Effects of Predictable Tax Liability Variation on Household Labor Income*

Naomi E. Feldman† and Peter Katuščák‡

April 2, 2009

Abstract

Economic theory assumes that taxpayers use their true marginal tax rate (MTR) to guide their economic decisions. However, complexity of the personal income tax system implies that taxpayers may incorrectly perceive true marginal prices and incentives. We first develop an updating model that formalizes this conjecture. A prediction of the model is that an unexpected increase in the previous year's tax liability pushes up the perception of the MTR in the current year, even though the MTR is not in fact changing. Then, assuming that taxpayers react to their *perceived* after-tax price as economic theory would suggest, we test this prediction empirically by examining whether household labor income responds to *predictable* (but not necessarily predicted) *lump-sum* variation in the previous year's tax liability due to loss of eligibility for the Child Tax Credit when the eligible child turns 17 in the given tax year. Using identification strategy based on eligibility discontinuity, we find that losing the credit *reduces* parental labor income in the year following the loss of the credit. This result is robust to a variety of tests and different data sources. This finding is inconsistent with the taxpayers being fully rational and fully informed, which suggests imperfect *ex-post* understanding of changes in the tax schedule.

JEL Classification: H21, H24, H31.

*We would like to thank Kate Antonovics, Eli Berman, Mike Christian, Julie Cullen, John Diamond, Libor Dušek, Roger Gordon, Štěpán Jurajda, Julie Lee, Jeff Liebman, Marco Manacorda, Karen Pence, Dan Shaviro, Joel Slemrod, seminar participants at UCSD, Hebrew University of Jerusalem, CEU and conference participants at the 2005 NTA and the 2006 AEA meetings for helpful comments and suggestions. We also thank Petr Chovanec for research assistance. Any remaining errors are our own.

†Department of Economics, Ben Gurion University, Be'er-Sheva, 84105, Israel, nfeldman@bgu.ac.il.

‡CERGE-EI, P.O.Box 882, Politických vězňů 7, 111 21 Praha 1, Czech Republic, Peter.Katuscak@cerge-ei.cz.

1 Introduction

Economic theory presumes that individuals respond to marginal prices when deciding on their labor supply, portfolio allocation, saving decisions, and many other behavioral margins. Because marginal prices are affected by marginal tax rates (MTRs), the latter have been recognized as important for behavioral responses. Indeed, there is now a voluminous empirical literature identifying significant behavioral responses to tax changes.[1]

Under the usual interpretation, these responses are attributed to changes in MTRs. This interpretation assumes, however, that taxpayers correctly perceive their MTRs and, as a result, marginal net-of-tax prices. The existing empirical evidence on such assumption is mixed, however.[2] One plausible reason why households may not have perfect information about their tax-induced incentives is due to complexity of the income tax. Indeed, the U.S. federal income tax code is filled with various deductions, credits, and exemptions, and knowing when one is or is not eligible for them requires a detailed knowledge of diverse and often arbitrary eligibility rules, phase-in and phase-out ranges, and possibly other details. For example, in their 2005 report, the President's Advisory Panel on Federal Tax Reform[3] laments:

> "There is no clearer proof of the complexity of the tax code than the collective anxiety felt by Americans every April as the tax filing deadline approaches. For many, filing taxes consists first of procrastination. Then there is the inevitable search for slips of paper containing once-meaningful but now unintelligible financial transactions. Then comes the maze of lengthy instructions complex enough that even highly schooled professionals have to reread the directions several times. Those directions send taxpayers on a search through baffling schedules and detailed worksheets requiring many illogical and counterintuitive computations. And in the end, most taxpayers give up, and visit a tax preparer who promises to make sense of the whole process - for a price."
>
> "To determine something as basic as figuring out the tax implications of having a child, you need to review numerous rules and complete many separate sets of computations. Figuring out whether you can claim the child tax credit, for example, requires

[1]See, for example, Eissa (1995) or Eissa and Liebman (1996) for labor force participation of women, Looney and Singhal (2004) for the intertemporal elasticity of labor earnings, Goolsbee (2000) for the timing of income realization, Poterba and Samwick (2003) for risk-taking and portfolio behavior, and Feldstein (1995), Auten and Carroll (1999), Gruber and Saez (2002), and Kopczuk (2005) for reported and taxable income. On the other hand, Saez (2004) finds that only the top 1% incomes show evidence of behavioral responses to taxation.

[2]See Section 2.

[3]In "Simple, Fair, and Pro-Growth: Proposals to Fix America's Tax System", a report of the President's Advisory Panel on Federal Tax Reform, November 2005. Source: www.taxreformpanel.gov.

the skills of a professional sleuth: You need to complete eight lines on a tax form, perform up to five calculations, and fill out as many as three other forms or schedules. Further research, reading, and computation may be needed to determine whether you can claim head of household filing status, an exemption for a dependent, the child and dependent care credit, the earned income tax credit, or tax credits related to your child's education, to name only some of the possibilities."

Similar warnings appear in academic work as well. For example, Kotlikoff and Rapson (2007) argue that "thanks to the incredible complexity of the U.S. fiscal system, it's impossible for anyone to understand her incentive to work, save, or contribute to retirement accounts absent highly advanced computer technology and software." However, experts are not the only ones who complain. According to a 2003 NPR/Kaiser Family Foundation/Kennedy School of Government Taxes Survey, 36 percent of respondents are more bothered by complexity of the federal income tax system than by the amount they pay in taxes or the feeling that rich people do not pay their "fair" share.[4] In addition, 90 percent of the respondents find the tax system very or at least somewhat complicated.[5] When asked what factors contribute to this complexity, the respondents named factors such as "too much record-keeping" (62 percent), "too many different tax rates" (59 percent), or "forms being too hard to fill" (56 percent). However, all of these percentages are overwhelmed by 96 percent of the respondents thinking that complexity is partially due to "so many different kinds of deductions and tax credits, and so many rules about how to take them." Moreover, 64 percent consider the latter to be the most important source of complexity.

Figure 1 gives a flavor of this complexity. It plots the effective and statutory federal MTR excluding payroll taxes for married couples filing jointly in 2002 as a function of household labor income, assuming no other income. It deliberately focuses on the income range up to $40,000, in which the actual effective MTR is highly non-monotone and quite variable. This is in contrast to the statutory tax schedule under which the MTR is an increasing step function of income with just a few brackets. In addition, both the effective and the statutory MTR schedules vary with the number of dependents claiming the personal exemption. On top of that, fixing the number of dependents, the effective MTR schedule, but not the statutory one, also varies with the number of dependents eligible for the CTC.

In response to this complexity, taxpayers are increasingly looking to experts or computer

[4]This finding is based on the following question: "Which of the following bothers you most about taxes: the large amount you pay in taxes, the complexity of the tax system, or the feeling that some wealthy people get away not paying their fair share?"

[5]This finding is based on the following question: "How complex do you think the current federal income tax system is? Do you think it is very complex, somewhat complex, not too complex or not complex at all?"

software for help.[6] To the extent that a preparer or software is used only as a tax compliance tool or an ex post minimizer of tax liability, it is not clear that the use of these tools leads to better informed taxpayers. On the contrary, tax preparers and software allow taxpayers to escape the complexity of the tax code to a large degree, which is likely to further reduce taxpayer knowledge of the tax system. Put differently, by going through their tax forms and instructions the old-fashioned way, line by line, taxpayers who use the traditional method of tax filing may actually be better informed about details of the tax system.

There are two fundamental reasons why this complexity may result in taxpayers having imperfect knowledge about the tax system with which they interact. First, complexity makes it costly for taxpayers in terms of cognitive abilities, time, or money to learn about the details. It is therefore plausible that many taxpayers are not aware of some or most tax law provisions that currently affect them, or that will affect them in the future. Note that this argument does not rely on bounded rationality. It simply stresses the fact that gathering and processing information is costly, so even fully rational economic agents may prefer to have less than perfect knowledge of the tax schedule. They then use any information they get from the interaction with the tax system, as well as any other signals, to update their beliefs. Following Feige and Pearce (1976) and Buiter (1980), this is referred to as *economically*, as opposed to *technically*, optimal belief formation. Second, taxpayers may be boundedly rational. In this case, a certain framing of tax changes makes them more salient, resulting in taxpayers being more responsive, and vice versa.

There are two degrees to which taxpayers may have misperceptions about the tax system they interact with. First, they may experience *ex ante misperceptions*. This is the case when taxpayers fail to predict a future change in the tax schedule that they face even though such change is predictable. For example, at the cross-sectional level, a household loses eligibility for the Child Tax Credit when their child turns 17. Alternatively, at the aggregate level, some tax cuts come with sunset provisions, such as the Economic Growth and Tax Relief Reconciliation Act of 2001. Such changes in the tax schedule are predictable, but may not necessarily be predicted by taxpayers ex ante. When experiencing ex ante misperceptions, households may fail to intertemporarily optimize their behavior and they may exhibit "excess behavioral sensitivity" to predictable changes in the tax schedule. This argument is closely related to empirical research

[6]For example, of about 130 million of individual tax returns filed for the tax year 2001, 72.5 million, or 56 percent, were prepared by a professional preparer. By 2003, this number jumped to about 79 million, or 61 percent. Many taxpayers are also turning to tax preparation software, which, beyond simplification, also brings about the benefits of electronic filing (if chosen) and a faster refund. For example, about 47 million, or 36 percent, of returns were e-filed in 2001. By 2003, this number jumped to 61 million, or 47 percent. This amount is split roughly equally between e-filing by individual taxpayers and tax preparers. As a result, about 84 percent of taxpayers relied either on a preparer or a software in 2003. Source: Statistics of Income Division of the Internal Revenue Service.

on the permanent income hypothesis (PIH), especially the finding that consumption tends to increase after predictable income increases despite the absence of credit constraints[7].

Second, taxpayers may experience *ex post misperceptions.* This is the case when taxpayers misperceive or misinterpret a change in the tax schedule they have already experienced. For example, suppose that a household receives an unexpected tax rebate check from the government or it realizes an unexpected increase in its after-tax income. With imperfect understanding of the source of the surprise, the household may interpret the surprise in multiple ways. First, the after-tax income increase could reflect a lump-sum tax decrease, as would be the case if a household gains an eligible child for the Child Tax Credit, or a tax cut, as would be the case following the 2001 tax rebate, for example.[8] Second, the surprise could reflect a tax decrease or a tax cut derived from an across the board decrease in the MTR, as would be the case following the Tax Reform Act of 1986, for example. Third, it could be nothing but a timing shift in the receipt of after-tax income, as would be the case following the 1991-92 reduction in income tax withholding, for example. As long as the household has imperfect prior information about both the level and the slope of the tax schedule, the unexpectedly high after-tax income will, in general, be partly interpreted in all three ways. Naturally, other interpretations are possible as well.

The first contribution of this paper is in developing a model that formalizes this intuition. The model generalizes the standard full-information rational agent model by allowing for imperfect knowledge of the income tax schedule. In the model, a household is subject to a linear income tax schedule that changes from year to year due to innovations that are predictable, but not necessarily predicted, well in advance.[9] The household perceives these innovations with noise due to information gathering and processing costs. As a result, the household is uncertain about the exact tax schedule it faces and it will use any signals generated by interaction with the tax system to update its beliefs. In particular, the model predicts that the beliefs about the current and future MTRs increase (decrease) with a surprisingly high (low) tax liability realized in the previous tax year.

The second contribution of the paper is in devising and implementing an empirical strategy that can identify whether households have a perfect understanding of their tax schedule changes (null hypothesis) or they experience some (ex ante or ex post) misperceptions, particularly of

[7]We discuss this point in more detail in Section 7

[8]The 2001 tax rebate originated from a 5 percentage point tax cut on the first \$12,000 (\$6,000) dollars of taxable income when filing as a married couple (single). Since such reduction in the MTR was inframarginal for a vast majority of households, the resulting tax cut in the given bracket was in fact lump-sum. In addition, many taxpayers experienced effective MTR cuts as well.

