How Taxing Is Tax Filing? Using Revealed Preferences to Estimate Compliance Costs.

Youssef Benzarti*
UCSB and NBER

Abstract

This paper uses a quasi-experimental design to estimate the cost of filing taxes. Using US tax returns, I observe how taxpayers choose between itemizing deductions and claiming the standard deduction. Taxpayers forgo large tax savings to avoid compliance costs, which provides a revealed preference estimate of such costs. I show that costs increase with income, consistent with an opportunity cost of time explanation. These estimates suggest substantial costs of filing federal income taxes, significantly larger than previously estimated using surveys.
JEL Classification: H24, H31, H83
Keywords: Public Finance, Taxes, Compliance Costs.

*UCSB and NBER (benzarti@ucsb.edu). I thank Alan Auerbach, Stefano DellaVigna, Matthew Rabin and Emmanuel Saez for invaluable advice, guidance, and encouragement throughout this project. I benefited from discussions with Miguel Almunia, Pierre Bachas, Dan Benjamin, Kimberly Clausing, Erik Eyster, Alex Gelber, Daniel Gross, Ben Handel, Nathaniel Hendren, Hilary Hoynes, Emiliano Huet-Vaughn, Damon Jones, Louis Kaplow, Marc Kaufmann, Henrik Kleven, Wojciech Kopczuk, Laurence Kotlikoff, David Laibson, Etienne Lehmann, Attila Lindner, Takeshi Murooka, Michaela Pagel, Colin Raymond, Alex Rees-Jones, Antonio Rosato, Jesse Rothstein, Josh Schwartzstein, Dan Silverman, Joel Slemrod, Charles Sprenger, Justin Sydnor, Alisa Tazhitdinova, Danny Yagan and seminar participants at the Tax Systems Conference (Oxford and U. of Michigan), Paris II, CREST, Aarhus University, SITE Psychology and Economics, the US Department of Treasury, Princeton, Harvard KSG, UCLA, McMaster, U. Mass Amherst and U. of Toronto. Funding from the Robert D. Burch Center for Tax Policy and Public Finance is gratefully acknowledged.

Income taxes represent the largest source of tax revenue in the United States. Today, 8.8% of GDP is transferred from individuals to the Federal Government through income taxes. While an extensive literature documents the efficiency costs of taxation, we know less about the costs of collecting taxes. But every year, 140 million taxpayers spend numerous hours gathering receipts and statements, filling out various tax schedules and forms and submitting them to the Internal Revenue Service (IRS). A large literature documents that individuals frequently leave "money on the table" in other domains because of transactional costs, which suggests that the compliance costs of taxation are likely to be very large.[1]

This paper provides the first estimate of this cost using quasi-experimental methods. I exploit the fact that taxpayers can choose between itemizing their tax deductions or claiming the standard deduction. Itemizing deductions requires some effort cost but can provide significant tax savings. Claiming the standard deduction saves time and effort but results in more taxes due.

With compliance costs, itemizing is beneficial only if it reduces the tax bill by more than the cost of itemizing. This implies that if compliance costs are non-zero, some taxpayers will claim the standard deduction, even though the sum of their deductions exceeds the amount of the standard deduction. The main identification challenge is to differentiate individuals who choose not to itemize because of compliance costs from those who claim the standard deduction because their total deductions are smaller than the standard deduction. 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. To estimate the cost of itemizing, I proceed in the following way. If individuals forgo tax benefits because of compliance costs, there should be a missing mass in the density of deductions just above the standard deduction threshold. I test this hypothesis by graphing the density of deductions for the years 1980 to 2006 using a stratified random sample of US tax returns, weighted to be representative of the population of itemizers. The density of itemizers exhibits a missing mass just above the standard deduction.

To confirm that this shape is due to taxpayers responding to the standard

[1] See, for example, Currie (2006), Bertrand et al. (2006) and, more recently, Bhargava and Manoli (2015).

deduction, I turn to a quasi-experimental design. Following a large increase in the standard deduction, I observe a drop in the mass of itemizers just above the post-reform standard deduction threshold. The post-reform density is systematically lower than the pre-reform density just above 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.[2] I use the missing mass to construct the distribution of forgone benefits. While related to bunching estimators, my approach is different: bunching estimators rely on one cross-section of data, while my approach compares two cross-sections before and after a reform.[3]

I find that the cost of itemizing ranges from 0.6% to 0.8% of adjusted gross income (AGI), i.e., the disutility derived from itemizing is equivalent to working 10 to 15 hours, which is substantially larger than previous estimates. 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 use the procedure outlined above on each subgroup. The results show an increasing relationship between forgone benefits and income - while controlling for marginal tax rates - consistent with the hypothesis that tax filing imposes a higher cost on richer individuals because they have a higher marginal value of time.

