

The Persistence of Tax Refunds: Evidence from Panel Data

Sara LaLumia*

October 2008

DRAFT: Preliminary and incomplete, comments welcome.

Abstract

This paper uses a 12-year panel of income tax return data to investigate patterns of refund receipt over time. I find that approximately 30% of non-elderly filers receive refunds in all twelve years of the panel. This share is higher for unmarried women than for unmarried men, and lowest for joint filers. Higher levels of income and wealth are associated with shorter spells of refund receipt. There is some evidence that taxpayers learn about their likely tax obligation over time and adjust their behavior accordingly. Owing a large balance in one year is associated with substantially higher withholding in the following year, and receiving a large refund is associated with a subsequently lower level of withholding. Even controlling for initial levels of wealth, persistent refund recipients experience lower growth in savings across the span of the panel.

Introduction

Each year approximately three quarters of tax returns generate refunds. For these returns, the money withheld from paychecks during the year (or remitted in the form of estimated tax payments) exceeds the tax liability computed at the time the tax return is filed. Average refund amounts are not trivial. In recent years, the average refund has been well over \$2000. Both the widespread receipt and the large dollar values of refunds are puzzling in a standard intertemporal model, in which an individual prefers receiving a dollar in the current period to receiving a dollar in the future. A taxpayer who receives a refund is acting in exactly the

*Department of Economics, Williams College.

opposite way, choosing to postpone the receipt of some income to a future period without receiving any compensating interest payment.

This paper investigates patterns of refund receipt using a panel of tax return data. Although this paper does not answer the question of *why* taxpayers receive refunds, it does provide a more complete picture of refund receipt. First, it describes the persistence of refunds across years. I find that approximately 30% of non-elderly filers receive refunds in all twelve years of the panel. This share is higher for unmarried women than for unmarried men, and lowest for joint filers. I estimate duration models to investigate how income levels and demographic factors affect the number of years of consecutive refund receipt. I find that higher levels of income and wealth are associated with shorter spells of refund receipt. Second, this paper investigates whether taxpayers learn about their likely tax obligation over time and adjust their behavior accordingly. I find that owing a large balance in one year is associated with substantially higher withholding in the following year, and that receiving a large refund is associated with a subsequently lower level of withholding. Third, the paper investigates the relationship between persistent refund receipt and savings. I document that, even controlling for initial levels of wealth, persistent refund recipients experience lower growth in savings across the span of the panel.

1 Background on Tax Refunds

1.1 Literature Review

Three hypotheses have been advanced to explain why such a large number of taxpayers receive refunds. The forced savings hypothesis, proposed by Thaler and Shefrin (1981), posits that individuals would like to save but have difficulty doing so because of self-control problems. Thus, individuals value commitment devices that essentially limit their short run consumption. The model of Thaler and Shefrin predicts that an individual who receives part of his income in the form of a substantial annual bonus will save more than an identical individual receiving the same total income, divided into equal monthly installments. They write, “This [bonus]

technology seems to have been recognized by the millions of taxpayers each year who claim too few exemptions in order to assure a tax refund.”

A clean test of the hypothesis that refunds promote saving by shifting money from high MPC (marginal propensity to consume) monthly income to low MPC annual income is provided by Feldman (forthcoming). She considers a 1992 change in the withholding tables that reduced the average refund amount without changing actual tax liability. Every \$100 decrease in withholding for the average household was correlated with a 0.012 percentage point decrease, or a 19.6% decrease, in the probability of contributing to an IRA. Supporting evidence shows that decreases in IRA contributions were not offset by increases in other shorter-term savings vehicles. Feldman’s findings contradict evidence presented in Feenberg and Skinner (1989) showing that filers who are in a balance due position are more likely to contribute to an IRA. They find that being in a balance due position is associated with an extra \$184 of IRA contributions for those with low initial levels of assets (asset income of \$250 or less) and with an extra \$518 of IRA contributions for those with higher initial levels of assets.

Neumark (1995) assumes that refund receipt is an indication of a consumer’s taste for forced saving. He tests whether refund recipients are more likely to use another type of forced saving, accepting lower initial wages in exchange for higher wage growth over time. An individual who would like to experience increases in consumption over time, but who cannot commit to saving out of current income, will prefer a pattern of increasing earnings over a flat earnings profile with the same present value. Neumark uses a single year of data, including information on tax refunds, earnings, and age. He finds that the estimated slope of the age-earnings profile is greater for those who received refunds than for those with positive tax balances due.

A second explanation for refund receipt is suggested by Highfill, Thorson, and Weber (1998). They suggest that overwithholding is a rational response to uncertainty. In their model income and deductible expenses have certain and uncertain components. A taxpayer must choose the amount to be withheld before his uncertain income and expenses are realized. Withholding too much generates the familiar opportunity cost of forgone interest. Withholding too little also

has a cost. If withholding falls below a threshold—the smaller of 90% of actual tax liability or 100% of the previous year’s tax liability—a penalty is incurred. The penalty is equal to some fraction (the penalty rate, set by the IRS) of the dollar amount underwithheld. A taxpayer chooses his withholding to equate the marginal forgone interest from overwithholding and the marginal penalty from underwithholding. The probability of receiving a refund is predicted to increase with the penalty rate and decrease with the interest rate. Highfill et al. use their model to predict the share of taxpayers receiving refunds in the years 1983 to 1992. On average, their prediction is a fairly close match to the actual share of taxpayers receiving refunds. They predict that 66.7% of taxpayers receive refunds, close to the actual level of 74.4%. Looking across years, though, their predicted share tends to be high when the actual share is low. The correlation between the predicted and actual shares is -0.59.