[9]Unpredictable tax schedule changes can be incorporated into the model at the cost of higher complexity without modifying the predictions of the model.

their MTR (alternative hypothesis). We identify a variation across households in the change of tax liability from one year to another that is *lump-sum, predictable in advance, and exogenous*, and examine how this variation affects parental labor income in married couple households in the year following the realization of the tax change.[10] This variation originates from age-discontinuity in eligibility for the Child Tax Credit (CTC) at the age of 17. In particular, in order to account for idiosyncratic heterogeneity in the level of labor income across households, we identify the effect by comparing the growth rate between years t and $t+1$ of parental labor income of married couples whose child turns 17 before the end of year t and who therefore lose the credit in year t to their counterparts whose child turns 17 early in year $t+1$ and who therefore do not lose the credit in year t. It is important for the variation to be lump-sum and predictable since, in case of no misperceptions, there should be no reaction in labor supply and labor income, except in the presence of liquidity constraints. In contrast, this would not be the case in the absence of misperceptions if the variation is unpredictable or it comes with changes in the MTR, when both income and substitution effects would play a role.

We implement the identification strategy using panel data from the 1996-1999 and 2000-2003 waves of the U.S. Census Bureau's Survey of Income and Program Participation (SIPP). We find that losing an eligible dependent has a negative impact on the *growth rate* of parental labor income in the year following the tax year in which the credit was lost. In order to examine whether this result is not just an artifact of the particular dataset or the relatively small number of observations used in the estimation, we replicate the identification strategy on the 2005-2007 repeated cross-sections from the U.S. Census Bureau's American Community Survey (ACS). Using this data, we find the same negative effect on the *level* of parental labor income in the year following the tax year in which the credit was lost.

These findings are obtained despite the fact that losing the CTC has no mechanical impact on the MTR for our selected sample.[11] We show, using a variety of robustness tests, that these findings are not driven by a direct effect of child aging or a spurious correlation between timing of birth and income level or its growth rate. We also argue that the finding cannot be driven by a presence of liquidity constraints, since the latter would imply a higher rather than a lower growth rate of labor income for households who lose the credit. Based on this finding, we reject the null hypothesis in favor of the alternative hypothesis that claims existence of misperceptions.

[10] We focus on labor earnings because this measure captures not only hours worked, but other types of effort (such as looking for a better job) that increase the value of one's labor input.

[11] This selection is, among other things, based on the base-year parental labor income being in a certain range. Because the ACS is a series of cross-sections, such restriction suffers from a potential endogeneity problem in that dataset. However, the result in the ACS data is robust to whether we do or do not apply the income restriction.

This conclusion is some respect similar to the "excess sensitivity" findings in the PIH literature. However, it stresses that when it comes to labor supply, it is not only the income effect, but also the substitution effect that is likely to play a role. In addition, if one makes the assumption that the loss of the CTC is completely unexpected ex ante (which is a particular type of an ex ante misperception), then the result suggests existence of ex post misperceptions, particularly an upward revision of beliefs about the MTR following a lump-sum increase in tax liability.

The rest of the paper is organized as follows. Section 2 reviews existing empirical evidence on taxpayer confusion, tax complexity, and tax salience. Section 3 presents a model that formalizes the intuition behind ex post misperceptions of tax schedule changes. Section 4 describes our identification strategy. Section 5 describes the datasets that we use and the estimating equations. Section 6 discusses our results as well as the robustness checks. Section 7 relates our findings to the empirical literature on the PIH and discusses further interpretation of the result. Finally, Section 8 concludes.

2 Existing Empirical Evidence on Taxpayer Confusion, Tax Complexity, and Tax Salience

There is a stream of literature in public finance that analyzes how well-informed taxpayers are about the tax system that they face. One strand of this literature focuses on documenting taxpayer perceptions of the income tax schedule. Brown (1968) compares self-reported MTRs of a group of UK taxpayers to their actual MTRs computed out of employer pay records and concludes that taxpayers "think they pay higher rates of tax than is in fact the case." Fujii and Hawley (1988), using the Survey of Consumer Finances, compare respondent self-reported MTRs to estimates of these MTRs based on the available survey demographic and income data. They find that individuals systematically underestimate their computed MTRs.[12] Romich and Weisner (2000) find that a high fraction of low-income households do not correctly perceive MTRs implied by the Earned Income Tax Credit (EITC) for hypothetical levels of income. In particular, the respondents' knowledge appears to be based on experience within their current income range, which they incorrectly extrapolate to other income ranges.[13]

Building on the idea of complexity, Liebman and Zeckhauser (2004) propose a simple hypothesis for how households interpret tax liability, or, equivalently, net income shocks. They suggest

[12] This interpretation is, however, sensitive to the assumption on the use of itemized deductions.

[13] For example, households who are in the phase-in portion of the EITC often assume that the amount of the credit increases linearly with the amount of labor income, even though the amount of the credit flattens out after a certain income threshold, and after another threshold it decreases.

that households "schmedule", that is, approximate their true MTR by the average tax rate realized in the previous year, and provide some supportive evidence for this claim. They do not conceptually distinguish between predictable and unpredictable income innovations, however. In fact, we will show in the next section that their hypothesis is a special case of a more general updating model.

Rather than focusing on misperceptions and complexity, another strand of the literature focuses on the hypothesis that certain taxes or certain ways of framing them may be more visible, or salient, to taxpayers in comparison with other taxes or other ways of framing them. For example, de Bartolome (1995) provides experimental evidence based on revealed choices that when the tax schedule is presented as a table mapping taxable income to the amount of tax entry by entry (as in the table accompanying the personal income tax form 1040), "there are at least as many individuals who use the average tax rate 'as if' it were the marginal tax rate, as individuals who use the true marginal tax rate."[14] Blumkin et al. (2007) provide experimental evidence that subject behavior is more sensitive to income as opposed to consumption taxes, despite the fact that the two are constructed to be theoretically equivalent. They attribute this finding in part to their experimental design where the income tax is designed to be more salient than the consumption tax. Chetty et al. (2007) find that consumer demand depends on whether the sales tax is included in the posted price despite the fact that the final after-tax price is the same in either case. Finkelstein (2007) goes a step further by arguing that if a particular tax or levy is less salient, then the tax base is less elastic to it, which in turn implies that the optimal tax or tax rate is higher. To support the claim, she documents that freeway toll charges are higher in places that use electronic toll debiting compared to places that collect tolls in cash.

3 Model

In this section, we formalize the intuitive hypothesis presented in Section 1 about how households interpret net income, or, equivalently, realized tax liability surprises. Formally, suppose that household faces a linear tax schedule in every period $t \in \{0, ..., T\}$ of its lifetime with the MTR given by τ_t and the intercept given by D_t. That is, the tax liability $T_t(y)$ of this household in period t based on the taxable income y is determined by $T_t(y) = D_t + \tau_t y$ for all $y \geq 0$. Although in reality many tax schemes are only piecewise linear, the proposed linear tax schedule can be thought of

[14] However, given that as many as 85 percent of taxpayers nowadays rely on a tax preparer or a tax preparation software, with the fraction growing over time, the significance of tax schedule framing for taxpayer decisions potentially affects only a relatively small and declining portion of taxpayers.

as an approximation of an otherwise more complicated tax scheme in its relevant range. This schedule varies from household to household because of different demographic characteristics such as the number of children and their age, taxpayers' age, disability status, type of income, etc. It also varies from year to year because of predictable and unpredictable changes in the tax schedule. The predictable changes are due to a variety of provisions related to the age of the taxpayers or their children, or due to tax consequences of planned actions such as mortgage interest payments. These changes are, under a stable tax system, predictable many years in advance. Unpredictable changes, on the other hand, are due to tax reforms as well as realization of states of the world that have tax consequences, such as medical expenditures, disability, number and timing of children, etc. In what follows, we will only focus on predictable changes. Unpredictable changes are obviously realistic, and they can easily be incorporated into the analysis without qualitatively affecting the results.

Formally, the parameters of the tax schedule affecting the household follow a process

$$\begin{pmatrix} \tau_{t+1} \\ D_{t+1} \end{pmatrix} = \begin{pmatrix} \tau_t \\ D_t \end{pmatrix} + \begin{pmatrix} \phi_{\tau t+1} \\ \phi_{Dt+1} \end{pmatrix}, \quad (1)$$

where $\phi_{t+1} \equiv (\phi_{\tau t+1}, \phi_{Dt+1})^T$ is a vector of predictable changes in the parameters of the tax schedule between years t and $t+1$. However, the household may perceive these changes with an error, resulting in its expectation of the change ϕ^e_{t+1} diverging from the actual change ϕ_{t+1}. In particular, from the point of the view of the household, $\phi^e_{t+1} - \phi_{t+1}$ is a realization of $N(0, S_t)$, with realizations in different time periods assumed to be independent. Although this simplifying assumption rules out the possibility that a tax liability surprise is perceived as a pure shift in the timing of taxes, it simplifies the exposition and allows us to focus on the confusion between changes in the intercept and the slope of the tax schedule. The matrix S_t measures the household's ability to correctly perceive the predictable changes. For a perfectly informed household, $S_t = 0_{2\times 2}$, and hence the predictable changes are in fact predicted without error. For a less than perfectly informed household, S_t is a non-zero positive semi-definite matrix, meaning that ϕ^e_{t+1} is only a crude measure of the predictable change in the parameters of the tax schedule between periods t and $t+1$. Although the normal distribution places a positive measure on the perceived MTR exceeding unity or falling below any arbitrary negative threshold, the stochastic specification in (1) may be thought of as a tractable approximation of beliefs over a bounded interval and we therefore overlook the problem of unboundedness in what follows.[15]

[15] An alternative modeling strategy would be to assume mean reversion in the parameters of the tax schedule.

We also assume that the household does not necessarily have an exact knowledge of the tax schedule when it first enters the labor force. In particular, its prior beliefs about the MTR and the intercept of the tax schedule at the end of period 0 are given by

$$\begin{pmatrix} \tau_0 \\ D_0 \end{pmatrix} \sim N\left[\mu_0, \Sigma_0\right], \tag{2}$$

where μ_0 is the vector of the actual parameters facing the household. Again, the matrix Σ_0 determines the extent to which the household is aware of details of the tax schedule when it first enters the labor force. For a fully informed household, $\Sigma_0 = 0_{2\times 2}$, while for a less that fully informed household, Σ_0 is a non-zero positive semi-definite matrix.

A few conceptual remarks are in place here. Real-world tax schedules are predominantly piecewise linear, and hence the series of linear schedules proposed here can be thought of as an approximation of the effective segments of the tax schedule facing the household each year. In light of this interpretation, *actual* year-to-year changes in the two tax parameters may reflect not only changes in the effective linear segment of the tax schedule for a fixed level of income, but also switches among different effective linear pieces of the tax schedule over time. Although the notation ignores household heterogeneity, this is just a notational convenience as the evolution of the tax parameters and the precision of their perceptions can vary from one household to another.

We also need to mention some technical remarks. We treat S_t and Σ_0 as exogenous, but in reality households have a control over how detailed their knowledge of the tax schedule and its changes is. This would suggest introducing an explicit cost of information acquisition and modeling the two variance matrices as outcomes of comparing marginal costs and benefits of information (Feige and Pearce, 1976; Buiter, 1980; Reis, 2006; Demery and Duck, 2007). On a different note, although not due to a conscious effort, the "size" of S_t may be an unintended consequence of income variation. For example, if the household taxable income fluctuates in a relatively narrow range from one year to another, switches among different segments of the tax schedule are not so frequent and hence S_t may be "small". On the other hand, S_t may be substantially "larger" if the household experiences large year-to-year taxable income variation. More details of taxable income history may matter as well in that the household may have more precise beliefs and information about segments of the tax schedule "familiar" from the past. Although these extensions are plausible and worth future exploration, the purpose of the current

This was done in a previous version of the paper and is available upon request. The exposition becomes more complicated with no effect on the qualitative results.

model is to analytically illustrate mechanics of updating based on realized tax liability in most simple terms, and we hence proceed with exogenous S_t and Σ_0. This can be understood as a reduced-form version of a more complete model with concious as well as accidental information acquisition. Yet another potential modeling extension it to allow ϕ_s^e for $s > t$ to be updated in period t based on the most recent available information. Incorporating this addition would not affect the central message of the model and we therefore omit it for simplicity.

