How Taxing Is Tax Filing? Leaving Money on the Table Because of Hassle Costs

Youssef Benzarti*

Abstract

I use a quasi-experimental design and a novel identification strategy to estimate the burden of filing taxes. Employing US income tax returns, I observe how taxpayers choose between itemizing deductions and claiming the standard deduction. Taxpayers forgo tax savings to avoid hassle costs, resulting in an average burden of itemizing of $644 per taxpayer, and suggesting large aggregate filing costs. This burden increases with income, consistent with a higher opportunity cost of time for richer households. I provide evidence consistent with taxpayers procrastinating on filing taxes and estimate that procrastination accounts for 65% of the cost.

JEL Classification: D03, H24, H31, H83

Keywords: Behavioral, Tax Filing, Household Economics

"Death, taxes and childbirth! There's never any convenient time for any of them."

Margaret Mitchell, *Gone with the Wind*

While there is a long tradition in public finance of assessing the magnitude of the efficiency cost of taxation, very little attention has been given to the burden of filing taxes. Every year, more than 140 million taxpayers have to file taxes in the US. With the tax code becoming increasingly complex, taxpayers have to spend a significant amount of time filling out the 1040 form, various schedules and keeping records of their transactions. The prevalence of hassle costs is a possible – yet unexplored – explanation for the incomplete take up of benefits such as the Earned Income Tax Credit (EITC) or Unemployment Insurance.[1] This has been emphasized by President Obama[2] who insists on the need to: "Help qualifying individuals (...) access public programs and benefits by (...) streamlining processes that may otherwise limit or delay participation – for example, removing administrative hurdles, shortening wait times, and simplifying forms"

How large is the burden of tax filing and are taxpayers forgoing benefits because of it? I answer this question by observing the choice of individuals over two tasks offering a trade-off between hassle costs and benefits. Itemizing deductions requires some effort cost but can provide large tax savings. Claiming the standard deduction saves time and effort but results in more taxes due.

With no hassle costs, taxpayers should itemize if the benefit of itemizing is greater than zero. With hassle costs, itemizing is only beneficial if it reduces the tax bill by more than the cost of itemizing. This implies that if hassle costs are non-zero, some taxpayers will claim the standard deduction even though the sum of their deductions is greater than the standard deduction amount. The main identification challenge is to differentiate between individuals who fail to itemize deductions because of hassle costs from individuals who claim

---

[1]See for example Blank and Card (1991) for UI and Currie (2006) for other programs.
[2]*Using Behavioral Insights to Better Serve the American People*, Executive Order, September 2015

the standard deduction because their total deductions are smaller than the standard deduction amount. This is particularly difficult because taxpayers who claim the standard deduction are not required to report their deductions, implying that their true level of deductions is not observable in tax data.

If individuals are forgoing tax benefits because of hassle costs, there should be a missing mass in the density of deductions immediately to the right of the standard deduction threshold. I test this hypothesis by graphing the density of deductions for years ranging from 1980 to 2003 using a representative sample of US tax returns. The shape of the density function suggests the presence of a missing mass in the neighborhood of the standard deduction. To confirm that this shape is due to taxpayers responding to the standard deduction, I turn to a quasi-experimental design. Following an increase in the standard deduction amount in 1971 and 1988, I observe a drop in the mass of itemizers in the neighborhood of the post-reform standard deduction threshold. The post-reform density is systematically lower than the pre-reform one in the neighborhood of the post-reform standard deduction threshold and the two densities overlap further away from the standard deduction. I ensure that no other reforms are affecting the densities of itemized deductions.[3]

I use the missing mass to construct the distribution of forgone benefits. I find significant heterogeneity among taxpayers. Some taxpayers still itemize even when savings are modest and some forgo large tax benefits, resulting in an average burden of itemizing of $644 per person.

If individuals switch to the standard deduction because they value their time more than the benefits they can derive from itemizing, richer households should forgo more tax benefits than poorer ones. To test this hypothesis, I break down individuals by income deciles and repeat the estimation using the same identification strategy outlined above. The results show an increasing relationship between forgone tax benefits and income - while controlling for the marginal tax rate - consistent with the hypothesis that tax filing imposes

---

[3]My estimates are not affected by the Alternative Minimum Tax, variation in marginal tax rates and the phase out of the personal interest deduction in 1987. Details are provided in section 5.5.

a higher burden on richer individuals because they have a higher marginal value of time. Using a revealed preference argument and back of the envelope calculations, I estimate that itemizing is perceived to be as painful as working 19 hours.

The existence of a missing mass in the neighborhood of the standard deduction is consistent with taxpayers forgoing benefits to avoid the cost of itemizing. However, there are three main alternative explanations to the presence of the missing mass. The first is that the standard deduction acts a concave kink point, effectively changing the price of a deduction. Behavioral responses to concave kink points predict that taxpayers will respond to variations in marginal tax rates but should not respond to variations in income while holding the marginal tax rate fixed. The fact that forgone benefits increase with income - while controlling for the marginal tax rate - supports the hassle costs explanation and rules out behavioral responses to concave kink points. A second alternative explanation for the presence of a missing mass is that some taxpayers mistakenly believe that IRS audits are more likely when itemizing and switch to the standard deduction to avoid the expected cost of an audit. To assess this explanation, I conduct a survey of taxpayers to elicit their beliefs over audit probabilities and audit costs. I find that the perceived expected cost of audits would explain at most 23% of the cost. A third alternative explanation is that the uncertainty taxpayers face over the amount of deductions they can claim drives them to not itemize. However, the cost I estimate is derived using taxpayers who were itemizing the year before the reform and deductions are stable over time, which implies that taxpayers should have a small uncertainty range. I show that for this theory to explain the result, this uncertainty range would need to be extremely large ($\pm\$14,000$).

Reasonable calibrations of the cost of itemizing suggest that it is unlikely that such a simple task requires so much time. Schedule A is one of the easiest forms to fill out as it does not require any complicated calculations or the use of tax tables. The taxpayer only needs to copy numbers from receipts and then sum them up, which is unlikely to require more than an hour of work.

Such high cost estimates could be consistent with an extreme aversion to

filing taxes. Using a revealed preference argument and survey estimates of the time required to file federal taxes, I estimate that taxpayers dislike working on taxes 4.2 times more than they dislike working. If this is the case, back-of-the-envelope calculations suggest that the overall burden of filing federal income taxes is 1.28% of GDP or 17% of total tax revenue raised. In contrast, the cost of operating Medicare is estimated to be 1 to 6% of total Medicare spending.[4] This suggests that tax collection is inefficient and imposes a large burden on taxpayers and implies the existence of large welfare saving opportunities.

However, there is compelling evidence that individuals are time inconsistent when saving for retirement,[5] searching for a job,[6] smoking[7] and other situations. Time inconsistency introduces a wedge between hassle costs and forgone benefits. A model of time inconsistency based on present bias shows that taxpayers forgo large benefits even when hassle costs are modest because they procrastinate on archiving receipts, eventually leading to large record keeping costs at the time of filing. This makes two testable predictions: first, procrastinators will delay filing until the deadline and second, taxpayers who file close to the deadline will forgo more deductions. I provide empirical evidence consistent with both predictions and show that late filing is a persistent behavior confirming that it is a systematic bias.

Overall, both aversion to tax filing and procrastination suggest that the burden of tax filing is significantly larger than previously estimated whether due to hassle costs per se or behavioral biases.

The results of this paper have implications in several dimensions. First, this is – to my best knowledge – the only paper to provide estimates of the burden of filing taxes by using a quasi-experimental design and directly observing the behavior of taxpayers from administrative tax data. Pitt and Slemrod (1989) address a similar question using tax data but cannot observe the preferences of taxpayers nor use quasi-experimental variation. Instead they estimate a censored model based on Gronau (1973) and Nelson (1977). Their approach

---

[4]Based on Medicare Board of Trustees report and National Health Expenditure Accounts.
[5]Madrian and Shea (2001).
[6]DellaVigna and Paserman (2005).
[7]Gruber and Köszegi (2001).

4

relies on key exclusion restriction assumptions necessary for identification[8]. Both measurement errors and failure of the exclusion restriction lead them to estimate a significantly lower cost of itemizing of $107. There is also a literature[9] that uses survey evidence to estimate hassle costs.[10] Although informative of the time spent filing taxes, it does not capture the preferences of taxpayers and in particular any aversion to filing taxes or any behavioral biases. It also suffers from the usual biases of surveys including high attrition rates and measurement errors.[11] The revealed preference estimates of the cost of itemizing derived in this paper are significantly larger than those estimated using surveys but are consistent with the amount of benefits forgone by individuals in other settings such as health insurance and retirement savings.[12] These results further emphasize the policy relevance of reducing hassle costs and advocate for a simplification of the tax code.

There is an extensive literature that documents low take-up rates of government provided benefits.[13] Three explanations are generally offered: lack of information about the program, stigma costs and hassle costs. This paper is the first to disentangle hassle costs from lack of information and stigma costs and to show that they have a significant effect on benefit take up. The literature has mostly focused on the role of information. Bhargava and Manoli (2015) for example show that failure to claim the EITC can be explained by lack of information about the program but do not address hassle costs. My findings also provide a plausible and additional explanation for other phenomena reported by the literature. Jones (2012) shows that taxpayers fail to adjust their tax withholding resulting in forgone interest payments. He explains his results with inertia but an additional explanation could be the cost of filling out form W4 and sending it to the IRS. Engström et al. (2013) and

---

[8]Discussed in section A

[9]See for example Slemrod (1989).

[10]The hassle costs estimated by this literature are listed in table J.9

[11]Slemrod and Sorum (1984) and Slemrod (1989) report an attrition rate of 71.3%

[12]See table J.10 listing research documenting the magnitude of forgone benefits in other settings.

[13]See Currie (2006) for a survey of the literature.

Rees-Jones (2013)[14] show that taxpayers who have a balance due are more likely to reduce their balance to zero by claiming additional deductions. They provide compelling evidence that this behavior is driven by loss aversion. My estimates show that the cost of sending a cheque to tax authorities could also be a channel for their result.

This paper is also related to a literature in marketing[15] and behavioral economics[16] documenting instances in which consumers fail to claim rebates. Some estimates suggest that only 1% of coupons are eventually redeemed.[17] Explanations of this findings are scarce. My results show that transaction costs (mailing the coupon etc.) are a plausible channel for this phenomenon.

Finally, this paper adds to a long tradition in public economics emphasizing the need to screen out applicants for welfare benefits by imposing high hassle costs[18] such as waiting in line, filling out forms etc. If poorer individuals value their time less – possibly because they are unemployed – then such policies can successfully target them by screening out richer individuals. My results show that this effect is indeed true because richer individuals tend to forgo more benefits than poorer ones. However, given how large hassle costs are, such policies could be screening out too many individuals. In addition, time inconsistency could lead to unwanted distortions such as screening out procrastinators versus non-procrastinators rather than rich versus poor individuals.

1 Data and Institutional Background

1.1 The Decision to Itemize Deductions

Taxpayers can reduce their taxable income by claiming deductions. Consider, for example, a single person with an income of $150,000. In 1989 her marginal tax rate is 28%. If the person spends a total of $10,000 on different expenses that she is allowed to deduct from her income, her tax liability is reduced by $2,800. If instead she decides to claim the standard deduction –

---

[14]See also Feenberg and Skinner (1989).
[15]Silk and Janiszewski (2008).
[16]Ericson (2011) and Letzler and Tasoff (2014).
[17]Inmar (2012).
[18]Nichols et al. (1971) and Duclos (1995).

6

which in 1989 was \$3,100 – her tax liability is reduced only by \$868.

The decision to itemize deductions requires comparing two numbers: the sum of itemized deductions to the standard deduction amount. Itemizing however is administratively burdensome as it requires collecting several documents and working through a separate tax form.

Approximately two thirds of the population claim the standard deduction. The standard deduction amount varies by filing status (single, joint, married fling separately and head of household) and by whether the person is blind or older than 65.

1.2   The Cost of Itemizing

Itemizing deductions is a two-step process. First, the taxpayer has to keep a record of all the expenses she wants to deduct during the year she is filing taxes for, year t. Second, she has to file a separate form when itemizing: Schedule A.

The majority of taxpayers itemize four types of deductions:

- State and local income taxes: these are taxes paid in year t to the state or to the locality. They are reported on form W2 received in January of year $t + 1$. On average they represent 17% of total deductions.
- Mortgage interest: this is the interest paid to finance the main or second home of the taxpayer. It is reported on form 1098 which is received in January of year $t+1$. On average they represent 40% of total deductions.
- Real estate taxes: these are taxes paid on real estate owned by the taxpayer. They can be found in financial records or by calling the county tax assessor. On average they represent 14% of total deductions.
- Charitable donations: any payment made for charitable purposes including to religious institutions. These payments are not subject to third-party reporting. The taxpayer has to keep records of her own receipts. On average they represent 12% of total deductions.

In addition, some taxpayers can also deduct other taxes (sales taxes in some years), other interest expenses (credit-card interest in some years), casualty or theft losses, medical and dental expenses and miscellaneous deductions.

Schedule A is relatively easy to fill out especially if the taxpayer only needs

to itemize the most common deductions outlined above. All she has to do is copy numbers from form 1098, form W2 or charitable contribution receipts, sum them up and copy the sum in the 1040 form. There are no complicated tax schedules nor intricate tax operations. Record keeping is more time consuming as one has to archive the various evidence of expenses to be able to recover them when the tax season arrives. It is however easier to keep track of deductions that are third-party reported given that taxpayers receive the W2 and 1098 in January of year $t + 1$.

1.3 Data

The dataset used to carry this analysis consists of annual cross sections of individual tax returns. It is constructed by the IRS and called the Individual Public Use Tax Files. They are commonly referred to as the Statistics of Income (SOI) files. The data is available annually for the periods that I am analyzing. The number of observation per year ranges from 80,000 to 200,000. The repeated cross sections are stratified random samples where the randomization occurs over the Social Security Number. The data over samples high-income taxpayers as well as taxpayers with business income but weights are provided by the IRS allowing my analysis to reflect population averages. In addition, I use a panel of tax returns known as the University of Michigan tax panel. The panel covers 1979 to 1990 and contains the same variables as the SOI files but has a smaller sample size (less than 40,000 observations per year). Sample restrictions for each figure and table are detailed in appendix section C.

1.4 How Representative Are Itemizers?

The population of joint filers who itemize – considered in this paper – is similar to the full population of joint filers. First, 48% of joint filers itemize deductions. Second, the median AGI for joint filers in 1989 was $63,000 and $68,000 for joint filers who itemize. Slemrod and Sorum (1984) provide evidence that the hassle costs of taxation are higher for very rich individuals, because they have to file more complicated schedules, and very poor ones, because they can face literacy hurdles or severe time constraints, compared to

the rest of the income distribution. It is hard to know how my estimates would extrapolate to these two income groups but they should be representative of households who do not lie at the extreme ends of the income distribution and are likely to be an underestimate for the rest because of the additional constraints poor individuals face and the additional forms rich ones have to file.

2   Missing Mass

If some taxpayers are claiming the standard deduction even though the sum of their itemized deductions is greater than the standard deduction there should be a missing mass in the neighborhood of the standard deduction threshold. I graph the density of deductions for all years ranging from 1980 to 2006 by bin sizes of $2,000[19] in figure 1. The bin closest to the standard deduction only includes itemizers whose deductions are strictly larger than the standard deduction amount. Notice that the density is systematically low in the neighborhood of the standard deduction and then increases and peaks 2 to 3 bins away. This is consistent across years and across filing status. Since I cannot observe the distribution of itemizers below the standard deduction, this cross-sectional evidence does not prove that the missing mass is caused by the standard deduction, and one could argue that is but a naturally occurring feature of the distribution (see appendix figure I.19).

To prove that the missing mass is a distortion due to the standard deduction, I turn to a quasi-experimental design. I exploit large increases in the standard deduction amounts. The largest of these changes happened in 1971 and 1988. Table J.4 reports that the standard deduction increased respectively by 50% and 33%.

I compare the pre-reform year to the post-reform year to account for lagged behavioral responses. Figures 2a and 2b graph the density of deductions in pre and post-reform years for the 1971[20] and 1988 reforms. Notice that the shape

---

[19]All dollar amounts are in 2016 dollars in the rest of the paper.

[20]For the 1971 reform, I compare the pre-reform year to the reform year because another reform of the standard deduction takes place in 1972. This means that the 1971 estimate is likely to be a lower bound as it does not account for any lagged response in 1972.

of the distribution in year t+1 mirrors that of year t-1 and that the missing mass precisely follows the new standard deduction threshold. This shows that some itemizers switch to the standard deduction once it is increased even though their deductions are larger than the standard deduction.

The fact that the missing mass closely follows the standard deduction establishes that there is a discontinuity in the distribution *caused* by the standard deduction. If this missing mass was a feature of the distribution and not due to the standard deduction, it should not track the standard deduction once it is increased.

3 Economic Interpretation of the Missing Mass

The interpretation of the missing mass in the neighborhood of the standard deduction relies on the intuition that taxpayers bear a cost when itemizing but no cost when claiming the standard deduction. I outline this explanation in this section, show that it is consistent with the empirical patterns of the distribution and contrast it - in section 5 - with other explanations by empirically testing predictions of each alternative explanation.

3.1 Hassle Cost

Denote by $f(.)$ the probability density function (PDF) of itemizers when facing no cost and $g(.)$ the PDF of itemizers when there is a cost to itemizing. $C(.)$ the cumulative distribution function (CDF) in the population defined over $[0, c_{max}]$, where c_{max} denotes the largest cost an individual can have, and $c(.)$ denotes the corresponding PDF. For every x, $C(x)$ is equal to the proportion of the population with a cost smaller than x and $c(x)$ is equal to the proportion of the population with cost equal to x.

Let d be the distance to the standard deduction, which is also equal to the benefit of itemizing. At a given point d, the mass of itemizers $g(d)$ is equal to the true (undistorted) mass of itemizers $f(.)$ minus the proportion of individuals with a cost greater or equal to d i.e $g(d) = f(d) - (1 - C(d))$.

Denote by T the after tax-benefit of itemizing deductions and by S the benefit of claiming the standard deduction. Notice that $T - S = d$.

If cost $c(.)$ is constant in the population with $c(x) = c$ for any x in $[0, c_{max}]$,

then any taxpayer who could derive $T - S < c$ from itemizing will claim the standard deduction because the cost of itemizing exceeds its net benefit. The observed distribution of itemizers will therefore be equal to:

$$g(d) = \begin{cases} 0, & \text{if } d < c_{max} \\ f(d), & \text{if } d \geq c_{max} \end{cases}$$

In this case, we would observe a missing mass for any taxpayer such that $T - S < c_{max}$ with nobody itemizing in the $[0, c_{max}]$ region.

If the cost is heterogenous in the population, then

$$g(d) = \begin{cases} f(d) - (1 - C(d)), & \text{if } d < c_{max} \\ f(d), & \text{if } d \geq c_{max} \end{cases}$$

where $f(.)$ is decreasing in the neighborhood of the standard deduction.[21] $C(.)$ is - by definition of a CDF - increasing in d, which implies that $f(d) - (1 - C(d))$ is increasing in d. This implies that $g'(d) > 0$ if $d < c_{max}$ and $g'(d) < 0$ otherwise. In other words - with heterogenous costs in the population - g(.) should be inverse U-shaped and peaks at $d = c_{max}$. This is consistent with the patterns observed in figures 1, 2a and 2b.

3.2 Recovering the Counterfactual Distribution Using the Reform Years

To calculate the distribution of forgone benefits in the population, I need to reconstruct the counterfactual distribution of itemizers. Using the pre-reform year as the counterfactual distribution would lead to an underestimate of the cost because the pre-reform distribution is distorted by its proximity to the standard deduction, as can be seen in figure 2a. This section explains how I reconstruct the counterfactual distribution.[22]

---

[21]This is due to the fact that approximately 65% of the population claims the standard deduction. Unless $f(.)$ is bi-modal (which I address and reject in section 5.3, the distribution has to be decreasing as illustrated in appendix figure I.19.