The missing mass just above the standard deduction is consistent with taxpayers forgoing benefits to avoid the cost of itemizing. However, there are three alternative explanations for the missing mass. The first is that the standard deduction acts as 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 marginal tax rates fixed. The fact that forgone benefits increase with income - while controlling for marginal tax rates - supports the compliance

[2] 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.

[3] My approach is related in spirit to a difference-in-differences approach where the treatment group includes taxpayers just above the standard deduction threshold and the control group are individuals far above the standard deduction threshold.

costs explanation. A second alternative explanation is that some taxpayers mistakenly believe that IRS audits are more likely when itemizing and, thus, 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 about audit probabilities and audit costs. The perceived expected cost of audits would explain, at most, one fifth of the cost. A third alternative explanation is that the uncertainty that taxpayers face over the amount of deductions they can claim drives them to not itemize. The cost that I estimate is based on taxpayers who itemized the year before the reform. And deductions tend to nominally increase over time, which implies that taxpayers should have a precise belief over the lower bound of the tax savings they can derive from itemizing, therefore contradicting the idea that uncertainty prevents them from itemizing.

While the large magnitude of the costs could be explained by high levels of aversion to filing taxes, I gather empirical evidence suggesting that taxpayers procrastinate on filing their taxes, which leads them to incur high costs. Procrastination provides 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.

This paper is related to several lines of prior work. It is the first and only paper to provide estimates of the cost of filing taxes using a quasi-experimental design. The most closely related paper is Pitt and Slemrod (1989): estimating the cost of itemizing using a censored model with unobserved censoring thresholds using maximum likelihood, they find a smaller cost of itemizing of \$107, which is equivalent to 0.12% of AGI and is 5 to 7 times smaller than my cost estimates. They use estimators from Gronau (1973) and Nelson (1977) to address the fact that the distribution below the standard deduction is unobservable. While our approaches are related, my method is able to provide a reduced-form demonstration of the existence of compliance costs without relying on a structural model. I discuss their approach and some of their assumptions in more detail in Appendix Section A.[4] There is also a literature that uses survey evidence to estimate compliance

[4]This paper is also related to a literature that estimates the effect of tax simplicity on

costs. Although informative of the time spent filing taxes, it does not capture the preferences of taxpayers and, in particular, any disutility of filing taxes.[5]

Finally, this paper adds to a long tradition in public economics emphasizing the need to screen out applicants using ordeal mechanisms (Nichols and Zeckhauser (1982)). If poorer individuals value their time less, then such policies can successfully target them by screening out richer individuals. My results lend support to this assumption because richer individuals tend to forgo more benefits than poorer ones. However, given how large costs are, such policies could be screening out too many individuals.

1 Data and Institutional Background

1.1 Institutional Background

Taxpayers can reduce their taxable income by choosing to itemize their deductions or claiming the standard deduction. The decision to itemize deductions requires comparing two numbers: the sum of itemized deductions and the standard deduction amount. Itemizing, however, is more costly: taxpayers need to keep a record of all the expenses they want to deduct and file Schedule A. Approximately two thirds of the population claim the standard deduction. The standard deduction varies by filing status and by whether the person is blind or older than 65.

Taxpayers claim four main deductions: (1) state and local income taxes, (2) mortgage interest, (3) real estate taxes and (4) charitable donations. They represent 17%, 40%, 14% and 12% of total deductions, respectively.

Schedule A is a one-page schedule and is relatively easy to fill out as it only requires copying numbers from receipts and statements and then summing them up. Record-keeping is more time consuming since it requires archiving various records of expenses.

1.2 Data

I use the the Statistics of Income (SOI) dataset. It consists of repeated annual cross-sections of individual tax returns. The number of observation per year

individual and firm behavior. See for example Abeler and Jäger (2015) and more recently Tazhitdinova (2016), Harju et al. (2019), Aghion et al. (2017) and Tazhitdinova (2018).

