

Inertia and Overwithholding: Explaining the Prevalence of Income Tax Refunds

By DAMON JONES*

Over three-quarters of US taxpayers receive income tax refunds, which are effectively zero-interest loans to the government. Previous explanations include precautionary and/or forced savings motives. I present evidence on a third explanation: inertia. I find that following a change in tax liability, prepayments are only adjusted by 29 percent of the tax change after one year and 61 percent after three years. Adjustment increases with income and experience, and for EITC recipients, I rule out adjustment greater than 2 percent. Thus, policies affecting default-withholding rules are no longer neutral decisions, but rather, may affect consumption smoothing, particularly for low-income taxpayers.

JEL: D14, H24, K34

Keywords: Income Tax Refunds, Inertia, Withholdings

A growing body of evidence suggests that the behavior of a substantial share of the population deviates from what is typically assumed in economic theory (see Matthew Rabin 1998, Stefano DellaVigna 2008 for overviews). Recent studies have shown that departures from "standard" behavior may be particularly important in the field of public finance, especially when it comes to calculating the welfare effects of various policies (Douglas Bernheim and Antonio Rangel 2009, Raj Chetty, Adam Looney and Kory Kroft 2009). This paper presents new evidence on "non-standard" behavior in the public finance domain, based on US income tax withholding patterns.

Every year approximately 100 million taxpayers (nearly 80 percent) receive a tax refund because they have overwithheld taxes in the previous year. Overwithholding generates \$155 billion in annual income tax refunds—on average 7 percent of adjusted gross income (AGI) (Internal Revenue Service 2004). Interestingly, those with lower incomes are most likely to overwithhold. Members of this group have limited (or even zero) tax liability, pay relatively high interest rates to finance consumption until their refund arrives and in some cases pay additional fees to accelerate the delivery of the refund via refund anticipation loans (Alan Berube, Anne Kim, Benjamin Forman and Megan Burns 2002, Gregory Elliehausen 2005).

Previous studies have offered two main explanations for overwithholding: precautionary behavior in light of uncertain tax liability and asymmetric penalties (Jannett Highfill, Douglas Thorson and William V. Weber 1998) and "forced savings" arising from time-inconsistent preferences and/or mental accounting (Richard H. Thaler 1994, David

* University of Chicago, The Harris School and NBER, 1155 E. 60th St., Chicago IL, 60605, damon-jones@uchicago.edu. I thank Emmanuel Saez, Ulrike Malmendier and David Card for research guidance. I am grateful for additional feedback from Alan Auerbach, Raj Chetty, Stefano DellaVigna, Jim Hines, Patrick Kline, Botond Koszegi, Bruce Meyer and Joel Slemrod.

Neumark 1995, Lee Anne Fennell 2006). Such models typically assume that tax filers actively choose their withholdings and frequently readjust as incentives change. In contrast, I explore an additional explanation based on inertia (or incomplete adjustment). Specifically, I consider cases in which there is an external force or "shock" that changes the level of one's withholdings relative to one's tax liability, thus altering one's expected refund level. I subsequently observe to what extent tax filers respond to this external shock. I find that tax filers only partially adjust their withholdings, offsetting less than one-third of the change in their refund level after one year and less than two-thirds after three years.

I begin by exploiting exogenous variation in withholding levels brought about by a Presidential Executive Order. In 1992, the Bush administration reduced the default level of income tax withholdings for wage earners below a specified income threshold, with the aims of stimulating the economy (Matthew D. Shapiro and Joel Slemrod 1995). Importantly, the level of tax liability for this group remained constant. Thus, in the absence of a behavioral response, the policy would result in a reduction in the refund level or increased balance due for treated tax filers. Using the relationship between withholdings and allowances, I estimate the counterfactual level of withholdings absent any adjustment and compare this to actual levels of withholdings. I conclude that tax filers offset this policy by only 25 percent in its first year.¹

I then consider the relationship between the number of child dependents and the refund level, using a panel of tax returns from the years 1979 to 1990. In an event study framework, I identify the change in tax liability following a change in the number of child dependents. Estimating the subsequent change in tax prepayments yields another test of the inertia hypothesis. I find that prepayments are adjusted to offset 29 percent of the change in tax liability in the first year. Three years following the shock, prepayments have adjusted to account for 61 percent of the change in tax liability. I also find suggestive evidence of heterogeneity in responses. First, it appears that tax filers are more likely to adjust their withholdings when the loss of a dependent causes an expected refund to become a balance due. In addition, it appears that the adjustment rate is increasing in income and experience with changes in the number of child dependents.

Finally, I turn attention to the population eligible for the Earned Income Tax Credit (EITC)—a refundable tax credit that directly reduces tax liability (see V. Joseph Hotz and John K. Scholz 2003 for an overview). Overwithholding for this group is on average 12 percent of income, which, in combination with potential borrowing constraints, may hinder the ability to smooth consumption. To test for inertia among these tax filers, I make use of variation in tax liability generated by the dramatic expansion of the EITC over the last quarter century. Using repeated cross sections of tax return data, I estimate the relationship between expected EITC amounts and average tax prepayments. I show that differential growth in EITC levels is a strong predictor of relative refund levels,

¹Naomi Feldman 2008 uses this 1992 change in default withholdings as an instrument in identifying the effect of the timing of income on IRA savings. A key identifying assumption is that individuals do not undo the 1992 change in defaults, or rather, that tax filers are substantially inert. She shows evidence that withholdings are affected by the change in defaults. I complement her findings by decomposing this change into a mechanical effect and behavioral response and comparing the relative magnitude of the two.

which suggests that tax prepayments are not adjusted much in response to this particular reduction in tax liability. For every \$1 increase in the EITC, I can rule out a response greater than \$0.02 in reduced tax prepayments. Thus, there is little evidence of offsetting behavior on the part of tax filers in this group.

The empirical results can be combined with a simple model of withholding to gain a better understanding of tax filer behavior. First, it is theoretically possible that uncertainty with respect to tax liability can generate overwithholding. However, a standard model with uncertainty requires either a high level of risk aversion or extreme beliefs about the cost of an error in withholding to fit the data. Introducing time-inconsistent preferences provides a slightly better fit of the data for a model with uncertainty, but not by much. Alternatively, time-inconsistency in combination with a borrowing constraint may generate overwithholding. However, testing such a model of forced savings requires data on borrowing constraints, which are not available in the tax data used here. Finally, a model with costs of adjusting withholdings may explain the pattern of inertia observed. However, the prevalence of overwithholding rather than underwithholding requires more. If defaults are biased in favor of overwithholding, or if adjustment costs are asymmetric, then inertia may explain the observed patterns of withholding. As it turns out, in most cases the default withholding level is high, and the data weakly support an asymmetric response.

These findings have at least three implications. First, caution must be taken when using the observed levels of income tax refunds to generate inferences about preferences. For example, the prevalence of overwithholding has been cited as evidence of time-inconsistent preferences and/or mental accounting (Neumark 1995, Thaler 1994, Fennell 2006). However, the presence of inertia confounds such an interpretation.² Second, to the extent that defaults drive the behavior of inert tax payers, the decisions made by a social planner in setting default withholdings may no longer have neutral effects. Similar conclusions have been made in other arenas where default effects have been detected (Brigitte Madrian and Dennis Shea 2001, James Choi, David Laibson, Brigitte Madrian and Andrew Metrick 2003, Eric J. Johnson and Daniel Goldstein 2003, Alberto Abadie and Sebastien Gay 2006, Stefano DellaVigna and Ulrike Malmendier 2006). Default withholding rules in the US generally predispose individuals toward refunds. This is especially relevant for tax filers in the lower tail of the income distribution, where sizeable refundable credits and a possibly higher incidence of inertia result in a significant share of income that is overwithheld. This phenomenon may be purposeful, increasing savings for these tax filers. On the other hand, given the evidence on inertia, it might also be the case that default withholding rules generate inefficiently high amounts of tax prepayments and result in costly constraints on liquidity throughout the year.

The rest of the paper proceeds as follows. Section I explains the US income tax withholding system. Next, I present an empirical framework for studying inertia in Section II. I then describe the data used in this study and provide descriptive statistics on overwithholding in Section III. I present empirical results on inertia in Section IV, and Section V concludes with a discussion.

²This point is similarly made by Michael S. Barr and Jane K. Dokko (2007).

I. Institutional Details

In the US, individuals are taxed on income as they receive it, in a so-called "pay-as-you-earn" system. Throughout the year tax filers make prepayments either through withholdings, which are taken out of each paycheck, or through quarterly, estimated payments to the Internal Revenue Service (IRS), which typically account for non-wage sources of income. At the end of the year, annual income has been fully realized, and tax liability is determined. If tax prepayments are too low, the tax filer must pay the remaining balance, with a possible interest penalty. If prepayments are too high, tax filers receive a refund, although no interest is earned on the overpayments. Given the uncertainty involved, it may prove difficult to exactly equate prepayments to tax liability. Nevertheless, clear feedback is received every year with the filing of a tax return, in the form of a refund or balance due. Lower-income tax filers may qualify for refundable credits, which can result in an unavoidable refund.³

In a traditional employment setting, the employer automatically withholds tax prepayments for an employee each pay period. Employees determine the withholding amount using a W-4 form (IRS 2009b). Specifically, the W-4 form involves choosing a number of allowances, which roughly reflect the anticipated number of exemptions to be claimed on the tax return. The higher the number of allowances, the lower are one's withholdings per pay period. The W-4 form provides guidelines for choosing a number of allowances based on the major factors affecting tax liability: number of dependents, deductions, marital status and number of jobs. In addition to choosing a number of allowances, tax filers may designate an additional dollar amount to be withheld from each paycheck, allowing in theory for a continuous menu of withholding amounts. Using the employee's W-4 form, the employee's level of earnings and an IRS-provided withholding schedule, the employer then computes withholdings. A W-4 form can be resubmitted at any time should tax liability be expected to change but is generally only required at the onset of employment. In the event that an employee submits an incomplete W-4 or no W-4, the employer is required to choose zero allowances, resulting in the maximum level of withholdings (IRS 2009a). This default rule may help explain why prepayments are initially set high. The evidence I present below on inertia and asymmetric adjustment may help to explain why prepayments tend to remain high overtime.

