treatment group, which would account for people who bunch due to high values of savings. The upper bound is obtained by trimming the lowest values for A_i from the treatment group, where this would be the absolute highest value for a treatment effect if only those with low values of savings were induced to bunch. The large range within which the true treatment effect lies shows again that the policy's differential impact for those receiving the higher credit versus the lower credit is weak at best, even with the bunching present. As a percentage of possible savings eligible for a credit, the increased credit rate elicits at most an additional 5% in contributions, though this estimate is imprecisely measured. Thus, in terms of the price effect the credit has on savings, the impact

[19] For a discussion of calculating the proportion to be trimmed see Lee (2002) and Leibbrandt et al. (2005). The problem posed in this paper is slightly different because the number of people that selected into the high credit is unknown. However, I use the counterfactual distribution as a means to determine the number of additional observations that were treated.

appears to be neither statistically nor economically significant.

2.7.2 Discussion

Price insensitivity with respect to retirement contributions has been found in past work looking at retirement contributions behavior.[20] Engelhardt and Kumar (2007) find that savings contributions respond inelastically to 401(k) matches, where a 25% increase in match is associated with a $365 increase in contributions. Although this estimate is modest, it is much larger than the one I obtained from this study. However, my study focuses only on low and middle income households who may be even less price sensitive due to credit constraints, or other reasons, when it comes to making retirement contributions.

Related work that focuses specifically on savings behavior of low income households includes work done on Individual Development Accounts (IDA), which are accounts that encourage low income individuals to save for a particular purpose such as buying a house or car. Deposits into IDAs are matched by non-profit organizations with the goal of aiding assets to build to achieve the set goal. These studies provide an additional comparison for how matching impacts the savings behavior of the targeted demographic of the Saver's Credit. Mills et al. (2006) look at how the match incentives from IDAs affect behavior and find a large take-up rate by those offered the program, but that roughly half of those participants withdrew their funds for non-matchable purposes. Again, the match appears to impact behavior on the extensive rather than the intensive margin. These findings are in line with a study by Schreiner et al. (2001) who provide a more comprehensive study on IDAs. Thus, my finding for a small increase in retirement contributions not statistically different from zero the notch is consistent with past studies of low income savers.

[20]Engelhardt and Kumar (2007), Bernheim (2003), Hubbard and Skinner (1996) and Poterba et al. (1996) all provide summaries of studies on how matching retirement contributions affects retirement savings.

But this is somewhat surprising in the context of the Saver's Credit. There is little response to the price change in retirement contributions even when people select into receiving the higher credit. Given that those with higher savings have the highest incentive to bunch, it's puzzling that when they are included in the sample, the effect of price on savings is still relatively small and insignificant. These results may reflect that everyone who saves has the incentive to bunch, and those who do so may simply have a lower cost to altering their reported income. If this is the case, then the bunching should not induce bias and the OLS results capture the true effect, that people appear to be insensitive to the price of retirement contributions.

2.8 Conclusion

The Saver's Credit, a policy designed to increase retirement savings, also creates notches, or discontinuities, in a household's budget constraint. These notches give households an incentive to misreport income in order to receive a higher credit rate, thereby creating inefficiencies. In this paper, I show that people respond to incentives created by the Saver's Credit by bunching their adjusted gross income (AGI) below the notch. I add to the literature by repurposing an econometric technique developed to test an identifying assumption for regression discontinuity, in order to estimate the behavioral response to a policy. In particular, I use a test developed by McCrary (2008) to evaluate whether people alter their AGI to get a discontinuous increase in credit.

The Saver's Credit presents a large economic incentive to bunch and I find that individuals indeed respond. The evidence of bunching is strong, with a statistically significant break in the density of AGI at the notch. This is in contrast to past studies that show there is no bunching when the incentives are small. However, the Saver's Credit presents large incentives with as much as $600 to be gained with a $1 change in AGI. As such, these findings contribute to the growing literature that argues people will bunch, but incentives explored previously are too small to induce behavioral change. However, relative to the size of the program, the number of individuals who bunch is small and given that most do not save the maximum amount, the cost to the government is negligible.

The regression discontinuity estimates I obtain show that the change in credit rate has a modest, imprecisely measured positive effect on the level of savings contributions for people within $1,000 of the notch. Given that the estimates change very little when bunchers are accounted for, this finding also suggests that bunchers make similar retirement contributions to those that do not bunch. This presents something of a puzzle. Economic theory dictates that people with the largest contributions should also have the most incentive to bunch since they have the most to gain from the increased credit rate. In addition the higher credit rate lowers the effective price of making contributions, which should also raise contribution levels. However, this behavior fails to materialize in the data. It may be the case that individuals do not factor in the credit rate when making contributions, and thus their behavior is unaltered by differing rates on savings. This suggests that people are not forward looking with respect to receiving the credit, and bunching may occur after the savings decision in order to minimize one's tax liability. Another possibility is that people are credit constrained. For example, if an individual has a very high discount rate due to credit constraints, increasing the return on contributions will have little to no effect. Future work should take a deeper look at the puzzle left in terms of the lack of response to the Saver's Credit with respect to savings contributions.

This paper finds strong evidence that some households respond to incentives by bunching their incomes so as to receive the high credit rate but they represent a relatively small fraction of all credits claimed. This paper also finds that households receiving a higher credit rate do not alter their savings behavior significantly relative to those that receive a lower credit rate. The implication of these findings is that the Saver's Credit induces a behavioral response but not the intended one. People who take the credit appear to understand the incentive to bunch and the incentive is large enough to do so, yet they are insensitive to changes in the price of retirement contributions.

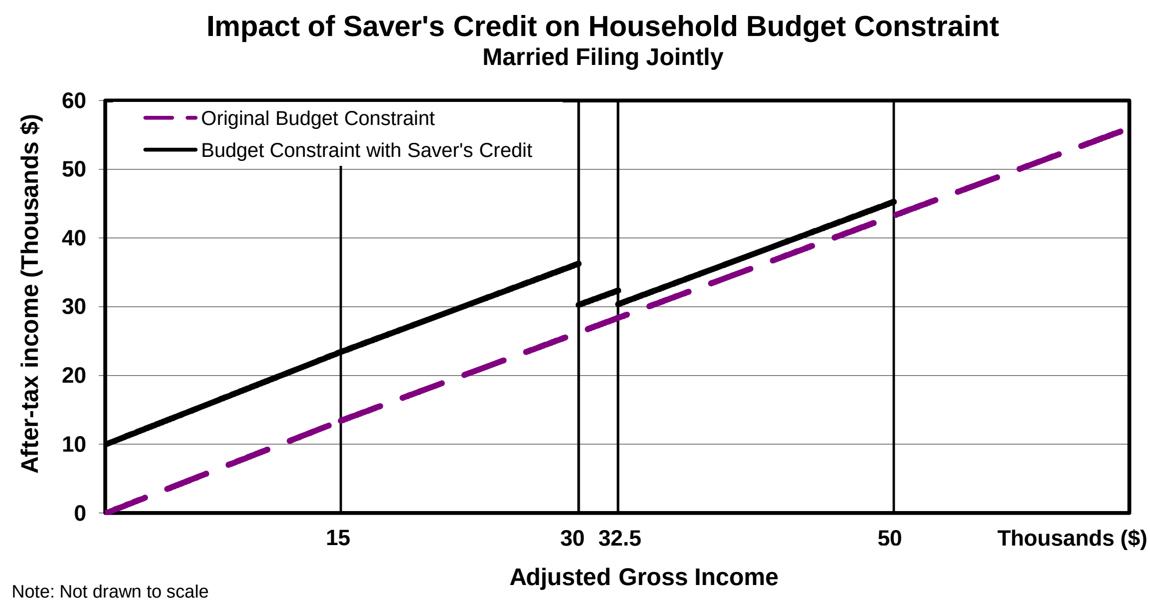
2.9 Figures and Tables

Table 2.1: Terms of Saver's Credit: Credit Rates by Filing Status and Adjusted Gross Income

	Household's Adjusted Gross Income ($)		
Credit (%)	50	20	10
Single	0-15,000	15,001- 16,250	16,251- 25,000
Head of Household	0-22,500	22,501- 24,375	24,376- 37,500
Married filing Jointly	0-30,000	30,001- 32,500	32,501- 50,000
Individual Max Credit	1000	400	200
Effective Match (%)	100	25	11

Notes.– Effective match rate is calculated by $\frac{s}{1-s}$, where s is the credit rate. Between 2002 and 2006 these income thresholds were not indexed for inflation, though this changed in 2007 after the passing of the Pension Protection Act of 2006. The individual maximum credit is earned per person and thus married couples filing jointly can earn twice the maximum listed.

Figure 2.1: Impact of the Saver's Credit on After-Tax Income for a Given Amount of Savings, for Taxpayers who are Married Filing Jointly



Notes.– Amount of credit has been exaggerated to illustrate notches in budget set.

Table 2.2: Aggregate Statistics on the Saver's Credit, 2002-2006

	2002	**2003**	**2004**	**2005**	**2006**
Percent of Total Credits Filed, by AGI					
No AGI	0.0	0.0	0.0	0.0	0.0
Between \$1 and \$5,000	0.0	.	.	0.0	0.0
Between \$5,001 and \$10,000	0.8	1.0	0.8	0.7	0.8
Between \$10,001 and \$15,000	4.9	4.9	4.6	5.6	5.0
Between \$15,001 and \$20,000	13.1	12.6	11.0	12.1	11.7
Between \$20,001 and \$25,000	20.2	20.1	20.7	20.9	22.8
Between \$25,001 and \$30,000	12.2	12.7	13.3	13.0	12.2
Between \$30,001 and \$40,000	23.9	24.2	25.0	24.6	25.0
\$40,000 under \$50,000	24.9	24.4	24.4	23.5	22.5
Percent Returns with AGI < \$50,000	71.2	70.6	69.1	67.9	66.7
Take up rate of returns with AGI < \$50,000	5.7	5.8	5.8	5.8	5.6
Average credit amount	199	195	191	178	172

Notes.– Source: IRS, Statistics of Income, Table 3.3: Individual Income Tax, All Returns: Tax Liability, Tax Credits, and Tax Payments, by Size of Adjusted Gross Income. Calculation for take up rate comes from taking the total number of Saver's Credits filed and dividing by the total number of returns with AGI below \$50,000. This number is a very rough estimate for the actual take up rate given that I cannot account for tax liability, savings or filing status.

Table 2.3: Aggregate Statistics on Taxpayers with IRAs

	2000	2001	2002	2004
AGI between $1 and $50,000				
Fraction of total pension participation	0.47	0.46	0.46	0.42
Fraction of Eligible that contribute to an IRA	0.17	0.21	0.23	0.23
Fraction of all IRA contributions	0.40	0.40	0.41	0.38
Average IRA contribution ($)	1677	1685	2023	2203
AGI between $1 and $25,000				
Fraction of total pension participation	0.18	0.16	0.16	0.16
Fraction of Eligible that contribute to an IRA	0.13	0.16	0.18	0.17
Fraction of all IRA contributions	0.17	0.17	0.18	0.17
Average IRA contribution ($)	1519	1523	1808	1811

Notes.– Source: IRS, Statistics of Income, Table 2.–Taxpayers with Individual Retirement Arrangement (IRA) Plans, by Size of Adjusted Gross.

Table 2.4: Summary Statistics on Saver's Credit Filers, 2002-2004

	All Savers Credit Filers	All Eligible Non-Credit Filers	Diff in Means
Proportion with Profit or Loss from Business (Sched. C)	0.19	0.11	0.08 *(0.000)*
Proportion with Capital Gains and Losses (Sched. D)	0.12	0.10	0.02 *(0.000)*
Proportion with Supplemental Income or Loss (Sched. E)	0.10	0.07	0.03 *(0.000)*
Proportion with Self-Employed Income	0.13	0.08	0.05 *(0.000)*
Proportion of Itemizers	0.25	0.14	0.11 *(0.000)*
Proportion of Tax Liability Withheld	3.68	3.73	-0.05 *(0.937)*
Mean amount of Saver's Credit ($)	196		
Mean amount of savings ($)	1,198		
Mean AGI ($)	30,296	21,636	8,661 *(0.000)*
Mean Taxes Owed ($)	1,643	1,073	570 *(0.000)*
Filed Electronically	0.65	0.49	0.16 *(0.000)*
Used a Paid Preparer	0.67	0.57	0.11 *(0.000)*
Unweighted Number of Observations	7,718	51,849	

Notes.– P Value of difference in parentheses.

Table 2.5: Factors Affecting the Probability of Filing for fhe Saver's Credit when Eligible

Conditional on Eligibility	
	Average Marginal Effect
Received 50% Credit	-0.016**
	(0.004)
Single/Married Filing Separately	-0.012**
	(0.004)
Head of Household	0.020**
	(0.004
Possible Credit Amount (thousands)	0.016**
	(0.003)
Wages and Salary (thousands)	0.004**
	(0.000)
Has Children	0.003
	(0.004)
Filed for EITC	0.020**
	(0.005)
Self-Employed	0.099**
	(0.005)
E-filed	0.039**
	(0.003)
Paid Preparer	0.021**
	(0.003)
Number of Observations	58,797
Pseudo-R^2	0.14

Notes.– Eligibility is based on falling below the threshold income level and having positive tax liability, it does not account for whether a person contributed to a retirement plan. Standard errors in parentheses, ** Significant at a 95% level.

Figure 2.2: Kernel Density of Normalized AGI, 2002-2004: Plug-In Bandwidth

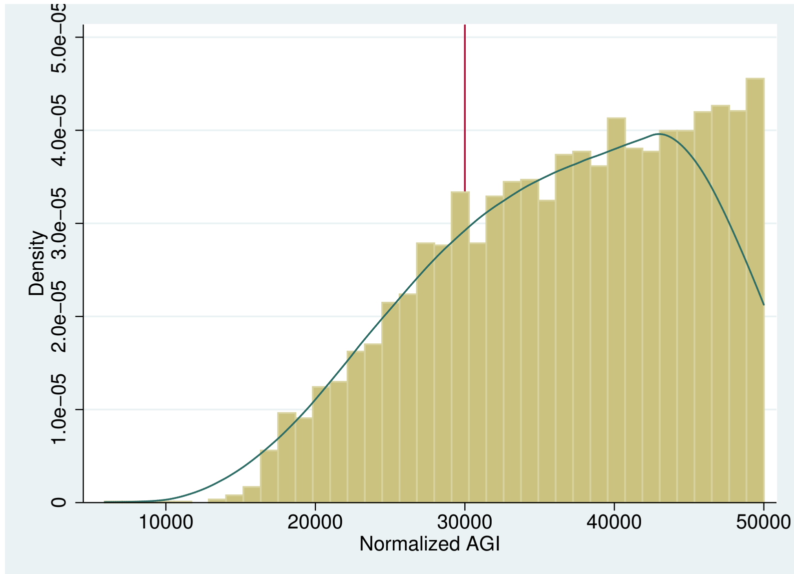
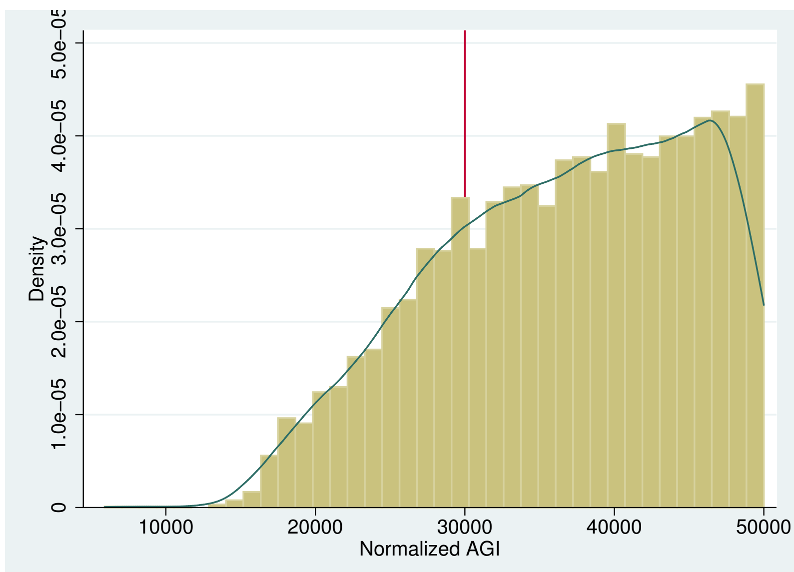


Figure 2.3: Kernel Density of Normalized AGI, 2002-2004: Half Plug-In Bandwidth



Notes.– Silverman's Plug-In Bandwidth of 3800 is used for kernel density along with Epanechenikov kernel weights. Binsize on histogram is 1100.

Figure 2.4: Carefully Defined Histogram

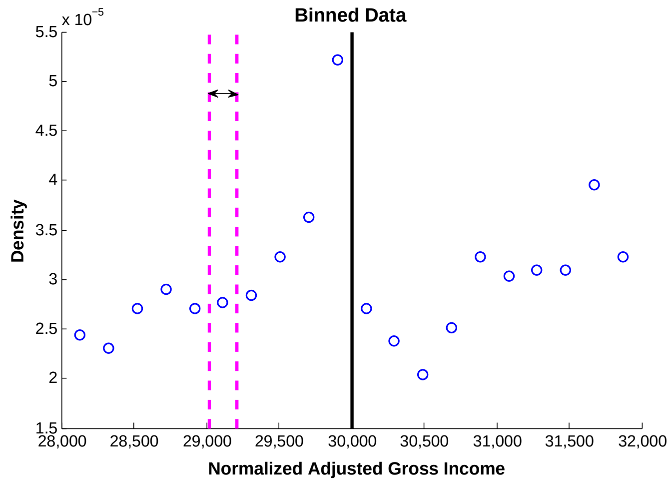


Figure 2.5: Smooth Histogram with Local Linear Regression

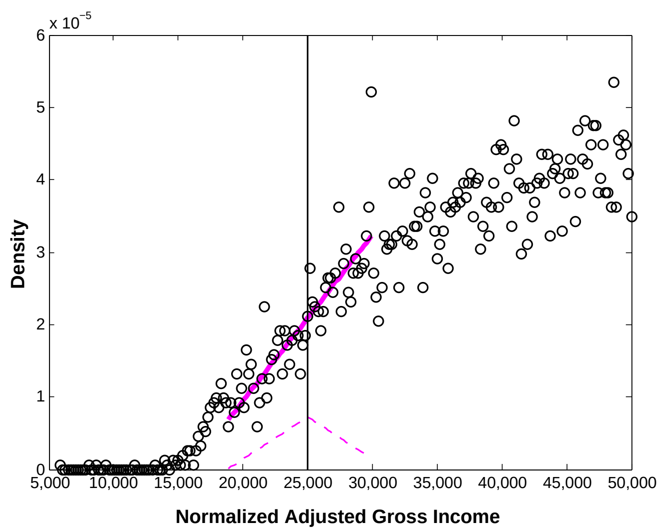


Figure 2.6: Estimated Density of Adjusted Gross Income, 2002-2004

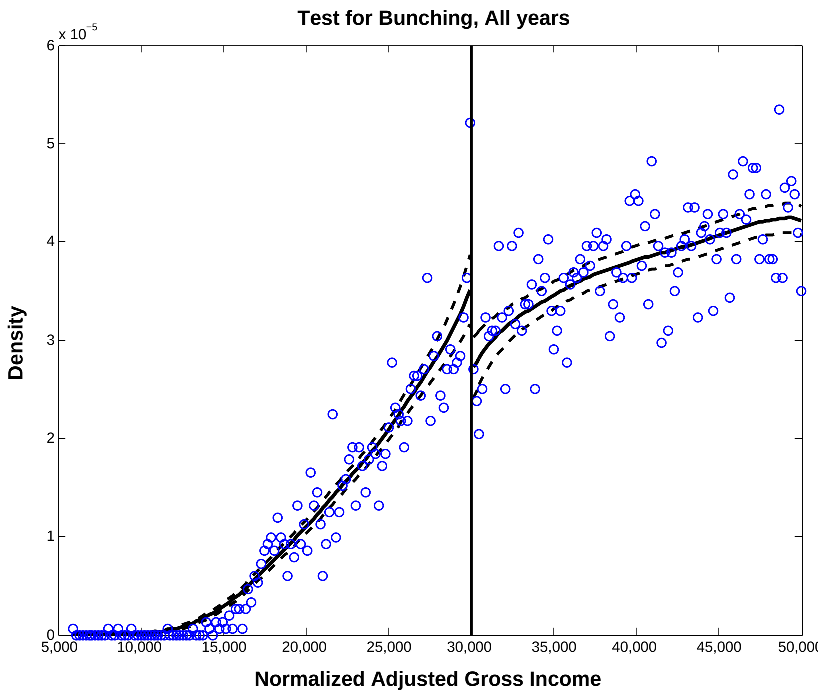


Table 2.6: Test for Break in the Estimated Density of AGI, 2002-2004

$\hat{\theta}$	-0.267*
	(0.08)
Bin Size	196.90
Bandwidth	6,523
Number of Observations	7701

Notes.– (Figure) Dashed line represents 95% confidence bands. Circles represent undersmoothed histogram of data. (Table) Standard errors in parentheses. *Significant at the 95% level.

Table 2.7: Test for Break in the Estimated Density of AGI using Alternative Bandwidths, 2002-2004

$\hat{\theta}$	-0.468*	-0.191*	-0.164*	-0.066
	(0.11)	*(0.06)*	*(0.06)*	*(0.04)*
Bin Size	196.89	196.89	196.89	196.89
Bandwidth	3500	10,000	12,000	20,000
Number of Observations	7701	7701	7701	7701

Notes.– Standard errors in parentheses. *Significant at the 95% level.

Figure 2.7: Test for Break at $40,000

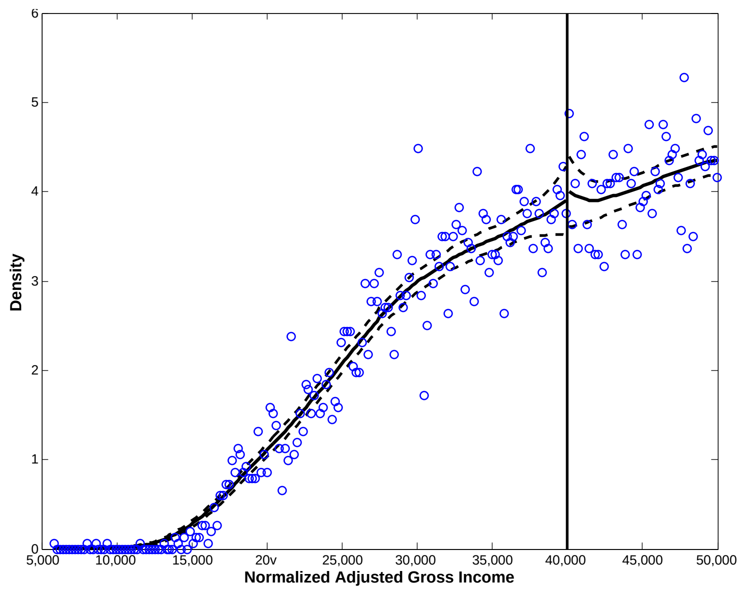
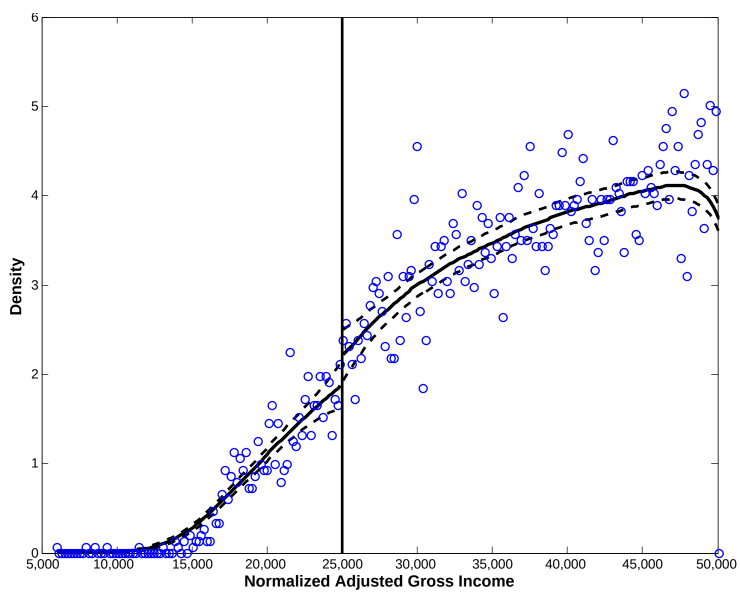


Figure 2.8: Test for Break at $25,000



Notes.– Dashed line represents 95% confidence bands. Circles represent undersmoothed histogram of data.

Figure 2.9: Schedule C Filers

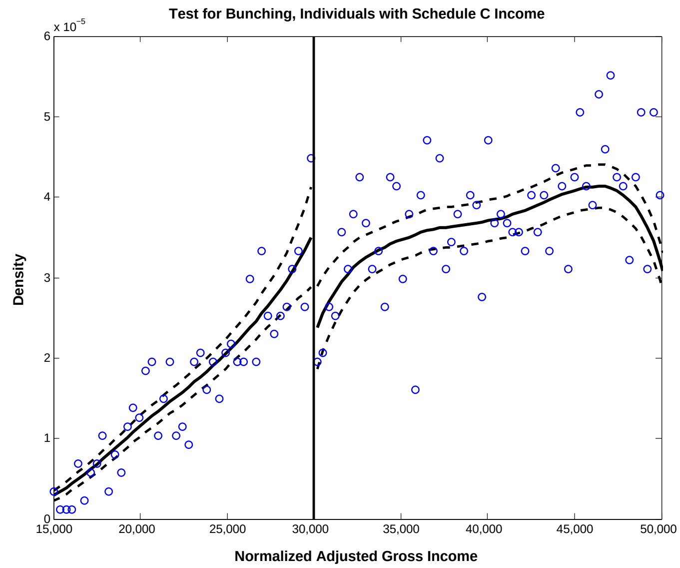
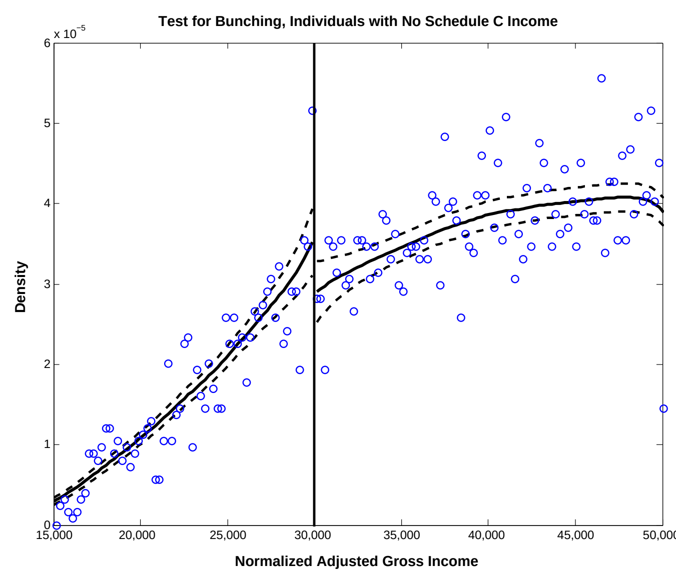


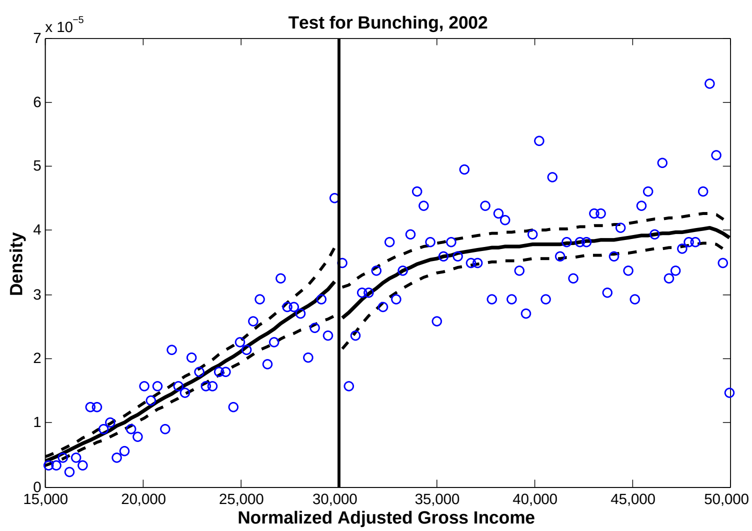
Figure 2.10: Non-Schedule C Filers



Notes.– Dashed line represents 95% confidence bands. Circles represent undersmoothed histogram of data.

Figure 2.11: Wage-Only Filers

Figure 2.12: Estimated Density of Adjusted Gross Income, 2002



Notes.– Dashed line represents 95% confidence bands. Circles represent undersmoothed histogram of data.