A third hypothesis for widespread refund receipt is that individuals tend to adhere to a default option. A taxpayer starting a new job may initially choose a level of withholding he expects to be close to his actual tax liability. When he eventually realizes that he is overwithholding, he may simply adhere to the status quo rather than change his withholding. Changing withholding involves paying an upfront, albeit small, cost (going to the payroll office at work and submitting a new form) while providing a deferred benefit (the additional interest earned on the money held in the taxpayer’s account instead of the federal government’s account during the year). Evidence of employees’ strong adherence to the default choice in pension plans suggests that even a small convenience cost may be enough to discourage many individuals from altering their withholding.

Recently, there has been particular interest in the use of refunds by low-income individuals. In part, this interest is motivated by the expansion of refundable, means-tested tax credits such as the Earned Income Tax Credit and the Child Tax Credit.¹ Smeeding et al. (2000) report results from a survey of low-income taxpayers in Chicago. A majority of respondents

¹While a substantial number of low-income filers receive refunds, they account for a small portion of all refund dollars. Orszag (2005) reports that the EITC accounted for less than 15% of total 2003 income tax refund dollars. This share is even lower in the years represented by my data.

expecting to receive a refund planned to divide the refund between “consumption” uses (such as paying rent or food bills) and “investment” uses (including saving, paying tuition, or paying for a car). About 34% planned to save at least part of their refund. The intent to use refund payments in a number of different ways suggests that taxpayers may value the ability to split a refund rather than receiving it in one lump sum. Beginning in 2006, the IRS has allowed filers to split direct-deposited refunds into as many as three different accounts. Beverly, Schneider, and Tufano (2007) describe a pilot study in Oklahoma that allowed filers to pre-commit to placing a portion of their tax refunds into a savings account. They find substantial demand for refund-splitting.

1.2 Withholding Rules

Much of the U.S. individual income tax is actually collected by employers rather than by the IRS. When an employee starts a new job he must file a W-4, Employee’s Withholding Allowance Certificate, with his employer. This form reports the employee’s marital status, number of withholding allowances, and any additional dollar amount to be withheld from each paycheck. The number of withholding allowances is not necessarily equal to the number of exemptions claimed on one’s federal tax return. Employees are instructed to use information on their likely itemized deduction, income from a second job, and income from a spouse’s job in order to calculate the appropriate number of allowances. It is the employer’s responsibility to use the information on the W-4, along with withholding rates published in IRS Circular E, to withhold income and FICA taxes from the employee’s paycheck. Employers deposit these funds with the federal government on a monthly or semi-weekly basis. For the simplest income tax situations, adhering to the basic instructions on the W-4 will result in a refund at the end of the tax year. The IRS may set the rules in this way to encourage filing. Taxpayers may be more likely to fulfill their filing obligation if doing so results in a refund than if it results in remitting a check to the IRS.

Technically, an individual is required to submit a new W-4 within 10 days of an event

that lowers the number of withholding allowances to which he is entitled. An individual is permitted to file a new W-4 at any time he chooses. There is very little evidence regarding the frequency with which employees actually file updated W-4 forms. These forms are held by employers. Although the IRS can request that an employer provide it with these forms, there is no systematic W-4 data collection. Shapiro and Slemrod (1995) describe a survey conducted shortly after a 1992 executive order reduced the taxes withheld from paychecks, with no corresponding change in tax liability. Fewer than 10% of the households with earned income reported that they had filed a new W-4 to offset this change. Barr and Dokko (2006) report that, among a group of households in low- and moderate-income census tracts in the Detroit metropolitan area, 47% of survey respondents would not choose to change their withholding.

There have been two instances, both following significant tax reforms, in which most employees were required to file new W-4 forms. The Revenue Act of 1964 lowered the top marginal tax rate from 91% to 70%, and employees had to submit new W-4 forms during 1966. Shortly after the passage of the Tax Reform Act of 1986, the IRS released an updated 4-page W-4, originally requiring that it be completed by employees before October of 1987. Concern about the difficulty of the form, and about low rates of compliance with it, led to a revised 2-page version being released in March of 1987. As an incentive, taxpayers who filed a new W-4 “in good faith” by June 1 of 1987 were exempt from paying the usual penalty assessed when withholding falls below 90% of current-year liability or 100% of previous-year liability.

Withholding at source has been a part of the U.S. individual income tax since the Current Tax Payment Act of 1943.² Even before this act was passed, it was predicted to lead to a sharp increase in the number of filers claiming refunds, and in corresponding administrative costs. At a time when many were bemoaning the administrative burden facing both taxpayers and the IRS, Lent (1942) points out a possible benefit. In an interesting precursor to the forced savings hypothesis, Lent notes, “The refunds due may in most instances be looked upon as a windfall

²There are earlier examples of temporary use of income tax withholding at source in the U.S. During the Civil War, income tax was withheld from the paychecks of government employees. When the modern income tax was established in 1913, there was a provision authorizing withholding at source. This provision was widely criticized and eliminated in 1917.

or savings account and could be exchanged for war stamps or bonds.”

2 Data

This paper uses a 12-year panel of tax return data compiled by the Office of Tax Policy Research (OTPR) at the University of Michigan, spanning tax years 1979 to 1990. The tax filing units included in the panel are a subset of the units included in the annual cross sectional data released by the Statistics of Income (SOI) division of the IRS. Like the cross-sectional files, each year of panel data is representative of all returns filed during a particular year, including late returns being filed for earlier tax years. I focus only on timely returns.³ While the cross-sectional SOI files are stratified by income, inclusion in the panel is random. Selection is based on the last four digits of the primary filer’s Social Security number. The number of four-digit combinations included in the panel varies from year to year, leading to wide variation in the number of tax returns included in different years. Approximately 45,000 returns are included in the years with the most widespread coverage, while slightly over 9000 returns are included in tax years 1982, 1984, and 1986. The number of tax filing units present in all 12 years of the panel is 4982.⁴

The dataset includes much of the information reported on a tax return and accompanying schedules. Hence, there is great deal of detail on income components and amounts, and reliable information on refund receipt. On the other hand, there is very little demographic information. Filing status can be used to proxy for marital status, and number of exemptions for household

³Slemrod et al. (1997) show that returns generating a refund tend to be filed earlier than returns with a balance due.