II. An Empirical Model of Withholding

A. General Framework

I will now motivate the empirical analysis with the following simple model of income tax refunds. Consider the refund level:

³Until 2011, individuals could even offset refundable credits by claiming the Advance EITC. This option is now discontinued, but was available to tax filers in the data used here.

$$(1) \quad R(A, E; Z) = P(A, E; Z) - L(A, E; Z),$$

where $R(\cdot)$, $P(\cdot)$ and $L(\cdot)$ are the refund, tax prepayment, and tax liability level respectively.⁴ There are two endogenous determinants of prepayments and liabilities, A and E . These can be thought of as the number of allowances and earnings. Finally, there is an exogenous policy parameter Z , which may represent some feature of the tax code. Now consider the change in the refund level given a change in the policy parameter Z :

$$(2) \quad \frac{\partial R}{\partial Z} = \underbrace{\left(\frac{\partial P}{\partial A} - \frac{\partial L}{\partial A} \right) \cdot \frac{\partial A}{\partial Z} + \left(\frac{\partial P}{\partial E} - \frac{\partial L}{\partial E} \right) \cdot \frac{\partial E}{\partial Z}}_{\text{behavioral response}} + \underbrace{\left(\frac{\partial P}{\partial Z} - \frac{\partial L}{\partial Z} \right)}_{\text{mechanical effect}},$$

where the first two terms on the right-hand side constitute a behavioral response by the taxpayer and the third term, the mechanical effect, represents the direct effect of the policy change. I make the following simplifying assumptions, which are relevant to the types of policy changes that I consider:

ASSUMPTION 1: *Allowances do not affect tax liability:*

$$\frac{\partial L}{\partial A} = 0$$

ASSUMPTION 2: *Changes in tax liability and tax prepayments brought about by an earnings response are offsetting:*

$$\frac{\partial P}{\partial E} = \frac{\partial L}{\partial E}$$

ASSUMPTION 3: *The policy change either only affects tax prepayments or only affects tax liability:*

$$\text{either } \frac{\partial P}{\partial Z} = 0 \text{ or } \frac{\partial L}{\partial Z} = 0$$

The first assumption describes the nature of allowances. Adjusting the number of allowances only affects withholdings. The second assumption captures the nature of automatic withholdings. If earnings change, withholdings from the paycheck are automatically adjusted in much the same way as tax liability via tax withholding schedules.

⁴See the online appendix for a theoretical discussion of how the tax filer has arrived at this preferred level of income tax refund or balance due.

The marginal withholding rate is (approximately) the same as the marginal tax rate.⁵ The final assumption describes a feature of the policy changes under consideration. In each case, either the default withholding level changes with no accompanying change in tax liability, or vice versa. Using these assumptions, the change in refund level in Equation (2) simplifies to:

$$(3) \quad \frac{\partial R}{\partial Z} = \underbrace{\frac{\partial P}{\partial A} \cdot \frac{\partial A}{\partial Z}}_{\Delta P_B} + \underbrace{\frac{\partial P}{\partial Z}}_{\Delta P_M}$$

when the policy affects default withholdings, or

$$(4) \quad \frac{\partial R}{\partial Z} = \underbrace{\frac{\partial P}{\partial A} \cdot \frac{\partial A}{\partial Z}}_{\Delta P_B} - \underbrace{\frac{\partial L}{\partial Z}}_{\Delta L_M}$$

when the policy affects tax liability. Here again, the changes are decomposed into the behavioral response via tax prepayments, ΔP_B , and the mechanical effects on prepayments and liabilities, ΔP_M and ΔL_M , respectively. In measuring the tax filer's response to the policy change, consider the following two extreme cases:

CASE 1 (Full Adjustment): *Under **full adjustment** the agent adjusts prepayments to fully offset the policy change:*

$$\frac{\partial R}{\partial Z} = 0,$$

and thus equations (3) and (4) can be rearranged as follows to define the **adjustment rate**, α , i.e. the ratio of the behavioral response to the mechanical effect:

$$\alpha_P \equiv -\frac{\Delta P_B}{\Delta P_M} = 1$$

$$\alpha_L \equiv \frac{\Delta P_B}{\Delta L_M} = 1.$$

CASE 2 (Full Inertia): *Under **full inertia** the agent does not offset the policy change at all:*

$$\frac{\partial A}{\partial Z} = 0,$$

and thus the above adjustment rates become:

$$\alpha_P = \alpha_L = 0.$$

⁵This assumption may not hold for all tax filers, especially those married filing jointly. Therefore, I estimate prepayment adjustments separately for single and married tax filers in each of the analyses below.

In practice, I estimate these adjustment rates by regressing an observed change in tax prepayment level on the expected mechanical change in prepayments or liabilities. Variation in the mechanical change is brought about by some policy change or other shock to the refund level, Z . Though the methodological details vary slightly, I generally use some variation of the following specification:

$$(5) \quad \Delta P_B = -\alpha_P \cdot \Delta P_M(Z) + \Gamma \mathbf{X} + \varepsilon$$

when the policy affects prepayments and

$$(6) \quad \Delta P_B = \alpha_L \cdot \Delta L_M(Z) + \Gamma \mathbf{X} + \varepsilon$$

when the policy affects tax liability. The vector \mathbf{X} includes a group of control variables. The key identifying assumption is that conditional on \mathbf{X} , the policy variable Z does not directly affect the underlying target refund level, and thus only affects tax prepayments via a change in default prepayments or tax liability.

B. Specific Applications

I use the preceding framework to estimate an adjustment rate, α , in three different settings. In each case, there is a unique shock that affects the expected refund level. I subsequently observe the taxpayers' response to this event. In the Section IV below, I outline the key features of the different sources of identification. In one case, the 1992 change in default withholdings, there is a change in default withholdings while holding liability constant. In the other two cases, the panel study of child dependents and the EITC expansion, tax liability changes without a compensating adjustment of default withholdings. In each case, I use a different econometric approach. I relate each of these approaches to the general empirical framework described above and also highlight the direction (up or down) in which the shock pushes the refund level in the absence of a behavioral response.

III. Data Description

A. Data Overview

The data used in this analysis come from the IRS Statistics of Income (SOI) Division. For almost every year since 1960, the IRS has released a public-use sample of income tax returns. Sample sizes range from 80,000 to 150,000. In addition to selected cross sections of the IRS public-use file, I use a panel of tax returns from the same source. The University of Michigan IRS Tax panel follows a subset of tax filers from 1979 to 1990. This unbalanced, longitudinal data set contains about 45,000 observations for the first three years, and then between 10,000 and 20,000 observations in each year thereafter. The data contain detailed information on sources of income, and include most of the information provided on the Form 1040 tax return. Most importantly, the data include tax

prepayments, disaggregated into withholdings from wages and estimated tax payments, tax liability and the level of refund/balance due. Demographic information is limited to marital status, number of children, other dependents and an indicator for age equal to or above 65 years. In Sections IV.A and IV.C I use repeated cross sections, while in Section IV.B I use panel data.

B. Descriptive Statistics

I provide summary statistics on overwithholding for tax filers in 2004 in Column (1) of Table 1. On average, individuals receive a refund of \$1,000 and the median ratio of prepayment to tax liability is 1.26. In addition, refunds comprise 7 percent of AGI for the average tax filer. Finally, the share of tax filers receiving a refund is just below 80 percent. Panel A of Figure 1 depicts a skewed right distribution of refunds in 2004 that visually reinforces the summary statistics. One may notice the mass of filers at a zero balance. This is mainly comprised of individuals with both zero tax liability and zero tax prepayments.⁶

[INSERT TABLE 1]

Further visual evidence reveals two significant patterns of overwithholding. First, individuals claim less than the total number of allowances to which they are entitled and are also clustered at zero allowances, which is the default level set for workers by employers. Panel B of Figure 1 illustrates this concept. The figure shows an estimated distribution of implied allowances along side a counterfactual distribution of allowances for wage earners. The number of implied allowances is estimated as follows. Consider an individual with a given level of wage and salary earnings, E^i . Using the withholding tables in IRS Publication 15, I can calculate the expected level of withholdings, $P(A, E^i)$, conditional on that person choosing a given level of allowances A . Next, I use observed withholdings, P^i , to infer the number of allowances that were likely chosen, i.e. $\hat{A}_i = \arg \min_A |P^i - P(A, E^i)|$. I label this the "implied" number of allowances. Repeating for all tax filers, I can estimate the discrete distribution of allowances, $\hat{F}(A)$. Because this process relies on a tight relationship between wages, allowance and withholdings, I restrict analysis to tax filers with more than 95 percent income in wages and salary that take the standard deduction.⁷

[INSERT FIGURE 1]

Next, I use demographic information from the tax return to calculate the total number of allowances that would be chosen if the individual strictly followed the instructions on IRS form W-4. I label this the number of "counterfactual" allowances. This comparison suggests that the level of overwithholding is not simply a mechanical artifact of

⁶This discontinuity in the distribution at a zero balance may also be evidence of tax evasion.

⁷See the online appendix for further details on this estimation procedure.

the IRS withholding tables. Had taxpayers simply followed the instructions on the W-4, they would typically have a higher number of allowances and, thus, a lower level of withholdings.