Figure 2.13: Estimated Density of Adjusted Gross Income, 2003

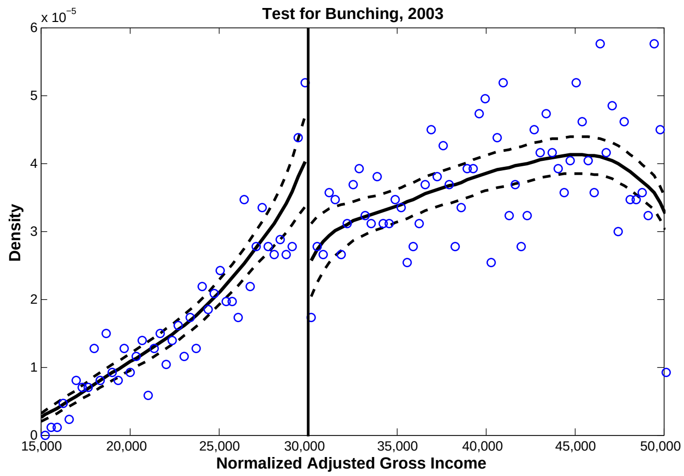
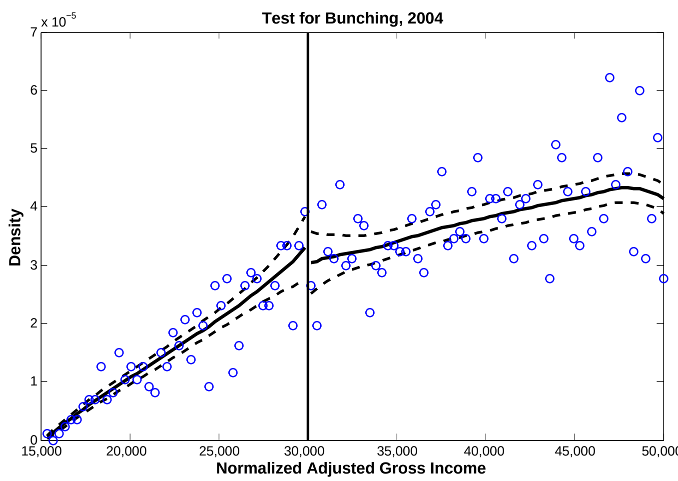


Figure 2.14: Estimated Density of Adjusted Gross Income, 2004



Notes.– Dashed line represents 95% confidence bands. Circles represent undersmoothed histogram of data.

Table 2.8: Test for Break in the Estimated Density of AGI, by year

	All	2002	2003	2004
$\hat{\theta}$	-0.267*	-0.212	-0.386*	-0.090
	(0.08)	*(0.13)*	*(0.13)*	*(0.10)*
Bandwidth	6,523	8,027	7,432	11,121
Joint test for equality (P value)	0.06			
Different from 2002 (P value)		.	0.16	0.00
Different from 2003 (P value)		.	.	0.05
Number of Observations	7,701	2,559	2,563	2,579

Notes.– Standard errors in parentheses, * Significant at the 95% level. Results from Wald tests for equality of all years together and then for each pair of years are included, with the p-values reported in the table.

Figure 2.15: Estimated Density of Adjusted Gross Income Excluding EITC Filers, 2002

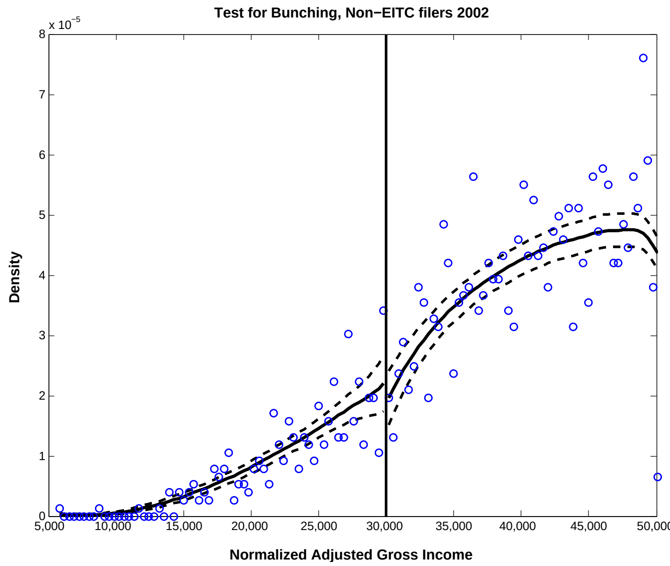
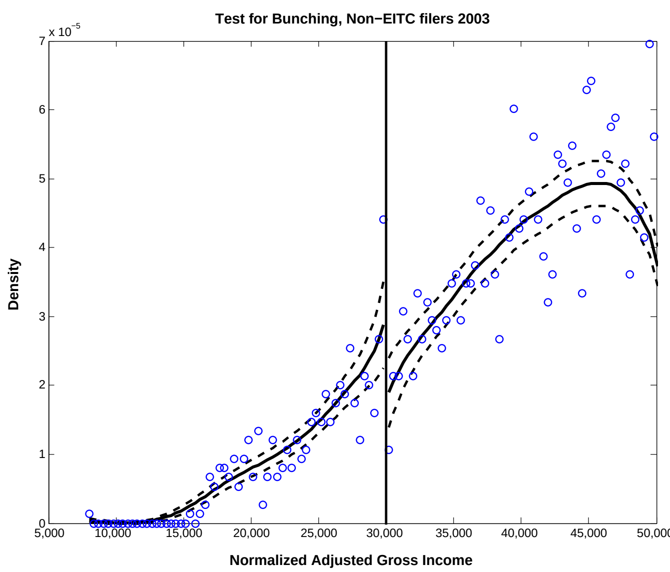


Figure 2.16: Estimated Density of Adjusted Gross Income Excluding EITC Filers, 2003



Notes.– Dashed line represents 95% confidence bands. Circles represent undersmoothed histogram of data.

Figure 2.17: Estimated Density of Adjusted Gross Income Excluding EITC Filers, 2004

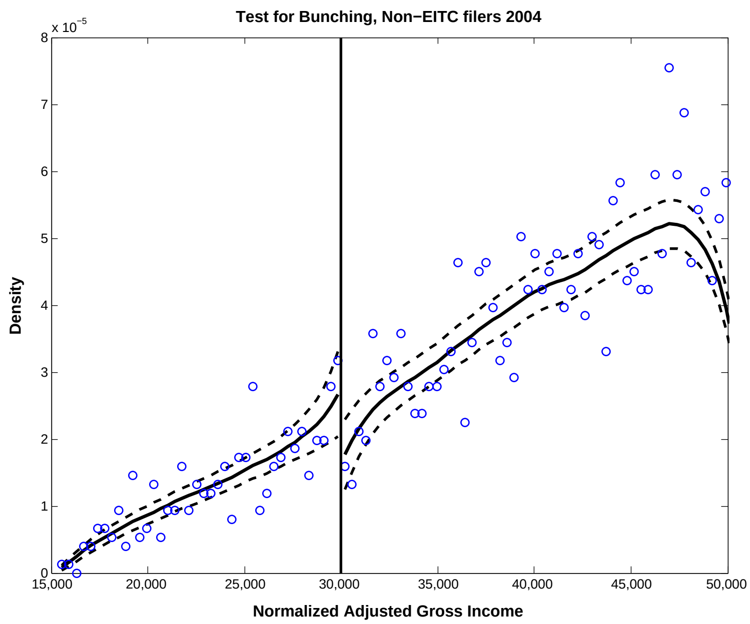


Table 2.9: Test for Break in the Estimated Density of AGI, excluding EITC Filers

	All	2002	2003	2004
All Taxpayers $\hat{\theta}$	-0.267*	-0.212	-0.386*	-0.090
	(0.08)	*(0.13)*	*(0.13)*	*(0.10)*
Non-EITC $\hat{\theta}$	-0.322*	-0.123	-0.349*	-0.353*
	(0.10)	*(0.16)*	*(0.15)*	*(0.19)*
Bandwidth	6,697	9,009	9,013	6,208
Number of Observations	6,225	2,065	2,092	2,068

Notes.– (Figure) Dashed line represents 95% confidence bands. Circles represent undersmoothed histogram of data. (Table) Standard errors in parentheses. *Significant at the 95% level.

Figure 2.18: Percentage of Bunchers Compared to a Counterfactual Distribution

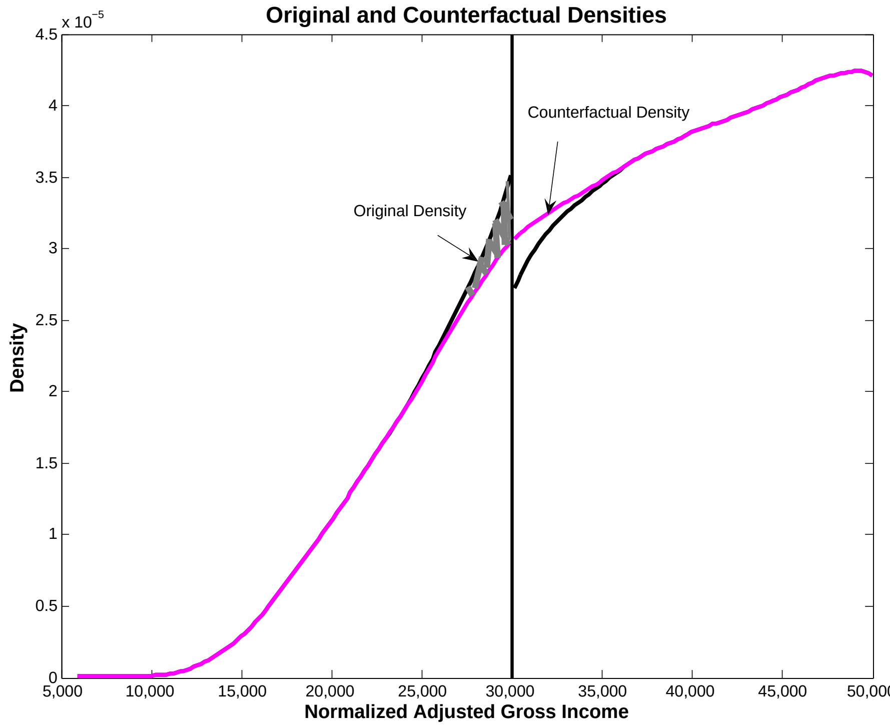


Figure 2.19: Average Retirement Contributions around Notch

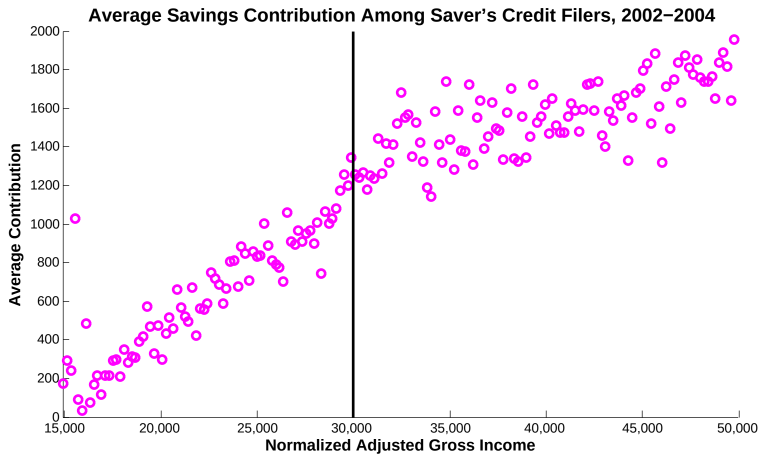


Table 2.10: Effect of a Change in Credit Rate on the Level of Retirement Contributions Close to the Notch

	Notch ± \$1,000	Notch ± \$2,000	Notch ± \$3,000	Notch ± \$4,000	Notch ± \$5,000
Avg contribution below the notch (α_{left})	1,173.96	1,089.08	1,016.04	1,017.39	979.66
	(83)	*(62)*	*(52)*	*(45)*	*(40)*
Avg contribution above the notch (α_{right})	1,077.21	1,028.30	1,011.33	1,045.90	1,048.40
	(114)	*(77)*	*(61)*	*(52)*	*(46)*
Average Treatment Effect	**96.75**	**60.78**	**4.71**	**-28.51**	**-68.74**
($\alpha_{left} - \alpha_{right}$)	*(140)*	*(99)*	*(80)*	*(69)*	*(61)*
N Left	322	586	819	338	406
N Right	270	575	316	419	524

Notes.– (Figure) Circles represent average level of retirement contributions that earned a Saver's Credit within bins of \$200. (Table) Standard errors in parentheses. None of the estimates are significant with 95% confidence.

Table 2.11: Trimmed Estimates ± $1000 Around the Notch

	Upper Bound	Lower Bound
α_{left}	1179.76	1056.09
	(83)	*(72)*
α_{right}	1,077.21	1,077.21
	(114)	*(77)*
Average Treat Effect	**102.55**	**-2.19**
$(\alpha_{left} - \alpha_{right})$	*(141)*	*(135)*
N Left	313	313
Obs Trimmed	9	9
N Right	270	270

Notes.– Standard errors in parentheses. None of the estimates are significant with 95% confidence.

CHAPTER III

Measuring the Change in Pre-Tax Wage Rates with Respect to Changes in the Personal Income Tax

3.1 Introduction

Do changes made to the personal income tax have explanatory power for pre-tax wage changes? Past empirical studies often assume the answer is no. For instance, studies that use changes in the marginal tax rate as exogenous variation in the net-of-tax wage often assume pre-tax wage rates are unaffected by the policy change. This exogenous variation is then used to measure behavioral parameters such as the labor supply elasticity.[1] If pre-tax wage rates indeed change when tax rates are altered, then the estimates obtained for these behavioral parameters would be biased. Despite the widespread use of the assumption that pre-tax wage rates are constant during a tax policy change, few empirical studies test whether this assumption holds true in the data. This paper aims to address this gap by empirically testing the nature of the relationship between pre-tax wage rates and the marginal tax rate on income.

A better understanding of the relationship between pre-tax wage rates and the marginal tax rate can also provide valuable insight into the incidence of the personal income tax. Assuming pre-tax wage rates are invariant to a tax change is akin to

[1] See Eissa (1995), Blundell et al. (1998),Eissa and Liebman (1996). Bosworth and Burtless (1992) offer a survey on studies estimating the various behavioral impacts of tax reforms.

assuming the entire burden of the income tax falls on the worker. This would happen, for example, in a general equilibrium framework if laborers were perfect substitutes for one another, which is shown more formally below. Studies that measure the progressivity of the overall tax system must assume the incidence of each tax (e.g. corporate tax, income tax, etc...) that make up the overall tax system in order to simulate the tax burden among different income classes.[2] Thus, knowing how pre-tax wage rates react to changes in the income tax has important implications for studies measuring the progressivity of the overall tax system.

This paper is the first study to use panel data to measure the effect of changes in marginal tax rates on pre-tax wage rates in the US. Past literature has used panel data to study a similar effect using policy changes in Scandanavian countries (Bingley and Lanot (2002) and Blomquist and Selin (2009)). Literature pertaining to the US is more sparse; a recent study analyzes the median worker in an occupation before and after the Tax Reform Act of 1986 (Kubik (2004)). We use panel data from the Survey of Income and Program Participations's (SIPP) 1986 panel, where we can observe one-year changes before and after the Tax Reform Act of 1986. Unlike previous studies, this paper provides separate estimates for continuing workers whose employment remains constant during the tax changes and for individuals who change jobs or occupations.[3] This distinction is important as it separates individuals whose wage rates are impacted by general equilibrium changes from those whose wage rates are driving the general equilibrium changes. Finally, this study uses a counterfactual tax rate to instrument for the actual tax rate to provide estimates that isolate the impact of the policy change on pre-tax wage rates. As such, this paper generates

[2]For example, Pechman and Okner (1974), Browning and Johnson (1979), Davies et al. (1984), and Pechman (1985) all estimate the overall progressivity of the US tax code by making the underlying assumption that labor taxes are borne solely by labor.

[3]Individuals who move in and out of non-employment are also excluded when excluding job and occupation changers as non-employment is coded as an occupation.

more complete and accurate estimates of the impact of tax rates on pre-tax wages for US workers than any previous study we are aware of.

One of the difficulties in estimating a relationship between pre-tax wage rates and the marginal tax rate on income is that an individual's tax rate is based in large part on wage earnings. When a person's wage rate increases, holding hours constant, her marginal tax rate could increase if she is bumped into a higher tax bracket. If such a relationship exists, this could mechanically bias our results. We deal with this issue by instrumenting for the change in the marginal tax rate using the change in a calculated counterfactual tax rate. Following Auten and Carroll (1999), we calculate a counterfactual tax rate by applying the new tax rules from the policy change to the pre-policy income. This method effectively allows us to rescale the change in marginal tax rate to reflect only the impact of the policy change. Thus, we are able to isolate the effect of the new tax rules on wage rates. In addition, we exclude individuals who changed jobs or occupations during the reform that potentially experienced a wage change for reasons unrelated to the tax change. By focusing on individuals that were more stable in their employment status, we obtain an estimate that is less likely to be biased by a potential mechanical relationship between changes in wage rates and changes in tax rates.

Following previous literature, we estimate the impact of the net-of-tax rate (1-τ) on pre-tax wage rates. We find that a 1% increase in the net-of-tax rate, where the tax rate is defined as the sum of both the federal and state marginal tax rates, is significantly associated with a 0.26% decrease in real wage for individuals remaining at the same job before and after the tax change. In other words, individuals moving from the highest bracket of 50% to a new bracket of 38.5%, would have faced a 6.0% decrease in their real wage. These results indicate that pre-tax wage rates are in fact affected by changes made to tax policy, contrary to the commonly used assumption. When only considering the change in federal tax rate, a 1% increase in the net-of-tax rate, is associated with a 0.31% decrease in real wage.

The rest of the paper is organized as follows: Section 3.2 provides background on the Tax Reform Act of 1986 and a brief overview of the literature on this topic; Section 3.3 outlines a theoretical framework; Section 3.4 describes the data; Section 3.5 presents our empirical estimation and results; Section 3.6 concludes.

3.2 Background

The 1980's marked a decade long era of tax reform in the United States. While there were numerous acts passed throughout the period, none were more drastic than the Tax Reform Act of 1986. The Tax Reform Act of 1986 (TRA86) completely restructured the US tax code by, among other changes, flattening the schedule of marginal tax rates on personal income.[4] Figure 3.1 illustrates the schedule of marginal tax rates before and after the reform. Marginal tax rates were simultaneously increased for the lowest income levels and lowered for the highest income levels. To counter the higher marginal rates for low incomes, the personal exemption and standard deduction amounts increased and the Earned Income Tax Credit was expanded. Overall, the changes made in the Tax Reform Act of 1986 were designed to be revenue neutral.

Since its inception, a large number of studies have used variation in tax rates generated by TRA86 to study behavioral responses to tax changes. While many studies have looked at the hours and the taxable income responses to tax changes, there are few studies that look at the change in pre-tax wage rates with respect to the

[4]For background on the Tax Reform Act of 1986 see Auerbach and Slemrod (1997) and Hausman and Poterba (1987).

tax rate.[5] The goal of this paper is to provide an empirical investigation into whether individuals facing changes in their marginal tax rate experienced subsequent changes in their pre-tax wage rates. Although there have been a number of changes made to the personal income tax since TRA86, TRA86 still provides the largest variation in rates for the broadest group of individuals. As such, we focus on changes to marginal tax rates made by TRA86 to study the relationship between pre-tax wage rates and taxes.

Hausman and Poterba (1987) simulate the overall impact of the 1986 reform on household behavior. They show, by extrapolating the 1983 tax returns using TAXSIM to determine the change in tax between 1988 and 1983, that only 13.3% of individuals faced the same marginal tax rate before and after the reform. Meanwhile, 47.7% of individuals faced a decline of 0-10 percentage points. Thus for most individuals, the marginal tax rate on personal income indeed changed, with the most drastic changes concentrated in the highest income brackets. Given that changes in marginal tax rate were faced by a majority of the population, a change in pre-tax wage rates may have also resulted from the policy change.

A small but growing body of literature focuses on measuring the impact that changes in marginal tax rates have on pre-tax wage rates (Moffitt and Wilhelm (2002), Bingley and Lanot (2002), Kubik (2004), Blomquist and Selin (2009)). Kubik (2004) looks at changes in the US wage structure that occurred in response to the Tax Reform Act of 1986 and is most closely related to the analysis conducted in this paper. In Kubik's study, there are two types of labor: skilled and unskilled. Kubik argues that when changes are made to the tax system that alter the relative supply of skilled and unskilled labor, then the pre-tax wage rate should increase for the

[5] See Eissa (1995) and Eissa and Liebman (1996), for example. Saez et al. (2009) provide a survey on this literature.

type of labor that becomes relatively scarce. For instance, if a labor tax on skilled workers decreases, which in turn, increases the supply of skilled workers relative to unskilled workers, then the pre-tax wage of skilled workers should fall.[6] This change in the relative supply between skilled and unskilled workers hinges on the fact that the change in marginal tax rate affects an individual's labor supply decision. Kubik offers two mechanisms through which labor supply can adjust: through changes in participation and hours decisions and switching occupations. Kubik focuses on the latter of the two mechanisms and controls for the former by restricting his sample to include only men as men are generally regarded as unresponsive in hours decisions with respect to a wage change. If an individual were to switch occupations in response to a tax change, then this would result in a change in the relative supply of workers within an occupation. The shift in relative supply among occupations would in turn result in changes to pre-tax wage rates. To test this hypothesis, Kubik uses repeated cross-section data from the CPS, which he collapses from the individual to the occupational level, to relate changes in the median wage within an occupation to changes in the median tax rate before and after the reform. He finds that, when the sample is restricted to include only men aged 25 to 55, occupations facing a 10 percentage point decline in median tax rate also faced a 2.5% decrease in median wage.

This paper extends the analysis in Kubik (2004) and similarly focuses on changes in wage rates and marginal tax rates before and after TRA86. There are a number of concerns stemming from the Kubik analysis that we aim to address by using panel data rather than repeated cross-section data. Kubik notes that if the tax change caused an increase of workers within a high income occupation, the median

[6] These results are derived making standard assumptions regarding the production function.

wage could be driven down by inexperience rather than decreases in the wage rates of existing workers. Kubik's solution is to restrict the sample to men, as women, and in particular married women, have been shown to respond to decreases in the marginal tax rate on income with increased labor force participation (Eissa (1995)). By focusing on the subset of individuals who remained at the same job before and after the tax change, we avoid the issue that an influx of inexperienced workers would bottom weight the distribution and thus change the median wage. In doing so, we also use information from the full distribution of wage rate and tax rate changes rather than focusing on one point in the distribution as in Kubik (2004).

Blomquist and Selin (2009) and Bingley and Lanot (2002) both provide estimates relating changes in pre-tax wage rates to changes in marginal tax rates on income. Blomquist and Selin (2009) use panel data from the Swedish Level of Living Survey to derive an elasticity of pre-tax hourly wage with respect to changes in the net-of-tax rate. With data on the same individuals in 1981 and in 1991, they use a first difference approach to obtain wage elasticities with respect to the net-of-tax rate. Their estimated elasticities lie between 0.14 and 0.16 for married men and between 0.41 and 0.57 for married women, which are large in light of the assumption that the elasticity is zero. These elasticities are all precisely estimated but, notably, they are of the opposite sign the results in Kubik (2004) and in contrast to a standard model. In particular Blomquist and Selin find that if the tax rate increases, making the net-of-tax rate decrease, then gross wage rates will actually decrease. However, this finding is in line with their argument that a decrease in the tax rate could have an impact on an individual's behavior including a movement towards more difficult and better compensated tasks, increased effort in wage bargaining, increased work intensity on the job, and changes in forms of compensation. Thus, a tax rate decrease could in fact drive an increase in the wage rate.

Bingley and Lanot (2002) use a Danish matched establishment-worker panel of individual workers spanning 1980 to 1991 to study the impact of tax changes on gross wage rates. Their data allow for tracking the same individuals at the same establishment over time. Because the data is collected at the firm level, if a person changes jobs, she is eliminated from their sample. The authors use a difference-in-difference approach which utilizes within-establishment variation in wage rates. They use local tax rates to instrument for the endogeneity of marginal tax rates and estimate an elasticity of gross wage rates with respect to the marginal rate of income tax of 0.44. Similar to Kubik (2004), Bingley and Lanot find that increasing the tax rate on income leads to an increase in gross wage rates.

In summary, the past literature that estimates a relationship between pre-tax wage changes and marginal tax rates shows mixed results. Whereas Kubik (2004) and Bingley and Lanot (2002) estimate a positive relationship, Blomquist and Selin (2009) instead estimate a negative relationship. However, Blomquist and Selin argue this reverse finding is not surprising and provide a story for why a tax decrease could lead to an increase in wage rate. Both Bingley and Lanot (2002) and Blomquist and Selin (2009) conduct their analysis with panel data in countries with tax systems drastically different to that found in the US. Also, both countries contain a stronger union presence than in the US, which makes generalizing their results to the US more difficult. Kubik (2004), on the other hand, analyzes pre-tax wage changes in the US. However, he focuses only on one point in the distribution of wage rates by occupation to determine how the median wage rate within an occupation changes with respect to changes to the median tax rate.

Our analysis contributes to this small literature by offering an additional estimate for changes in wage rates with respect to marginal tax rates on income in the US. We make use of the full distribution of changes in wage rates to measure the effect of changes in the tax rates for continuing workers. By focusing on people who are relatively stable in their employment, we avoid the concern raised in Kubik (2004) that identifying the impact of tax rates off of the median wage within an occupation may also include the impact of changes in the distribution of experience within an occupation. Our study complements Kubik (2004) in that we find similar results even after excluding individuals who could be experiencing changes in wage rates for reasons other than the policy change.

3.3 Theory

As seen in Figure 3.1, the Tax Reform Act of 1986 had a differential impact on individuals in the higher income brackets compared to those in the lower brackets. This variation in tax changes allows for identifying the impact of taxes on wage rates. However, a typical simplification that workers are identical no longer holds as we observe many wage rates and many tax changes. To motivate the empirical section, we consider a one sector model with the simplifying assumption that there are two types of labor to analyze the effects of a type-specific labor tax.[7] The model we present is a static model that we use to analyze short-run responses of wage rates to tax changes. We show that the effect of a tax change on wage rates for one type of worker will largely depend on the firm's elasticity of substitution over the different types of labor as well as each labor type's preferences for working.

[7] We use a framework and notation similar to that of Fullerton and Metcalf (2002). The model presented in this paper has one sector with two types of labor and ignores the role of capital, whereas Fullerton and Metcalf present both a one-sector and a two-sector model each with capital and one type of labor. Kubik (2004) includes a model with two types of labor; however, one type of labor is supplied inelastically.

3.3.1 Firm's Problem

Suppose there are two types of labor, L_1 and L_2, used as inputs into a production function, $F(L_1, L_2)$, that produces output, X. The production function is assumed to exhibit constant returns to scale and the price of X is normalized to one. We assume that the labor market is perfectly competitive and workers are paid their marginal product, so that

$$F_1(L_1, L_2) = w_1$$

and

$$F_2(L_1, L_2) = w_2.$$

With constant returns to scale technology, profits must be zero, giving us the following condition:

$$X = w_1 L_1 + w_2 L_2.$$

Letting $\hat{Z}$ denote the percent deviation from equilibrium (dZ/Z) and taking the log differential of the production function gives us,

$$\hat{X} = \left(\frac{w_1 L_1}{X}\right) \hat{L}_1 + \left(\frac{w_2 L_2}{X}\right) \hat{L}_2.$$

Letting $\theta_1 = \frac{w_1 L_1}{X}$ and $\theta_2 = \frac{w_2 L_2}{X}$, we can rewrite this equation as

$$\hat{X} = \theta_1 \hat{L}_1 + \theta_2 \hat{L}_2.$$

From the constant returns condition we also have that

$$\hat{X} = \theta_1(\hat{w}_1 + \hat{L}_1) + \theta_2(\hat{w}_2 + \hat{L}_2).$$

After substituting for $\hat{X}$ from the production function, we can derive an equation for the $\hat{w}_1$ in terms of $\hat{w}_2$.

$$\hat{w}_1 = -\frac{\theta_2}{\theta_1}\hat{w}_2$$

The definition for the elasticity of substitution, σ_x between L_1 and L_2 is given by,

$$\sigma_x(\hat{w}_1 - \hat{w}_2) = \hat{L}_2 - \hat{L}_1.$$

Finally, substituting for $\hat{w}_1$ and using $\theta_1 + \theta_2 = 1$, gives an equation for the relative demand for $\hat{L_2}$ and $\hat{L_1}$

$$-\frac{\sigma_x}{\theta_1}\hat{w}_2 = \hat{L}_2 - \hat{L}_1. \tag{3.1}$$

3.3.2 Workers' Problem

For the worker's problem, we assume each type of worker is endowed with one unit of time. The worker's time is allocated between market work, L, which is an input into the production of X, and leisure, l, so that

$$L_i + l_i = 1,$$

where $i \in 1, 2$. Workers derive utility from consumption of output, X, and leisure l, $(1 - L)$. Both types of workers are assumed to have the same CES utility with the following functional form:

$$U(X_i, (1 - L_i)) = \left((X_i)^{\left(\frac{\gamma_i - 1}{\gamma_i}\right)} + (1 - L_i)^{\left(\frac{\gamma_i - 1}{\gamma_i}\right)}\right)^{\left(\frac{\gamma_i}{\gamma_i - 1}\right)}$$

where γ_i represents the type-specific elasticity of substitution between X and l. Each worker's objective is to maximize utility subject to the budget constraint

$$X_i = w_i L_i (1 - \tau_i) + G,$$

where G is a lump sum transfer from the government and τ_i is a tax on wage income for type i.

After solving for type i's respective maximization problem and taking log derivatives we obtain:[8]

$$\hat{X}_i - \hat{l}_i = \gamma_i(\hat{w}_i - \hat{\tau}_i).$$

[8]The economy is assumed to have no initial taxes.

