

## Reexamining Income Tax Overwithholding as a Response to Uncertainty

Ashvin Gandhi  
Harvard University

Michael Kuehlwein  
Pomona College

September 2013

### Abstract

This paper reexamines the proposition of Highfill *et al* (1998) that income tax overwithholding in the US can be explained as rational risk-neutral taxpayers trying to avoid penalties for underwithholding when faced with uncertain tax liability. We first adjust their model to account for interest accumulated on underwithheld income and to enforce consistent boundary conditions. We then incorporate a relevant tax rule into the model. Finally, we replace two distributional assumptions with ones that allow more realistic levels of tax liability uncertainty. Each of these modifications reduces the predicted level of overwithholding. Together, they imply that penalty avoidance can explain only a fourth to a fifth of the true refund rate on average. This suggests that penalty-aversion in a risk-neutral framework is not sufficient to explain observed overwithholding in the US.

## 1. Introduction

In the United States, income tax on wages is automatically withheld from one's paycheck. While the default rate of automatic withholding tends to exceed one's tax liability, informed employees can adjust their automatic withholding by filing a Form W-4 with their employer. Additionally, one must make estimated payments for any tax liability stemming from non-salary income.<sup>1</sup>

Surprisingly, each year a significant majority of taxpayers choose to withhold more than they owe in taxes. They also overshoot their tax liability by a significant margin: the average taxpayer's overpayment is 7% of their adjusted gross income (Jones 2010). While the government does eventually return this money as a tax refund, it amounts to a \$340 billion zero-interest loan made voluntarily by 83% of America's 140 million income tax filers (Internal Revenue Service 2011). This high refund rate is not new, and the root of taxpayers' persistent willingness to make this zero-interest loan rather than earn positive market returns is an area of some debate.<sup>2</sup>

The significance of this phenomenon is underscored by the fact that low income taxpayers are the most likely to withhold too much, and they do so by the largest relative margin: the average low income taxpayers' overpayment is 13% of their adjusted gross income (Jones 2010). This group tends to borrow at high interest rates—such as from credit cards and payday loans—in order to smooth consumption over the year, so the implications of withholding too much extend beyond forgone interest. Jones (2012) estimates that the cost to consumption smoothing can be as high as 14% of income for taxpayers in the bottom quintile.

Highfill, Thorson, and Weber (1998), henceforth HTW, propose that taxpayers withhold more than necessary because they are uncertain about their tax liability and hope to avoid penalties for withholding too little. In the period we examine, 1983-1992, the Internal Revenue Service (IRS) did not impose a penalty if the taxpayer either withheld more than a preset percentage of his current-year tax liability or more than his previous year's tax liability.<sup>3</sup> For the purposes of this paper, a taxpayer who is penalized is said to be "underwithholding." We consider a taxpayer whose withholding exceeds at least one of the

above cutoffs to be “overwithholding” because he could lower his withholding without being penalized. It is useful to observe the distinction between someone who is overwithholding and someone who receives a refund. One can be the former without being the latter. In fact, this arguably is the taxpayer’s ideal situation to be in: withholding less than he owes, but not so little as to be penalized.

HTW present a model in which risk-neutral taxpayers, uncertain about their tax liability, choose the level of withholding that minimizes the sum of their expected penalties from underwithholding and their expected forgone interest from overwithholding. In that model, taxpayers withhold at rates which cause them to overwithhold with great frequency and receive refunds with only slightly less frequency. In fact, their model predicts refund rates that approach, and sometimes even exceed, the actual refund rate.

Their model is a novel and valuable contribution to the withholding literature. Nonetheless, it contains two oversights and some simplifications that may influence their results. This paper addresses those oversights and extends their model to see whether their key results still hold. First, we account for the interest that the taxpayer might earn on underwithheld income. Second, we impose consistent boundary conditions required by HTW’s assumptions. Both of these changes decrease the model’s predicted probability of overwithholding. When we further account for the fact that taxpayers are not penalized as long as their withholding exceed last year’s tax liability, the model’s estimates of the refund rate fall to about half of actual rates. Finally, we separately relax two strong assumptions about the PDF of tax liability. First, we allow taxpayers to have some certain level of tax liability. More certain tax liability produces less overwithholding. Under reasonable assumptions about the ratio of certain to uncertain liability, the model predicts refund rates that average a quarter of the actual ones. Alternatively, we allow uncertain tax liability to be normally, rather than uniformly, distributed. This can also be thought of as allowing for greater taxpayer precision in estimating their tax liability. When tax liability is normally distributed, under reasonable parameter assumptions the model’s predicted refund rate drops to 16% from 1983-1992, about a fifth of the real average of 75%. This suggests that penalty-aversion in a risk-neutral framework is not sufficient to explain observed overwithholding in the US.

The remainder of the paper is organized as follows. Section 2 discusses the literature on income tax overwithholding. Section 3 summarizes the model and results from HTW. Section 4 presents our amendments and extensions to the model and their implications. Section 5 concludes. Proofs of all non-immediate results are contained in an appendix.

## **2. Literature**

Much of the recent literature on overwithholding views it as more of a behavioral phenomenon than the result of optimization under uncertainty. For example, overwithholding to receive a refund has been considered a forced savings mechanism. Neumark (1995) finds that overwithholding is correlated with steeper earnings profiles, which he takes as support for the hypothesis that rising earnings profiles are a means of forced savings.

Barr and Dokko (2007) use survey data to examine the relationship between portfolio liquidity and withholding for low to moderate income taxpayers. They find that survey participants generally report a preference for overwithholding enough to receive a refund, and that this preference is correlated with putting savings in less accessible assets. They conclude that overwithholding is used as a commitment device to restrict consumption in the face of dynamic inconsistency. Further, they provide evidence against precautionary savings, the mental accounting of refunds as windfalls, and loss aversion as reasons for overwithholding.

Chambers and Spencer (2008) estimate that taxpayers are less likely to spend income tax refunds received in an annual lump-sum compared with being evenly divided over 12 months. They attribute this to mental accounting and different perceptions of the two streams of income. This argument is also made explicitly in Thaler (1996) and evidenced in Shefrin and Thaler (1998).

Finally, Jones (2012) proposes that inertia plays a large role in overwithholding. By analyzing taxpayer responses to changes in dependents, the expansion of the Earned Income Tax Credit, and the 1992 mandated change in default withholding,<sup>4</sup> Jones determines that taxpayer adjustments to changes in

tax liabilities or default withholding levels are often slow and either weak or non-existent. This lines up with the “default effects” seen in Madrian and Shay (2001).

Hence, recent empirical research has called into question the assertion that income tax overwithholding can be reconciled with strict consumer rationality. Nonetheless, little has been done at the theoretical level to challenge that assertion, leading to a disconnect between the theory and empiricism. That is the motivation for this paper: to reexamine HTW’s important model to determine whether it robustly predicts that significant overwithholding is a rational response to uncertainty.

### **3. Model and Results from Highfill *et al***

HTW model a risk-neutral taxpayer optimizing his level of withholding ( $W$ ) under two tax rules. A taxpayer can avoid a penalty by satisfying either of the following rules.

Rule 1: Withhold at least a set percentage ( $\alpha$ ) of the current year’s tax liability.

Rule 2: Withhold at least last year’s tax liability.

HTW refer to Rule 1 as the “90% Rule”, since they assume  $\alpha = .9$  for the years they examine (1983-1992). In reviewing historical IRS Form 1040 Instructions, however, we found that while  $\alpha$  has been .9 since 1987, prior to 1987  $\alpha$  was .8.

A taxpayer who withholds less than the minimum required amount to satisfy these rules is said to have “underwithheld.” Underwithholding taxpayers must pay a per-dollar penalty ( $j$ ) on the difference between their withholding and the minimum required level of withholding. On the other hand, if a taxpayer withholds more than enough to avoid being penalized there is a per-dollar opportunity cost ( $k$ ) that applies to that overwithholding. One can think of this opportunity cost as the return taxpayers could have earned in the market should they have invested their excess withholding instead.<sup>5</sup> The IRS sets  $j > k$  to discourage underwithholding.

In HTW’s model, the taxpayer minimizes his expected penalty or opportunity cost (PC) from withholding.<sup>6</sup> The taxpayer would like to set his withholding at the minimum rules-satisfying level.