We also see in Panel C of Figure 1 that refunds are persistent. Here I use the 1979-1990 Michigan panel of tax filers, calculate the share of time that a refund is received for each individual and plot the distribution of this statistic. To reduce noise, I restrict analysis to those with at least 3 years in the panel. Contrary to the idea that individuals may fluctuate between under- and overwithholding, nearly half of all tax filers always receive a refund.

IV. Results

A. 1992 Change in Default Withholdings

In his 1992 State of the Union Address, President Bush announced a decrease in default withholdings aimed at stimulating a sluggish economy (Shapiro and Slemrod 1995). New withholdings tables were issued in February of that year and employers were instructed to incorporate the new tables as soon as possible (IRS 1992). The withholdings were reduced by as much as \$187 and \$423 per job for single and married wage earners with taxable wages below \$64,000 and \$110,000 respectively.⁸ Panel A of Figure 2 provides a stylized illustration of the change in withholdings. Importantly, there was no concurrent reduction in tax liability. Within the framework presented of Section II, Z corresponds to the default withholding rules. There is no change in tax liability due to the policy change, $\partial L/\partial Z = 0$, and thus I am estimating the adjustment rate α_P . The mechanisms, A , by which individuals offset the policy are (1) submitting a new W-4 with a lower number of allowances to raise withholdings or (2) increasing estimated payments. For this analysis, I use repeated cross section data from the IRS SOI public use samples for the years 1991 and 1992. Analysis here is restricted to tax filers who take the standard deduction and have more than 95 percent of income in wage and salary earnings. These tax filers represent 45 percent of the population. Those with other significant sources of income and itemized deductions may choose allowances in a different manner than what is assumed below, and thus are not included.⁹ Table 1, Column (1) provides descriptive statistics on the sample used. In terms of income and refund propensity, this sample falls somewhere between the general population of tax filers and the EITC population.

[INSERT FIGURE 2]

Taxpayers are made aware of the 1992 policy change through two main avenues. First, individuals receive a higher after-tax paycheck every pay period once the employer implements the change in withholdings tables. Shapiro and Slemrod (1995) find that about

⁸These amounts are presented in terms of year 2000 dollars and represent the maximum changes. Actual changes may vary for individuals in the phase-in or phase-out region of the withholding adjustment, as depicted in Figure 2, Panel A.

⁹Those with income above \$200,000 are also dropped, as their income data is additionally blurred in the IRS SOI data set. The high income tax filers represent less than 1% of tax filers. Additional details on sample selection are included in the online appendix.

one-third of survey respondents noticed a reduction in withholdings a month after the policy took effect. Second, when the tax return is filed, the tax filer should receive a lower refund or owe a higher balance than usual. In addition, employers were instructed to directly notify their employees of the change in withholdings, and also to instruct them on how to offset the reduction in withholdings. The new Employer's Tax Guide reads

If some of your employees do not want their withholding changed, they should complete new Forms W-4 (IRS 1992, pg. 1).

In comparison to the other shocks that I analyze, this policy change generates downward pressure on the refund level. In the absence of adjustment, the tax filer will be more likely to owe a balance at the end of the year.

I use information on the relationship between withholdings, wages and allowances to arrive at an estimate of α_P . This method of estimating the mechanical effects, behavioral responses and adjustment rates requires the following three elements:

- $P_0(A_0^i, E^i)$: baseline withholdings prior to the policy change
- $P_1(A_0^i, E^i)$: withholdings following the policy change, holding allowances fixed
- $P_1(A_1^i, E^i)$: withholdings after the policy change and change in allowances.

where withholdings, $P(\cdot)$, are a function of allowances, A^i , and wage and salary earnings, E^i , as described in IRS withholding tables. The 0 and 1 subscripts denote pre- and post-policy variables respectively, for the i th individual in 1992. I observe post-policy withholdings and earnings, and thus can infer the distribution of post-policy allowances. However, I do not observe pre-policy 1992 withholdings and thus cannot make direct inferences regarding pre-policy allowances, A_0^i . Therefore, I make the following assumption:

ASSUMPTION 4: *In the absence of the policy change, the distribution of allowances would have remained constant between 1991 and 1992:*

$$F_0(A_0)|_{t=91} = F_0(A_0)|_{t=92}$$

If this holds, I can estimate the distribution of allowances in 1991 and use this as a proxy for the pre-policy distribution of allowances in 1992. Using the same method described in Section III.B, I estimate the discrete distribution of allowances in 1991 generating $\hat{F}_0(A)|_{t=91}$. In order to gain predictive power, this is done separately for tax filers grouped by marital status and income quantile, indexed by the vector θ^i .¹⁰ In other words, for a married couple earning \$100,000 in 1992, the distribution of allowances among married filers with similar income in 1991 is likely a more relevant counterfactual

¹⁰To be exact, the data are split into 16 quantiles of AGI, by marital status. Additional details regarding the estimation of these distributions are provided in the online appendix.

than the unconditional distribution of allowances among married and single filers of all incomes in 1991. I similarly use data from 1992 to estimate the distribution of post-policy allowances in 1992. Under Assumption 4 above, I arrive at estimates of the conditional distributions, $\hat{F}_0(A_0|\theta)$ and $\hat{F}_1(A_1|\theta)$ for individuals in 1992. Using these conditional distributions, I simulate withholdings for individuals in 1992 as follows:

$$(7) \quad \hat{P}_0^i(A_0^i, E^i) = \int P_0(a, E^i) d\hat{F}_0(a|\theta^i)$$

$$(8) \quad \hat{P}_1^i(A_0^i, E^i) = \int P_1(a, E^i) d\hat{F}_0(a|\theta^i)$$

$$(9) \quad \hat{P}_1^i(A_1^i, E^i) = \int P_1(a, E^i) d\hat{F}_1(a|\theta^i).$$

For a given individual, then, the mechanical effect and behavioral response are defined as follows:

$$(10) \quad \Delta P_M^i = \hat{P}_1^i(A_0^i, E^i) - \hat{P}_0^i(A_0^i, E^i)$$

$$(11) \quad \Delta P_B^i = \hat{P}_1^i(A_1^i, E^i) - \hat{P}_1^i(A_0^i, E^i).$$

The former isolates the change in the withholdings driven by the policy-induced change in the withholding schedule, $P(\cdot)$, while the latter isolates changes in withholdings driven by the behavioral change in allowances chosen. Finally, I use the estimated mechanical effects and behavioral responses in the following regression:

$$(12) \quad \Delta P_B^i = -\alpha_P \cdot \Delta P_M^i + x^i \beta + \varepsilon^i,$$

where x includes a cubic polynomial in tax liability. For this procedure I report both standard errors clustered within each income-by-marital cell and bootstrap standard errors.

Panel B of Figure 2 lends credence to this method. The graph shows the estimated distribution of allowances from 1990 to 1993, using the same methods as in Figure 1. First, we see that the distribution is relatively stable between 1990 and 1991, suggesting that in the absence of a policy change, the distribution of allowances would have remained constant from 1991 to 1992. We also see that the distribution shifts in 1992 in the direction toward lower allowances and thus higher withholdings, which would be expected of individuals attempting to offset the policy change. This is consistent with a behavioral response beginning in 1992.

In Table 2, I estimate the fraction by which this behavioral response offsets the mechanical effect of the policy shock. Using Equation (10), I estimate an average mechanical decrease in withholdings of \$237, with conditional averages of \$181 and \$392 for single and married filers respectively. In contrast, I estimate an average behavioral response of only \$57 in additional withholdings using Equation (11). Estimating Equation (12), this translates into an estimate of 0.25 for α_P . Tax filers only offset 25 percent of

the decrease in withholdings during the first year of the policy change.

[INSERT TABLE 2]

One concern may be that these estimates are biased due to differential trends in prepayments across the affected and non-affected groups. To address this, I estimate Equations (10), (11) and (12) using data from years prior to the policy change. Hypothetical mechanical effects are imputed for individuals in 1990 and 1991 based on the rules of the 1992 policy change. These "placebo" estimates of α_P will pick up preexisting differences in withholding trends among those in the earlier years who would have been affected by the 1992 policy change. As seen in Table 2 Column (3), the placebo estimates are indistinguishable from zero.

B. Panel Study of Child Dependents

I further explore inertia by estimating the adjustment of tax prepayments in response to a change in the number of child dependents and therefore tax liability. Adding a child increases the number of exemptions that a taxpayer can claim, reducing taxable income. In addition, tax credits such as the EITC become available for households within certain income ranges. Thus, when one either loses or gains a child dependent, tax liability will rise or fall in a predictable manner. Returning to the general empirical framework, the so-called policy variable, Z , is now the number of child dependents. While there is a change in tax liability via the number of exemptions claimed when filing taxes, the automatic withholding from wages does not adjust unless a new W-4 form is filed. We have a case where $\partial P / \partial Z = 0$, and I am therefore estimating α_L .

Following a change in the number of child dependents, tax filers receive direct feedback on the change in tax liability when the tax return is filed. The loss or gain of a child will result in a lower or higher refund level, respectively. In addition, if a new W-4 form is filed for any reason, the tax payer is explicitly directed to take into account any changes in the number of children that are claimed (IRS 2009b). Within this context, I can directly compare the adjustment prepayments after being pushed toward a refund or toward owing a balance.

To examine this phenomenon I use panel data on tax returns spanning 1979 to 1990 from the Michigan IRS Tax Panel. I perform an event study of the loss or gain of a child dependent, and thus restrict analysis to taxpayers who undergo a change in the number of dependents while in the panel. In particular, I include all observations that are within a seven-year window of an observed change in child dependents. I exclude those with income above \$200,000, as their number of children is censored, and those over age 65, since they rarely have child dependents. The remaining subsample represents 23 percent of the population covered by this panel.¹¹

¹¹In the case of multiple events, individuals are assigned to the event with the nearest proximity in time. Those with both a loss and gain in the same seven-year window are excluded. Additional details regarding the sample selection are included in the online appendix.