⁴Exclusion from the balanced panel because the final digits of one’s Social Security are not tracked in all years is randomly determined. Exclusion from the balanced panel based on not filing a return in every year is not randomly determined. Relative to those who exit, those who appear in all 12 years of the panel are more likely to be married, have higher incomes, and are more likely to claim an age exemption. Some events that are commonly associated with dropping out of other longitudinal datasets are not an issue here. There is no reason to expect that movers would exit from the panel, as cross-sectional files are linked by Social Security number rather than by address. Changes in marital status do not automatically lead to exclusion from the panel. Suppose a couple is married in 1979 and becomes divorced in 1980. Information from the 1979 joint return will be linked to subsequent single returns filed by the person who was listed as the primary taxpayer on the 1979 return.

size. The sex of the primary taxpayer is also known, as this information was included in the 1979 and 1980 SOI cross-sections. State of residence is reported if adjusted gross income (AGI) is less than \$200,000.

I use information on interest and dividend income to impute measures of taxable wealth for each tax filing unit, following the approach of Feenberg and Skinner (1989). For each filing unit, interest income is averaged across the first three years of the panel and this average is divided by the average annual interest rate of that period. Similarly, dividend income is averaged over the first three years of the panel and divided by the average dividend yield during this period. Information on interest rates and dividend yields comes from the *Economic Report of the President*. The imputed interest-bearing principal and dividend-yielding principal is summed to arrive at a measure of taxable wealth. For each household I compute four wealth measures, each using averaged data from a three-year period. Because dividend income in particular can be volatile across years, this averaging produces a more stable wealth measure.

The uncertainty hypothesis described above suggests that taxpayers who are less able to accurately predict their annual tax liability may have higher rates of withholding. I proxy for a filing unit's uncertainty about tax liability by calculating a measure of wage dispersion over the 12-year panel. For each filing unit, I compute the standard deviation of wage income as a share of the unit's mean wage income.

Throughout the paper I focus on non-elderly tax filing units. The tax return includes checkboxes indicating whether the primary and secondary filer are over 65 years old. Those over 65 can claim a higher standard deduction.⁵ I classify as elderly any return claiming at least one age (or blindness) exemption in 1990, the last year of the panel. This means that the oldest individuals in the non-elderly group were 53 at the start of the panel. Not surprisingly, the probability of reporting wage income and the amount of wage income as a share of AGI are substantially lower for elderly filers. Approximately 95% of returns in the non-elderly sample

⁵A taxpayer who is legally blind qualifies for the same increase in the standard deduction as one who is over 65. In tax years 1988 through 1990 there is a single variable indicating the number of age and blindness exemptions claimed. In earlier years age and blindness exemptions are reported separately.

report wage income in the first year of the panel, and the share with wage income never falls below 92%. In contrast, the share of elderly returns with non-zero wage income falls steadily from 71% in the first year of the panel to 34% in the last year. Because employer withholding out of wage income is the most common way to generate a tax refund, I expect that refund receipt will be substantially lower among the elderly.

Basic descriptive statistics for the sample are reported in Table 1. Every filing unit that appears in all twelve years of the panel is represented in the first column. The second column is restricted to the non-elderly, and corresponds to the sample that I use in most of the results that follow. The final two columns divide the sample into “frequent” refund recipients with 10 or more years of refund receipt, and filers with 9 or fewer years of refund receipt. Frequent refund recipients are less likely to be married, have lower incomes, and are less likely to have ever reported capital gains or income from self-employment (Schedule C income).⁶ Only 7.3% of frequent refund recipients made estimated tax payments in any year of the panel, while 40.6% of filers with nine or fewer refunds did so. Frequent refund recipients are somewhat less likely to have ever itemized deductions or used a paid preparer, perhaps reflecting less complex returns.

Information on early withdrawal penalties and on credit card interest deductions may reveal information about taxpayers’ self-control in savings decisions. A taxpayer can subtract from his income any amount paid as a penalty on early withdrawals from a time savings deposit, such as a CD. Paying a penalty on an early withdrawal may indicate impatience or lack of self-control in savings decisions. If so, the forced savings hypothesis would suggest that penalties would be reported more often by people with a preference for refunds. This is not the case. Approximately 7% of frequent refund receivers report an early withdrawal penalty in at least one year of the panel, while 13% of those with 9 or fewer refunds do so. On the other hand, it is only possible to incur an early withdrawal penalty if one has previously established a time savings deposit. If frequent refund receivers find it difficult to save voluntarily, they may rarely

⁶Both capital gains and self-employment income can take on negative values. Here, I make no distinction between gains and losses. The capital gains and self-employment income dummies are equal to one if the corresponding income amount is non-zero.

open these accounts in the first place. In 1979 and 1980, itemizers could deduct any credit card interest paid. It is possible that paying credit card interest indicates impatience and, perhaps, a taste for forced saving. If so, this behavior should be more common among frequent refund receivers. However, those with 10 or more refunds are less likely to deduct credit card interest. This is in part due to the fact that this group is less likely to itemize deductions. Conditional on itemizing, 79% of frequent refund receivers deducted credit card interest while 74% of others did so.