However, this is complicated by uncertainty about his tax liability. HTW model this uncertainty by assuming that the taxpayer has certain and uncertain income ( $Y_C$  and  $Y$  respectively) and certain and uncertain allowances ( $A_C$  and  $A$  respectively). They further assume that  $Y$  and  $A$  are bounded below by 0 and above by  $V$  and  $U$  respectively, and that the variables have the joint probability distribution function  $f(A, Y)$ . Also, the taxpayer faces a fixed tax rate ( $t$ ), so that his tax liability is  $t(Y_C + Y - A_C - A) = t(Y - A) + m$ , where  $m \equiv t(Y_C - A_C)$ . For mathematical convenience they assume  $\alpha m < W < \alpha m + \alpha t(V - U)$ .<sup>7</sup>

### 3.1 Rule 1

For simplicity, HTW consider Rule 1 alone at first. The assumptions above imply the following penalty or opportunity cost (PC) functions associated with a given level of withholding ( $W$ ):

$$PC = \begin{cases} j[\alpha(t(Y - A) + m) - W] & \text{if } W \leq \alpha[t(Y - A) + m] \\ k[W - \alpha(t(Y - A) + m)] & \text{if } W > \alpha[t(Y - A) + m] \end{cases} \quad (1)$$

Since  $Y$  and  $A$  are random variables, the risk-neutral taxpayer sets  $W$  to minimize:

$$E[PC] = \int_0^U \int_{\frac{W}{\alpha t} - \frac{m}{t} + A}^V j [\alpha(t(Y - A) + m) - W] f(A, Y) dY dA \quad (2)$$

$$+ \int_0^U \int_0^{\frac{W}{\alpha t} - \frac{m}{t} + A} k [W - \alpha(t(Y - A) + m)] f(A, Y) dY dA$$

Solving the model yields the result that the taxpayer optimally sets  $W$  so that:

$$\Pr(\text{underwithholding}) = \frac{k}{j+k} \quad (3)$$

$$\Pr(\text{overwithholding}) = \frac{j}{j+k} \quad (4)$$

Figure 1, reproduced from HTW, gives a graphical interpretation of this result. The line  $Y = \frac{W}{\alpha t} - \frac{m}{t} + A$  is the "no penalty-cost" line: the line where the taxpayer neither overwithholds nor underwithholds. When  $(A, Y)$  is above this line (i.e.  $W < m + t(Y - A)$ ), the taxpayer has underwithheld. When  $(A, Y)$  is below this line (i.e.  $W > m + t(Y - A)$ ), the taxpayer has overwithheld. The optimizing taxpayer sets  $W$  (shifting the no

penalty-cost line) so that the probabilities of the areas above and below the line satisfy equations (3) and (4). As the penalty to underwithholding ( $j$ ) increases, the taxpayer withholds more, shifting the line up, decreasing the probability of underwithholding. As the opportunity cost to overwithholding ( $k$ ) increases, the taxpayer withholds less, shifting the line down, decreasing the probability of overwithholding. Given realistic values for  $j$  and  $k$ , the probability of overwithholding is substantial. For example, if  $j=.07$  and  $k=.035$ , the taxpayer will select  $W$  such that his probability of overwithholding is  $2/3$ .

### 3.2 Rules 1 and 2

Under Rules 1 and 2 together, a penalty of  $j$  percent is assessed on the deficit of one's withholding from  $M = \min\{\alpha[t(Y-A) + m], N\}$ ,<sup>8</sup> the lesser of the minimum withholding required under each rule. The penalty-cost functions are then:

$$PC = \begin{cases} j(M - W) & \text{if } W \leq M \\ k(W - M) & \text{if } W > M \end{cases} \quad (5)$$

By setting up an analogous minimization problem to that used under Rule 1, it can be shown that the taxpayer sets  $W$  according to the following rules:<sup>9</sup>

$$\begin{aligned} \text{Pr(overwithholding)} = \text{Pr(Area 3)} &= \frac{j}{j+k} \text{ if } \text{Pr(Area 1)} \leq \frac{k}{j+k} \\ W = N &\text{ if } \text{Pr(Area 1)} > \frac{k}{j+k} \end{aligned} \quad (6)$$

where Areas 1 and 3 are as depicted in Figure 2, again reproduced from HTW.

In Figure 2, the bottom line is the no penalty-cost line, which represents the set of  $(A, Y)$  for which the taxpayer just satisfies either Rule 1 or Rule 2. Below it, Area 3, the taxpayer has overwithheld. In addition, we have a "rules equal line",  $Y = \frac{N}{\alpha t} - \frac{m}{t} + A$ , which derives its name from the fact that along it the two rules are equivalent. Above it,  $N$  is less than  $\alpha t(A - Y)$  and Rule 2 requires less withholding than Rule 1. Below it,  $N$  is greater than  $\alpha t(A - Y)$  and Rule 1 requires less withholding than Rule 2. HTW note that rational taxpayers will never set  $W > N$ , implying that the highest the rules equal line gets is the no penalty-

cost line. But they may set  $W < N$  if  $N$  is high enough and there is a significant probability that  $\alpha$  of this year's tax liability is less than the previous year's tax liability. If they do, the underwithholding penalty depends on whether  $(A, Y)$  puts the taxpayer in Area 2 or Area 1. One important implication here is that a taxpayer's probability of overwithholding is weakly lower under Rules 1 and 2 than just Rule 1, since  $W$  can only be lower (not higher) under both rules.

A surprising point of note is that taxpayers do not seem to adhere to the rule that one should set  $W \leq N$ . Using the Michigan Tax Panel of IRS Statistics of Income data, we found that for most years from 1980-1990, between 63 and 72% of taxpayers set  $W > N$ ! The sole exception was 1987, when there were major changes in taxes and withholding rates due to the Tax Reform Act of 1986.<sup>10</sup>

### 3.3 The Refund Rate

The refund rate in the United States has historically been quite high. In the time period HTW examine the average refund rate was 74.4%. One of the major results achieved by HTW is that their model predicts similarly high rates. They solve for a theoretical range for the refund rate under three key assumptions. First, they assume that Rule 2 is unlikely to bind, allowing them to just impose Rule 1. Second, they assume that  $f(A, Y)$  is uniform. Third, they assume that  $m=0$ , so the taxpayer has no certain tax liability. Under these assumptions, the optimal level of withholding is:

$$W = \frac{\alpha t(2jV - (j+k)U)}{2(j+k)} \quad (7)$$

Since taxpayers receive a refund when they have withheld more than their liability, i.e. when  $W > t(Y-A)$ , the refund rate ( $R$ ) is then:

$$R = \frac{\alpha j}{j+k} + \frac{(1-\alpha)U}{2V} \quad (8)$$

HTW point out that since they assume  $V > U > 0$ , it must be that  $0 < \frac{U}{V} < 1$ , so:

$$\frac{\alpha j}{j+k} < R < \frac{\alpha j}{j+k} + \frac{1-\alpha}{2} \quad (9)$$



They compare the above theoretical interval for the refund rate with actual refund rates from 1983 to 1992 in a table reproduced in Table 1. The average predicted range of refund rates is 64.2-69.4, only a few percentage points below the actual average of 74.4. Hence their conclusion that penalty avoidance in the presence of uncertainty largely explains income tax overwithholding and the refund rate.

#### 4. Adjusted Model and Results

In this section we modify and extend HTW's model. Proofs for all non-immediate results are in the Appendix.

##### 4.1 Interest Earned While Underwithholding

One factor overlooked in HTW's model is the interest that taxpayers can earn on their withholding deficit. While the IRS will penalize this deficit at rate  $j$ , this money can be invested to earn interest at rate  $k$ . Therefore, the effective penalty on any dollar underwithheld is  $j-k$ , not  $j$ . Accounting for this, the penalty-cost functions under Rule 1 and Rules 1 and 2 are respectively:

$$PC = \begin{cases} (j - k)[\alpha(t(Y - A) + m) - W] & \text{if } W \leq \alpha[t(Y - A) + m] \\ k[W - \alpha(t(Y - A) + m)] & \text{if } W > \alpha[t(Y - A) + m] \end{cases} \quad (10)$$

$$PC = \begin{cases} (j - k)(M - W) & \text{if } W \leq M \\ k(W - M) & \text{if } W > M \end{cases} \quad (11)$$

This implies that under Rule 1 a taxpayer sets  $W$  to satisfy:

$$\Pr(\text{underwithholding}) = \frac{k}{j} \quad (12)$$

$$\Pr(\text{overwithholding}) = 1 - \frac{k}{j} \quad (13)$$

Under Rules 1 and 2 a taxpayer now sets  $W$  to satisfy:

$$\Pr(\text{overwithholding}) = \Pr(\text{Area 3}) = 1 - \frac{k}{j} \text{ if } \Pr(\text{Area 1}) \leq \frac{k}{j} \quad (14)$$

$$W = N \text{ if } \Pr(\text{Area 1}) > \frac{k}{j}$$

In comparing these results to those in Section 3 we see that the probability of overwithholding has decreased for positive  $j$  and  $k$ . This difference can be substantial. For example, if  $j=.07$  and  $k=.035$ , then  $1 - \frac{k}{j} = \frac{1}{2} < \frac{2}{3} = \frac{j}{j+k}$ . We continue to apply these new results in the rest of the paper.