In Column (3) of Table 1, we see that, compared to the other cases that I consider, this sample has higher incomes, in part due to the restriction in data to tax filers that at some point have dependents. While 84 percent of the changes in child dependents from year to year involve one child, I pool all changes, which may include two or more dependents lost or gained. Losses and gains are equally likely to occur in the sample. Nonetheless, losses and gains of children may not be directly comparable events. The former tends to happen later in the life cycle. Furthermore, the loss of a child may be commonly preceded by a divorce or negative shock to income. For example, Figure 3 shows a dip in earnings at the time of a loss of a dependent, which is most likely associated with a concurrent divorce. I discuss below control variables designed to capture the effects of income and marital dynamics.

[INSERT FIGURE 3]

MAIN ESTIMATES. — I will estimate the adjustment rate within an event study framework, where the event is a change in number of child dependents. I estimate the following system of equations over a seven-year window surrounding the event:

$$(13) \quad liability_{it} = \sum_{j=-3}^3 \Delta L_l^j \cdot Loss_{i,t-j} - \sum_{j=-3}^3 \Delta L_g^j \cdot Gain_{i,t-j} + \eta_i + \gamma_t + \Gamma \mathbf{X}_{it} + \varepsilon_{it}$$

$$(14) \quad prepayment_{it} = \sum_{j=-3}^3 \Delta P_l^j \cdot Loss_{i,t-j} - \sum_{j=-3}^3 \Delta P_g^j \cdot Gain_{i,t-j} + \tilde{\eta}_i + \tilde{\gamma}_t + \tilde{\Gamma} \mathbf{X}_{it} + \tilde{\varepsilon}_{it}$$

Each observation is indexed by individual, i , and time t . The \mathbf{X}_{it} are vectors of time-varying characteristics: a 10-piece linear spline in income interacted with marital status, a similar spline in lagged income, marital status, lagged marital status and a dummy for transitions from married to single.¹² The $Loss_{i,t-j}$ and $Gain_{i,t-j}$ are a set of dummy variables indicating that at time t a change in dependents has taken place j periods in the past, $j \in [-3, 3]$. I then estimate an adjustment rate for each of the three years after the

¹²The sensitivity of the estimates to functional form is explored in the online appendix Table A.1. The estimates change very little whether one uses a simple linear income control or a spline, whether lagged income is included, whether marital status alone is included or the more flexible specification here, and whether one includes individuals over age 65 or not. Furthermore, the results are robust to restricting the sample to those who have a constant marital status throughout the panel. The inclusion of time trends in event time does result in a modest increase in the adjustment rate estimates, as seen in online appendix Table A.5. Finally, the unbalanced nature of panel results in missing lagged variables for a subset of the data. I impute these lagged variables for this subset, but show in appendix Table A.5 that dropping these observations has a negligible effect on the estimates. Moreover, dropping lagged variables altogether leaves the results unchanged.

event using the following nonlinear constraint:

$$(15) \quad \alpha_L^j = \frac{\Delta P_l^j}{\Delta L_l^j} = \frac{\Delta P_g^j}{\Delta L_g^j}, \quad j \in \{1, 2, 3\}.$$

Equations (13) and (14), under the constraint in (15), are estimated via feasible generalized nonlinear least squares (FGNLS).¹³

The parameters ΔL^j and ΔP^j measure the change in liabilities or prepayments between event year 0 and event year j . As such, the coefficients ΔL_l^j and ΔL_g^j in equation (13) can be thought of as the mechanical effect on current tax liability of a change in child dependents j periods ago for losers and gainers respectively. Likewise, the coefficients ΔP_l^k and ΔP_g^k in (14) can be thought of as the behavioral response by taxpayers. In Table 3 I summarize these changes using the following weighted averages:

$$(16) \quad \Delta L_M^j = \pi_l \cdot \Delta L_l^j + \pi_g \cdot \Delta L_g^j$$

$$(17) \quad \Delta P_B^j = \pi_l \cdot \Delta P_l^j + \pi_g \cdot \Delta P_g^j,$$

where π_l and π_g are the share of losers and gainers in the sample. The ΔL_M^j and ΔP_B^j are measures of the average mechanical effect and behavioral response. Finally, the parameter α_L^j measures the relative magnitude of the two. Put another way, α_L^j is the response of tax prepayments to changes in tax liability, driven by a change in number of dependents j periods ago.

Identification of α_L^j is illustrated in Panel B of Figure 3. The graph shows the level of tax liability and prepayments in a seven-year window around the event, with the event year normalized to zero. Outcomes are shown separately for losers and gainers. The solid line plots the coefficients from Equations (13) and (14), and thus are adjusted for the η_i , γ_i and \mathbf{X}_{it} . The horizontal axis measures event time and the vertical axis measures outcomes relative to the year in which the number of child dependents changes. There is a sharp increase in tax liability when a dependent is lost. The inverse is true for gains in dependents. However, we see prepayments do not change as sharply. One thing to note is that there are pre-existing trends in event time, which most likely are due to life-cycle effects correlated with gaining or losing a child dependent. Unfortunately, a continuous measure of age is not available in the tax data, which would help to control for this. Nonetheless, the sharp change in tax liability at event time zero appears to represent a break from this trend.¹⁴

In Table 3, I report the point estimates underlying these figures. As can be seen in Column (1), a change in the number of dependents translates into an immediate change in tax liability of about \$524 dollars. This change in tax liability persists over the next

¹³Specifically, from (15), we have that $\Delta P_k^j = \alpha_L^j \Delta L_k^j$, $k \in \{l, g\}$. Substituting for ΔP_k^j in (14) imposes the nonlinear restriction across the two equations.

¹⁴Indeed, one can control for the trend in event time using the pre-event data. Results from controlling for a trend in event time are available in online appendix Table A.5, and result in a slightly more accelerated adjustment over time.

three years. In Column (2) of Table 3, we see that the response of tax prepayments is not as large: \$154 following a change in the number of dependents. This response gradually increases over time. The adjustment rates estimated from Equation (15), α_L^j , are reported in Column (3). In the first year following the change in tax liability the adjustment rate is 0.29. Tax prepayments do not fully adjust; three years after the change in dependents, only 61 percent of the shock has been offset. It is important to note that the sample is an unbalanced panel. The construction of the Michigan tax panel is such that it is not common for an observation to have missing years. As such, the difference in adjustment rates across years may signify a gradual increase in adjustment, but it may also be driven by sample imbalances over the seven-year window.¹⁵

[INSERT TABLE 3]

HETEROGENEITY IN ADJUSTMENT RATES. — Though the results thus far demonstrate that tax filers have a limited response to changes in tax liability or default withholdings, inertia alone does not explain a bias toward refunds. What is also needed is possibly a bias in default withholdings or asymmetry in the inertia. One can examine the asymmetry hypothesis by separately estimating adjustment rates for those who lose a child and those who gain a child and seeing whether the adjustment rate is larger for the former group. Table 4 presents adjustment rates separately for losers and gainers in Columns (2) and (3). The two groups have fairly similar responses to changes in tax liability. Thus, evidence of an asymmetric response does not show up for the general sample.

An alternative conjecture is that tax filers respond differently to increases and decreases in tax liability when near a balance of zero. Given that most tax filers initially have excess withholding, we may not pick up the effect of a zero balance in the general population. Thus, in Columns (4) and (5), I restrict the sample to tax filers that have an initial refund level or balance due less than \$1,000, a so called "Zero Balance" sample. For this sample, a loss of a child dependent is likely to cause a tax filer who had previously received a refund to owe a balance due. The inverse is true for a tax filer in this sample who gains a child. Now, losers have an adjustment rate between 0.93 and 1.89 in the first three years following a change in dependents, while gainers' adjustment rates are indistinguishable from zero.

[INSERT TABLE 4]

The results are consistent with the idea that transitioning from receiving a refund to owing a balance is particularly noticeable to tax filers and prompts a larger response. However, there are other, equally plausible explanations. First, there may be unobserved differences between the general population and the Zero Balance sample and

¹⁵It would be ideal to use a balanced sample. However, given the construction of the Michigan Tax Panel it is not uncommon for observations to drop in and out of the sample over time. Restricting analysis to a subset that is present the entire 7 years leaves only 10% of the original sample. Appendix Table A.2 shows that point estimates for this smaller subsample are higher, but much more imprecisely estimated.

across losers and gainers.¹⁶ For example, it may be the case that the opportunity cost of time increases for gainers and decreases for losers. This pattern may be more pronounced for those in the Zero Balance sample, who tend to have lower income.

Table 4 explores other possible dimensions of heterogeneity. In Columns (6) and (7) the adjustment rate is estimated separately for those with below- and above-median income in the year of the event. The point estimates suggest that high-income filers adjust more and increasingly so over time, though the confidence intervals overlap between the two groups. Similarly, we see in Columns (8) and (9) that tax filers appear to adjust more rapidly the more experience they have with a change in number of children.¹⁷

C. EITC Expansion

In the final case, I use the expansion of the EITC as a source of variation in tax liability. Introduced in 1975, the EITC is a tax credit available to low income, working households. The earning subsidy may constitute as much as 40 percent of income, with a maximum benefit of \$5,751 in 2011. The maximum earnings thresholds are \$43,998 for single filers with three or more children, \$40,964 for single filers with two children, \$36,052 for single filers with one child and \$13,660 for single filers with no children. For married couples, the earnings threshold is relaxed by an additional \$5,080. The credit is refundable—meaning once it has reduced tax liabilities to zero, the remaining credit is paid out as a transfer (see V. Joseph Hotz and John K. Scholz 2003 for an overview). The maximum EITC amount nearly tripled during the 1990s, growing from \$1,255 in 1990 to \$3,888 in 2000 (Committee on Ways and Means 2004). For eligible households, this created a significant downward trend in tax liability over the same period. However, IRS withholding tables do not account for EITC eligibility, and the W-4 form used to determine withholdings makes no explicit mention of the need to adjust withholdings in expectation of an EITC refund.