[22]Intuition and examples for the counterfactual reconstruction process are provided in appendix section D.

3.2.1   Identification Assumptions

I need two assumptions to reconstruct the counterfactual:

- **A1**: The cost is constant across years.

- **A2**: The cost does not increase with the level of deductions.

Assumption A1 can be verified by graphing two densities in years with no reforms and ensuring that they are overlapping. This is confirmed in figure I.15.

Assumption A2 is not testable. Overall, a failure of A2 introduces a bias in the cost estimate but the size of the bias is small: in appendix section B, I provide an upper bound for the size of the bias and show that the estimated cost would lie between \$572 and \$644 if A2 fails.

3.2.2   Counterfactual Distribution

I generate bins of a given size n. Denote by i the distance of a given bin to the standard deduction and by t the year I am considering. b_i^t is the number of taxpayers who have total deductions in the range $[(i-1) \times n + S, i \times n + S]$ where S denotes the standard deduction amount in year t.

For example, if $S = \$10,000$ and $n = \$2,000$ then b_5^{1987} counts the number of itemizers who are located 5 bins away from the standard deduction in 1987, i.e. their total deductions fall in the range $[\$18,000, \$20,000]$.

Denote by $m \geq 1$ the number of bins by which the standard deduction increases after the reform. If the standard deduction increases by \$4,000, and the bin size $n = \$2,000$, then $m = 2$. Recall that $C(.)$ denotes the CDF of the cost of itemizing with support $[0, c_{max}]$. Define $c_i = \frac{C(i+n) - C(i)}{1 - \tau}$, where τ is the marginal tax rate.[23] c_i denotes the proportion of individuals who have a cost of itemizing in the range $[i, i + n]$. If $m < c_{max}$ i.e. the increase in the standard deduction created by the reform is smaller than the range of the cost, I cannot use the density in year t as a counter-factual for year $t + 2$ as the year t density is likely to be distorted and will yield an underestimate of the cost. For this reason, I need to reconstruct the counterfactual density.

Denote by $\widetilde{b}_i^t$ the counterfactual density of itemizers in bin i and in year t. Assuming nothing else but the cost of itemizing affects the densities between

---

[23]It is divided by $1 - \tau$ so that the cost of itemizing is in pre-tax dollars as for deductions.

the pre and post reform years, the true counterfactual densities pre and post reform should overlap, implying:

$$\widetilde{b}^t_{i+m} = \widetilde{b}^{t+2}_i. \tag{1}$$

By definition and for any i and t

$$c_i = \widetilde{b}^t_i - b^t_i. \tag{2}$$

Equations 1 and 2 imply:

$$c_i = \widetilde{b}^t_{i+m} - b^{t+2}_i. \tag{3}$$

Denote by J the smallest j such that $c_j = 0$. J is the bin at which no taxpayer is willing to forgo deductions anymore. If the cost of itemizing is finite, J exists and is unique. In addition, for any $j \geq J$, $c_j = 0$.

$c_J = 0$ and equation 3 imply:

$$\widetilde{b}^t_{J+m} = b^{t+2}_J \tag{4}$$

We know from equation 2 and $c_j = 0$ for any $j > J$ that:

$$\widetilde{b}^t_{J+m} = b^t_{J+m} \tag{5}$$

Equations 4 and 5 therefore imply that:

$$b^t_{J+m} = b^{t+2}_J \tag{6}$$

Equation 6 holds true - by induction - for any $j \geq J$. It states that J can be identified empirically: it is the bin at which the pre and post reform densities overlap and keep overlapping thereafter.

From equation 2 and the fact that for any $j \geq J$ $c_j = 0$, it follows that for any $j \geq J$ $\widetilde{b}^t_j = b^t_j$. In particular:

$$\widetilde{b}^t_{J+m-1} = b^t_{J+m-1}$$

Equation 3 for $i = J - 1$ is:

$$c_{J-1} = \widetilde{b}^t_{J+m-1} - b^{t+2}_{J-1} = b^t_{J+m-1} - b^{t+2}_{J-1} \tag{7}$$

Equation 7 states that to calculate c_{J-1} one only needs to take the difference between the *observed* pre and post-reform densities. This holds true because $c_{J+m-1} = 0$ implying that b_{J+m-1} is the true counterfactual for b^{t+2}_{J-1}.

By induction, it follows that as long as $j + m \geq J$:

$$c_j = \widetilde{b}_{j+m}^t - b_j^{t+2} = b_{j+m}^t - b_j^{t+2} \tag{8}$$

Equation 8 provides an expression for c_j as long as $c_{j+m} = 0$.

For any j such that $c_{j+m} > 0$, equation 8 is replaced by:

$$c_j = \widetilde{b}_{j+m}^t - b_j^{t+2} = b_{j+m}^t + c_{j+m} - b_j^{t+2} \tag{9}$$

Equation 9 has an additional term c_{j+m}, which corrects b_{j+m}^t to account for the fact that it is distorted by its proximity to the standard deduction.

Equation 9 defines c_j as a function of c_{j+m} and the empirically observed densities b_{j+m}^t and b_j^{t+2}. c_{j+m} is calculated using equation 8. Using backwards induction, I can therefore derive each of the $c_J, c_{J-1}, c_{J-2}, ..., c_0$.

This is the process I follow to reconstruct the counterfactual density of itemizers used to calculate the average forgone benefit in the following section. This reconstructed counterfactual is graphed in figure 4 for the 1989 distribution. Using the shape of the reconstructed counterfactual, I extrapolate it to the density below the standard deduction and find that it would account for the 65% of the population claiming the standard deduction.

To calculate the standard errors of the difference between the bins in the 1987 and 1989 densities and 1970 and 1971 densities, I use a bootstrap procedure. The results are reported in table J.5 and table J.6. The difference between the first and second bins is statistically significant with large z statistics (6.55 and 3.47). The rest of the bins are all overlapping with differences that are not significant even at the 10% level, at the exception of bin 10, 11 and 13 that are statistically significantly different at the 5 and 10% level, with differences of very small magnitude (less than 10 times that of the first or second bins).

3.3 Estimation of the Burden of Itemizing Deductions

3.3.1 Distribution of the Burden

The counterfactual density that I constructed allows me to observe the number of taxpayers who itemize in every bin and compare it with the number of taxpayers who should have been itemizing had there been no cost to

14

itemizing.

Following notations from the previous sections, denote by d the amount of tax savings a given taxpayer i can derive from itemizing and c_i the burden of itemizing. d is a random variable: it depends on the mortgage interest, state and local taxes etc. of individual i and the level of the standard deduction, c_i on the other hand is inherent to each taxpayer. What I am able to observe is whether taxpayer i itemizes for a given level of savings d. For a given realization of d a taxpayer who itemizes has a burden $c_i < d$. Denote by d_k the amount of savings a taxpayer derives from itemizing when located in bin k. I can observe the proportion p_k of the population who itemizes when assigned savings d_k: $p_k = \Pr(c_k \leq d_k) = 1 - F(d_k)$, where $F(.)$ is the cumulative distribution function (CDF) of the burden. Hence,

$$F(c_k) = 1 - p_k. \tag{10}$$

In bin 1 for example we know that $d_1 \in (0, 2000]$. By using the difference between the mass of itemizers and the missing mass, I observe that $p_1 = 53\%$. This implies that $F(2000) = 1 - 0.53 = 47\%$ i.e. 47% of the population has a burden that is lower than \$2000 and greater than \$0.[24]

By repeating the same procedure for the remaining bins, I can construct the CDF of c_i. I need the PDF to calculate the average perceived burden of itemizing, which can be derived from the CDF by taking the difference of the proportion of itemizers between each subsequent bin. Denote by m_k the PDF in bin k, then:

$$m_k = p_k - p_{k-1}. \tag{11}$$

The PDF and CDF are shown in table 1(a). Table 1(b) reports the PDF and CDF for a smaller bin size (\$1,000).[25]

---

[24]Unless some individuals enjoy filing taxes, it is safe to assume that $c_i > 0$ for any i i.e. $p_0 = 0$.

[25]Using a smaller bin size yields similar results because as the bin size is reduced, the proportion of taxpayers in a given bin changes: there are less taxpayers who itemize in the first bin when considering a bin size of \$1,000 for example.

3.3.2   Hassle Cost Calculation Using the 1988 Reform

Besides the standard deduction reform, the only reform happening in 1988 that could affect the amount of deductions is the phase out of the personal interest deduction, which I control for (details in section 5.5). There were no other reforms affecting deductions in 1988 or 1989 and the reforms affecting the 1987 distribution do not have lagged effects.[26] I restrict the sample to taxpayers with the same marginal tax rate (28%) and who are not subject to the Alternative Minimum Tax (AMT). There is a marginal tax rate decrease for married filing jointly with income above $45,000 (in 1987 dollars) in 1988. I control for this change by only considering taxpayers with income below $45,000.

I use 1989 as the post-reform year rather than 1988 because the reform occurs in 1988. If taxpayers learn about the increase in the standard deduction when filing their taxes, we should observe the full response in 1989. Figure 3 confirms that the effect is smaller during the reform year.

Table 1 implies that taxpayers forgo large amounts of deductions, resulting in a burden of itemizing of $644 (s.e. 54.1). The average net annual wage for households in the neighborhood of the standard deduction is equal to $92,743. I assume that these households work on average 40 hours a week and 50 weeks a year. Given that all these households fall in the 28% marginal tax bracket, this results in a net wage of $33 for the household. A revealed preference argument implies that taxpayers perceive the task of itemizing to be equivalent to 19 hours of work on average.[27] In contrast, the IRS estimates that itemizing deductions requires 4 hours and 34 minutes to itemize deductions.[28]

---

[26]See section 5.5 for the full list of reforms and appendix section E for the TRA'86 reforms

[27]This is a lower bound as I am assuming that there is only one earner in the household and that their hourly wage is equal to $33. If there are two earners and they work the same amount of hours, their individual wage would be $16.5. Filing the tax return only requires one person, implying that if there are two earners, they would perceive the task of itemizing to be as costly as working 37 hours.

[28]Guyton et al. (2003) describe the methods used by the IRS to calculate the cost of filing taxes which are based on surveys of taxpayers and the Individual Taxpayer Burden Model.

3.3.3 Hassle Cost Estimate Using the 1971 Reform

To calculate the average burden of itemizing deductions using the 1971 reform I need to control for the change in the parallel system of standard deduction which increases from 10% to 13% of AGI. The details of the adjustment are in section F. In addition, and contrary to the 1988 reform, I can only calculate the average burden of itemizing using the average marginal tax rate because there were 25 marginal tax brackets in 1970 and 1971, which could bias the estimate. I also focus on married taxpayers filing jointly to simplify the average marginal tax rate calculations and use bins of $3,065 ($500 in 1970 dollars) to reduce the noise. I estimate that the average burden of itemizing deductions is $996 (s.e. 126).

4 Anatomy of the Missing Mass

4.1 The Burden of Itemizing Deductions Increases With Income

If rich taxpayers value their time more than poor ones because their hourly wage is higher, we should expect them to forgo more deductions. I can test this using the income reported on tax returns. I break down the sample by deciles of income. Because this would significantly reduce the sample size, I consider a moving average of each income decile. For example, the lower income group consists of every individual with income below the second decile threshold. And the second group consists of taxpayers with income above the first decile and below the third decile etc.

Once the groups are constructed and because I have less data points, I fit a polynomial of degree 3 through each deduction bin. I construct confidence intervals around each bin. Any bins for which the confidence intervals overlap are considered as overlapping bins. Using the predicted bins from this polynomial, I calculate the forgone benefits for each group by repeating the procedure developed in the previous section: I compare the distribution in 1987 to that in 1989, reconstruct the counterfactual distribution of itemized deductions and calculate the distribution of the burden of itemizing by comparing the counterfactual distribution to the true one. I only report results for the first six

groups because deductions and income are positively correlated implying that there are very few high income individuals close to the standard deduction threshold. In figure 5(a), the x-axis represents the average income and the y-axis the average burden of itemizing for each income group. All taxpayers in figure 5(a) fall in the 28% marginal tax bracket, implying that the positive relationship between income and forgone benefits is not due to marginal tax rate variation but rather to income per-se. The relationship is increasing and statistically significant (with a t-statistic of 5.62): as income increases taxpayers forgo more benefits consistent with the idea that they value their time relatively more.

Notice that even though itemized deductions increase with income, this is not what drives the increasing relationship between income and forgone benefits. Because I am using a quasi-experimental design, and compare the same income groups before and after the reform, I am implicitly controlling for the relationship between income and deductions.

Figure 5(b) shows the relationship between income and the perceived hours required to itemize deductions.[29] To construct it, I assume that taxpayers work on average forty hours a day and fifty weeks a year and I divide their wages by the number of hours worked per year. By dividing the estimated burden of tax filing by this measure of their hourly wage, I get the perceived hours required to itemize deductions. The relationship between hours and income is increasing and statistically significant (with a t-statistic of 4.6) but considerably less steep than the relationship between forgone benefits and income, consistent with a value of time interpretation.

4.2 Tax Preparers and Electronic Filing

Electronic filing and the use of tax preparers may reduce the cost of filling out forms as one need not file schedule A. However, it does not affect the cost of record keeping. Therefore, if record keeping costs are non-zero, electronic filing or the use of tax prepapers will not reduce costs to zero and one will still observe a missing mass close to the standard deduction threshold. That record

---

[29]The y-axis in figure 5(b) is of the same magnitude as in figure 5(a), this ensures that the two figures are comparable.

keeping is the driver of the cost of itemizing has been consistently documented by survey estimates of the cost of filing taxes.[30]

To test for whether electronic filing or using a tax preparer eliminates the burden of itemizing, I graph the density of itemizers who use a tax preparer and those who use electronic filing in graph 6 and look for whether there is still a missing mass close to the standard deduction threshold. The missing mass is still present implying that tax preparers or electronic filing does not eliminate the burden of itemizing.

Figure 6(b) compares the density of taxpayers who use electronic filing to those who do not. It shows a slightly smaller – but statistically significant – missing mass for taxpayers who file electronically than those who do not. The difference is statistically significant.[31] This is consistent with the missing mass being driven by taxpayers claiming the standard deduction to avoid the cost of itemizing. However, electronic filing only slightly reduces the cost of itemizing and does not eliminate the missing mass, consistent with record-keeping being the main driver of hassle costs.

Unfortunately, I cannot perform a similar test for taxpayers who use tax preparers as the two densities do not overlap away from the standard deduction – possibly because the population that uses tax preparers is intrinsically different from the population that does not – making a comparison of the missing mass impossible. Figure 6(a) shows however that the use of tax preparers does not eliminate the cost of itemizing.

5 Alternative Explanations

5.1 Lack of Information

Information or cognitive abilities are unlikely to explain the forgone deductions. I focus on taxpayers who switch from itemizing to claiming the standard deduction, therefore they should be well aware of the decision to itemize and have the cognitive abilities to do so. In addition, taxpayers are reminded on the 1040 form of the fact that they can itemize deductions as they have to make

---

[30]See for example Guyton et al. (2003), Slemrod and Sorum (1984), Slemrod and Bakija (2008) and Blumenthal and Slemrod (1992).

[31]Bootstrapped standard errors are reported in table J.7

an active decision between itemizing and claiming the standard deduction.

5.2 Evasion

An alternative explanation for the presence of a missing mass is that taxpayers are concerned with being audited by the IRS. They mistakenly believe that audit probabilities are higher when itemizing.[32] Their beliefs over audit probabilities could lead them to switch to the standard deduction once it increases to avoid the expected cost of an audit.

Audit probabilities are very low, therefore for this behavior to explain the missing mass, taxpayers would need to mistakenly believe that audit probabilities are high or that audit costs are large. To address this, I carry a survey of 200 individuals in a wealthy neighborhood of the city of Los Angeles to capture as many itemizers as possible. Appendix section G details the survey instrument. The survey allows me to elicit their beliefs over both the audit probabilities for itemizers and the perceived costs of undergoing an audit. The results of the survey are reported in figure 9.

On average, individuals believe that audits occur with a probability of 8.72%, which is 7.9 times the true audit probabilities.[33] This accounts for at most 23% of the $644 estimated forgone benefits.[34]

5.3 Concave Kink Points

When claiming the standard deduction, taxpayers are paying the full cost of charitable donations, mortgage interest etc. they may have incurred below the standard deduction. When they itemize however, they only pay a portion of it because deductions are subsidized by 1 minus the marginal tax rate. The standard deduction acts as a *concave* kink point: the price of charitable donations is lower when itemizing than when claiming the standard deduction. The indifference curve of a given taxpayer can be tangent at two points of the concave kinked budget set (see figure 7(a)), possibly inducing some taxpayers

---

[32]Unless the ratio of total itemized deductions to income is abnormally large, itemizers are not audited at a higher frequency.

[33]This is consistent with Bhargava and Manoli (2015) who find that EITC filers believe that audit probabilities are 8 times larger than the true ones.

[34]On average, their willingness to pay to avoid an audit is $1,748, which implies that the expected cost of an audit is $147, with a 95% confidence interval of [126, 169].

to be indifferent between two points, one above the standard deduction and one below. Depending on the curvature of the indifference curve, this could create a bi-modal distribution with a missing mass both to the right and to the left of the standard deduction (see figure 7(b)).

However, according to the assumption that taxpayers respond to concave kink points, the size of the missing mass should not respond to variations in income when controlling for the marginal tax rate. The only reason taxpayers should adjust their deductions in response to a concave kink point is because of the marginal tax rate and income should not matter per se in this case. On the other hand, a behavioral response due to hassle costs predicts that richer taxpayers will forgo more money because they have a higher opportunity cost of time even controlling for the marginal tax rate. Figure 5 graphs the relationship between forgone benefits and income - *controlling for the marginal tax rate* - and finds an increasing relationship, rejecting that taxpayers are responding to concave kink points in this setting.

In addition, behavioral responses to concave kink points lead individuals to locate away from the concave kink point. This mechanism is illustrated in figure 7. If behavioral responses to concave kink points were leading to the observed missing mass, as the standard deduction increases, the bi-modal distribution should track the new standard deduction threshold as illustrated in figure 7(c) and the pre and post-distribution peaks should not overlap. The observed pre and post distributions in figure 2a and 2b contradict the prediction of figure 7(c): the pre and post distribution peaks are overlapping rejecting again that the missing mass is caused by behavioral responses to concave kink points.

Overall, both the fact that forgone deductions increase with income and the shape of the post reform distribution of deductions rule out responses to concave kink points. The absence of a behavioral response to concave kink points is consistent with the extensive empirical public finance literature that documents behavioral responses to tax systems which does not find any evidence of responses to concave kink points. Saez (2010), Kleven and Waseem (2013) and Tazhitdinova (2015) directly test the predictions of a behavioral

response to both concave and convex kink points, find responses to convex kink points but no responses to concave kink points.[35]

5.4 Rational Inattention

Can uncertainty over the level of deductions lead a taxpayer to switch to the standard deduction and explain the observed missing mass?

A simple calibration – in appendix section H – of a model illustrative of this type of behavior with varying levels of risk aversion shows that a taxpayer would need an uncertainty range of $\pm\$14,000$ for her to forgo $600 when their true deductions are $12,000. This is especially unlikely given that I focus on taxpayers who were itemizing the year before and because total deductions are highly serially correlated across years for a given individual since 71% of total deductions are mortgage interest, state taxes and real estate taxes which are relatively stable for a given person year after year.