3 Duration of Refund Receipt

Figures 1 and 2 show the number of years of refund receipt for younger and older filers. On average, younger filers received refunds in 9.3 out of 12 years, while older filers received refunds in 6.4 years. Approximately 29% of younger filers (1185 out of 4031) received refunds in all 12 years of the panel. This is substantially higher than the 3% of filers who could be expected to receive 12 consecutive years of refunds if 75% of returns generated a refund in each year and refund receipt in one year was completely independent of refund receipt in other years. Nearly 58% of younger filing units receive a refund in 10 or more years. Figures 1 and 2 show that the group of persistent refund receivers lacks a counterpart at the other end of the distribution. There is not a substantial group of people who persistently withhold less than their tax liability. Only 1% of younger filers and 3% of older filers never receive a refund. This is particularly striking because it indicates that extremely few taxpayers behave in the way predicted by the standard intertemporal model. In each year of the panel, persistent refund recipients (those who received refunds in all 12 years) account for 35% to 42% of all refund recipients and for 31% to 40% of refund dollars received. Pooling together all cases of refund receipt, the mean real dollar amount of a refund going to a 12-refund recipient is \$1343 while the mean amount going to a filer with fewer than 12 refunds is \$1533.

Dividing the sample of younger filers into men who were unmarried in 1979, women who

were unmarried in 1979, and those who filed jointly in 1979 shows that the modal number of refunds is 12 for all three groups. However, the share with 12 refunds is 23% for the married group, 33% for the single men, and 50% for the single women. This may reflect a higher degree of financial risk aversion among women. Jianakoplos and Bernasek (1998) use data from the 1989 Survey of Consumer Finances to show that the positive relationship between the level of wealth and the share of wealth held in risky assets is weaker for single women than for single men or for married couples.

The histograms of refund receipt show that about 30% of non-elderly filers have the same refund status in all 12 years of the panel, with 1185 receiving refunds in all years and 42 never receiving refunds. How frequently do the remaining filers change refund status? Approximately 11% (444 filers) change refund status once, 22% (900 filers) change twice, and 37% (the remaining 1460 filers) change three or more times. The share of non-elderly filing units changing refund status is below 20% in all years before 1987, with an average of 17.3%. The requirement that employees file new W-4 forms after the Tax Reform Act of 1986 essentially lowered the cost of altering one's withholding. Not surprisingly, the share of filing units changing refund status subsequently increases. This share is 22.4% for tax year 1987 and remains above 22% in each of the three remaining years of the panel.

For the 3295 non-elderly filing units receiving a refund in the first year of the panel, I measure the length of the first "spell" of refund receipt. Figure 3 shows, for non-elderly filers receiving a refund in 1979, the distribution of observed spell length. The average length of this first spell is 7.6 years. This certainly understates the average length of all refund spells, for two reasons. First, my data are subject to right censoring. The panel continues only until 1990, so I cannot observe the true length of spells that extend beyond 1990. Second, my data are also subject to left censoring. I do not know if 1979 is the first year of a refund spell, or for how many consecutive years before 1979 a filing unit was receiving a refund. Thus, the large number of filing units with an observed refund spell length of 12 years certainly includes many cases in which the actual number of years with consecutive refund receipt exceeds 12. The effects of

the Tax Reform Act of 1986 and the resulting increase in W-4 filing is evident in Figure 3 as well. There is a slight spike in the number of filing units with an observed spell length of eight refunds. These filing units received a refund in every year from 1979 until 1986.

I use duration analysis to model the probability of exiting from (initial) refund receipt as a function of income and demographic variables. I estimate proportional hazard models of the form

$$\lambda(t; X) = \kappa(X)\lambda_0$$

where λ is the hazard rate, the probability of not receiving a refund at time t conditional on having received a refund in each year from 1979 to $t - 1$. The baseline hazard, λ_0 , is assumed to be the same for all filing units. Individual hazards differ proportionately, based on components of the vector X such as average AGI, initial imputed wealth, and demographics. I use the Cox proportional hazard model which does not require estimation of the baseline hazard.

Hazard ratios indicate how the relative probability of exiting from the refund receipt state changes when one covariate changes. A hazard ratio greater than one indicates that an increase in the regressor is associated with an increased probability of no longer receiving a refund at time t , conditional on having previously received refunds. A hazard ratio less than one indicates a reduced probability of exit from refund receipt status. In other words, ratios greater than one are associated with shorter spells of refund receipt and ratios less than one are associated with longer spells of refund receipt.

Table 2 presents hazard ratios from a number of specifications. The first column of the table uses only regressors that are constant over the course of the panel. The hazard ratios show that higher levels of income and wealth are associated with shorter spells of refund receipt. An increase of \$1000 in a filer's 12-year average AGI is associated with a 0.2 percentage point increase in the probability of exiting from refund receipt. An increase of \$1000 in imputed 1979-1981 wealth is associated with a 0.1 percentage point increase in the probability of exiting from refund receipt. Even holding constant a filing unit's average income, having ever been below 200% of the poverty line is associated with a substantially lower probability of exiting from

refund receipt. This pattern of results could reflect a greater degree of financial sophistication among higher-income individuals. Perhaps higher-income, higher-wealth taxpayers are the ones who recognize that receiving a refund means forgoing some interest income. Another possibility is that higher-income individuals have better access to banks and other financial institutions, and can expect to earn higher average interest rates on their savings. In this case refund receipt will actually be more costly for these individuals, involving greater amounts of forgone interest.

Filers with higher levels of wage dispersion are more likely to exit from refund receipt. This is not consistent with the hypothesis that uncertainty about tax liability raises withholding and hence increases refunds. My result may be due to sample selection bias. By considering refund spells of only those who received a refund in 1979, taxpayers with high levels of wage dispersion and below-average 1979 wage realizations are more likely to appear in my sample than are taxpayers with high levels of wage dispersion and above-average 1979 wage realizations.