In terms of the general framework for inertia, the policy variable, Z , is now the level of the EITC for eligible tax filers. In this case, there is a change in tax liability but no accompanying change in withholding defaults: $\partial P/\partial Z = 0$ and I am again estimating α_L . The mechanism, A , for offsetting the policy is again the lowering of withholdings through the W-4 form or the lowering of estimated payments. Individuals could have also signed up for Advance EITC payments in order to offset the change in tax liability, though Damon Jones (2010) shows that very few tax filers typically made use of this option. Note that this approach differs from using general changes in tax liability, for many types of tax liability are accounted for in updated withholdings tables. This is not true for the EITC.

The frequency of feedback provided by the EITC is generally at the annual level. Over

¹⁶In online appendix Table A.3, I compare the demographics of these groups. Gainers tend to have lower incomes than losers, probably due to life cycle effects. Furthermore, the "Zero Balance" sample is comprised of lower incomes than the general population.

¹⁷Note that we only measure the first observed change in child dependents in the sample, which need not be the first change experienced by tax filers. Online appendix Table A.4 presents additional results on heterogeneity. Comparing single and married tax filers yields results that are less conclusive.

time, eligible households are presented with larger and larger refunds. Further signals of EITC expansion may result from the marketing and outreach efforts of tax preparers, both free and commercial, who encourage eligible households to file a tax return and claim the EITC. A complete understanding of the connection between the EITC and tax liability, however, may yet be elusive for recipients. For example, EITC recipients generally do not bunch at kink points in the EITC schedule (Emmanuel Saez 2010), though explicitly informing individuals about the schedule may increase bunching (Raj Chetty and Emmanuel Saez 2009). As compared to the other cases under consideration, the EITC expansion drives eligible tax filers toward receiving a larger refund in the absence of any behavioral response.

To estimate the effect of the EITC on prepayments, I make use of repeated cross sections of tax return data from 1980 to 2004. I restrict analysis to the group of tax filers eligible for the EITC, where EITC eligibility is defined based on one's income and the number of child dependents.¹⁸ Next, tax filers are further split into three groups: EITC-eligible tax filers with zero children, one child, or two or more children. In order to account for changes in group composition that occur due to changes in EITC eligibility rules over time, income variables are inflation-adjusted to year 2000 levels and EITC eligibility is defined based on year 2000 criteria using the National Bureau of Economic Research (NBER) Internet TAXSIM model.¹⁹ Note that eligibility depends on the level of income and the number of children present, so that, for example, tax filers with no children become ineligible for the EITC at a much lower income level than those with two or more children. Thus, we can expect systematic differences in income levels across the groups. To account for this, I control for group level fixed effects below, which capture time-invariant differences between the groups. Next, I calculate group-by-year averages and estimate the following model:

$$(18) \quad \bar{P}_{gt} = \eta_g + \gamma_t - \alpha_L \cdot \overline{EITC}_{gt} + \Gamma \cdot \bar{X}_{gt} + \varepsilon_{gt},$$

where g indexes the three groups, t is a year index, the η 's and γ 's are group and year fixed effects, \overline{EITC}_{gt} is the average EITC received by group g in year t and \bar{X}_{gt} is a vector of average observable controls including a cubic in income, tax liability, and the child tax credit. The outcome, \bar{P}_{gt} , measures average tax prepayments for group g in year t . There is a negative sign in front of α_L since $\partial liability / \partial EITC = -1$. Given the presence of time and group dummies, the coefficient α_L is identified off of variation in EITC levels over time, within eligibility groups.

As shown in Column (4) of Table 1, this sample represents a little more than a quarter of the entire tax filing population and occupies a lower segment of the income distribution than the tax filers in the previous two cases. As such, the costs of overwithholding may be the greatest for this group, especially if they are facing liquidity constraints. It is surprising, then, that these tax filers are particularly prone to overwithholding. We see

¹⁸For more details on EITC eligibility rules during this time period, see Committee on Ways and Means (2004).

¹⁹For more on the TAXSIM model see Daniel Feenberg and Elizabeth Coutts (1993) or visit the NBER website at <http://www.nber.org/taxsim/>.

in Table 1 that the median tax prepayments for this group is more than twice as much as tax liability. This ties up an average of 12 percent of income in overwithholdings throughout the year. As I will show, this high propensity to overwithhold is in part due to the interaction of growing tax credits and high levels of inertia.

[INSERT FIGURE 4]

As demonstrated in the Panel A of Figure 4, the credit underwent significant expansions during the early 1990s, especially for families with two or more children. I use this variation in tax liability to test for inertia in prepayments. Panel B of Figure 4 illustrates a strong positive correlation between EITC levels and refund levels across the groups and over time. This visual evidence suggests that there was little to no adjustment of tax prepayments in response to increases in EITC levels. In Panel C of Figure 4, I have plotted tax prepayments over the same time period. Tax prepayments do not appear to decline in response to the EITC increases. During the 1990s, when the EITC underwent its most pronounced growth, the level of tax prepayments among eligible tax filers is relatively flat. In 1992 there are noticeable declines in prepayments, which, as has been shown, is due to a 1992 Executive Order. Included in this graph for comparison are a group of low-income tax filers who do not qualify for the EITC.²⁰ To the extent that there is a decreasing trend in tax prepayments, it is nearly identical for eligible and non-eligible tax filers. This underscores the notion that changes in tax prepayments over this period were not in response to EITC growth.

Table 5 reports the coefficients estimated from Equation (18). After controlling for a cubic in income, tax liability and the child tax credit, the change in tax prepayments in response to EITC growth is not statistically significant. Controlling for group or time fixed effects does little to change this result, nor does splitting the sample into married and single tax filers. Thus, there is strong evidence of nearly full inertia with respect to EITC growth. I can rule out an adjustment rate, α_L , larger than 0.02.

[INSERT TABLE 5]

V. Discussion and Conclusion

I observe estimates of an adjustment rate that range from nearly 0 in the case of the EITC to about 0.29 in the first year following a change in the number of dependents. The effect of these shocks on the refund level appears to persist for some time. In Table 3 we see that 3 years after a change in the number of dependents, tax filers appear to adjust prepayments by only 61 percent of the change in liability. There is also limited evidence of heterogeneity. First, there are results that suggest an asymmetric response of tax filers when going from receiving a refund to owing a balance due. This is found when comparing adjustment rates among tax filers who lose or gain a child dependent. When focusing on tax filers near the threshold of a zero balance due, the former group

²⁰These tax filers are not included in the regressions below that are ultimately used to test for inertia.

exhibits a larger adjustment rate. In addition, tax filers appear to adjust more quickly the more experience they have with a change in tax liability.

Another pattern that emerges is that inertia is greatest among the lower-income population. First, there is evidence consistent with the idea that higher income tax filers respond more quickly when adjusting withholdings in response to a change in child dependents. Secondly, the adjustment rates among the EITC eligible population are particularly low. This is made clear in Table 5, Column (4) where I rule out an adjustment rate greater than 0.02. Note that the low adjustment rates in the case of the EITC may either be due to the low income of the sample used or the specific nature of the EITC. Nonetheless, these results are intriguing given the fact that the benefit of reducing withholdings is likely to be the greatest among lower-income tax filers, who may face liquidity constraints. At the same time, the cost of adjusting withholdings and uncertainty with respect to tax liability may also be the greatest among this group, which may more than outweigh the benefits. In any event, defaults will tend to affect outcomes the most for this group.

A model of withholding may shed some further light on tax filer behavior when combined with the empirical results above.²¹ The simplest model that generates overwithholding is a standard model with uncertain tax liability. However, calibrations reveal that in order to match the observed odds of overwithholding, one either needs a very high level of risk aversion or extreme beliefs regarding the cost of a withholding error. The fit of this model can be approved by allowing for time-inconsistent preferences, but not by much. An alternative approach combines time-inconsistent preferences with borrowing constraints to generate forced savings via overwithholding. However, a test of forced savings requires data not here available on whether a tax filer faces a borrowing constraint. A final approach introduces costs to adjusting withholding. Such a model is consistent with the presence of inertia highlighted above. However, inertia alone does not predispose tax filers to over- or underwithholding. Asymmetric adjustment costs may explain an overwithholding bias. However, the results above only show such an asymmetry for a subset of tax filers. Alternatively, the fact that default withholding levels tend to be high in conjunction with inertia may help explain withholding patterns.

The evidence that I have documented has two additional implications. First, the observed prevalence of income tax refunds is traditionally attributed to precautionary behavior in response to uncertain tax liability or commitment savings resulting from time-inconsistency. However, to disentangle these alternative theories, one must first account for the inertia that partially breaks down the link between outcomes (refund levels) and active decisions (preferences). Second, the presence of inertia changes the interpretation of default withholding rules designed by a social planner. If taxpayers fully and frequently adjust their withholdings, defaults are essentially neutral. However, the evidence presented here suggests that these default withholdings rules may actually affect outcomes such as the timing of income and perhaps the ability to smooth consumption.