#### 4.2 Enforcing Consistent Boundary Conditions

HTW calculate the range of refund rates by assuming  $0 < \frac{U}{V} < 1$ . However, to derive the theory underlying those calculations, they make a different assumption:  $\alpha m < W < \alpha m + \alpha t(V-U)$ . That different assumption dictates that even tighter bounds must be applied to  $\frac{U}{V}$ . To see this, note that when  $m = 0$ ,  $0 < W < \alpha t(V-U)$ , bounds which are depicted in Figure 3. To ensure that  $W > 0$ , it must be the case that  $\Pr[\text{Area } T_2] < 1 - \frac{k}{j}$ . Similarly, to ensure that  $W < \alpha t(V-U)$ , it must be the case that  $\Pr[\text{Area } T_1] < \frac{k}{j}$ . But under the assumption that  $A$  and  $Y$  are uniformly distributed,  $\Pr[\text{Area } T_1] = \Pr[\text{Area } T_2] = \frac{U}{2V}$ . So to predict refund rates, the boundary conditions from the theory imply that:<sup>11</sup>

$$0 < \frac{U}{V} < 2 \min\left\{\frac{k}{j}, 1 - \frac{k}{j}\right\} \quad (15)$$

These bounds for  $\frac{U}{V}$  are tighter than those used by HTW in their refund calculations. For example, when  $j=.067$  and  $k=.052$ , as was the case in 1984, the bounds require  $U/V < .45$ . Equation (8) then tells us that the lower limits on  $U/V$  reduce the maximum predicted refund rate.<sup>12</sup>

#### 4.3 Rephrasing the Problem

So far we have maintained consistency with HTW by phrasing the withholding problem in terms of income ( $Y_c + Y$ ) and allowances ( $A_c + A$ ). However, one can rephrase the withholding problem in a way that simplifies the mathematics and elucidates some of the underlying assumptions. To start, observe that the taxpayer is

not separately concerned with either income or allowances. Rather, he is concerned with his tax liability:  $t(Y_c - A_c + Y - A) = m + t(Y - A)$ , where  $m$  is his liability stemming from certain income and allowances, and  $t(Y - A)$  is his liability stemming from uncertain income and allowances. By defining  $L = t(Y - A)$  so that  $L$  is the stochastic part of tax liability, the taxpayer's problem can be distilled into one of trying to set  $W$  given uncertainty about his current tax liability ( $m + L$ ). Under Rule 1, the taxpayer is penalized for withholding less than  $\alpha(m + L)$ , and under Rules 1 and 2, the taxpayer is penalized for withholding less than  $M = \min\{\alpha(m + L), N\}$ .

The HTW assumption that  $f(A, Y)$  is uniform over  $[0, U] \times [0, V]$  translates into the assumption that the PDF of  $L$ , which we denote  $f(L)$ , is:

$$f(L) = \begin{cases} \frac{1}{tV} + \frac{L}{t^2UV} & \text{if } -tU \leq L \leq 0 \\ \frac{1}{tV} & \text{if } 0 < L < t(V - U) \\ \frac{1}{tU} - \frac{L}{t^2UV} & \text{if } t(V - U) \leq L \leq tV \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

In other words,  $f(L)$  has the trapezoidal distribution depicted in Figure 4. The tails of the trapezoidal distribution correspond to  $T_1$  and  $T_2$  in the previous section, which we have also marked in Figure 4. Additionally, we have marked the range of  $W$  implied by assumptions with  $W_{\min}$  and  $W_{\max}$ . Area 3 from Figure 2 is now represented by the interval  $[-tU, \frac{W}{\alpha}]$ , and Areas 2 and 1 are represented by the intervals  $[\frac{W}{\alpha}, \frac{N}{\alpha}]$  and  $[\frac{N}{\alpha}, tV]$  respectively. Therefore, the point  $L = \frac{W}{\alpha}$  is analogous to the no penalty-cost line because it represents the only case in which the taxpayer has neither under nor overwithheld. Moreover, if  $L < \frac{W}{\alpha}$  the taxpayer has overwithheld, and if  $L > \frac{W}{\alpha}$ , the taxpayer has underwithheld.<sup>13</sup> Similarly, the point  $L = \frac{N}{\alpha}$  is analogous to the rules equal line.

Our model still, however, cannot satisfactorily estimate a refund rate under Rules 1 and 2. The piecewise form of  $f(L)$  requires either for the analysis to be broken down into so many cases as to be intractable, or for variables such as  $W$  to be severely limited, as they were in HTW. The limited range of  $W$ ,

though, is problematic because it allows  $T_2$  to be an area in which the taxpayer always receives a refund, no matter what the taxpayer's preferences. This can generate perverse results. For example, under the given assumptions,  $\Pr[\text{Area } T_1] = \frac{U}{2V}$ , so as  $\frac{U}{V} \rightarrow 1$ ,  $\Pr[\text{Area } T_1] \rightarrow \frac{1}{2}$ . In other words, as  $U$  approaches  $V$  (or as allowance uncertainty approaches income uncertainty), the model by construction predicts that the upper-bound for the refund rate is at least 50%, a very high number.

Fortunately, we can circumvent these problems by assuming that  $L$  is uniformly distributed between 0 and  $T = t(V-U) > 0$ , where  $T$  is the taxpayer's maximum possible tax liability stemming from uncertain income and allowances. This new  $f(L)$  does not presuppose either a minimum rate of overwithholding or a minimum rate of underwithholding, and it does not require that we place any bounds on  $W$ . We depict  $f(L)$  under this assumption in Figure 5. Area 3 is now the interval  $[0, \frac{W}{\alpha}]$  and Area 1 is the interval  $[\frac{N}{\alpha}, T]$ , while Area 2 remains the same.

Importantly, this distributional assumption does not differ greatly from that made by HTW, either in form or effect. One can think of it as truncating the previous distribution so that  $\Pr(\text{Area } T_1) = \Pr(\text{Area } T_2) = 0$ . Hence we have simply removed the problematic cases. In doing so, we have not greatly changed the results: under the new distributional assumptions, the refund rate calculated under Rule 1 is equivalent to the lower-bound calculated under Rule 1 given the previous assumptions. It can easily be shown that the difference between the upper and lower bounds of the refund rate when enforcing consistent boundary conditions from the previous section is only  $(1-\alpha)\min\{\frac{k}{j}, 1-\frac{k}{j}\} \leq \frac{1-\alpha}{2} = .05$  when  $\alpha=.9$ , or less than five percentage points. So the new assumed distribution for  $L$  does not dramatically change refund estimates and it importantly makes the model tractable enough to impose both Rules 1 and 2.

#### 4.4 Refund Rate Under Rules 1 and 2

HTW restrict their refund analysis to Rule 1, arguing that one needn't examine the refund rate under Rule 2 because Rule 2 will bind infrequently. The exact condition for Rule 2 to bind – so Rules 1 and 2 yield different levels of withholding than Rule 1 alone – is  $\Pr(\text{Area 1}) > 1 - \frac{k}{j}$ . Since now

$$\Pr(\text{Area 1}) = \Pr(L > \frac{N}{\alpha}) = \Pr(\alpha L > N),$$

we can think of  $\Pr(\text{Area 1})$  as the probability that  $\alpha$  of this year's tax liability exceeds last year's tax liability. Given worker salaries rising with experience and inflation, it does not seem improbable that a taxpayer will experience such an increase in nominal liability. In fact, using the Michigan Federal Income Tax Panel we found that in each of the years from 1980 to 1990 (excluding the tax reform year of 1987), between 27 and 42% of taxpayers saw such an increase in their liability. So there may be merit to examining the refund rate under Rule 2 too. Therefore, in this section, we calculate the refund rate under Rules 1 and 2.

For our analysis we assume that  $N$  has the same PDF as  $m+L$ , or that taxpayers face the same range of tax liabilities this year as they did last year. Because this implies stagnant nominal earnings, while most people face rising nominal earnings over time, we are understating the frequency with which Rule 2 binds and overstating the frequency of overwithholding as a response to uncertainty.

Under these conditions, the taxpayer sets  $W$  as follows:

$$W = \begin{cases} \alpha T (1 - \frac{k}{j}) & \text{if } N \geq (1 - \frac{k}{j}) \alpha T, \\ N & \text{if } N < (1 - \frac{k}{j}) \alpha T. \end{cases} \quad (17)$$

Therefore, the probability of receiving a refund given  $N=n$ , which we shall denote  $R_n$ , is:

$$R_n = \begin{cases} \alpha (1 - \frac{k}{j}) & \text{if } n \geq (1 - \frac{k}{j}) \alpha T, \\ \frac{n}{T} & \text{if } n < (1 - \frac{k}{j}) \alpha T. \end{cases} \quad (18)$$

So the refund rate can be calculated as:

$$R = \int_0^T R_n \Pr[N = n] dn \quad (19)$$

We solve and find:

$$R = \alpha \left(1 - \frac{k}{j}\right) - \frac{\alpha^2}{2} \left(1 - \frac{k}{j}\right)^2 . \quad (20)$$

This refund rate is lower than under Rule 1 whenever  $0 < k < j$ . Table 2 is analogous to Table 1, and it demonstrates that when we add Rule 2 and the effective penalty  $(j-k)$ , there is a substantial gap between the theoretical and actual refund rates. Now, on average, the model explains only half of the actual refund rate.

#### 4.5 Nonzero Certain Liability

Another assumption made in HTW's calculations of refund rates is that taxpayers have no certain tax liability, i.e.  $m=0$ . But this seems unrealistic because taxpayers with stable employment probably anticipate having some minimum tax liability. Mathematically, this assumption implies a great deal of uncertainty in tax liability: it implies that the coefficient of variation of  $L$ ,  $\frac{\sigma_L}{\mu_L}$ , is almost 58%. In this section, we relax the assumption of zero certain liability, letting  $m \geq 0$ , and find that it matters.

We maintain the assumption that the taxpayer's current and previous year's liabilities are drawn from the same distribution. Therefore,  $N$  is uniformly distributed between  $m$  and  $m+T$ . Under these assumptions:

$$\Pr(\text{Area 3}) = \frac{W-m}{T},$$

$$\Pr(\text{Area 1}) = \frac{m+T-N}{T},$$

From equation (14), the taxpayer sets  $W$  in the following manner:

$$W = \begin{cases} \alpha \left(1 - \frac{k}{j}\right) T + \alpha m & \text{if } N \geq \alpha \left(1 - \frac{k}{j}\right) T + \alpha m \\ N & \text{if } N < \alpha \left(1 - \frac{k}{j}\right) T + \alpha m \end{cases} \quad (21)$$

Therefore, the refund rate is:

$$R = \begin{cases} R^\circ - (1 - \alpha)(1 - \alpha D) \frac{m}{T} - \frac{(1 - \alpha)^2}{2} \left(\frac{m}{T}\right)^2 & \text{if } D \geq \frac{(1 - \alpha) m}{\alpha T} \\ 0 & \text{if } D < \frac{(1 - \alpha) m}{\alpha T} \end{cases} \quad (22)$$

where  $D = (1 - \frac{k}{j})$  and  $R^\circ$  denotes the refund rate given in equation (20).