Having ever used a paid preparer is associated with a marginally significant increase of 8.5 percentage points in the probability of exiting from refund receipt. Perhaps tax preparers urge their clients to adjust their withholding to move closer to a zero balance position. Having ever paid a penalty on an early withdrawal is associated with a substantial increase in the probability of exiting from refund receipt. If making early withdrawals is an indication of impatience, the forced saving hypothesis would predict the opposite—impatient types should be more likely to persist in receiving refunds. As mentioned previously, it is only possible to make an early withdrawal after a time saving deposit has been established, so an early withdrawal may not be a reliable measure of impatience. An alternative story is that making an early withdrawal is a sign of financial distress. In times of a negative financial shock, taxpayers may turn to relatively illiquid sources of funds, both withdrawing money from time deposits that have yet to mature and choosing not to defer income receipt via a refund.

Turning to demographic variables, filing status in 1979 has a strong effect on the duration of refund receipt but household size (as measured by the number of exemptions claimed) does not. Compared to men who are unmarried in 1979, women who are unmarried in 1979 are 27.5

percentage points less likely to exit from refund receipt, while married filers are 34.4 percentage points more likely to exit.

The next three columns of Table 2 add time-varying regressors. In column 2, I add year-specific dummy variables for itemizing deductions, reporting any Schedule C income, reporting any capital gains or losses, and reporting unemployment insurance income. The hazard ratios for these variables are generally consistent with intuition. Conditional on income, itemizers have larger deductions from taxable income than do those who take the standard deduction, and hence are more likely to receive a refund. Column 2 shows that itemizing is associated with a 10.2 percentage point decline in the probability of exiting from refund receipt. Reporting income that is not generally subject to third-party withholding, either self-employment income or capital gains, substantially increases the probability of exiting from refund receipt. I have included the indicator for unemployment income to probe the possibility that individuals stop receiving refunds in times of financial hardship. However, the effect of having unemployment income is insignificant, and its inclusion has little effect on the hazard rate associated with paying an early withdrawal penalty. In general, the addition of time-varying characteristics in column 2 does little to change the estimated effects of the time-invariant characteristics included in both columns 1 and 2. The notable exception is that having ever used a paid preparer no longer has a significant effect on the probability of exiting from refund receipt.

In column 3 and column 4 of Table 2, I add measures of changes in demographic variables from one year to the next. A change in marital status substantially increases the probability of exiting from refund receipt. Column 3 shows that experiencing any change in the number of exemptions lengthens refund spells, and column 4 shows that, as expected, this effect is driven by those who are able to claim an additional exemption (say, due to the birth of a child) rather than those who lose an exemption (say, to a child aging out of the dependent range).

The duration models I have estimated consider only the first refund spell observed for each filing unit. However, many filing units in my sample have multiple spells of refund receipt. In future versions of this paper I plan to make use of data from subsequent refund spells.

4 Learning

The tax code is notorious for its difficulty. The process of completing and filing one year's tax return should give a filer information about his own likely future tax liability, even if a paid preparer does most of the work and simply informs the taxpayer of his current liability. In this section I look for evidence that taxpayers learn about their tax liability over time. One simple piece of evidence that would suggest taxpayers learn over time would be a steady decline in the share of returns generating a refund. Aggregate data from the IRS show no such decline. The share of returns generating a refund has been between 70% and 81% in every year from 1979 to 2005, with some of the highest shares observed in the last five years. It is possible that an individual's propensity to receive a refund declines over time, but that such a pattern is not apparent in pooled cross-sections of data. New, inexperienced taxpayers enter the tax-filing population every year (presumably bringing up the share of returns generating refunds) while taxpayers with many years of experience (and presumably the lowest propensity to generate a refund) exit. In the balanced panel I use there is no entry or exit over time, and taxpayer learning should result in a declining share of returns generating a refund. The share of returns in my sample receiving a refund fluctuates between 78% and 82% over the first few years of the panel, reaches a peak of 82.5% in 1983, and then steadily declines to 71% by 1990. The decline in refund receipt is especially pronounced among those who initially filed joint returns. For this group, the share receiving refunds falls steadily from a 1983 peak of 80.9% to 66.5% by 1990.

Panel data offer the possibility of more convincing tests for taxpayer learning. If the process of filing a return in one year imparts information to a taxpayer that he can use to finetune his withholding for the following year, prior refunds should be associated with reductions in withholding and cases of prior balance due should be associated with increases in withholding. Unfortunately I do not have any information from W-4 forms which would directly indicate changes in withholding. Instead, I construct a variable equal to total withholding of federal income tax divided by wage income. The presence of a few extreme outliers distorts the distri-

bution of this variable, and in what follows I drop observations for which the withholding share is greater than or equal to 2, as well as filing units that have fewer than 12 years with non-zero wage income.⁷ After applying these restrictions, the mean of my constructed withholding share is 0.143.

Making use of the panel structure of my data, I estimate a fixed effects specification in which the withholding share in year t depends on year t income and demographics as well as on the year $t - 1$ refund amount. My baseline specification is:

$$WHshare_t = f(WHshare_{t-1}, Exemptions_t, Married_t, AGI_t, Refund_{t-1})$$

I restrict my sample to non-elderly filing units that report positive wage income in all years of the panel. I use information on lagged refund amount to group observations into six categories: those with refunds of more than \$1000, refunds of \$500 to \$1000, refunds of less than \$500, balance due of less than \$500, balance due of \$500 to \$1000, and balance due of more than \$1000. I include dummies for these categories, with the smallest refund amount as the omitted category. Below I report results in which I include a set of categorical AGI variables, but the pattern of results is similar if I replace these variables with the amount of AGI and AGI squared.