At the end of period t, the household files its tax return for that period.[16] Conditional on pre-tax income y_t in period t, the household observes its tax liability $T_t = D_t + \tau_t y_t$ which serves as a signal for (τ_t, D_t). The following proposition characterizes the evolution of beliefs about the parameters of future tax schedules based on past and current realizations of tax liability.

Proposition 1 *Suppose that S_t is positive definite in all time periods, or that three of its elements are zero and the remaining diagonal element is positive. Then the beliefs about the parameters of the tax schedule in period $s \in \{t+1, .., T\}$ at the end of period t are given by a normal distribution with mean*

$$
\begin{aligned}
E_t\left[(\tau_s, D_s)^T\right] &= E_t\left[(\tau_t, D_t)^T\right] + \sum_{u=t+1}^{s} \phi_u^e \\
&= \mu_0 + \sum_{u=1}^{t} \Gamma_u \left[T_u - E_{u-1}\left(T_u | y_u\right)\right] + \sum_{u=1}^{s} \phi_u^e \qquad (3)
\end{aligned}
$$

and variance

$$
Var_t\left[(\tau_s, D_s)^T\right] = \Sigma_t + \sum_{u=t+1}^{s} S_t, \qquad (4)
$$

where Σ_u is defined recursively by

$$
\Sigma_u = \frac{\det(\Sigma_{u-1} + S_u)}{(y_u, 1)\left(\Sigma_{u-1} + S_u\right)(y_u, 1)^T} \begin{bmatrix} 1 & -y_u \\ -y_u & y_u^2 \end{bmatrix}, \; u = 1, ..., t, \qquad (5)
$$

and

$$
\Gamma_u \equiv \frac{\left(\Sigma_{u-1} + S_u\right)(y_u, 1)^T}{(y_u, 1)\left(\Sigma_{u-1} + S_u\right)(y_u, 1)^T}, \; u = 1, ..., t. \qquad (6)
$$

Proof. See the Appendix. ■

Intuitively, in each time period t relative to the previos time period $t-1$, the mean of beliefs over $(\tau_s, D_s)^T$ for $s > t$ is adjusted based on realization of the tax liability surprise $T_t - E_{t-1}\left(T_t | y_t\right)$,

[16] In reality, households file their tax returns in the early part of the year following the tax year in question. However, as long as such filing has a potential to affect the behavior in the year of filing, the exact timing of the filing is less important. We assume it happens at the end of period t for a simplicity of notation.

with the slopes of the adjustment given by Γ_t. Note that the surprise is only due to unexpected changes in the tax parameters since any possible income change is conditioned out. [17] The analytical form of Σ_t underlines the fact that at the end of period t, $\tau_t y_t + D_t = T_t$ is known with certainty, and hence $\Sigma_t (y_t, 1)^T = 0$.

Signs of the effects of the realized tax surprise in period t on the expected value of the beliefs about the parameters of future tax schedules are given by the signs of the elements of Γ_t. Given the assumptions on S_t, $(y_t, 1)(\Sigma_{t-1} + S_t)(y_t, 1)^T$ is positive, and hence the signs of the elements Γ_t depend on the signs of the elements of $(\Sigma_{t-1} + S_t)(y_t, 1)^T$. As we informally discussed before, one would expect that an unexpectedly high realization of tax liability would lead the household to revise upwards its belief about both the MTR and the intercept. However, this prediction hinges on the covariance between the prior beliefs about the two parameters from the previous period as well as on the covariance in the realization, relative to the expectation, of their changes in the current period. The following proposition provides sufficient conditions for the intuitively appealing signs.

Proposition 2 *Suppose that*

$$-[S_t]_{12} < \min \left\{ y_t [S_t]_{11}, \frac{[S_t]_{22}}{y_t} \right\}. \tag{7}$$

Then there exists an $\varepsilon_t > 0$ such that if $|\Delta y_t| < \varepsilon_t$, then $[\Gamma_t]_{11}, [\Gamma_t]_{21} > 0$. That is, an unexpected positive shock in the realized tax liability increases the mean of the belief about both τ_t and D_t.

Proof. See the Appendix. ■

That is, a sufficient condition for the intuitively appealing sign pattern is that the covariance in the perceived noise of predictable changes in τ and D is not too negative and year-to-year taxable income changes are moderate. This applies even if the surprise comes from a change in the intercept only, without any real change in the MTR, and it may induce a negative substitution effect on labor supply. This is an important observation since the natural experiment employed in our identification strategy presented in the next section constitutes a lump-sum change in tax liability.

[17] If one assumes that the household is only confused about the MTR, but not about the intercept, then all of the elements of S_t and Σ_0 except for the element $(1,1)$ are equal to zero. In this case Proposition 1 implies that all of the elements of Σ_{t-1} except for the element $(1,1)$ are equal to zero, and hence $\Gamma_t = (1/y_t, 0)^T$. As a result, any unexpected hike in the tax liability is reflected in an increase in the expectation of future MTRs by the magnitude of the surprise in the realized *average tax rate*. If coupled with the assumption that there is no intercept in the tax schedule at any time period, this case corresponds to the "schmeduling" hypothesis considered by Liebman and Zeckhauser (2004). When schmeduling, a household predicts its MTR for the current period to coincide with the average tax rate realized in the previous period.

However, tax reforms often match signs of changes in MTR and the intercept of the tax schedule segment. The same is often true when switching among different tax brackets as well. Likewise, many households face significant income variation from one year to another. As a result, it is not clear that the sufficient condition is satisfied. We therefore put forward only a more general hypothesis: rational and well-informed households do not change their perception of current and future tax schedule parameters upon experiencing tax schedule or segment changes. The argument is particularly appealing for tax parameter changes that are predictable well in advance. On the other hand, boundedly rational or less well-informed households may misinterpret realized tax liability surprises and may change their perception of future tax schedule parameters in a way that does not accord with their actual changes.

4 Identification Strategy

We are interested in examining how parental labor income reacts to predictable lump-sum variation in after-tax income in the previous year. Our identification strategy is based on variation generated by the eligibility rules for the Child Tax Credit (CTC). Beginning in 1998, taxpayers with a dependent below 17 years of age on December 31 of the tax year in question could claim a credit of $400 per eligible child. This credit was generally non-refundable and only households with a sufficiently high tax liability were able to take a full advantage of the credit.[18] At the same time, the Additional Child Tax Credit (ACTC) was introduced. This credit provided for a limited refundability of the non-refundable part of the CTC for families with three or more qualifying children.[19] The CTC was increased to $500 for the 1999 and 2000 tax years, $600 for the 2001 and 2002 tax years, and $1,000 for the 2003 tax year, where it currently stands. At the same time, beginning in 2001, the ACTC was expanded to allow *any* family to claim the non-refundable part of the CTC up to one tenth of the excess of their earned income over $10,000.[20] The CTC has historically been phased out with adjusted gross income above $110,000 for married couples filing

[18]There are several provisions in the tax code that make the tax schedule a function of whether a dependent child did or did not reach a certain age in a given tax year. One such provision is the loss in the eligibility for the personal exemption and the Earned Income Tax Credit for a dependent child who turns 19 (or 24, if a full time student). This provision has been exploited by Looney and Singhal (2004) and Dokko (2005) in order to estimate the effect of marginal tax rates on labor supply.

[19]These families could claim the non-refundable part of the CTC up to the amount of employee contributed social security and medicare taxes less any earned income tax credit they received.

[20]The $10,000 threshold has been indexed for inflation over time. In addition, starting in 2004, the ACTC limit was increased to 15 percent of earned income in excess of the threshold. Families with three or more eligible children could still claim the non-refundable part of the CTC up to the amount of employee contributed social security and medicare taxes less any earned income tax credit they received if this limit turned out to be higher.

a joint tax return[21] at the rate of 5 percent.[22] In this study, we restrict attention only to married couples and assume that they file a joint tax return.[23]

Three features of the CTC make it a good natural experiment for testing our hypothesis of interest. First, to be eligible, the dependent child must not have reached 17 years of age by December 31 of the tax year in question. Because the timing of a child's 17th birthday is perfectly predictable, so is the implied timing of the net income loss. Second, due to the ACTC, virtually any household with up to two dependents above the labor income level of $30,000 or more can take advantage of the full amount of the CTC within the time period we consider (2001-2007). As a result, the loss of the CTC constitutes a pure lump-sum change in the tax liability and hence after-tax income. On the other end of the income spectrum, the argument extends all the way to the adjusted gross income of $110,000, after which the phase-out range begins. The loss of the CTC therefore constitutes *lump-sum* variation in tax liability and net income for all intermediate income households.[24] Third, it can be difficult to plan the timing of birth for a particular quarter, month, or day. As a result, among families whose children turn 17 before the end of year t or at the beginning of year $t + 1$, eligibility for the CTC is virtually exogenous.[25] Put together, losing the CTC generates a *exogenous, predictable* and *a lump-sum* variation in net income. As a result, we can identify the effect of predictable lump-sum variation in tax liability in year t on parental labor income in year $t + 1$ by comparing the growth rate between years t and $t + 1$ for households whose child turned 17 in a fixed time window at the end of year t (the treatment group) to households whose child turned 17 in the time window of the same length at the beginning of year $t + 1$ (the control group). Likewise, we can identify the effect of predictable lump-sum variation in tax liability on the level of parental labor income by comparing this level in year $t + 1$ for households whose child turned 17 in a fixed time window at the end of year t (the treatment group) to households whose child turned 17 in the time window of the same length at the beginning of year $t + 1$ (the control group).

The null hypothesis is that households are fully rational and fully informed, meaning that they have no misperceptions of the tax schedule that they face, either ex ante or ex post. Because

[21] The thresholds are $75,000 and $55,000 for single/head of household taxpayers and married taxpayers filing separately, respectively. None of these thresholds are indexed for inflation.

[22] That is, a household loses $0.05 of the credit for every extra dollar of adjusted gross income above the threshold.

[23] In general, based on IRS statistics, over 95% of married households file jointly.

[24] In fact, in the SIPP data we select the sample such that the base year (before the CTC loss is realized) labor income is in the range of $30,000 to $110,000 in order to minimize the possibility that losing the CTC has a *mechanical* impact on the effective MTR of a household. In the ACS data, where such restriction may potentially lead to an endogenous sample selection problem due to data being a cross-section, we show that our qualitative findings are robust to whether we apply such restriction or not.

[25] In section 6, we discuss evidence why this may not be the case and test the robustness of our results to the potential endogeneity in the timing of birth.

the change is lump-sum and predictable, we should observe no impact on parental labor income in the subsequent year, i.e., no difference in this variable between the treatment and the control groups. One possible exception may come from the presence of liquidity constraints, in which case we would expect a positive effect of the credit loss. The alternative hypothesis is that households do experience misperceptions, in which case the effect on parental labor income will depend on the resulting income and substitution effects, and possibly also on the liquidity constraint effect. As a result, finding of a zero or a positive effect of the treatment is consistent with both the null hypothesis and the alternative hypothesis, whereas a negative effect of the treatment is only consistent with the alternative hypothesis.

5 Data and Empirical Implementation

Our identification strategy requires a dataset that contains information on household labor and non-labor income, number of children and their dates of birth, as well as basic household demographic characteristics. We use two sources of data in this study. Baseline results come from the 1996-1999 and 2001-2003 waves of the U.S. Census Bureau Survey of Income and Program Participation (SIPP). The SIPP is a nationally representative longitudinal survey of households in which each household is interviewed every four months over the course of three years for a total of nine waves. This survey collects information on income, employment, and detailed demographic information on all family members. The SIPP data also contain very specific information on the year and month of birth of each child in the household. Based on this data, we compute tax liabilities and tax rates using the NBER's TAXSIM (Feenberg and Coutts, 1993) calculator. An advantage of this dataset is its panel structure, which allows controlling for heterogeneity in the level of household labor income using fixed effects and focusing on (presumably less noisy) growth rate of this income as the outcome variable. A disadvantage of this dataset is its smaller size. In particular, after reducing the data to the set suitable households, we are left with a few hundred observations in each specification. Given that the average potential impact of misperception on the growth rate of labor income may be small, it may become difficult to distinguish between economic and statistical significance of the estimates.