One can think of  $\frac{m}{T}$  as the ratio of certain liability to uncertain liability. Figure 6 depicts  $R$  as a function of  $\frac{k}{j}$  for different values of  $\frac{m}{T}$  when  $\alpha = .9$ . Figure 7 shows the same functions when  $\alpha = .8$ . It is clear from these figures that allowing for some certain tax liability makes a difference. It reduces the probability of receiving a refund, and the greater the proportion of certain to uncertain liability, the less likely taxpayers are to receive a refund. This makes sense. An increase in certain liability widens the margin in which the taxpayer may satisfy Rule 1 without receiving a refund, but does not yield a corresponding increase in uncertainty about liability. In effect it makes it easier for the taxpayer to avoid underwithholding, but without withholding so much as to receive a refund.

Furthermore, the effect on the refund rate is significant. Table 3 compares actual refund rates to theoretical ones when  $\frac{m}{T} = 2$  for the period 1983-1992. Assuming that  $\frac{m}{T} = 2$  means that there is a 5% chance that a taxpayer's liability will deviate more than 19% from his expected liability. For most taxpayers, such large unexpected changes in liability are unlikely. One possible reason for a large unexpected change in liability is involuntary job loss. Letting  $\frac{m}{T} = 2$  actually aligns reasonably well with real rates of unexpected job loss: Boisjoly *et al* (1998) found that the rate of involuntary job loss averaged 3.2% annually for men in the period 1980-1992. The rate of unexpected job gain is probably similar, but when individuals are hired, they are asked to file a new Form W-4, allowing them to adjust their withholding rate appropriately.

The theoretical estimates in Table 3 are considerably lower than those in Tables 1 or 2. This is especially true for the years 1983-1986 in which  $\alpha = .8$ . In fact, for 1983, 1984, and 1986, the theoretical

refund rate is zero. For the entire period 1983-1992, the theoretical refund rate averages 19.3%, or roughly a quarter of the actual refund rate.

#### 4.6 Applying a Normal Distribution

Until now we have assumed uniform distributions for L and N. But under a uniform distribution, all tax liabilities, no matter how far or close to the taxpayer's expectation, are equally probable. So for instance, having zero tax liabilities this year would be just as likely as having the same level of tax liabilities that you had last year. One might, instead, expect that tax liabilities closer to the taxpayer's expectation (and probably last year's level) are more probable than those farther away. In this section, we impose that condition. We let  $m=0$ , but assume  $L \sim N(\mu_L, \sigma_L)$ , and  $N \sim N(\mu_N, \sigma_N)$ , so that taxable income and previous year's tax liability are normally distributed. These assumptions can be viewed as a different approach to adding more certainty to a taxpayer's tax liabilities.<sup>14</sup>

Given these assumptions, a taxpayer sets W in the following manner under Rules 1 and 2:

$$W = \begin{cases} \alpha \left[ \sigma_L \Phi^{-1} \left( 1 - \frac{k}{j} \right) + \mu_L \right] & \text{if } N \geq \alpha \left[ \sigma_L \Phi^{-1} \left( 1 - \frac{k}{j} \right) + \mu_L \right] \\ N & \text{if } N < \alpha \left[ \sigma_L \Phi^{-1} \left( 1 - \frac{k}{j} \right) + \mu_L \right] \end{cases} \quad (23)$$

where  $\Phi$  is the CDF of the standard normal distribution. Therefore, the probability of receiving a refund given  $N=n$ , which we shall denote  $R_n$ , is:

$$R_n = \begin{cases} \Phi \left( \frac{\alpha \sigma_L \Phi^{-1} \left( 1 - \frac{k}{j} \right) - (1-\alpha)\mu_L}{\sigma_L} \right) & \text{if } n \geq \alpha \left[ \sigma_L \Phi^{-1} \left( 1 - \frac{k}{j} \right) + \mu_L \right] \\ \Phi \left( \frac{n - \mu_L}{\sigma_L} \right) & \text{if } n < \alpha \left[ \sigma_L \Phi^{-1} \left( 1 - \frac{k}{j} \right) + \mu_L \right] \end{cases} \quad (24)$$

The refund rate can thus be calculated as:

$$R = \int_{-\infty}^{\infty} R_n \Pr[N = n] dn. \quad (25)$$



We can substitute in  $\Pr[N=n] = \phi\left(\frac{n-\mu_N}{\sigma_N}\right)$ , where  $\phi$  is the PDF of a standard normal distribution, and  $R_n$  from equation (24) to find:

$$R = \int_{-\infty}^{\alpha[\sigma_L\Phi^{-1}\left(1-\frac{k}{j}\right)+\mu_L]} \Phi\left(\frac{n-\mu_L}{\sigma_L}\right) \phi\left(\frac{n-\mu_N}{\sigma_N}\right) dn + \Phi\left(\frac{\alpha\sigma_L\Phi^{-1}\left(1-\frac{k}{j}\right)-(1-\alpha)\mu_L}{\sigma_L}\right) \left(1 - \Phi\left(\frac{\alpha[\sigma_L\Phi^{-1}\left(1-\frac{k}{j}\right)+\mu_L]-\mu_N}{\sigma_N}\right)\right) \quad (26)$$

In order to meaningfully compare the refund rate when liability (L) is normally distributed with when it is uniformly distributed we assume that  $\mu_L = \frac{T}{2}$  and define  $s = \frac{\mu_L}{\sigma_L}$  to be the inverse of the coefficient of variation of L. Thus, L is normally distributed with the same mean as the uniform distribution and with a variance such that 2s standard deviations fit in [0,T], the support under the uniform distribution.

As in the previous section we assume that N and L are identically distributed (i.e.  $\mu_N = \mu_L$  and  $\sigma_N = \sigma_L$ ), which should overstate the refund rate. Substituting into equation (26) and changing the variable of integration to  $x = \frac{n-\mu_L}{\sigma_L} = \frac{n-\mu_N}{\sigma_N}$ , we find:

$$R = \int_{-\infty}^{\alpha\Phi^{-1}\left(1-\frac{k}{j}\right)-(1-\alpha)s} \Phi(x)\phi(x)dx + \Phi\left(\alpha\Phi^{-1}\left(1-\frac{k}{j}\right)-(1-\alpha)s\right) \left(1 - \Phi\left(\alpha\Phi^{-1}\left(1-\frac{k}{j}\right)-(1-\alpha)s\right)\right) \quad (27)$$

Figure 8 depicts how the refund rate is affected by different values of s when  $\alpha=.9$ .<sup>15</sup> Figure 9 shows the same function when  $\alpha=.8$ . It is clear that switching to a normal distribution makes a significant difference. One can think of s as a measure of a taxpayer's confidence that his income will be close to its expected value. The two figures indicate that as a taxpayer's confidence in the accuracy of his prediction increases, the probability of receiving a refund decreases.<sup>16</sup> They also show that assuming normality potentially yields much lower estimates of the refund rate, especially when s is large. Predicted refund rates are also much smaller when  $\alpha = .8$ .

Table 4 gives the theoretical and actual refund rates for 1983-1990 assuming  $s=10$ . When  $s=10$ , a taxpayer perceives only a 5% chance that his tax liability will diverge more than 20% from his expected

liability – a similar case to when  $\frac{m}{T} = 2$  with a uniform distribution. As in the previous section, the theoretical refund rates are a small fraction of the actual ones, especially for the years 1983-1986 in which  $\alpha=.8$ . For 1983-1992 as a whole, the theoretical refund rate averages 15.6% or about a fifth of the actual refund rate. That is similar to our 19.3% estimate when we allowed  $m>0$  with a uniform distribution. Hence, these two approaches to adding more certainty to tax liabilities are generally consistent.

## 5. Conclusion

In our analysis we have scrutinized the proposition that the high rate of income tax refunds in the US can largely be explained as a risk-neutral taxpayer's rational response to uncertain tax liabilities and underwithholding penalties. Accounting for interest earned on underwithheld income and enforcing consistent boundary conditions both decrease predictions of refund rates. So does incorporating the previous year's tax liability rule and either relaxing the assumption of no certain tax liabilities or replacing the uniform distribution for uncertain tax liabilities with a normal distribution. We estimate that penalty avoidance is more likely to produce refunds rates between 15-20% than the 75% observed in our sample. These estimates are also probably conservative, as they assume no normal growth in nominal income over time, which would further depress predicted refund rates.

Given that this model assumes risk neutrality, one might wonder whether risk aversion might prove more favorable to the theory of overwithholding as a response to uncertainty. Jones (2012), however, found that when the cost to overwithholding from consumption smoothing was taken into account, an implausibly high degree of risk aversion was required to justify withholding patterns seen in the United States. Hence, our findings complement Jones' result and provide evidence for behavioral analyses of overwithholding. They also suggest that further research is warranted to determine which behavioral phenomena contribute to income tax overwithholding and refunds in the US.