Results are shown in Table 3. Not surprisingly, the single best predictor of year t withholding share is the withholding share of year $t - 1$. Demographic variables affect withholding share in sensible ways. Both an additional personal exemption and joint filing status reduce average tax liability, and in the regression both are associated with a lower withholding share. As expected under a progressive tax, withholding share increases with AGI. Withholding shares are significantly lower in years after 1986, consistent with the reduction in tax rates experienced by many filers. The pattern of coefficients on the set of refund dummies is consistent with taxpayer learning. Receiving a refund of over \$1000 in one year is associated with a 0.3 percentage point decline in the following year's withholding share. Owing a balance of over \$1000 is associated

⁷Dropping observations with a withholding share of 2 or more eliminates only 39 observations. The value of wages is low in all of these cases, with a maximum wage of \$3636.

with a 0.8 percentage point increase in the subsequent withholding share. In the second column of Table 3, I restrict the sample to filing units for whom wage income was 50% or more of AGI in each year of the panel. These are presumably individuals who are best able to predict their future tax liability, and who may therefore show the greatest reaction to previous refunds. As the table shows, this sample restriction makes very little difference in the results.

In Table 4, column 1, I consider whether filers who used a paid preparer show more responsiveness to prior-year refunds or balances due. I interact the set of lagged refund dummies with an indicator for having ever used a paid preparer. Only one of the five interaction terms is significant at the 10% level. For most refund or balance due amounts, the effect on subsequent withholding is the same regardless of whether a paid preparer is used. The one exception to this pattern is that only those taxpayers who never use a paid preparer respond to a medium-sized balance due, one between \$500 and \$1000, by increasing their subsequent withholding. In column 2 I consider whether learning differs with income. Here I interact the lagged refund dummies with an indicator for having had AGI less than 200% of the federal poverty line. There is evidence of a difference between low- and higher-income groups only in the case of a medium-sized balance due. Only higher-income filers respond to owing a medium-sized balance by increasing their subsequent withholding.

5 Refund Receipt and Saving

If indeed many taxpayers choose to receive refunds because they would like to save more, and perceive refunds as a way to help them commit to that goal, it is natural to ask whether refunds in fact increase saving levels. Tables 5 and 6 compare measures of saving for filers receiving refunds in all twelve years of the panel and for filers receiving fewer refunds. I sort filers into four categories based on imputed initial wealth. I do this because comparisons of changes across groups are difficult to interpret when the groups begin at different initial levels, and filers with more refunds have systematically lower initial wealth levels. The average initial

wealth for filers receiving refunds in all 12 years is \$2125 while the average initial wealth for other filers is \$21070. The four wealth categories I consider are zero taxable wealth, taxable wealth between \$1 and \$625, taxable wealth between \$626 and \$5562, and taxable wealth of \$5563 or more. These categories correspond roughly to quartiles of wealth for all non-elderly taxpayers, although slightly more than a quarter of those in my sample have imputed initial wealth of zero.

Table 5 shows the mean amount of interest income reported in 1979 and in 1990. This type of income is indicative of savings held in a traditional, taxable savings account or interest-bearing checking account. Interest income from 1979 is the upper entry in each cell. By construction, 1979 interest income is similar for persistent and non-persistent refund receivers within each initial wealth category. The table shows that, within each initial wealth category, persistent refund receivers experience smaller increases in taxable interest income between 1979 and 1990. In the lowest imputed wealth category, average interest income rises from \$0 to \$123 for persistent refund receivers but from \$0 to \$343 for other filers. In all four categories the dollar increase among non-persistent refund receivers is at least two and a half times larger than the dollar increase for persistent refund receivers. This strongly suggests that persistent refund receipt is correlated with lower levels of saving. However, this pattern of results does not say anything about the *causal* impact of refunds on saving. Instead this pattern could be due entirely to the selection of low-savings types into refund receipt.

Table 6 shows the mean amount of IRA contributions made over the course of the panel, as well as the share of filers ever contributing to an IRA. Within each initial wealth category, the combined IRA contributions made over 12 years are lower for persistent refund recipients. The difference is most pronounced in the lower two wealth quartiles. The probability of ever contributing to an IRA is also lower for persistent refund receivers than for others, and again the gap is larger in the bottom two wealth quartiles.

Conclusion

While puzzling within standard intertemporal economic models, income tax refunds enjoy great popularity among taxpayers. In this paper I use a panel of tax return data to establish that about 30% of filers receive a refund in twelve consecutive years, far more than would be the case if refund receipt in one year was independent of refund receipt in other years. I show that persistent refund receipt is more common among those with lower levels of income and wealth. I show some evidence of learning about tax liability. Taxpayers who owe large amounts of money in one year tend to have a higher rate of tax withholding in the following year, while taxpayers who receive large refunds in one year have lower subsequent tax withholding. Finally, I show that persistent refund receivers experience less growth over time in taxable and IRA savings.

References

- Barr, Michael S. and Jane K. Dokko. 2006 “Tax Filing Experiences and Withholding Preferences of Low- and Moderate-Income Households: Preliminary Evidence from a New Survey.” *IRS Research Conference Proceedings*.
- Beverly, Sondra, Daniel Schneider and Peter Tufano. 2006. “Splitting Tax Refunds and Building Savings: An Empirical Test.” *Tax Policy and the Economy*. 20: 111-162.
- Feenberg, Daniel and Jonathan Skinner. 1989. “Sources of IRA Saving.” *Tax Policy and the Economy*. 3: 25-46.
- Feldman, Naomi E. Forthcoming. “Mental Accounting Effects of Income Tax Shifting.” *Review of Economics and Statistics*.
- Highfill, Jannett, Douglas Thorson, and William V. Weber. 1998. “Tax Overwithholding as a Response to Uncertainty.” *Public Finance Review*. 26(4): 376-391.
- Jianakoplos, Nancy Ammon and Alexandra Bernasek. 1998. “Are Women More Risk Averse?” *Economic Inquiry*. 36(4): 620-630.
- Lent, George E. 1942. “Collection of the Personal Income Tax at the Source.” *Journal of Political Economy*. 50(5): 719-737.
- Neumark, David. 1995. “Are Rising Earnings Profiles a Forced-Saving Mechanism?” *Economic Journal*. 105(428): 95-106.
- Orszag, Peter R. 2005. “Individual Income Tax Refunds.” *Tax Notes* January 31: 599.
- 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.
- Slemrod, Joel, Charles Christian, Rebecca London, and Jonathan Parker. 1997. “April 15 Syndrome.” *Economic Inquiry* 35(4): 695-709.
- Smeeding, Timothy M., Katherin Ross, and Michael O’Connor. 2000. “The EITC: Expectation, Knowledge, Use, and Economic and Social Mobility.” *National Tax Journal*. 53(4): 1187-1209.
- Thaler, Richard H. and H. M. Shefrin. 1981. “An Economic Theory of Self-Control.” *Journal of Political Economy*. 89(2): 392-406.