To address this potential problem and to examine the sensitivity of results to the usage of a particular dataset, we also use the Census' Integrated Public Use Microdata Series American Community Survey (ACS) from 2005-2007.[26]. ACS is a collection of repeated cross-sections, each

[26]Unfortunately, we could not use data from the same period as in SIPP dataset because a crucial quarter-of-birth variable is not available in the ACS data during 1990-2004

with over a million households. Like SIPP, it records detailed demographic information on all family members. However, unlike SIPP, it records quarter of birth of household members and their age as of the last birthday. Even after reducing the data to the set of suitable households, we are typically left with around 14,000 observations in each specification. This large number of observations is the main advantage of the ACS dataset. However, this comes at a price of three disadvantages. First, given that this dataset is a series of cross-sections rather than a panel, we are not able to eliminate unobserved heterogeneity in the level of household labor income like we do in SIPP by focusing on the income growth rate. Second, for the same reason we are also not able to restrict the sample by a base-year income to a range in which the CTC loss has no impact on the MTR since this income is potentially affected by the treatment, raising endogenous sample selection issues. Hence we use all available data in the estimation, although the results are robust to applying the sample restriction. Third, because we only observe age as of the last birthday as opposed to observing the year of birth, the classification into treatment and control groups suffers from a degree of imprecision discussed below.[27] Due to these limitations, the ACS-based results should be interpreted purely as robustness tests of the SIPP-based results.

5.1 SIPP Data

To implement our identification strategy using the SIPP data, we require at least two complete consecutive years of data for each household in the sample. The responses of households interviewed in each wave refer to the previous quarter. Because there is no wave in the first quarter of 2004, data is missing for the final quarter of 2003. In order to be able to use the 2003 data on an annual basis, we compute the 2003 annual income as 12/9 times the sum of income in the first three quarters of 2003.[28] The coverage of the 1996-1999 panel, on which some of the robustness tests are based, starts midway through 1996 and ends midway through 1999. As a result, we have income data for only two quarters in 1996 and 1999. We therefore use only years 1997 and 1998 that give us complete information on income.

[27] A potential disadvantage of both the SIPP and the ACS data is that we only have survey as opposed to administrative measures of household labor income. An alternative dataset with more precise information on income and tax variables would be the IRS Statistics of Income (SOI) Tax Return Microfiles. However, this dataset is unsuitable for our analysis because it does not provide explicit information on dependents' ages (and it provides only very coarse information on household demographics). In case a panel of tax returns were available, we could potentially obtain necessary information from tracking eligibility for child-related credits and exemptions over time. However, the last period for which such panel is publicly available is late 1980s, which is well before the introduction of the Child Tax Credit, the main identification instrument of our study.

[28] The subsequent results are robust to excluding the 2003 data and using only data from 2001 and 2002, for which we have complete data for all 4 quarters. These results are available from the authors upon request. Note that the results are identical for any imputation procedure that assumes that the 2003 income is a multiple, fixed across all households, of the income earned in the first 9 months of 2003. This would be the case if, for example, households earn end-of-the-year bonuses proportional to their base pay.

The treated group consists of households that have at least one child who turns 17 in the last k months of year t and the control group consists of households that have at least one child who turns 17 in the first k months of year $t+1$.[29] We define a "cohort" to be the year in which the eligible dependent turns 17. Then given the data availability, our baseline results based on the 2001-2003 panel use cohorts $t = 2001$ and $t = 2002$, whereas the robustness tests based on the 1997-1998 panel use the cohort $t = 1997$.

The choice of the time window is driven by a tradeoff between the sharpness of the regression discontinuity design and the identification power due to available number of observations. In order to minimize any omitted variable bias, it is desirable to define the time window to be as narrow as possible. However, we have on average only about 35 observations per month in each of the two groups. As a result, our baseline results are based on a larger time windows of +/-6 months. However, in order to evaluate the sensitivity of the findings to the choice of the time window, we also conduct analogous estimations in time windows of +/-1 and +/-12 months.

The 2001-2003 panel originally contains data on 36,700 households. We apply five restrictions to this data. First, we use data only on married couples that have at least one child who turns 17 between 2001 and 2004 (some of the households are only used in alternative specifications to our baseline results or in subsequent robustness tests). Second, we restrict attention only to households with at least two consecutive years of complete data on yearly income[30] and information on the following control variables: age and the highest achieved education level (high school diploma, associate degree, or college degree) for each spouse, and the number of dependents (children under the age of 24 living at home with parents) in the household. Third, we drop any households in which at least one of the parents is above 62 years of age in 2001 (or 2002 if data does not exist for 2001) because we do not want to confound our results by retirement decisions. Fourth, we only use data on households who have adjusted gross income, as computed by TAXSIM, in the range of $30,000 to $110,000 in the base year (i.e., before filing the tax return on which the loss of the Child Tax Credit is realized). This is done in order to eliminate the phase-in and the phase-out ranges of the credit, in which the loss of the CTC does impact the MTR. However, this selection on base-year income may raise issues of endogenous selection in case some of the sample households had a child turn 17 in years 1999 or 2000, the first two years in which it was possible to lose the credit.[31] As a result, we finally drop households who had at least one child turn 17 in

[29]There are a few households where more that one child turns 17 in the relevant time window. Although this leads to a loss of multiple credits, we still indicate the treatment or control group in the same way as for households who only lose one eligible dependent.

[30]As mentioned earlier, we impute income in the last quarter od 2003 in case the information for the other three quarters of that year is available.

[31]Recall that the Child Tax Credit first applied in 1998. Therefore the first year in which a household could lose

1999 or 2000 to mitigate this problem.

These data restrictions significantly reduce the universe of the remaining data. Depending on the specification, our sample sizes run from as small as 57 observations in the one-month sample to as large as 883 in the 12-month sample. Moreover, a number of robustness tests contain other distinct subsamples of the data generally containing around 500 observations. Panel A of Table 1 presents the means of the demographic control variables and labor income for the treatment and the control groups for the baseline six-month subsample. Results of standard t-tests show that the equality of means of control variables between these two groups cannot be rejected at any conventional level of significance for any of the variables. As a result, if there is a difference in the growth rate of labor income between the two groups, it is not driven by the heterogeneity of the two groups in terms of observable demographic characteristics.

Using analogous sample restrictions on the 1997-1998 panel, the sample is reduced to 192 unique households (the complete 1996-1999 panel originally contained 40,188 households). Panel B of Table 1 provides summary statistics for this dataset with the treatment and control group definition based on the six-month time window (104 observations). Again, the equality of means of control variables between the two groups cannot be rejected at any conventional level of significance for any of the variables.

With observations selected as discussed, we estimate the following equation by OLS, clustering the standard errors at the household level:

$$\Delta \ln Y_{it+1} = \beta_0 + \beta_1 T_{it} + \gamma I_{2002} + \pi' X_{it+1} + u_{it+1}, \tag{8}$$

In this equation, $t \in \{2001, 2002\}$, T_{it} is the indicator of the treatment group, i.e., whether a household i has a child who turn 17 in the last k months of year t, I_{2002} is an indicator for $t = 2002$, and X_{it+1} is a vector of household demographic characteristics in year $t+1$. The intercept and the indicator variable for $t = 2002$ control for a secular non-linear time trend in labor income. As discussed before, we use time windows of $k \in \{1, 6, 12\}$.

5.2 ACS Data

In this dataset, the treated group would ideally consist of households that have at least one child who turns 17 in the last quarter of year t and the control group would ideally consist of

it was 1999.

households that have at least one child who turns 17 in the first quarter of year $t+1$.[32] We define a "cohort" to be the year in which eligible dependents it the treatment group turn 17. Given the data availability, we use cohorts $t = 2004$, $t = 2005$ and $t = 2006$.

A difficulty with the ACS dataset arises because while we know quarter of birth of the dependent and the age as of the last birthday, we cannot determine when the household was interviewed. [33] This means that we potentially misclassify households in our treatment and control groups. In particular, focusing on the 2005 cohort, we classify as belonging to the treatment group any household which, when interviewed during 2005, reports having a child born in the fourth quarter and being 17 as of the last birthday. In case the interview had taken place anytime during the first three quarters of 2005, the household is classified correctly. However, if the interview had taken place in the last quarter of 2005, our classification includes households with children who turned 17 in the fourth quarter of 2005 before the interview took place, and hence these households are classified incorrectly. In particular, these households' children are a full year behind our intended treatment group. An analogous problem arises for the control group, which includes households whose children turn 17 in the first quarter of 2004 as opposed to the first quarter of 2005 as intended, and hence are a full year ahead of our intended treatment group. If we hypothetically assume uniform distribution of births and interviews across the year, approximately one eighth of observations in both groups are misclassified. Both misclassifications may introduce bias into our estimates which is hard to quantify. The results should therefore be taken with caution and perhaps understood as another robustness test of the results based on SIPP data.

The original 2005-2007 dataset contains information on nearly 3,660,000 households. We apply five restrictions to this data. First, we use data only on married couples living together with their children and no other family members that have at least one child who is reported to be 17 as of the last birthday and is reported to be born in the first or the fourth quarter (with some households only being used in robustness tests). Second, we only use households with available data on age and educational level of parents and the number of dependents in the household. Third, we drop all households where at least one of the parents is above 62 years of age as of the preceding birthday. Fourth, we drop all households in states in which the school start cutoff date in 1993-1996 was later than (and including) September 30th. This is done in order to minimize the possibility that our treatment group would be one grade ahead of our control group, something that may affect labor income as children are leaving to college or moving out on their own.[34]

[32]As in with the SIPP data, if more that one child turns 17 in the relevant time window, we indicate the treatment or control groups in the same way as for households who only lose one eligible dependent.

[33]This information is held confidential and is not released to the general public, despite our repeated requests.

[34]Of course, we ideally would like to know where each household lived when the child turned six years old but

Fifth, we only use households where it is the first child who is turning 17, and we apply the similar restriction also in case of placebo tests.

Applying these restrictions, we obtain approximately 7,250 observations in either the treatment or the control group across all the cohorts. Panel C of Table 1 contains the summary statistics. Standard t-tests do not reject the equality of means of individual control variables between the two groups at conventional levels of statistical significance. Hence, as in the SIPP data, if there is a difference in the labor income between the two groups, it is not driven by differences in observable demographic characteristics.

With observations selected as discussed, we estimate the following equation by OLS, robustifying the standard errors for heteroscedasticity:

$$\ln Y_{it+1} = \beta_0 + \beta_1 T_{it} + \gamma_1 I_{2005} + \gamma_2 I_{2006} + \pi' X_{it+1} + u_{it+1}, \tag{9}$$

In this equation, $t \in \{2004, 2005, 2006\}$, T_{it} is the indicator of the treatment group, i.e., whether a household i has a child who turned 17 in the last quarter of year t, I_{2005} and I_{2006} are indicators for $t = 2005$ and $t = 2006$, respectively, and X_{it+1} is a vector of household demographic characteristics in year $t + 1$. The intercept and the indicator variables for $t = 2005$ and $t = 2006$ control for a secular non-linear time trend in labor income.

The next section presents the results and estimates of β_1 based on (8) and (9). It also discusses a series of robustness checks that evaluate to what extent these estimates may be affected by a potential correlation between T_{it} and u_{it+1} originating from an omitted variable bias due to a direct effect of child aging on parental labor income and from a potential spurious correlation between the timing of birth and income.

6 Results

6.1 Main Result

Table 2 reports our baseline results. Panel A estimates based on (8) using the SIPP data for the time window of plus or minus six months in column (1), one month in column (2), and 12 months in column (3). In column (1), the estimate of β_1 shows that households whose children turn 17 in the last six months of year t have (approximately) a 3 percentage points *lower* growth rate of labor income between years t and $t + 1$ compared to households whose children turn 17

this is unobserved. Thus, we assume that the state in which the child lived when turning 17 is the same state as when the child turned six.

in the first six months of year $t+1$ (p-value of 0.059). We obtain a somewhat larger (in absolute value) and statistically more significant result for the 12-month window in column (3), with the growth rate differential being 3.8 percentage points (p-value of 0.001). On the other hand, the estimate for the one-month window is virtually zero (-0.009 with the p-value of 0.86). We also reestimated each specification by the median regression in order to examine the robustness of the OLS estimates to outliers, with P-values computed by a nonparametric bootstrap with 1000 replications and clustering at household level. The estimates (p-values) are -0.026 (0.045) for the six-month window, 0.014 (0.860) for the one-month window, and -0.034 (0.001) for the twelve-month window.