5.5 Other Reforms Affecting Deductions?

Other changes happened in 1988. In this section, I describe these changes and explain how I adjust for the ones that are likely to affect my estimates. The estimates derived in section 3.3.2 already accounted for these adjustments. The fact that the pre and post-reform densities overlap away from the standard deduction threshold shows that the pre-reform density is a relevant counterfactual for the post-reform density in figure 2a and that – after adjusting for these changes – the missing mass estimates are not affected by these changes.

The personal interest deduction was phased out starting from 1986. In 1987, taxpayers could only deduct 65% of their personal interest, 40% in 1988 and 20% in 1989. This is likely to affect the distribution of deductions from 1987 to 1989. To control for this effect, I adjust the 1987 distribution - which is the counterfactual for 1989 - by recalculating the personal interest deduction as if only 20% of it could be deducted. This leads some taxpayers to have deductions below the standard deduction whom I drop. To ensure that there is no behavioral effect associated with the phasing out of the personal interest deduction, I compare the distribution of deductions for individuals below the

---

[35]Kleven (2016) in survey of the bunching literature confirms this finding.

28% marginal tax rate bracket and above. If there was a behavioral effect, we should observe more deductions for individuals above the 28% marginal tax bracket. Graph I.20 shows that there is no discontinuity at the marginal tax rate change at $30,950 in 1989. This is consistent with the fact that the majority of the personal interest deduction is claimed for interest on student loans which are hard to adjust once they are contracted. In addition, after making this correction, I can compare the overlap between the pre and post-reform densities. Away from the standard deduction, the two graphs overlap implying that the post-reform density is an appropriate counterfactual for the 1989 density.

In 1988, the third and fourth marginal tax brackets were removed in favor of two marginal tax brackets (and a 33% rate bubble). To control for this, I only consider taxpayers who were in the 28% MTR bracket in 1987 and in 1989.

The 1971 changes are discussed in appendix section F.

6 Hassle Costs or Behavioral Costs?

A revealed preference argument implies that $644 corresponds to the true hassle cost. This means that when a rational taxpayer is faced with the decision to itemize or claim the standard deduction, she will only itemize if she can save more than $644. The fact that taxpayers experience such a large disutility from itemizing could be due to an extreme aversion to filing taxes and would imply large aggregate filing costs.

However, the behavioral economics literature has documented several instances in which the axiom of revealed preferences fails, in particular because of procrastination. A failure of this axiom introduces a wedge between forgone benefits and hassle costs, reconciling the large magnitude of my estimates with the survey evidence.

In what follows, I discuss the welfare implications of my result in light of both perspectives and show that procrastination accounts for 65% of the forgone benefits.

6.1 Aversion to Filing Taxes

The IRS estimates that itemizing requires less than 4.5 hours. A revealed preference argument implies that taxpayers perceive the task of itemizing to be as costly as working 19 hours. My estimates are 4.2 times larger than the ones provided by the IRS. Their estimates are based on surveys of the time spent filing taxes. However, they do not ask how much taxpayers dislike filing taxes. If my estimates are driven by aversion to filing taxes then – taking the IRS estimates as given – my results suggest that spending one hour preparing taxes is 4.2 times more painful than spending one hour working. Taking this estimate as given, back-of-the-envelope calculations can inform us on the overall burden of filing taxes. These figures are only suggestive as I am inferring preferences over filing taxes from taxpayers who itemize deductions who are richer than non-itemizers and not necessarily representative of the population.

If these are the true preferences of taxpayers then I can use this estimate and the survey estimates of the time required to file the various income tax forms to calculate the aggregate cost of filing taxes. If the wedge between survey estimates and revealed preference estimates is due to the aversion to filing taxes which cannot be captured by surveys, then multiplying the survey estimates by 4.2 would account for the full burden of tax filing including the aversion taxpayers experience when filing taxes. Table 3a shows the results of these calculations. Overall, the cost of filing federal income taxes amounts to 1.28% of GDP in 1989. In comparison, Feldstein (1999) estimates that the efficiency cost of Personal Income Tax and the Payroll Tax ranges between 2 and 5% of GDP. These orders of magnitude emphasize the importance of hassle costs.

6.2 Procrastination

There is extensive evidence that individuals are time inconsistent and tend to procrastinate.[36] If taxpayers procrastinate on filing their taxes, one should observe a large proportion of taxpayers filing on April 15th. And those taxpayers should forgo more deductions. I formalize this argument in appendix

---

[36]See DellaVigna (2009) for a survey of the literature.

section I and show that procrastination can lead to high record keeping costs resulting in individuals failing to itemize.

First, consistent with individuals procrastinating on filing their taxes, I observe that taxpayers bunch at the deadline of April 15th. Figure 8a graphs the volume of Google search of the term *1040* by week and figure 8b uses data from irs.gov[37] and graphs the number of tax returns filed by week.[38] Both exhibit a clear spike in the weeks that include April 15th.

Second, taxpayers who file close to the deadline tend to forgo more deductions consistent with procrastination accounting for a significant portion of the estimated forgone deductions. Figure 8c[39] shows that the missing mass for close to the deadline filers (first two weeks of April) is larger than for March filers.

Note that rational taxpayers should not file close to the deadline for two reasons: by delaying filing, they forgo interest on their refunds and they expose themselves to higher filing costs. Indeed, the sample I use to generate figure 8c only includes taxpayers who are owed a refund by the IRS and therefore have an incentive to file as early as possible to save on interest.[40] Second, filing costs are substantially higher closer to the deadline because lines at the post office and tax preparers are longer and it is harder to get tax help from the IRS because their phone lines are busier than usual.[41]

Notice also that late filing is hard to reconcile with the option value of waiting for low cost realizations. One could argue that taxpayers who bunch at the deadline are rational taxpayers who wait for a low cost realization and face a series of idiosyncratic shocks that force them to file hastily at the very last moment and lead them to forgo benefits. If that is the case, then we should observe that taxpayers who file late in year t are likely to file earlier in year

---

[37]https://www.irs.gov/uac/2016-and-prior-year-filing-season-statistics

[38]A similar finding has been reported by Hoopes et al. (2014) using the volume of calls made to the IRS.

[39]Appendix section I.0.1 explains how the graph is constructed.

[40]Slemrod et al. (1997) estimates that taxpayers forgo $46 million in interest by not claiming their refund as soon as possible.

[41]Redelmeier and Yarnell (2012) for example report that there are more road crash fatalities on April 15th and argue that is due to taxes.

$t+1$. To test for this, I graph the average week in which returns are processed in year $t + 1$ by week of processing in year t, in figure 8d. If taxpayers who bunch at the deadline are doing so for rational reasons, the relationship should be constant as we should observe mean reversion. If they are doing so because of a systematic bias, the relationship should be increasing as year t week of processing should predict year $t + 1$ week of processing. Figure 8d shows an increasing relationship – with a t-statistic equal to 49 – between processing week in year t and year $t + 1$ consistent with the explanation that late filing is due to a systematic bias.

6.2.1 Hassle Costs When Taxpayers Are Naive Present Biased

I use the difference in size of the missing mass for taxpayers who file in March versus taxpayers who file in April to estimate the proportion of the forgone benefits that is due to procrastination.

In section 3.2, I defined $g(.)$ as the observed density of itemizers and I reconstructed $f(.)$ which is the counterfactual density of itemizers if they were facing no cost of itemizing. Therefore, $f(x) - g(x)$ is the proportion of individuals who fail to itemize in bin x.

Denote by f_{march} and f_{april} the observed density of March and April filers. $f_{march}(x) - f_{april}(x)$ corresponds to the difference in proportion of standard filers vs. procrastinators. Therefore $[f_{march}(x) - f_{april}(x)]/[f(x) - g(x)]$ is the percentage of the missing mass due to procrastinators in bin x. On average, procrastinators forgo \$417 (standard errors of the difference between the two densities are reported in table J.8) more than non-procrastinators i.e. 65% of the cost is due to procrastination.

In table 3b, I calculate the aggregate cost of filing taxes assuming taxpayers are naive present-biased using the parameters derived above. The costs of filing taxes when the taxpayer is assumed to be naive present-biased amounts to 0.5% of GDP versus 1.28% if we assume that forgone benefits are only due to rational behavior.

7   Conclusion

Using a quasi-experimental design and a novel method to recover the counterfactual density of deductions, I find that taxpayers forgo large amounts of deductions, resulting in an average burden of itemizing of $644. This implies tax filing costs of a much larger magnitude than previously estimated and raises important policy implications.

Should the burden of tax filing be reduced? The IRS faces a tradeoff between requiring less forms and receipts (and therefore reducing filing costs) versus reducing evasion. In light of the large magnitude of my estimates, it seems to be welfare improving to reduce reporting even if it leads to higher evasion costs.[42]

Filing costs can also be reduced without reducing reporting. This is especially true if taxpayers tend to procrastinate. The IRS can ensure that the deadline for filing taxes falls on a day when people are likely to be less busy such as the weekend. The IRS actually has the opposite policy: if April 15th is a weekend day, the deadline is postponed to the following Monday. In addition, the IRS could shift the burden of filing taxes to firms since they are less likely to be subject to procrastination. This can be achieved through informational reporting as is the case with the mortgage deduction. For example, in the case of charitable donations, requiring charitable organizations to report donations and then having the IRS send a statement to taxpayers (or pre-populate Schedule A) will result in less hassle costs on aggregate because firms are less likely to be time inconsistent and more likely to have a system of information that deals with receipts in a systematic way.

However, some hassle cost can be efficient when designing a tax system and can be used as a policy instrument. This is especially true when the social gains of deductions or loopholes are small and if political economy concerns prevent the government from repealing these deductions. One way of ensuring that taxpayers do not claim them is to impose large hassle costs.[43]

Finally, the identification strategy used in this paper can be exported to

---

[42]Tazhitdinova (2014) explores this tradeoff in the case of charitable donations.

[43]This is discussed in Kaplow (1998)

estimate other hassle costs when individuals have a choice between a low cost low benefits option versus a high cost high benefit one as is the case for example when deciding whether to file a lawsuit in small claims court versus regular court. It can also be used when identifying responses from a censored distribution above or below a certain threshold.

References

Bhargava, S. and D. Manoli (2015). Psychological frictions and the incomplete take-up of social benefits: Evidence from an irs field experiment. *The American Economic Review 105*(11), 3489–3529.

Blank, R. and D. Card (1991). Recent trends in insured and uninsured unemployment: Is there an explanation? *Quarterly Journal of Economics 106*(4), 1157–1189.

Blumenthal, M. and J. Slemrod (1992). The compliance cost of the us individual income tax system. *National Tax Journal*, 185–202.

Currie, J. (2006). The take-up of social benefits. *Public Policy and the Income Distribution*, 80.

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic Literature 47*(2), 315 –372.

DellaVigna, S. and M. D. Paserman (2005). Job search and impatience. *Journal of Labor Economics 23*(3).

Duclos, J.-Y. (1995). Modelling the take-up of state support. *Journal of Public Economics 58*(3), 391–415.

Engström, P., K. Nordblom, H. Ohlsson, and A. Persson (2013). Tax compliance and loss aversion. *American Economic Journal: Economic Policy*, Forthcoming.

Ericson, K. M. M. (2011). Forgetting we forget: Overconfidence and memory. *Journal of the European Economic Association 9*(1), 43 –60.

Feenberg, D. and J. Skinner (1989). Sources of IRA saving. *Tax Policy and the Economy*, 25 –46.

Feldstein, M. (1999). Tax avoidance and the deadweight loss of the income tax. *Review of Economics and Statistics 81*(4), 674 – 680.

Gronau, R. (1973). The effect of children on the housewife's value of time. *The Journal of Political Economy*, S168–S199.

Gruber, J. and B. Köszegi (2001). Is addiction rational? theory and evidence. *The Quarterly Journal of Economics 116*(4), 1261–1303.

Guyton, J., J. O'Hare, M. Stavrianos, and E. Toder (2003). Estimating the

compliance cost of the US individual income tax. *National Tax Journal*, 673 –688.

Handel, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *The American Economic Review 103*(7), 2643–2682.

Holt, C. A. and S. K. Laury (2002). Risk aversion and incentive effects. *American Economic Review 92*(5), 1644–1655.

Hoopes, J. L., D. H. Reck, and J. B. Slemrod (2014). Taxpayer search for information: Implications for rational attention. *AEJ: Economic Policy*.

Inmar (2012). 2013 coupon trends. Year end report.

Jones, D. (2012). Inertia and overwithholding: Explaining the prevalence of income tax refunds. *American Economic Journal: Economic Policy 4*(1), 158 –185.

Kaplow, L. (1998). Accuracy, complexity, and the income tax. *Journal of Law, Economics, and Organization 14*(1), 61–83.

Keys, B. J., D. G. Pope, and J. C. Pope (2014). Failure to refinance. *Working Paper*.

Kleven, H. J. (2016). Bunching. *Annual Review of Economics 8*(1).

Kleven, H. J. and M. Waseem (2013). Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from pakistan. *The Quarterly Journal of Economics 129*, 669–723.

Letzler, R. and J. Tasoff (2014). Everyone believes in redemption: Overoptimism and nudges. *Journal of Economic Behavior and Organization*, 107–122.

Levit, K. R. (1985). Personal health care expenditures, by state: 1966-82. *Health Care Financing Review 6*(4), 1.

Little, A. D. (1988). Development of methodology for estimating the taxpayer paperwork burden.

Madrian, B. and D. Shea (2001). The power of suggestion: Inertia in 401(k) participation and savings behavior. *Quarterly Journal of Economics 116*(4), 1149–1187.

Nelson, F. D. (1977). Censored regression models with unobserved, stochastic censoring thresholds. *Journal of Econometrics 6*(3), 309–327.

Nichols, D., E. Smolensky, and T. N. Tideman (1971). Discrimination by waiting time in merit goods. *The American Economic Review*, 312 –323.

Pitt, M. and J. Slemrod (1989). The compliance cost of itemizing deductions: Evidence from individual tax returns. *The American Economic Review*, 1224–1232.

Redelmeier, D. A. and C. J. Yarnell (2012). Road crash fatalities on us income tax days. *The Journal of the American Medical Association 307*(14), 1486–1488.

Rees-Jones, A. (2013). Loss aversion motivates tax sheltering: Evidence from US tax returns. Working paper.

Saez, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, 180–212.

Silk, T. and C. Janiszewski (2008). Managing mail-in rebate promotions. Working paper.

Slemrod, J. (1989). The return to tax simplification: An econometric analysis. *Public Finance Review 17*(1), 3–27.

Slemrod, J. and J. Bakija (2008). Taxing ourselves: a citizen's guide to the debate over taxes. *MIT Press Books*.

Slemrod, J., C. Christian, R. London, and J. A. Parker (1997). April 15 syndrome. *Economic Inquiry 35*(4), 695 –709.

Slemrod, J. and N. Sorum (1984). The compliance cost of the US individual income tax system. *National Tax Journal*, 461 –474.

Steuerle, E., R. McHugh, and E. M. Sunley (1978). Who benefits from income averaging? *National Tax Journal*, 19–32.

Sydnor, J. (2010). (over) insuring modest risks. *American Economic Journal: Applied Economics 2*(4), 177–199.

Tazhitdinova, A. (2014). Optimal reporting threshold: Theory and evidence from charitable contributions. Manuscript, UC Berkeley.

Tazhitdinova, A. (2015). Tax breaks for low earners: Who benefits from them? evidence from mini-jobs in germany. *Working Paper*.

Wicks, J. H. (1965). Taxpayer compliance costs from the montana personal income tax. *Montana Business Quarterly 3*, 36–42.

Wicks, J. H. and M. N. Killworth (1967). Administrative and compliance costs of state and local taxes. *National Tax Journal*, 309–315.

Figure 1: Missing Mass In the Neighborhood of the Standard Deduction
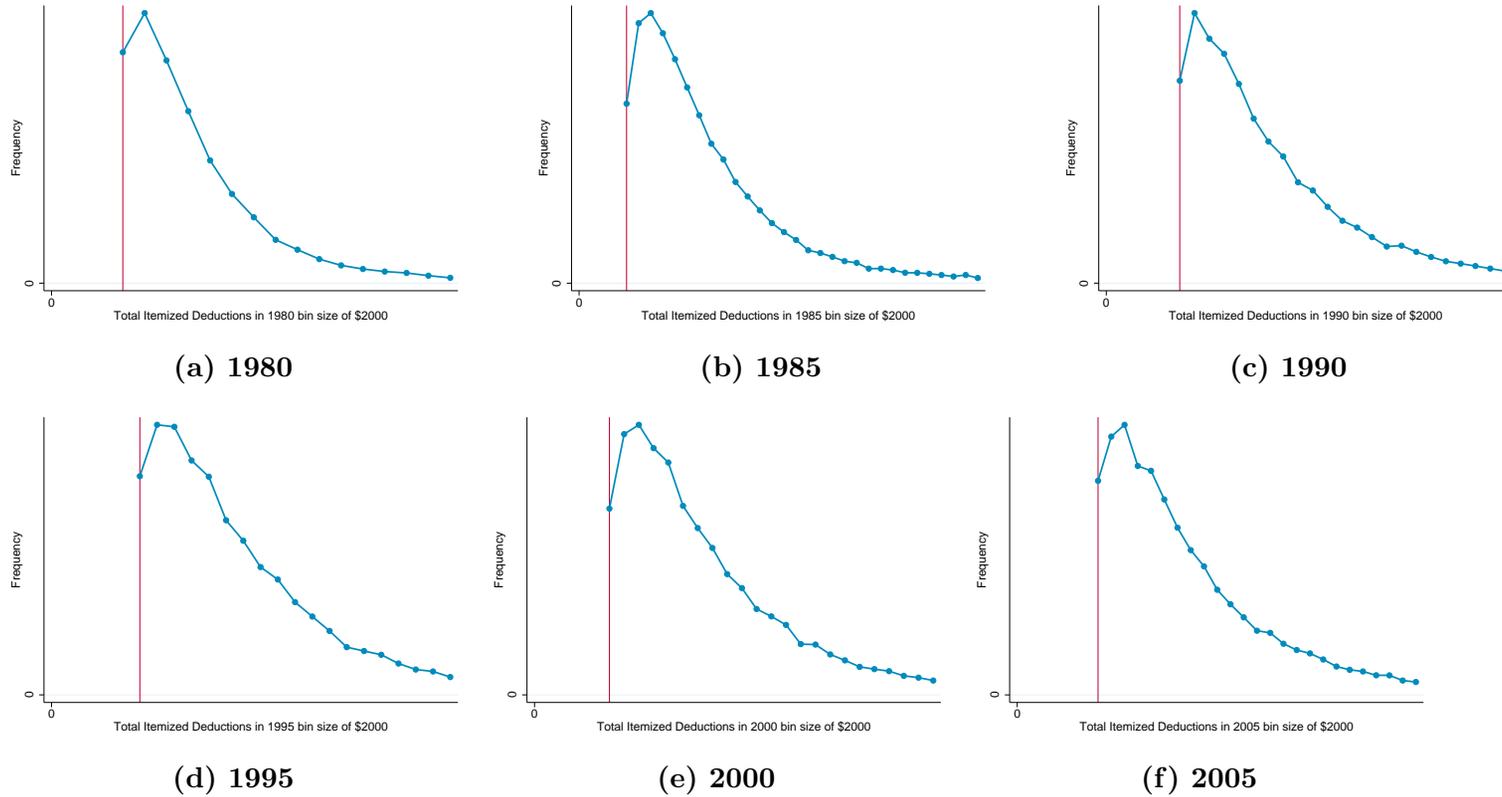
(a) 1980

(b) 1985

(c) 1990

(d) 1995

(e) 2000

(f) 2005

Notes: The figures above plot the density of deductions for itemizers filing jointly. The bin size is $2,000 and the vertical line represents the standard deduction threshold for each year. Notice the missing mass in the neighborhood of the standard deduction threshold. Additional years are reported in appendix figures I.10, I.11, I.12 and I.13 and figure I.14 for single filers.