## References

- Barr, Michael S. and Jane K. Dokko. 2008. "Paying to Save: Tax Withholding and Asset Allocation Among Low- and Moderate-Income Taxpayers." Finance and Economics Discussion Series Working Paper, Federal Reserve Board.
- Boisjoly, Johanne, Greg Duncan, and Timothy Smeeding. 1998. "Shifting Incidence of Involuntary Job Losses from 1968 to 1992." *Industrial Relations: A Journal of Economy and Society*, 37(2): 207-231.
- Chambers, Valrie and Marilyn Spencer. 2008. "Does Changing the Timing of a Yearly Individual Tax Refund Change the Amount Sent vs. Saved?" *Journal of Economic Psychology*, 29(6): 856-862.
- Fennell, Lee Anne (2006), "Hyperopia in Public Finance." in *Behavioral Public Finance* (Edward J. McCaffery and Joel Slemrod, eds.), Russell Sage: 141-172.
- Highfill, Jannett, Douglas Thorson, and William V. Weber. 1998. "Tax Overwithholding as a Response to Uncertainty." *Public Finance Review*, (26): 376-391.
- Internal Revenue Service. 2011. "IRS Operations, Budget & Compliance." Available at <http://www.irs.gov/taxstats/index.html>.
- Jones, Damon. 2010. "Withholdings, Saliency and Tax Policy." Stanford Institute of Economic Policy Research Policy Brief.
- \_\_\_\_\_. 2012. "Inertia and Overwithholding: Explaining the Prevalence of Income Tax Refunds." *American Economic Journal: Economic Policy*, 4(1): 158-85.
- Madrian, Brigitte and Dennis F. Shay. 2001. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics*, 116(4): 1141-1187.
- Neumark, David. "Are Rising Earnings Profiles a Forced-Savings Mechanism?" *The Economic Journal*, 105: 95-106.
- Shapiro, Matthew D. and Joel Slemrod. 1995. "Consumer Response to the Timing of Income: Evidence from a Change in Tax Withholding." *American Economic Review*, 85(1), 274-283.

Shefrin, Hersh M. and Richard H. Thaler. 1998. "The Behavioral Life-Cycle Hypothesis." *Economic Inquiry*, 26(4): 609-643.

Thaler, Richard H. 1994. "Psychology and Savings Policies." *American Economic Review*, (84)2: 186-192.

## Mathematical Appendix

### Proof of Equations 12 and 13:

The taxpayer sets  $W$  to minimize:

$$E[PC] = \int_0^U \int_{\frac{\alpha t}{t} - \frac{m}{t} + A}^V (j - k)[\alpha(t(Y - A) + m) - W]f(Y, A)dYdA$$

$$+ \int_0^U \int_0^{\frac{W}{\alpha t} - \frac{m}{t} + A} k[W - \alpha(t(Y - A) + m)]f(Y, A)dYdA$$

The results follow from taking the first order conditions with respect to  $W$ . A more detailed proof is given of equation (14).

### Proof of Equation 14:

The taxpayer sets  $W$  to minimize:

$$E[PC] = \int_0^U \int_0^{\frac{W}{\alpha t} - \frac{m}{t} + A} k(W - \alpha t(Y - A) - \alpha m)f(Y, A)dYdA$$

$$+ \int_0^U \int_{\frac{\alpha t}{t} - \frac{m}{t} + A}^{\frac{N}{\alpha t} - \frac{m}{t} + A} (j - k)(\alpha t(Y - A) + \alpha m - W)f(Y, A)dYdA$$

$$+ \int_0^U \int_{\frac{\alpha t}{t} - \frac{m}{t} + A}^V (j - k)[N - W]f(Y, A)dYdA \quad \text{subject to } N \geq W$$

The Lagrangian is:

$$\mathcal{L} = \int_0^U \int_0^{\frac{W}{\alpha t} - \frac{m}{t} + A} k(W - \alpha t(Y - A) - \alpha m)f(Y, A)dYdA$$

$$+ \int_0^U \int_{\frac{\alpha t}{t} - \frac{m}{t} + A}^{\frac{N}{\alpha t} - \frac{m}{t} + A} (j - k)(\alpha t(Y - A) + \alpha m - W)f(Y, A)dYdA$$

$$+ \int_0^U \int_{\frac{\alpha t}{t} - \frac{m}{t} + A}^V (j - k)(\alpha t(Y - A) + \alpha m - W)f(Y, A)dYdA - \lambda(N - W)$$

The first-order condition with respect to  $W$  is:

$$\int_0^U \left[ \frac{1}{\alpha t} \left( k \left[ W - \alpha t \left( \frac{W}{\alpha t} - \frac{m}{t} + A - A \right) - \alpha m \right] f \left( A, \frac{W}{\alpha t} - \frac{m}{t} + A \right) \right) + \int_0^{\frac{W}{\alpha t} - \frac{m}{t} + A} k f(A, Y) dY \right] dA$$

$$+ \int_0^U \left[ -\frac{1}{\alpha t} \left\{ (j - k) \left[ \alpha t \left( \frac{W}{\alpha t} - \frac{m}{t} + A - A \right) + \alpha m - W \right] f \left( A, \frac{W}{\alpha t} - \frac{m}{t} + A \right) \right\} \right]$$

$$- (j - k) \int_0^U \int_{\frac{W}{\alpha t} - \frac{m}{t} + A}^{\frac{N}{\alpha t} - \frac{m}{t} + A} f(A, Y) dY dA - (j - k) \int_0^U \int_{\frac{N}{\alpha t} - \frac{m}{t} + A}^V f(A, Y) dY dA + \lambda = 0$$

$$\text{Note: } \frac{1}{\alpha t} \left( k \left[ W - \alpha t \left( \frac{W}{\alpha t} - \frac{m}{t} + A - A \right) - \alpha m \right] f \left( A, \frac{W}{\alpha t} - \frac{m}{t} + A \right) \right) = 0$$

$$-\frac{1}{\alpha t} \left\{ (j - k) \left[ \alpha t \left( \frac{W}{\alpha t} - \frac{m}{t} + A - A \right) + \alpha m - W \right] f \left( A, \frac{W}{\alpha t} - \frac{m}{t} + A \right) \right\} = 0$$

Therefore the first order condition is:

$$k \int_0^U \int_0^{\frac{W}{\alpha t} - \frac{m}{t} + A} f(A, Y) dY dA - (j - k) \int_0^U \int_{\frac{W}{\alpha t} - \frac{m}{t} + A}^{\frac{N}{\alpha t} - \frac{m}{t} + A} f(A, Y) dY dA$$

$$- (j - k) \int_0^U \int_{\frac{N}{\alpha t} - \frac{m}{t} + A}^V f(A, Y) dY dA + \lambda = 0$$

This is equivalent to the following:  $k\text{Pr}[\text{Case 3}] - (j-k)\text{Pr}[\text{Case 2}] - (j-k)\text{Pr}[\text{Case 1}] = -\lambda$

We find consistency when  $N-W > 0$  and  $\lambda=0$ . In this case our first order condition is equivalent to:

$$k\text{Pr}[\text{Case 3}] - (j-k)\text{Pr}[\text{Case 2}] - (j-k)\text{Pr}[\text{Case 1}] = 0 \Leftrightarrow k\text{Pr}[\text{Case 3}] - (j-k)(\text{Pr}[\text{Case 2}] + \text{Pr}[\text{Case 1}]) = 0$$

$$\Leftrightarrow k\text{Pr}[\text{Case 3}] - (j-k)(1 - \text{Pr}[\text{Case 3}]) = 0 \Leftrightarrow k\text{Pr}[\text{Case 3}] = j - k - (j-k)\text{Pr}[\text{Case 3}] \Leftrightarrow j\text{Pr}[\text{Case 3}] = j - k$$

$$\Leftrightarrow \text{Pr}[\text{Case 3}] = 1 - \frac{k}{j} \Rightarrow \text{Pr}[\text{Case 1}] \leq \frac{k}{j}$$

Additionally, we find the opposite consistency that holds when  $W = N$  and  $\Pr[\text{Case 1}] > \frac{k}{j}$ . It is worth noting that if  $1 - \frac{k}{j} < \int_0^U \int_0^A f(A, Y) dY dA$ , the taxpayer sets  $W=0$ . This is intuitive: if one would optimally like to set  $W < 0$ , one will set  $W=0$ . This case, however, is ignored in HTW and this section of our analysis.

### Proof of Equation 20

For notational simplicity we denote  $D = 1 - \frac{k}{j}$ . Then, from equation (19):

$$R = \int_0^T R_n \Pr[N = n] dn = \int_0^{D\alpha T} R_n \Pr[N = n] dn + \int_{D\alpha T}^T R_n \Pr[N = n] dn.$$

We note that  $P[N=n] = \frac{1}{T}$  and substitute from equation (18) and find that:

$$R = \int_0^{D\alpha T} \frac{n}{T^2} dn + \int_{D\alpha T}^T \alpha D \frac{1}{T} dn = \frac{D^2 \alpha^2}{2} + \alpha D(1 - \alpha D) = \alpha D - \frac{\alpha^2 D^2}{2}.$$

By substituting  $D = 1 - \frac{k}{j}$ , we have:  $R = \alpha \left(1 - \frac{k}{j}\right) - \frac{\alpha^2}{2} \left(1 - \frac{k}{j}\right)^2$ .