Panel B, or column (4), of Table 2 presents an estimate based on (9) using the ACS data. The estimate of β_1 shows that households whose children are born in the last quarter and are 17 as of the last birthday when interviewed in year $t+1$ have a 2.5 percentage points *lower* level of parental labor income in year $t+1$ compared to households whose children are born in the first quarter and are 17 as of the last birthday when interviewed in year $t+1$ (p-value of 0.011). When estimated by the median regression, the estimate (p-value) is -.019 (0.065).

These findings, with the exception of column (2), contradict the null hypothesis of $\beta_1 = 0$ and support an alternative hypothesis of $\beta_1 < 0$. In light of the discussion in Section 4, this empirical result indicates the presence of a substitution effect, and, hence, imperfect ex post understanding of the surprise in net after-tax income.[35] On the other hand, the SIPP-based result for the one-month window in column (2) is not statistically significantly different from zero. Although the one-month specification most closely approaches the ideal identification design based on age discontinuity in CTC eligibility, it comes at the cost of only 57 observations, and hence large standard errors. This invites a question of how the one-month estimate would look like if we had more observations in that time window. The results for the other two larger time windows are suggestive in this respect. However, the larger the window, the larger are the possible unobserved differences between our treatment and control groups as well. In the next subsection, we test the robustness of our baseline results for several possible alternative explanations.

[35] Some of this effect may reflect a shift of labor supply to an untaxed informal sector rather than to leisure or household production. However, since neither of our datasets contain separate information on informal labor market earnings, we are not able to address this hypothesis. Another possibility is that rather than changing labor supply, households strategically shift the timing of income realization. We come back to this possibility in the next subsection.

6.2 Robustness Tests

Identification of the effect of losing the CTC on parental labor income by means of estimating β_1 in (8) or (9) by OLS rests on two basic assumptions: (1) eligibility for the credit is not correlated with any variables which we do not control for and that may have a systematic effect on the outcome variable; and (2) there is no spurious non-tax correlation between the outcome and the treatment variable. The purpose of this subsection is to examine potential violations of these assumptions and their impact on our interpretation of the results.

First, despite the age discontinuity design, one may envision that the treatment variable T_{it} may be correlated with unobserved changes in tastes for supplying labor. This is because, especially with larger time windows in SIPP data, the dependents on which the identification is based are somewhat older in the treatment group compared to the control group. Combined with the possibility that parental labor income responds to the age of their children, the estimate of β_1 may confound the effect of losing the CTC with a direct effect of the child's age.

Second, Bound et al. (1995) cite a number of references documenting, among other things, that the season of birth may be directly correlated with income due to children from high income families being less likely to be born in the winter months. This argument would suggest a spurious correlation between the timing of birth and the household's income class. Although this argument seems, to the first order, to affect only the ACS-based results, as documented by Autor et al. (2005), wage inequality among the U.S. households has been increasing over the period 1990-2005. As a result, the argument of Bound et al. (1995) may extend to the *growth rate* of labor income as well.

Third, Dickert-Conlin and Chandra (1999) argue that if a child is to be born around the turn of the year, parents may have a preference to speed up the birth on the margin so that they can claim tax benefits for the ending calendar year. The authors also find that such behavior is more prevalent among higher income households, raising another potential spurious correlation problem, especially when using the ACS data. However, this problem is less of a concern for our interpretation of the results since it suggests that we tend to underestimate β_1 in absolute value, suggesting that the true coefficient is even more negative.

All of these criticisms suggest that the baseline estimate of β_1 in six- and twelve-month windows may at least partly be driven by factors that have nothing to do with the causal effect of the loss in the CTC on parental labor income. However, these three criticisms are equally applicable to ages other than 17, and the first two are also applicable to time cutoffs other than the end of a calendar year. As a result, if our baseline estimates in Table 2 are mostly driven by these correlations, we should obtain similar estimates of β_1 at other ages cutoffs when the "treatment" has no CTC consequences, such as 15 or 16, or at other time thresholds when the CTC consequence is identical for the "treatment" and the "control" group, as a middle of a calendar year.

Table 3 presents placebo tests of this kind. In particular, we investigate the impact on the growth rate of parental labor income between years t and $t+1$ (in SIPP data, based on the 6-month window) and level of this income in year $t+1$ (ACS data, based on a 3-month window) of a child: (1) turning 15 in in year t versus year $t+1$; (2) turning 16 in year t versus year $t+1$; and (3) turning 17 in the first half of year t versus the second half of year t.[36] In each case, the estimated coefficient is statistically indistinguishable from zero at even 20 percent critical level and most estimates are numerically close to zero as well.[37] We also estimate the effect of: (4) turning 18 in in year t versus year $t+1$; and (5) turning 19 in in year t versus year $t+1$. The estimates in column (4) are positive and, in case of SIPP, statistically significant at 5 percent level. Since at this age threshold it is the control rather than the treatment group that is affected by the loss of the CTC in the preceding tax year, the positive estimated coefficients corroborate the baseline results presented in Table 2 and their interpretation. Note that this is true even if the loss in the dependent tax *deduction* for non-students in the year the dependent turns 19, if not a full-time student, that affects the treatment group in this case is taken into account. This is because if households are rational and well-informed, such loss is fully expected and hence has no income effect, although it does result in an increase in taxable income and hence potentially an increase in MTR. If the latter is the case, it should reduce labor income by the substitution effect. As a result, we should not observe a positive estimate. To abstract from the dependent loss issue, the results in column (5) are based only on households in which the 19-year-old dependent is a full-time student (in which case the household is still eligible for the deduction).[38] Again, the estimated coefficient is statistically insignificant and close to zero.

This set of robustness tests shows that the baseline estimates presented in Table 2 do not appear to be driven by a direct effect of child age or spurious correlation between income growth

[36]Due to the problems with exact 17th birthday classification in the ACS data that becomes particularly severe with the middle-the-year cutoff, we perform the latter placebo test only using the SIPP data.

[37]Regarding the result in column (2) for SIPP data, note that in 2003 the amount of the CTC was increased from \$600 to \$1,000. Because the information about this change was available before 2003, households in the control group for $t = 2002$ may have experienced a positive income shock since their children would be eligible for the (increased) credit in 2003, whereas this was not the case for the treatment group households. This income surprise may have reduced the growth rate of labor income of the control relative to the treatment group households, implying a positive bias in the estimate of β_1. The same criticism does not apply to $t = 2001$, though, since both the treatment and the control group children turned 17 before 2003. Consequently, we have also rerun the robustness test based on children turning 16 using only the data for $t = 2001$, with the coefficient estimate (p-value) being 0.002(0.96). As a result, this bias does not appear to be significant.

[38]Because we do not have sufficiently detailed information on student status in the SIPP data, we conduct this estimation only using the ACS data.

and the timing of birth as long as such correlations apply equally to ages other than 17 and are not dependent upon the end of the year cutoff date. One may still argue, however, that there may be a direct timing of birth effect that is particular to the 17th birthday and to the end of the year cutoff. For example, if a local school district uses the December 31 or a nearby cutoff date for the 6th birthday of a child in order to let the child enroll in the first grade in the preceding fall, children in the treatment group are much more likely to be in the senior rather than the junior year of their high school compared to children in the control group. As a result, the former are more likely to start college or work in the subsequent fall. This may have a direct impact on the growth rate of parental labor income, although we are not aware of any systematic evidence on the direction or size of this effect.[39] We do not face this problem in ACS data since we restrict to households living only in states where the 6th birthday cutoff date for first-grade enrollment is before September 30. However, given that we use 6- and 12-month windows in case of SIPP data, such restriction is not possible due to a very small resulting sample size.

Presumably, though, a direct timing of birth effect that is particular to the 17th birthday should be present regardless of the presence of the CTC. As a result, this identification problem in the SIPP data can be addressed by estimating a placebo effect of a child turning 17 before or after the end of a year in the absence of the CTC and compare it with an analogous effect in the presence of the CTC. We can implement this idea by applying our estimation strategy to the time period before 1998 (when the CTC was introduced), and comparing the estimate with the 2001-2003 estimate. To do this, we apply the same estimating procedure to the 1997-1998 panel and compare the growth rate of labor income between 1997 and 1998 for households whose child turns 17 at the end of 1997 versus at the beginning of 1998.[40] Panel A of Table 4 reports estimates of the pre-CTC "treatment" effect based on time windows of six and twelve months. The estimated coefficients (p-values) displayed in columns (1) and (2), are 0.009 (0.607) for the six-month window and 0.001 (0.892) for the 12-month window. These estimates are an order of magnitude smaller than the ones in the baseline specifications, and statistically highly insignificant. Panel B of this table then compares the pre-CTC and the post-CTC treatment effects using a triple-difference estimator. In particular, we estimate

$$\Delta \ln Y_{it+1} = \beta_0 + \beta_1 T_{it} + \beta_2 T_{it} I_{\{t \geq 2001\}} + \gamma_1 I_{\{t \geq 2001\}} + \pi' X_{it+1} + u_{it+1}, \tag{10}$$

[39]There is some evidence, though, on the effect of a child going to college on household consumption. In particular, Souleles (2000) finds that the effect is negligible.

[40]Note that the introduction of the CTC in 1998 is not an issue because neither the treatment nor the control group children would be eligible for it.

where $t \in \{1997, 2001, 2002\}$ and $I_{\{t \geq 2001\}}$ is an indicator for years after the introduction of the CTC. As before, T_{it} is an indicator for a household with a child turning 17 in the last k months of year t and the set of observations used to estimate this equation is restricted only to households in which at least one child turns 17 either in the last k months of year t, or in the first k months of year $t+1$, with $k \in \{6, 12\}$. We add an indicator for 2001-2002 (the time period after the introduction of the CTC) to control for a secular time trend in labor income. Recall that in 1997, the households are not affected by the CTC, so β_1 measures the direct effect of a child turning 17 at the end of year t as opposed to at the beginning of year $t+1$. On the other hand, β_2 measures how this effect changes between the pre-CTC period ($t = 1997$) and the post-CTC period ($t \in \{2001, 2002\}$). As a result, the effect measured by β_2 can be attributed to the introduction of the CTC, assuming that all other omitted factors that affected the growth rate of labor income between these two periods affected both types of households similarly. We estimate this equation by OLS and adjust the standard errors for clustering at the household level.

The estimate of β_2 is -0.038 using the six-month window (column (3)) and -0.042 using the 12-month window (column (4)), with the p-values equal to 0.126 and 0.012, respectively. In other words, the loss on the CTC is estimated to reduce the growth rate of parental labor income by 3.8 to 4.2 percentage points, but the only the latter estimate based on the 12-month sample is statistically significant at conventional levels. These estimates are similar in sign and significance to the baseline estimates, and, if anything, slightly larger in magnitude. The estimates of the direct effect β_1 of a child turning 17 before the turn of the year are positive and statistically highly insignificant, suggesting in yet another way that a direct effect of the child's date of birth does not appear to drive the baseline results.

Put together, these robustness tests document that a direct effect of child age or timing of birth on the growth rate of parental labor income or a spurious correlation between the latter and the timing (or season) of birth cannot account for the baseline estimates, which seem to be robust to multiple plausible critiques of the identification strategy. and hence they also suggest that the results in the one-month window are predominantly driven by the lack of more observations.

7 Discussion

7.1 How Large are the Implied Elasticity Estimates and Welfare Losses?

Our estimates from the previous section imply that households tend to reduce the growth rate of labor income in response to losing eligibility for CTC, and we interpret this finding as a result of a substitution effect due to a perceived increase in the MTR. In this subsection, we calibrate the implied estimates of the uncompensated elasticity of labor income with respect to the net-of-the-tax-rate share using the theoretic model presented in Section 3. This parameter is commonly estimated in the tax literature and we can therefore examine how our implied estimate compares to the ones obtained from traditional identification strategies. In the next step, we attempt to calibrate the additional deadweight loss resulting from such MTR misperception.