Policy makers have at different times attempted to capitalize on the inertia and low salience of withholdings. A leading case is the 1992 Executive Order mentioned above. This policy relied on the assumption that taxpayers would not undo a withholdings

²¹ See the online appendix for more details on the various models and results.

change and furthermore spend the extra income despite having to owe back the money at the years end. Survey evidence suggests that about 43 percent did indeed do just that (Shapiro and Slemrod 1995). The American Recovery and Reinvestment Act of 2009 included a tax credit that is disbursed via a reduction in withholdings, which is coupled with an equal reduction in tax liability. It has been argued that distributing stimulus payments via withholdings is more likely to stimulate demand than one-time rebate checks, since the former is less salient or subject to different mental accounting rules (James Surowiecki 2009), though Claudia R. Sahm, Matthew D. Shapiro and Joel B. Slemrod (2010) find evidence to the contrary. Most recently, the state of California increased withholdings in November 2009 by 10 percent, with no accompanying increase in tax liability. The state's explicit aims are to fill budgetary gaps in the short run via zero-interest loans from wage earners (Shane Goldmacher and W. J. Hennigan 2009). Again, the policy hinges on the assumption that taxpayers will not readjust their withholdings. In each of these cases, the affect of withholding policies on taxpayers can vary greatly.

If these withholding-based policies have differential effects, one may wonder what the distribution of costs is across incomes. The costs depend on the distribution of consumer debt, investment opportunities and credit access, which can be estimated with the Survey of Consumer Finances (SCF). I use the 2004 SCF data to impute interest rates for taxpayers in the 2004 IRS SOI data set. Next, I calculate the opportunity cost of overwithholding in terms of lost interest, which serves as a lower bound for the cost. Table 6 shows that these costs are fairly modest at an average of \$63 per year. At the other extreme, overwithholding can be much more costly if individuals cannot borrow or draw on savings to smooth consumption. As an upper bound on the cost, I calculate the welfare loss of an uneven consumption profile throughout the year due to overwithholding. Depending on the curvature of utility, these costs can be of an order of magnitude larger, with an average of about \$1,000 as seen in Table 6. These types of costs, which stem from credit constraints, are most relevant for lower-income groups. As a share of income, the costs of consumption smoothing range from 14 percent among individuals in the bottom quintile of income to 1 percent for the top quintile.²²

[INSERT TABLE 6]

It is also worth noting that the status quo of a refund-biased withholding system is by no means a universal phenomenon. Consider the Working Tax Credit (WTC), the UK analog of the EITC. The WTC, similarly a tax credit for low-income workers, is disbursed on a bi-weekly or monthly basis, and thus its timing is more similar to the Advance EITC in the US (Mike Brewer, Emmanuel Saez and Andrew Shephard 2010). An interesting question, then, is why and how have the UK and US systems come to be so different in the timing of refundable credits? Furthermore, do UK taxpayers share the same affinity for large income tax refund payments? In the presence of strong preferences for large refunds, we would expect to observe many UK workers demanding a lump sum payment in lieu of the more frequent WTC. However, this does not appear to

²²More details on the imputation of interest rates and calculation of consumption smoothing costs are provided in the online appendix.

be the case in the UK (Brewer, Saez and Shephard 2010). Thus, identifying preferences over large refunds and determining the optimal setting of withholding defaults remains an open debate. In light of the findings presented in this study, future inquiry into the subject must account for the presence of inertia.

TABLE 1—DESCRIPTIVE STATISTICS

	(1)	(2)	(3)	(4)
	Sample Used In Analysis			
	Tax Year 2004	1992 Change in Default Withholdings	Panel Study of Child Dependents	EITC Expansion
<i>Adjusted Gross Income</i>				
10th Percentile	4,516	3,176	9,581	1,938
Median	27,047	18,118	37,205	8,146
90th Percentile	89,965	46,577	84,295	23,390
Mean	46,745	22,027	43,793	10,423
Standard Deviation	329,022	17,903	35,597	21,176
<i>Refund</i>				
Mean	1,070	775	932	1,157
Median	747	555	905	587
<i>Prepayment to Liability Ratio</i>				
10th Percentile	0.60	0.93	0.79	0.44
Median	1.26	1.29	1.23	2.20
90th Percentile	3.31	3.40	2.38	11.32
<i>Refund to AGI Ratio</i>				
Mean	0.07	0.06	0.04	0.12
Median	0.03	0.04	0.03	0.08
<i>Refund Probability</i>	0.79	0.88	0.80	0.93
<i>Time Period</i>	2004	1991-1992	1979-1990	1990-2004
<i>Share of Total Filers</i>	1.00	0.45	0.23	0.27
<i>N</i>	150,047	32,049	62,604	249,650

Note: Descriptive statistics are estimated for Columns (1), (2) and (4) using IRS SOI public use files and the Michigan tax panel for Column (3). Dollar amounts are reported in year 2000 levels.

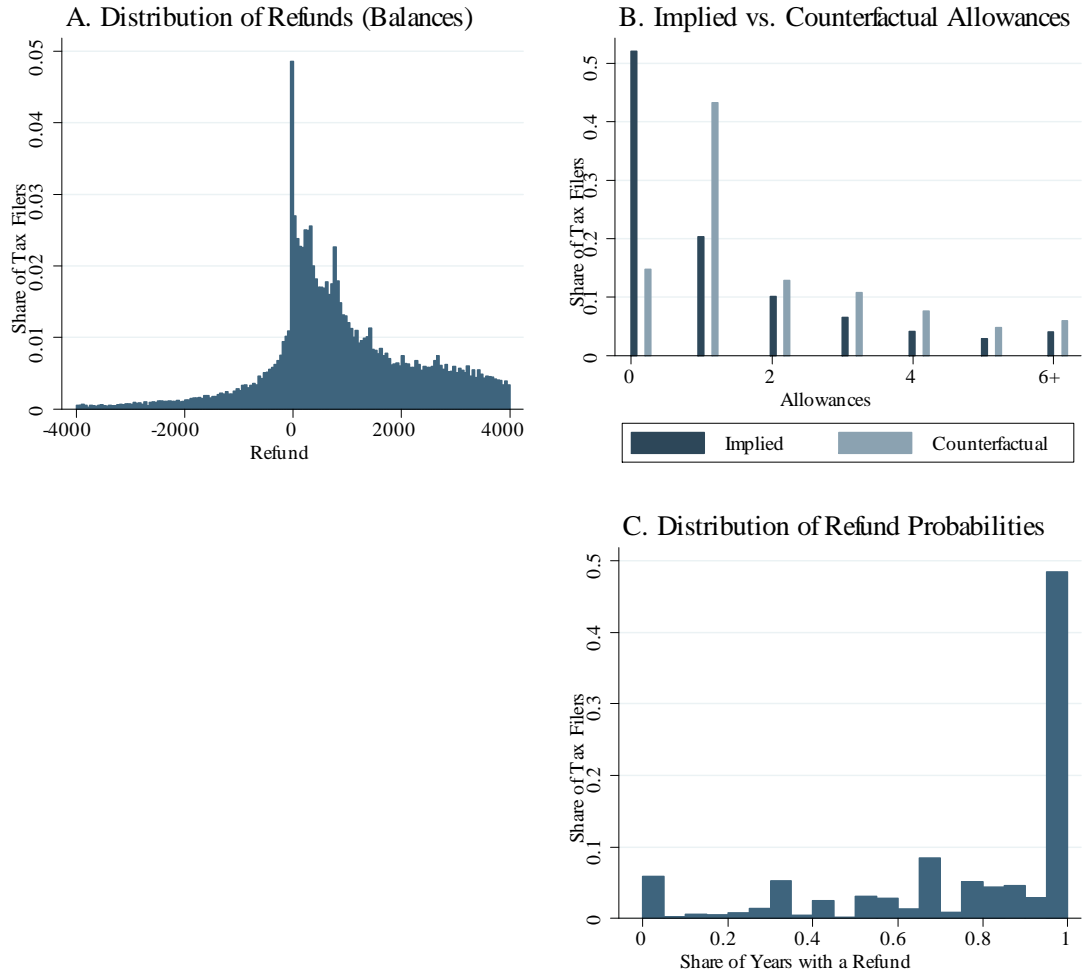


FIGURE 1. DESCRIPTIVE STATISTICS

Note: (A) The distribution of refunds and balances are for US tax filers in 2004, taken from the IRS SOI public use file. (B) The implied number of allowances is estimated using the amount of withholdings reported on the tax return in conjunction with wages, marital status, AGI and IRS withholding tables. See Section III.B for further details on this procedure. Counterfactual allowances were calculated using income and demographic information reported on the tax return in conjunction with the instructions on the W-4 form. Data are for US tax filers in 2004, taken from the IRS SOI public use file. The sample is restricted to tax filers with more than 95 percent of income from wages, who used the standard deduction and had an AGI of less than \$200,000. (C) The figure presents the distribution of individual refund probabilities for US tax filers from 1979-1990, estimated using the Michigan tax panel. Analysis is restricted to individuals with at least three years of data.