From the worker's budget constraint, we also know that

$$\hat{X}_i = \theta_i(\hat{L}_i + \hat{w}_i - \hat{\tau}_i).$$

Using the worker's time constraint, we can derive an equation for labor as a function of leisure,

$$\hat{L}_i = \left(\frac{(1 - L_i)}{L_i}\right)\hat{l}_i.$$

Letting $\phi_i = \frac{1-L_i}{L_i}$ and substituting for $\hat{X}_i$,

$$\gamma_i(\hat{w}_i - \hat{\tau}_i) - \phi_i\hat{L}_i = \theta i(\hat{L}_i + \hat{w}_i - \hat{\tau}_i),$$

which gives the labor supply of each type,

$$\hat{L}_i = \left(\frac{\gamma_i - \theta_i}{(\theta_i + \phi_i)}\right)(\hat{w}_i - \hat{\tau}_i).$$

Finally we can derive an expression for the change in relative labor supply between type 1 and type 2 workers as a function of the model's parameters, wage changes, and tax changes,

$$\hat{L}_2 - \hat{L}_1 = \left(\frac{\gamma_i - \theta_2}{\theta_2 + \phi_2}\right)(\hat{w}_2 - \hat{\tau}_2) - \left(\frac{\gamma_i - \theta_1}{\theta_1 + \phi_1}\right)(\hat{w}_1 - \hat{\tau}_1). \tag{3.2}$$

For each worker, $\left(\frac{\gamma_i - \theta_i}{\theta_i + \phi_i}\right)$ represents the uncompensated labor supply elasticity, where $\left(\frac{\gamma_i}{\theta_i + \phi_i}\right)$ is the substitution effect and $\left(\frac{-\theta_i}{\theta_i + \phi_i}\right)$ is the income effect. Substituting in for $\hat{w}_1$, we can derive an expression in terms of changes in only $\hat{w}_2$,

$$\hat{L}_2 - \hat{L}_1 = \left(\frac{\gamma_i - \theta_2}{\theta_2 + \phi_2}\right)(\hat{w}_2 - \hat{\tau}_2) + \left(\frac{\gamma_i - \theta_1}{\theta_1 + \phi_1}\right)\left(\frac{\theta_2}{\theta_1}\hat{w}_2 + \hat{\tau}_1\right).$$

Assuming that the government returns tax revenue to the households as a lump sum payment, the income effect is ignored and instead we are left with the compensated labor supply elasticity, $\eta_i^c = \left(\frac{\gamma_i}{\theta_i + \phi_i}\right)$, which gives[9]

$$\hat{L}_2 - \hat{L}_1 = \eta_2^c(\hat{w}_2 - \hat{\tau}_2) + \eta_1^c\left(\frac{\theta_2}{\theta_1}\hat{w}_2 + \hat{\tau}_1\right)$$

[9] TRA86 was designed to be revenue neutral, which lessens the potential impact of income effects.

3.3.3 Equilibrium

Now, substituting Equation (3.2) into Equation (3.1) to equate labor supply and labor demand, we can solve for $\hat{w}_2$ in terms of the tax changes,

$$\hat{w}_2 = \frac{-\eta_2^c \hat{\tau}_2 + \eta_1^c \hat{\tau}_1}{-\frac{\sigma_x}{\theta_1} - \eta_2^c - \eta_1^c \frac{\theta_2}{\theta_1}}. \tag{3.3}$$

Thus, we are left with an expression that relates the change in w_2 to changes in the labor tax on both type 1 and type 2 workers. Once we know how w_2 changes, we can also derive the change in w_1 using the relationship $\hat{w}_1 = -\frac{\theta_2}{\theta_1}\hat{w}_2$

To gain intuition, we can posit the special cases in production of perfect substitutes, perfect complements, and Cobb-Douglas and determine how wage rates will change for a given value of σ. For simplicity, suppose the change in τ_1 is zero. For a 1 percent change in τ_2,

$$\frac{\hat{w}_2}{\hat{\tau}_2} = \frac{\eta_2^c}{\frac{\sigma_x}{\theta_1} + \eta_2^c + \eta_1^c \frac{\theta_2}{\theta_1}}.$$

This shows that $\hat{w}_2$ will depend on the compensated elasticities of supply for both types of workers, the elasticity of substitution of the two types of workers in production, and their relative value shares in the output market. If L_1 and L_2 are perfect substitutes, that is $\sigma = \infty$, then the change in w_2 with respect to a change in τ_2 is zero. This also means the change in wage of type 1 workers is zero. Thus, type 2 workers bear the full burden of the tax with a take home wage of $w_2(1-\tau_2)$. This makes intuitive sense in that if a tax is imposed on type 2 workers, then the firm will respond by substituting to type 1 workers. Thus, the pre-tax wage change will be zero since the firm is only willing to pay one wage.

Now, suppose L_1 and L_2 are perfect complements, so that $\sigma = 0$. Then,

$$\frac{\hat{w}_2}{\hat{\tau}_2} = \frac{\theta_1}{\theta_1 + \left(\frac{\eta_1^c}{\eta_2^c}\right)\theta_2}.$$

The change in w_2 will depend on the relative compensated labor supply elasticities of type 1 and type 2 workers and the relative value shares that each type has in the output market. If $\eta_1^c > \eta_2^c$, then the change in w_2 with respect to $\hat{\tau}_2$ will be less than θ_1. On the other hand, if $\eta_1^c < \eta_2^c$, then the change in w_2 with respect to $\hat{\tau}_2$ will be greater than θ_1. Finally, if $\eta_1^c = \eta_2^c$, then $\frac{\hat{w}_2}{\hat{\tau}_2} = \theta_1$ and $\frac{\hat{w}_1}{\hat{\tau}_2} = -\theta_2$. Overall, the change in the wage of type 2 workers resulting from a tax on type 2 workers will be bounded between 0 and 1 and will depend on the compensated elasticity of supply between type 1 and type 2 workers if they are perfect complements in production.

An intermediate case between perfect complements and perfect substitutes is that of Cobb-Douglas production. When a Cobb-Douglas production function is used, $\sigma = 1$. This means that

$$\frac{\hat{w}_2}{\hat{\tau}_2} = \frac{\theta_1}{1 + \theta_1 + \left(\frac{\eta_1^c}{\eta_2^c}\right)\theta_2}.$$

If $\eta_1^c > \eta_2^c$, then the change in w_2 with respect to $\hat{\tau}_2$ will be less than $\frac{\theta_1}{2}$. On the other hand, if $\eta_1^c < \eta_2^c$, then the change in w_2 with respect to $\hat{\tau}_2$ will be greater than $\frac{\theta_1}{2}$. Finally, if $\eta_1^c = \eta_2^c$, then $\frac{\hat{w}_2}{\hat{\tau}_2} = \frac{\theta_1}{2}$ and $\frac{\hat{w}_1}{\hat{\tau}_2} = -\frac{\theta_2}{2}$.

The theory presented here shows that pre-tax wage rates need not remain constant when a tax is imposed on labor. One exception includes the case when different types of labor are perfect substitutes for one another in production. However, if this is not the case, then there will be a resulting change in pre-tax wage rates. In these cases the model suggests the change in the wage of type i resulting from a tax on only type i will be bounded between 0 and 1, and the magnitude of the change depends on the elasticity of substitution between the types of labor, the relative labor supply elasticities of the workers, and each type's share of value in output. Thus, a testable implication of the model is that an individual's wage rate should have a positive relationship with their own marginal tax rate (τ), or a negative relationship with their net-of-tax rate $(1 - \tau)$. To test this implication, we decompose the data into types defined by income, gender, and marital status. We then focus on the short-run response of wage rates to changes in the marginal tax rate on income.

There are a number of ways in which the data can be parsed in order to test the model's implications. In the model, labor types are distinguished by differing marginal productivities, which, in turn, determine wage rates and tax rates. In the data, a worker's true marginal productivity is unobserved. If an individual is actually paid her marginal product, as would be the case in a perfectly competitive market, then income can serve as a proxy for productivity. Thus, a natural empirical counterpart for labor types that vary in productivity is to separate the data by income.

The model also allows for differences in labor supply elasticity among labor types. Women in this era were generally believed to have a higher labor supply elasticity than men.[10] Thus, differences in labor supply elasticities between types can be examined by analyzing men and women separately. But grouping the data by sex highlights a simplification of the model, which is that in the model households are comprised of one individual whose tax rate is derived based on his or her type. However, in the US, married households are taxed based on their joint income, which may be earned by two individuals of differing types. For example, if one household member has a high wage rate, while the other has a low wage rate, then the low wage individual's marginal tax rate will not correspond to that member's productivity. As such, the model may be better suited to predict the behavior of single tax filers.

[10] See Pencavel (1986) and Killingsworth and Heckman (1986) for surveys of the literature on labor supply elasticities for men and women, respectively.

3.4 Data

The Survey of Income and Program Participation (SIPP), conducted by the U.S. Census Bureau, is a panel survey of households about income sources, labor supply, welfare participation, and other economic and family outcomes. Each individual in the household is surveyed every four months regarding each intervening month. We focus entirely on the 1986 SIPP panel, which covers the period between October 1985 and March 1988, as this panel includes monthly data for the full calendar years 1986 and 1987, the year before TRA86 and the year of its implementation.[11]

The SIPP collects information on each employed individual's jobs (up to two per person per interview wave), including hours worked, wage and/or salary, occupation, industry, and tenure in the job. About 59 percent of employees report only their monthly earnings in the job, and not the hourly wage; for these individuals, we divide monthly earnings by the product of weeks worked that month and the usual weekly hours worked in the four-month interview wave to get an imputed hourly wage. To get an annual wage figure, our outcome variable in most of our analysis, we take the mean over each monthly wage (reported where available, but often only the imputed wage) for that month's primary job, defined as the job with the highest hourly wage. Because this imputation process could result in some hourly wage figures that are unrealistically high (if, for example, few hours are reported, or if labor and non-labor income are confused), we winsorize wage rates at the 95th percentile, capping the 1986 hourly wage at $17 per hour and the 1987 wage at $18/hr.[12]

[11]Other SIPP panels also cover this period, but survey households for only part of these two key years. It may be useful to observe households during the full calendar year 1988 as well, in case the full effect of TRA86 is not felt until after taxpayers have adjusted to the new regime, but only the 1987 SIPP Panel covers this full year, and it lacks monthly data for most of the pre-period.

[12]The accuracy of the hourly wage measures in the SIPP data becomes an important question, as our results rest on using self-reported data that may be rife with measurement error. Stinson (2002) estimates the measurement error in SIPP data using administrative earnings data from the Social Security Administration (SSA) and finds that 18 percent of the variation in SIPP annual job earnings can be attributed to measurement error, comparable to the 21 percent measurement error found in the Detailed Earnings Records from the SSA. That the measurement error in

Given the large number of calculated observations, we may be worried this might impact the results. In particular, Borjas (1980) discusses the spurious negative relationship that can occur when calculating the wage by dividing earnings by hours worked, and then estimating using hours as the dependent variable. This division bias poses less of a concern in the context of this paper as wage is used as the dependent variable. However, the measurement error may be larger for calculated observations. Although we are able to distinguish true reported hourly wage rates from imputed wage rates, the imputations are predominately made for individuals with higher incomes. This means, estimating the results for only the reported sample would eliminate much of the variation in tax rates that was concentrated in the higher income brackets. Thus, we are reluctant to exclude imputed observations, and instead we proceed with the caveat that a large portion of our observations are calculated using labor income and hours.

We use the NBER TAXSIM database to calculate marginal federal and state tax rates.[13] We group individuals within a family into tax filing units, consisting of the respondent, his or her spouse, and their dependents. Because the 1986 SIPP panel includes most of the information needed to calculate the tax filing unit's itemized deductions, except for short- and long-term capital gains, we do not assume that tax filing units take the standard deduction.[14]

In some specifications, we exclude workers who change occupations and/or jobs

the SIPP earnings data is smaller than that from the SSA is promising. Both types of data are found to have serial correlation in their measurement error, and specifically for the SIPP data the correlation worsens the attenuation bias resulting from measurement error. We take first-differences of the data, which helps to alleviate some concern raised by the serial correlation. Also, given that we expect a negative coefficient, the attenuation bias is against finding a result.

[13]Because we are interested in both federal and state marginal rates, our sample excludes individuals with missing information about their state of residence and those living in the unidentifiable states in the 1986 SIPP: Alaska, Idaho, Iowa, Maine, Montana, North Dakota, South Dakota, Vermont, and Wyoming.

[14]The TAXSIM procedure assumes that the taxpayers opt for the standard deduction only when itemized deductions are not worthwhile. The 1986 SIPP panel collected information on tax filing in a topical module near the April 15 filing deadline in each year, but this data is no longer publicly available, so we cannot be certain that taxpayers are paying their minimum possible tax, nor do we know their actual filing status (e.g., whether married individuals filed one joint return or separate individual returns).

sometime during 1987. We define "occupation" as the type of job an individual holds and "job" as the given position an individual holds for a particular employer. SIPP asks respondents for their occupation at each interview, and attempts to identify one's employer across waves. It is possible that the same occupation or employer could be coded differently across waves, inducing false transitions and eliminating valid observations from our sample.[15]

3.4.1 Impact of TRA86

The individual changes between 1986 and 1987 in the marginal tax rate on personal income are summarized in Figure 3.2. Roughly 29.3% of the population experienced an increase in combined federal and state tax rate while 55.9% experienced a decrease, and the remaining 14.8% experienced no change. The 1987 income, which is used to calculate the 1987 tax rate, includes changes made to wage rates that might have occurred for reasons other than the tax change. Thus, we also compare a counterfactual tax rate, derived using 1986 income with the 1987 rules, to the actual 1986 rate. Figure 3.3 shows that roughly 14.9% had an increase, 59.9% had a decrease, and 25.2% had no change. The counterfactual comparison is relatively close to the actual change in terms of whether the change is positive, negative, or zero; however, the counterfactual changes are more concentrated between a decrease in tax rate of 0.10 and an increase in tax rate of 0.10.

For those taxpayers whose marginal rate decreased, TRA86 resulted in a substantial reduction in the federal tax burden, more than $1,700 (Table 3.1). The increase in federal taxes for those whose actual marginal rate increased is due largely to tax-

[15] Another concern with panel data like the SIPP is seam bias (Czajka (1983), Ham et al. (2009)), where transitions that occur during non-interview months are reported during the month of interview. Seam bias is more of a concern in event studies and likely won't affect our results, but it could affect our sample, if people who report a job or occupation transition in their first interview in 1987 actually changed jobs during 1986. As with the false transitions due to coding inconsistencies, this will only make our sample smaller than it should be.

payers who earned more in 1987 than in 1986, and thus move up one or more tax brackets; for those whose rate increased, their tax burden would have increased by an average of only $262 using their 1986 income and the 1987 tax rules.

Table 3.1 also suggests a positive correlation between wage changes and tax rate changes, which matches the prediction of our theoretical model. Taxpayers whose marginal rate decreased saw an average 93 cent decrease in their real pre-tax (hourly or imputed) wage, while wage rates rose by more than a dollar per hour for those whose rates increased. While this correlation could be driven by the tax rate's own dependence on wages, the patterns are similar (though smaller in magnitude) when looking at the change in marginal tax rate holding income constant at the 1986 level. Figure 3.4, which plots the change in net-of-tax rate (so the direction of correlation is reversed) against the change in wage, indicates that this correlation is consistent throughout the wage distribution.

Table 3.2 provides summary statistics on the sample grouped by 1986 tax rates. As expected, there is substantial variation of demographic characteristics for each grouping of brackets. For example, the higher brackets include a larger proportion of college graduates compared to the lower brackets, while women are more likely to fall into a lower tax bracket. By breaking down the data into subsets of the population, we can observe how differences in marginal product and in the elasticity of labor supply could potentially impact changes in wage rates. Income differences are used to proxy for differences in marginal product, while gender and marital status are used to proxy for differences in labor supply elasticity. Thus, we focus on three different ways to partition the data: income, gender, and marital status. By slicing the data for the regression analysis, we hope to gain a better sense for the source of variation in the wage response to tax changes.

3.5 Empirical Model and Results

The model presented in Section 3.3 predicts that each worker type's wage rate could be affected when changes are made to a tax on a specific type of labor. We now measure the extent to which the theory holds true in the data. A challenge arises when empirically estimating the impact of taxes on wage rates, which is that wage rates also influence the tax rate that an individual faces. To mitigate this simultaneity problem, we instrument for the actual tax change using a counterfactual tax rate, which is constructed by applying the 1987 tax rules to income earned in 1986.

We focus on short-run changes that occur between 1986 and 1987. This time span may be rather short for general equilibrium effects to fully materialize. Thus, we expect any effect taxes have on wage rates to be relatively small in magnitude. However, the Tax Reform Act of 1986 was a highly publicized event that had been discussed at length before going into effect in 1987. Therefore, the existence of an effect could be expected, especially in light of the estimates found in Kubik (2004).

3.5.1 Empirical Model

To measure the relationship between wage changes and tax changes, we start with a model describing the relationship between an individual i's pre-tax wage and the net-of-tax rate:

$$\ln w_{it} = \alpha_{it} + \beta \ln(1 - \tau_{it}) + \gamma X_i + \varepsilon_{it}, \tag{3.4}$$

where w_{it} is the real wage of person i at time t, τ_{it} is the marginal tax for person i at time t, X_i is a set of time-invariant demographic characteristics, and ε_{it} is the sum of an unobservable fixed effect, η_i, and an error term with standard assumptions, μ_{it}. The unobserved fixed effect will likely bias cross-sectional OLS estimates. We can eliminate this individual fixed effect by estimating a first-differences model,

$$\Delta \ln w_i = \alpha_i + \beta \Delta \ln(1 - \tau_i) + \Delta \varepsilon_i. \tag{3.5}$$

Estimating the transformed equation could also lead to biased results due to the endogeneity of marginal tax changes to wage changes. For instance, if real wage rates increase for a reason unrelated to taxes, this can push an individual into a higher bracket. The increase in tax rate would decrease her net-of-tax rate and mechanically force a negative estimate of β. Given that the theoretical model presented in Section 3.3 predicts β to be negative (or the coefficient on τ to be positive), this would bias β in the expected direction.[16] To minimize the impact of this bias, we use an instrumental variables approach, where the change in a counterfactual tax rate serves as an instrument for the change in the actual tax rate. This counterfactual tax rate is the marginal (federal plus state) rate that a taxpayer faces in 1987 using that taxpayer's income from 1986. Following Auten and Carroll (1999), we estimate the first stage regression:

$$\Delta \ln(1 - \tau_i^a) = a_i + \gamma \Delta \ln(1 - \tau_i^{cf}) + \xi_i, \tag{3.6}$$

where τ_i^{cf} corresponds to the counterfactual tax rate and τ_i^a refers to the actual tax rate for person i. We then use the predicted change in tax rate, $\Delta \ln(1 - \tilde{\tau}_i)$, as the independent variable in Equation (3.4). This method, by construction, removes the change in tax rate that would have occurred due to changes in taxable income in order to isolate the impact of the change in tax policy.[17]

[16] Blomquist and Selin (2009) give reasons why the coefficient on the net-of-tax rate could be positive.

[17] Using a counterfactual tax rate as an instrument in the context of estimating the elasticity of taxable income has received some criticism due its potential to violate the exogeneity assumption (Moffitt and Wilhelm (2002) and Blomquist and Selin (2009)). If the changes in tax rate are monotonic in income (e.g. higher incomes get bigger tax rate changes), then the exogeneity assumption could be violated (Weber (2010)). In particular, if tax changes are highest for higher incomes, then, because the base year income is used to predict a counterfactual tax rate, any shocks to transitory income are larger for higher incomes and therefore correlated with the larger tax rate change. However, our results focus on labor market earnings, which are likely less subject to transitory shocks (other than unemployment spells, and we exclude those with more than three months not working), and we use reported hourly wage rates where available.

We can now write the percent change in wage as a function of the predicted tax change,

$$\Delta \ln w_i = \alpha + \beta \Delta \ln(1 - \tilde{\tau}_i) + \varepsilon_i. \tag{3.7}$$

The theoretical model suggests that β should be negative – a decrease in marginal tax rate, as experienced by more than half of the sample, which is an increase in the marginal net-of-tax rate, should be associated with a decrease in the real pre-tax wage.

Because the location on the income distribution is driving the counterfactual tax rate, we may be concerned that this location may also predict wage changes. This would result in a coincidental systematic relationship between tax changes and wage changes, when in truth, the correlation was with income. Figure 3.5 shows the difference between the counterfactual and actual tax rates by 1986 income.[18] This figure highlights that the tax changes generated by the new tax policy varied across the income distribution with no discernable pattern based on income. As long as income in previous years does not share a similar pattern with wage changes, then, it is unlikely any measured relationship between changes in wage rates and tax changes would have occurred absent the tax change. Figure 3.6 shows the log wage changes in the previous year (1985-1986) by 1985 income.[19] The figure shows that location on the income distribution does not appear to predict wage changes in any meaningful way. Thus, measuring a coincidental impact of tax rates on wage rates is less of a concern.

[18] Income is binned in $100 increments and the mean difference in counterfactual and actual tax rate is graphed.
[19] Income is binned in $100 increments and the mean log wage change is graphed.

3.5.2 Estimation Results

The IV results from estimating Equation (3.7), where τ is the sum of federal and state marginal tax rates, are summarized in Table 3.3. Column (1) displays results using the full sample without controlling for individual or firm characteristics. We find that a 1% increase in the net-of-tax rate $(1 - \tau)$ leads to a 0.51% decrease in real pre-tax wag rates; that is, a decrease in an individual's marginal tax rate is associated with decreases in her real wage, consistent with the theoretical prediction from Section 3.3. For these results, we report the 95 percent confidence interval around our point estimate; we can reject the null hypotheses that this coefficient is equal to zero or to negative one, suggesting that the burden of this tax change is borne at least in part by both workers and firms.[20]

Column (2) displays results when individual and firm level controls are included. The coefficient of interest, that on the change in log net-of-tax rate, decreases slightly compared to column (1), from 0.51 to 0.43, but zero and negative one remain outside the 95 percent confidence interval. Controlling for individual characteristics does not appear to effect our qualitative result.[21]

We conduct the same analysis for the sample excluding job and occupation changers in columns (3) and (4). We focus on individuals who stay at the same job and occupation because this group is least likely to see wage changes other than the general equilibrium effect of the tax change, allowing us to potentially avoid attributing wage changes caused by non-tax factors to taxes (e.g., changes in union coverage or different benefit/wage tradeoffs at the new job). The estimate of the impact of taxes on wage rates for this group should be more conservative, as we neglect to

[20] Table 3.3 also includes the first stage results, where the F-statistic indicates that the weak instrument problem is not a concern.

[21] Most of the coefficients on the individual and firm controls are not statistically significant in the second stage, except marital status and the non-white indicator, which are significant at the 10 percent level. Marital status and its interaction with gender and age are the only statistically significant variables in the first stage.

account for taxpayers who switch jobs to take advantage of the new tax regime. This is confirmed by our results where, after excluding job and occupation changers, the coefficient on the net-of-tax rate (federal+state) change is, as expected, smaller in magnitude. However, the coefficient is still negative and statistically significant. In particular, we find that, after controlling for demographic characteristics and excluding job and occupation changers, a 1% increase in net-of-tax rate led to a 0.26% decrease in real wage rates.[22] Table 3.4 shows results when the same regressions are run using only the federal tax rate (instead of federal and state tax). The results are similar to Table 3.7, but the magnitudes are larger. Table 3.4 shows, after including demographic controls and excluding job and occupation changers, that a 1% increase in the net-of-tax rate is associated with a statistically significant 0.31% decrease in wage.[23]

To assess the magnitude of these results, suppose an individual (who did not change jobs between 1986 and 1987) makes $10/hr in 1986 and faces a marginal federal tax rate of 20 percent, making her net-of-tax rate 0.80. Now, suppose that her net-of-tax rate increases to 0.90, that is, her net-of-tax rate increases by 12.5%. The results in column (4) suggest that her 50% (or 10 percentage point) decrease in federal tax rate will be associated with a 3.875% pre-tax real wage decrease, to $9.61/hr. While she will keep more of her income post-tax, she will not reap the full benefit of the tax reduction, as would have been the case if her pre-tax wage remained the same.

[22]We also estimated the model for just the sample of job or occupation changers. The point estimates are higher, approximately -0.8 both with and without controls and for both federal taxes only and federal plus state taxes. The coefficients are significantly different from zero, but with the smaller sample size and the larger estimated magnitude, we can no longer reject the null hypothesis that the coefficient is equal to negative one.

[23]The OLS results for Equation 3.5 are not presented in the paper. The OLS results, as expected, overestimate the impact of the net-of-tax changes on changes in wage rates. The OLS estimate on the net-of-tax change when considering both the federal and state tax rates is -0.63 while the IV coefficient for the comparable sample is -0.26. When considering only the change in federal tax rate, the OLS coefficient on the net-of-tax change is -0.71, while the IV coefficient for the comparable sample is -0.31.

Table 3.5 shows the results when dividing the sample into those above and below the median annual income in 1986 ($17,712). The results show that the negative relationship between the net-of-tax rate and wage changes is concentrated within the higher income group, particularly when excluding job and occupation changers. In fact, the point estimate on the net-of-tax rate for low income individuals is positive, though not significantly different from zero, after excluding job and occupation changers. This reversal of sign may seem surprising, but is actually consistent with the predictions of the model presented above. In particular, if income is a proxy for skill, then we can observe two types of labor: low skill and high skill. The model expects that the change in wage for high skilled workers has the opposite sign as the wage changes for the low skilled workers. Given that the tax changes were most pronounced at the highest income levels, this is akin to a relative tax change for only the high skilled workers. Thus, according to the model, the decrease in tax rate for most high income individuals is associated with a wage decrease, but for low income workers, this would result in a wage increase under reasonable restrictions for the elasticity of substitution between low and high skilled workers. As such, our results are in line with the model's predictions for high income individuals. For low income individuals, the point estimate on the coefficient is consistent with the theoretical predictions, but we cannot reject the null of no change in the pre-tax wage rate.

An interesting exercise is to observe how the results vary by gender, given that women are believed to supply labor more elastically than men. When breaking down the data by sex (Table 3.6), the coefficient on tax changes is larger in magnitude for men than women. These results seem surprising given that if women have a larger elasticity of supply, we might expect that women's wages should be more responsive than their male counterparts. Further decomposition of the data by marital status reveals little difference between single men and married men, with the results remaining negative and statistically different from zero. However, the differences among married and single women is stark. Although the coefficient on the net-of-tax rate for married women remains negative, it is no longer statistically significant. Moreover, the magnitude of the coefficient is roughly 10 times smaller for married women compared to their single counterparts. On the other hand, wage rates for single women share a similar coefficient on the net-of-tax rate to that of men. That married women's wage rates appear to be unaffected by tax rates may stem from the fact that if a married woman is the secondary earner, then when filing jointly her marginal tax rate may not necessarily correspond to her wage rate. Thus, the model's predictions may be less suited particularly for married women as it fails to account for a joint system of taxation. An alternative explanation is that the elasticity of demand in production for married women is large, with the limiting case of perfect substitutes illustrated in Section 3.3. If labor demand for married woman is relatively more elastic than their compensated labor supply elasticities, then the estimated coefficient actually should move closer to zero, where zero is the limiting case when married women are perfect substitutes for one another.

3.6 Summary and Conclusion

Using panel data from the SIPP, this paper looks at whether pre-tax wage rates were impacted by changes made to the net-of-tax rate before and after the Tax Reform Act of 1986. This paper provides strong evidence that pre-tax wage rates are in fact altered in response to change in tax policy. We find that a 1% increase in the net-of-tax rate is associated with a 0.26% decrease in real wage for individuals who changed neither job nor occupation. This estimate is derived inclusive of the state tax rate, avoiding bias from their exclusion as state taxes were also changing during the reform period. When considering only the federal tax rate change, a 1% increase in the net-of-tax rate is associated with a 0.31% decrease in real wage.

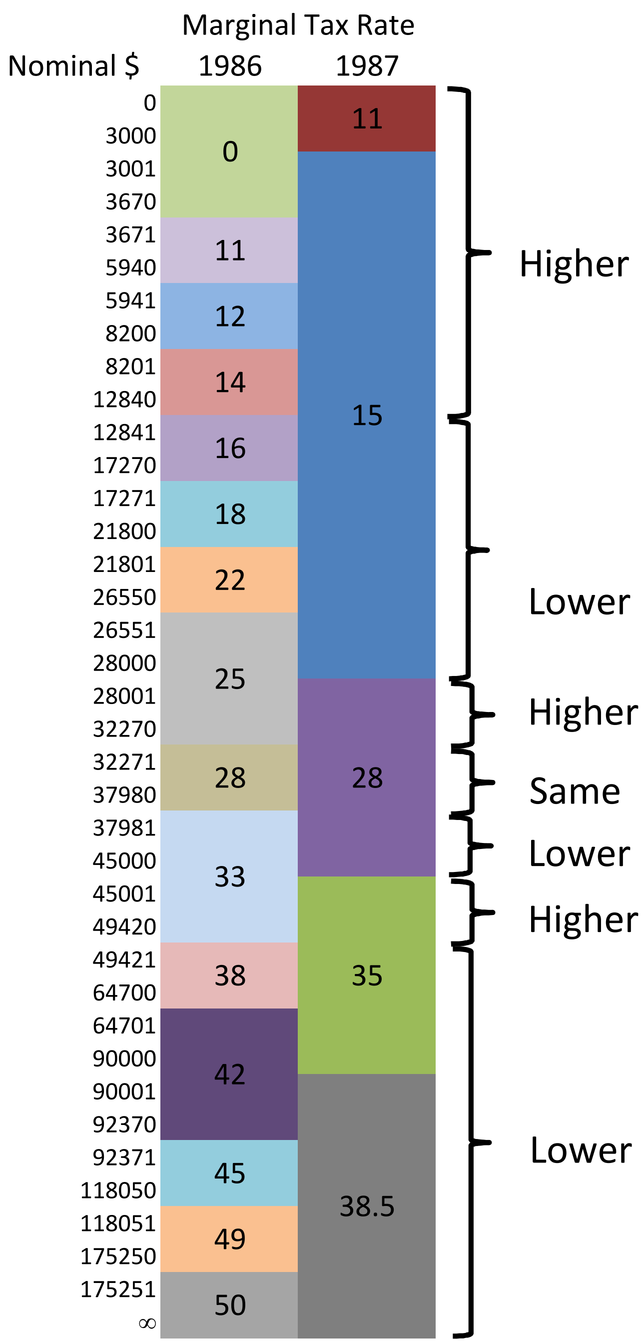
This is not the first study to attempt to document such a relationship between pre-tax wage rates and the marginal tax rate. However, this study is the first to use panel data with a rich set of controls in conjunction with TAXSIM to study this particular question for the US. As such, our estimates are more conservative than previous studies as we are able to exclude variation in wage rates that is not attributable to the tax change. Also, our data allow for a more detailed exploration of results since we are able to observe individuals and not just points in a changing distribution.