The relevant elasticity can be approximated by $\Delta \ln Y_{t+1}/\Delta \ln(1-\tau_{t+1})$ if the marginal tax rates τ_t and τ_{t+1} are known, where $\Delta \ln Y_{t+1}$ is the tax-driven component in the growth rate of labor income. In our application, $\Delta \ln Y_{t+1}$ is given by various estimates of β_1. However, given that we assume that households may misperceive their MTR, we do not exactly know what they believe about τ_t and τ_{t+1} as of the beginning of each respective period. In order to operationalize the calibration of $\Delta \ln(1-\tau_{t+1})$, we assume certainty equivalence behavior in that as far as choosing the level of labor activity goes, households pay attention only to the means of their beliefs. In particular, we assume that households behave according to the expected MTR as of the end of the previous period, i.e., according to $E_{t-1}(\tau_t)$ in period t and $E_t(\tau_{t+1})$ in period $t+1$. As a result, we calibrate the elasticity by

$$\widehat{\varepsilon} = \frac{\widehat{\beta}_1}{\ln\left[1-E_t(\tau_{t+1})\right] - \ln\left[1-E_{t-1}(\tau_t)\right]}. \tag{11}$$

Around $\overline{x} \in (0,1)$, we can approximate $\ln(1-x)$ to the first order by $\ln(1-\overline{x}) - (x-\overline{x})/(1-\overline{x})$. Therefore approximating both $\ln\left[1-E_t(\tau_{t+1})\right]$ and $\ln\left[1-E_{t-1}(\tau_t)\right]$ around the same point $\overline{\tau}_t$, we obtain that

$$\widehat{\varepsilon} \simeq -\frac{\widehat{\beta}_1\,(1-\overline{\tau}_t)}{E_t(\tau_{t+1}) - E_{t-1}(\tau_t)}. \tag{12}$$

We are now going to turn to the theoretical model presented in Section 3 in order to calibrate $E_t(\tau_{t+1}) - E_{t-1}(\tau_t)$. In order to proceed, we make three simplifying assumptions. First, we will focus only on househods who have a "similar" (to be made more exact later) taxable income in years $t-1$ and t and hence assume that $y_{t-1} = y_t$. Under this assumption, Γ_t simplifies to

$$\Gamma_t = \frac{S_t(y_t,1)^T}{(y_t,1)S_t(y_t,1)^T}. \tag{13}$$

Second, assume that the household expects no changes in the tax schedule between $t-1$ and t

and between t and $t+1$, i.e., $\phi_t^e = \phi_{t+1}^e = 0$. This assumption implies that

$$\begin{aligned} E_{t-1}(T_t|y_t) &= E_{t-1}(\tau_t y_t|y_t) + E_{t-1}(D_t|y_t) \\ &= \left[E_{t-1}(\tau_{t-1}) + \phi_{t\tau}^e\right] y_t + E_{t-1}(D_{t-1}) + \phi_{tD}^e \\ &= E_{t-1}(\tau_{t-1})y_t + E_{t-1}(D_{t-1}) \\ &= E_{t-1}(T_{t-1}|y_t). \end{aligned} \tag{14}$$

When combined with the first assumption, this result implies that

$$\begin{aligned} E_{t-1}(T_t|y_t) &= E_{t-1}(T_{t-1}|y_{t-1}) \\ &= T_{t-1}. \end{aligned} \tag{15}$$

Third, assume that, conditional on y_t, the loss of the CTC is the only *actual* source of tax change between years $t-1$ and t. That is, there are no other changes in either the MTR or the intercept of the tax schedule. This is a reasonable assumption given that we only focus on households with similar income in years $t-1$ and t. Analytically, this implies $T_t - (\tau_{t-1}y_t + D_{t-1}) = C_t$, where C_t is the amount of the CTC in year $t-1$ (that is lost in year t). When combined with the first assumption, we obtain

$$\Delta T_t = C_t. \tag{16}$$

Using the three assumptions and their implications (13), (15) and (16), it then follows that

$$\begin{aligned} E_t\left[(\tau_{t+1}, D_{t+1})^T\right] &= E_t\left[(\tau_t, D_t)^T\right] + (\phi_{t+1\tau}^e, \phi_{t+1D}^e)^T \\ &= E_t\left[(\tau_t, D_t)^T\right] \\ &= E_{t-1}\left[(\tau_t, D_t)^T\right] + \Gamma_t\left[T_t - E_{t-1}(T_t|y_t)\right] \\ &= E_{t-1}\left[(\tau_t, D_t)^T\right] + \Gamma_t \Delta T_t \\ &= E_{t-1}\left[(\tau_t, D_t)^T\right] + \frac{S_t(y_t,1)^T}{(y_t,1)S_t(y_t,1)^T}C_t, \end{aligned} \tag{17}$$

and hence

$$E_t(\tau_{t+1}) - E_{t-1}(\tau_t) = \frac{[S_t]_{11}\, y_t + [S_t]_{12}}{[S_t]_{11}\, y_t^2 + 2\,[S_t]_{12}\, y_t + [S_t]_{22}} C_t. \tag{18}$$

Combining this with (12) then gives

$$\widehat{\varepsilon} \simeq -\frac{\widehat{\beta}_1\,(1-\overline{\tau}_t)\left\{[S_t]_{11}\, y_t^2 + 2\,[S_t]_{12}\, y_t + [S_t]_{22}\right\}}{C_t\left\{[S_t]_{11}\, y_t + [S_t]_{12}\right\}} \tag{19}$$

Focusing on households with a similar level of taxable income in $t-1$ and t (see below), this estimate can be computed at various income levels y_t, $\overline{\tau}_t$ can be taken to be the actual MTR in year t, $\widehat{\beta}_1$ can be taken from the previous section and C_t is given by \$600 during 2001-2003, and by \$1,000 afterwards. However, in order to operationalize this estimate, for each chosen level of y_t we need to calibrate the matrix S_t. Recall that S_t is the variance of $\phi_t^e - \phi_t$, which, under our assumption that $\phi_t^e = 0$, is simply equal to the variance of $\phi_t \equiv (\Delta\tau_t, \Delta D_t)^T$. We therefore estimate S_t as the variance of $(\Delta\tau_t, \Delta D_t)^T$ by using SIPP data on households with a similar level of income in $t-1$ and t, with $t \in \{2002, 2003\}$ (the ACS data is not amenable to this analysis since we observe each household only in one year). However, in order to avoid the estimate of S_t reflecting an endogenous response to the loss of the CTC, we focus on households whose children are turning 17 only in 2004 and 2005.In order to compute D_t, we use data on taxable income y_t and data on the effective MTR τ_t and the actual tax liability T_t as computed by Taxsim. We then set $D_t = T_t - \tau_t y_t$, i.e., (τ_t, D_t) are the parameters of the linear segment of the tax schedule effectively faced by a household. We use both within-household and across-household variation in $(\Delta\tau_t, \Delta D_t)^T$ to estimate S_t. However, we do so separately for four taxable income classes as of year t (in 2003 dollars): between \$30,000 and \$50,000, between \$50,001 and \$66,000, between \$66,001 and \$81,000, and between \$81,001 and \$115000. These cutoffs are approximately equal to the quartile cutoffs of the taxable income distribution for households with nominal taxable income not exceeding \$110,000 in any of the years. Within these income classes, we evaluate the elasticity estimate at \$40,000, \$60,000, \$75,000 and \$93,000, respectively, which approximately correspond to the median taxable income within each income range. We conduct the estimation using four notions of "similarity" of taxable income in $t-1$ and t (in 2003 dollars): $|\Delta \ln y_t| \leq 0.05$, $|\Delta \ln y_t| \leq 0.1$, $|\Delta \ln y_t| \leq 0.15$ and $|\Delta \ln y_t| \leq 0.20$.

Table 5 presents the implied estimates of the uncompensated elasticity of labor income with respect to the net-of-tax-rate share. The estimates are based on $\widehat{\beta}_1 = -0.03$, the baseline estimate for the six-month window presented in Table 2.[41] The estimates for the income range \$30,000 to \$50,000 are in the range of 0.50 to 0.66. These are in the middle of the range of existing estimates of the elasticity of overall (labor and non-labor) taxable income with respect to the net-of-tax share based on conventional identification using panel data, which mostly range from 0 to 1.3 (see Gruber and Saez (2002) for a survey). The estimates for the income range \$66,001 to \$81,000 range from 1.18 to 1.36, which is toward the top of the conventional range of estimates. The

[41] We have also reestimated β_1 using SIPP data for households with a similar level of income in years $t-1$ and t. The resulting estimates are -0.027 if $|\Delta \ln y_t| \leq 0.05$ and -0.024 if $|\Delta \ln y_t| \leq 0.10$. Hence our results could be understood as upper bounds on the underlying elasticities.

estimates for the income range $81,001 to $115,000 are all above 3, hence very high compared to conventional estimates. Finally, estimates for the income range $50,001 to $66,000 are not only large in absolute value, but they also have the wrong sign. These elasticity calibrations are still a work in progress, and we hope to be able to provide further comment on these results in the next version of the paper.

7.2 Relation to the Permanent Income Hypothesis Literature

The results presented in the previous section are related to the tests of the permanent income hypothesis (PIH) conducted on a cross-sectional and panel data. The basic difference between this literature and our empirical application is that the former focuses on household consumption as the outcome variable, whereas we focus on household labor income, which is a proxy for the labor supply. However, both applications share the prediction that predictable lump-sum changes in disposable income should not, barring credit constraints, have any effect on the behavior of the household, be it consumption or labor supply.

The existing evidence on this prediction within the PIH literature is mixed. On the one hand, there are studies that do not find evidence in contrary to the PIH. For example, Browning and Collado (2001), using Spanish panel data, find that consumption does not respond to whether labor income is paid in equal instalments each month or with semi-annual "bonuses". Coulibaly and Li (2006), using the Consumer Expenditure Survey, examine the reaction of consumption to termination of mortgage payments and find no response either. On the other hand, there are studies that find that consumption increases with positive predictable income shocks, which is in contrary to the PIH. For example, using the Consumer Expenditure Survey, Parker (1999) analyzes the reaction of consumption to predictable after-payroll-tax income variation due to cross-sectional variation in social security withholding and due to within taxpayer variation in when the taxpayer hits the taxable income cap. He finds that consumption increases with after-payroll-tax income. Similar findings using various waves of the same survey are obtained by Souleles (1999) in reaction to the annual federal income tax refund, Souleles (2002) in reaction to *pre-announced* tax cuts in mid-1980s, Stephens (2003) in reaction to the timing of social security income, and Johnson et al. (2006) in reaction to the timing and amount of the 2001 tax rebate checks. The latter finding is also obtained by Shapiro and Slemrod (2003) using the Michigan Survey of Consumers. Using a different wave of the same survey, Shapiro and Slemrod (1995) find that consumers report a consumption increase in response to a pure forward timing shift in the after-tax income originating from the 1991-92 change in federal income tax withholding. Stephens

(2006), using the UK Family Expenditure Survey, finds that consumption does respond to the timing of paycheck arrival. His is the only study of the ones cited here that can fully account for the findings by the presence of liquidity constraints proxied by age or the level of (liquid) assets.

What can account for the variety these findings? One possibility is that the results may be to some extent driven by different samples and data sources used in different studies. However, Hsieh (2003) provides a striking evidence to the contrary. He documents that Alaskan households *do not* increase their consumption when paid from the (oil revenue-based) Alaska Permanent Fund, but the *very same* households *do* increase their consumption in response to the annual federal income tax refund. This finding implies that the heterogeneity of the PIH test results is unlikely to be driven by differences in samples. Rather, Hsieh comments on his results: "This evidence suggests that households will take anticipated income changes into account in their consumption decisions when the income changes are large, regular, and easy to predict, but will not do so when they are small and irregular." He also states: "...many tax and fiscal policy measures will probably have an effect on aggregate consumption as long as people find it difficult and costly to understand precisely how their incomes are affected by these policies."