Figure 2: Density of Deductions for Itemizers Filing Jointly Pre and Post Reform

(a) The 1988 Reform
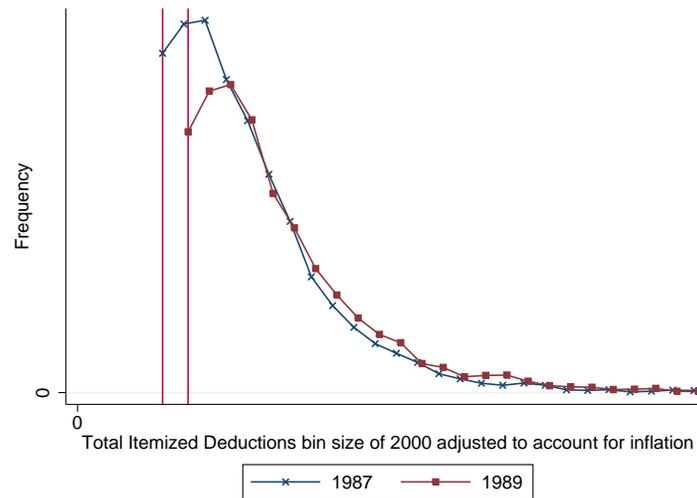
(b) The 1971 Reform
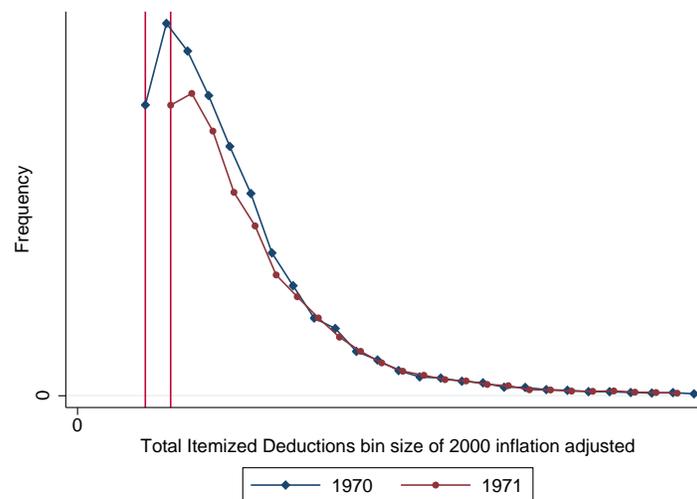
Notes: The first graph plots the density of deductions for the 1988 reform and the second one for the 1971 reform. Notice that the pre-reform density is higher than the post-reform density specifically in the neighborhood of the standard deduction, whereas the two densities are very similar when comparing them further away from the standard deduction. The statistical difference between the two densities is reported on appendix tables J.5 for 1988 and J.6 for 1971.

Figure 3: Lagged Response: Smaller Effect During Reform Year (1987-1988)
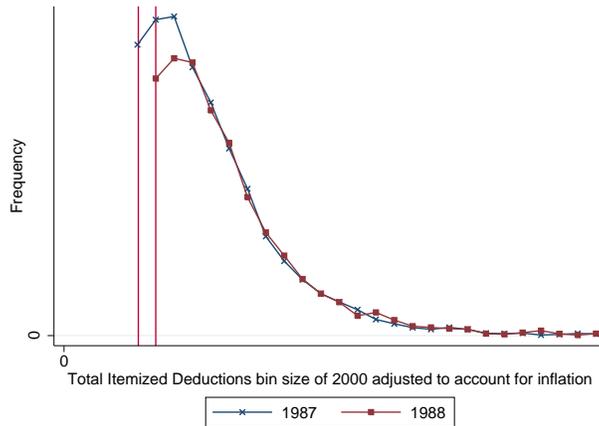
Notes: This graph plots the distribution of deductions for itemizers filing jointly in 1987 and 1988. Notice that the missing mass is smaller than in figure 2a showing that there is a lagged response to the reform.

Figure 4: Reconstructed Counterfactual
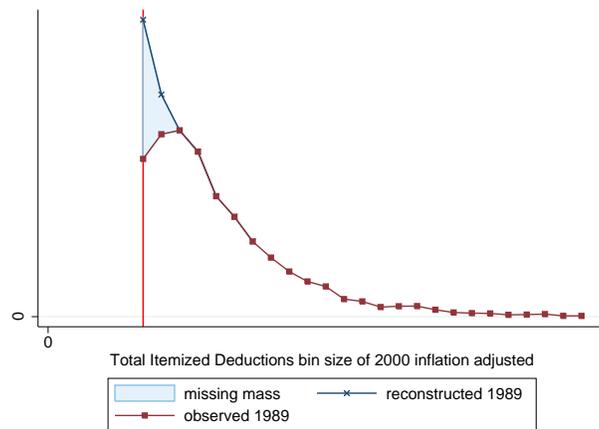
Notes: This graph plots the reconstructed counterfactual in 1989 using the method outlined in section 3.2 and the observed density for 1989. The missing mass used to estimate the burden of itemizing is given by the area lying between the two curves. The distribution of the burden of itemizing is provided in table 1(a) for a bin size of $2,000 and table 1(b) for a bin size of $1,000.

33

Figure 5: Relationship Between Income and the Burden of Itemizing Deductions

(a) Burden of Itemizing and Income
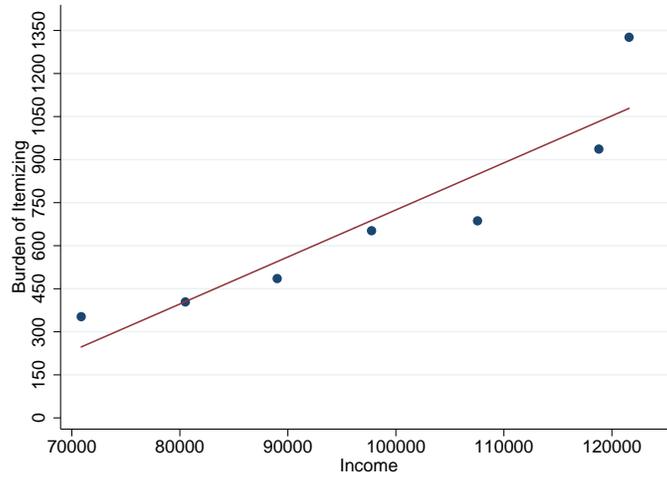
(b) Hours Spent Itemizing and Income
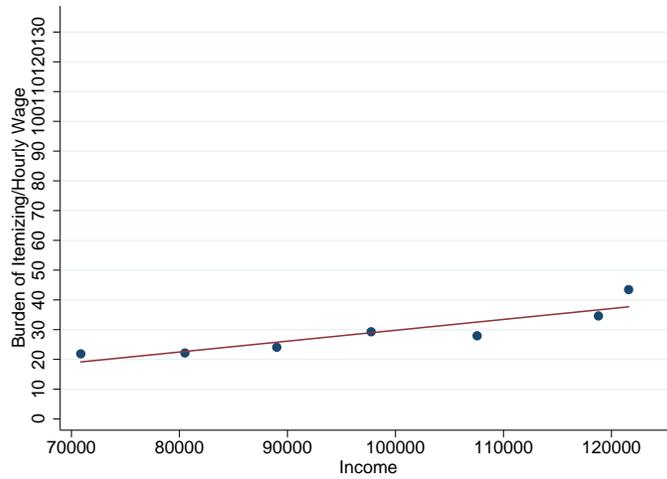
Notes: (a) The first graph shows the increasing relationship between income and the burden of itemizing: richer households are more likely to forgo deductions. This relationship controls for the variation in MTR across the different income groups. (b) The second graph divides the burden of itemizing by the hourly wage and shows the implied hours spent itemizing by each income group. The y-axes of each graphs have scales of the same magnitude.

Figure 6: Use of Tax Preparer and Electronic Filing

(a) Tax Preparer
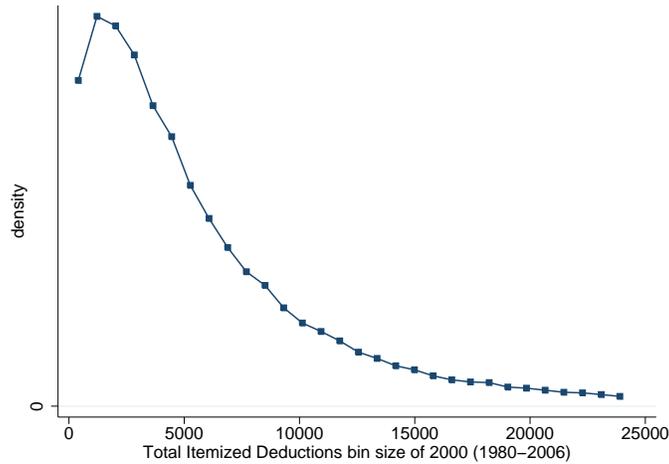
(b) Electronic Filing
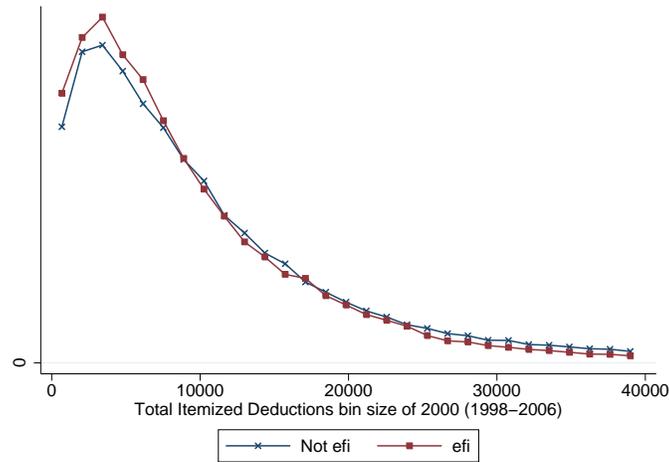
Notes: The x-axis is normalized such that 0 corresponds to the standard deduction threshold. Graph (a) plots the density of total deductions for taxpayers who use tax preparers from 1980 to 2006 by bin size of $2000. Graph (b) plots the density of total deductions for taxpayers who file returns electronically from 1998 to 2006 by bin size of $2000 and compares it to the density of taxpayers who do not file returns electronically. Both graphs exhibit a missing mass close to the standard deduction implying that neither tax preparers nor electronic filing eliminate the burden of itemizing. The use of electronic filing slightly reduces the missing mass consistent with hassle costs being the driver of the missing mass and record-keeping being the largest portion of the cost of itemizing.

Figure 7: Concave Kink Point: Densities Following Reform Should Not Overlap

(a) Concave Kink Point

(b) Missing Mass Due to Concave Kink Point

Disposable Income

Deductions allowed above Standard Deduction

No tangency point

Standard Deduction

Charitable Donation

Frequency

Pre-reform density

Standard Deduction

Total Deductions

(c) Pre and Post Reform Densities

Frequency

Pre-reform density

Post-reform density

SD increase

Total Deductions
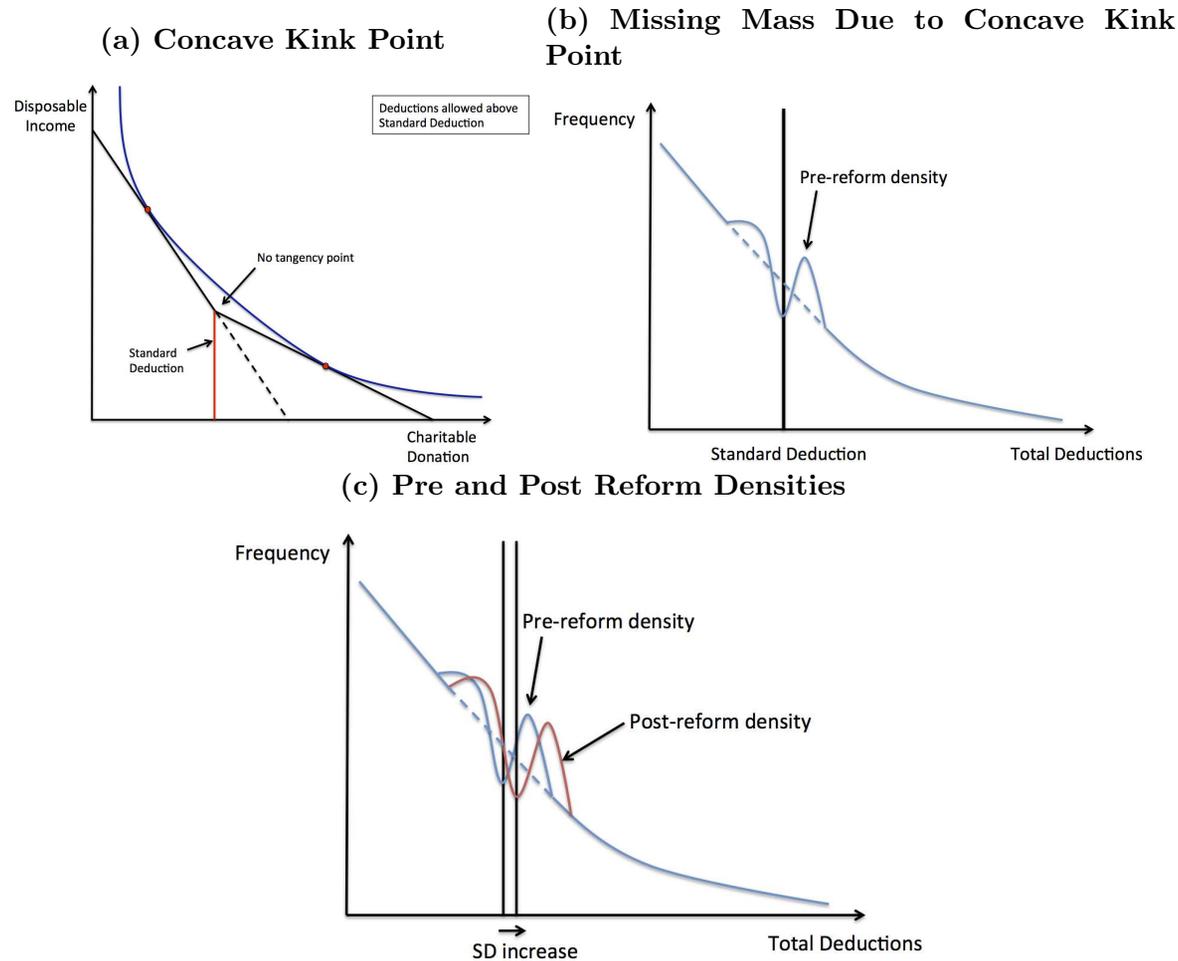
Notes: Panel (a) displays a budget set with a concave kink point. Panel (b) shows the effect that a concave kink point could in theory have on the density of itemizers. Panel (c) shows that if itemizers were responding to the concave kink point, we should observe that the pre and post reform densities are not overlapping in the neighborhood of the standard deduction. This is contradicted by figure 2a, therefore ruling out a behavioral response to a concave kink point.

Figure 8: Deadline Effects

(a) Google Search of the Term 1040 **(b) Number of Returns Filed by Week**
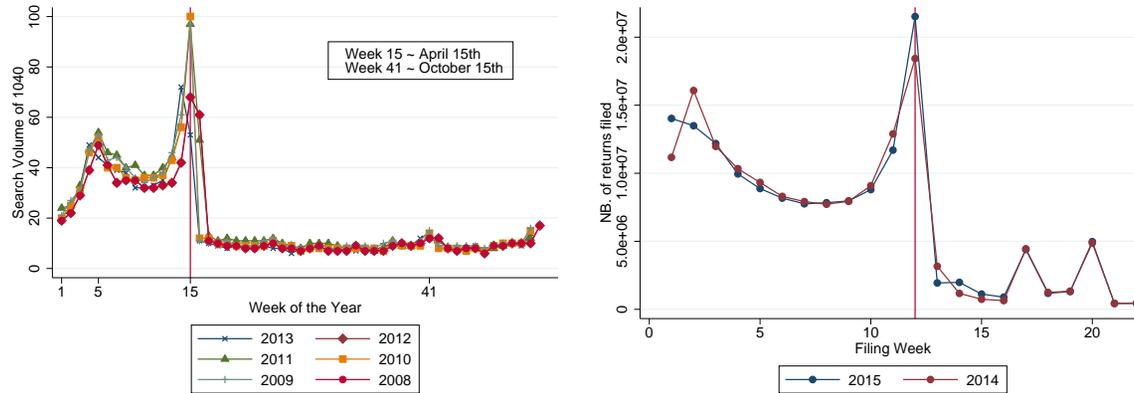
(c) March Itemizers v.s. April Itemizers (d) Processing Week in Year t v.s. $t-1$
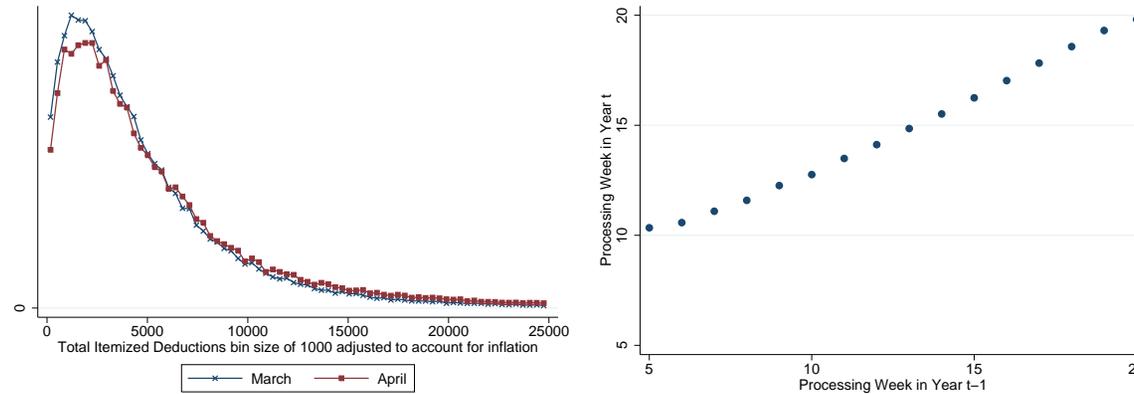
Notes: Panel (a) plots the volume of search of the term "*1040*" in Google and panel (b) plots the volume of tax returns filed by week in 2014 and 2015. The red vertical line corresponds to the week of April 15. Panel (c) plots the density of itemizers who file in March versus in April, the x-axis is normalized such that 0 corresponds to the standard deduction. Panel (d) plots the average week in which a return is processed in year t on the y-axis and the average week in which a return is processed in year $t-1$ on the x-axis.

