INERTIA AND OVERWITHHOLDING: EXPLAINING THE PREVALENCE OF INCOME TAX REFUNDS – ONLINE APPENDIX

By DAMON JONES *

A1. Modeling Withholding Behavior

BASELINE MODEL. — A simple approach to modeling withholding behavior is with a two-period model. In period one, the agent receives income, w_1, and makes a tax prepayment, $\hat{\tau}$. In addition, savings are determined, s. The remaining income is consumed. In period two, the agent receives income, w_2 and interest on savings. In addition, actual taxes, τ_0, are paid. If the prepayment is higher than actual tax liability, a refund is received. If the prepayment is lower than actual tax liability, the differenced is paid, and an additional penalty, π is levied on the underpaid tax liability. Prepayments are restricted to being non-negative. The maximization problem can be summarized as follows:

$$\text{(A1)} \qquad \max_{s,\hat{\tau}} U = u(w_1 - s - \hat{\tau}) + \delta \cdot u\Big(w_2 + (1+r)s - \left(1 + \pi\left(\hat{\tau}\right)\right)\left(\tau_0 - \hat{\tau}\right)\Big)$$

$$\text{(A2)} \qquad \text{s.t. } \hat{\tau} \geq 0$$

where

$$\text{(A3)} \qquad \pi\left(\hat{\tau}\right) = \begin{cases} \pi & \text{if } \hat{\tau} < \tau_0 \\ 0 & \text{if } \hat{\tau} \geq \tau_0, \end{cases}$$

where $u(\cdot)$ is an increasing and strictly concave function and δ is the per-period discount factor. In this simple set up, the decision boils down to deciding which of two riskless assets to use for savings: one through the private sector and the other through withholdings. The returns to these assets are r, the interest rate, and π, the savings in avoided penalties, respectively. The latter effectively has a cap of τ_0. Actually, withholdings are allowed to exceed τ_0, but at that point, there is no longer a benefit of avoiding penalties. The solution can be summarized by the following:[23]

s^* satisfies:

*University of Chicago, The Harris School and NBER, 1155 E. 60th St., Chicago IL, 60605, damonjones@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.

[23] I have omitted a fourth and not very interesting case where $\pi = r = 0$.

$$u'\left(w_1 - s^* - \hat{\tau}^*_{Base}\right) = \delta(1+r)\cdot u'\Big(w_2 + (1+r)\,s^* - (1+\pi)\left(\tau_0 - \hat{\tau}^*_{Base}\right)\Big) \tag{A4}$$

and $\hat{\tau}^*_{Base}$, the level of withholding chosen in this baseline example is as follows:

$$\hat{\tau}^*_{Base} = \begin{cases} 0 & \text{if } \pi < r \\ \tau_0 & \text{if } \pi > r \\ [0, \tau_0] & \text{if } \pi = r. \end{cases} \tag{A5}$$

Note that in this simple setup, there are no intentional refunds. One exception arises when the interest rate is zero or negative. This may capture situations where a secure source of saving is not available, when other benefits are means tested based on after-tax income, or when cash-on-hand is exposed to intra-household bargaining. The majority of these cases are probably more relevant for lower income tax filers, who are most likely to overwithhold. Nevertheless, overwithholding is prevalent even among middle-income households.

BASELINE MODEL WITH BORROWING CONSTRAINTS. — In this section, we explore the effect of a borrowing constraint on withholding in the baseline case. This will be useful for comparison with the case below of borrowing constraints under time-inconsistent preferences. The borrowing constraint is captured by introducing an additional constraint:

$$s \geq 0. \tag{A6}$$

Now, the first order condition for saving is:

$$s : -u'\left(c_1^*\right) + \delta(1+r)\,u'\left(c_2^*\right) \leq 0 \tag{A7}$$

where the conditions hold with equality when $s^* > 0$ and

$$c_1^* = w_1 - s^*_{BC} - \hat{\tau}^*_{BC} \tag{A8}$$

$$c_2^* = \begin{cases} w_2 + (1+r)\,s^*_{BC} - (1+\pi)\left(\tau_0 - \hat{\tau}^*_{BC}\right) & \text{if } \hat{\tau}^*_{BC} < \tau_0 \\ w_2 + (1+r)\,s^*_{BC} - \left(\tau_0 - \hat{\tau}^*_{BC}\right) & \text{if } \hat{\tau}^*_{BC} \geq \tau_0. \end{cases} \tag{A9}$$

In this case, when savings are positive, the optimal withholding level is the same as in the case of no borrowing constraints, and $\hat{\tau}^*_{BC} = \hat{\tau}^*_{Base}$. Here the subscript denotes withholding under a borrowing constraint. Alternatively, when savings equal zero the borrowing constraint is binding. Now, the tax filer may be willing to reduce withholdings, i.e. borrow from the IRS depending on the penalty rate π. First, if $\pi \leq r$, then $\hat{\tau}^*_{BC} = 0$. In the case that $\pi > r$ there are two possibilities. If the following condition holds:

$$u'(w_1 - \tau_0) \leq \delta (1 + \pi) u'(w_2) \tag{A10}$$

then $\hat{\tau}^*_{BC} = \tau_0$. Otherwise, $\hat{\tau}^*_{BC}$ satisfies

$$u'\left(w_1 - \hat{\tau}^*_{BC}\right) \leq \delta (1 + \pi) u'\left(w_2 - (1 + \pi)\left(\tau_0 - \hat{\tau}^*_{BC}\right)\right) \tag{A11}$$

and $\hat{\tau}^*_{BC} < \tau_0$. Thus, the introduction of a borrowing constraint weakly reduces withholding relative to the baseline case, $\hat{\tau}^*_{BC} \leq \hat{\tau}^*_{Base}$. We will see that this prediction does not always hold in the case of time-inconsistent preferences.