The reasoning in Hsieh (2003) suggests that not all predictable changes in net income may in fact be predicted, or understood by taxpayers *ex ante*. However, it is silent about whether the source and nature of these changes is fully understood *ex post*. Indeed, if an imperfect ex ante understanding of the income change is driven by costs of gathering and processing information or by bounded rationality, then there is no reason to expect perfect ex post understanding either. This is because the cost of information or bounded rationality are present ex post as well as ex ante. The results reported in the previous section suggest that taxpayers do not appear to understand changes in their tax schedule ex post, in particular in relation to the change in the marginal tax rate.

The finding of ex-post misunderstanding has implications for interpreting the findings of the PIH literature as well. If consumers do not perfectly understand the source of a surprise in their disposable income ex post, their reaction may differ from the case when they do. For example, if a household receives an unexpectedly large tax refund check, then this may be due to having had too much money withdrawn from the household members' paychecks, but it may also be due to a decrease in taxes, either due to a tax cut, or due to a change in the household tax situation (e.g., a new child or a new mortgage). The former does not represent a change in lifetime resources of the household, whereas the latter does. As a result, whether or not a consumption increase should be interpreted as an "excess sensitivity" depends on the way households interpret a surprise in their disposable income ex post. We believe more work in this direction is necessary.

8 Conclusion

Due to the complexity of the income tax system, taxpayers may have difficulties recognizing their true marginal tax rate. As a result, they may turn to rules of thumb in approximating how much of an additional dollar of income is taken away in taxes. We present a formal model in which households have only a limited understanding of the tax schedule they face and update their estimate of the current year's marginal tax rate based on the previous year's unexpected innovation in the realized tax liability. This in general leads to ex post tax schedule misperceptions, particularly misperceptions about the MTR. Under the assumption that taxpayers react to perceived after-tax incentives as predicted by economic theory, we examine the validity of the misperception hypothesis by measuring taxpayer labor income responses to an exogenous and predictable variation in the tax liability due to losing eligibility for the Child Tax Credit when the child turns 17. The main advantage of this identification strategy is that variation in the tax liability is exogenous, lump-sum, and predictable well in advance. A fully rational and fully informed taxpayer should not react to such change. On the other hand, an imperfectly informed or a boundedly rational taxpayer may mistakenly believe that an increase in the household's tax bill reflects a higher MTR, resulting in a possible reduction of labor supply and labor income due to the conventional substitution effect. This CTC-induced variation in the tax liability therefore allows us to examine whether and to what extent taxpayers may not be fully informed or fully rational.

Using an identification strategy based on CTC eligibility discontinuity in age, we find that households who lose the credit due to having their child turn 17 at the end of a calendar year have a *lower* growth rate and level of parental labor income in the subsequent year compared to households that have their child turn 17 at the beginning of the following calendar year. This finding is robust to a variety of tests that include placebo effects at various other age and calendar cutoffs, a placebo effect based on pre-CTC data, and a triple-difference estimator that compares the effect of a child turning 17 around the end of the year before and after the CTC was introduced.

We interpret this finding as evidence for the presence of the substitution effect on labor supply and evidence for imperfect ex post understanding of the CTC loss. Taken at face value, this result suggests that tax policy changes that are not well-understood or predicted by the affected population, despite being predictable, may have unintended behavioral and welfare consequences.

In particular, changes that affect the level but not the slope of the tax schedule may result in a substitution effect that is not intended, hence increasing or reducing the deadweight loss relative to the full information case. On the other hand, changes that mostly affect the marginal tax rate may be partly interpreted as changes in the level of the tax schedule, with analogous implications for the deadweight loss. Complexity of the tax system may therefore interact with tax changes to create departures from conventionally understood welfare effects. This reasoning suggests that whenever households are likely to overshoot, relative to reality, their beliefs about the marginal tax rate, providing more and better information may be beneficial. On the other hand, just the opposite is the case when households are likely to undershoot.

The simple theoretical model presented in this paper leaves several open areas for future research. First, if households face a tax schedule about which they have imperfect knowledge, they may in principle experiment in order to obtain more information. That is, there may be a feedback effect from the choice of labor supply to the process of belief evolution over time. Second, as mentioned before, it is likely that the error variance in predicting changes in the tax schedule is determined endogenously by conscious information-gathering actions. For example, this variance can be reduced by investing time to learn about the tax code or hiring a tax advisor.

Likewise, our empirical findings invite a question of whether consumer purchasing decisions are affected by marginal price misperceptions and whether design consumer pricing schemes takes this into account. For example, whenever buyers overestimate the marginal price, providing more information benefits the seller, while just the opposite is the case if buyers underestimate it. For example, in policymaking, such considerations may have implications for designing regulations on information disclosure.

References

Auten, G. and R. Carroll, "The Effect of Income Taxes on Household Income," *Review of Economics and Statistics*, 1999, *81* (4), 681–693.

Autor, D., L. Katz, and M. Kearney, "Trends in U.S. Wage Inequality: Re-assesing the Revisionists," NBER Working Paper 11627, September 2005.

Blumkin, T., Y. Ganun, and B. J. Ruffle, "Are Income and Consumption Taxes Ever Really Equivalent? Evidence from a Real-Effort Experiment," Manuscript, Ben Gurion University, 2007.

Bound, J., D. A. Jaeger, and R. M. Baker, "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak," *Journal of the American Statistical Association*, 1995, *90* (430), 443–450.

Brown, C. V., "Misconceptions About Income Tax and Incentives," *Scottish Journal of Political Economy*, 1968, *15* (1), 1–21.

Browning, M. and M. D. Collado, "The Response of Expenditures to Anticipated Income Changes: Panel Data Estimates," *American Economic Review*, 2001, *91* (3), 681–692.

Buiter, W. H., "Macroeconomics of Pangloss - Critical Survey of the New Classical Macroeconomics," *Economic Journal*, 1980, *90* (357), 34–50.

Chetty, R., A. Looney, and K. Kroft, "Salience and Taxation: Theory and Evidence," NBER Working Paper 13330, August 2007.

Coulibaly, B. and G. Li, "Do Homeowners Increase Consumption After the Last Mortgage Payment? An Alternative Test of the Permanent Income Hypothesis," *Review of Economics and Statistics*, 2006, *88* (1), 10–19.

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

DeGroot, Morris, *Optimal Statistical Decisions*, McGraw-Hill, 1970.

Demery, D. and N. W. Duck, "The Theory of Rational Expectations and the Interpretation of Macroeconomic Data," *Journal of Macroeconomics*, 2007, *29* (1), 1–18.

Dickert-Conlin, S. and A. Chandra, "Taxes and the Timing of Births," *Journal of Political Economy*, 1999, *107* (1), 161–177.

Dokko, Jane, "The Effect of Taxation on Lifecycle Labor Supply: Results from a Quasi-Experiment," Manuscript, University of Michigan, October 2005.

Eissa, N. and J. B. Liebman, "Labor Supply Response to the Earned Income Tax

Credit," *Quarterly Journal of Economics*, 1996, *111* (2), 605–637.

Eissa, Nada, "Taxation and Labor Supply of Married Women: The Tax Reform Act of 1986 as a Natural Experiment," NBER Working Paper 5023, February 1995.

Feenberg, D. and E. Coutts, "An Introduction to the Taxsim Model," *Journal of Policy Analysis and Management*, 1993, *12* (1), 189–194.

Feige, E. L. and D. K. Pearce, "Economically Rational Expectations - Are Innovations in Rate of Inflation Independent of Innovations in Measures of Monetary and Fiscal Policy?," *Journal of Political Economy*, 1976, *84* (3), 499–522.

Feldstein, M., "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act," *Journal of Political Economy*, 1995, *103* (3), 551–572.

Finkelstein, Amy, "E-ZTax: Tax Salience and Tax Rates," Manuscript, NBER, February 2007.

Fujii, E. T. and C. B. Hawley, "On the Accuracy of Tax Perceptions," *Review of Economics and Statistics*, 1988, *70* (2), 344–347.

Goolsbee, A., "What Happens When You Tax the Rich? Evidence from Executive Compensation," *Journal of Political Economy*, 2000, *108* (2), 352–378.

Gruber, J. and E. Saez, "The Elasticity of Taxable Income: Evidence and Implications," *Journal of Public Economics*, 2002, *84* (1), 1–32.

Hsieh, C. T., "Do Consumers React to Anticipated Income Changes? Evidence from the Alaska Permanent Fund," *American Economic Review*, 2003, *93* (1), 397–405.

Johnson, D. S., J. A. Parker, and N. S. Souleles, "Household Expenditure and the Income Tax Rebates of 2001," *American Economic Review*, 2006, *96* (5), 1589–1610.

Kopczuk, W., "The Optimal Elasticity of Taxable Income," *Journal of Public Economics*, 2005, *84* (1), 91–112.

Kotlikoff, L. and D. Rapson, "Does It Pay, at the Margin, to Work and Save? Measuring Effective Marginal Taxes on Americans' Labor Supply and Saving," in James M. Poterba, ed., *Tax Policy and the Economy*, Vol. 21, The MIT Press, 2007, pp. 81–143.

Liebman, Jeffrey B. and Richard J. Zeckhauser, "Schmeduling," Manuscript, Harvard, October 2004.

Looney, Adam and Monica Singhal, "The Effect of Anticipated Tax Changes on Intertemporal Labor Supply and the Realization of Taxable Income," Manuscript, Harvard, November 2004.

Parker, J. A., "The Reaction of Household Consumption to Predictable Changes in Social

Security Taxes," *American Economic Review*, 1999, *89* (4), 959–973.

Poterba, J. M. and A. A. Samwick, "Taxation and Household Portfolio Composition: US Evidence from the 1980s and 1990s," *Journal of Public Economics*, 2003, *87* (1), 5–38.

Reis, R., "Inattentive Consumers," *Journal of Monetary Economics*, 2006, *53* (8), 1491–1530.

Romich, J. L. and T. Weisner, "How Families View and Use the EITC: Advance Payment versus Lump Sum Delivery," *National Tax Journal*, 2000, *53* (4), 1245–1265.

Saez, E., "Reported Incomes and Marginal Tax Rates, 1960-2000: Evidence and Policy Implications," in James M. Poterba, ed., *Tax Policy and the Economy*, Vol. 18, The MIT Press, 2004, pp. 117–173.

Shapiro, M. D. and J. Slemrod, "Consumer Response to the Timing of Income - Evidence from a Change in Tax Withholding," *American Economic Review*, 1995, *85* (1), 274–283.

__ **and** __ , "Consumer Response to Tax Rebates," *American Economic Review*, 2003, *93* (1), 381–396.

Souleles, N. S., "The Response of Household Consumption to Income Tax Refunds," *American Economic Review*, 1999, *89* (4), 947–958.

__ , "College Tuition and Household Savings and Consumption," *Journal of Public Economics*, 2000, *77* (2), 185–207.

__ , "Consumer Response to the Reagan Tax Cuts," *Journal of Public Economics*, 2002, *85* (1), 99–120.

Stephens, M., ""3rd of the Month": Do Social Security Recipients Smooth Consumption Between Checks?," *American Economic Review*, 2003, *93* (1), 406–422.