### Proof of Equation 22:

Note that if  $W < m$ ,  $R=0$ . This can only happen if  $\alpha(DT+m) < m$ , which is when  $D < \frac{(1-\alpha)m}{\alpha T}$ . For the

following analysis assume that  $D$  is sufficiently large when compared to  $\frac{m}{t}$  that  $W \geq m$ . Then,

when  $N=n$ , we have that the refund rate  $R_n$  can be given as:

$$R_n = \begin{cases} \alpha D - (1 - \alpha) \frac{m}{T} & \text{if } n \geq \alpha(DT + m), \\ \frac{n - m}{T} & \text{if } n < \alpha(DT + m). \end{cases}$$

Substituting in  $n' = n - m$ , we have:

$$\begin{aligned}
R &= \int_m^{m+T} R_n \Pr[N = n] dn = \int_0^{\alpha DT - (1-\alpha)m} \frac{n'}{T^2} dn' + \int_{\alpha DT - (1-\alpha)m}^T \left( \frac{\alpha D}{T} - (1-\alpha) \frac{m}{T^2} \right) dn' \\
&= \frac{1}{2T^2} (\alpha DT - (1-\alpha)m)^2 + (1 - \alpha D + (1-\alpha) \frac{m}{T}) (\alpha D - (1-\alpha) \frac{m}{T}) \\
&= \frac{1}{2} (\alpha D - (1-\alpha) \frac{m}{T})^2 + \alpha D - (1-\alpha) \frac{m}{T} - (\alpha D - (1-\alpha) \frac{m}{T})^2 \\
&= \alpha D - (1-\alpha) \frac{m}{T} - \frac{1}{2} (\alpha^2 D^2 - 2\alpha(1-\alpha)D \frac{m}{T} + (1-\alpha)^2 (\frac{m}{T})^2) \\
&= R^\circ - (1-\alpha)(1-\alpha D) \frac{m}{T} - \frac{(1-\alpha)^2}{2} (\frac{m}{T})^2.
\end{aligned}$$

**Proof of Equation 23:**

Rule 2 does not bind only when:

$$P\left(L < \frac{N}{\alpha} \mid N = n\right) = \Phi\left(\frac{\frac{n}{\alpha} - \mu_L}{\sigma_L}\right) \geq 1 - \frac{k}{j} \Leftrightarrow N \geq \alpha \left[ \sigma_L \Phi^{-1}\left(1 - \frac{k}{j}\right) + \mu_L \right].$$

In such a case, the taxpayer sets  $W$  so that:

$$P\left(L < \frac{W}{\alpha}\right) = \Phi\left(\frac{\frac{W}{\alpha} - \mu_L}{\sigma_L}\right) = 1 - \frac{k}{j} \Leftrightarrow W = \alpha \left[ \sigma_L \Phi^{-1}\left(1 - \frac{k}{j}\right) + \mu_L \right].$$

**Proof of Equation 26:**

We see that:

$$\begin{aligned}
R &= \int_{-\infty}^{\alpha[\sigma_L \Phi^{-1}(1-\frac{k}{j}) + \mu_L]} \Phi\left(\frac{n - \mu_L}{\sigma_L}\right) \phi\left(\frac{n - \mu_N}{\sigma_N}\right) dn \\
&\quad + \int_{\alpha[\sigma_L \Phi^{-1}(1-\frac{k}{j}) + \mu_L]}^{\infty} \Phi\left(\frac{\alpha \sigma_L \Phi^{-1}(1-\frac{k}{j}) - (1-\alpha)\mu_L}{\sigma_L}\right) \Phi\left(\frac{n - \mu_L}{\sigma_L}\right) dn.
\end{aligned}$$

Equation (26) then follows by pulling the CDF out of the second integral.



## Tables and Figures

Table 1: HTW's Actual and Theoretical Refund Rates, 1983-1992

Year	Actual Refund Rate	Underwithholding Penalty	Overwithholding Opportunity Cost	Theoretical Refund Rate
1983	77.9	10.3	5.4	59.0-64.0
1984	76.1	6.7	5.2	50.7-55.7
1985	76.1	8.4	3.9	61.5-66.5
1986	75.8	7.6	4.1	58.5-63.5
1987	74.6	5.6	2.2	64.6-69.6
1988	72.1	6.4	2.6	64.0-69.0
1989	72.2	6.7	3.3	60.3-65.3
1990	73.4	5.6	2.1	65.5-70.5
1991	74.6	6.0	1.2	75.0-80.0
1992	71.3	5.0	0.4	83.3-88.3
Mean	74.4	6.8	3.0	64.2-69.2
SD	2.02	1.5	1.56	

Table 2: Actual and Theoretical Refund Rates Recalculated, 1983-1992

Year	Actual Refund Rate	Underwithholding Penalty	Overwithholding Opportunity Cost	Theoretical Refund Rate
1983	77.9	10.3	5.4	30.8
1984	76.1	6.7	5.2	16.3
1985	76.1	8.4	3.9	33.7
1986	75.8	7.6	4.1	30.1
1987	74.6	5.6	2.2	39.7
1988	72.1	6.4	2.6	39.2
1989	72.2	6.7	3.3	35.2
1990	73.4	5.6	2.1	40.4
1991	74.6	6.0	1.2	46.1
1992	71.3	5.0	0.4	48.5
Mean	74.4	6.8	3.0	36.0
SD	2.02	1.5	1.56	9.18

Table 3: Actual and Theoretical Refund Rates when  $\frac{m}{T} = 2$ , 1983-1992

Year	Actual Refund Rate	Underwithholding Penalty	Overwithholding Opportunity Cost	Theoretical Refund Rate
1983	77.9	10.3	5.4	0
1984	76.1	6.7	5.2	0
1985	76.1	8.4	3.9	2.8
1986	75.8	7.6	4.1	0
1987	74.6	5.6	2.2	28.6
1988	72.1	6.4	2.6	27.9
1989	72.2	6.7	3.3	22.4
1990	73.4	5.6	2.1	29.7
1991	74.6	6.0	1.2	38.5
1992	71.3	5.0	0.4	43.1
Mean	74.4	6.8	3.0	19.3
SD	2.02	1.5	1.56	17.0

Table 4: Actual and Theoretical Refund Rates when  $s=10$ , 1983-1992

Year	Actual Refund Rate	Underwithholding Penalty	Overwithholding Opportunity Cost	Theoretical Refund Rate
1983	77.9	10.3	5.4	2.0
1984	76.1	6.7	5.2	0.05
1985	76.1	8.4	3.9	2.6
1986	75.8	7.6	4.1	1.8
1987	74.6	5.6	2.2	20.0
1988	72.1	6.4	2.6	19.3
1989	72.2	6.7	3.3	15.0
1990	73.4	5.6	2.1	21.0
1991	74.6	6.0	1.2	32.3
1992	71.3	5.0	0.4	42.2
Mean	74.4	6.8	3.0	15.6
SD	2.02	1.5	1.56	14.3

Figure 1: HTW's Taxpayer's Decision Under Rule 1 (Replicated from HTW)

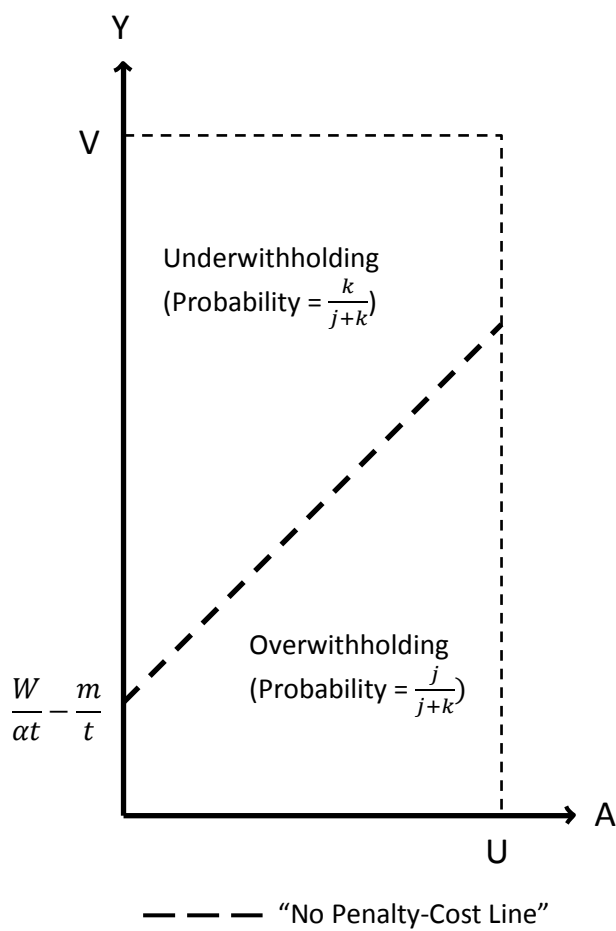


Figure 2: HTW's Taxpayer's Decision Under Rules 1 and 2 (Replicated from HTW)