Table 1: Summary Statistics, Returns of Filing Units in All 12 Years of Panel

	Full Sample	No Age Ex in 1990	Non-Elderly Only	
			9 or Fewer Refunds	10 or More Refunds
Number of Refunds	8.8	9.3	6.6	11.3
<i>Demographics</i>				
Filed jointly in 1979	60.4	58.1	67.3	51.3
Average exemptions	2.7	2.8	2.8	2.7
Ever claimed age exemption	20.2	1.4	2.4	0.7
Ever changed joint filing status	26.1	30.2	26.7	32.7
<i>Income (amounts in 1990 dollars)</i>				
Average AGI	39337	39986	49400	33113
Average wage income	32095	36018	40124	33020
Average interest	2471	1209	2311	405
Average capital gains	1086	863	1838	151
Ever had capital gain or loss	37.7	33.7	49.5	22.1
Ever had Schedule C income	33.6	35.6	48.3	26.3
Ever made estimated tax payments	29.7	21.3	40.6	7.3
Ever had AGI < 2*Poverty Line	58.4	56.1	49.3	61.1
Ever had unemployment income	29.3	33.1	27.7	37.1
<i>Financial Sophistication</i>				
Ever itemized	74.6	76.6	83.2	71.7
Ever used paid preparer	70.8	70.0	74.3	66.8
Ever paid early withdrawal penalty	11.7	9.5	13.2	6.8
Ever deducted credit card interest	31.9	34.4	40.2	30.2
Ever deducted CC interest, itemizers	70.8	76.5	73.8	79.4
N	4982	4031	1701	2330

Table 2: Duration of Refund Receipt

	(1)	(2)	(3)	(4)
Average AGI, 1000s	1.002	1.002	1.002	1.002
	(4.85)	(4.80)	(4.80)	(4.81)
1979-81 Wealth, 1000s	1.001	1.001	1.001	1.001
	(2.32)	(1.85)	(1.91)	(1.89)
Wage Dispersion	1.010	1.009	1.009	1.009
	(18.76)	(16.82)	(16.74)	(16.70)
Ever Used Paid Preparer	1.085	1.068	1.066	1.065
	(1.68)	(1.34)	(1.31)	(1.30)
Ever Had AGI < 2·Poverty	0.757	0.756	0.759	0.760
	(-5.65)	(-5.31)	(-5.25)	(-5.20)
Ever Paid Early Penalty	1.280	1.245	1.248	1.249
	(3.37)	(2.97)	(3.00)	(3.01)
Exemptions, 1979	1.011	1.007	1.012	1.009
	(0.56)	(0.35)	(0.63)	(0.44)
Single Woman, 1979	0.725	0.731	0.747	0.744
	(-4.02)	(-3.92)	(-3.61)	(-3.66)
Joint Filer, 1979	1.344	1.301	1.313	1.315
	(4.12)	(3.64)	(3.75)	(3.76)
Itemized		0.898	0.900	0.903
		(-2.05)	(-2.00)	(-1.96)
Any Schedule C Income		1.607	1.613	1.615
		(7.32)	(7.38)	(7.39)
Any Capital Gains		1.472	1.465	1.467
		(5.29)	(5.23)	(5.24)
Any Unemployment Income		0.963	0.958	0.960
		(-0.41)	(-0.47)	(-0.45)
Change in Marital Status			1.592	1.617
			(3.80)	(3.91)
Change in Exemptions			0.866	
			(-1.86)	
Gain Exemption				0.805
				(-2.16)
Lose Exemption				0.930
				(-0.76)

Note: The table shows hazard ratios, with z-statistics in parentheses.

Table 3: Fixed Effects Regressions Predicting Withholding Share

	Wage in All Years	Wage> 0.5AGI
Lagged WH Share	0.518 (0.005)	0.510 (0.005)
Exemptions	-0.003 (0.0003)	-0.002 (0.0002)
Joint Filer	-0.016 (0.001)	-0.018 (0.0007)
AGI < 0	0.006 (0.005)	-0.003 (0.004)
20K < AGI < 35K	0.017 (0.001)	0.018 (0.0005)
35K < AGI < 50K	0.025 (0.001)	0.027 (0.0006)
AGI >50K	0.030 (0.001)	0.034 (0.0008)
Refund, Over 1000	-0.003 (0.001)	-0.002 (0.0004)
Refund, 500 to 1000	-0.001 (0.0006)	-0.0008 (0.0004)
Balance Due, 0 to 500	0.0008 (0.0008)	0.001 (0.0005)
Balance Due, 500 to 1000	0.002 (0.001)	0.003 (0.0008)
Balance Due, Over 1000	0.008 (0.001)	0.006 (0.0007)
Post-1986	-0.008 (0.0004)	-0.011 (0.0003)
Constant	0.072 (0.001)	0.072 (0.0009)
N	38071	33836

Note: Standard errors are in parentheses.

Table 4: Does Learning Differ with Taxpayer Characteristics?