__ , "Paycheque Receipt and the Timing of Consumption," *Economic Journal*, 2006, *116* (513), 680–701.

A Appendix

Proof of Proposition 1. Choose an arbitrary time period $u \in \{0, ..., T\}$. Suppose that, based on the initial beliefs in period 0 and all the signals accumulated up until the end of period $u-1$, the household beliefs about (τ_{u-1}, D_{u-1}) are given by $N(\mu_{u-1}, \Sigma_{u-1})$. Due to expected changes in the tax schedule, the beliefs about (τ_u, D_u) at the beginning of period u are given by $N(\mu_{u-1} + \phi_u^e, \Sigma_{u-1} + S_u)$. Then the joint distribution of τ_u, D_u, and T_u is given by

$$\begin{pmatrix} \tau_u \\ D_u \\ T_u \end{pmatrix} \sim N \left\{ \begin{pmatrix} \mu_{u-1} + \phi_u^e \\ (y_u, 1)(\mu_{u-1} + \phi_u^e) \end{pmatrix}, \begin{bmatrix} \Sigma_{u-1} + S_u & (\Sigma_{u-1} + S_i)(y_u, 1)^T \\ (y_{iu}, 1)(\Sigma_{iu-1} + S_i) & (y_u, 1)(\Sigma_{u-1} + S_u)(y_u, 1)^T \end{bmatrix} \right\}. \tag{A-1}$$

Based on observing the realization of T_u, the posterior belief about (τ_u, D_u) is then given by (DeGroot, 1970)

$$\begin{pmatrix} \tau_u \\ D_u \end{pmatrix} \sim N \left\{ \mu_{u-1} + \phi_u^e + \Gamma_u \left[T_u - E_{u-1}(T_u | y_u) \right], \Sigma_u \right\}, \tag{A-2}$$

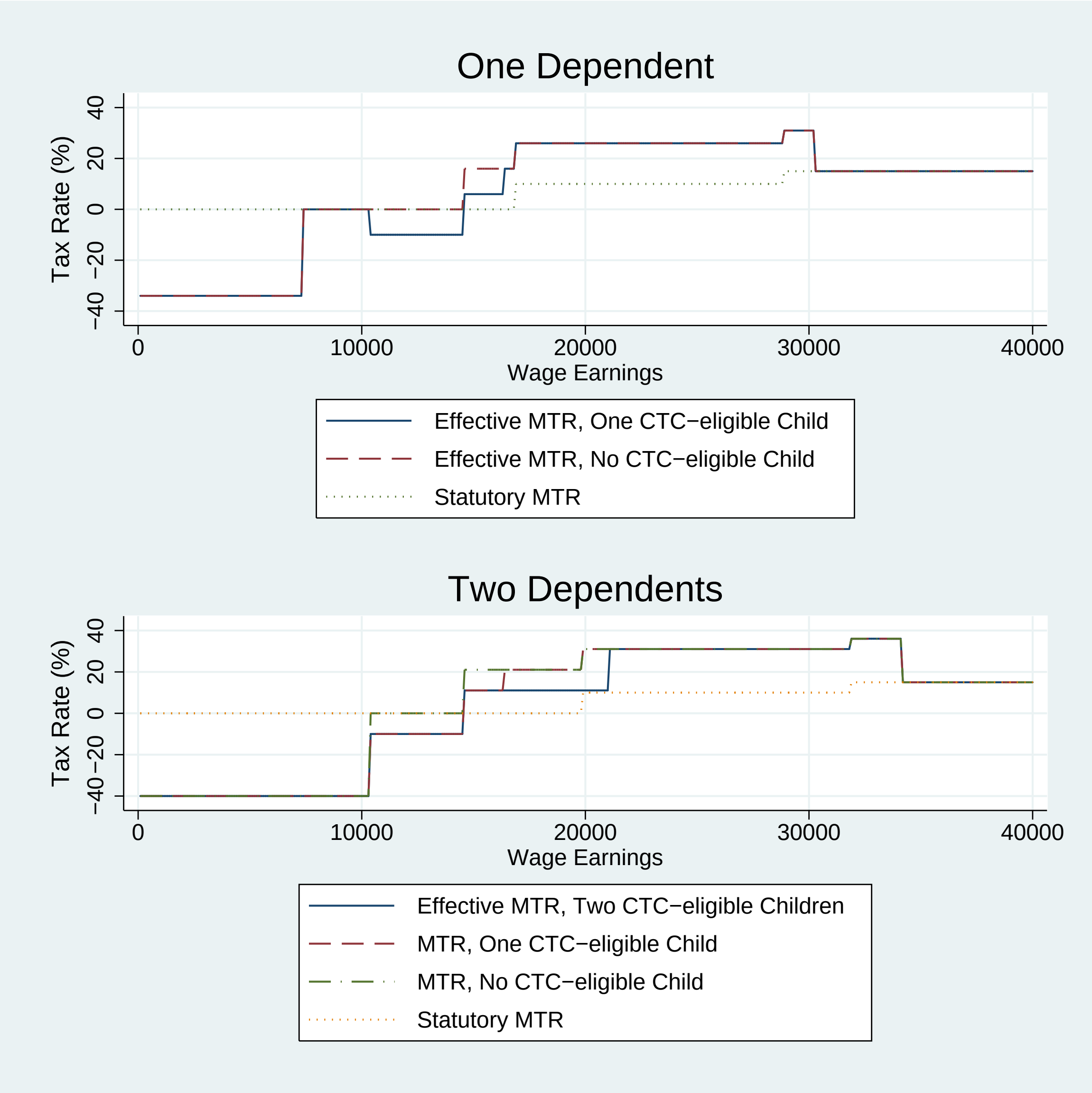
where Σ_u and Γ_u are given by (5) and (6). Recursive application of this formula then gives (3) for any $u \leq t$. For $u > t$, the mean and the variance of the beliefs are only affected by addition of independent increments of tax parameter changes. ■

Proof of Proposition 2. Note that (5) implies that $\Sigma_{t-1}(y_{t-1}, 1)^T = 0$, and hence

$$\begin{aligned} (\Sigma_{t-1} + S_t)(y_t, 1)^T &= \Sigma_{t-1}(y_t, 1)^T + S_t(y_t, 1)^T \\ &= \Sigma_{t-1}(y_{t-1}, 1)^T + \Sigma_{t-1}(\Delta y_t, 0)^T + S_t(y_t, 1)^T \\ &= \Sigma_{t-1}(\Delta y_t, 0)^T + S_t(y_t, 1)^T. \end{aligned}$$

Given (7), both elements of $S_t(y_t, 1)^T$ are strictly positive. As a result, if $|\Delta y_t|$ is small enough, the same sign pattern applies to the elements of $(\Sigma_{t-1} + S_t)(y_t, 1)^T$, and hence, by (6), also the elements of Γ_t. ■

Figure 1: Effective and Statutory MTR Schedules for Married Couples Filing Jointly in 2002



Notes:

1. The plots are based on the assumption that a household earns only labor income, does not itemize its deductions, and that all household members, including dependents, are eligible for a personal exemption.
2. Source: TAXSIM and Urban/Brookings Tax Policy Center.

Table 1: Summary Statistics for the Treatment and Control Groups

	A: SIPP 2001-2003 Sample				B: SIPP 1997-1998 Sample				C: ACS 2005-2007 Sample			
	Treatment		Control		Treatment		Control		Treatment		Control	
	Mean	S.d.	Mean	S.d.	Mean	S.d.	Mean	S.d.	Mean	S.d.	Mean	S.d.
Wife's Age	43.89	5.02	43.72	5.38	43.70	4.10	42.81	4.38	44.51	5.27	44.61	5.23
Husband's Age	46.12	5.29	45.83	5.49	46.72	4.35	44.52	5.50	46.60	5.71	46.72	5.69
Wife High School	0.37	0.48	0.34	0.48	0.33	0.47	0.38	0.49	0.27	0.44	0.27	0.44
Wife Associate Degree	0.34	0.48	0.33	0.47	0.30	0.47	0.34	0.48	0.22	0.42	0.23	0.42
Wife College	0.20	0.40	0.27	0.44	0.28	0.46	0.21	0.41	0.44	0.50	0.44	0.50
Husband High School	0.28	0.45	0.29	0.46	0.33	0.47	0.29	0.46	0.29	0.45	0.28	0.45
Husband Associate Degree	0.34	0.47	0.30	0.46	0.35	0.48	0.40	0.49	0.21	0.40	0.20	0.40
Husband College	0.27	0.45	0.31	0.46	0.30	0.47	0.26	0.44	0.42	0.49	0.43	0.50
Dependents	2.44	1.06	2.35	1.10	2.46	0.89	2.29	1.06	2.04	1.06	2.02	1.04
Wage Income	55,997	21,061	61,735	20,981	57,391	18,960	52,597	20,267	87,314	70,571	90,238	75,051
Observations	244		209		46		58		7,222		7,281	

Notes:

1. Unweighed means and standard deviations (in parentheses).
2. MTR refers to the effective marginal tax rate as computed by TAXSIM expressed in percentage points.
3. The summary statistics in Panels A and B are based on six-month time window.
4. The number of unique households is 442 in the 2001-2003 sample and 104 in the 1997-1998 sample.
5. Source: SIPP 2001-2003 and 1997-1998, ACS 2005-2007.

Table 2: The Effect of Losing CTC Eligibility on Parental Labor Income

	A: SIPP 2001-2003 Sample			B: ACS 2005-2007 Sample
Time Window +/-	6 Months	1 Month	12 Months	3 Months
	(1)	(2)	(3)	(4)
Treatment	-0.030	-0.009	-0.038	-0.030
	(0.059)	(0.860)	(0.001)	(0.021)
$t = 2002$	-0.021	-0.045	-0.014	
	(0.192)	(0.324)	(0.212)	
$t = 2005$				0.020
				(0.195)
$t = 2006$				0.073
				(0.000)
Observations	453	57	883	14,503
Number of Households	442	57	671	
R-squared	0.031	0.022	0.023	0.19

Notes:

1. Columns (1)-(3) present OLS estimates of (8) using the SIPP 2001-2003 sample, with $t \in \{2001, 2002\}$. Column (4) presents the OLS estimate of (9) using the ACS 2005-2007 sample, with $t \in \{2004, 2005, 2006\}$. In both cases, the additional control variables (estimates not displayed) are a constant, age and age squared of both parents, education level indicators of both parents, and the total number of dependents in the household.
2. P-values based on standard errors clustered at household level in parentheses.

Table 3: Post-CTC Placebo Effects

Event	Turning 15	Turning 16	Turning 17	Turning 18	Turning 19
Time Threshold	Dec. 31	Dec. 31	June 30	Dec. 31	Dec. 31
	(1)	(2)	(3)	(4)	(5)
SIPP 2001-2003	0.006	-0.009	0.007	0.045	
	(0.705)	(0.614)	(0.594)	(0.025)	
ACS 2005-2007	0.017	-0.013		0.020	0.007
	(0.227)	(0.315)		(0.141)	(0.714)

Notes:

1. The SIPP row presents OLS estimates of β_1 in (8) with $t \in \{2001, 2002\}$ based on a time window of +/- 6 months. The ACS row presents the OLS estimates of β_1 in (9) with $t \in \{2004, 2005, 2006\}$ based on a time window of +/- 3 months. In both cases, T_{it} is defined as an indicator for an event defined in the heading of the table.
2. The estimates in column (5) is based only on households where the 19-year-old dependent is a college student.
3. P-values based on standard errors clustered at household level in parentheses.

Table 4: Pre-CTC Placebo Effects and Triple Difference Estimates

	A: Pre-CTC Placebo Effects		B: Triple Difference	
Event	Turning 17	Turning 17	Turning 17	Turning 17
Time Window +/-	6 Months	12 Months	6 Months	12 Months
	(1)	(2)	(3)	(4)
Treatment_{it}	0.009	0.001	0.009	0.005
	(0.607)	(0.892)	(0.645)	(0.650)
$\text{Treatment}_{it} \times (t \geq 2001)$			-0.038	-0.042
			(0.126)	(0.012)
$(t \geq 2001)$			0.027	0.023
			(0.027)	(0.078)
Observations	104	192	557	1,079
Number of Households	104	192	546	867
R-squared	0.093	0.021	0.022	0.019

Notes:

1. Columns (1) and (2) present OLS estimates of (8) with $t = 1997$. Specifications (3) and (4) present OLS estimates of (10), with $t \in \{1997, 2001, 2002\}$. In all specifications, the additional control variables are (estimates not displayed) a constant, age and age squared of both parents, education level indicators of both parents, and the total number of dependents in the household.
2. P-values based on standard errors clustered at household level in parentheses.
3. Source: SIPP 2001-2003 and SIPP 1997-1998.

Table 5: Implied Uncompensated Elasticities of Labor Income with Respect to the Net-of-tax-share

	Upper limit on $\Delta ln(y_t)$			
Income Range (in 2003 $)	5%	10%	15%	20%
30,000-50,000	0.66	0.509	0.593	0.644
50,001-66,000	-10.387	-5.446	-10.001	-6.343
66,001-81,000	1.357	1.351	1.238	1.179
81,001-115,000	3.048	5.132	4.932	5.363