By further analyzing the data by subgroups, we are able to explore where in the income and demographic distributions the major wage changes are taking place. In particular, the negative relationship between changes in wage rates and tax changes exists in the upper half of the income distribution after excluding job and occupation changers, while the coefficient on the net-of-tax rate for the lower half of the income distribution is positive but not significantly different from zero. This result is consistent with a one-sector model with two types of labor where a tax is levied on only one type. Breaking the data down by gender and marital status reveals that married women have the smallest association between wage changes and tax changes that is not significantly different than zero. Thus pre-tax wage rates for married women appear invariant to policy changes, consistent with the standard assumption. That the results do not hold for married women is less surprising given that the model does not account for a joint system of taxation. Additionally, the lack of response in married women's pre-tax wage rates could also result when there exists a large elasticity of demand for their labor. However, differentiating between the two hypotheses is beyond the scope of this paper, but raises an interesting question for future research.

The estimates derived in this paper provide evidence contrary to the commonly used assumption in previous studies that wage rates do not change when the marginal tax rate on income changes. Our results suggest that assuming income taxes are fully borne by labor is not supported by the data, with the possible exception of married women. In addition, we provide evidence supporting this claim with a more conservative estimate that excludes individuals who changed jobs or occupations during the tax reform. The results from this paper have implications for past empirical studies that use changes in tax policy to estimate behavioral parameters. Given that changes in wage rates work to counter the tax change, past work using variation in marginal tax rates to obtain estimates of labor supply elasticities may be understated. Also, past work using variation in marginal tax rates to measure the elasticity of taxable income could also be understated as general equilibrium changes in wage rates might dampen the predicted changes in taxable income. Overall, this paper contributes, with additional evidence, to a growing literature that has shown pre-tax wage rates are in fact affected by changes in marginal tax rates, and thus highlights that caution should be used when measuring behavioral parameters through variation from changes in tax policy.

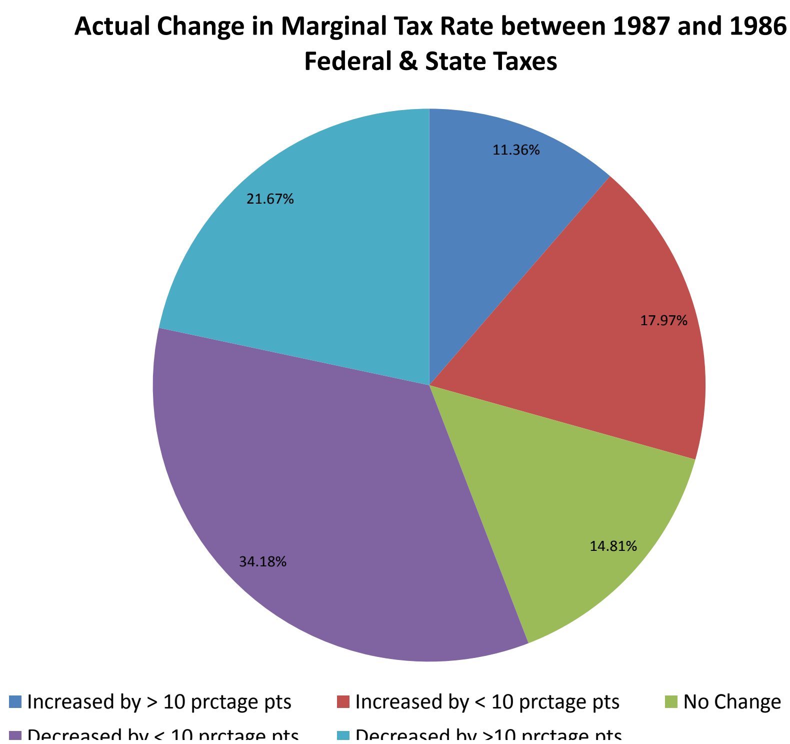
3.7 Figures and Tables

Figure 3.1: Tax Reform of 1986: Bracket Changes



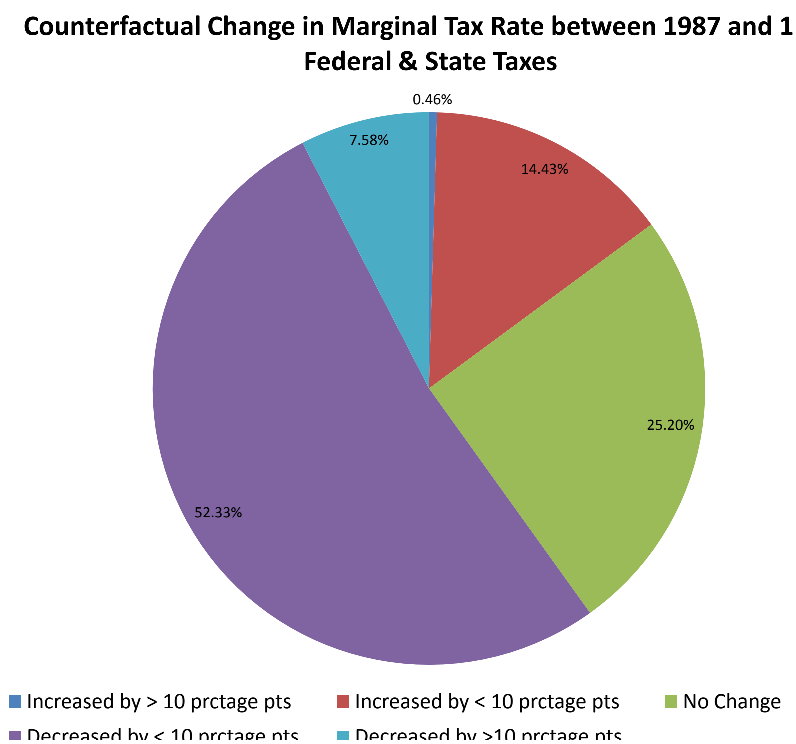
Source.–www.taxfoundation.org

Figure 3.2: Distribution of Tax Changes between 1987 and 1986



Notes.– Federal and state marginal tax rates are calculated using NBER's TAXSIM.

Figure 3.3: Distribution of Change between Counterfactual rate and 1986 rate



Notes.– Federal and state marginal tax rates are calculated using NBER's TAXSIM. Counterfactual tax rate is calculated using 1986 income with 1987 tax rules.

Table 3.1: Impact of TRA86 on Overall Tax Burden

	Change in Actual Marginal Tax Rate, 1987-1986		
	Lower Rate	No Change	Higher Rate
Mean Change in Fed Tax Burden ($)	-1721	-10	2042
	(3706)	(310)	(3579)
Mean Change in State Tax Burden ($)	-230	10	443
	(741)	(69)	(803)
Mean Change in Wage ($)	-0.93	-0.06	1.14
	(3.11)	(1.58)	(2.67)
Mean Percent Change in Wage	-5	24	56
	(98)	(253)	(245)
Mean Percent Change in Log Wage	-13	3	22
	(56)	(85)	(55)
% of Total	60.1	6.5	33.4
Number of Observations	7656	835	4258

Notes.– All variables other than tax rates come from 1986 SIPP. Federal and state marginal tax rates are calculated using NBER's TAXSIM.

Figure 3.4: Relationship Between Log Real Wage Changes and Log Net-of-tax Rate Changes

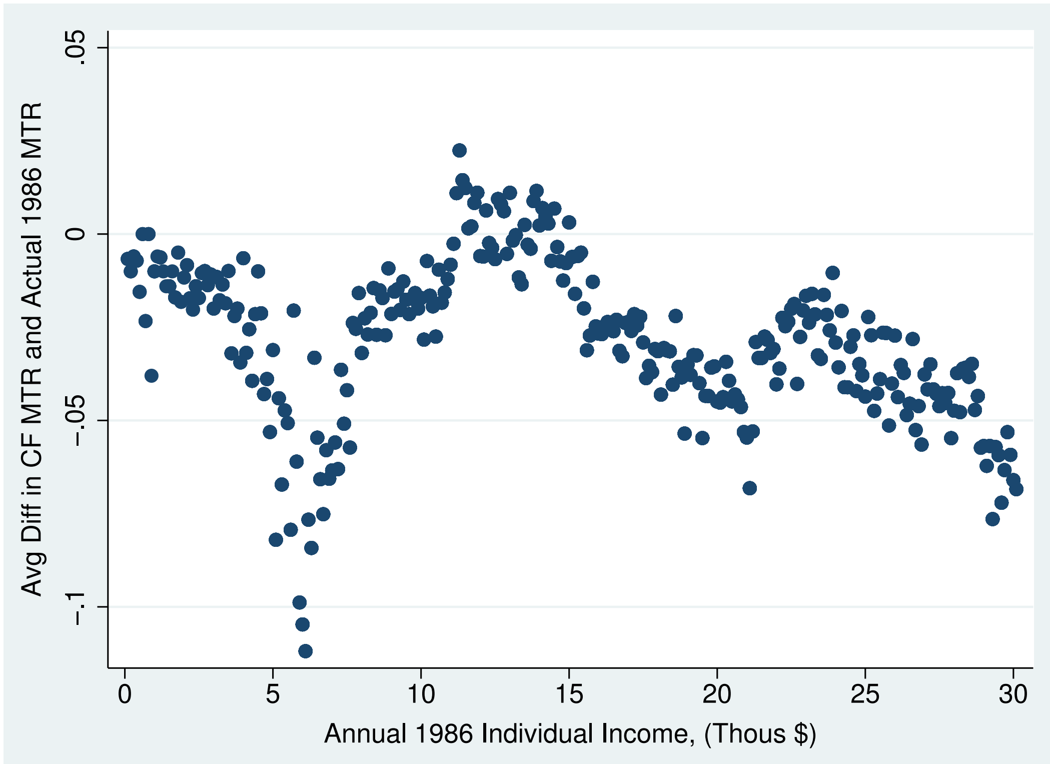
Notes.– Wage data come from 1986 SIPP. Federal and state marginal tax rates are calculated using NBER's TAXSIM. Each point represents the mean log change in net-of-tax rate for bins of 1 percentage point for the log change in real wage.

Table 3.2: Summary Statistics, by 1986 Federal Tax Bracket

	1986 Bracket					
	0	11-18	22-28	33-38	42-50	All
Mean 1986 Income ($)	3204	11798	21132	36800	61018	14897
	(3529)	(5949)	(10192)	(17423)	(32240)	(14746)
Mean 1986 Fed Tax ($)	-47	838	3199	9408	23000	2378
	(123)	(635)	(1687)	(3541)	(8209)	(4166)
Mean 1986 State Tax ($)	-14	176	647	1650	2881	430
	(93)	(240)	(650)	(1480)	(2665)	(876)
Mean Real 1986 Wage ($)	1.46	5.41	9.03	12.74	13.94	6.05
	(2.04)	(3.00)	(4.23)	(4.88)	(5.46)	(5.03)
Age	37	37	38	42	46	38
	(14)	(13)	(12)	(11)	(10)	(13)
Married (0/1)	0.54	0.68	0.62	0.71	0.76	0.62
	(0.50)	(0.46)	(0.48)	(0.45)	(0.43)	(0.49)
Female (0/1)	0.68	0.53	0.38	0.34	0.33	0.51
	(0.47)	(0.50)	(0.48)	(0.47)	(0.47)	(0.50)
Nonwhite (0/1)	0.26	0.21	0.13	0.09	0.05	0.20
	(0.44)	(0.41)	(0.34)	(0.29)	(0.23)	(0.40)
Less than HS (0/1)	0.21	0.14	0.08	0.04	0.02	0.13
	(0.40)	(0.35)	(0.27)	(0.19)	(0.15)	(0.34)
Some College (0/1)	0.22	0.23	0.25	0.22	0.21	0.23
	(0.41)	(0.42)	(0.43)	(0.42)	(0.41)	(0.42)
College (0/1)	0.11	0.14	0.26	0.49	0.59	0.20
	(0.31)	(0.35)	(0.44)	(0.50)	(0.49)	(0.40)
Fed Tax Diff, 1987-1986 ($)	569	588	-552	-3483	-11366	-363
	(2034)	(2438)	(2860)	(5693)	(10549)	(3767)
State Tax Diff, 1987-1986 ($)	130	148	-43	-499	-1068	3
	(469)	(558)	(689)	(1250)	(2122)	(764)
Wage Change, 1987-1986 ($)	0.35	-0.15	-0.42	-0.98	-1.28	-0.19
	(2.16)	(2.77)	(3.19)	(3.89)	(4.17)	(2.94)
Wage Change, 1987-1986 (%)	74	7	3	1	-3	16
	(365)	(108)	(63)	(115)	(69)	(174)
Log Wage Change, 1987-1986	25	-3	-6	-10	-11	0
	(93)	(57)	(46)	(46)	(47)	(60)

Notes.– All variables other than tax rates come from 1986 SIPP. Federal and state marginal tax rates are calculated using NBER's TAXSIM.

Figure 3.5: Difference in Actual 1986 Tax Rate and Counterfactual Tax Rate, by 1986 Income



Notes.– Wage data come from 1986 SIPP. Federal and state marginal tax rates are calculated using NBER's TAXSIM. Counterfactual tax rate is calculated using 1986 income with 1987 tax rules. Income is binned in $100 increments.

Figure 3.6: Log Difference in Wage from 1985 to 1986, by 1985 Income

Notes.– Data come SIPP. Income is binned in $100 increments.

Table 3.3: IV Regression of Change in Pre-Tax Wage on Change in (Federal + State) Tax Rate

	$\Delta \ln w$, 1987-1986			
	Full Sample (1)	Full Sample (2)	Excl Job & Occ Chgs (3)	Excl Job & Occ Chgs (4)
$\Delta \ln(1-\tilde{\tau})$	-0.509	-0.428	-0.345	-0.264
(S.E.)	(0.089)	(0.097)	(0.098)	(0.106)
[95%CI]	$[-0.683, -0.335]$	$[-0.618, -0.238]$	$[-0.537, -0.154]$	$[-0.473, -0.055]$
R^2	0.074	0.076	0.051	0.051
N	9475	9475	6138	6138
Include Controls	No	Yes	No	Yes
First Stage	$\Delta \ln \tau^a$, 1987-1986			
$\Delta \ln \tau^{cf}$	0.628	0.589	0.670	0.634
(S.E.)	(0.021)	(0.021)	(0.025)	(0.026)
R^2	0.086	0.097	0.101	0.110
F-stat	888.83	50.61	692.97	39.63

Notes.– Marginal tax rate includes both federal and state tax rates. Control variables include gender interacted with 1986 marital status, age, age squared, non-white indicator, and categorical variables indicating education level and industry. Individuals moving in and out of non-employment are also excluding when excluding job and occupation changers.

Table 3.4: IV Regression of Change in Pre-Tax Wage on Change in Federal Tax Rate

	$\Delta \ln w$, 1987-1986			
	Full Sample (1)	Full Sample (2)	Excl Job & Occ Chgs (3)	Excl Job & Occ Chgs (4)
$\Delta \ln(1-\tilde{\tau})$	-0.543	-0.446	-0.406	-0.311
(S.E.)	(0.107)	(0.117)	(0.118)	(0.129)
[95%CI]	$[-0.753, -0.333]$	$[-0.675, -0.217]$	$[-0.637, -0.175]$	$[-0.564, -0.059]$
R^2	0.068	0.070	0.051	0.050
N	9475	9475	6138	6138
Include Controls	No	Yes	No	Yes
First Stage	$\Delta \ln \tau^a$, 1987-1986			
$\Delta \ln \tau^{cf}$	0.570	0.533	0.606	0.571
(S.E.)	(0.02)	(0.02)	(0.024)	(0.025)
R^2	0.079	0.090	0.092	0.101
F-stat	807.33	46.97	621.36	36.13

Notes.– Marginal tax rate includes only federal tax rate. Control variables include gender interacted with 1986 marital status, age, age squared, non-white indicator, and categorical variables indicating education level and industry. Individuals moving in and out of non-employment are also excluding when excluding job and occupation changers.

Table 3.5: IV Results by Income

	$\Delta \ln w$, 1987-1986			
	Low Income		High Income	
	Full Sample (1)	Excl Job & Occ Chgs (2)	Full & Sample (3)	Excl Job & Occ Chgs (4)
$\Delta \ln(1 - \tilde{\tau})$	-0.162	0.078	-0.242	-0.282
(S.E.)	(0.267)	(0.343)	(0.134)	(0.132)
[95%CI]	$[-0.499, 0.175]$	$[-0.314, 0.471]$	$[-0.504, 0.021]$	$[-0.541, -0.022]$
R^2	0.023	.	0.057	0.066
N	4658	2563	4817	3575
First Stage F-stat	307.86	257.73	293.16	234.56

Notes.– Marginal tax rate includes only federal tax rate. Regression does not include controls for individual characteristics. Individuals moving in and out of non-employment are also excluding when excluding job and occupation changers.

Table 3.6: IV Results by Gender and Marital Status

	$\Delta \ln(1-\tilde{\tau})$	(S.E.)	95% CI	N	R^2	1st Stage F
Female						
All	-0.249	(0.117)	[-0.478,-0.019]	4084	0.025	396.49
Single	-0.769	(0.337)	[-1.429,-0.109]	1587	0.117	73.54
Married	-0.077	(0.109)	[-0.292,0.137]	2497	0.006	388.69
Male						
All	-0.742	(0.137)	[-1.012,-0.473]	5391	0.134	462.13
Single	-0.645	(0.412)	[-1.452,0.162]	1604	0.152	44.66
Married	-0.686	(0.147)	[-0.974,-0.397]	3787	0.106	459.95

Notes.– Marginal tax rate includes both federal and state tax rates. Sample includes job and occupation changers. Regression does not include controls for individual characteristics.

CHAPTER IV

Beyond Taxes: Understanding the Labor Wedge

4.1 Introduction

A recent literature has focused on examining the determinants of the long-run behavior of aggregate labor hours per working-age population (H/P) within and across countries.[1] This literature's analytical framework is based on a standard macroeconomic neoclassical model. The theory behind this model implies that equilibrium H/P is implicitly defined through a static optimality condition that equates the marginal rate of substitution of consumption for leisure with the marginal product of labor. The extent to which this condition fails to hold has been coined the labor wedge. More concisely, the labor wedge is simply a residual that captures the percent difference between model-predicted H/P and its empirical counterpart. The labor wedge has been found to be substantial across a large sample of OECD countries. Recent studies focus on why the labor wedge exists and what factors can account for it.[2] These studies argue that across countries, a considerable fraction of the labor wedge can be explained by taxes. Hence, when the standard neoclassical model is enhanced to incorporate taxes, the model's predictions regarding the long-run behavior of H/P improve considerably. However, this improvement is for all purposes limited

[1] See, for example, Prescott (2004) and Ohanian et al. (2008).
[2] See, for example, Shimer (2009) in addition to the earlier cited literature.

to European countries. In particular, the model's predictions are contrary to US data, and also fail considerably for Canada. The aim of this paper is to understand what factors, in addition to taxes, can account for the labor wedge.

In this paper, we argue that the failures of the standard model to account for the long-run behavior of H/P in Canada and the US are not exceptions, but rather the norm. The standard model yields an equation that implicitly defines total equilibrium labor hours. When testing the model, it is standard to normalize all variables by the working-age population.[3] Hence, the model is assumed to yield predictions in terms of H/P. Of course, H/P is equal to the product of hours per worker (H/E) and the fraction of employed individuals (E/P). We present evidence that the standard model has no long-run explanatory power regarding E/P. However, once taxes are accounted for, the model is capable of accurately predicting H/E. Therefore, whenever E/P does not change much relative to H/E, which has been the case for most European countries, the empirical behavior of H/E and H/P are indistinguishable and the model gives the impression of correctly predicting H/P. On the other hand, when E/P *does* change considerably relative to H/E, as has been the case in Canada and the US, the standard model implicitly reveals its limitations regarding E/P and fails.

Since the standard model lacks predictive power regarding changes in the extensive margin of labor supply, as we will show, the E/P ratio is absorbed by the residual. Given that the labor wedge is implicitly defined as a residual, or the portion of the data that predicted hours can not explain, the E/P ratio will automatically comprise part of the labor wedge. This finding represents an important contribution in terms of the interpretation of the labor wedge. The research in Prescott (2004)

[3]From now on, for simplicity we refer to population and working-age population interchangeably.

and Ohanian et al. (2008) successfully identified that part of the labor wedge can be attributed to ignoring taxes. We complement this research by showing that another part of this wedge actually holds by construction. In other words, a portion of the wedge exists because of the model's inability to predict extensive margin changes in labor supply.

The limitations of the standard model with regards to predicting E/P imply that the model may be incapable of fully capturing the impact of taxes on H/P. Hence, we develop a model that allows for heterogeneity in terms of employment status. In our framework, a household planner maximizes the joint utility of all household members by optimally choosing the fraction of the population that is employed, the hours that each employed individual works, and the distribution of household consumption across individuals conditional on employment status. We incorporate a time-varying fixed cost associated with employment that provides a natural motivation for the existence of voluntary non-employment.[4] Our model rationalizes why the standard model is better suited towards predicting hours per worker than hours per population. In particular, when the choice of employment is not available, the theory is implicitly denied the ability to endogenously make an assertion regarding the scaling of work hours.

In addition, we focus on the model's relevance as a tool for assessing policy. We analyze how changes to the net-of-tax rate impact hours per worker and the number of workers. We find that both respond in the same direction and therefore have an unambiguous impact on hours per population, which is that hours per population are increasing in the net-of-tax rate. We also assess the differential impact that changes in average tax rates versus marginal taxes can have on hours per population. We

[4]That is, in the neoclassical spirit of market clearing, we do not focus on involuntary aspects of unemployment.

find that a decrease in the average tax rate will lead to a decrease in hours per worker accompanied by an increase in employment per population. Thus, a decrease in the average tax rate could potentially have an ambiguous impact on hours per population.

The rest of this paper is organized as follows. Section 4.2 reviews related literature. In order to build intuition, in Section 4.3 we review the standard neoclassical model's explanatory power in terms of hours per population both with and without taxes. In Section 4.4 we present evidence that the standard model lacks predictive power in regards to changes in E/P. In turn, Section 4.5 develops a model in which a stand-in household's optimal decisions include both the extensive and intensive margins of labor supply. This section examines the model's implications regarding tax policy. Then, Section 4.6 addresses the model's ability to match data on H/E and E/P. Finally, Section 4.7 concludes.

4.2 Related Literature

Conceptually, the labor wedge discussed in this paper stems from the analysis in Prescott (2004). Prescott seeks to explain differences in hours per population between the US and a set of European countries. Using aggregate data, Prescott derives effective tax rates for a group of OECD countries between 1970 and 1974, and also between 1993 and 1996. He then uses the first-order conditions derived from a standard neoclassical model with taxes to generate data on hours per population. Prescott argues that over his reference periods, cross-country differences in effective tax rates can account for a considerable fraction of the level differences in hours per population. However, he notes that the model predicts that in the US, H/P should have gone down, when in reality this ratio has gone up. He suggests that this failure of the model owes to the fact that the US experienced an increase in married women's labor force participation along with a flattening of the tax schedule during the 1980's. Thus, the marginal tax rate for large changes in income when moving to a two-earner household was significantly higher in the earlier period compared to the later period even though the calculated marginal tax rate used for predicting hours remained the same.

Alesina et al. (2005) argue that the assumptions underpinning the model in Prescott (2004) actually drive the results. In particular, they argue that the choice of log utility function as well as an implied labor supply elasticity with respect to the tax rate roughly equal to 3 are what allow for taxes to explain most of the differences in labor hours across countries. Moreover, the authors suggest that an omitted variable in Prescott's analysis is cross-country differences in the degree of unionization. Alesina et al. posit that in the absence of changes in market regulations imposed by unions, changes in effective tax rates would not have affected hours worked to the extent implied by Prescott (2004).

More recently, Ohanian et al. (2008) extend the analysis in Prescott (2004) by studying a larger set of countries over a longer time frame and using a slightly different functional form for utility. The analysis uses annual effective tax rates derived in McDaniel (2007). This allows for a broad and detailed documentation of the long-run behavior of H/P, as well as extensive testing of the standard macroeconomic model's explanatory power. Ohanian et al. agree with Prescott in that augmenting the neoclassical model with taxes can broadly account for changes within country changes in H/P over time. This conclusion is robust to controlling for institutional differences. However, Ohanian et al. note that for Canada and the US, model-generated hours per population fail to match their empirical counterparts.

Shimer (2009) reviews literature on the labor wedge. The central points of Shimer's paper involve shifting attention to the cyclical behavior of the labor wedge. In particular, he backs out the US labor wedge and notes its strong procyclicality. He then considers plausible explanations for this procyclicality. As argued in Prescott (2004) and Ohanian et al. (2008), a fraction of the labor wedge can be accounted for by taxes. Thus, one explanation for the cyclical behavior of the labor wedge are cyclical fluctuations in taxes. However, Shimer argues that this explanation is unreasonable as it would be consistent with taxes being strongly countercyclical. This means that taxes would increase substantially during recessions. Instead, he suggests that a more plausible explanation for these cyclical fluctuations may involve noncompetitive aspects of the labor market.

This paper shifts attention to the puzzle created by the model's predictions for the US and Canada. We search for explanations of the discrepancy within the model itself by disaggregating the model's predictions and examining the extent to which it is capable of predicting behavior on multiple dimensions. To motivate our study on the labor wedge, we start by presenting additional background information on its existence and the impact of including taxes in the standard neoclassical model. Also through this discussion, we will highlight the failure of the model to account for labor supply behavior in the US and Canada.

4.3 The Labor Wedge

In a standard neoclassical macroeconomic model, setting labor supply equal to labor demand yields a straightforward equation for hours worked. Hours are derived to be a function of output, consumption, and the structural parameters of the model. It is therefore possible to generate a predicted series of hours using aggregate data on consumption and output, along with assumed parameter values. The extent to which the predicted hours differ from the actual data on hours is captured by the "labor wedge". The labor wedge is defined as the ratio of model predicted hours and the actual data on hours. If the model perfectly predicted hours, then this ratio would be one. Realistically, many of the assumptions of the standard neoclassical model are likely to not hold. However, the exercise of comparing the generated hours with the data is useful for measuring the accuracy of the model with its simplifying assumptions. In what follows, we derive the equation for labor hours that stems from the standard neoclassical model. For ease of exposition, we focus on Canada, France, Germany, and the US to measure the accuracy of the model's predictions. These countries are representative of the differences that earlier research has found between European and North-American countries in terms of hours per working-age population. Whereas in Europe, on average, H/P has decreased slightly over the last several decades, in North America H/P has slightly increased. We calculate the resulting labor wedge when the equation is applied to data for the US, Canada, France and Germany over the span of roughly 40 years. We then relax the assumption that tax distortions are non-existent in the standard model and re-derive the model assuming a broad set of taxes on labor, capital, investment, and consumption. The effective tax rate that had been previously excluded from the standard model factors into the prediction for hours worked.

4.3.1 The Standard Model

The standard neoclassical macroeconomic model assumes a representative household that maximizes its present discounted value of utility subject to an intertemporal budget constraint. The infinitely lived household derives utility from household consumption C and disutilty from household labor hours H. Thus, the household seeks to maximize

$$\mathbb{E}_t \sum_{t=0}^{\infty} \beta^t U(C_t, H_t),$$

subject to

$$W_t H_t + R_t K_t \geq C_t + I_t.$$

Above, β is the discount factor, W is the real wage, I is investment, and the price of consumption is normalized to 1. The household is assumed to own the economy's capital, K, and a representative firm rents the capital from the household at a rental rate of R. The capital accumulation equation is given by

$$K_{t+1} = I_t + (1-\delta)K_t,$$

where δ is the capital depreciation rate.

Following Shimer (2009), we assume that the household's instantaneous utility is given by

$$U(C_t, H_t) = \ln C_t - \gamma \frac{\varepsilon}{1+\varepsilon} H_t^{\frac{1+\varepsilon}{\varepsilon}}, \tag{4.1}$$

where ε is the Frisch (marginal value of real wealth held constant) elasticity of labor supply, and γ is a positive constant.[5]

Output Y is determined by a representative firm with Cobb-Douglas production function

$$\begin{aligned} Y_t &= Z_t F\left(K_t, H_t\right) \\ &= Z_t K_t^{\alpha} H_t^{1-\alpha}, \end{aligned}$$

where $\alpha \in (0,1)$ and Z is technology. The representative firm chooses capital and

[5]The assumed functional form for instantaneous utility differs from that used in Prescott (2004) and Ohanian et al. (2008). However, as shown in Shimer (2009), the choice of utility function has little impact on both the qualitative and quantitative results. Chetty (2009) offers a detailed comparison between the Frisch elasticity, which is commonly used in macroeconomic literature, and the compensated elasticity, which is frequently used in the public finance and labor literature.

work hours to maximize profits, given by

$$\Pi_t = Z_t K_t^{\alpha} H_t^{1-\alpha} - W_t H_t - R_t K_t.$$

Assuming all markets are perfectly competitive, in equilibrium, labor supply must equal labor demand. Therefore, combining the household's first-order conditions for labor supply and consumption with the firm's first-order condition for labor demand yields an equation for the equilibrium level of hours worked

$$H_t^{NM} := \left(\frac{(1-\alpha)Y_t}{\gamma C_t} \right)^{\frac{\varepsilon}{1+\varepsilon}}, \tag{4.2}$$

where NM stands for neoclassical model.[6] This is a static condition that, in theory, must hold within any time period.