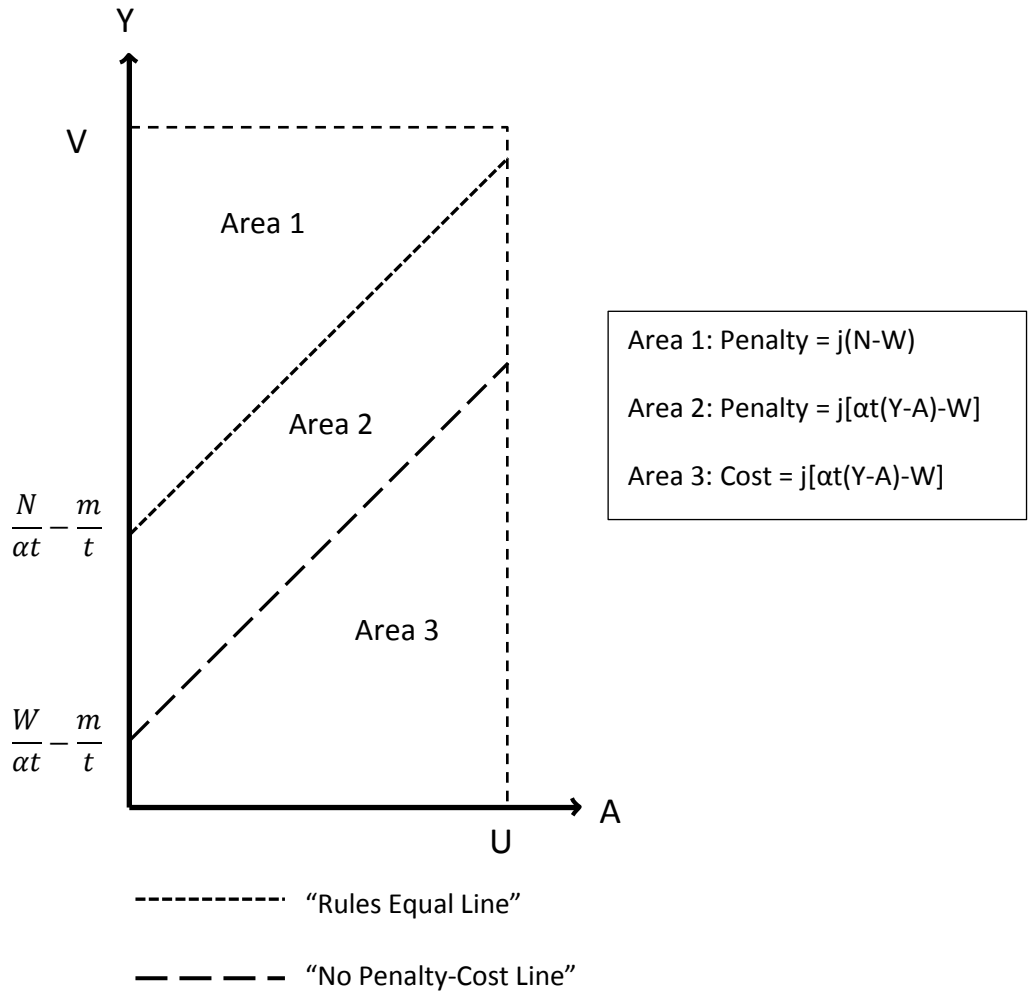


Figure 3: Boundary Conditions ( $m = 0$ )

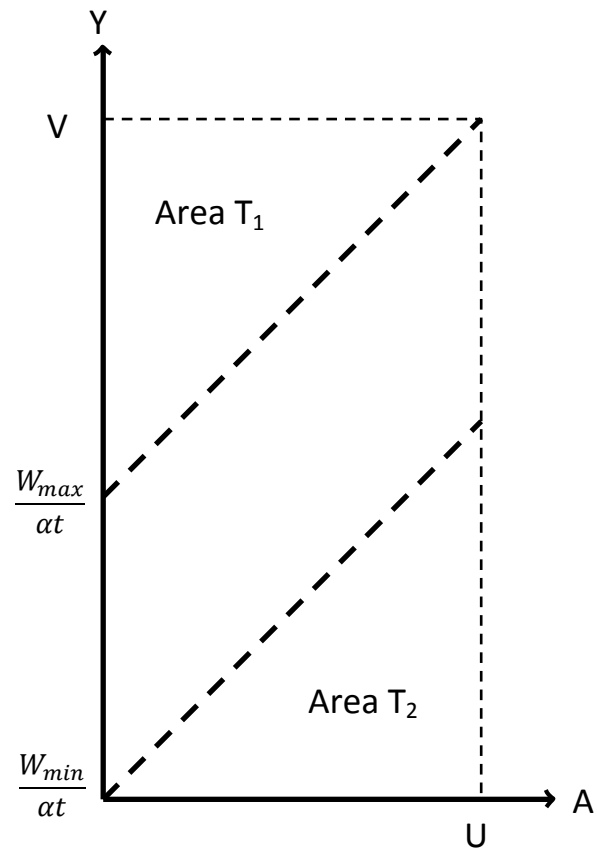
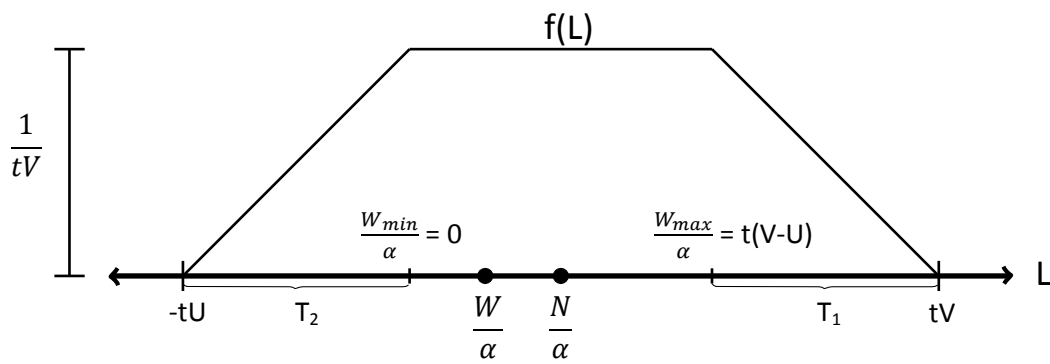


Figure 4:  $f(L)$  Under HTW's Assumptions



tV

Figure 5: Uniformly Distributed Liability ( $m = 0$ )

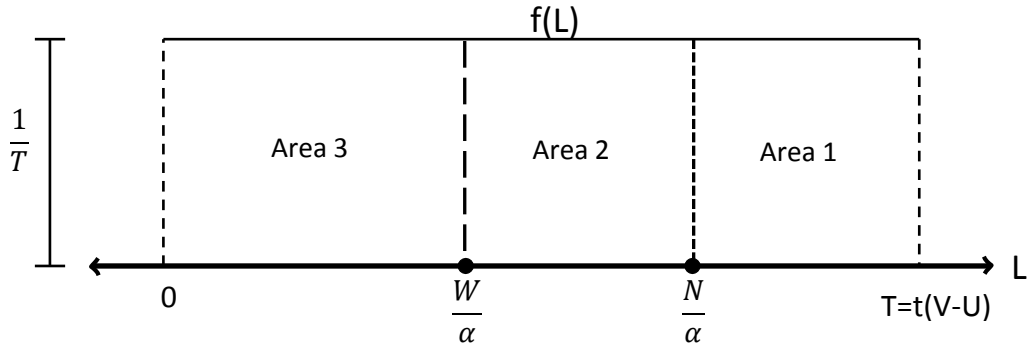


Figure 6: Refund Rates When  $m \geq 0$  ( $\alpha = .9$ )

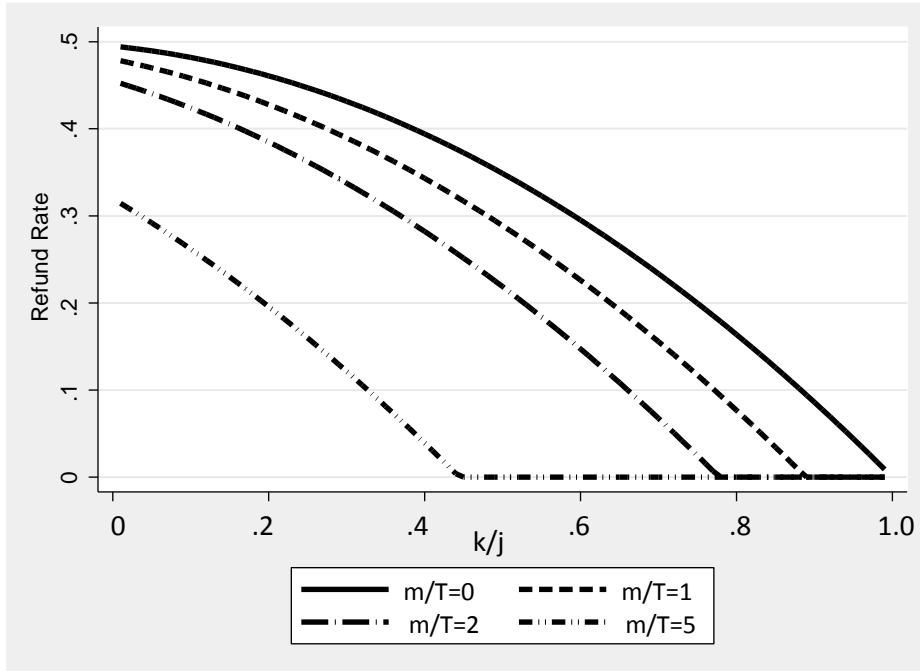


Figure 7: Refund Rates When  $m \geq 0$  ( $\alpha = .8$ )