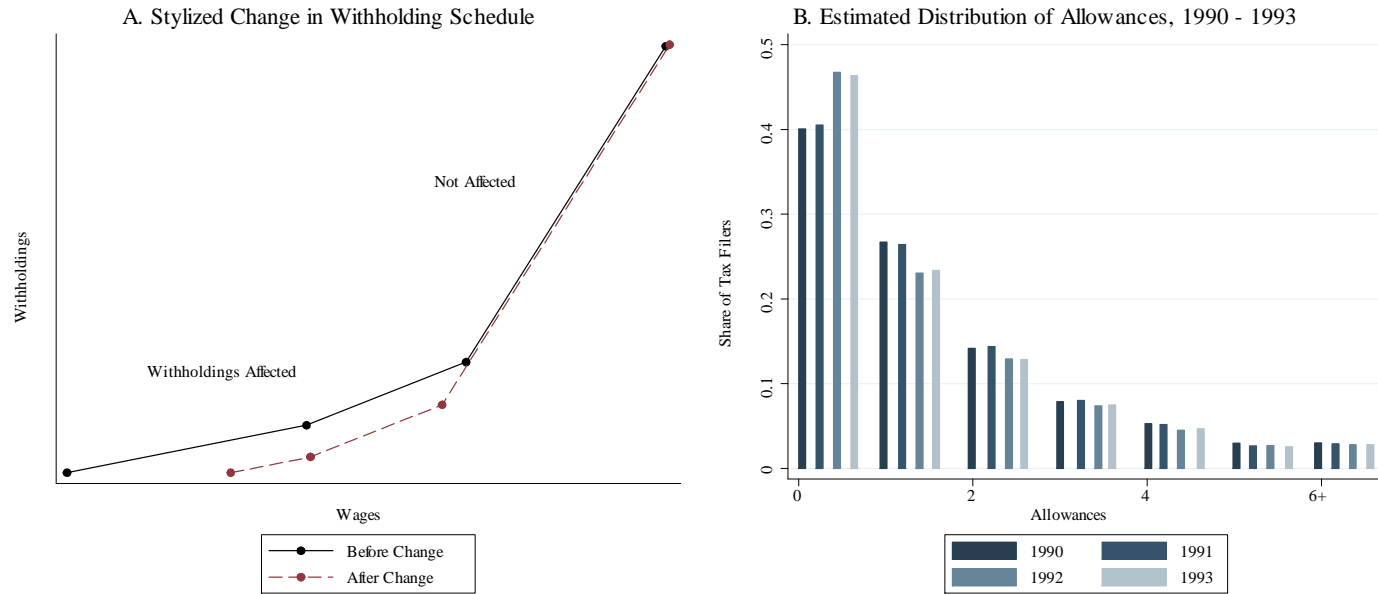


FIGURE 2. EFFECT OF DEFAULT CHANGE IN WITHHOLDINGS

Note: (A) Stylized demonstration of the adjustments made to withholding tables following the 1992 Executive Order. Not Drawn to Scale. (B) The distribution of allowances claimed by US tax filers for the years 1990 to 1993 is estimated using IRS SOI public use files. See Section III.B for further details on this procedure. The sample is restricted to tax filers with more than 95 percent of income from wages, who used the standard deduction and had an AGI of less than \$200,000.

TABLE 2—1992 WITHHOLDINGS CHANGE - MECHANICAL EFFECT, BEHAVIORAL RESPONSE AND ADJUSTMENT RATE ESTIMATES

	(1)	(2)	(3)
	Mechanical Effect: ΔP_M	Behavioral Effect: ΔP_B	Adjustment Rate: α_P
Full Sample	237.10 {18.23}*** [18.03]***	56.81 {8.85}*** [8.87]***	0.25 {0.07}*** [0.08]***
<i>N</i>	14,083	14,083	14,083
Single	180.88 {13.36}*** [13.2]***	45.14 {9.84}*** [9.65]***	0.30 {0.05}*** [0.14]**
<i>N</i>	9,284	9,284	9,284
Married	391.74 {17.08}*** [16.74]***	88.91 {16.37}*** [16.22]***	0.15 {0.06}** [0.18]
<i>N</i>	4,799	4,799	4,799
Placebo 1 1991	274.03 {23.09}*** [22.75]***	13.83 {7.61}* [7.35]*	0.05 {0.05} [0.05]
<i>N</i>	17,966	17,966	17,966
Placebo 2 1990	235.90 {18.03}*** [17.83]***	-11.91 {8.72} [8.42]	0.07 {0.06} [0.07]
<i>N</i>	16,058	16,058	16,058

Note: Mechanical effects, behavioral responses and adjustment rates are estimated using Equations (10), (11) and (12). Data are from the repeated cross sections of the 1989-1992 IRS SOI public use files. The sample is restricted to tax filers with more than 95 percent of income originating from wages or salary. The mechanical effect for Placebo samples reports the hypothetical effect had the policy change taken place in earlier years. Standard errors, clustered at the income group-by-marital status level are reported in braces, while bootstrap standard errors are reported in brackets. One, two and three stars denote statistical significance at the 10%, 5% and 1% level respectively. Dollar amounts are reported in year 2000 levels.



FIGURE 3. CHANGE IN LIABILITY AND PREPAYMENTS ASSOCIATED WITH A CHANGE IN CHILD DEPENDENTS

Note: The expected change in AGI, tax liability and prepayments at the time of a change in dependents is estimated using the Michigan tax panel spanning 1979 to 1990. Coefficients are obtained in an event study regression, as specified in Equations (13) and (14).

TABLE 3—CHANGE IN CHILD DEPENDENTS - MECHANICAL EFFECT, BEHAVIORAL RESPONSE AND ADJUSTMENT RATE ESTIMATES

	(1)	(2)	(3)
	Mechanical Effect: ΔL_M	Behavioral Response: ΔP_B	Adjustment Rate: α_L
Year 1	523.72 (24.91)***	154.08 (34.79)***	0.29 (0.06)***
Year 2	559.82 (33.70)***	267.46 (43.26)***	0.48 (0.08)***
Year 3	574.59 (46.42)***	350.17 (55.87)***	0.61 (0.10)***
<i>N</i>	62,604	62,604	62,604

Note: Estimates of mechanical effect, behavioral response and adjustment rate are obtained using Equations (13)-(15). Data are from the Michigan tax panel spanning 1979-1990. Controls include a 10-piece linear spline in income by marital status, a similar spline in lagged income, marital status, lagged marital status, a dummy for transitions from married to single, individual and time fixed effects. Robust standard errors are reported in parentheses. One, two and three stars denote statistical significance at the 10%, 5% and 1% level respectively. Dollar amounts are reported in year 2000 levels.

TABLE 4—CHANGE IN CHILD DEPENDENTS - HETEROGENEITY IN ADJUSTMENT RATES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Baseline	Full Sample		Zero Balance Sample		Low Income	High Income	1st Event	2nd+ Event
		Loss	Gain	Loss	Gain				
Year 1									
α_L	0.29 (0.06)***	0.20 (0.15)	0.37 (0.13)***	0.54 (0.17)***	-0.13 (0.08)	0.21 (0.07)***	0.33 (0.09)***	0.26 (0.07)***	0.38 (0.14)***
Year 2									
α_L	0.48 (0.08)***	0.58 (0.22)***	0.56 (0.21)***	1.29 (0.34)***	-0.30 (0.17)	0.34 (0.12)***	0.56 (0.09)***	0.44 (0.09)***	0.55 (0.15)***
Year 3									
α_L	0.61 (0.10)***	1.22 (0.70)*	0.66 (0.21)***	1.71 (0.57)***	-0.21 (0.20)	0.43 (0.09)***	0.74 (0.14)***	0.55 (0.14)***	0.72 (0.16)***
<i>N</i>	62,604	30,517	32,087	11,554	15,585	28,523	34,081	48,058	14,546

Note: Estimates of the adjustment rate are obtained using Equations (13)-(15). The "Zero Balance" sample is restricted to tax filers with a refund or balance due less than \$1,000 in the base year. "Low" and "High" income are defined relative to the median, event-year level of income. The "1st Event" sample include the first observed event for an individual, while the "2nd+ Event" sample includes all subsequent events observed. Data are from the Michigan tax panel spanning 1979-1990. Controls include a 10-piece linear spline in income by marital status, a similar spline in lagged income, marital status, lagged marital status, a dummy for transitions from married to single, individual and time fixed effects. Robust standard errors are reported in parentheses. One, two and three stars denote statistical significance at the 10%, 5% and 1% level respectively. Dollar amounts are reported in year 2000 levels.

TABLE 5—EITC EXPANSION - ADJUSTMENT RATE ESTIMATES

	(1)	(2)	(3)	(4)
	Adjustment Rate: α_L			
Full Sample	-0.34 (0.03)***	0.05 (0.04)	0.00 (0.04)	-0.06 (0.04)*
Share of EITC filers	1.00	1.00	1.00	1.00
Single	-0.32 (0.03)***	0.05 (0.03)*	0.06 (0.04)	0.00 (0.03)
Share of EITC filers	0.83	0.83	0.83	0.83
Married	-0.12 (0.06)**	0.06 (0.03)**	0.06 (0.03)*	0.01 (0.03)
Share of EITC filers	0.17	0.17	0.17	0.17
Controls	No	Yes	Yes	Yes
Group Fixed Effects	No	No	Yes	Yes
Year Fixed Effects	No	No	No	Yes
<i>N</i>	72	72	72	72

Note: The effect of EITC changes on prepayments is estimated using data for US tax filers from the years 1981-2004. Tax filers are aggregated by year into three groups of EITC-eligible tax filers depending on number of children. Controls include a cubic in AGI, level of child tax credit and tax liability. Robust standard errors are reported in parentheses. One, two and three stars denote statistical significance at the 10%, 5% and 1% level respectively.

TABLE 6—AVERAGE PRIVATE COST OF INCORRECT WITHHOLDING BY INCOME QUINTILE

	(1)	(2)	(3)	(4)	(5)	(6)
	Adjusted Gross Income					
	\$0 to \$11,010	\$11,010 to \$22,650	\$22,650 to \$39,530	\$39,530 to \$69,590	Above \$69,590	Full Sample
Interest Costs:	\$20	\$78	\$87	\$91	\$41	\$63
Consumption Smoothing Costs:						
$\gamma = 1$	\$389	\$850	\$572	\$523	\$1,051	\$677
$\gamma = 2$	\$525	\$1,166	\$870	\$843	\$1,643	\$1,009
$\gamma = 3$	\$546	\$1,314	\$1,051	\$1,052	\$2,018	\$1,196
$\gamma = 4$	\$589	\$1,405	\$1,177	\$1,214	\$2,334	\$1,343

Note: The first row reports the cost of overwithholding in terms of lost interest. The next four rows report the equivalent variation of deviating from a constant monthly consumption profile to one where income timing is distorted by overwithholding and agents face borrowing constraints. The four cases vary by the curvature of utility as parameterized by the coefficient of relative risk aversion, γ . More details are provided in the online appendix.