There are diverse ways to test the validity of models such as structural estimation and numerical simulation. However, given that the condition in equation (4.2) is static, we can easily test its accuracy by using aggregate data on output and consumption to generate the model's prediction for hours worked, subject to choices for the model's parameters. This approach has been used in past literature including Parkin (1988), Rotemberg and Woodford (1999), and Mulligan (2002). The extent to which the model's predicted hours H^{NM} differ from actual hours H^{ACTUAL} is captured by the labor wedge, $(1 - \Delta_t)$, which satisfies

$$(1 - \Delta_t) := \frac{H_t^{ACTUAL}}{H_t^{NM}}.$$

The difference between the labor wedge and unity, $-\Delta_t$, measures the percent deviation between actual hours and model hours. We follow Prescott (2004), Ohanian et al. (2008), and Shimer (2009) in focusing on the long-run behavior of labor hours

[6] The notation := means that the object on the left-hand side of this symbol is defined by the object on its right-hand side.

by using yearly data. In order to gauge the performance of the standard model we compare model-predicted hours to actual hours.

A country's model-predicted hours H^{NM} are generated using equation (4.2) as follows. As is standard in the literature, we normalize all within-country variables by its working-age population; thus, the standard model is assumed to predict hours per working age population.[7] We use annual data from 1960 through 2006, detailed in Appendix A. Using data on real output and consumption, and assuming a value for ε, we generate the series $(Y_t/C_t)^{\frac{\varepsilon}{1+\varepsilon}}$ for each country. The final version of model hours requires scaling this series by $(1-\alpha)/\gamma$. Let $\kappa = (1-\alpha)/\gamma$. In the model, κ is constant over time. Therefore, we can use κ as a free parameter to calibrate the model to hit a predetermined target. Following Shimer (2009), we define

$$\kappa^{NM} := \frac{mean\left((H_t/P_t)^{ACTUAL}\right)}{mean\left((Y_t/C_t)^{\frac{\varepsilon}{1+\varepsilon}}\right)},$$

which implicitly allows for cross-country heterogeneity in κ. Hence, for each country we choose the scaling parameter κ such that mean model-generated hours H^{NM} is equal to that country's mean actual hours H^{ACTUAL}.

Although micro estimates of the Frisch elasticity of labor supply usually imply values of ε less than unity, macro estimates are on average slightly higher than 1. Some studies develop explanations by which these difference can be reconciled (Chetty (2009a), Rogerson and Wallenius (2009)). However, a line of research, especially that regarding real business cycle analysis, tends to impute ε by choosing a value for this parameter that makes the model-predicted cyclical fluctuations in labor hours most closely match the cyclical fluctuations in the true data. This approach leads to much higher choices of ε than those mentioned earlier. A similar approach regarding the trend behavior of labor hours rather than their cyclical fluctuations leads

[7] From here on out we will use working age population and population interchangeably.

Prescott (2004) to impute $\varepsilon = 3$, and Shimer (2009) to impute $\varepsilon = 4$.

We compare actual hours per working-age population with model H/P generated alternatively with $\varepsilon = 1$ and $\varepsilon = 4$. Figure 4.1 shows the model generated hours per working age population with ε set to 1 and and to 4. As noted above, the appropriate parameter value for the Frisch labor supply elasticity is debatable. However, the results in Figure 4.1 show that the values for ε under consideration make little difference in terms of their effect on the trend behavior of model-generated hours. Henceforth, we follow Shimer (2009) and set ε equal to 4. Figure 4.2 shows the actual hours per working age population along with their model-generated counterparts using ε equal to 4. The figure illustrates that although the model performs well when predicting the trend behavior in Canada, this is not the case in the other countries. The residual of the model's predictions relative to the data is captured by the labor wedge graphed in Figure 4.3. The models's accuracy is best when the wedge is closest to one.

4.3.2 The Model with Taxes

That trends in hours per working age population are vastly different across countries stimulated interest into potential causes for reconciling this stylized fact. One explanation is that taxes contribute to these differences both across countries and within-country over time. Thus a growing body of literature (Prescott (2004), Ohanian et al. (2008), Shimer (2009)) incorporates taxation into the standard neoclassical model. This literatures argues that a fraction of the labor wedge is accounted for by taxes. Below, we rederive the equation for labor hours including a broad set of taxes.

Following past literature, we assume that the statutory incidence of all taxes is on consumers making the household's budget constraint

$$(1+\tau_t^c)C_t + (1+\tau_t^i)I_t \leq (1-\tau_t^h)W_tH_t + (1-\tau_t^k)R_tK_t.$$

Above, τ^c, τ^i, τ^h, and τ^k, are respectively consumption, investment, labor, and capital taxes. The counterpart of equation (4.2) is now

$$H_t^{NMT} := \left((1-\tau_t)\frac{(1-\alpha)Y_t}{\gamma C_t}\right)^{\frac{\varepsilon}{1+\varepsilon}}, \tag{4.3}$$

where $\tau = \left(\tau^h + \tau^c\right) / (1+\tau^c)$ is the effective tax rate and NMT stands for neoclassical model with taxes. The tax-inclusive model reveals that $(1-\tau)^{\frac{\varepsilon}{1+\varepsilon}}$ is included in the labor wedge of the standard model when taxes are ignored. Hence when taxes are included, the labor wedge satisfies,

$$(1-\Delta_t) := \frac{H_t^{ACTUAL}}{H_t^{NMT}}. \tag{4.4}$$

We generate model hours for the period 1960-2006 using equation (4.3) normalized by the working-age population, along with data on C and Y. Following Ohanian et al. (2008), we use the effective marginal tax series created by McDaniel (2007), which includes calculated taxes on both income and consumption. McDaniel's methods are similar to that of Mendoza et al. (1994), though the data is mainly derived from national accounts publications. Both income and expenditure data and tax revenue are all categorized into labor or capital income and consumption or private investment. Tax rates are then calculated by dividing the tax revenue by either the income or expenditure for that category. This method for calculating tax rates has the nice feature that taxes can be derived independent of tax return data using only aggregate data. However, a trade off is made where strong assumptions are required for classifying the data into categories, which necessary impact the results. An addition drawback from this method is that the calculated labor income tax rates are average tax rates rather than marginal tax rates. However, McDaniel provides a comparison of the average tax rates and average marginal tax rates series calculated from past studies for the US and finds a similar trend behavior in each of the two series. The McDaniel (2007) tax data is summarized in Appendix A.[8] The normalized model is once again assumed to predict hours per working-age population. To generate model hours, we continue to set ε to 4. For each country in our OECD sample we generate the series $((1-\tau)Y_t/C_t)^{\frac{\varepsilon}{1+\varepsilon}}$. In this case note that for any given country the scaling parameter $\kappa = (1-\alpha)/\gamma$ must now satisfy

$$\kappa^{NMT} := \frac{mean\left((H_t/P_t)^{ACTUAL}\right)}{mean\left((1-\tau_t)\left(Y_t/C_t\right)^{\frac{\varepsilon}{1+\varepsilon}}\right)}.$$

Figure 4.4 presents model-generated and actual hours per working-age population. For comparison, Figure 4.5 shows the wedges generated by the standard model and the wedges generated by standard model augmented with taxes. The wedge generated by the model with taxes is closer to one for France and Germany. This highlights the improvement made in terms of predicting the long-run behavior of H/P relative to the standard model for these countries. However, as noted by both Prescott (2004) and Ohanian et al. (2008), when taxes are included, the model's predictions for the US and Canada fail to account for the true data.[9]

4.4 The Role of the E/P Ratio

Given that $H/P = (H/E)\cdot(E/P)$, understanding why the standard model with taxes fares poorly for some countries may be illuminated by understanding on which margin it is failing: H/E, E/P, or both. Thus, it is useful to disentangle the relative influence of H/E and E/P in shaping the observed patterns in H/P. Figure 4.6,

[8] See McDaniel (2007) for a more detailed explanation for how tax rates are calculated.

[9] We calculate a sum of the squared differences, where the difference is between the actual data on hours and the model predicted hours with and without taxes. The sum of squared differences is lower when taxes are included in the model for France and Germany, but higher when taxes are included in the model for Canada and the US.

shows the actual hours per population for Canada, France, Germany, and the US. The graph illustrates the behavior of H/P, had E/P remained fixed at its 1960 value and only H/E changed. Also included is the behavior of H/P, had H/E remained fixed at its 1960 value and only E/P changed. In Canada and the US, the hours per population have been increasing while hours per worker have been decreasing. Thus, in Canada and the US, the long-run trend in H/P is predominately driven by changes in E/P. This stands in contrast to France and Germany, where both hours per population and hours per worker have been decreasing while E/P has remained mostly constant.

Recall that the standard model with taxes was shown to provide accurate predictions of H/P for France and Germany, but not for Canada and the US. Combined with the patterns in Figure 4.6, this suggests that it may be the case that the standard model is inherently incapable of predicting changes in employment and instead a better predictor of hours per worker. If this is true, when E/P does not change much relative to H/E, then the model will give the impression of being successful in predicting H/P only as a matter of coincidence.

The theoretical predictions stemming from NM and NMT regarding equilibrium hours of work do not specify whether they are in per worker or per population terms. However, it has become standard in the literature to normalize all variables in the model by population, P. As noted above, the model is therefore assumed to predict H/P. Normalizing by population implicitly assumes that all household members equally share consumption utility as well as work-hours disutility. In both the NM and NMT, assuming that disutility from work-hours is shared across the population is the same as assuming both that everyone works the same amount of hours and that the entire population is employed.

Alternatively, suppose that equation (4.3) actually satisfied

$$\frac{H_t}{E_t} := \left((1-\tau_t) \frac{(1-\alpha)Y_t}{\gamma C_t} \right)^{\frac{\varepsilon}{1+\varepsilon}}. \tag{4.5}$$

That is, suppose that the model's prediction for hours were explicitly in per-worker terms. In this case, model hours per worker can be generated for each country by creating the series $((1-\alpha)Y_t/\gamma C_t)^{\frac{\varepsilon}{1+\varepsilon}}$ and then scaling it by the parameter

$$\kappa := \frac{mean\left((H_t/E_t)^{ACTUAL} \right)}{mean\left((1-\tau_t)\,(Y_t/C_t)^{\frac{\varepsilon}{1+\varepsilon}} \right)}.$$

Figure 4.7 shows the actual hours per worker for Canada, France, Germany, and the US along with hours per worker generated using equation (4.5). If the model is assumed to predict H/E, as in equation (4.5), then the model-generated data is virtually identical to the actual data.

If the standard model with taxes provides good predictions of hours per worker, then it should also be the case that multiplying these hours by a correct prediction of E/P would yield correct predictions of H/P for all countries, including the US and Canada. Unfortunately, the standard model only provides an equation for labor hours. However, if equation (4.5) is a good approximation to actual behavior of H/E, then multiplying the implied model hours by each country's actual E/P ratio should yield a largely correct approximation of each country's actual H/P ratio. Let

$$\left(\frac{H_t}{P_t}\right)^{HYBRID} := \left(\frac{E_t}{P_t}\right)^{ACTUAL} \left(\frac{H_t}{E_t}\right),$$

where H/E are model-generated hours per worker as implied by equation (4.5). Hybrid hours per working-age population as well as actual H/P are shown in Figure 4.8 . Hybrid hours per working-age population perform extremely well in accounting for actual H/P for each country. In particular, for Canada and the US, the trend behavior of H/P is correct.

The analysis thus far, suggests that the standard neoclassical model when it includes taxes, is better suited for predicting H/E rather than H/P. The degree to which this is true can be gauged by one final test. If the standard neoclassical model is incapable of predicting E/P, then E/P will in practice fall into the labor wedge defined in equation (4.4). Figure 4.9 graphs the labor wedge implied by equation (4.4) along with each country's E/P ratio. Except for a scaling constant, the long-run behavior of these two series track one another surprisingly well. This suggests that a significant fraction of the labor wedge in equation (4.4) is the E/P ratio.

Prescott (2004), Ohanian et al. (2008), and Shimer (2009) correctly identify taxes as part of the labor wedge. Our research complements these previous findings in that we show the residual generated by the standard model with taxes is comprised of the E/P ratio. This is largely due to the fact that the standard model lacks predictive power for changes in employment.

Our conclusions have mixed implications for the success of the standard model. The bad news is that successes attributed to the standard model with taxes regarding predictions of the long-run behavior of H/P exist when the E/P ratio does not change much relative to H/E. In other words, the model is successful when the behaviors of H/E and H/P are virtually indistinguishable. This gives the model the appearance of correctly predicting H/P. However, when E/P does change significantly relative to H/E, as has been the case over the last several decades in Canada and the US, the model's predictions fail. For example, suppose all countries had experienced comparable changes in E/P relative to H/E as Canada and the US. Then, research based on the tax-enhanced standard model might have incorrectly implied that taxes have a negligible impact on predicting the long-run behavior of work hours. On the other hand, the good news are two-fold. First, our analysis shows that the current theory produces a relatively accurate prediction for H/E, which implies a strong theoretical understanding of the determinants for this variable. Second, we show that the E/P ratio accounts for a large portion of the trend behavior of the labor wedge that remains after taxes are accounted for in the standard model. This is helpful for guiding the development of future research on the behavior of H/P, which is necessary for understanding the implications of tax policy. Tax policy can have differential impacts on the extensive and intensive margins for labor supply decisions. In particular, the average tax rate is associated with changes on the extensive margin while the marginal tax rate impacts the intensive margin. Thus in the next section, we develop a model that disentangles the household's choice between employment and labor hours. We then analyze how each margin of labor supply is impacted by changes in the marginal tax rate and additionally the impact of changes in the average and marginal tax rates when they differ.

4.5 Heterogeneity

We have argued that the equation for hours implied by the standard theory is more successful for predicting hours per worker rather than hours per working-age population. When the model is normalized by employment, the consumption utility will also be on a per worker basis. On the other hand, if the model is normalized by population, then a within-household distribution is implicitly established where all household members share utility from consumption and disutility from work-hours equally. In other words, normalizing by population establishes that all household members consume the same amount and work the same number of hours. This implicitly dampens the representative household's labor disutility since aggregate hours are normalized by a group that includes non-workers.

The differences stemming from choice in normalization highlight an aspect of the model that is lacking, which is its ability to distinguish between the intensive and extensive margin. As an alternative, we develop a model that explicitly incorporates the possibility that some individuals are employed and others (voluntarily) are not.[10] That is, we allow for heterogeneity in employment status. If there were no cost to employment, then all individuals would be employed in order to maximize utility. Thus, we include a fixed cost to employment that motivates the heterogeneity in employment status. In our model, hours per worker and employment are disentangled as choice variables to be optimized by a household planner. The result is an equation for equilibrium labor hours in terms of hours per worker. Interestingly, this equation is almost identical to the one stemming from the employment-normalized standard model. In fact, we show that the hours equation stemming from both the population-normalized and the employment-normalized models are special cases of the hours equation from the model developed below. This model provides a theoretical rationalization for our earlier finding that the standard model is relatively better at predicting hours per worker.

In addition, we derive conditions to show the impact that taxes and the employment fixed cost have on both hours per worker and employment. We show that the dominating effect depends on the household planner's respective weights on employed and non-employed individuals. We then continue our discussion by relaxing the assumption of a flat tax on wage income to allow for a more realistic graduated wage income tax. By allowing for a graduated wage tax, we can isolate the differential impact that average verses marginal taxes have on labor supply. In particular, we derive conditions for how both average tax rates and marginal tax rates each impact

[10] In the spirit of market clearing, we focus on non-employment rather than unemployment.

hours per worker and employment.

4.5.1 Theory

Suppose that the within-period utility of an employed individual is given by

$$\ln(c_t) - \gamma\left(\phi_t - \frac{\varepsilon}{1+\varepsilon}\left(h_t\right)^{(1+\varepsilon)/\varepsilon}\right),$$

where c is the individual's consumption, ϕ is a fixed cost associated with employment, and h is the individual's work hours. As before γ is a labor disutility parameter and ε is the Frisch elasticity of labor supply.[11] Note that the within-period utility of a non-employed individual is simply given by $\ln(c_t)$.

In standard representative agent macroeconomic models, the population consists of a continuum of infinitely divisible individuals that is normalized to 1. Given that individuals are assumed to be identical, the household's instantaneous utility is equivalent to that of a single representative agent multiplied by the number of individuals (i.e., 1). The model we develop maintains the assumption of an infinitely divisible population (that we will normalize to unity), but extends the household planner's problem to include the possibility for non-employment. Let P denote the population and, as in the standard model, assume that all individuals in the population are grouped in a single household in which resources are pooled. We assume individuals are altruistic and their joint objective is to maximize the household's utility. Thus, suppose a household planner maximizes the joint utility $\mathcal{U}$ of the household's P members, taking prices, and government policy, as given. In particular, the household

[11] Of course, modeling the determinants of employment is a non-trivial task as idiosyncratic worker aspects and search frictions are just some of the many potential factors that can affect this variable. Here, we have taken a parsimonious approach by assuming that the disutility from employment enters the household's utility as a per worker time-varying fixed cost. With the exception of allowing for a time-varying employment fixed cost, the instantaneous utility function we use is the same as that used in Kimball and Shapiro (2008).

planner's objective is to maximize

$$\mathcal{U} = \mathbb{E}_t \sum_{t=0}^{\infty} \beta^t \left\{ \psi^E E_t \left(\ln \left(\frac{C_t^E}{E_t} \right) - \gamma \phi_t - \gamma \frac{\varepsilon}{1+\varepsilon} \left(\frac{H_t}{E_t} \right)^{(1+\varepsilon)/\varepsilon} \right) \right.$$
$$\left. + \psi^N \left(P_t - E_t \right) \ln \left(\frac{C_t^N}{P_t - E_t} \right) \right\}$$

subject to

$$\left(1+\tau_t^c\right)\left(C_t^E + C_t^N\right) + \left(1+\tau_t^i\right) I_t \leq \left(1-\tau_t^l\right) W_t H_t + \left(1-\tau_t^k\right) R_t K_t + T_t.$$

Above, ψ^E is the weight the planner places on the utility of employed individuals, E is the number of employed individuals, C^E is the fraction of total household consumption, C, that each employed individual receives, H is total work hours, $\psi^N = 1 - \psi^E$ is the weight the household places on the utility on non-employed individuals, and C^N is the fraction of total household consumption that each non-employed individual receives. All other variables as well as the capital accumulation equation are described earlier in the paper in Section 4.3.[12]

The choice of employment versus non-employment matters on two important dimensions. First, note that since non-employed individuals do not work, they contribute no labor disutility to $\mathcal{U}$. Also, what matters for total labor income is total labor hours H. Once there is a choice between employment E and hours per worker H/E, where explicit disutility from the former is linear and from the latter is convex, the decision over which margin to adjust total hours given changes in economic conditions becomes relevant. Indeed, note that the relative disutilities of hours-per-worker and employment change at different rates.

Let λ_t be the Lagrange multiplier associated with the household planner's problem. We continue to focus on the labor market. The first-order conditions for

[12]Of course, the number of people employed is not a continuous variable. However, in this context, by choosing employment, the household planner is implicitly choosing the fraction of the population that is employed. Therefore treating employment as a continuous variable is inoccuous.

consumption, total work hours, and employment, after rearranging, imply that

$$\psi^E E_t / C_t^E = \lambda_t \left(1 + \tau_t^c\right), \tag{4.6}$$

$$\psi^N \left(P_t - E_t\right) / C_t^N = \lambda_t \left(1 + \tau_t^c\right), \tag{4.7}$$

$$\psi^E \gamma \left(\frac{H}{E_t}\right)^{1/\varepsilon} = \lambda_t \left(1 - \tau_t^l\right) W_t, \tag{4.8}$$

and

$$\begin{aligned}\psi^E \left(\ln\left(\frac{C_t^E}{E_t}\right) - \gamma\phi_t\right) - \psi^N \ln\left(\frac{C_t^N}{\left(P_t - E_t\right)}\right) \\ - \psi^E + \psi^N + \psi^E \frac{\gamma}{1+\varepsilon}\left(\frac{H_t}{E_t}\right)^{(1+\varepsilon)/\varepsilon} = 0.\end{aligned} \tag{4.9}$$

Combining the consumption first-order conditions, it follows that within any period

$$C_t^N = \frac{\psi^N \left(P_t - E_t\right)}{\psi^E E_t} C_t^E.$$

Using this, and the fact that the sum of C_t^E and C_t^N must equal total household consumption, C_t, implies that

$$C_t^E = \zeta_t C_t, \tag{4.10}$$

where

$$\zeta_t = \psi^E E_t / \left(\psi^E E_t + \psi^N \left(P_t - E_t\right)\right)$$

is the fraction of total household consumption that the planner assigns to employed individuals.

In order to close the model, we must once more consider the firm's problem. The firm chooses total work hours H and capital K in order to maximize

$$\begin{aligned}\Pi_t &= Y_t - W_t h_t E_t - R_t K_t \\ &= Z_t K_t^{\alpha} \left(H_t\right)^{1-\alpha} - W_t H_t - R_t K_t.\end{aligned} \tag{4.11}$$

The first-order conditions for hours per worker and employment both result in

$$(1-\alpha) Z_t K_t^{\alpha} (H_t)^{-\alpha} = W_t$$
$$\implies (1-\alpha) Y_t = W_t H_t. \tag{4.12}$$

Combining (4.12) with (4.8) yields

$$\psi^E \gamma \left(\frac{H_t}{E_t}\right)^{(1+\varepsilon)/\varepsilon} = \lambda_t \left(1-\tau_t^l\right)(1-\alpha)\frac{Y_t}{E_t}.$$

Substitute in for λ_t using (4.6). It follows that

$$\gamma \left(\frac{H_t}{E_t}\right)^{(1+\varepsilon)/\varepsilon} = (1-\alpha)(1-\tau_t)\frac{Y_t}{C_t^E},$$

where τ is the effective tax rate, defined earlier in the paper in Section 4.3.2. Substituting in for C_t^E, using (4.10), and rearranging yields

$$h_t = \left(\left(\frac{\psi^E E_t + \psi^N (P_t - E_t)}{\psi^E E_t}\right) \cdot \left((1-\tau_t)(1-\alpha)\frac{Y_t}{\gamma C_t}\right)\right)^{\varepsilon/(1+\varepsilon)}, \tag{4.13}$$

where $h_t = H_t/E_t$ is hours per worker.

The right-hand side of equation (4.13) defines the hours per worker that are theoretically consistent with the household's optimal choice of employment, taxes, and output to consumption ratio. Note that as $\psi^E \to 1$,

$$\left(\psi^E E_t + \psi^N (P_t - E_t)\right)/\psi^E E_t \to 1.$$

Therefore, as $\psi^E \to 1$, equation (4.13) converges to the prediction of hours that the standard model enhanced with taxes yields when all variables are normalized by the level of employment. On the other hand, when $\psi^E = \psi^N = 0.5$, slight rearrangement of equation (4.13) implies that the prediction for hours is the same as for the model with all variables normalized by population. Thus, for $\psi^E = \psi^N = 0.5$, the first-order condition from our model is the equivalent to that of the standard model normalized by population.

We now derive an expression for employment per population. Consider once more equation (4.9). Substituting in for C_t^E, C_t^N, and rearranging yields

$$\psi^N \ln\left(\frac{\psi^N}{\psi^E E_t + \psi^N (P_t - E_t)} C_t\right) - \psi^E \left(\ln\left(\frac{\psi^E}{\psi^E E_t + \psi^N (P_t - E_t)} C_t\right) - \gamma\phi_t\right)$$
$$= \psi^E \frac{\gamma}{1+\varepsilon} h_t^{(1+\varepsilon)/\varepsilon} + \psi^N - \psi^E.$$

Further rearrangement implies that

$$F(E_t) = \psi^E \frac{\gamma}{1+\varepsilon} h_t^{(1+\varepsilon)/\varepsilon} - \left(\psi^E - \psi^N\right)(1 - \ln(C_t)) - \psi^E \gamma\phi_t, \tag{4.14}$$

where

$$F(E_t) = \psi^N \ln\left(\psi^N / \left(\psi^E E_t + \psi^N (P_t - E_t)\right)\right) - \psi^E \ln\left(\psi^E / \left(\psi^E E_t + \psi^N (P_t - E_t)\right)\right).$$

The right-hand of equation (4.14) implicitly defines the level of employment that is theoretically consistent with the household's optimal choices regarding aggregate consumption C and hours per worker h, given the employment fixed cost ϕ. Note that as $\psi^E \to 1$, equation (4.14) converges to

$$\ln(E_t) = \frac{\gamma}{1+\varepsilon} h_t^{(1+\varepsilon)/\varepsilon} - (1 - \ln(C_t)) - \gamma\phi_t,$$

which implies complementarity between employment and hours per worker and, as expected that employment is decreasing in ϕ.[13]

4.5.2 Comparative Statics

In the standard model, the impact of the effective tax rate is captured through only one margin of adjustment – aggregate hours worked. If the effective tax rate were

[13] Intuitively, the non-employment state can be thought of as equivalent to resting. Although, some non-employment is optimal at the household level, the utility of employed and non-employed individuals differs, which begs the question: why would some individuals accept a relatively lower utility level and not deviate from the household planner's decision? Because the household utility is maximized by a (benevolent) household planner, the resulting utility is at its max. In terms of individual utilities, the household could engage in an optimization subproblem, which we ignore for simplicity. The optimal solution to the subproblem would be rotating individuals between states of non-employment and employment so that everyone rests and works the same amount. This would allow no one household member to be stuck forever in non-employment with different consumption. In other words, the individual expected flow utility streams would be equalized across the population.

increased, then aggregate hours would respond by decreasing. In our model, there are two margins that can adjust to changes in the effective tax rate including both hours per worker and changes in the number of workers. To examine how changes in the effective tax rate and the fixed cost associated with employment impact hours per worker and the number of workers, we totally differentiate the equations presented above. We can then isolate four objects of interest, which are the elasticities of both hours per worker and employment with respect to the net-of-tax rate and the fixed cost, holding all other variables fixed.

As shown in Appendix B,

$$\begin{aligned}\left(1+\frac{\psi^N\psi^E}{\left(\psi^E-\psi^N\right)^2\left(E_t/P_t\right)}\frac{\gamma}{\varepsilon}h_t^{(1+\varepsilon)/\varepsilon}\right)d\ln\left(h_t\right)=\left(\frac{\psi^N}{\left(\psi^E-\psi^N\right)^2\left(E_t/P_t\right)}\right)\left(\psi^E\gamma\phi_t\right)d\ln\phi_t\\+\frac{\varepsilon}{1+\varepsilon}d\log\left(1-\tau_t\right)+\frac{\varepsilon}{1+\varepsilon}d\ln\left(Y_t/C_t\right)-\left(\frac{\psi^N}{\left(\psi^E-\psi^N\right)\left(E_t/P_t\right)}\right)d\ln\left(C_t/P_t\right)\end{aligned} \tag{4.15}$$

and

$$\begin{aligned}\left(1+\left(\frac{\psi^N}{\left(\psi^E-\psi^N\right)^2\left(E_t/P_t\right)}\right)\left(\psi^E\frac{\gamma}{\varepsilon}h_t^{(1+\varepsilon)/\varepsilon}\right)\right)d\ln\left(E_t/P_t\right)=\\-\left(\frac{\psi^E E_t/P_t+\psi^N\left(1-\left(E_t/P_t\right)\right)}{\left(\psi^E-\psi^N\right)^2\left(E_t/P_t\right)}\psi^E\gamma\phi_t\right)d\ln\left(\phi_t\right)\\+\left(\frac{\psi^E(E_t/P_t)+\psi^N\left(1-\left(E_t/P_t\right)\right)}{\left(\psi^E-\psi^N\right)^2\left(E_t/P_t\right)}\right)\left(\psi^E\frac{\gamma}{\varepsilon}h_t^{(1+\varepsilon)/\varepsilon}\right)\frac{\varepsilon}{1+\varepsilon}\left(d\ln\left(1-\tau_t\right)+d\ln\left(Y_t/C_t\right)\right)\\+\left(\frac{\psi^E E_t/P_t+\psi^N\left(1-\left(E_t/P_t\right)\right)}{\left(\psi^E-\psi^N\right)\left(E_t/P_t\right)}\right)d\ln\left(C_t/P_t\right).\end{aligned} \tag{4.16}$$

This means that hours per worker are increasing in the employment fixed costs, ϕ, while employment is decreasing in ϕ. This result intuitively illustrates the substitution between hours per worker and employment. Both hours per worker and employment are increasing in the net-of-tax rate, $(1-\tau)$. Thus, an increase in the effective tax rate τ causes the household to decrease both hours per worker and employment, ceteris paribus.

From equations (4.15) and (4.16), we can obtain elasticities of hours per worker and of employment with respect to ϕ holding all other variables fixed. For hours per worker,

$$\frac{d\ln h_t}{d\ln \phi_t} = \frac{\left(\frac{\psi^N}{(\psi^E-\psi^N)^2(E_t/P_t)}\right)\left(\psi^E\gamma\phi_t\right)}{1+\frac{\psi^N\psi^E}{(\psi^E-\psi^N)^2(E_t/P_t)}\frac{\gamma}{\varepsilon}h_t^{(1+\varepsilon)/\varepsilon}},$$

while for employment,

$$\frac{d\ln(E_t/P_t)}{d\ln \phi_t} = \frac{\left(\frac{\psi^E(E_t/P_t)+\psi^N(1-(E_t/P_t))}{(\psi^E-\psi^N)^2(E_t/P_t)}\psi^E\gamma\phi_t\right)}{\left(1+\left(\frac{\psi^N}{(\psi^E-\psi^N)^2(E_t/P_t)}\right)\left(\psi^E\frac{\gamma}{\varepsilon}h_t^{(1+\varepsilon)/\varepsilon}\right)\right)}.$$

Given a 1 percent increase in ϕ the percentage employment decrease is greater than the percentage increase in hours per worker if and only if

$$\frac{\left(\frac{\psi^E(E_t/P_t)+\psi^N(1-(E_t/P_t))}{(\psi^E-\psi^N)^2(E_t/P_t)}\psi^E\gamma\phi_t\right)}{\left(1+\frac{\psi^N}{(\psi^E-\psi^N)^2(E_t/P_t)}\psi^E\frac{\gamma}{\varepsilon}\left(h_t^{(1+\varepsilon)/\varepsilon}\right)\right)} > \frac{\left(\frac{\psi^N}{(\psi^E-\psi^N)^2(E_t/P_t)}\right)\left(\psi^E\gamma\phi_t\right)}{1+\frac{\psi^N\psi^E}{(\psi^E-\psi^N)^2(E_t/P_t)}\frac{\gamma}{\varepsilon}h_t^{(1+\varepsilon)/\varepsilon}}$$
$$\iff \psi^E > \psi^N.$$

Thus, the employment response to a change in fixed cost is unambiguously greater than that of hours per worker when $\psi^E > \psi^N$.