Figure 9: Audit Survey

(a) Expected Cost of Audit

(b) Audit Probability

(c) Cost of Audit

(d) Income
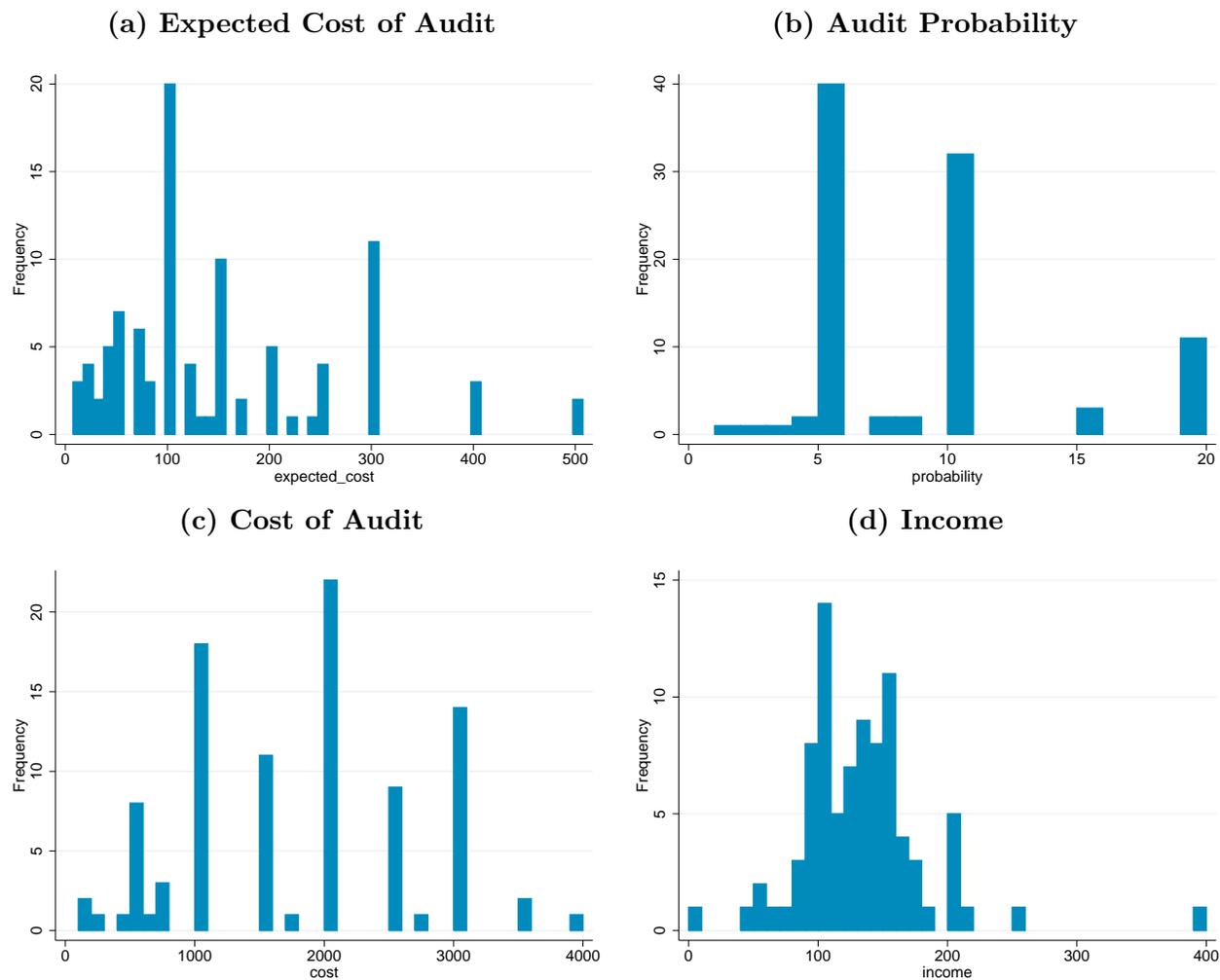
Notes: The number of observations for each panel is 95 individuals. Panel (a) is the distribution of expected cost of audit and is equal to the product of audit probabilities by cost of audit. Panel (b) is the distribution of perceived audit probabilities. Panel (c) is the distribution of cost of audit. Panel (d) is household income in brackets of $10,000. 195 individuals were surveyed, of which 95 files their taxes themselves and itemize deductions.

Table 1: Cumulative Distribution Function of the Burden of Itemizing

(a) Bin size of $2,000

Deduction Interval (b_k)	Average Deduction	Average Benefit	CDF (p_k)	PDF (m_k)
(0, 2000]	$1000	$280	53%	53%
(2000, 4000]	$3000	$840	82%	29%
(4000, 6000]	$5000	$1400	100%	18%

(a) Bin size of $1,000

Deduction Interval (b_k)	Average Deduction	Average Benefit	CDF (p_k)	PDF (m_k)
(0, 1000]	$500	$140	43%	43%
(1000, 2000]	$1500	$420	63%	20%
(2000, 3000]	$2500	$700	79%	16%
(3000, 4000]	$3500	$980	86%	7%
(4000, 5000]	$4500	$1260	100%	14%

Notes: These two tables report the Cumulative Distribution Function (CDF) and Probability Density Function (PDF) of the perceived burden of itemizing deductions. Table 1(a) uses a bin size of $2,000 and table 1(b) uses a bin size of $1,000. The first column corresponds to deductions, and the second to the after tax deductions. For example, the second row of table 1(b) corresponds to taxpayers who can save $1,000 to $2,000 of deductions, which is on average $1,500 of tax deductions and corresponds to $420 with 28% marginal tax rate. The CDF is calculated by comparing the proportion of taxpayers who itemize and those who fail to itemize. In the first row for example, 43% of taxpayers itemize implying that their perceived burden of itemizing is less than $1,000 of deductions. The average burden of itemizing is a weighted average given by the product of the average benefit and the PDF. b_k, m_k and c_k refer to the notation used in section 3.2.

Table 3: Aggregate Burden of Filing Taxes

(a) Assuming Taxpayer Is Rational

Form	Hours (from IRS)	Hourly Wage (in $)	Tax-Aversion Coefficient	Individual Burden (in $)	Nb. of Taxpayers (million)	Aggregate Burden (in $b.)	% of GDP
1040	9.40	17.69	4.32	718.90	0.11	80.52	0.74
Sch. A	4.53	34.70	4.32	679.37	0.03	21.74	0.20
Sch. B	1.28	22.26	4.32	123.13	0.01	1.53	0.01
Sch. C	9.63	22.12	4.32	920.62	0.01	12.89	0.12
Sch. D	3.75	37.05	4.32	600.48	0.01	5.12	0.05
Sch. E	5.83	35.67	4.32	898.71	0.01	12.76	0.12
Sch. F	16.10	21.43	4.32	1491.05	0.00	3.52	0.03
Sch. SE	1.13	17.00	4.32	83.05	0.01	0.96	0.01
Total						**139**	**1.28**

(b) Assuming Taxpayer Procrastinates

Form	Hours (from IRS)	Hourly Wage (in $)	Tax-Aversion Coefficient	Individual Burden (in $)	Nb. of Taxpayers (million)	Aggregate Burden (in $b.)	% of GDP
1040	9.40	17.69	1.69	281.36	0.11	31.51	0.29
Sch. A	4.53	34.70	1.69	265.88	0.03	8.51	0.08
Sch. B	1.28	22.26	1.69	48.19	0.01	0.60	0.01
Sch. C	9.63	22.12	1.69	360.30	0.01	5.04	0.05
Sch. D	3.75	37.05	1.69	235.01	0.01	2.00	0.02
Sch. E	5.83	35.67	1.69	351.73	0.01	4.99	0.05
Sch. F	16.10	21.43	1.69	583.55	0.00	1.38	0.01
Sch. SE	1.13	17.00	1.69	32.50	0.01	0.38	0.00
Total						**54.4**	**0.5**

Notes: These tables report the aggregate burden of filing each item of the income tax return. The hours are reported by the IRS in the documentation accompanying the 1040 form. The hourly wage is estimated from the reported wage in the SOI dataset by restricting the sample to individuals with positive wages and dividing the annual wage by 2000 hours. The number of taxpayers is reported in the SOI files. The aggregate burden is equal to the product of the hours required to file the form, the estimated aversion to filing coefficient, the average hourly wage and the number of taxpayers. The GDP in 1989 was 10.9 trillion dollars (in 2016 dollars).

WEB APPENDIX
FOR ONLINE PUBLICATION

A Pitt and Slemrod (1989)

Pitt and Slemrod (1989) very elegantly apply the methods of Gronau (1973) and Nelson (1977) to assess the hassle cost of itemizing deductions by estimating a censored model with unobserved censoring thresholds using maximum likelihood.

To do so they estimate a cost and benefit function of itemizing deductions. The benefit of itemizing is given by $TS_i = X_i\beta + u_i$ where X_i are exogenous and observed characteristics, β is a vector of parameters and u_i an error term. Similarly, the cost of itemizing is assumed to be $C_i = Z_i\gamma + v_i$, where Z_i are exogenous and observed characteristics, γ a vector of parameters and v_i an error term. A person will itemize if $TS_i \geq C_i$. TS_i is only observed when $TS_i \geq C_i$ but C_i is never observed. Gronau (1973) and Nelson (1977) show that if u_i and v_i are uncorrelated or if there are some characteristics present in X_i but not in Z_i then the model is identified and a likelihood function can be maximized to estimate both TS_i and C_i. Pitt and Slemrod (1989) acknowledge that there is no reason to assume that the errors are uncorrelated but that there are some characteristics that are likely to be present in X_i but not in Z_i, therefore arguing that identification should be valid.

The set of exogenous and observable characteristics they consider to estimate both β and γ are whether a person is married, her AGI, the square of AGI, whether a person owns a farming business, the number of age exemptions a person claims and the number of exemptions claimed. The set of exogenous characteristics specific to β are positive investment income, the average state income and sales taxes for an income of $40,000, the average property tax rate in a given state and an index of medical costs in a given state.

If the assumptions from Gronau (1973) and Nelson (1977) hold and given these exogenous and observed characteristics, they can estimate the cost and benefit function. They find that the average cost of itemizing is $107 (in 2016

dollars) i.e. 6 times lower than the cost I estimate.

Since Pitt and Slemrod (1989) acknowledge that u_i and v_i are likely to be correlated, for the Gronau (1973) and Nelson (1977) estimators to be consistent, the exclusion restriction imposed on X_i and Z_i becomes necessary for identification.

Unfortunately, the authors do not directly test the exclusion restriction. Their main argument for why it should hold is that the dollar value of a deduction should not affect how costly it is to deduct it. For example a $1,000 medical bill should be equally costly to deduct as a $2,000 medical bill. Although this argument is very reasonable in the case of state income tax deductions or mortgage interest deductions for which taxpayers receive one statement with the total deductible amount, it is hard to argue that larger charitable deductions or medical expenses are not also correlated with larger record-keeping. If a person only goes to the hospital once, then their assumption holds, but the more a person will go to a hospital, the larger the number of receipts she will have to keep track of, the larger her medical expenses will be and the costlier it will be to deduct them.

In addition, there are good reasons to believe that the observable characteristics from X_i that Pitt and Slemrod (1989) claim are correlated with TS_i are in fact also very likely to be correlated with C_i which would lead to a failure of the exclusion restriction. I discuss them for each identifying variable:

- Positive investment income: the correlation between positive investment income and AGI can be directly tested in the data. I use the 1982 cross section of the SOI files to construct an investment income variable by taking the sum of interest income, dividends received and capital gains included in AGI. I then regress the investment income variable conditional on it being positive on AGI and using the appropriate population weights, I find a coefficient of 0.49 and a t-statistic of 223.80, implying high correlation. This implies that positive investment income fails the exclusion restriction.

- Medical cost index: it is described as being expenses per day at commu-

nity hospitals in the state and is taken from Levit (1985). This index is likely to be higher in states that have an older population since medical expenses increase with age. But age is used in the cost estimation equation C_i, which would lead a failure of the exclusion restriction for this variable.

- The average rate of state income and sales tax at $40,000 (1979 dollars) of AGI. This variable could be correlate with both age and income. If older people move to states with no state income tax – such as Florida – there would be a correlation with age. In addition, if the incidence of income taxes falls on income we would also observe a correlation with AGI. These two reasons are likely to lead to a failure of the exclusion restriction. This variable suffers from an additional problem which is that it is estimated for individuals who have $40,000 of income. Pitt and Slemrod (1989) show on page 1229 the distribution of income for itemizers: 75% of them have an annual income lower than $40,000. This measurement error means that the estimated coefficient for the states sales and income tax is biased downwards and further explains why they find a smaller cost than I do.

- Property taxes: states with low property taxes are less likely to attract family with children which would imply a negative correlation between marital status, number of dependents and property taxes. Similarly, older individuals without kids looking to relocate will look for states with lower property taxes also leading to possible correlations between property taxes and age, leading again to a failure of the exclusion restriction.

Does the failure of the exclusion restriction bias the estimates from Pitt and Slemrod (1989) upwards or downwards? Some coefficients from their main estimation in table A1 are significantly smaller than intuition would predict. In particular, state taxes and property taxes have coefficients of 1.1 and 0.8 compared to 25.8 for the square of AGI even though an increase in state taxes or property taxes results in a mechanical increase in the benefit

43

of itemizing since they result in a one-to-one increase in total deductions. This suggests that the failure of the exclusion restriction biases the estimates downwards.

B   Assumption A2

Assumption A2 states that the cost should not increase with the level of deductions. It makes sense to assume that the cost of deducting \$10,000 worth of mortgage interest is the same as deducting \$100,000 because total mortgage interest is reported on form 1098. However, it is also reasonable to assume that an individual who donates \$100,000 to charity is more likely to donate to more charities than an individual who donates \$10,000. This would imply that cost is likely to increase with the level of deductions and A2 is likely to fail.

Assumption A2 is important for equation 9, it allows me to use c_{j+m} to recover c_j. Intuitively, it allows me to infer the distortion imposed by the standard deduction on the pre-reform distribution in bin j from bin $j+m$ when the pre- and post-reform standard deduction thresholds are m bins away. If the cost of itemizing is orthogonal to the total amount of deductions, I can use c_{j+m} to correct b^t_{j+m}. A2 can fail if the cost of itemizing decreases with the size of total deductions which would bias my cost estimate downwards. But more importantly it can fail if the cost of itemizing increases with the size of total deductions, which would overestimate the cost. There is an easy way to provide an upper bound for the bias introduced by a failure of A2: by using the pre-reform distribution (b^t_{j+m}) as the true counterfactual i.e. replacing equation 9 by $c_j = b^t_{j+m} - b^{t+2}_j$ and dropping the adjustment term c_{j+m}. This is a generous upper bound because it assumes that the pre-reform distribution is undistorted in the neighborhood of the standard deduction in spite of figure 2a showing a clear distortion. In this case, the estimated cost would be \$563 instead of \$717. Therefore if A2 fails, the cost of itemizing would lie between \$563 and \$717.

C   Sample Restrictions

C.1   Figure 1

The sample used for figure 1 are joint filers who itemize deductions. I focus on joint filers because they represent more than 50% of the population and the standard deduction is specific to the filing status. This means that I cannot show every tax filing status on the same graph because they would have different standard deductions. Joint filers provide the highest power because they have a larger sample size than of the other filing status.

Figure I.14 shows the same patterns for single taxpayers.

C.2   Figures 2a, 2b and 3

In figure 2a and 3, I restrict attention to taxpayers who are married filing jointly for the reasons outlined in section C.1. In addition, in 1988 and 1989 there were two tax brackets (15% and 28%) and a tax rate "bubble" (33%). Most taxpayers who itemize deductions fall in the 28% marginal tax bracket. Therefore, to control for the effect of the marginal tax rate, I only consider taxpayers who fall in the 28% marginal tax rate bracket. This allows me to precisely calculate the amount of after tax forgone benefit.

In figure 2b, I focus on married filing jointly as well but I do not control for the marginal tax rate. This is because there are 25 different marginal tax brackets in 1970 ranging from 14% to 70%. Selecting taxpayers who have the same marginal tax rate will reduce the sample size too much rendering the estimates too imprecise.

C.3   Figure 5

In figure 5, I use the same sample restrictions as in figure 2a and 3 and break down the sample into deciles of income.

C.4   Figure 6

To generate figure 6, I consider joint filers as explained in section C.1. In figure (a), I consider all years from 1980 to 2006 but exclude 1985 and 1990 because the tax preparer variable is missing in those years. In figure (b), I consider all years from 1998 to 2006 because few taxpayers used electronic

filing prior to 2006.

C.5 Figure 8c

The variable indicating the week in which a return is processed by the IRS is only present in the SOI files in year 1980 to 1999. Thus, to generate figure 8c, I restrict attention to those years. I use the same sample restrictions as in figure C.1 in addition to dropping taxpayers who have a balance due to the IRS. If taxpayers owe money to the IRS, it is rational to wait as much as possible so as to save on interest.

C.6 Taxpayers Who Have To Claim the Standard Deduction

In rare cases, taxpayers have to claim the standard deduction even when their itemized deductions exceed the standard deduction. This happens in the following four cases:

1. A married taxpayer whose spouse files separately and itemizes deduction.
2. In some states, a taxpayer who wants to itemize on her state tax return has to itemize on her federal tax return as well.
3. A taxpayer who is neither a citizen nor a permanent resident of the United States.
4. A taxpayer who can benefit from itemizing for alternative minimum tax purposes even though the standard deduction is greater than the sum of her itemized deductions.

D Counterfactual Distribution Reconstruction Using an Illustrative Example

To calculate the burden of itemizing deductions I create bins of $2000.[44] I calculate the weighted frequency of individuals located in those bins. I subtract the mass of the 1989 bin from the mass of the corresponding bin in 1987 after adjusting the amounts to account for inflation.

To calculate the standard errors of the difference between the bins in the 1987 and 1989 densities and 1970 and 1971 densities, I use a non-parametric

---

[44]I also consider $1,000 bin sizes in table 1(b), which yields similar results.

bootstrap[45] with 100 replications. The results are reported in table J.5 and table J.6. The difference between the first and second bins is statistically significant with large z statistics (6.55 and 3.47). The rest of the bins are all overlapping with differences that are not significant even at the 10% level, at the exception of bin 10, 11 and 13 that are statistically significantly different at the 5 and 10% level, with differences of very small magnitude (less than 10 times that of the first or second bins).

This approach allows me to measure the percentage of individuals that claim the standard deduction even though their total itemized deductions exceed the standard deduction amount by multiples of $2,000.

Once I get those percentages, I need to adjust the 1987 distribution to get the true counterfactual as it might be distorted by its proximity to the standard deduction threshold. For clarity I associate each bin with a number that denotes its distance from the standard deduction amount. For example, in 1987 the standard deduction amount is $7,865. This means that bin [7865, 9865] is called bin number 1 in 1987 and bin [9865, 11865] is called bin number 2 in 1987. Bins in 1989 are defined in a similar way relative to the standard deduction amount of $9,991: bin [9991, 11991] is bin number 1 and bin [11991, 13991] is bin number number 2.

To recover the counterfactual distribution of deductions I use the fact that the distribution of costs should be the same in 1987 and 1989.[46] I consider the first bins for which the 1987 and 1989 densities are overlapping. Figure 2a shows that these are bin number 3 in 1989 and bin 4 in 1987. In 1989, taxpayers located in bin number 3 can save from $4,000 to $6,000 worth of deductions. Given that the 1987 and 1989 densities are overlapping in the 1989 bin 3, no taxpayer is willing to forgo more than $4,000 worth of deductions. In 1987, taxpayers who can save from $4,000 to $6,000 of deductions are located in bin 3. Since I assume that the cost is the same across years, this means

---

[45]For every bin, I resample observations with replacement, estimate the bin size for each repetition and then use the variance of all the repetitions to calculate the standard errors.

[46]It is reasonable to assume that the time required to itemize deductions in 1987 and 1989 is the same because there were no changes to the Schedule A form or to the record keeping requirements.

that 1987 is the undistorted counterfactual for the 1989 density in bin 2. Bin 1 of the 1989 density is compared to bin 2 of the 1987 density. However, from observing the difference between the 1987 bin 3 and 1989 bin 2, I know that some taxpayers will forgo deductions, biasing the 1987 bin 2 downward. By using the difference between the 1987 bin 3 and 1989 bin 2, I can calculate this bias and correct bin 2 in 1987 to get the true counterfactual for the 1989 bin 1.

The following hypothetical example illustrates the adjustment process and formalizes the approach I use to recover the true counterfactual density. I generate an undistorted hypothetical density of deductions in figure I.18. Each bin size is equal to $100. I assume that the distribution of the burden of itemizing in the population is given by the following:

- 40% have a burden lower than $100
- 70% have a burden lower than $200
- 85% have a burden lower than $300
- 95% have a burden lower than $400

I introduce a standard deduction in the second bin in figure I.18 and apply the distribution assumed above to the density. To calculate the distribution of the burden in this scenario, I would simply compare the percentage difference between the true density and the distorted one. However, the true density is unobserved. In order to reconstruct it, I use an exogenous increase in the standard deduction. Figure I.18 assumes that the distribution of the burden is the same across years and introduces a reform that increases the standard deduction amount by $200 (2 bins). I denote by d_i the distortion introduced by the standard deduction in bin i. 40% of the population experiences a burden smaller than $100. This means that $1 - 40\% = 60\%$ will claim the standard deduction in the first bin. This implies that the first bin is distorted by 60% i.e. $d_1 = 60\%$. Similarly, $d_2 = 30\%$, $d_3 = 15\%$ and $d_4 = 5\%$ and $d_i = 0$ for any $i > 4$.