BASELINE MODEL WITH UNCERTAINTY. — Under full certainty, the tax filer never intentionally overwithholds. However, when tax liability becomes uncertain, precautionary motives may lead to overwithholding. Assume that tax liability is unknown in period 1, but it's CDF, $F(\cdot)$ with support $\left[\underline{\tau}, \overline{\tau}\right]$ is known. Now, the maximization problem is:

$$\begin{aligned} \text{(A12)} \quad \max_{s,\hat{\tau}} \mathbb{E}[U] &= u(w_1 - \hat{\tau} - s) \\ &\quad + \delta \left[\int_{\hat{\tau}}^{\overline{\tau}} u\Big(w_2 + (1+r)s - (1+\pi)\left(t_0 - \hat{\tau}\right)\Big)\, dF(t_0) \right. \\ &\quad \left. + \int_{\underline{\tau}}^{\hat{\tau}} u\Big(w_2 + (1+r)s - \left(t_0 - \hat{\tau}\right)\Big)\, dF(t_0) \right] \\ \text{(A13)} \quad \text{s.t. } \hat{\tau} &\geq 0 \end{aligned}$$

The first order conditions for savings and tax prepayments are now:

$$s: \quad -u'\left(c_1^*\right) + \delta(1+r) \cdot \int_{\underline{\tau}}^{\overline{\tau}} u'\left(c_2^*\right) dF(t_0) \quad = 0 \tag{A14}$$

$$\hat{\tau}: \quad -u'\left(c_1^*\right) + \delta\left[\int_{\underline{\tau}}^{\overline{\tau}} u'\left(c_2^*\right)\, dF(t_0) + \pi \int_{\hat{\tau}^*_U}^{\overline{\tau}} u'\left(c_2^*\right) dF(t_0)\right] \quad \leq 0 \tag{A15}$$

where the second condition holds with equality if $\hat{\tau}^*_U > 0$, and c_1^* and c_2^* are defined as before. Here the U subscript denotes withholding under uncertainty.

With the introduction of the uncertainty, the choice of prepayment is no longer a binary outcome (i.e. 0 or τ_0). Furthermore, the level of risk aversion affects the agent's decision. When $\hat{\tau}^*_U > 0$, we can equate (A14) and (A15) and rearrange terms to obtain the following expression for the likelihood of receiving a refund, $F\left(\hat{\tau}^*_U\right)$:

$$F\left(\hat{\tau}^*_U\right) = 1 - \frac{r\mathbb{E}\left[u'\left(c_2^*\right)\right]}{\pi\,\mathbb{E}\left[u'\left(c_2^*\right) \middle|\, \tau_0 > \hat{\tau}^*_U\right]}. \tag{A16}$$

Since c_2^* is lower when $\tau > \hat{\tau}^*_U$ and utility is concave, the probability of receiving a refund increases with the curvature of $u(\cdot)$. Furthermore, let us denote the threshold $\underline{\pi}$. We can see from (A14) and (A15) that a necessary condition for $\hat{\tau}_U^* > 0$ is that $\pi > \underline{\pi} = r$.[24] Finally, note that (A16) can be seen as a more generalized version of a result shown by Highfill, Thorson and Weber (1998). They present a model with risk-neutral preferences. In the case of risk-neutral preferences, (A16) reduces to: $F\left(\hat{\tau}^*\right) = (\pi - r)/\pi$. This is identical to the result of Highfill, Thorson and Weber (1998), where the penalty rate is set to $\pi' = \pi + r$.

TIME-INCONSISTENCY AND FORCED SAVINGS. — Besides a precautionary motive or inertia, an alternative explanation for overwithholding involves Time-inconsistency and forced savings. Time-inconsistency will be modeled with sophisticated "$\beta\delta$" preferences (David Laibson 1997, Ted O'Donoghue and Matthew Rabin 1999). We will first show that time-inconsistency alone will not generate forced savings and then consider two modifications that, along with time-inconsistency, may result in forced savings. In order to demonstrate time-inconsistency with $\beta\delta$ preferences one needs at least three time periods. Thus, we will begin with the baseline model of Section A.A1 with the following modification: the level of withholding, $\hat{\tau}_\beta$, will be chosen in a pre-period, Period 0 (here the β subscript indicates the level of withholding chosen under $\beta\delta$ preferences). In Period 1, the individual makes a savings decision conditional on the level of withholdings, and Period 2 is carried out as before. Because preferences are time-inconsistent, the decision maker will be treated as different "selves" in each period, as preferences systematically change over time.

Working through backward induction, there are no choices made in the final period. Returns to savings are received, taxes and penalties are paid, and the remainder is consumed. In Period 1, the individual makes a decision similar to (A1), except that the only choice variable is s. Furthermore, since we are now dealing with $\beta\delta$ preferences, the discount factor used in the Period 1 decision problem is slightly different:

$$\text{(A17)} \quad \max_{s} \ U_1 = u(w_1 - s - \hat{\tau}) + \beta\delta \cdot u\Big(w_2 + (1+r)\,s - \left(1 + \pi\left(\hat{\tau}\right)\right)\left(\tau_0 - \hat{\tau}\right)\Big).$$

The choice of withholdings, $\hat{\tau}_\beta$ is now made in Period 0 and solves the following

[24]To see this, assume that $\hat{\tau}_U^* > 0$ and that $\pi < r$. We can use the fact that $\int_{\hat{\tau}_U^*}^{\overline{\tau}} u'\left(c_2^*\right) dF\left(t_0\right) \leq \int_{\underline{\tau}}^{\overline{\tau}} u'\left(c_2^*\right) dF\left(t_0\right)$ to show that $\pi < r$ implies (A15) < (A14). However, due to the Kuhn-Tucker conditions, (A15) < (A14) = 0 implies that $\hat{\tau}_U^* = 0$, which is a contradiction.

optimization problem:

$$\begin{aligned}
\text{(A18)} \qquad \max_{\hat{\tau}} U_0 &= u(w_1 - s(\hat{\tau}) - \hat{\tau}) + \delta \cdot u\Big(w_2 + (1+r)\, s(\hat{\tau}) \\
&\quad -\big(1+\pi(\hat{\tau})\big)\big(\tau_0 - \hat{\tau}\big)\Big) \\
&\text{s.t.:} \\
\text{(A19)} \qquad \hat{\tau} &\geq 0 \\
\text{(A20)} \qquad u'(w_1 - s(\hat{\tau}) - \hat{\tau}) &= \beta\delta(1+r) \cdot u'\Big(w_2 + (1+r)\, s(\hat{\tau}) \\
&\quad -\big(1+\pi(\hat{\tau})\big)\big(\tau_0 - \hat{\tau}\big)\Big).
\end{aligned}$$