Next, we consider the elasticities of hours per worker and of employment with respect to the net-of-tax rate $(1-\tau)$, again holding all other variables fixed. The elasticity of hours per worker with respect to $(1-\tau)$ is

$$\frac{d\ln h_t}{d\ln(1-\tau_t)} = \frac{\frac{\varepsilon}{1+\varepsilon}}{1+\frac{\psi^N\psi^E}{(\psi^E-\psi^N)^2(E_t/P_t)}\frac{\gamma}{\varepsilon}h_t^{(1+\varepsilon)/\varepsilon}},$$

and the elasticity of employment with respect to $(1-\tau_t)$ is

$$\frac{d\ln(E_t/P_t)}{d\ln(1-\tau_t)} = \frac{\left(\frac{\psi^E(E_t/P_t)+\psi^N(1-(E_t/P_t))}{(\psi^E-\psi^N)^2(E_t/P_t)}\right)\left(\psi^E\frac{\gamma}{\varepsilon}h_t^{(1+\varepsilon)/\varepsilon}\right)\frac{\varepsilon}{1+\varepsilon}}{\left(1+\left(\frac{\psi^N}{(\psi^E-\psi^N)^2(E_t/P_t)}\right)\left(\psi^E\frac{\gamma}{\varepsilon}h_t^{(1+\varepsilon)/\varepsilon}\right)\right)}.$$

This means a 1 percent increase in $(1-\tau)$ causes a greater percentage increase in employment than in hours per worker if and only if

$$\frac{\left(\frac{\psi^E(E_t/P_t)+\psi^N(1-(E_t/P_t))}{(\psi^E-\psi^N)^2(E_t/P_t)}\right)\left(\psi^E\frac{\gamma}{\varepsilon}h_t^{(1+\varepsilon)/\varepsilon}\right)\frac{\varepsilon}{1+\varepsilon}}{\left(1+\left(\frac{\psi^N}{(\psi^E-\psi^N)^2(E_t/P_t)}\right)\left(\psi^E\frac{\gamma}{\varepsilon}h_t^{(1+\varepsilon)/\varepsilon}\right)\right)} > \frac{\frac{\varepsilon}{1+\varepsilon}}{1+\frac{\psi^N\psi^E}{(\psi^E-\psi^N)^2(E_t/P_t)}\frac{\gamma}{\varepsilon}h_t^{(1+\varepsilon)/\varepsilon}}$$

$$\iff \left(\frac{\psi^E(E_t/P_t)+\psi^N\left(1-(E_t/P_t)\right)}{\left(\psi^E-\psi^N\right)^2(E_t/P_t)}\right)\left(\psi^E\frac{\gamma}{\varepsilon}h_t^{(1+\varepsilon)/\varepsilon}\right) > 1.$$

Thus, the relative response between hours per worker and employment to a tax change is ultimately ambiguous. For example, consider a country with $\varepsilon = 4$, $E_t/P_t = 0.7$ and suppose $h_t = 40$. Then,

$$\left(\frac{\left(\psi^E-\psi^N\right)0.7+\psi^N}{\left(\psi^E-\psi^N\right)^2 0.7}\right)\left(\psi^E\frac{\gamma}{4}40^{1.25}\right).$$

As $\psi^E \to 1$

$$\left(\frac{\left(\psi^E-\psi^N\right)0.7+\psi^N}{\left(\psi^E-\psi^N\right)^2 0.7}\right)\left(\psi^E\frac{\gamma}{4}40^{1.25}\right) \to (25.1487)\,\gamma$$

As $\psi^E \to \psi^N$

$$\left(\frac{\left(\psi^E-\psi^N\right)0.7+\psi^N}{\left(\psi^E-\psi^N\right)^2 0.7}\right)\left(\psi^E\frac{\gamma}{4}40^{1.25}\right) \to \infty$$

As $\psi^N \to 1$

$$\left(\frac{\left(\psi^E-\psi^N\right)0.7+\psi^N}{\left(\psi^E-\psi^N\right)^2 0.7}\right)\left(\psi^E\frac{\gamma}{4}40^{1.25}\right) \to 0$$

This suggests that the relative magnitudes depend on γ and in this particular example, for $\gamma \geq 1/25.1487$, an increase in $(1-\tau_t)$ will cause a greater percentage increase in participation than hours per worker as long as ψ^E is not particularly small.

4.5.3 Average vs. Marginal Tax Rates

So far, we have assumed that wage income is taxed at a flat rate, making the marginal tax rate on wage income equivalent to its respective average tax rate. However, the true tax system may be a graduated tax schedule on personal income, where the marginal tax rate increases with income. For simplicity, consider a graduated schedule for wage income with two marginal tax rates, τ^A and τ^B, where $\tau_B > \tau_A$. The lower rate τ^A, applies to the first WH_1 dollars of wage income earned and WH_1 is the income cut off where the marginal rate changes, while τ^B applies to the next $W(H - H_1)$ dollars. Also, to isolate responses stemming from differences in the average and marginal tax rates on wage income, we focus on the case where $\tau^c = \tau^i = \tau^k = 0$. The representative household's budget constraint is now written as,

for $H \leq H_1$

$$\left(C_t^E + C_t^N\right) + I_t \leq W_t H_t - W_t H_1 \tau^A + R_t K_t + T_t$$

and for $H > H_1$

$$\left(C_t^E + C_t^N\right) + I_t \leq W_t H_t - W_t H_1 \tau^A - W_t (H_t - H_1) \tau^B + R_t K_t + T_t.$$

Thus, a graduated tax on wage income induces a kink in the household's budget constraint. This means the maximization is non-differentiable at WH_1. Nonetheless, the first-order conditions will remain valid for each segment of the household's problem. We focus on the maximization that occurs on the segment where $H > H_1$, as it highlights the household's problem when the marginal tax rate is not equivalent to the average tax rate.

Given that the average and marginal tax rates are no longer equivalent, we can analyze the impact of the average tax rate on h and E/P holding the marginal tax rate fixed. The average tax rate $\bar{\tau}$ is defined as

$$\bar{\tau}_t = \frac{W_t H_1 \tau_t^A + W_t (H_t - H_1) \tau_t^B}{W_t H_t},$$

where τ^B is the marginal tax rate for $H > H_1$. If there is an increase in τ^A with no subsequent change in τ^B, then for $H > H_1$, this represents an increase in the average tax rate with no change to the marginal tax rate. The household budget constraint can be rewritten as

$$(1+\tau_t^c)\left(C_t^E+C_t^N\right)+\left(1+\tau_t^i\right)I_t \leq W_tH_t(1-\tau^B)+W_tH_1(\tau^B-\tau^A)+\left(1-\tau_t^k\right)R_tK_t+T_t.$$

Given that WH_1 is constant, $WH_1(\tau^B-\tau^A)$ is also constant. This means that changes to the average tax rate that do not impact the marginal tax rate will have only a pure income effect. For instance, a decrease to τ^A will allow the household to increase total consumption for a given amount of investment. Thus, the impact of the average tax rate on h, and E/P is realized through changes in consumption. Equations 4.15 and 4.16 show that changes to consumption impact both changes to h and E/P where an increase in consumption leads to a decease in h and an increase E/P. However, a 1% increase in C/P will cause a greater percentage-wise increase participation than percent decrease in hours per worker if and only if

$$\frac{\left(\frac{\psi^E E_t+\psi^N(1-E_t)}{(\psi^E-\psi^N)E_t}\right)}{1+\left(\frac{\psi^N}{(\psi^E-\psi^N)^2E_t}\right)\left(\psi^E\frac{\gamma}{\varepsilon}h_t^{(1+\varepsilon)/\varepsilon}\right)} > \frac{\left(\frac{\psi^N}{(\psi^E-\psi^N)E_t}\right)}{\left(1+\frac{\psi^N\psi^E}{(\psi^E-\psi^N)^2E_t}\frac{\gamma}{\varepsilon}h_t^{(1+\varepsilon)/\varepsilon}\right)}$$
$$\iff \psi^E > \psi^N.$$

Thus, a decrease in τ^A, which in turn decreased $\bar{\tau}$, will lead to an overall increase in H/P even though h falls, if the household planner weights the employed greater than the non-employed.

To analyze the impact of the marginal tax rate, suppose τ^B decreased, while τ^A increased, keeping the average tax rate constant. This assumption allows us to abstract from the additional income effect that would occur if the average tax rate had also changed. Conditional on $H > H_1$, the planner seeks to maximize the household's utility subject to the budget constraint with varying marginal rates. The new first order condition for hours is,

$$H_t : \psi^E \gamma \left(\frac{H}{E_t}\right)^{1/\varepsilon} = \lambda_t \left(1 - \tau_t^B\right) W_t.$$

After taking first-order conditions for consumption and setting labor supply equivalent to labor demand, we obtain a new hours per worker equation in terms of only the marginal tax rate,

$$h_t = \left(\frac{\psi^E E_t + \psi^N (P_t - E_t)}{\psi^E E_t}\right) \left((1 - \tau^B) \frac{(1-\alpha)}{\gamma} \frac{Y_t}{C_t}\right).$$

The equations used in the comparative statics above can be used to study the impact of the marginal tax rate on h and E/P, though replacing $(1-\tau)$ with $(1-\tau^B)$, since we have assumed only wage taxes. The comparative statics illustrated that a decrease in τ^B (and thus an increase in $(1-\tau^B)$ will increase both h and E/P. The response with a larger percentage increase is ambiguous and depends largely on γ, which represents the distaste for working.

Overall we find that a decrease in the average tax rate, holding the marginal tax rate constant, increases the overall H/P ratio by increasing E/P more than the fall in h when $\psi^E > \psi^N$. A decrease in the marginal tax rate, holding the average tax rate fixed, will increase both h and E/P. Although the relative magnitudes of their responses is ambiguous, the overall effect will be a subsequent increase to H/P. The sensitivity of h and E/P with respect to the average and marginal tax rates depends on ψ^E/ψ^N because the weights influence how much additional disutility rhe household endures by increasing employment relative to increasing hours per worker.

4.6 Application of the Model

Relative to the standard model, our theory adds two additional parameters that affect the behavior of hours per population: ψ^E (the household planner's weight placed on employed individuals) and ϕ (the fixed cost endured by employed individuals). Although we are unable to empirically estimate either parameter, we can examine what values for each of these parameters are consistent with the observed data.

4.6.1 Hours Per Worker

As shown above, ψ^E is an important parameter for determining the direction of several comparative statics. In order to discern the values of ψ^E consistent with each country's data, we generate model predicted hours per worker using empirical measures for the variables on the right-hand side of equation (4.13). We apply the same data generating methods detailed in Section 4.3. Figures 4.10-4.13 show the actual and model-predicted hours per worker for Canada, France, Germany, and the US. The model-predicted hours per worker are graphed for $\psi^E \in \{0.1, 0.5, 0.9\}$. For each country, the model-predicted hours per worker are a closer match to their empirical counterparts as ψ^E increases.

Table 4.2 extends the overall analysis to account for 15 OECD countries for which the McDaniel (2007) tax series is available and also considers a wider set of values for ψ^E. The second column shows the mean of the annual percent changes in actual hours per worker from 1960 through 2006 in each country. The remainder of the table shows the mean annual percent change in hours per worker generated using equation (4.13) over the same period and for different values of ψ^E. The boxed values represent model-predicted results that are relatively better at matching the actual data. For two thirds of the countries, values of $\psi^E \geq 0.5$ are associated with predictions that better match the data. The exceptions are Belgium, Japan, Spain, and Switzerland, where $\psi^E = 0.01$ is a relatively better fit, and Italy where $\psi^E = 0.25$ is a relatively better fit. The relatively better fits for Canada and the US are at the other end of the spectrum with $\psi^E = 0.75$ and $\psi^E = 0.99$, respectively. The results in Table 4.2 suggest that, on average and across countries, households tend to act as if optimization is carried out by a household planner that puts a relatively heavier weight on the utility of employed individuals.

The substantial cross-country variation in the implied measures of ψ^E that best reconcile the model with the data is somewhat surprising. Although modeling the factors that determine the planner's weights is beyond the scope of this paper, it may be helpful to think of the ratio ψ^E/ψ^N as reflecting factors affecting the relative differences in the utility among employed and non-employed individuals that are not explicitly accounted for by the model. For example, extra utility arising from work activities that are enjoyable would raise the relative utility stemming from employment through channels that we have not modeled. Future research may focus on endogenizing the household planner's weights to further understand these large differences across countries.

For the purpose of comparison to the neoclassical model enhanced with taxes (NMT), Table 4.4 presents the mean of the year-to-year percent changes in actual H/P, H/E, and E/P for the period 1960-2004. Also included are the predictions of NMT regarding H/P. The fifth and sixth columns of Table 4.4 show NMT's predictions relative to the actual data on H/P and H/E, respectively. When NMT is assumed to predict H/P, on average, it explains around 60% of the data. On the other hand, when NMT is assumed to predict H/E, its explanatory power improves considerably, on average explaining 74% of the data. This improvement is especially large for countries like the US and Canada where there have been large changes in E/P. The second to last column of Table 4.4 shows our model's best predictions regarding H/E from Table 4.2. The last column shows our model's predictions relative to the actual data on H/E. On average, our model can account for roughly 84% of the the data.

4.6.2 Employment

As shown above, the trend behavior of E/P can have an important impact on the trend behavior of H/P. Since changes in ϕ directly affect the determination of E/P, it is of interest to understand the relative changes in ϕ necessary to reconcile the model with the data. Using equation (4.14), we can generate the fixed cost series consistent with the actual data on employment, hours per worker, and consumption. Note that

$$\phi_t = \frac{F\left(E_t\right)}{\psi^E \gamma} - \frac{1}{1+\varepsilon} h_t^{(1+\varepsilon)/\varepsilon} + \frac{\left(\psi^E - \psi^N\right)\left(1 - \ln\left(C_t\right)\right)}{\psi^E \gamma}. \tag{4.17}$$

It is not obvious what γ should be set to, or for that matter if it is appropriate or not to assume that γ varies across countries. However, since γ is assumed to be a time invariant scaling parameter, we can focus on the series ϕ_t/ϕ_{1960}. Hence, for a given country we scale the model-generated fixed costs series by its 1960 value. Table 4.3 shows the actual mean of the year-to-year percent change in the E/P ratios for the same period and countries as in Table 4.2. Table 4.3 also shows the mean of the year-to-year percent changes in ϕ, scaled by its 1960 value, that are consistent with explaining changes in E/P for different values of ψ^E. For $\psi^E \geq 0.25$, the model suggests that the mean of the year-to-year percent changes in ϕ_t should have decreased on average less than 1% per year in order to account for the data. This means that on a year-to-year basis a relatively small decline in the costs associated with employment is sufficient to reconcile the model with the data. Given that the fixed costs are calculated as a residual, they can also include factors affecting hours behavior that the model ignores. That the implied annual changes in fixed cost are relatively small suggests that the model may indeed accurately account for major determinants that drive changes in labor supply behavior.

4.7 Conclusion

This paper questions the ability of the standard neoclassical model augmented with taxes to predict hours per population across country. Past work has argued that by including taxes in the standard neoclassical model, much of the long-run differences in hours per population both within and across country can be accounted for. However, two countries stand out as exceptions – Canada and the US. We delve deeper into these puzzling exceptions and highlight that unlike the other countries studied, Canada and the US both experienced large increases in their employment to population ratio. Upon further inspection, we conclude that the failure of the standard neoclassical model with taxes to predict hours per population in Canada and the US stems from its inability to predict accurate changes in the employment to population ratio.

After identifying the shortcomings of the model with taxes with respect to employment changes, we develop a model that explicitly incorporates employment as a choice variable. In our model, a household planner maximizes the weighted utility over employed and non-employed individuals. Our model features a fixed cost associated with employment. This fixed cost is such that the household's optimal decisions with regards to aggregate hours per population involve a trade-off between the (linear) costs associated with employment and the (convex) costs associated with work hours. This leads hours and employment to be substitutes with regards to the employment fixed cost. In particular, an increase in the fixed cost induces a decrease in employment and an increase in hours per worker. We find that as the planner increases the weight on the utility of employed individuals, the relative magnitude of changes in employment given a change in fixed costs also increases. Thus, when the weight on employed individuals is greater, the impact on hours per population will be dominated by the employment response with respect to changes in the fixed cost to employment. In other words, for a large enough weight on employed individuals, an increase in fixed cost will lead to a decrease in hours per population, which is driven by a decrease in the employment per population. This decrease in hours per population is mitigated by the increase in hours per worker. We also analyze changes to the net-of-tax rate and find that the relative magnitude between the hours per worker response and the employment response is ambiguous. However, both respond in the same direction and therefore have an unambiguous impact on hours per population, which is that hours per population move in the same direction as the net-of-tax rate.

Additionally, we highlight the model's relevance as a tool for analyzing policy by exploring the differential impact that changes in average tax rates versus marginal taxes can have on hours per population. We find that a decrease in the average tax rate leads to a decrease in hours per worker, while employment per population increases. Similar to a change in employment fixed cost, the dominating response will depend on how the household planner weights employed individuals. Thus, a decrease in the average tax rate could potentially lead to either an increase or decrease in hours per population. This finding is particularly interesting in the US, where there were large decreases in the average tax rate, with increases in hours per population. When viewing each component of hours per population separately, as in Figure 4.6, hours per worker and employment per population in the US behave as the model predicts, with hours per worker falling while employment per population is rising.

Finally, we reconcile our model with the data to obtain reasonable parameter values for the household planner's respective weights on employed and non-employed individuals and for the long-run behavior of the employment fixed cost. If the household planner weights employed and non-employed individuals equally, then the predictions of our model regarding labor supply are the same as the standard model with population as the normalization. However, as the weight on employed individuals tends to one, our model gives the same prediction as the standard model if it were assumed to predict hours per worker. By allowing for an intermediate case on weights other than 0.5 or 1, we find the model best matches the data on hours per worker when the planner's weight on employed individuals is between .5 and 1. However, there are exceptions for some countries, which suggests that by allowing for more flexibility in the weights, we can derive more accurate predictions of hours. In terms of the the fixed cost to employment, we find that it varies little over time. Thus, only small changes in the fixed cost are necessary to explain long-run changes in the employment to population ratio.

Intuitively, the large differences across countries in the planner's weights may capture differences in non-market benefits such as unemployment compensation, welfare, and social security. Given that the household planner takes government policy as given, the differences in weights may be the result of differences in government policy across countries, among other things. Allowing for endogenous weighting of employed and non-employed individuals is beyond the scope of this paper, though future work may include an option for weights to respond to changes in government policy.

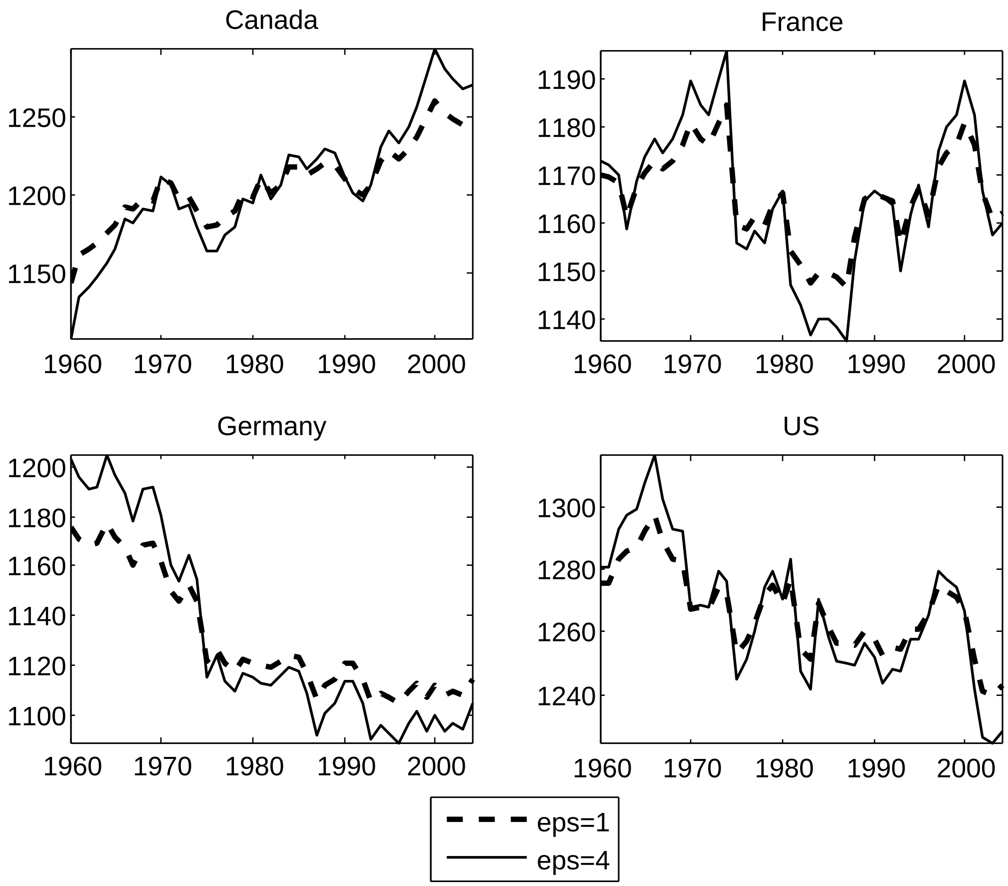
The fixed cost to employment can include any number of costs that impact an individuals decision to work. Past work has found that the manner in which the government uses its revenue can have an important impact on the incentive to work

(Ragan (2006) and Rogerson (2007)). For example, if the government were to subsidize child care, as is the case in certain Scandanavian countries, then there would be an increased incentive to substitute home production with market work (Shimer (2009)). This lowered cost to entering the work force can be interpreted as a decrease in ϕ, as it makes working relatively more attractive. The degree to which the government subsidizes market work over home production varies across countries, as well as other differences impacting the decision to work, and can thus be captured by differences in ϕ. The importance of both ψ^E and ϕ for understanding how policy impacts trends in labor hours could warrant future research to focus on obtaining a deeper understanding on the relationship between social programs within a country and its respective parameter values.

Overall, this paper makes two contributions to the literature. First, we highlight the standard neoclassical model's inability to predict long-run changes in the employment to population ratio. We show that the result is the model's failure to accurately predict long-run changes in hours per population for countries with large changes in employment. Second, we develop a model that disentangles the choice between hours per worker and employment per population in optimizing an economy's overall labor force. Thus, we allow for a better understanding of how policy changes can influence aggregate behavior by separately analyzing both the extensive and intensive margin of the labor supply response.

4.8 Figures and Tables

Figure 4.1: Neoclassical Model Generated Hours with Different ϵ



Notes.– Source of data used to create all figures and tables come from The Groningen Growth and Development Centre (www.ggdc.net), the Source OECD Database (www.sourceoecd.org), and the Penn World tables (http://pwt.econ.upenn.edu). The tax data calculated in McDaniel (2007) is available at (www.caramcdaniel.com). A more detailed description of the data can be found in Appendix A.

Figure 4.2: Actual Data on Hour per Population versus Neoclassical Model Predicted Hours

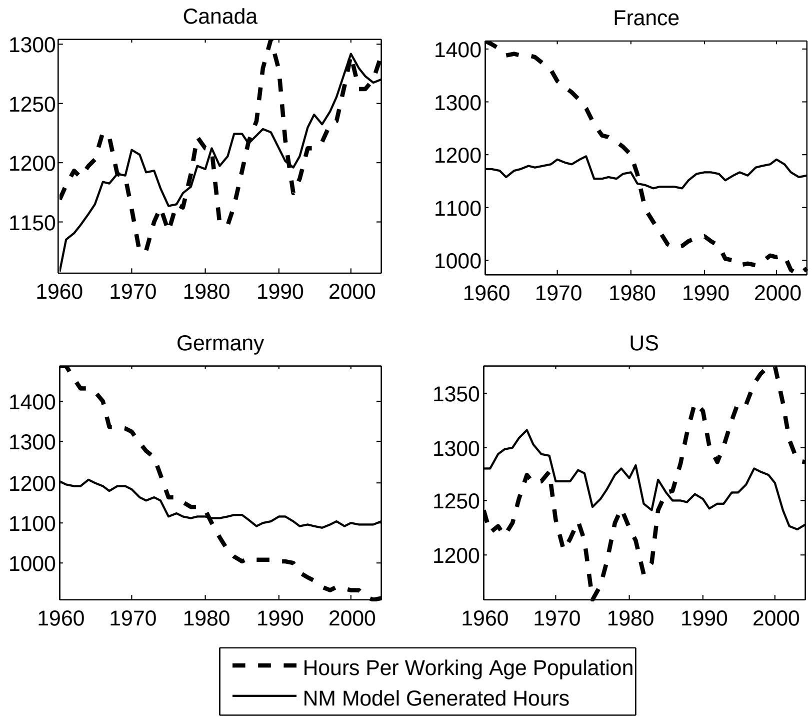


Figure 4.3: Labor Wedge from Neoclassical Model

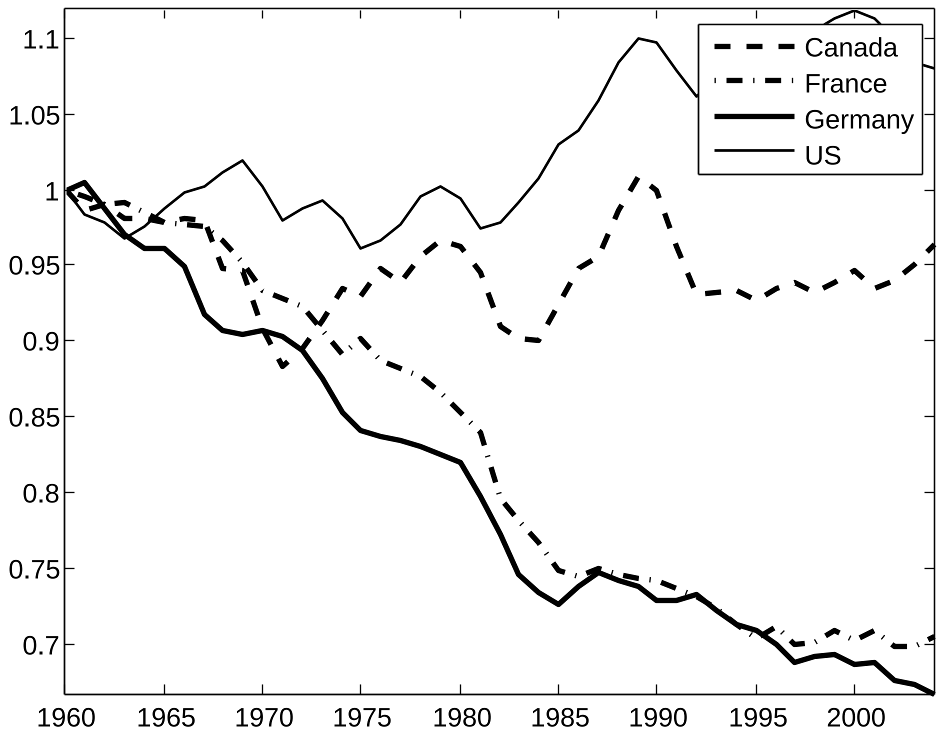


Figure 4.4: Actual Data on Hour per Population versus Neoclassical Model with Taxes Predicted Hours