Denote by b_i^t the bin density, where i is the distance (in bins) to the standard deduction and t is the year. Year t corresponds to the pre-reform year and year $t+1$ to the post-reform year. When overlapping the density of deductions

for year t and year $t + 1$, b_i^t will be at the same location as b_{i-2}^{t+1} because the standard deduction increases by 2 bins because of the reform. If $b_i^t - b_{i-2}^{t+1} = 0$ then $d_{i-2} = 0$. I then start with the first undistorted bin. In graph I.18 it corresponds to bin 7 in year t:

- $b_7^t - b_5^{t+1} = 0$ implies that $d_5 = 0$. This means that for both year t and $t + 1$, b_5, b_6, b_7 etc. are undistorted as $d_5 = 0$ means that nobody has a burden greater than 500. This also means that I can use b_6^t as the true counterfactual to calculate d_4.

- $b_6^t - b_4^{t+1} = 5\%$ implies that $d_4 = 5\%$. Given that b_6^t is the true density (from the previous bullet point), I can use $\frac{b_4^t}{1-d_4}$ as the *true* counterfactual to calculate d_2.

- $b_5^t - b_3^{t+1} = 15\%$ implies that $d_3 = 15\%$. Given that b_5^t is the true density (from the first bullet point), I can use $\frac{b_3^t}{1-d_3}$ as the *true* counterfactual to calculate d_1.

- To calculate d_2 I need to use b_4^t. But I know from above (second bullet point) that b_4^t is distorted. This implies that the counterfactual density that I need to use to calculate d_2 is $\frac{b_4^t}{1-d_4}$ rather than b_4^t. Hence, $d_2 = \frac{b_4^t}{1-d_4} - b_2^{t+1} = 30\%$.

- Similarly, to calculate d_1 I need to use b_3^t. But I know from the third bullet point that b_3^t is distorted. This implies that the counterfactual density that I need to use to calculate d_3 is $\frac{b_3^t}{1-d_3}$ rather than b_3^t. Hence, $d_1 = \frac{b_3^t}{1-d_3} - b_1^{t+1} = 60\%$.

The distribution of the burden d_i derived using this method allows me to precisely recover the distribution of the burden that I assumed above. This example shows that I am able to recover the true (unobserved) density by using the pre-reform and post-reform densities. The cost distribution d_1, d_2 etc. allows me to calculate the distribution of the burden and recover the true density of deductions in figure 4.

49

E   Tax Reform Act of 1986 and Lagged Responses

Could there be any other exogenous variation altering the distribution of itemized deductions in 1989 affecting my main identification strategy? The majority of tax reforms happened following the TRA'86 and were enacted in 1987. Among those, there were some deduction reforms. Because I am comparing 1987 to 1989, I am implicitly controlling for the Tax Reform Act of 1986 (TRA'86) reforms. But there might be slow adjustments and lagged responses in 1988 or 1989. To rule these out, I consider all the reforms enacted by TRA'86 that could affect the level of deductions and show that it is reasonable to assume that the adjustment is immediate. Because all of the reforms reduced the amount of eligible deductions, they have no lagged response. To see this consider a hypothetical example: assume the charitable donation deduction is capped at $10,000. A taxpayer who was donating $15,000 will now only be able to deduct $10,000. Will the taxpayer reduce her donations? She might reduce them up to $10,000 but there is no reason to expect that she will reduce them any further. What does this imply for the level of deductions? We should observe a drop in deductions to $10,000 in 1987 and then *no further* drop in 1988 or 1989, ruling out any lagged responses. Since I am comparing 1987 to 1989, any reform that caps the amount of deductions should not affect my estimates. The deduction reforms enacted in 1987 are the following (source: IRS):

- Prior to 1987, medical deductions in excess of 5% of the AGI are deductible. In 1987, this threshold is increased to 7.5% of AGI, further limiting the allowable amount of medical deductions. There is no reason to assume that there will be a slow adjustment that spills over into 1988 or 1989 in this case.

- Sales taxes are not deductible anymore. For similar reasons, one should observe a drop in the total deductions in 1987 as sales taxes were a large portion of it but there should be no lagged effect.

- The home mortgage interest deduction is subject to a new limit. The

home mortgage interest deductions for a given year are capped at the value of one's house (plus renovations). Anything in excess of the value of the house have to be deducted as personal interest for which only 65% of the total value can be deducted. First, the IRS estimated that very few taxpayers were affected by this reform since it is very rare that one's home mortgage interest in one given year exceeds the total value of one's house. Second, there is no reason to expect a drop in levels in the subsequent years. If a person is affected by this reform, in 1987 she will be forced to claim less deduction than she was previously claiming.

- Any interest for home mortgages in excess of 1 million dollars is not deductible anymore. Again, there is no reason to expect any lagged effects due to this reform because it caps the amount of deductions.

There are no other reforms affecting directly or indirectly the amount of itemized deductions an individual can qualify for.

F   The 1971 reform

In 1970 taxpayers could claim as a standard deduction the smaller of $6,130 or 10% of their income. In 1971, both thresholds were increased to $8,809 or 13% of income if income is greater than $46,983, and the larger of $6,166 or 13% of income for taxpayers with income smaller than $46,983.

If I were to only look at the density of itemizers above $6,130 in 1970 and compare it to the density of itemizers above $8,809 in 1971, my estimates would be biased because some taxpayers who have deductions greater than $8,809 in 1971 are likely to stop itemizing – not because of hassle costs – but only because their deductions are now smaller than 13% of their income. To control for this, I only consider taxpayers whose deductions exceed 13% of income and $6,166 in 1970. This provides an accurate counterfactual for 1971.

In the 1988 reform I compare the pre-reform year (1987) to the post-reform year (1989). However in this case, the standard deduction is further increased in 1972 making it impossible to compare pre and post-reform years. The 1971 reform estimates are likely to be a lower bound because they do not account for lagged responses.

G   Audit Survey

The survey was carried outside a health food supermarket in Santa Monica, California. The location was chosen to attract as many wealthy individuals as possible to increase the proportion of itemizers. 195 individuals were surveyed of which 114 file their taxes themselves. Of those 95 itemize deductions, which constitutes the final sample. They were asked the following questions:

1. Do you file taxes yourself?

2. Do you itemize deductions or claim the standard deduction?

3. Per year, what do you think the chances of being audited are?

4. Assume the IRS wants to audit you. What is the highest amount you would pay a lawyer that would deal directly with the IRS and prevent you from being audited?

5. What is the annual income of your household? (Brackets of $10,000)

H   Rational Inattention

Could taxpayers forgo large amounts of deductions because they are uncertain of whether their total deductions are larger than the standard deductions threshold?

Most of the deductions are relatively stable from year to year as they mostly consist of items that vary very little such as mortgage payments, real estate taxes or state income taxes. This means that taxpayers should have an accurate signal of their true deductions. In addition, the expenses associated with deductions are an active decision: if deductions increase or decrease by a large percentage, taxpayers are likely to be aware of this change because they caused it.

Therefore, for rational inattention to explain the magnitude of the estimated hassle costs, one would need to assume that taxpayers receive a very noisy signal which is unlikely given that deductions vary little from year to year. I formalize this argument in what follows:

Assume that the taxpayer has a Constant Relative Risk Aversion (CRRA) utility function given by $U(x) = \frac{1}{1-\theta}x^{1-\theta}$ if $\theta \neq 1$ and $U(x) = \log(x)$ if $\theta = 1$.

Denote by τ the after tax amount of deductions the taxpayer can claim (deduction multiplied by marginal tax rate) and by S the after tax amount of the standard deduction. Assume that the taxpayer has beliefs over τ that follow a normal distribution with mean μ and standard deviation σ. Denote by c the cost incurred by the taxpayer to calculate the total amount of deductions τ. The cost is only incurred when she itemizes, not when she claims the standard deduction.

The taxpayer will decide to itemize if the expected benefit from itemizing given her beliefs over τ exceeds the cost of figuring out the level of τ i.e. c. This occurs when the following equation is satisfied:

$$\mathbb{E}\left[\frac{1}{1-\theta}(\tau - c)^{1-\theta}\right] \geq \frac{1}{1-\theta}S^{1-\theta}. \tag{12}$$

This equation does not have a closed form solution, so I use a Taylor expansion of second degree around the mean of $\tau - c$, as follows:

$$\frac{1}{1-\theta}(\mu - c)^{1-\theta} - \frac{1}{2}\theta(\mu - c)^{-1-\theta}\sigma^2 \geq \frac{1}{1-\theta}S^{1-\theta}. \tag{13}$$

And for $\theta = 1$, it is equal to:

$$\log(\mu - c) - \frac{\sigma^2}{2(\mu - c)^2} \geq \log(S). \tag{14}$$

The first term in equation 14 is the expected benefit that the taxpayer derives from itemizing. The second term is a correction for the risk aversion of the taxpayer: she will itemize deductions if the benefit of itemizing corrected for her risk aversion is greater than the benefit she derives from itemizing. Holt and Laury (2002) find a θ that ranges between -0.95 and 1.37. I assume here that $\theta = 1$ but also consider $0 < \theta \leq 2^{47}$ in table J.11. I fix the standard

---

[47]Negative values of θ are not considered because they imply risk lovingness and would trivially reject rational inattention.

deduction at \$10,000 for joint filers. The cost estimated by the IRS of the time required to itemize deductions is $c = 149$. I can calculate a lower bound on the standard deviation of the taxpayer's beliefs over τ (σ). Using these parameters, I find that for rational inattention to explain the magnitude of the forgone benefits, the standard deviation of after tax deductions σ has to be greater than \$1,814 (which corresponds to \$6,479 worth of deductions with a 28% marginal tax rate). This means that the taxpayer has a range of uncertainty of deductions of more than \$6,479. This implies very high uncertainty in the beliefs of the benefits that the taxpayer can save from itemizing which is unlikely given that deductions are relatively stable from year to year as they are mostly constituted of mortgage payments and state taxes and are the results of active decisions. If a taxpayer's total deductions were to increase or decrease dramatically, she would most likely know about it because it would be due to for example to large income variations, the take up of a mortgage etc. which are salient.

If I assume a standard deviation of $\sigma = 200$ – which corresponds to a standard deviation of deductions of \$714 – then rational inattention with $\theta = 1$ predicts that taxpayers would claim the standard deduction up to total deductions of \$10,557 and forgo an average of \$557 worth of deductions, i.e. 156 of after tax dollars given a cost c=\$149. With reasonable parameters, rational inattention predicts that taxpayers will forgo an additional \$7 in excess of the cost of \$149.

I   Time Inconsistency: Model

I assume that the cost of record keeping continuously increases for every day that the receipt is not archived as soon as it is received. When the taxpayer is issued a receipt for a charitable donation and fails to archive it, the cost of keeping track of this receipt increases continuously because it is more likely to be lost or it could take more time to look for it. The rational taxpayer archives the receipt as soon it is issued. The naive present-biased taxpayer plans on archiving the receipt but fails to do so, leading to high record keeping costs.

Assume for simplicity that the taxpayer only needs to itemize one deduction

for example for a charitable contribution she made. The taxpayer is facing two distinct costs when considering the decision to itemize deductions. The first one is that of record keeping, denoted here by c. The second one is filling out Schedule A itself which is denoted by k.

If the taxpayer succeeds in performing the two tasks she receives a one time benefit b in the subsequent period. Once the taxpayer gets the receipt for her charitable contribution, she can decide to archive it immediately by incurring a cost c or archive it later and incur a larger cost $c(1+r)$ next period where r is the rate at which the cost of record keeping grows if the receipt is not archived.

δ is the time-discount factor, β the present-bias parameter, t the period in which the record keeping is performed and Schedule A is filled out and $(t+1)$ the period in benefit b is received.

In what follows, I use two definitions:

Definition 1: For given β, δ, c, k, $(1+r)$ and t a task is said to be β-worthwhile if $-c(1+r)^t - k + \beta b > 0$.

Similarly:

Definition 2 For given δ, c, k, $(1+r)$, and t a task is said to be δ-worthwhile if $-c(1+r)^t - k + \delta b > 0$.

The rational taxpayer has a standard utility function where per-period utility is discounted by δ in the future.

The decision to itemize or claim the standard deduction for the rational taxpayer can be written as follows:

$$\max_t \delta^t(-c(1+r)^t - k + \delta b),$$

conditional on itemizing being δ-worthwhile.

Cost c is incurred as soon as the taxpayer starts the record keeping. If she waits an additional t periods before archiving the receipt, the cost of record keeping is multiplied by $(1+r)$ for every additional period i.e. $(1+r)^t$ overall. Therefore, to minimize the cost of record keeping, the rational taxpayer will choose $t = 0$, this means that she will archive the receipt as soon as it is

received and will incur a record keeping cost of c rather than $c(1+r)^t$.

The taxpayer is left with choosing t such that:

$$\max_{t} \delta^t(-c(1+r)^t - k + \delta b)$$

Assume the taxpayer is contemplating the decision to perform the record keeping task in the first period yielding utility: $-c - k + \delta b$. She will only perform it if $-c - k + \delta b > 0$. And if she waits an additional period she will receive $\delta(-c(1+r) - k + \delta b)$, which is smaller than the utility she would have enjoyed if the task had been performed in the first period. This means that the rational taxpayer will either archive the receipt immediately or never archive it because she does not plan on itemizing her deductions.

The naive present biased taxpayer can perform the record keeping in period t or can wait and perform it in period $t + 1$. She will prefer performing it in period $t + 1$ if the following inequality is satisfied:

$$-c(1+r)^t - k + \beta b < \beta[-c(1+r)^{t+1} - k + b].$$

This inequality simplifies to:

$$-c(1+r)^t - k < \beta[-c(1+r)^{t+1} - k]. \tag{15}$$

A sufficient condition for equation 15 to hold is:

$$(1+r)\beta < 1. \tag{16}$$

Intuitively, for the naive present-biased taxpayer to procrastinate on archiving her receipt, it is sufficient that the rate at which the record keeping cost increases be smaller than the rate at which she discounts the future.

Provided that condition 15 holds in period $t = 0$, it will also hold in any subsequent period $t > 0$ i.e. if itemizing is worthwhile but not performed in the very first period, the taxpayer will procrastinate until she reaches the deadline.

Testable Prediction 1: Naive present-biased taxpayers will file their returns at the deadline of April 15th when condition 15 holds.

Testable Prediction 2: The cost of record keeping for naive present-biased taxpayers is greater than for rational ones. This predicts that taxpayers who file close to the deadline are likely to forgo more deductions.

I.0.1 Week of Filing Variable

The SOI files contain a variable that indicates the week in which a return is processed by the IRS. Slemrod et al. (1997) have access to the internal IRS files that record the filing date and compare it to the processing date from the SOI files. They find that the order in which returns are processed matches the order in which they are filed. Knowing the order is sufficient for my purposes because what I am interested in is comparing taxpayers who file close to the deadline to those who file earlier. I can therefore use the processing time variable to identify late filers and verify the predictions of the naive present bias model. The IRS promises that returns are processed within 6 weeks. This constraint is likely to be binding for returns that are filed close to the deadline given that a lot of returns are processed at the time. Therefore, I assume that the processing time has a lag of 6 weeks.

I restrict the sample used to generate this graph to taxpayers who are owed refunds by the IRS and who do not have to file any other schedule but Schedule A. This allows me to rule out taxpayers who rationally delay filing to save on interest on the amount they owe to the IRS and taxpayers who cannot file early because others schedules sometimes require additional paperwork that only becomes available later in the year.

I.0.2 Bound on Record Keeping Costs

At any point in time, taxpayers have access to tax preparers. For a given sum of money, the taxpayer can get a tax specialist to fill out her 1040 and Schedule A forms. However, the tax preparer cannot perform the record keeping for her. The tax preparer fee provides an upper bound on the cost of filling out Schedule A for the taxpayer: if the cost of filling out Schedule A is larger than the fee, she can go to a tax preparer.

I can identify this fee in the dataset: individuals who itemize their deductions are allowed to deduct the tax preparer fee from their income. The average tax preparer fee for individuals who file the 1040 and Schedule A but

not Schedule B, C, D etc. is \$220. This is the fee for filling out both the 1040 form and Schedule A, submitting the documents and helping with audits if the need arises. This means that \$220 is a generous upper bound. If the burden of itemizing deductions is driven by the cost of filling out Schedule A then taxpayers have the outside option of paying someone to perform this task and - for some of them - save large sums of money. This suggests that any cost in excess of \$220 should be attributed to record keeping. Since the estimated cost is equal to \$644, the record keeping cost accounts for more than 64% of the burden of itemizing.[48] This is consistent with taxpayers having large record keeping costs.

---

[48]This is consistent with Slemrod and Bakija (2008) who use survey evidence to argue that most of the burden of filing taxes is due to record keeping.

Figure I.10: Missing Mass In the Neighborhood of the Standard Deduction 1998-2003
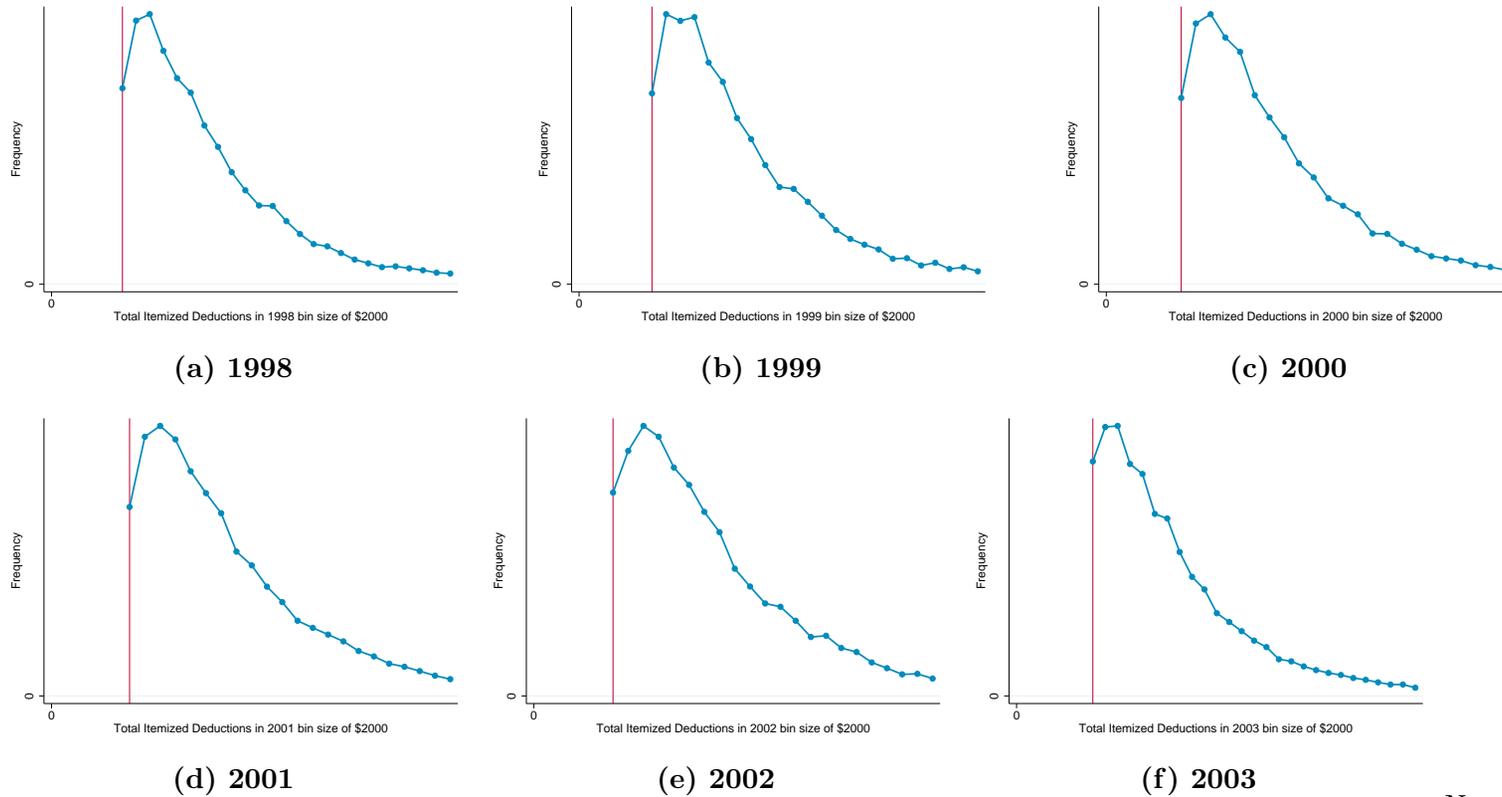
(a) 1998

(b) 1999

(c) 2000

(d) 2001

(e) 2002

(f) 2003

Notes: The figures above plot the density of deductions for itemizers filing jointly. The bin size is $2,000 and the vertical line represents the standard deduction threshold for each year. Notice the missing mass in the neighborhood of the standard deduction threshold.