The additional constraint is the first order condition of (A17). It takes into account the fact that in Period 1, savings will be chosen as an optimal response to the level of withholding, hence the $s(\hat{\tau})$. Also, note the time-inconsistency: in Period 0, the discount factor between Periods 1 and 2 is δ, while in Period 1, the discount factor $\beta\delta$ places a smaller weight on Period 2 relative to Period 1. Forced savings involves action taken by the Period 0 decision maker to bring Period 1 decisions closer in line with Period 0 preferences. Specifically, we will refer to forced savings as a higher level of withholding, chosen by the Period 0 self than what would have been chosen by the Period 1 self (i.e $\hat{\tau}^*_\beta > \hat{\tau}^*_{Base}$).

The solution to the baseline model with $\beta\delta$ preferences will be the same as in the previous case in Section A.A1. To see this recall that the decision to withhold 0 or τ_0 in Section A.A1 is one of wealth maximization. Any deviation from this choice, for instance a higher level of withholdings, will simply reduce the net present value of wealth for the Period 1 decision maker. This will in turn lead to a reduction in Period 1 consumption. One may think this is the goal of choosing a higher level of withholding in Period 0. However, the Euler equation in (A20) must still hold. As Period 1 consumption is reduced, so must Period 2 consumption be reduced. This reduction in consumption levels decreases the Period 0 objective function (A18). Thus, the Period 0 decision maker chooses the same level of withholding as would the Period 1 "self," and no forced savings nor overwithholding take place. The key is that the Period 1 self can just undo the Period 0 decision by dissaving, and thus the forced saving method is ineffective.

TIME-INCONSISTENCY WITH BORROWING CONSTRAINTS. — What is needed to generate a forced savings results is some friction in Period 1 self's ability to offset overwithholding by dissaving. One possible source of friction is a borrowing constraint. Suppose, as earlier, we introduce an additional constraint on the Period 1 self's maximization problem in Section A.A1:

$$\text{(A21)} \qquad s \geq 0.$$

Now, the constraint for the Period 0 decision is slightly altered:

$$u'(w_1 - s(\hat{\tau}) - \hat{\tau}) \geq \beta\delta(1+r)\cdot u'\Big(w_2 + (1+r)\,s(\hat{\tau}) - \big(1+\pi(\hat{\tau})\big)(\tau_0 - \hat{\tau})\Big). \tag{A22}$$

The Period 0 self can potentially use a binding borrowing constraint to gain some leverage on Period 1 and Period 2 outcomes. As $\hat{\tau}$ is increased by the Period 0 self, the Period 1 self will reduce s. Denote $\tilde{\tau}$ as the level of withholding at which dissaving hits the borrowing constraint: i.e. $s(\hat{\tau}) = 0\;\forall\hat{\tau} \geq \tilde{\tau}$. Note that if the Period 0 self raises withholding beyond this level, then she completely dictates the levels of consumption in Period 1 and Period 2. In particular, if

$$u'(w_1 - \tilde{\tau}) < \delta(1+\pi(\tilde{\tau}))\,u'(w_2 + (1+\pi(\tilde{\tau}))(\tau_0 - \tilde{\tau})) \tag{A23}$$

then there will be forced savings, in that $\hat{\tau}_{\beta,BC} \geq \hat{\tau}_{Base}$, where $\hat{\tau}_{\beta,BC}$ is the level of withholding chosen by a tax filer with $\beta\delta$ preferences who is facing a borrowing constraint. Furthermore, if $\tilde{\tau} \leq \tau_0$, and $u'(w_1 - \tau_0) < \delta u'(w_2)$ then we will have $\hat{\tau}_{\beta,BC} \geq \tau_0$, i.e. intentional overwithholding. As mentioned earlier, with time-consistent preferences, the introduction of a borrowing constraint weakly reduces withholding, while in the time-inconsistent case, the effect is ambiguous.

TIME-INCONSISTENCY AND UNCERTAINTY. — Another source of friction for the Period 1 decision can be found in uncertainty. Without a certain level of tax liability, the Period 1 self will not completely offset higher withholdings due to risk aversion. Even though the Period 1 self discounts Period 2 consumption more than the Period 0 self, she still takes into account Period 2 outcomes and will leave some precautionary savings for Period 2. The maximization problem for the Period 0 self can now be characterized as follows:

(A24)

$$\begin{aligned}\max_{\hat{\tau}} \mathbb{E}[U]_0 \;=\;& u(w_1 - s(\hat{\tau}) - \hat{\tau}) \\ &+\delta\Bigg[\int_{\hat{\tau}}^{\overline{\tau}} u\Big(w_2 + (1+r)\,s(\hat{\tau}) - (1+\pi)\,(t_0 - \hat{\tau})\Big)\,dF(t_0) \\ &+\int_{\underline{\tau}}^{\hat{\tau}} u\Big(w_2 + (1+r)\,s(\hat{\tau}) - (t_0 - \hat{\tau})\Big)\,dF(t_0)\Bigg] \\ &\text{s.t.}\end{aligned}$$

$$\hat{\tau} \geq 0 \tag{A25}$$

$$u'\left(w_1 - s\left(\hat{\tau}\right) - \hat{\tau}\right) = \beta\delta\left(1+r\right)\cdot\int_{\underline{\tau}}^{\overline{\tau}} u'\Big(w_2 + \left(1+r\right)s\left(\hat{\tau}\right) - \left(1+\pi\right)\left(t_0 - \hat{\tau}\right)\Big)dF\left(t_0\right), \tag{A26}$$

where the second constraint again accounts for the savings decision of the Period 1 self in response to the level of withholding. Note again that the Period 1 self has a different discount factor between Period 1 and Period 2. The first order condition for withholding will now be:

$$u'\left(c_1^*\right)\left[1+s'\left(\hat{\tau}^*_{\beta,U}\right)\right] \geq \delta\Bigg[\left[1+\left(1+r\right)s'\left(\hat{\tau}^*_{\beta,U}\right)\right]\int_{\underline{\tau}}^{\overline{\tau}} u'\left(c_2^*\right)\, dF\left(t_0\right) + \pi\int_{\hat{\tau}^*_{\beta,U}}^{\overline{\tau}} u'\left(c_2^*\right)dF\left(t_0\right)\Bigg], \tag{A27}$$

where the condition holds with equality if $\hat{\tau}^*_{\beta,U} > 0$ (the β, U subscript denotes withholdings with $\beta\delta$ preferences and uncertainty). Similar to before

$$c_1^* = w_1 - s\left(\hat{\tau}^*_{\beta,U}\right) - \hat{\tau}^*_{\beta,U} \tag{A28}$$

$$c_2^* = \begin{cases} w_2 + \left(1+r\right)s\left(\hat{\tau}^*_{\beta,U}\right) - \left(1+\pi\right)\left(t_0 - \hat{\tau}^*_{\beta,U}\right) & \text{if } \hat{\tau}^*_{\beta,U} < t_0 \\ w_2 + \left(1+r\right)s\left(\hat{\tau}^*_{\beta,U}\right) - \left(\tau_0 - \hat{\tau}^*_{\beta,U}\right) & \text{if } \hat{\tau}^*_{\beta,U} \geq t_0. \end{cases} \tag{A29}$$

We now have an analog to (A16) in the case of $\beta\delta$ preferences. Substituting (A26) into (A27), we have

$$F\left(\hat{\tau}^*_{\beta,U}\right) = 1 - \frac{\tilde{r}\left(\beta\right)\cdot\mathbb{E}\left[u'\left(c_2^*\right)\right]}{\pi\,\mathbb{E}\left[u'\left(c_2^*\right)\middle|\,\tau_0 > \hat{\tau}^*_{\beta,U}\right]}, \tag{A30}$$

when $\hat{\tau}^*_{\beta,U} > 0$. The new term $\tilde{r}\left(\beta\right)$ is defined as:

$$\tilde{r}\left(\beta\right) = \beta r - \left(1-\beta\right)\left[1+\left(1+r\right)s'\left(\hat{\tau}^*_{\beta,U}\right)\right] \tag{A31}$$

$$= r\cdot\left(\frac{u''\left(c_1^*\right) + \beta\delta\left(1+r\right)^2\mathbb{E}\left[u''\left(c_2^*\right)\right]\left(\beta + \left(1-\beta\right)\frac{\pi}{r}\lambda\right)}{u''\left(c_1^*\right) + \beta\delta\left(1+r\right)^2\mathbb{E}\left[u''\left(c_2^*\right)\right]}\right), \tag{A32}$$

and

$$\lambda = \frac{\int_{\hat{\tau}^*_{\beta,U}}^{\overline{\tau}} u''\left(c_2^*\right)dF\left(t_0\right)}{\int_{\underline{\tau}}^{\overline{\tau}} u''\left(c_2^*\right)\, dF\left(t_0\right)} < 1. \tag{A33}$$

The second line (A32) follows from starting with (A26), using the implicit function theorem to solve for $s'\left(\hat{\tau}^*_{\beta,U}\right)$ and substituting into (A31). First, note that $\tilde{r}(1) = r$, i.e. when $\beta = 1$, preferences are time-consistent and (A30) collapses to (A16). Second, we use similar logic as in Section A.A1 to show a necessary condition for positive withholdings, $\hat{\tau}^*_{\beta,U} > 0$, is $\pi > \tilde{r}(\beta)$. This condition can be rewritten as

$$\text{(A34)}\quad \begin{aligned} \pi &> \underline{\pi}_\beta \\ &= r \cdot \left(\frac{u''\left(c_1^*\right) + \beta\delta(1+r)^2\,\mathbb{E}\left[u''\left(c_2^*\right)\right](\beta)}{u''\left(c_1^*\right) + \beta\delta(1+r)^2\,\mathbb{E}\left[u''\left(c_2^*\right)\right](\beta + (1-\beta)(1-\lambda))} \right). \end{aligned}$$

We see in (A34) that compared to the results in Section (A.A1), $\underline{\pi}_\beta < r = \underline{\pi}$. So, we have forced savings in the following limited sense. Holding constant the distribution of tax liability, wages, interest rate r and δ, withholdings are guaranteed to be zero for a larger range of penalties (i.e. $\pi \in [0, \underline{\pi})$) when the individual is time-consistent as compared to the time-inconsistent case (i.e. $\pi \in [0, \underline{\pi}_\beta)$).

BASELINE MODEL WITH TRANSACTIONS COSTS. — We return to the initial model with no uncertainty, but now we introduce a fixed cost of adjusting withholdings. In this case, high transactions costs may cause an individual to underwithhold or overwithhold. The thought experiment is as follows: If we endow an agent with a withholding level, will they find it worthwhile to adjust their withholding? We will assume that the individual faces an additive disutility of adjustment, φ_i. The answer depends on whether the optimal prepayment absent adjustment costs, $\hat{\tau}^*\big|_{\varphi=0}$, is 0 or τ_0. The answer also depends on whether the endowed level of withholding will result in a refund or balance due, τ^R and τ^B respectively.

First consider the case where absent adjustment costs, the individual would make full prepayments: $\hat{\tau}^*\big|_{\varphi=0} = \tau_0$. Now, let us endow the agent with a prepayment of $\tau^R > \tau_0$, such that $\left|\tau^R - \tau_0\right| = \triangle\tau$. Inaction will result in a refund, hence the superscript. In effect, if the taxpayer takes no action, it is as if she was forced to give a zero interest loan of size $\triangle\tau$ to the government. Thus, if we denote $W_0|_{\hat{\tau}^*=\tau_0} = w_1 + \frac{w_2}{1+r} - \tau_0$ as the baseline level of after-tax wealth absent this forced loan, then the agent will change the default withholding level if:

$$\text{(A35)}\qquad V(W_0) - V\left(W_0 - \frac{r\triangle\tau}{1+r}\right) > \varphi_i,$$

where $V(\omega)$ is the indirect utility function associated with (A1), given an after-tax level of wealth ω.