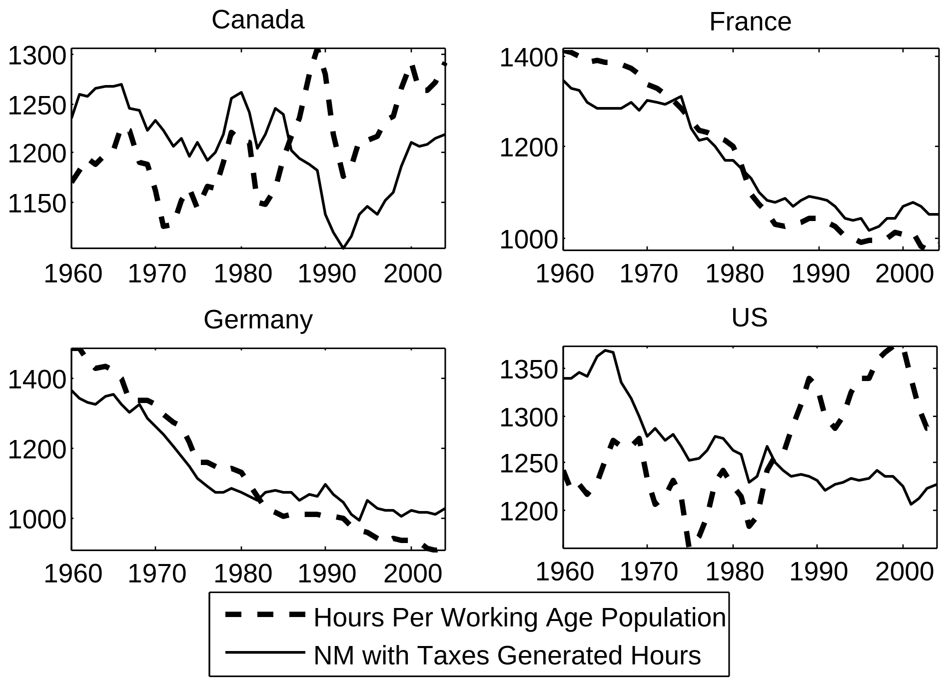


Figure 4.5: The NM Wedge and the NMT Wedge

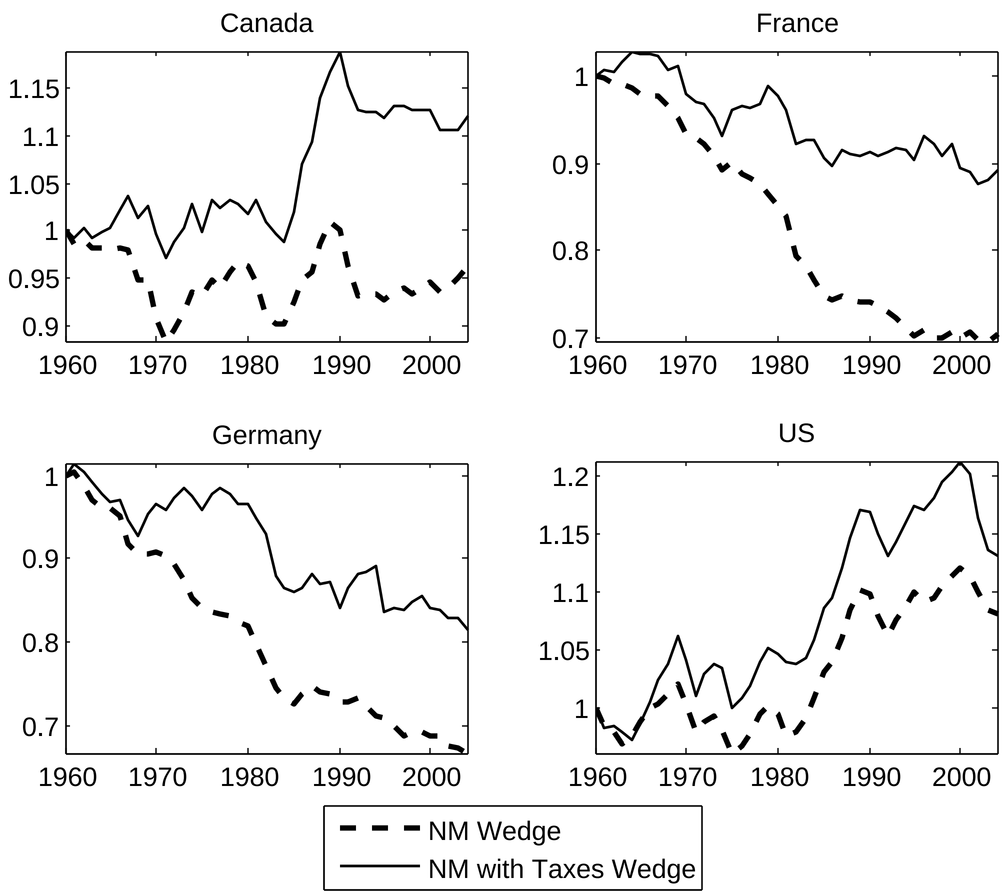


Figure 4.6: Hours Per Worker and Employment to Population's Contribution to Hours Per Population

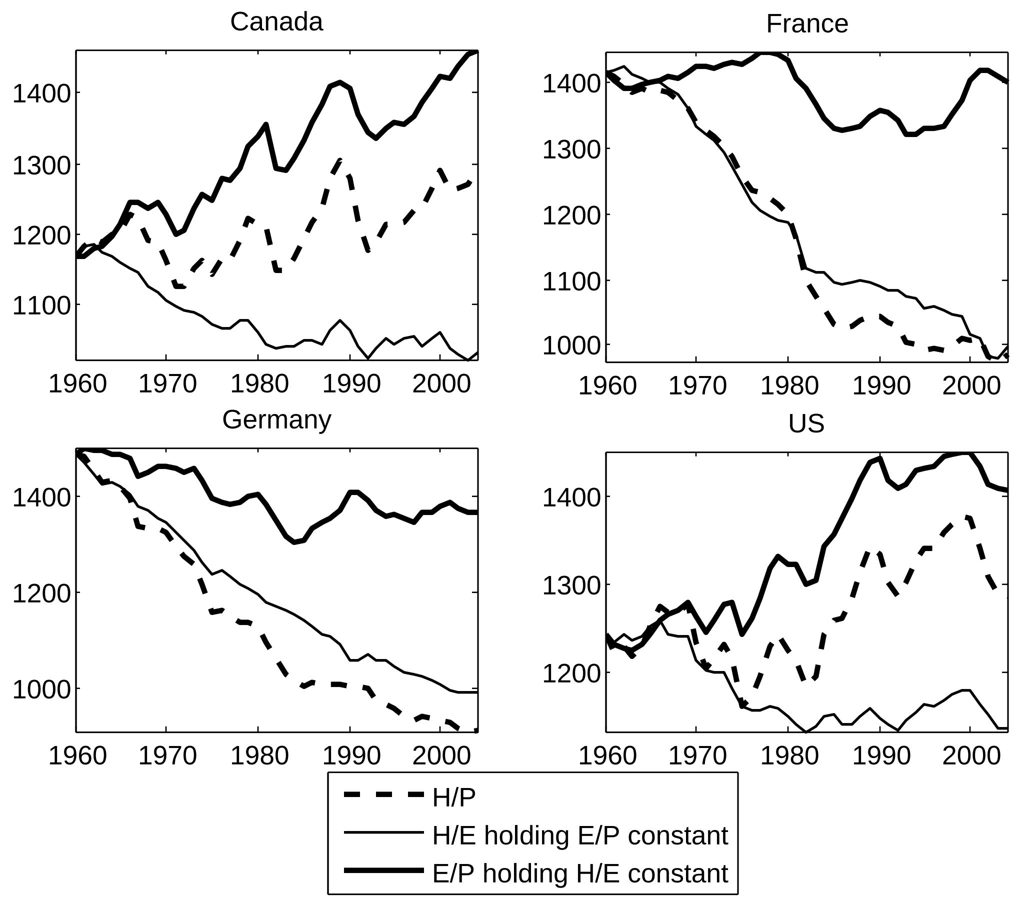


Figure 4.7: Actual Hours Per Worker and the Neoclassical Model with Taxes Prediction for H/E

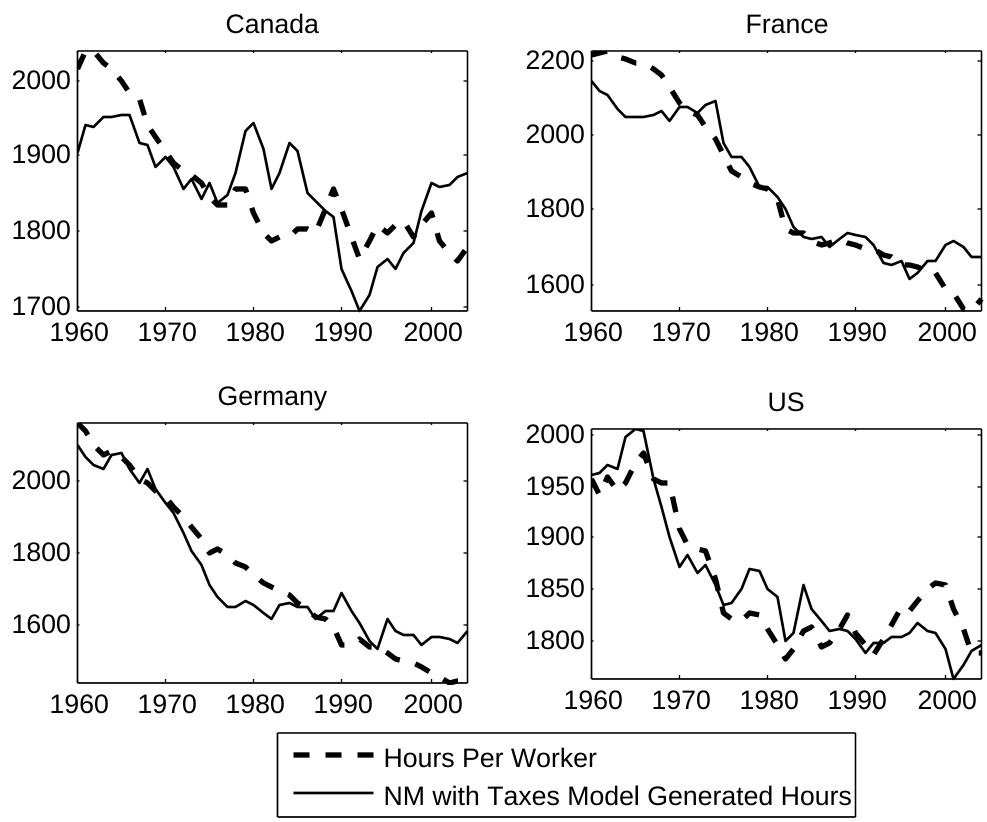


Figure 4.8: Actual Hours per Population and Hybrid Hours per Population

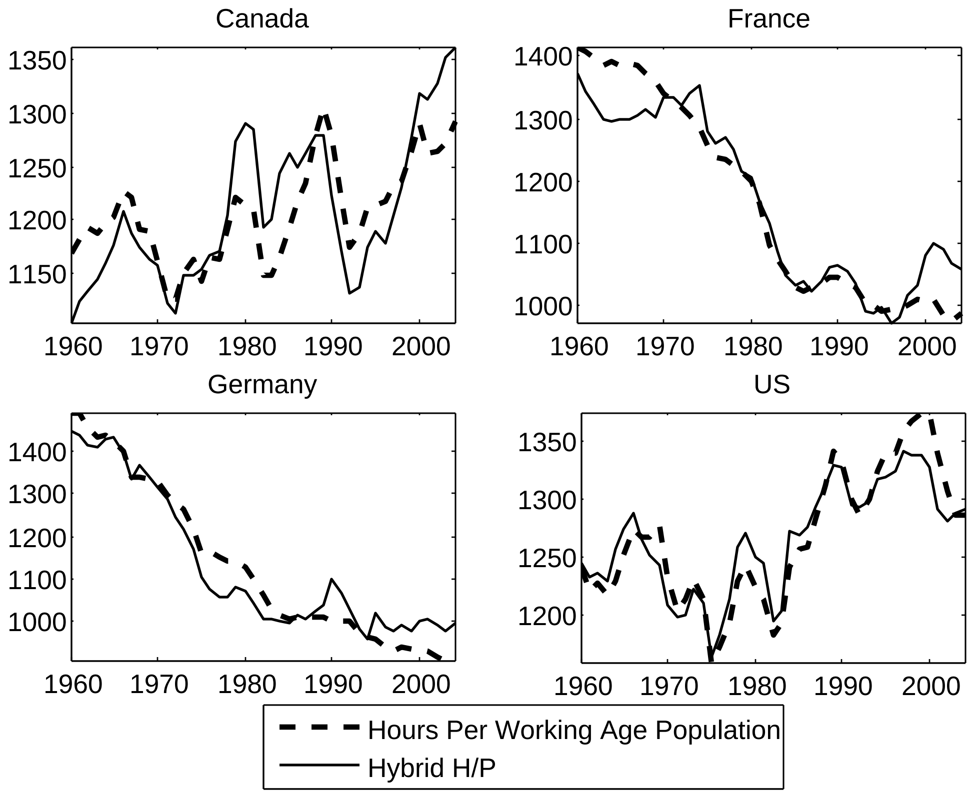


Figure 4.9: The Neoclassical Model with Taxes Wedge and the E/P ratio

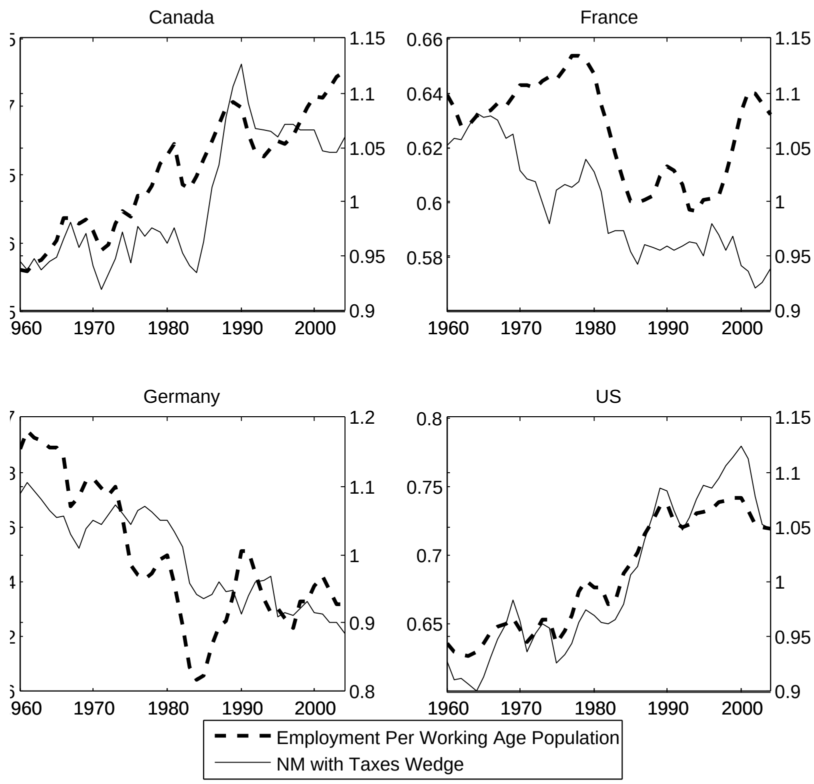


Figure 4.10: Actual Hours Per worker and Our Model's Predictions, Canada

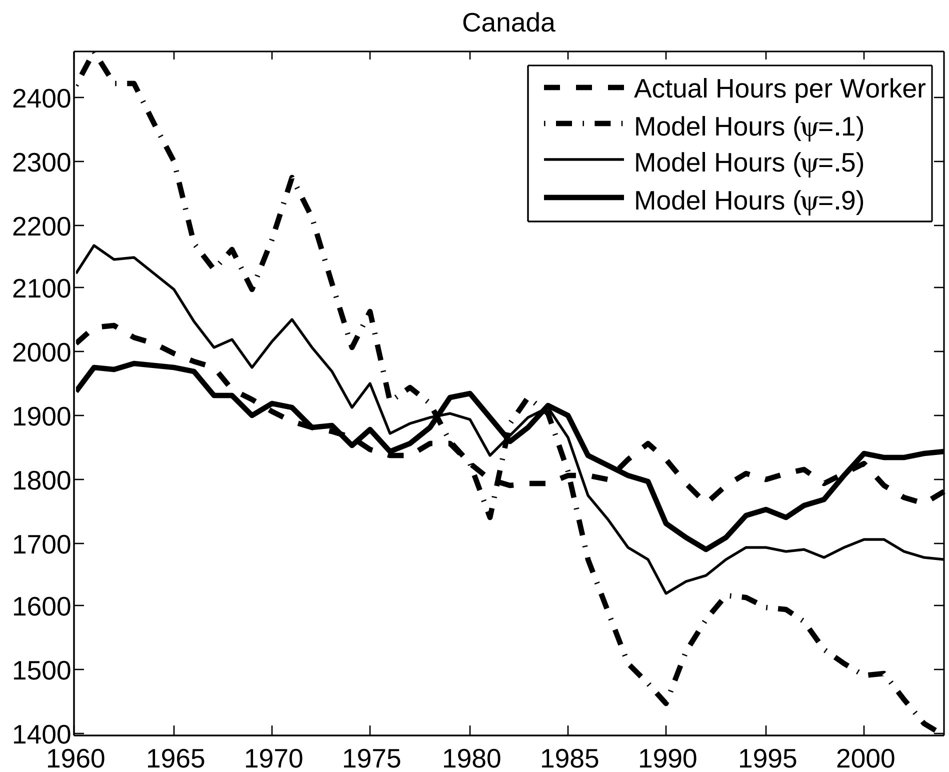


Figure 4.11: Actual Hours Per worker and Our Model's Predictions, France

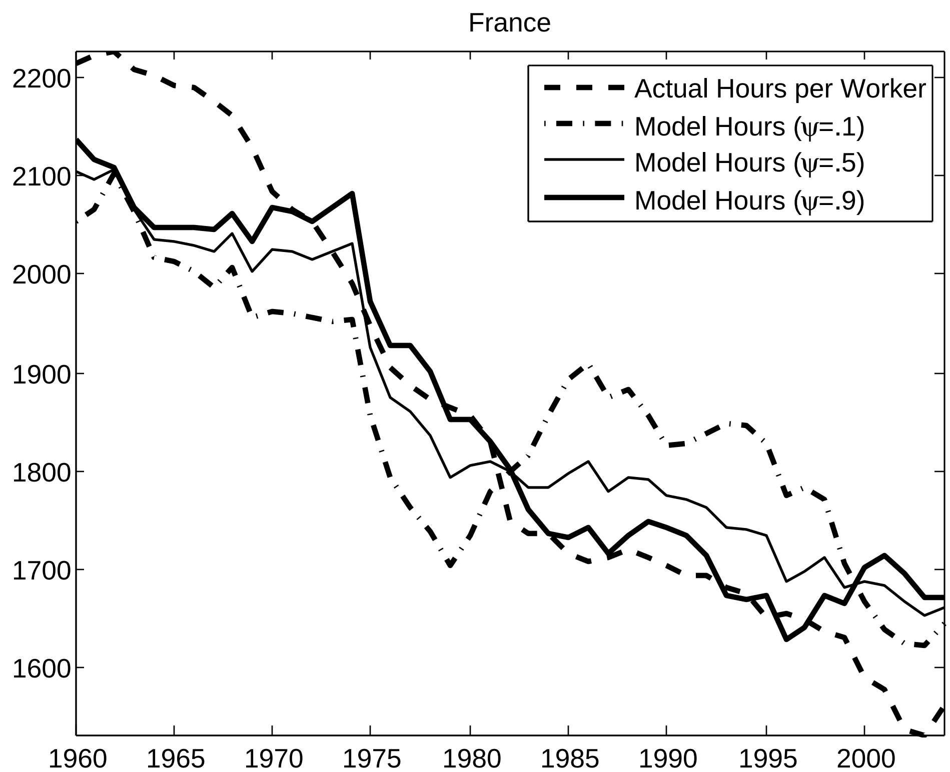


Figure 4.12: Actual Hours Per worker and Our Model's Predictions, Germany

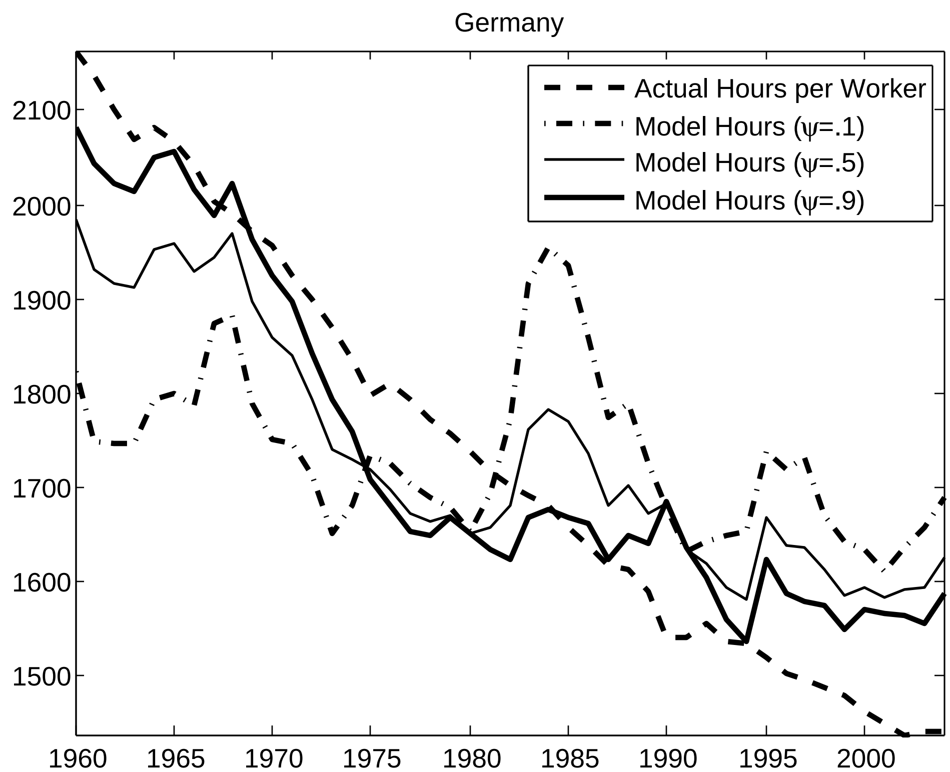


Figure 4.13: Actual Hours Per worker and Our Model's Predictions, US

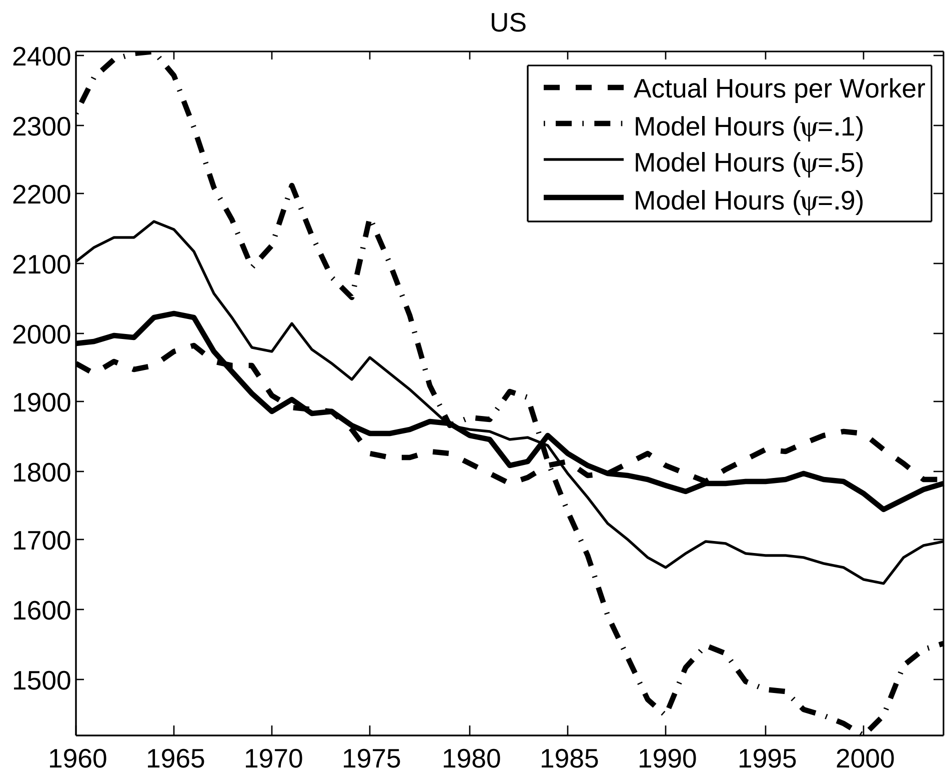


Table 4.1: Data Description of Consumption to Output Ratio, 1960-2004

	Mean	St Dev	Min	Max
Australia	0.56	0.02	0.52	0.59
Austria	0.58	0.01	0.56	0.61
Belgium	0.55	0.02	0.52	0.59
Canada	0.57	0.02	0.52	0.62
Finland	0.49	0.02	0.46	0.53
France	0.56	0.01	0.55	0.58
Germany	0.58	0.02	0.54	0.61
Italy	0.55	0.02	0.49	0.59
Japan	0.51	0.02	0.48	0.57
Netherlands	0.53	0.02	0.50	0.58
Spain	0.57	0.02	0.54	0.61
Sweden	0.55	0.02	0.50	0.59
Switzerland	0.57	0.01	0.54	0.60
UK	0.61	0.02	0.57	0.66
US	0.67	0.01	0.64	0.70

Table 4.2: Hours per worker 1960-2004: Actual Data vs. Model with Non-Employment (mean percent change)

		Model with Non-Employment				
Country	Actual	$\psi^E = 0.01$	$\psi^E = 0.25$	$\psi^E = 0.5$	$\psi^E = 0.75$	$\psi^E = 0.99$
Australia	-0.226	-0.571	-0.370	$\boxed{-0.238}$	-0.148	-0.085
Austria	-0.715	0.328	-0.067	-0.296	-0.441	$\boxed{-0.539}$
Belgium	-0.810	$\boxed{-0.546}$	-0.537	-0.529	-0.523	-0.518
Canada	-0.281	-1.190	-0.751	-0.438	$\boxed{-0.208}$	-0.038
Finland	-0.406	0.577	-0.045	$\boxed{-0.371}$	-0.569	-0.697
France	-0.795	-0.507	-0.527	-0.542	-0.553	$\boxed{-0.562}$
Germany	-0.921	-0.196	-0.368	-0.491	-0.579	$\boxed{-0.644}$
Italy	-0.771	-0.753	$\boxed{-0.774}$	-0.790	-0.803	-0.812
Japan	-0.496	$\boxed{-0.249}$	-0.229	-0.218	-0.211	-0.207
Netherlands	-0.961	-1.657	-1.189	$\boxed{-0.876}$	-0.657	-0.500
Spain	-0.355	$\boxed{-0.171}$	-0.168	-0.167	-0.165	-0.164
Sweden	-0.381	-0.616	-0.590	-0.575	-0.567	$\boxed{-0.561}$
Switzerland	-0.449	$\boxed{-0.495}$	-0.344	-0.285	-0.254	-0.236
UK	-0.634	-0.421	-0.422	-0.422	-0.422	$\boxed{-0.423}$
US	-0.204	-0.889	-0.613	-0.426	-0.296	$\boxed{-0.203}$
Mean	-0.560	-0.490	-0.466	-0.444	-0.427	-0.413

Table 4.3: 1960-2004 actual E/P and Model with Non-Employment ϕ_t mean percent change

Country	Actual E/P	Model with Non-Employment ϕ_t				
		$\psi^E = 0.01$	$\psi^E = 0.25$	$\psi^E = 0.5$	$\psi^E = 0.75$	$\psi^E = 0.99$
Australia	0.193	-0.630	-0.286	-0.282	-0.281	-0.280
Austria	-0.308	-1.720	-0.904	-0.894	-0.891	-0.889
Belgium	0.014	-1.784	-1.022	-1.013	-1.010	-1.008
Canada	0.508	-0.740	-0.356	-0.352	-0.351	-0.350
Finland	-0.413	-0.918	-0.513	-0.507	-0.505	-0.504
France	-0.026	-1.815	-1.004	-0.994	-0.990	-0.989
Germany	-0.195	-2.206	-1.164	-1.151	-1.147	-1.146
Italy	-0.028	-1.757	-0.974	-0.964	-0.961	-0.960
Japan	0.014	-1.176	-0.627	-0.620	-0.618	-0.617
Netherlands	0.478	-2.447	-1.214	-1.201	-1.197	-1.196
Spain	0.003	-0.896	-0.450	-0.444	-0.442	-0.441
Sweden	0.019	-0.986	-0.482	-0.476	-0.474	-0.474
Switzerland	0.062	-1.186	-0.568	-0.562	-0.560	-0.559
UK	-0.001	-1.511	-0.801	-0.793	-0.790	-0.789
US	0.283	-0.622	-0.259	-0.255	-0.253	-0.253
Mean	0.040	-1.360	-0.708	-0.700	-0.698	-0.697

Table 4.4: 1960-2004 actual and model mean percent changes

	Actual			NMT			Model with Non-Employment	
Country	H/P	H/E	E/P	H/P	Rel. H/P	Rel. H/E	H/E	Rel. H/E
Australia	-0.032	-0.226	0.193	-0.083	2.595	0.369	-0.238	1.055
Austria	-1.023	-0.715	-0.308	-0.542	0.530	0.759	-0.539	0.754
Belgium	-0.796	-0.810	0.014	-0.518	0.650	0.639	-0.546	0.674
Canada	0.227	-0.281	0.508	-0.031	-0.139	0.112	-0.208	0.741
Finland	-0.819	-0.406	-0.413	-0.702	0.857	1.730	-0.371	0.915
France	-0.821	-0.795	-0.026	-0.562	0.685	0.707	-0.562	0.707
Germany	-1.116	-0.921	-0.195	-0.647	0.579	0.702	-0.644	0.700
Italy	-0.799	-0.771	-0.028	-0.813	1.017	1.054	-0.774	1.004
Japan	-0.482	-0.496	0.014	-0.207	0.429	0.417	-0.249	0.502
Netherlands	-0.483	-0.961	0.478	-0.494	1.022	0.514	-0.876	0.912
Spain	-0.352	-0.355	0.003	-0.164	0.466	0.462	-0.171	0.480
Sweden	-0.362	-0.381	0.019	-0.560	1.547	1.471	-0.561	1.472
Switzerland	-0.388	-0.449	0.062	-0.235	0.607	0.524	-0.495	1.102
UK	-0.635	-0.634	-0.001	-0.423	0.666	0.666	-0.423	0.666
US	0.080	-0.204	0.283	-0.203	-2.552	0.997	-0.203	0.997
Mean	-0.520	-0.560	0.040	-0.412	0.597	0.741	-0.457	0.845

4.9 Appendix A: Data Sources and Summary

All of the data is yearly. Data on hours per worker H/E and employment E are from The Groningen Growth and Development Centre (www.ggdc.net). This data is used to back out total hours H. Data on working-age population P is taken from the Source OECD Database (www.sourceoecd.org), and data on consumption C and output Y are from the Penn World tables (http://pwt.econ.upenn.edu). The McDaniel (2007) tax data is available at (www.caramcdaniel.com), and is derived using similar methods as in Mendoza et al. (1994). All of the data we use is summarized in Tables (a) through (g) over the period 1960-2006 for the 15 OECD for which the McDaniel (2007) tax series is available.