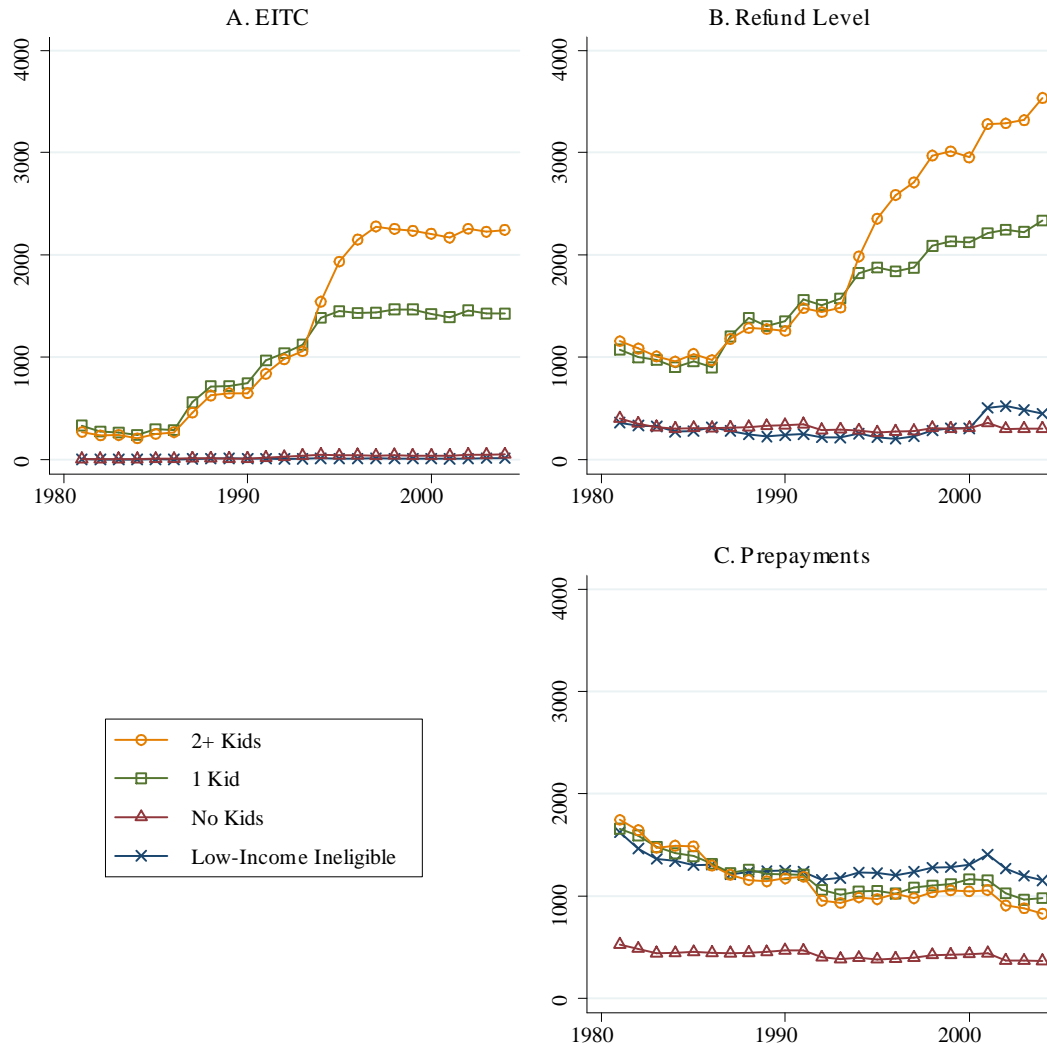


FIGURE 4. MEAN EITC, REFUND LEVEL AND PREPAYMENTS, 1980-2004

Note: Mean EITC, refund levels and prepayments are estimated for US tax filers from 1981 - 2004, using repeated cross sections of IRS SOI public use files. The first three categories include individuals who qualify for the EITC and have zero, one, or two or more children. The fourth category, "Low-income Ineligible" corresponds to individuals who have AGI below 75 percent of the maximum EITC income threshold and who do not qualify for the EITC for some other reason (e.g. age below 25, too much investment income, etc.).

REFERENCES

- Abadie, Alberto, and Sebastien Gay.** 2006. "The Impact of Presumed Consent Legislation on Cadaveric Organ Donation: A Cross-Country Study." *Journal of Health Economics*, 25(4): 599–620.
- Barr, Michael S., and Jane K. Dokko.** 2007. "Paying to Save: Tax Withholding and Asset Allocation Among Low- and Moderate-Income Taxpayers." Unpublished.
- Bernheim, Douglas, and Antonio Rangel.** 2009. "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics." *The Quarterly Journal of Economics*, 124(1): 51–104. Forthcoming.
- Berube, Alan, Anne Kim, Benjamin Forman, and Megan Burns.** 2002. "The Price of Paying Taxes: How Tax Preparation and Refund Loan Fees Erode the Benefits of the EITC." *The Brookings Institution, Metropolitan Policy Program, Washington, D.C.*
- Brewer, Mike, Emmanuel Saez, and Andrew Shephard.** 2010. "Means-Testing and Tax Rates on Earnings." In *The Mirrlees Review: Reforming the Tax System for the 21st Century.*, ed. Sir James Mirrlees, Tim Besley, Richard Blundell, Malcolm Gammie and James Porterba. Oxford University Press.
- Chetty, Raj, Adam Looney, and Kory Kroft.** 2009. "Salience and Taxation: Theory and Evidence." *The American Economic Review*, 99(4): 1145–1177.
- Chetty, Raj, and Emmanuel Saez.** 2009. "Teaching the Tax Code: Earnings Response to an Experiment with EITC Recipients." *NBER Working Paper No. 14836.*
- Choi, James, David Laibson, Brigitte Madrian, and Andrew Metrick.** 2003. "Optimal Defaults." *American Economic Review Papers and Proceedings*, 93: 180–185.
- Committee on Ways and Means.** 2004. *Green Book.* Washington, D.C.: Government Printing Office.
- DellaVigna, Stefano.** 2009. "Psychology and Economics: Evidence from the Field." *Journal of Economic Literature*, 47: 315–372. Forthcoming.
- DellaVigna, Stefano, and Ulrike Malmendier.** 2006. "Paying Not to Go to the Gym." *The American Economic Review*, 96(3): 694–719.
- Eliehausen, Gregory.** 2005. "Consumer Use of Tax Refund Anticipation Loans." *Credit Research Monograph #37, Credit Research Center, McDonough School of Business, Georgetown University.*
- Feenberg, Daniel, and Elizabeth Coutts.** 1993. "An Introduction to the TAXSIM Model." *Journal of Policy Analysis and Management*, 12(1): 189–194.
- Feldman, Naomi.** 2010. "Mental Accounting Effects of Income Tax Shifting." *Review of Economics and Statistics*, 92(1): 70–86.
- Fennell, Lee Anne.** 2006. "Hyperopia in Public Finance." In *Behavioral Public Finance.*, ed. Edward J. McCaffery and Joel Slemrod, 141–172. Russell Sage.
- Goldmacher, Shane, and W. J. Hennigan.** 2009. "California to withhold a bigger chunk of paychecks." *Los Angeles Times Online.* <http://www.latimes.com>.

- Highfill, Jannett, Douglas Thorson, and William V. Weber.** 1998. "Tax Overwithholding as a Response to Uncertainty." *Public Finance Review*, 26(4): 376–391.
- Hotz, V. Joseph, and John K. Scholz.** 2003. "The Earned Income Tax Credit." In *Measn-Tested Transfer Programs in the U.S.*, ed. Robert Moffitt, 141–198. Chicago:Universeity of Chicago Press.
- Internal Revenue Service.** 1992. *Employer's Tax Guide (Revised February 1992)*. Washington, D.C.:Department of the Treasury.
- Internal Revenue Service.** 2004. "2004 Statistics of Income Public Use Tax File."
- Internal Revenue Service.** 2009a. *Employer's Tax Guide*. Washington, D.C.:Department of the Treasury.
- Internal Revenue Service.** 2009b. *Form W-4, Employees Withholding Allowance Certificate*. Washington, D.C.:Department of the Treasury.
- Johnson, Eric J., and Daniel Goldstein.** 2003. "Do Defaults Save Lives." *Science*, 302(5649): 1338–1339.
- Jones, Damon.** 2010. "Information, Preferences and Public Benefit Participation: Experimental Evidence from the Advance EITC and 401(k) Savings." *American Economic Journal: Applied Economics*, 2(2): 147–163.
- Madrian, Brigitte, and Dennis Shea.** 2001. "The Power of Suggestion: Inertia in 401(K) Participation and Savings Behavior." *The Quarterly Journal of Economics*, 116(4): 1149–1188.
- Neumark, David.** 1995. "Are Rising Earnings Profiles a Forced-Saving Mechanism?" *The Economic Journal*, 105(428): 95–106.
- Rabin, Matthew.** 1998. "Pyschology and Economics." *Journal of Economic Literature*, 36(1): 11–46.
- Saez, Emmanuel.** 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy*, 2(3): 180–212.
- Sahm, Claudia R., Matthew D. Shapiro, and Joel B. Slemrod.** 2010. "Check in the Mail or More in the Paycheck: Does the effectiveness of fiscal stimuls depend on how it is delivered?" *NBER Working Paper No. 16246*.
- Shapiro, Matthew D., and Joel Slemrod.** 1995. "Consumer Response to the Timing of Income: Evidence from a Change in Tax Withholding." *The American Economic Review*, 85(1): 274–283.
- Surowiecki, James.** 2009. "A Smarter Stimulus." *The New Yorker Online*. <http://www.newyorker.com>.
- Thaler, Richard H.** 1994. "Psychology and Savings Policies." *American Economic Review Papers and Proceedings*, 84(2): 186–192.