Instead suppose that the endowed payment is below the tax level, so that no action will result in a balance due and penalty. That is $\tau^B < \tau_0$ and $\left|\tau^B - \tau_0\right| = \triangle\tau$, where the superscript now denotes a default balance due. Here, it is as if the agent is force to take a loan of size $\triangle\tau$ from the government at an interest rate of π. Now, the agent will change the withholding level if:

$$V(W_0) - V\left(W_0 - \frac{(\pi - r)\triangle\tau}{1+r}\right) > \varphi_i. \tag{A36}$$

Alternatively, assume that absent adjustment costs, the individual would make no prepayments: $\hat{\tau}^*\big|_{\varphi=0} = 0$. Consider an identical experiment as in the previous two cases. If the endowed prepayment would result in a refund, the individual will change the default withholding level if:

$$V(W_0) - V\left(W_0 - \frac{r\tau^R - \pi\tau}{1+r}\right) > \varphi_i, \tag{A37}$$

where now $W_0|_{\hat{\tau}^*=0} = \omega_1 + \frac{\omega_2}{1+r} - \frac{(1+\pi)\tau_0}{1+r}$. Withholdings will be changed with a low endowment if:

$$V(W_0) - V\left(W_0 - \frac{(r - \pi)\tau^B}{1+r}\right) > \varphi_i. \tag{A38}$$

The withholding level is more likely to deviate from the default in each case the greater the inconvenience of the default. With respect to a bias towards receiving refunds or owing a balance, one is more likely to adjust withholdings in the event of an endowed balance due, *ceteris paribus*, if $\pi - r > r$. Intuitively, when one gives an interest free loan to the government via overwithholding, what's lost is the interest that could otherwise be gained: r. On the other hand, when one is forced to take a loan from the government via lower withholdings, the interest charged on this loan is π. However, the amount borrowed can be save, earning an interest rate of r, hence the quantity $\pi - r$ on the left hand side of the inequality. In the event that the inequality holds, an overwithholding is less costly than underwithholding. Whether or not this condition is realistic depends on the actual penalty and interest rate faced by taxpayers.

Asymmetric adjustment costs can be introduced by allowing φ_i to vary with the position of the endowed balance relative to some reference point, such as no balance. In general, heterogeneity in inertia may be driven by heterogeneity in adjustment costs φ_i. For example, if φ_i is constant, then inertia will be decreasing in income. On the other hand, if φ_i is proportional to W_0, then inertia may be relatively constant across income groups.

One final point is that with adjustment costs, agents will respond differently to shocks that are permanent than to shocks that are transitory. In general, transitory shocks to the default level of withholding will trigger less of a response, if agents think that in time the default withholding level will revert to its previous state. Thus, things like a one-time policy change in tax liability may be less accounted for by adjustments in withholdings than an addition of a new dependent that will remain in the household for some time.

DISCUSSION. — To summarize, in a simple model of withholding with no uncertainty, tax filers are generally predicted to at most withhold exactly as much as their tax liability, i.e. there is no intentional overwithholding. This baseline model is unreasonable in at least two ways: it assumes that agents can very accurately predict their tax liability and it ignores adjustment costs. We see that if either of these assumptions is relaxed, then at least some share of agents may overwithhold.

Are these additional features enough to explain the data? In the case of uncertainty, we can conduct a back of the envelope calculation using the identity for the likelihood of overwithholding in (A16). Rearranging (A16) we have:

$$\frac{F\left(\hat{\tau}_U^*\right)}{1-F\left(\hat{\tau}_U^*\right)} = \frac{\pi - r}{r} \cdot \frac{\mathbb{E}\left[u'\left(c_2^*\right)\middle|\,\tau_0 > \hat{\tau}_U^*\right]}{\mathbb{E}\left[u'\left(c_2^*\right)\middle|\,\tau_0 \leq \hat{\tau}_U^*\right]}. \tag{A39}$$

Here we see the odds of overwithholding are proportional to the ratio of expected marginal utility in the event that one has either over- or underwithheld. We can plug in for $F\left(\hat{\tau}_U^*\right)$ the observed probability of overwithholding 0.8, a reasonable value for π of 0.03 and a risk-free rate of return of 0.02 for r. With these parameters the expected jump in marginal utility when underwithheld relative to when overwithholding is 8. This implies a significant drop in after-tax earnings. For example, with a CRRA utility function of form $u(c) = \frac{c^{1-\gamma}-1}{1-\gamma}$ and a coefficient of relative risk aversion $\gamma = 2$, it takes a 65 percent drop in consumption to generate an eight-fold increase in marginal utility. The model may also be salvaged by assuming a very high level of risk aversion or a significant upward bias in beliefs about π. Nonetheless, it would appear that uncertainly alone cannot entirely explain the observed patterns of overwithholding.

One possible way to reconcile the data with the model is to introduce time-inconsistent preference, modeled by $\beta\delta$ preferences. We have shown that in certain cases, a time-inconsistent agent may choose a higher level of overwithholding than a time-consistent agent. However, in the baseline case with no uncertainty, the time-inconsistent agent is not distinguishable from a time-consistent agent based on the level of withholding. The $\beta\delta$ preferences must be combined with either a borrowing constraint or uncertainty to generate a forced savings result. In particular, we have shown in the case of certainty that the introduction of a borrowing constraint will weakly reduce withholdings for a time-consistent agent. In contrast, the effect of a borrowing constraint on a tax filer with $\beta\delta$ preferences is ambiguous, and may even result in higher withholdings. This provides a limited test of time-inconsistency. Unfortunately, in the IRS data used here, the presence of a borrowing constraint is unobserved, and therefore one is not able to perform such a test.