Table a: H/P summary statistics 1960-2006

Country	Mean	St. Dev.	Min	Max
Australia	1226	36.57	1128	1283
Austria	1253	181.91	1019	1605
Belgium	1070	149.01	916	1389
Canada	1206	46.176	1125	1305
Finland	1339	178.01	1051	1675
France	1165	161.80	974	1416
Germany	1131	187.35	910	1489
Italy	1034	133.82	908	1420
Japan	1474	82.08	1327	1648
Netherlands	1077	108.94	944	1291
Spain	1076	169.83	841	1296
Sweden	1250	57.10	1177	1395
Switzerland	1409	105.51	1287	1616
UK	1263	134.42	1111	1517
US	1266	57.38	1159	1375

Table b: H/E summary statistics 1960-2006

Country	Mean	St. Dev.	Min	Max
Australia	1818	44.773	1751	1945
Austria	1762	178.25	1495	2073
Belgium	1802	194.15	1603	2289
Canada	1858	82.562	1760	2040
Finland	1868	115.31	1719	2074
France	1857	227.58	1532	2227
Germany	1736	224.57	1438	2163
Italy	1783	186.17	1590	2234
Japan	2054	142.50	1786	2224
Netherlands	1680	228.67	1399	2135
Spain	1954	134.85	1737	2137
Sweden	1648	116.76	1508	1900
Switzerland	1719	127.74	1551	1936
UK	1810	162.42	1614	2134
US	1854	62.76	1782	1981

Table c: E/P summary statistics 1960-2006

Country	Mean	St. Dev.	Min	Max
Australia	0.67	0.02	0.63	0.71
Austria	0.70	0.03	0.66	0.77
Belgium	0.59	0.02	0.54	0.62
Canada	0.65	0.04	0.57	0.72
Finland	0.71	0.05	0.59	0.81
France	0.62	0.02	0.59	0.65
Germany	0.64	0.03	0.60	0.69
Italy	0.57	0.02	0.55	0.63
Japan	0.71	0.02	0.68	0.75
Netherlands	0.64	0.05	0.59	0.76
Spain	0.54	0.06	0.45	0.61
Sweden	0.76	0.04	0.71	0.83
Switzerland	0.81	0.03	0.76	0.87
UK	0.69	0.02	0.64	0.72
US	0.68	0.04	0.62	0.74

Table d: Y/P summary statistics 1960-2006

Country	Mean	St. Dev.	Min	Max
Australia	32269	8148	19429	49697
Austria	32351	9816	14925	48353
Belgium	30970	9202	14802	47030
Canada	33540	7665	20811	49010
Finland	26785	8110	13961	42984
France	31117	8342	15438	44959
Germany	33398	8391	17778	47212
Italy	27679	8694	12497	41902
Japan	28541	10905	8613	43757
Netherlands	32723	8029	19486	47891
Spain	24291	7710	9235	38675
Sweden	31442	7177	18253	45455
Switzerland	42195	6839	27555	51676
UK	28964	8086	17305	45697
US	41187	10705	24672	61156

Table e: C/P summary statistics 1960–2006

Country	Mean	St. Dev.	Min	Max
Australia	17881	3952	11428	26628
Austria	18684	5565	8541	26976
Belgium	16869	4804	8737	24531
Canada	18906	3587	12848	26088
Finland	13129	3897	6790	20615
France	17589	4733	8651	25545
Germany	19695	5580	9595	28379
Italy	15476	5167	6156	23536
Japan	14585	5549	4904	22950
Netherlands	17378	3956	9832	24429
Spain	13798	4054	5672	21159
Sweden	16990	3216	10740	22841
Switzerland	23828	3764	15734	29051
UK	17851	5639	10592	29735
US	27739	7585	16325	42632

Table f: C/Y summary statistics 1960-2006

Country	Mean	St. Dev.	Min	Max
Australia	1.79	0.06	1.68	1.91
Austria	1.73	0.03	1.63	1.79
Belgium	1.83	0.05	1.69	1.91
Canada	1.76	0.07	1.58	1.91
Finland	2.03	0.07	1.87	2.17
France	1.77	0.02	1.71	1.82
Germany	1.71	0.07	1.63	1.85
Italy	1.80	0.08	1.70	2.04
Japan	1.94	0.07	1.75	2.07
Netherlands	1.87	0.07	1.72	1.98
Spain	1.74	0.05	1.62	1.84
Sweden	1.83	0.08	1.69	1.99
Switzerland	1.77	0.04	1.65	1.85
UK	1.63	0.06	1.52	1.75
US	1.49	0.03	1.42	1.56

Table g: $(1-\tau)$ summary statistics 1960–2006

Country	Mean	St. Dev.	Min	Max
Australia	0.77	0.04	0.73	0.84
Austria	0.54	0.06	0.47	0.67
Belgium	0.53	0.08	0.44	0.69
Canada	0.70	0.05	0.63	0.80
Finland	0.55	0.09	0.41	0.73
France	0.52	0.06	0.44	0.62
Germany	0.54	0.05	0.48	0.64
Italy	0.58	0.07	0.49	0.69
Japan	0.75	0.06	0.67	0.82
Netherlands	0.51	0.06	0.43	0.68
Spain	0.66	0.08	0.51	0.76
Sweden	0.48	0.08	0.37	0.68
Switzerland	0.79	0.04	0.73	0.86
UK	0.63	0.03	0.59	0.73
US	0.74	0.02	0.71	0.79

4.10 Appendix B: Derivations

Normalize all variables by P_t, and let $e_t = E_t/P_t$ and $c_t = C_t/P_t$. First, consider

$$\begin{aligned} F(e_t) &= \psi^E \frac{\gamma}{1+\varepsilon} h_t^{(1+\varepsilon)/\varepsilon} - \left(\psi^E - \psi^N\right)(1 - \ln(c_t)) - \psi^E \gamma \phi_t \\ \implies dF(e_t) &= \left(\psi^E \frac{\gamma}{\varepsilon} h_t^{1/\varepsilon}\right) dh_t + \left(\psi^E - \psi^N\right) d\ln(c_t) - \left(\psi^E \gamma\right) d\phi_t \\ &= \left(\psi^E \frac{\gamma}{\varepsilon} h_t^{1/\varepsilon}\right) h_t \frac{dh_t}{h_t} + \left(\psi^E - \psi^N\right) d\ln(c_t) - \left(\psi^E \gamma\right) d\phi_t \\ &= \left(\psi^E \frac{\gamma}{\varepsilon} h_t^{(1+\varepsilon)/\varepsilon}\right) d\log h_t + \left(\psi^E - \psi^N\right) d\ln(c_t) - \left(\psi^E \gamma \phi_t\right) d\log \phi_t \end{aligned}$$

Now, consider

$$\begin{aligned} F(e_t) &= \psi^N \log\left(\frac{\psi^N}{\psi^E e_t + \psi^N(1-e_t)}\right) - \psi^E \log\left(\frac{\psi^E}{\psi^E e_t + \psi^N(1-e_t)}\right) \\ &= \psi^N \log\left(\psi^N\right) - \psi^N \log\left(\psi^E e_t + \psi^N(1-e_t)\right) \\ &\qquad - \psi^E \log\left(\psi^E\right) + \psi^E \log\left(\psi^E e_t + \psi^N(1-e_t)\right) \\ \implies dF(e_t) &= \frac{-\psi^N\left(\psi^E - \psi^N\right) de_t}{\psi^E e_t + \psi^N(1-e_t)} + \frac{\psi^E\left(\psi^E - \psi^N\right) de_t}{\psi^E e_t + \psi^N(1-e_t)} \\ &= \left(\frac{\left(\psi^E - \psi^N\right)^2}{\psi^E e_t + \psi^N(1-e_t)}\right) de_t \\ &= \left(\frac{\left(\psi^E - \psi^N\right)^2 e_t}{\psi^E e_t + \psi^N(1-e_t)}\right) d\log(e_t) \end{aligned}$$

where the last line follows from applying $dx/x = d\log(x)$.

Combining the previous:

$$\begin{aligned} \left(\frac{\left(\psi^E - \psi^N\right)^2 e_t}{\psi^E e_t + \psi^N(1-e_t)}\right) d\log(e_t) &= \left(\psi^E \frac{\gamma}{\varepsilon} h_t^{(1+\varepsilon)/\varepsilon}\right) d\log h_t \\ &\quad + \left(\psi^E - \psi^N\right) d\ln(c_t) - \left(\psi^E \gamma\right) d\phi_t \\ \implies d\log(e_t) &= \left(\frac{\psi^E e_t + \psi^N(1-e_t)}{\left(1 - \psi^N\right)^2 e_t}\right)\left(\psi^E \frac{\gamma}{\varepsilon} h_t^{(1+\varepsilon)/\varepsilon}\right) d\log h_t \\ + \left(\frac{\psi^E e_t + \psi^N(1-e_t)}{\left(\psi^E - \psi^N\right)^2 e_t}\right)\left(\psi^E - \psi^N\right) d\ln(c_t) &- \left(\frac{\psi^E e_t + \psi^N(1-e_t)}{\left(\psi^E - \psi^N\right)^2 e_t}\right)\left(\psi^E \gamma \phi_t\right) d\log \phi_t, \end{aligned}$$

where the last lines follow from multiplying both sides by $\left(\frac{\left(\psi^E-\psi^N\right)^2 e_t}{\psi^E e_t+\psi^N(1-e_t)}\right)$.

Now, consider

$$\begin{aligned}
h_t &= \left(\frac{\psi^E e_t + \psi^N\left(1-e_t\right)}{\psi^E e_t}\right)\left(\frac{\left(1-\tau_t\right)\left(1-\alpha\right)}{\gamma}\frac{Y_t}{C_t}\right)^{\varepsilon/(1+\varepsilon)} \\
\implies \log\left(h_t\right) &= \log\left(\psi^E e_t + \psi^N\left(1-e_t\right)\right) - \log\left(\psi^E\right) - \log\left(e_t\right) \\
&\quad + \frac{\varepsilon}{1+\varepsilon}\log\left(1-\tau_t\right) + \frac{\varepsilon}{1+\varepsilon}\log\left(1-\alpha\right) \\
&\quad - \frac{\varepsilon}{1+\varepsilon}\log\left(\gamma\right) + \frac{\varepsilon}{1+\varepsilon}\log\left(Y_t/C_t\right) \\
\implies d\log\left(h_t\right) &= \left(\frac{\psi^E-\psi^N}{\psi^E e_t + \psi^N\left(1-e_t\right)}\right) de_t \\
&\quad - d\log\left(e_t\right) + \frac{\varepsilon}{1+\varepsilon}d\log\left(1-\tau_t\right) + \frac{\varepsilon}{1+\varepsilon}d\log\left(Y_t/C_t\right) \\
&= \left(\frac{\left(\psi^E-\psi^N\right)e_t}{\psi^E e_t + \psi^N\left(1-e_t\right)}\right) d\log\left(e_t\right) \\
&\quad - d\log\left(e_t\right) + \frac{\varepsilon}{1+\varepsilon}d\log\left(1-\tau_t\right) + \frac{\varepsilon}{1+\varepsilon}d\log\left(Y_t/C_t\right) \\
&= \left(\frac{\left(\psi^E-\psi^N\right)e_t}{\psi^E e_t + \psi^N\left(1-e_t\right)} - 1\right) d\log\left(e_t\right) \\
&\quad - d\log\left(e_t\right) + \frac{\varepsilon}{1+\varepsilon}d\log\left(1-\tau_t\right) + \frac{\varepsilon}{1+\varepsilon}d\log\left(Y_t/C_t\right) \\
&= \left(\frac{-\psi^N}{\psi^E e_t + \psi^N\left(1-e_t\right)}\right) d\log\left(e_t\right) \\
&\quad + \frac{\varepsilon}{1+\varepsilon}d\log\left(1-\tau_t\right) + \frac{\varepsilon}{1+\varepsilon}d\log\left(Y_t/C_t\right)
\end{aligned}$$

Therefore, the relevant two equations in the two "unknowns" $d\log\left(e_t\right)$ and $d\log\left(h_t\right)$ are

$$\begin{aligned}
d\log\left(e_t\right) &= \left(\frac{\psi^E e_t + \psi^N\left(1-e_t\right)}{\left(\psi^E-\psi^N\right)^2 e_t}\right)\left(\psi^E\frac{\gamma}{\varepsilon}h_t^{(1+\varepsilon)/\varepsilon}\right) d\log h_t \\
&\quad + \left(\frac{\psi^E e_t + \psi^N\left(1-e_t\right)}{\left(\psi^E-\psi^N\right)^2 e_t}\right)\left(\psi^E-\psi^N\right) d\ln\left(c_t\right) \\
&\quad - \left(\frac{\psi^E e_t + \psi^N\left(1-e_t\right)}{\left(\psi^E-\psi^N\right)^2 e_t}\right)\left(\psi^E\gamma\phi_t\right) d\log\phi_t
\end{aligned} \tag{4.18}$$

and

$$
\begin{aligned}
d\log(h_t) = {} & \left(\frac{-\psi^N}{\psi^E e_t + \psi^N (1-e_t)}\right) d\log(e_t) \\
& + \frac{\varepsilon}{1+\varepsilon} d\log(1-\tau_t) + \frac{\varepsilon}{1+\varepsilon} d\log(Y_t/C_t).
\end{aligned} \tag{4.19}
$$

Insert (4.19) in (4.18):

$$
\begin{aligned}
d\log(e_t) = {} & \left(\frac{\psi^E e_t + \psi^N (1-e_t)}{(\psi^E - \psi^N)^2 e_t}\right) \left(\psi^E - \psi^N\right) d\ln(c_t) \\
& - \left(\frac{\psi^E e_t + \psi^N (1-e_t)}{(\psi^E - \psi^N)^2 e_t}\right) \left(\psi^E \gamma \phi_t\right) d\log\phi_t \\
& - \left(\frac{\psi^E e_t + \psi^N (1-e_t)}{(\psi^E - \psi^N)^2 e_t}\right) \left(\psi^E \frac{\gamma}{\varepsilon} h_t^{(1+\varepsilon)/\varepsilon}\right) \left(\frac{\psi^N}{\psi^E e_t + \psi^N (1-e_t)}\right) d\log(e_t) \\
& + \left(\frac{\psi^E e_t + \psi^N (1-e_t)}{(\psi^E - \psi^N)^2 e_t}\right) \left(\psi^E \frac{\gamma}{\varepsilon} h_t^{(1+\varepsilon)/\varepsilon}\right) \frac{\varepsilon}{1+\varepsilon} \left(d\log(1-\tau_t) + d\log(Y_t/C_t)\right).
\end{aligned}
$$

After rearranging and simplifying this yields

$$
\begin{aligned}
\left(1 + \left(\frac{\psi^N}{(\psi^E - \psi^N)^2 e_t}\right) \left(\psi^E \frac{\gamma}{\varepsilon} h_t^{(1+\varepsilon)/\varepsilon}\right)\right) d\log(e_t) = {} & \left(\frac{\psi^E e_t + \psi^N (1-e_t)}{(\psi^E - \psi^N) e_t}\right) d\ln(c_t) \\
& - \left(\frac{\psi^E e_t + \psi^N (1-e_t)}{(\psi^E - \psi^N)^2 e_t} \psi^E \gamma \phi_t\right) d\log(\phi_t) \\
& + \left(\frac{\psi^E e_t + \psi^N (1-e_t)}{(\psi^E - \psi^N)^2 e_t}\right) \left(\psi^E \frac{\gamma}{\varepsilon} h_t^{(1+\varepsilon)/\varepsilon}\right) \frac{\varepsilon}{1+\varepsilon} \left(d\log(1-\tau_t) + d\log(Y_t/C_t)\right).
\end{aligned}
$$

Now, insert (4.18) in (4.19):

$$
\begin{aligned}
d\log(h_t) = {} & -\left(\frac{\psi^N}{\psi^E e_t + \psi^N (1-e_t)}\right) \left(\frac{\psi^E e_t + \psi^N (1-e_t)}{(\psi^E - \psi^N)^2 e_t}\right) \left(\psi^E \frac{\gamma}{\varepsilon} h_t^{(1+\varepsilon)/\varepsilon}\right) d\log h_t \\
& - \left(\frac{\psi^N}{\psi^E e_t + \psi^N (1-e_t)}\right) \left(\frac{\psi^E e_t + \psi^N (1-e_t)}{(\psi^E - \psi^N)^2 e_t}\right) \left(\psi^E - \psi^N\right) d\ln(c_t) \\
& + \left(\frac{\psi^N}{\psi^E e_t + \psi^N (1-e_t)}\right) \left(\frac{\psi^E e_t + \psi^N (1-e_t)}{(\psi^E - \psi^N)^2 e_t}\right) \left(\psi^E \gamma \phi_t\right) d\log\phi_t \\
& + \frac{\varepsilon}{1+\varepsilon} d\log(1-\tau_t) + \frac{\varepsilon}{1+\varepsilon} d\log(Y_t/C_t).
\end{aligned}
$$

After simplifying and rearranging, this yields

$$\left(1+\frac{\psi^N\psi^E}{\left(\psi^E-\psi^N\right)^2 e_t}\frac{\gamma}{\varepsilon}h_t^{(1+\varepsilon)/\varepsilon}\right)d\log\left(h_t\right)=-\left(\frac{\psi^N}{\left(\psi^E-\psi^N\right)e_t}\right)d\ln\left(c_t\right)$$
$$+\left(\frac{\psi^N}{\left(\psi^E-\psi^N\right)^2 e_t}\right)\left(\psi^E\gamma\phi_t\right)d\log\phi_t+\frac{\varepsilon}{1+\varepsilon}d\log\left(1-\tau_t\right)+\frac{\varepsilon}{1+\varepsilon}d\log\left(Y_t/C_t\right).$$

CHAPTER V

Conclusion

This dissertation provided an empirical investigation into how tax policy impacts behavior. One of my key findings was that unintended incentives from tax policy can impact an individual's behavior, while the intended incentives can fail to generate a response. This is illustrated in the first essay, which analyzed a tax credit given to low and middle income households for contributing to a retirement savings plan. The policy creates a notch, or discontinuous jump, within a household's budget constraint that creates an incentive to misreport income. I found that individuals responded to the incentive to misreport income, yet they failed to increase retirement contributions, the intended goal of the program.

A second finding of this dissertation is that taxation can have an impact on market wage rates. Specifically, past empirical studies often assume that pre-tax wage rates are unchanged when changes are made to marginal tax rates. The second essay, co-written with Matthew Rutledge, found the contrary to be true. We tested whether marginal tax rates are related to pre-tax wage rates using evidence drawn from changes made during the Tax Reform Act of 1986. We found that the net-of-tax rate is negatively associated with pre-tax wages, which implies that marginal tax rates indeed have an impact on pre-tax wage rates.

A third finding of this dissertation is that tax policy impacts the overall economy through margins that the standard neoclassical model is unable to capture. In the the third essay, co-written with Brendan Epstein, we showed that the standard neoclassical model with taxes is a better predictor of hours per worker due to its inability to accurately generate hours changes on the extensive margin. We then developed a model that allows both hours per worker and employment per population to vary. We found that the impact taxes have on individuals' decisions to work can be better understood at an aggregate level when the hours decision is separated into an intensive and an extensive margin.

Overall, this dissertation offered empirical evidence to foster a better understanding of the specific impacts of taxation analyzed using both a micro- and macroeconomic framework. Each essay tested an aspect of economic theory as it applies to problems in public finance. Throughout this dissertation I contributed empirical analyses that provided insight into areas where previous empirical work had been sparse. These findings introduced evidence to help shed light on the role tax policy plays in shaping decisions at an individual, a national, and an international level.

BIBLIOGRAPHY

Alesina, Alberto, Edward Glaeser, and Bruce Sacerdote, "Work and Leisure in the United States and Europe: Why So Different?," *NBER Macroeconomics Annual*, 2005, *20*, 1–64.

Auerbach, Alan J. and Joel Slemrod, "The Economic Effects of the Tax Reform Act of 1986," *Journal of Economic Literature*, 1997, *35* (2), 589–632.

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.

Bernheim, Douglas B., "Taxation and Savings," in Alan J. Auerbach and Martin Feldstein, eds., *Handbook of Public Economics*, Vol. 3, North Holland, Amsterdam, 2003, chapter 18, pp. 1173–1249.

Bingley, Paul and Gauthier Lanot, "The incidence of income tax on wages and labor supply," *Journal of Public Economics*, 2002, *83*, 173–194.

Blomquist, Soren and Hakan Selin, "Hourly Wage Rate and Taxable Labor Income Responsiveness to Changes in Marginal Tax Rates," January 2009.

Blundell, R., A. Duncan, and C. Meghir, "Estimating labor supply responses using tax reforms," *Econometrica*, 1998, pp. 827–861.

Borjas, George J., "The Relationship between Wages and Weekly Hours of Work: The Role of Division Bias," *The Journal of Human Resources*, 1980, *15* (3), 409–423.

Bosworth, Barry and Gary Burtless, "Effects of Tax Reform on Labor Supply, Investment, and Saving," *The Journal of Economic Perspectives*, 1992, *6* (1), 3–25.

Browning, Edgar K and William R Johnson, *The Distribution of the Tax Burden*, Washington, 1979.

Cameron, A.C. and P.K. Trivedi, *Microeconometrics: methods and applications*, Cambridge Univ Pr, 2005.

Chetty, Raj, "Bounds on Elasticities with Optimization Frictions: An Application to Taxation and Labor Supply," *NBER Working Paper 15616*, 2009.

——, "Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance," *American Economic Journal: Economic Policy*, August 2009, *1* (2), 31–52.

—— **and Emmanuel Saez**, "Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients," April 2009.

——**, John N. Friedman, Tore Olsen, and Luigi Pistaferri**, "The Effect of Adjustment Costs and Institutional Constraints on Labor Supply Elasiticities: Evidence from Denmark," July 2009.

Czajka, J., "Subannual Income Estimation?," in M. David, ed., *Technical, Conceptual and Administrative Lessons of the Income Survey Development Program (ISDP)*, Social Science Research Council, 1983, pp. 87–97.

Davies, James, France St-Hilaire, and John Whalley, "Some Calculations of Lifetime Tax Incidence," *The American Economic Review*, Sep 1984, *74* (4), 633–649.

Duflo, Esther, William Gale, Jeffrey Liebman, Peter Orzag, and Emmanuel Saez, "Savings Incentives for Low- and Middle-Income Families: Evidence from a Field Experiment with H&R Block," *The Quarterly Journal of Economics*, November 2006, *121* (4), 1311–1346.

___ , ___ , ___ , ___ , and ___ , "Savings Incentives for Low- and Moderate- Income Families in the United States: Why is the Saver's Credit Not More Effective?," *Journal of the European Economic Association*, April-May 2007, *5* (2-3), 647–661.

Eissa, N., "Taxation and labor supply of married women: the Tax Reform Act of 1986 as a natural experiment," *NBER working paper*, 1995.

___ and J.B. Liebman, "Labor supply response to the earned income tax credit," *The Quarterly Journal of Economics*, 1996, pp. 605–637.

Engelhardt, Gary V. and Anil Kumar, "Employer matching and 401(k) saving: Evidence from the health and retirement survey," *Journal of Public Economics*, 2007, *91*, 1920–1943.

Fan, J and I Gijbels, *Local Polynomial Modelling and its Applications*, Chapman & Hall, New York, 1996.

Feldstein, Martin, "Tax Avoidance and the Deadweight Loss of the Income Tax," *The Review of Economics and Statistics*, November 1999, *81* (4), 674–680.

Fullerton, D. and G.E. Metcalf, "Tax Incidence," *Handbook of Public Economics*, 2002, *4*, 1788–1872.

Gale, William G., J. Mark Iwry, and Peter R. Orzag, "The Saver's Credit: Expanding Retirement Savings for Middle- and Lower-Income Americans," The Retirement Security Project March 2005.

Ham, John C., Xianghong Li, and Lara Shore-Sheppard, "Seam Bias, Multiple-State, Multiple-Spell Duration Models and the Employment Dynamics of Disadvantaged Women," *NBER Working Paper 15151*, 2009.

Hausman, Jerry A. and James M. Poterba, "Household Behavior and the Tax Reform Act of 1986," *The Journal of Economic Perspectives*, 1987, *1* (1), 101–119.

Hubbard, Glenn and Jonathan S. Skinner, "Assessing the Effectiveness of Saving Incentives," *The Journal of Economic Perspectives*, Autumn 1996, *10* (4), 73–90.

Imbens, Guido W. and Thomas Lemieux, "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, 2008, *142*, 615–625.

Killingsworth, Mark R. and James J. Heckman, "Female Labor Supply: A Survey," in O. Ashenfelter and R. Layard, eds., *Handbook of Labor Economics*, 1986.

Kimball, M.S. and M.D. Shapiro, "Labor Supply: Are the Income and Substitution Effects Both Large or Both Small?," *NBER Working Paper*, 2008.

Koenig, Gary and Robert Harvey, "Utilization of the Saver's Credit: An Analysis of the First Year," *National Tax Journal*, December 2005, *LVIII* (4), 787–806.

Kubik, Jeffrey D., "The incidence of personal income taxation: evidence from the tax reform act of 1986," *Journal of Public Economics*, 2004, *88*, 1567–1588.

Lee, David S., "Trimming for Bounds on Treatment Effects with Missing Outcomes," June 2002. NBER Technical Working Paper No. 277.

Leibbrandt, Murray, James Levinson, and Justin McCrary, "Incomes in South Africa Since the Fall of Apartheid," May 2005. NBER Working Paper No. 11384.

McCrary, Justin, "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 2008, *142*, 698–714.

McDaniel, C., "Hours worked in the OECD 1960-2003: Driving forces and propagation mechanisms," 2007.

Mendoza, Enrique G., Assaf Razin, and Linda L. Tesar, "Effective tax rates in macroeconomics: Cross-country estimates of tax rates on factor incomes and consumption," *Journal of Monetary Economics*, 1994, *34* (3), 297–323.

Mills, Gregory, William G. Gale, Rhiannon Patterson, and Emily Apostolov, "What Do Individual Accounts Do? Evidence from a Controlled Experiment," 2006.

Moffitt, Robert A. and Mark O. Wilhelm, "Taxation and the Labor Supply Decisions of the Affluent," in Joel B. Slemrod, ed., *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*, Russell Sage Foundation, 2002.

Mulligan, Casey B., "A Century of Labor-Leisure Distortions," *NBER Working Paper 8774*, 2002.

Ohanian, L., A. Raffo, and R. Rogerson, "Long-term changes in labor supply and taxes: Evidence from OECD countries, 1956–2004," *Journal of Monetary Economics*, 2008, *55* (8), 1353–1362.

Parkin, Michael, "A Method for Determining Whether Parameters in Aggregative Models Are Structural," *Carnegie-Rochester Conference Series on Public Policy*, 1988, *29*, 215–252.

Pechman, Joseph A., *Who paid the taxes, 1966-1985*, Washington D.C., 1985.

___ **and Benjamin A. Okner**, *Who Bears the Tax Burden?*, Washington, 1974.

Pencavel, J., "Labor Supply of Men: A Survey," in O. Ashenfelter and R. Layard, eds., *Handbook of Labor Economics*, 1986.

Poterba, James M., Steven F. Venti, and David A. Wise, "How Retirement Saving Programs Increase Saving," *The Journal of Economic Perspectives*, Autumn 1996, *10* (4), 91–112.

Prescott, E.C., "Why Do Americans Work So Much More Than Europeans?," *Federal Reserve Bank of Minneapolis Quarterly Review*, 2004, *28* (1), 2–13.

Ragan, Kelly S., "Taxes, Transfers, and Time Use: Fiscal Policy in a Household Production Model." PhD dissertation, University of Chicago 2006.

Rogerson, Richard, "Taxation and Market Work: Is Scandinavia an Outlier?," *Economic Theory*, 2007, *32* (1), 59–85.

___ **and Johanna Wallenius**, "Micro and Macro Elasticities in a Life Cycle Model with Taxes," *Journal of Economic Theory*, 2009, *144* (6), 2277–2292.

Rotemberg, Julio J. and Michael Woodford, "The Cyclical Behavior of Prices and Costs," in John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics*, Vol. 1B, Cambridge, MA: MIT Press, 1999, p. 10511135.

Saez, Emmanuael, "Do Taxpayers Bunch at Kink Points?," *American Economic Journal: Economic Policy*, 2009, *Forthcoming.*

Saez, Emmanuel, Joel Slemrod, and Seth H Giertz, "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review," *NBER Working Paper 15012*, 2009.

Sallee, James M., "The Incidence of Tax Credits for Hybrid Vehicles," 2009.

Schreiner, Mark, Michael Sherraden, Margaret Clancy, Lissa Johnson, Jami Curly, Michal Grinsten-Weiss, Min Zhan, and Sondra Beverly, "Savings and Asset Accumulation in Individual Development Accounts," 2001.

Shimer, R., "Convergence in macroeconomics: the labor wedge," *American Economic Journal: Macroeconomics*, 2009, *1* (1), 280–297.

Slemrod, Joel, "A General Model of the Behavioral Response to Taxation," *International Tax and Public Fiance*, 2001, *8*, 119–128.

Stinson, M., "Estimating measurement error in sipp annual job earnings: A comparison of census survey and ssa administrative data," *US Census Bureau Technical Report TP-2002-24, Suitland, MD (September)*, 2002.

Weber, Caroline, "Obtaining a Consistent Estimate of the Elasticity of Taxable Income Using Difference-in-Differences." PhD dissertation, University of Michigan 2010.