Figure I.11: Missing Mass In the Neighborhood of the Standard Deduction 1992–1997

(a) 1992   (b) 1993   (c) 1994

(d) 1995   (e) 1996   (f) 1997

Notes: The figures above plot the density of deductions for itemizers filing jointly. The bin size is $2,000 and the vertical line represents the standard deduction threshold for each year. Notice the missing mass in the neighborhood of the standard deduction threshold.

Figure I.12: Missing Mass In the Neighborhood of the Standard Deduction 1986-1991

(a) 1986

(b) 1987

(c) 1988

(d) 1989

(e) 1990

(f) 1991

Notes: The figures above plot the density of deductions for itemizers filing jointly. The bin size is $2,000 and the vertical line represents the standard deduction threshold for each year. Notice the missing mass in the neighborhood of the standard deduction threshold.

Figure I.13: Missing Mass In the Neighborhood of the Standard Deduction 1980-1985
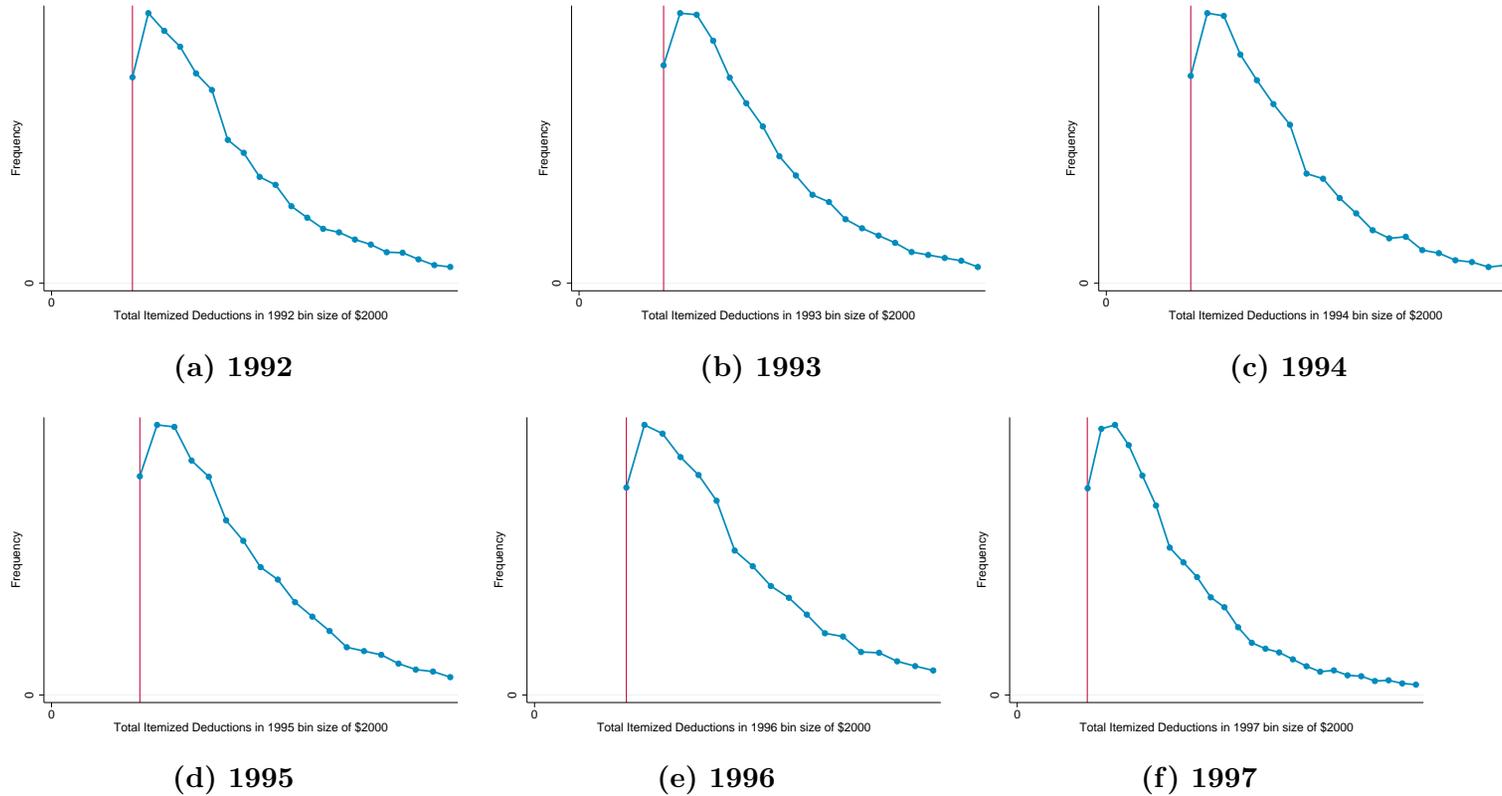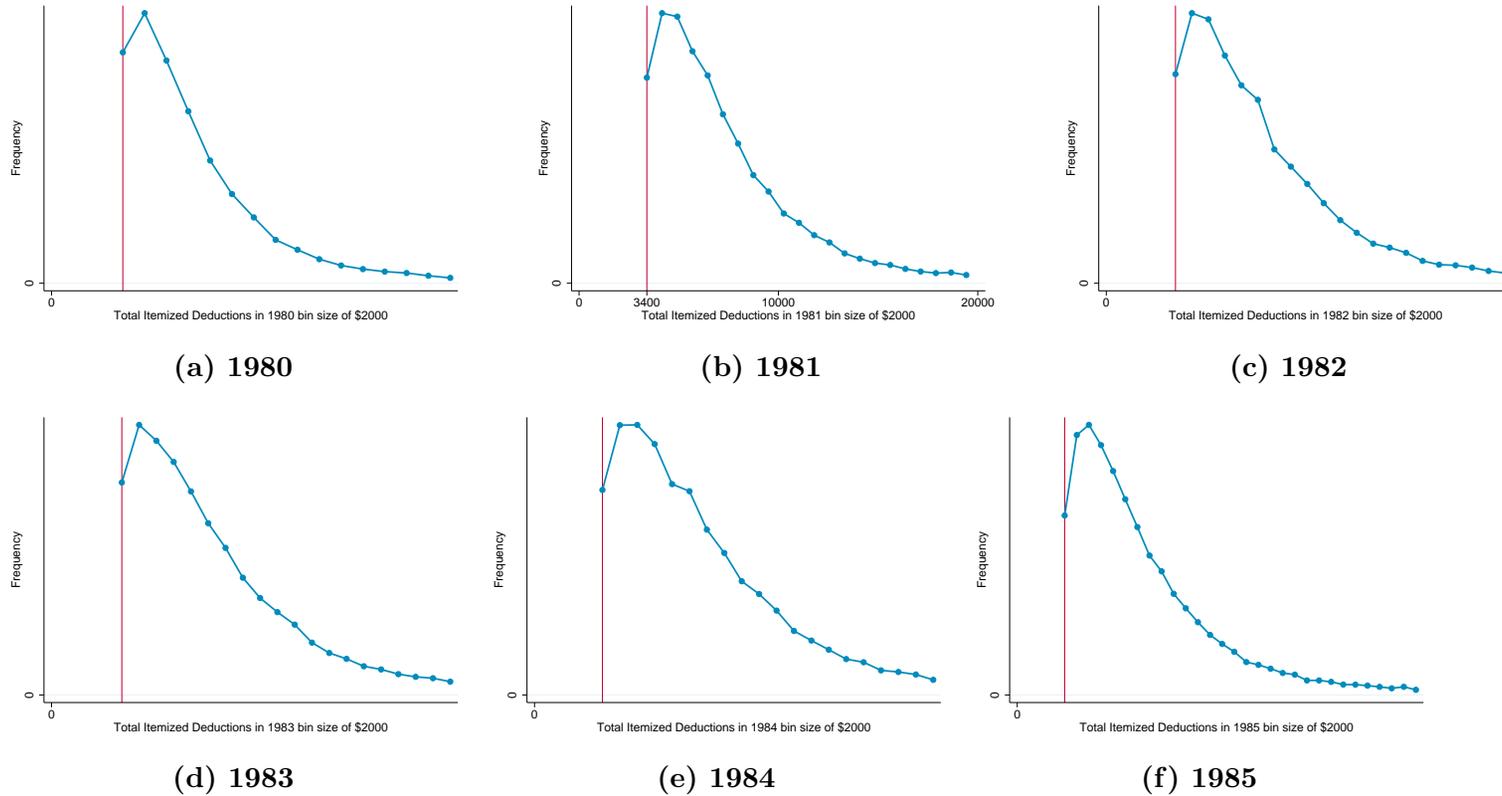

(a) 1980
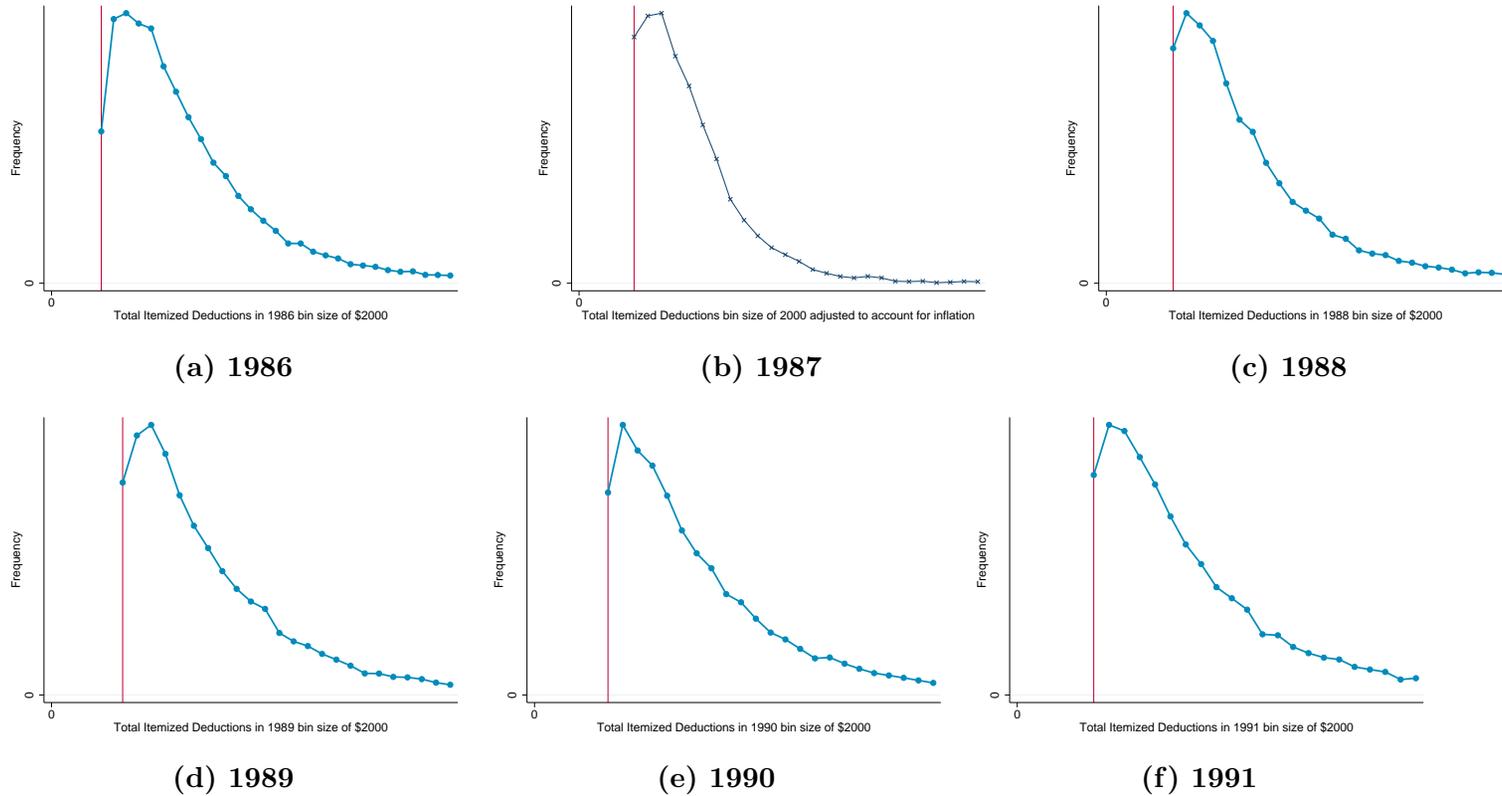
(b) 1981

(c) 1982

(d) 1983

(e) 1984

(f) 1985

Notes: The figures above plot the density of deductions for itemizers filing jointly. The bin size is $2,000 and the vertical line represents the standard deduction threshold for each year. Notice the missing mass in the neighborhood of the standard deduction threshold.

Figure I.14: Missing Mass In the Neighborhood of the Standard Deduction (Single Filers)

(a) 1980

(b) 1985

(c) 1990

(d) 1995

(e) 2000

(f) 2005
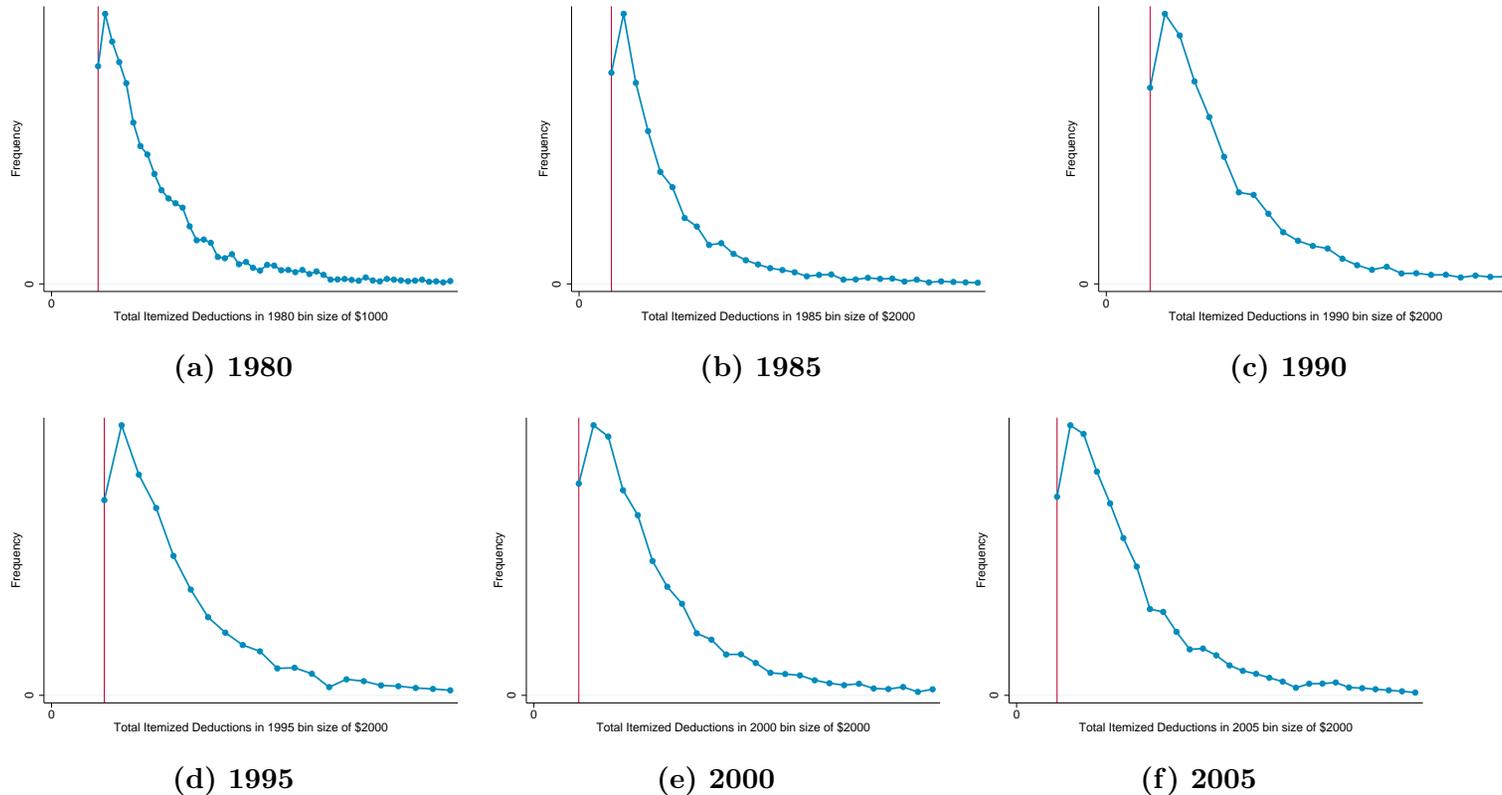
Notes:
The figures above plot the density of deductions for single filers who itemize deductions. The bin size is $2,000 and the vertical line represents the standard deduction threshold for each year. Notice the missing mass in the neighborhood of the standard deduction threshold.

Figure I.15: Placebo Test: Overlapping Densities In Years With No Reforms

(a) 1990-1992

(b) 1991-1993

(c) 1992-1994

(d) 1993-1995

Notes: The figures above plot the density of deductions for itemizers filing jointly in years with no reforms of the standard deduction. Notice that there is no missing mass in the neighborhood of the standard deduction.