In the case of uncertainty, the $\beta\delta$ model does not go far in providing a better fit to the data. The analog to (A39) in the time-inconsistent case is

$$\text{(A40)} \qquad \frac{F\left(\hat{\tau}^*_{\beta,U}\right)}{1-F\left(\hat{\tau}^*_{\beta,U}\right)} = \frac{\pi - \tilde{r}\left(\beta\right)}{\tilde{r}\left(\beta\right)} \cdot \frac{\mathbb{E}\left[u'\left(c_2^*\right)\middle|\,\tau_0 > \hat{\tau}^*_{\beta,U}\right]}{\mathbb{E}\left[u'\left(c_2^*\right)\middle|\,\tau_0 \leq \hat{\tau}^*_{\beta,U}\right]},$$

where $\tilde{r}\left(\beta\right)$ is defined as before in (A31). For a given probability of withholding, the model with time-inconsistent preferences will require a smaller drop in expected marginal utility if $\tilde{r}\left(\beta\right) < r$. Looking at (A32), we can see that this will be true when:

$$\text{(A41)} \qquad \frac{\pi}{r}\lambda < 1$$

$$\text{(A42)} \qquad \text{or } \frac{F\left(\hat{\tau}^*_{\beta,U}\right)}{1-F\left(\hat{\tau}^*_{\beta,U}\right)} < \frac{\pi - r}{r} \cdot \frac{\mathbb{E}\left[u''\left(c_2^*\right)\middle|\,\tau_0 > \hat{\tau}^*_{U}\right]}{\mathbb{E}\left[u''\left(c_2^*\right)\middle|\,\tau_0 \leq \hat{\tau}^*_{U}\right]}.$$

Here again the model runs into problems fitting the data. Plugging in values from before, and assuming CRRA preferences again, we need at least a 50 percent drop in consumption when underwithheld relative to when overwithheld. This is still a rather large swing in consumption. As such, the $\beta\delta$ model cannot salvage an explanation of overwithholding based only on the model of uncertainty presented above.

An alternative explanation introduces adjustment costs associated with choosing a new level of withholding. Inertia alone does not predispose one toward overwithholding or underwithholding. However, if the payoff to correcting too low of a default withholding level is higher than the payoff to adjusting too high a default withholding level, i.e. $\pi - r > r$, then individuals will be more likely to exhibit inertia when overwithholding. Again, plugging in the values of $\pi = 0.03$ and $r = 0.02$ does not satisfy such a condition, though there may exist heterogeneity in both r and the perceived value of π among taxpayers. Alternatively, the model of inertia may generate a bias toward overwithholding if default withholding levels are more like to be higher than tax liability. This is true at least in the case where an individual does not file an initial W-4 form with the employer; the employer is instructed to choose zero allowances on behalf of the employee, with typically will result in overwithholding.

A2. Estimating the Distribution of Allowances

I use data on wage and salary earnings, withholdings and the IRS withholding tables to estimate the distribution of chosen allowances. This is used to create Figure 1 Panel B and in the analysis of the 1992 change in default withholdings. The distribution of allowances $\hat{F}_0\left(A_0^i\right)$ and $\hat{F}_1\left(A_1^i\right)$ are estimated as follows. The data for tax filers from 1991 and 1992 are restricted to individuals who claimed a standard deduction, with wage and salary income comprising more than 95 percent of AGI. This eliminates other sources of income that may confound the relationship between wages and withholdings. In addition, the top one percent of tax filers (those with AGI greater than $200,000) were dropped due to additional blurring of their incomes.

Next, for a given level of wage and salary earnings, E^i, a level of withholdings, $P\left(A, E^i\right)$, for each number of allowances, A, was computed using IRS Publication 15, *Circular E: Employer's Tax Guide* for the given year. The number of allowances that generate the closest match to actual withholdings is assigned to the tax filer: $\hat{A}_i = \arg\min_A \left(P^i - P\left(A, E^i\right)\right)$. Essentially, I invert the $P(\cdot)$ withholding functions. In the analysis of the 1992 withholding policy, the discrete distribution of these estimated allowances are calculated for each year-by-income group, separately for married and single tax filers, where the income groups are defined by 16 AGI quantiles. These conditional distributions are used to increase the accuracy of the imputation. 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 than the unconditional distribution of allowances among married and single filers of all incomes in 1991. Under Assumption (4), I arrive at estimates of the conditional distributions, $\hat{F}_0\left(A_0^i\middle|\theta^i\right)$ and $\hat{F}_1\left(A_1^i\middle|\theta^i\right)$, where θ^i is a vector containing income group and marital status.

A3. Additional Details on Sample Construction

1992 CHANGE IN DEFAULT WITHHOLDINGS SAMPLE. — For the analysis of the 1992 change in default withholdings, I use public use tax return data from the years 1991 and 1992. The initial sample includes 208,856 observations. I first drop tax returns in the data set that are filed for years other than 1991 or 1992. Because IRS Statistics of Income (SOI) group applies additional blurring of income for those with income above $200,000, I drop these observations from the sample. There are 145,278 observations remaining after this. Note that the SOI data oversample those with high income, so after using sampling weights, these high-income observations represent less than 1 percent of the population. Next, I restrict analysis to those with more than 95 percent of their income from wage earnings and who take a standard deduction. This is because the analysis in this section uses withholdings and withholdings tables to draw inference about the number of allowances chosen on the W4. The need to adjust withholdings for other sources of income or anticipated, itemized deductions confounds the relationship between wages, withholdings and allowances that is exploited in estimating the distribution of allowances. Restricting the sample to this group results in a sample size of 32,049. Again, since the IRS data set oversamples high-income tax filers, the sample weights imply that the remaining observations represent 45 percent of the population. As a placebo check, I replicate this analysis using time periods 1990 - 1991 and 1989 - 1990. The same method of sample selection is used in these placebo exercises.

PANEL STUDY OF CHILD DEPENDENTS SAMPLE. — The sample used in this analysis is taken from the University of Michigan panel of tax return data spanning 1979 - 1990. The panel is constructed from a random sample of tax returns during this period, and due to the process of sample construction, it is not uncommon for individuals to drop out and reappear in the panel. As such, constructing a balanced subsample requires a massive loss of data. Therefore, I used an unbalance panel to conduct the analysis. The initial sample size is 294,175. I begin by dropping individuals with income above $200,000. The IRS SOI department that prepares public use data censors the number of children for tax returns in this range. I also drop tax returns that are coded as earlier than 1979. This leaves 293,011 observations.