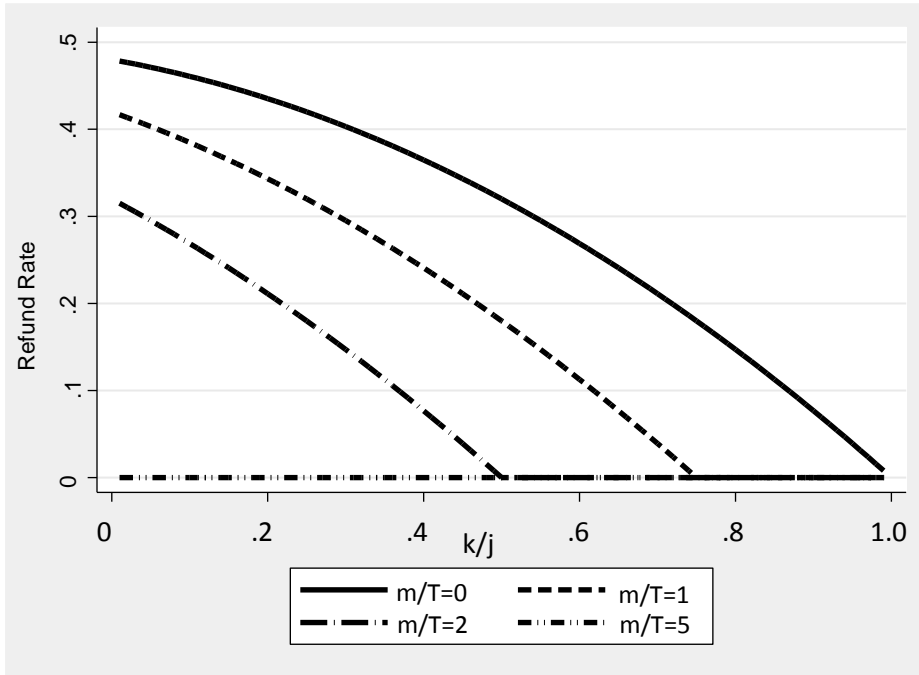


Figure 8: Uniform Distribution Versus Normal Distribution ( $\alpha = .9$ )

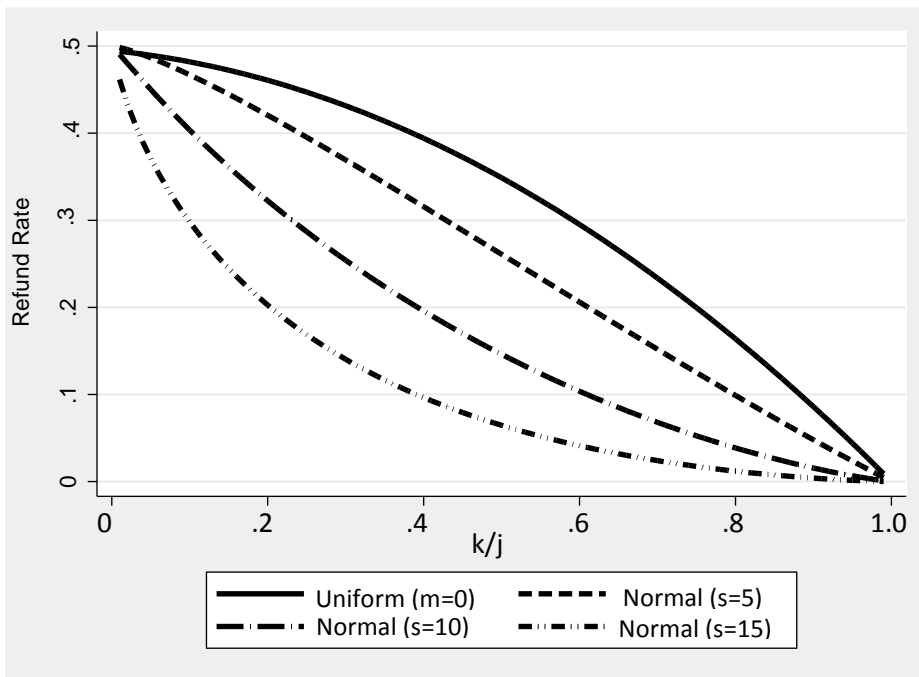
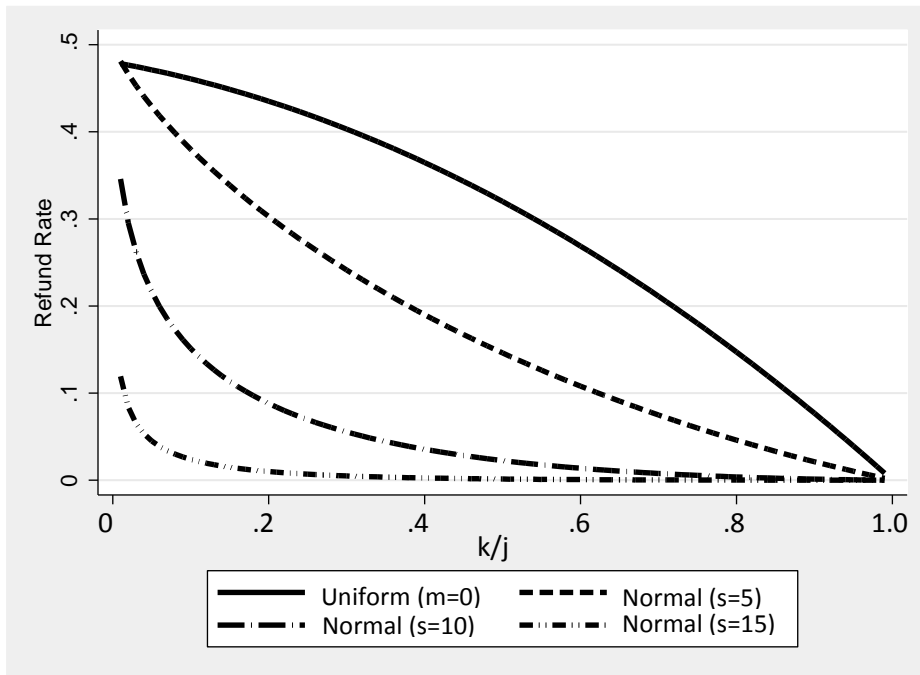


Figure 9: Uniform Distribution Versus Normal Distribution ( $\alpha = .8$ )





## Endnotes

<sup>1</sup> Since one can think of estimated tax payments as a manual form of withholding, we take “withholding” to include both automatic withholding and estimated payments.

<sup>2</sup> Fennell (2006) terms this decision a “hyperopic choice.”

<sup>3</sup> The exact tax law on when penalties will be assessed is actually rather complicated. However, these rules are an accurate distillation of the most important laws that were applied during the time period analyzed in this paper.

<sup>4</sup> Shapiro and Slemrod (1995) implement surveys to determine the effect on consumption of President George H. W. Bush’s 1992 mandated decrease in default withholding. Their results suggest that “myopic or rule-of-thumb decision-making” (as opposed to liquidity constraints) may be behind some household’s decisions to spend their newfound income, a result which they conclude is compelling in the context of the puzzle of overwithholding.

<sup>5</sup> In some cases, however, this interpretation of  $k$  is insufficient. Some taxpayers are willing to pay high rates of interest to smooth consumption, so this interpretation of  $k$  may understate the true costs to overwithholding.

<sup>6</sup> Note that the taxpayer is not directly concerned with withholding to meet his tax liability. By assumption, he will meet that liability no matter what his level of withholding.

<sup>7</sup> This implicitly assumes  $V > U$ .

<sup>8</sup> At this point in their exposition HTW set  $m=0$ . Though it does not affect the probability of overwithholding, we continue to allow  $m \geq 0$  for the sake of generalization.

<sup>9</sup> There was a small typo in the original article which read:  $W=N$  if  $\Pr(\text{Area 2}) > \frac{k}{j+k}$ .

---

<sup>10</sup> Another empirical difficulty with the model is that it seems to significantly overpredict the fraction of taxpayers that underwithhold and are penalized. Applying the annual values of  $k$  and  $j$  from Table 1 to equation (3), one can calculate that the average underwithholding rate during those years should have been 29%. Aggregate underwithholding data from the 1980s were unavailable, but in 1990 and 1991 only 3% of taxpayers underwithheld and were penalized.

<sup>11</sup> Alternatively, we could enforce these boundary conditions by assuming that if the taxpayer's desired  $W < m$ , then  $W = m$ , and when the desired  $W > \alpha t(V-U)$ , then  $W = \alpha t(V-U)$ . This method, however, seems unrealistic when  $m > 0$ , because in such a case there is no non-arbitrary reasoning why one cannot set  $W < m$  and  $W > \alpha t(V-U)$ . Further, the model would violate the strictness of the inequalities of the bounds.

<sup>12</sup> Relaxing the assumption that  $W < \alpha t(V-U)$  yielded very similar refund results.

<sup>13</sup> This again assumes that the taxpayer sets  $W \leq N$ .

<sup>14</sup> While it may seem problematic that normal distributions allow for cases in which  $L < 0$  and  $N < 0$ , our later assumptions will insure that such cases occur with at most a few percent probability. Furthermore, truncating these distributions only strengthens the result that applying a normal distribution decreases the theoretical refund rate.

<sup>15</sup> We thank Christopher Overstreet for his advice on computational methods here.

<sup>16</sup> Consistent with our assumption of rationality, we assume that the taxpayer's perceived distribution is the same as the true probabilistic distribution of liability.