	Paid Preparer	Ever Had AGI < 2·Pov
Refund, Over 1000	-0.003 (0.001)	-0.003 (0.001)
Refund, 500 to 1000	-0.001 (0.001)	-0.001 (0.001)
Balance Due, 0 to 500	0.000003 (0.001)	0.0002 (0.001)
Balance Due, 500 to 1000	0.005 (0.002)	0.0005 (0.002)
Balance Due, Over 1000	0.006 (0.002)	0.009 (0.001)
(Refund, Over 1000)·Prep	0.0000007 (0.001)	
(Refund, 500 to 1000)·Prep	0.0005 (0.001)	
(Balance Due, 0 to 500)·Prep	0.001 (0.002)	
(Balance Due, 500 to 1000)·Prep	-0.004 (0.002)	
(Balance Due, Over 1000)·Prep	0.002 (0.002)	
(Refund, Over 1000)·LMI		0.001 (0.001)
(Refund, 500 to 1000)·LMI		0.0008 (0.001)
(Balance Due, 0 to 500)·LMI		0.003 (0.002)
(Balance Due, 500 to 1000)·LMI		0.005 (0.002)
(Balance Due, Over 1000)·LMI		-0.002 (0.002)

Note: Each column reports results from a regression that also includes lagged withholding share, current year exemptions and marital status, a set of current year AGI indicators, and a post-1986 dummy. Standard errors are in parentheses.

Table 5: Mean Interest Income in 1979 and 1990, By Years of Refund Receipt

Imputed Wealth, 1979-1981	Under 12 12 Refunds	
	Under 12 Refunds	12 Refunds
Zero	0 343	0 123
1 - 625	14 614	13 251
626 - 5563	114 1440	115 515
5564 and over	1633 5585	641 1041

Note: The top entry in each cell is for 1979, and the bottom entry in each cell is for 1990. All dollar amounts are expressed in real 1990 dollars.

Table 6: IRA Contributions, By Years of Refund Receipt

Imputed Wealth, 1979-1981	Mean Contributions		Ever Contribute	
	Under 12 Refunds	12 Refunds	Under 12 Refunds	12 Refunds
Zero	858	338	23.4	11.9
1 - 625	1550	522	35.5	18.8
626 - 5563	3899	3152	55.4	48.3
5564 and over	8764	6615	72.9	69.7

Note: Cell entries in the second and third columns represent the sum of IRA contributions made in tax years 1979 through 1990.

Figure 1: Years of Refund Receipt, Younger Sample (N=4031)

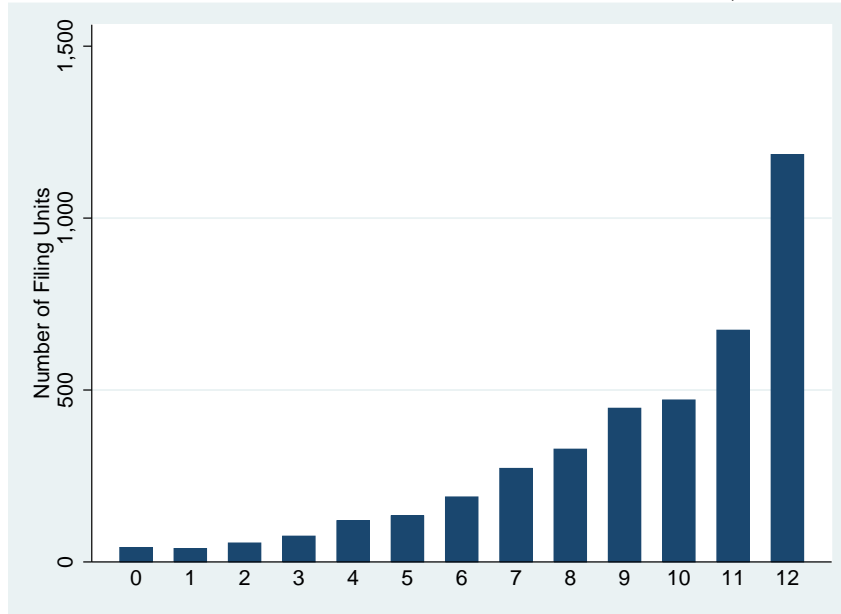


Figure 2: Years of Refund Receipt, Older Sample (N=951)

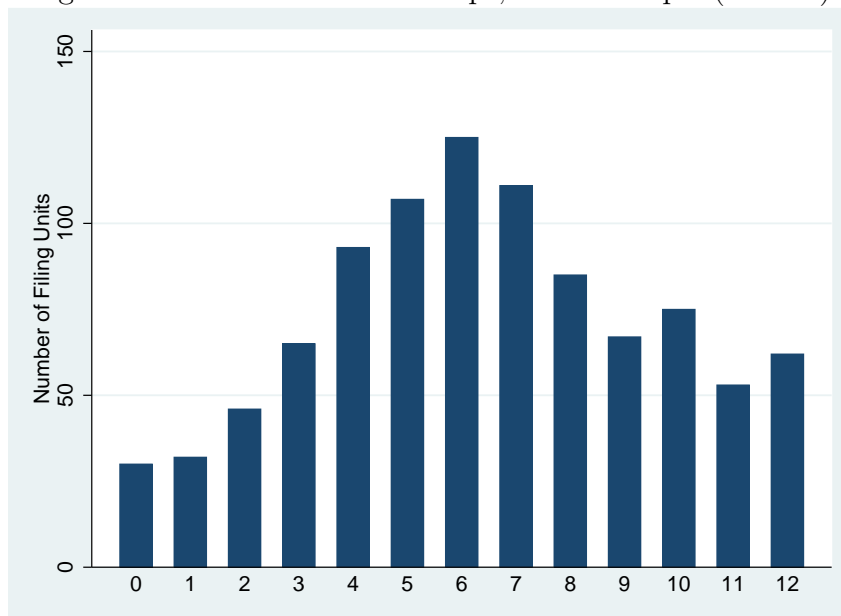


Figure 3: Duration of First Refund Spell