In the next step, I identify individuals who are observed undergoing a change in the number of children. This restricts analysis to individuals who have a different number of children in two adjacent years during the panel. Those with a lower number of children in the latter year are labeled as having a "loss" and those with a higher number of children are labeled as having a "gain" in dependents. Note, the number of children may change for various reasons, most likely due to a birth, a child aging out of the household or a divorce. The subsample of individuals who have experienced an event includes 87,904 observations.

Dummy variables that measure time since the event ($Loss_{i,t-j}$ and $Gain_{i,t-j}$) are defined for each observation in this subset. Event year zero is defined as the last year before the number of child dependents change. In the case of multiple events, the dummy variables are defined based on the event with the nearest proximity in time. Because estimates become dramatically noisier when done more than three years away from the event, I restrict analysis to a seven-year window surrounding the event, resulting in a sample of 68,702. In addition, because analysis is sometimes done separately on those who lose and those who gain, observations that are near both a loss and gain event are dropped. Finally, I drop individuals who are above age 65, as they are very unlikely to have children dependents. The remaining sample includes 62,604 observations. Using the sampling weights, this subsample is represents 23 percent of the population covered by the panel.

In order to account for income and marital dynamics in the neighborhood of the event, I include lagged values of splined income and marital status as control variables . However, due to the nature of the sampling procedure, some panel observations may not have a prior year observation in the data set. Thus, I impute lagged income and marital status for these individuals by fitting a fixed-effects regression of lagged income (or marital status) on the current value. The regressions include interactions with event time, and controls for calendar year, and in the case of martial status, controls for income. Columns (3) and (4) of Appendix Table A5 shows that dropping those observations that require imputation has a negligible effect on the estimates.

EITC EXPANSION SAMPLE. — The EITC sample pools observations from cross sections of public use tax return data from 1981 to 2004. For each observation, I normalize income to year 2000 dollars and calculate EITC eligibility using the year 2000 EITC eligibility parameters. This is done using the NBER TAXSIM tax simulator. Thus, eligibility is defined this way so as to abstract from the changes in the composition of the EITC-eligible pool over time. This initial pool of observations has a sample size of 2,616,330. Analysis is restricted to those eligible for the EITC, which results in a sample size of 270,603. Next, the remaining observations are aggregated within each year by number of children: zero children, 1 child or 2 or more children. After the data is collapsed, there remain 72 group-by-year cells.

A4. Calculating the Private Cost of Incorrect Withholding

A lower bound on the cost to the tax filer of overwithholding is measured as the money lost by giving the government an interest-free loan during the year. The relevant interest rate used in calculating the opportunity cost of overwithholdings depends on whether or not individuals are holding debt and the types of investment opportunities that are available to them. To calculate the "Interest Cost" of withholding I use the 2004 Survey of Consumer Finance (SCF) to impute interest rates for individuals in the 2004 IRS SOI data set.

For each observation in the SCF I record the maximum of (1) credit card interest rates for those with positive credit card debt, (2) the July 2004 rate of 1.06 percent for 9-month Certificates of Deposit (CD) for those with positive CD holdings or (3) a rate of 0.4 percent for those with a positive savings account balance. An interest rate of zero is recorded for individuals who hold none of the previous debts or assets. I then split the SCF into married and non-married households and further into income deciles, based on the IRS SOI income distribution. Next, for each observation in the IRS data set, I randomly draw an interest rate from their corresponding marital status by income decile pool in the SCF. The imputed interest rate is then multiplied by the individual's income tax refund or balance due. Those with a refund have a cost of overwithholding, while those with a balance due receive a benefit of underwithholding. The average costs in terms of loss or gained interest is reported for each income quintile and the total sample in Table 6.

If individuals face imperfect credit markets and/or have no savings, then an upper bound on the cost of overwithholding will be based on the inability to smooth consumption. To calculate these costs, I consider a case where individuals have a discount rate of zero and face an interest rate of zero. Income is received in T equal installments y. In this case, an individual with concave utility will desire a flat consumption profile. Now assume that tax prepayments are likewise paid in T equal installments, p and tax liability is also due in T equal installments l. Denote the monthly net refund as $r \equiv p - l$. If the individual overwithholds every period for T periods, she will receive a refund of $T \cdot r$ in month T (net withholdings of r are still incurred in month T). Finally, assume that individuals cannot borrow, so that consumption is equal to income minus net withholdings for overwithholders. The cost of overwithholding is the equivalent variation, $\triangle y$, of deviating from a constant consumption profile to one where the timing of income is distorted by overwithholding and satisfies the following:

$$\text{(A43)} \qquad \sum_{i=1}^{T} u\left(y - \triangle y\right) = \left[\sum_{i=1}^{T-1} u\left(y - r\right)\right] + u\left(y + (T-1)\,r\right).$$

For individuals who underwithhold $\triangle y$ is set to zero, as these tax filers can achieve the optimal, flat consumption profile by saving net withholdings until the last period and paying all taxes owed then. I assume a Constant Relative Risk Aversion (CRRA) functional form for utility: $u(c) = \frac{c^{1-\gamma}-1}{1-\gamma}$ or $u(c) = \ln(c)$ when $\gamma = 1$. Solving for $\triangle y$ we have:

$$\text{(A44)} \quad \triangle y = \begin{cases} y - \left[\frac{T-1}{T}(y-r)^{1-\gamma} + \frac{1}{T}(y+(T-1)r)^{1-\gamma}\right]^{\frac{1}{1-\gamma}} & \text{when } \gamma \neq 1 \\ y - \exp\left[\frac{T-1}{T}\ln(y-r) + \frac{1}{T}\ln(y+(T-1)r)\right] & \text{when } \gamma = 1. \end{cases}$$

The average, annual cost, $T \cdot \triangle y$, is calculated for each individual in the IRS SOI data set. Time periods are set to one month, $T = 12$, y is one-twelfth of AGI and r is one-twelfth of the refund level. The average cost within each income quintile is reported in Table 6 for different values of γ.