Figure I.16: Placebo Test: Overlapping Densities In Years With No Reforms
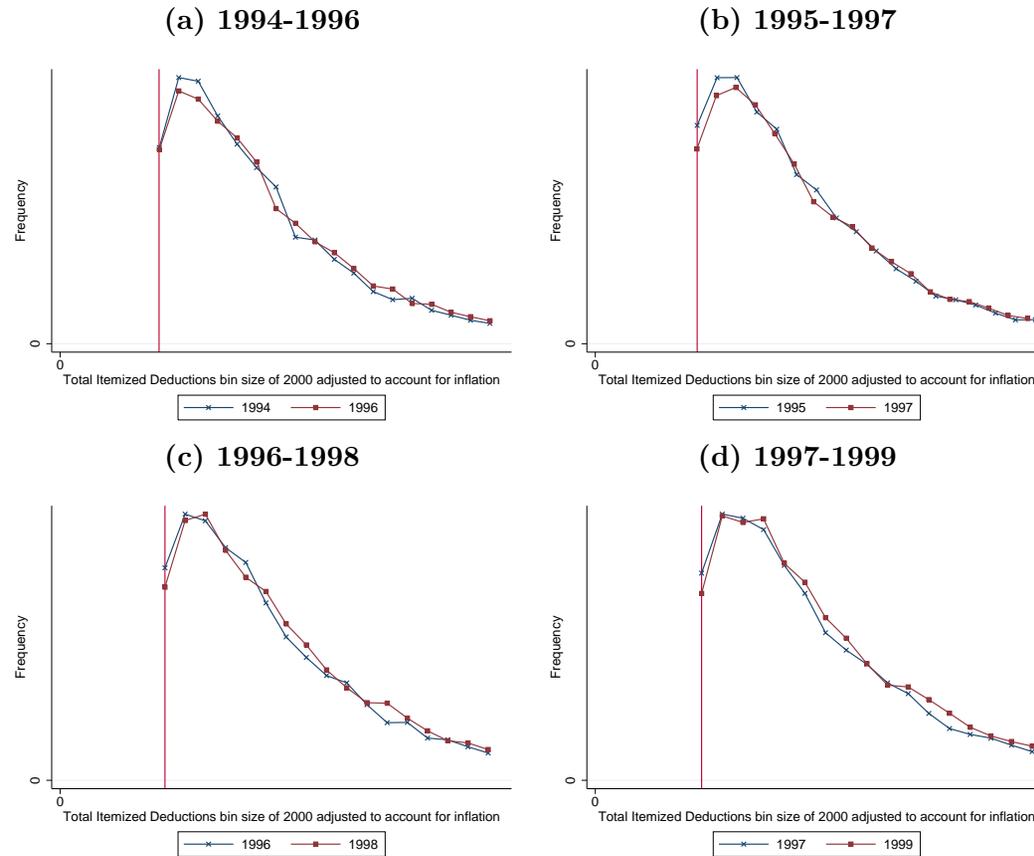
(a) 1994-1996

(b) 1995-1997

(c) 1996-1998

(d) 1997-1999

Notes: The figures above plot the density of deductions for itemizers filing jointly in years with no reforms of the standard deduction. Notice that there is no missing mass in the neighborhood of the standard deduction.

Figure I.17: Placebo Test: Overlapping Densities In Years With No Reforms
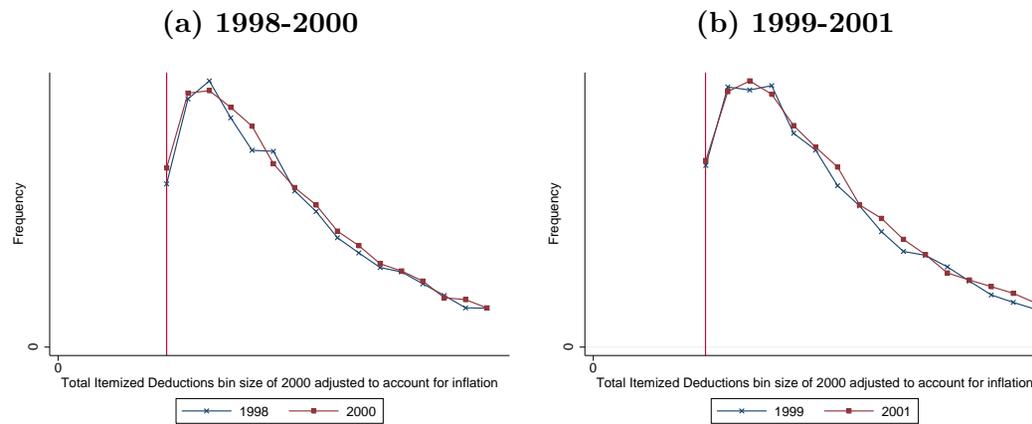
(a) 1998-2000 **(b) 1999-2001**

Notes: The figures above plot the density of deductions for itemizers filing jointly in years with no reforms of the standard deduction. Notice that there is no missing mass in the neighborhood of the standard deduction.

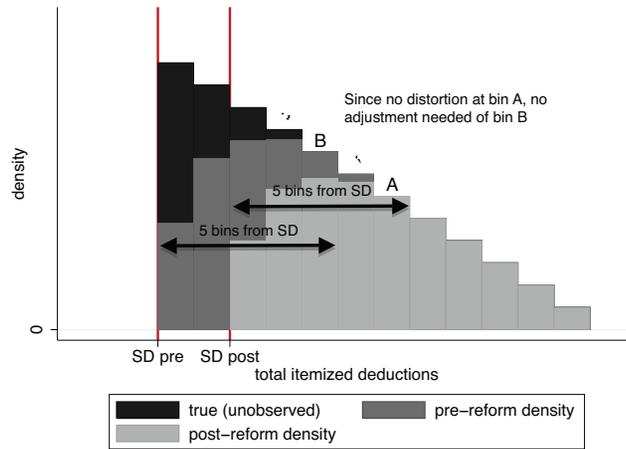## Figure I.18: Reconstructing the Counterfactual Density

(a) No Distortion

(b) Distortion
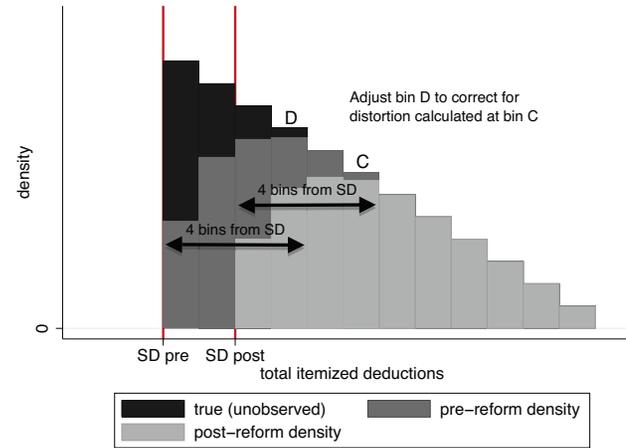
Notes: These graphs illustrate the method that I use to reconstruct the counterfactual density. The darkest histograms correspond to the true density of deductions when assuming that there is no cost. The next shade corresponds to the pre-reform year and the lightest one to the post-reform year. The vertical lines show the standard deduction threshold. In figure (a), I consider the first bin – bin A – for which the pre-reform and post-reform years overlap. There is no distortion for this bin because the two densities are overlapping. This means that 5 bins away from the pre-reform standard deduction – bin B – there should be no distortion. Which implies that 3 bins away from the post-reform standard deduction, the pre-reform density is the true density. On the other hand, in figure (b), when looking 4 bins away from the post-reform standard deduction (bin C), I find a distortion. This implies in turn that 4 bins away from the pre-reform density (bin D), there should be a distortion of equal proportion to the one that I calculated 4 bins away from the post-reform standard deduction. I adjust the density that is 2 bins away from the post-reform standard deduction – bin D – by this amount and repeat this process for all bins thereafter. This adjustment allows me to recover the true (unobserved) density.

Figure I.19: Different Scenarios Below the Standard Deduction

(a) Increasing: Impossible

65% of taxpayers claim
Standard Deduction

Total Itemized Deductions in 2006 bin size of $2000

Hypothetical   Observed

(b) Double Peaked: Unlikely

65% of taxpayers claim
Standard Deduction

Total Itemized Deductions in 2006 bin size of $2000

Hypothetical   Observed

(c) Missing Mass

65% of taxpayers claim
Standard Deduction

Total Itemized Deductions in 2006 bin size of $2000

Hypothetical   Missing
Observed

Notes: The graphs above plot the different scenarios that could be happening below the standard deduction. Graph (a) assumes that the density is strictly increasing, which is impossible given that 65% of taxpayers claim the standard deduction. This scenario would fail to account for most of the population of taxpayers. Graph (b) accounts for most of the population and is continuous at the standard deduction but the density is double peaked. This is possible but unlikely given that densities are usually single peaked. This however does not rule out densities that are double-peaked *because of the standard deduction*. Graph (c) assumes that there is a discontinuity at the standard deduction threshold because of hassle costs creating a missing mass.

Figure I.20: No Behavioral Response For Personal Interest Deduction
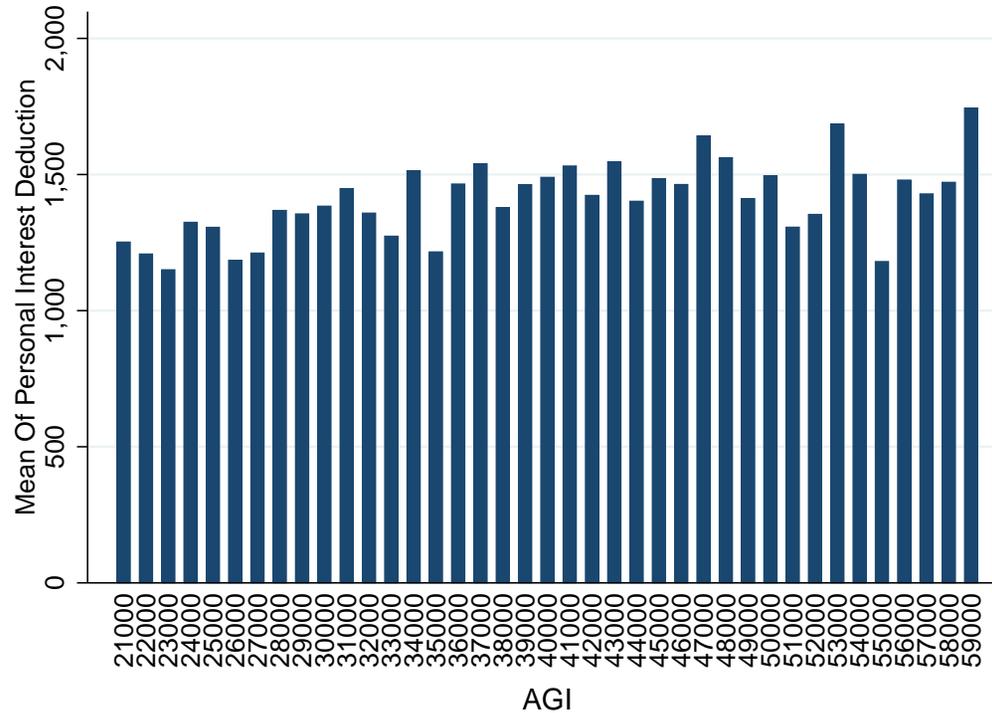
Notes: This figure plots the average personal interest deduction claimed by income bins of $1000 in 1989. Below $30,950, the marginal tax rate is 15% for married filing jointly and above it is equal to 28%. If taxpayers were responding to tax incentives when claiming the personal interest deduction, one would observe a discontinuity at the MTR threshold. None is observed here.

J   APPENDIX TABLES

Table J.4: Standard Deduction By Year For Joint Filers

Year	Standard deduction	S.D. in 2014 $	Growth Rate	Year	Standard deduction	S.D. in 2014 $	Growth Rate
1961	1000	7968	0.00%	1984	3400	7796	0.00%
1962	1000	7889	0.00%	1985	3540	7838	4.12%
1963	1000	7786	0.00%	1986	3670	7978	3.67%
1964	1000	7686	0.00%	1987	3760	7886	2.45%
1965	1000	7564	0.00%	**1988**	**5000**	**10070**	**32.98%**
1966	1000	7353	0.00%	1989	5200	9991	4.00%
1967	1000	7133	0.00%	1990	5450	9935	4.81%
1968	1000	6846	0.00%	1991	5700	9971	4.59%
1969	1000	6492	0.00%	1992	6000	10189	5.26%
1970	1000	6140	0.00%	1993	6200	10223	3.33%
1971	**1500**	**8824**	**50.00%**	1994	6350	10208	2.42%
1972	2000	11400	33.33%	1995	6550	10240	3.15%
1973	2000	10732	0.00%	1996	6700	10174	2.29%
1974	2000	9665	0.00%	1997	6900	10243	2.99%
1975	2600	11514	0.30%	1998	7100	10378	2.90%
1976	2800	11724	0.08%	1999	7200	10293	1.41%
1977	3200	12580	0.14%	2000	7350	10169	2.08%
1978	3200	11693	0.00%	2001	7600	10515	3.40%
1979	3400	11158	0.06%	2002	7850	10560	3.29%
1980	3400	9831	0.00%	2003	9500	12301	21.02%
1981	3400	8911	0.00%	2004	9700	12234	2.11%
1982	3400	8394	0.00%	2005	10000	12199	3.09%
1983	3400	8133	0.00%	2006	10300	12173	3.00%

Notes: The table shows the standard deduction amounts from 1961 to 2006 for joint filers and its growth rate. The years that I use to identify the burden of itemizing deductions are in bold.

Table J.5: Standard Errors of the Difference Between the 1987 and 1989 Densities (figure 2a)

Bin	Deduction Range	Difference	Standard Errors	z-stat
1	[9991, 11991]	0.00311***	0.00047	6.55
2	(11991, 13991]	0.00190***	0.00044	3.47
3	(13991, 15991]	0.00000	0.00040	0.02
4	(15991, 17991]	-0.00047	0.00041	-1.13
5	(17991, 19991]	0.00022	0.00038	0.59
6	(19991, 21991]	-0.00010	0.00033	-0.31
7	(21991, 23991]	-0.00041	0.00028	-1.45
8	(23991, 25991]	-0.00042	0.00025	-1.67
9	(25991, 27991]	-0.00032	0.00020	-1.60
10	(27991, 29991]	-0.00042**	0.00018	-2.24

Table J.6: Standard Errors of the Difference Between the 1970 and 1971 Densities (figure 2b)

Bin	Deduction Range	Difference	Standard Errors	z-stat
1	[6140, 9140]	0.00373***	0.00102	3.64
2	(9140, 12140]	0.00288***	0.00090	3.20
3	(12140, 15140]	0.00307***	0.00074	4.11
4	(15140, 18140]	0.00083*	0.00046	1.81
5	(18140, 21140]	0.00019	0.00037	0.54
6	(21140, 24140]	0.00039	0.00027	1.45
7	(24140, 27140]	-0.00025	0.00018	-1.41
8	(27140, 30140]	-0.00001	0.00015	-0.09
9	(30140, 33140]	-0.00007	0.00011	-0.63
10	(33140, 36140]	-0.00010	0.00010	-0.94

Notes: These tables show the bootstrapped standard errors for the difference between bins in 1987 and 1989 and 1970 and 1971 for taxpayers with deductions below $30,000. Notice that only the first bins are statistically significantly different at the 99% level: the first two for the 1988 reform and the first three for the 1971 reform. * denotes significance at the 10% level, ** at the 5% level and *** at the 1% level. I use 100 replications for the bootstrap estimation.

Table J.7: Standard Errors of the Difference Between the Density of Electronic Filers v.s. Paper Filers (Figure 6(b))

Bin	Deduction Range	Difference	Standard Errors	z-stat
1	[0, 2000)	7.08e-06***	1.44e-06	4.92
2	[2000, 4000)	3.02e-06*	1.55e-06	1.95
3	[4000, 6000)	5.91e-06***	1.39e-06	4.25
4	[6000, 8000)	3.44e-06**	1.54e-06	2.23
5	[8000, 10000)	5.10e-06***	1.49e-06	3.42
6	[10000, 12000)	1.47e-06	1.41e-06	1.04
7	[12000, 14000)	2.37e-07	1.42e-06	0.17
8	[14000, 16000)	-1.73e-06	1.18e-06	-1.47
9	[16000, 18000)	-1.93e-07	1.04e-06	-0.19
10	[20000, 22000)	-1.88e-06*	1.03e-06	-1.82

Table J.8: Standard Errors of the Difference Between the Density of March vs. April Filers (Figure 8c)

Bin	Deduction Range	Difference	Standard Errors	z-stat
1	[0, 2000)	-.000012***	2.87e-06	-4.17
2	[2000, 4000)	-.0000114***	2.83e-06	-4.02
3	[4000, 6000)	-5.08e-06	3.09e-06	-1.64
4	[6000, 8000)	-.0000141***	3.39e-06	-4.17
5	[8000, 10000)	-9.22e-06***	3.41e-06	-2.71
6	[10000, 12000)	-8.12e-06**	3.42e-06	-2.51
7	[12000, 14000)	-4.21e-06	1.42e-06	-1.23
8	[14000, 16000)	-5.94e-06	3.05e-06	-1.94
9	[16000, 18000)	-4.86e-07*	3.36e-06	-0.14
10	[20000, 22000)	-5.57e-06**	2.81e-06	-1.98

Notes: These two tables show the bootstrapped standard errors for the difference between the density of itemizers who use electronic filing versus paper filing in the first table and the difference between April and March Filers in the second table. * denotes significance at the 10% level, ** at the 5% level and *** at the 1% level.

Table J.9: Survey Based Estimates of the Hassle Costs of Taxation in the US

Article	Methodology	Cost of Itemizing Deductions	Aggregate Costs of Filing Taxes
Wicks (1965) and Wicks and Killworth (1967)	Survey of Montana residents	Not reported	32% of state and 11.5% of federal tax revenue
Slemrod and Sorum (1984)	Survey of 2000 Minnesota residents	Not reported	5% to 7% of total tax revenue
Little (1988), Commissionned by IRS	Two separate surveys of 750 and 6200 taxpayers	Not reported	1.59 billion hours
Slemrod (1989)	Estimate structural model based on survey of 2000 Minnesota residents	3.2 to 3.5 hours	Not reported
Blumenthal and Slemrod (1992)	Survey of 2000 Minnesota households in 1990	9 hours	85 billion dollars
Guyton et al. (2003)	Survey and ITBM* simulations	9.9 hours	18.7 billion hours dollars

Notes: This table reports the results of several research article documenting the cost of tax filing using survey evidence.
*ITBM stands for the Individual Tax Burden Model.

Table J.10: Articles Documenting Low Take-Up Rates/Large Forgone Benefits

Article	Setting	Forgone Benefits
Steuerle et al. (1978)	Tax Benefits/Income Averaging	$666
Blank and Card (1991)	Unemployment Insurance Benefits	Take-up rate of less than 30% of eligible unemployed individuals
Madrian and Shea (2001)	Retirement Savings	50% match of retirement savings up to 6% of contributions
Sydnor (2010)	Home Insurance	Five times the insurance premium
Bhargava and Manoli (2015)	Taxes	Earned Income Tax Credit Benefits
Handel (2013)	Health Insurance	$2,032 per year
Keys et al. (2014)	Mortgage Refinancing	Present discounted cost of $11,500

Table J.11: Calibration of Rational Inattention Model

	Precision of Beliefs About Level of Savings (σ)							
	10	50	100	200	500	1000	2000	3000
CRRA coefficient								
0.1	149	149	149	150	154	177	219	301
0.25	149	149	149	151	160	193	316	501
0.5	149	149	150	153	171	235	461	774
0.8	149	149	150	154	184	283	611	1029
1	149	150	151	156	193	313	696	1164
1.1	149	150	151	157	197	328	735	1223
1.25	149	150	151	158	203	349	789	1302
1.5	149	150	152	160	213	382	868	1411
1.8	149	150	152	162	225	419	948	1513
2	149	150	153	163	233	442	993	1566

Notes: This table shows the results of a calibration of the Rational Inattention model derived in section 5.4. Rational inattention cannot explain the magnitude of the forgone benefits unless one assumes that the standard deviation of the savings is greater than $2000, which implies a standard deviation of deductions of $7,143. This corresponds to a 95% confidence interval of deductions of $\pm 14,000$, implying that a taxpayer with total deductions of $12,000 needs a 95% confidence interval equal to $[-2,000, 26,000]$ in order to forgo more than $600. Such high uncertainty is extremely unlikely given that deductions are stable and changes are usually due to active decisions on the part of the taxpayer (increase in income, take up of home mortgage etc.).