A5. Robustness Checks

TABLE A1—CHANGE IN CHILD DEPENDENTS - ALTERNATE SPECIFICATIONS

	(1)	(2)	(3)	(4)	(5)	(6)
	Adjustment Rate: α_L					
	Baseline	Alternative Specifications				
Year 1	0.28	0.29	0.29	0.31	0.29	0.29
	(0.06)***	(0.06)***	(0.06)***	(0.07)***	(0.07)***	(0.06)***
Year 2	0.38	0.40	0.39	0.47	0.53	0.48
	(0.08)***	(0.08)***	(0.08)***	(0.08)***	(0.08)***	(0.08)***
Year 3	0.49	0.52	0.51	0.61	0.66	0.61
	(0.09)***	(0.09)***	(0.10)***	(0.10)***	(0.10)***	(0.10)***
N	63,782	62,604	62,604	62,604	62,604	62,604
Drop 65+		Yes	Yes	Yes	Yes	Yes
Marital Dynamics		No	Yes	No	No	Yes
Income Spline		No	No	Yes	Yes	Yes
Lagged Income Spline		No	No	No	Yes	Yes

Note: Estimates of the adjustment rate are obtained using Equations (13)-(15). Data are from the Michigan tax panel spanning 1979-1990. Robust standard errors are reported in parentheses. Controls not mentioned include 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 A2—CHANGE IN CHILD DEPENDENTS - ALTERNATE SAMPLES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Sample Used								
						Zero Balance I		Zero Balance II	
	Full Sample	Stable Marriage	Balanced Sample	Losers	Gainers	Losers	Gainers	Losers	Gainers
Year 1	0.29	0.43	0.86	0.20	0.37	0.54	−0.13	0.48	−0.02
	(0.06)***	(0.17)***	(0.22)***	(0.15)	(0.13)***	(0.17)***	(0.08)	(0.15)***	(0.08)
Year 2	0.48	0.68	0.88	0.58	0.56	1.29	−0.30	1.14	−0.09
	(0.08)***	(0.17)***	(0.23)***	(0.22)***	(0.21)***	(0.34)***	(0.17)	(0.28)***	(0.17)
Year 3	0.61	0.59	0.87	1.22	0.66	1.71	−0.21	1.74	−0.03
	(0.10)***	(0.12)***	(0.22)***	(0.70)*	(0.21)***	(0.57)***	(0.20)	(0.53)***	(0.16)
N	62,604	16,048	5,476	30,517	32,087	11,554	15,585	16,617	20,598

Note: Estimates of the adjustment rate are obtained using Equations (13)-(15). The "Stable Marriage" sample includes tax filers who do not experience a change in marital status while in the panel. The "Balanced" sample only includes observations that are present for all seven years surrounding the event. The "Zero Balance I" sample is restricted to tax filers with a refund or balance due less than $1,000 in the base year, while the "Zero Balance II" sample uses a threshold of $1,500. 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, martial status, lagged martial 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 A3—CHANGE IN CHILD DEPENDENTS - ALTERNATE SAMPLE DESCRIPTIVE STATISTICS

	(1)	(9)	(8)	(2)	(3)	(4)	(5)	(6)	(7)
	Sample Used								
						Zero Balance I		Zero Balance II	
	Full Sample	Balanced Sample	Gainers	Losers	Gainers	Losers	Gainers	Losers	Gainers
Adjusted Gross Income									
10th Percentile	9,581	14,193	14,060	11,166	8,399	8,187	5,515	8,930	6,312
Median	37,205	45,108	43,125	42,531	32,303	35,292	24,441	35,132	25,783
90th Percentile	84,295	89,841	85,669	90,688	76,535	77,234	61,947	76,935	63,544
Mean	43,793	49,941	47,822	48,354	39,330	40,313	30,292	40,263	31,737
Standard Deviation	35,597	35,919	31,793	37,538	32,585	31,704	25,170	30,608	25,652
Refund									
Mean	932	1,019	1,035	762	1,099	285	666	426	771
Median	905	966	908	801	991	357	605	499	737
Prepayment to Liability Ratio									
10th Percentile	0.79	0.81	0.85	0.73	0.86	0.77	0.89	0.79	0.90
Median	1.23	1.22	1.22	1.17	1.31	1.08	1.23	1.11	1.26
90th Percentile	2.38	2.16	2.08	2.13	2.62	1.73	2.57	1.87	2.62
Refund to AGI Ratio									
Mean	0.04	0.03	0.03	0.03	0.04	0.02	0.04	0.03	0.04
Median	0.03	0.03	0.03	0.02	0.04	0.01	0.03	0.02	0.03
Refund Probability	0.80	0.79	0.79	0.75	0.85	0.69	0.84	0.73	0.85
Share of Total Filers	0.23	0.08	0.03	0.12	0.12	0.04	0.06	0.06	0.08
N	62,604	16,048	5,476	30,517	32,087	11,554	15,585	16,617	20,598

Note: See Table A2 for a description of the samples. Dollar amounts are reported in year 2000 levels.

TABLE A4—CHANGE IN CHILD DEPENDENTS - HETEROGENEITY IN ADJUSTMENT RATE ESTIMATES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Baseline	Low Income	High Income	1st Event	2nd+ Event	Single	Married
Year 1							
α_L	0.29	0.21	0.33	0.26	0.38	0.22	0.28
	(0.06)***	(0.07)***	(0.09)***	(0.07)***	(0.14)***	(0.06)***	(0.09)***
Year 2							
α_L	0.48	0.34	0.56	0.44	0.55	0.53	0.43
	(0.08)***	(0.12)***	(0.09)***	(0.09)***	(0.15)***	(0.08)***	(0.10)***
Year 3							
α_L	0.61	0.43	0.74	0.55	0.72	0.55	0.58
	(0.10)***	(0.09)***	(0.14)***	(0.14)***	(0.16)***	(0.10)***	(0.12)***
N	62,604	28,523	34,081	48,058	14,546	18,129	44,475

Note: Estimates of the adjustment rate are obtained using Equations (13)-(15). "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. Single tax filers are those that are single in the year of the event, while married tax filers are married in the event year. 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, martial status, lagged martial 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.