[5]See Appendix Table I.7 for a list of papers using survey-based estimates.

ranges from 80,000 to 200,000. The repeated cross-sections are stratified random samples in which the randomization occurs over the Social Security number. The data oversamples high-income taxpayers and taxpayers with business income, but the IRS provides weights that are used in this analysis to produce estimates representative of the total US population. All of the analysis in this paper uses these weights to reflect population averages. No particular sample restrictions are made, except for properly assigning individuals to filing types (single, joint etc.) and marginal tax brackets (details in Appendix Section C). In both cases, I rely on the information provided in tax returns. In addition, I use a panel of tax returns known as the University of Michigan tax return panel. The panel covers 1979 to 1990 and contains the same variables as the SOI files but has a smaller sample size.

2 Missing Mass

If some taxpayers are claiming the standard deduction when the sum of their itemized deductions is greater than the standard deduction, there should be a missing mass just above the standard deduction threshold. I graph the density of deductions for several years in Figure 1 (the remaining 27 years are shown in Appendix Figures H.8-H.12).[6] The bin closest to the standard deduction includes only those itemizers whose deductions are strictly larger than the standard deduction. The density is systematically low just above the standard deduction and then increases and peaks two to three bins away. This is true across all years from 1980 to 2006 and for all filing statuses. Since I cannot observe the distribution of itemizers below the standard deduction, this cross-sectional evidence is only suggestive.[7]

To prove that the missing mass is a distortion due to the standard deduction, I turn to a quasi-experimental design. There were four large increases in the standard deduction amounts since 1960. These changes occurred in 1971, 1975, 1988 and 2003. I use the 1988 reform to estimate the cost of itemizing because other changes occurred at the same time as the 1971, 1975 and 2003 reforms. The 2003 reform is likely to provide a lower bound on the cost of itemizing

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

[7] Appendix Figure H.22 shows different alternative scenarios that could create a missing mass.

since there were changes in marginal tax rates and deduction rules that made it more attractive to itemize. In 1971 and 1975, there were changes to the parallel standard deduction system.[8] Although the magnitudes of the estimated costs for the 1971, 1975 and 2003 reforms are inaccurate, they still provide reduced-form evidence of the *existence* of compliance costs.

I compare the pre-1988 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 1988 reform.[9] The shape of the distribution in year t+1 mirrors that of years t and t-1 and 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 were a feature of the distribution and not due to the standard deduction, it would not track the standard deduction once it is increased.

3 Cost Estimation

3.1 Cost Estimation Methodology

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 shown in Figure 2a. This section explains how I reconstruct the counterfactual distribution. Importantly, this estimation method is *model-free*: the estimated distribution of forgone benefits does not require nor depends on any assumptions made over the determinants of the forgone benefits except for assumptions A1 and A2 below. No assumptions about the drivers of the cost are needed in this section: the distribution of costs can be due to the costs of record

[8] More details about the parallel system and other changes are provided in Appendix Section E.

[9] Appendix Figure H.14 reports these densities for the 1971, 1975 and 2003 reforms and shows that the changes are qualitatively consistent with the 1988 reform.

keeping and filing, or due to fear of audits and uncertainty. While alternative explanations – discussed in Section 5 – could change the interpretation of the estimated dollar amount of forgone benefits, they would not change the dollar amount itself.

Denote by $f(\cdot)$ the unobserved p.d.f. of itemizers, assuming that there is no standard deduction and no cost of itemizing, as illustrated in Figure 3.[10] Denote by $g_S(\cdot)$ the observed probability density function (p.d.f.) of itemizers when the standard deduction is equal to S. Then, $g_0(\cdot)$ and $g_\delta(\cdot)$ correspond, respectively, to the pre- and post-reform p.d.f. of itemizers when the standard deduction increases from 0 to δ. The cumulative distribution function (c.d.f.) of the cost of itemizing is denoted by $C_S(\cdot)$ and is defined over $[0, c_{max}]$, where c_{max} denotes the largest cost an individual can have. Individuals whose total deductions exceed the standard deduction by less than the cost of itemizing choose the standard deduction. Formally,

$$\forall S = \{0; \delta\} : \quad g_S(d) = \begin{cases} 0, & \text{if } d \leq S \\ f(d)C_S(d-S), & \text{if } S < d \leq c_{max} + S \\ f(d), & \text{if } d > c_{max} + S. \end{cases} \tag{1}$$

By rearranging (1) over $d \in [0, c_{max}]$:

$$C_S(d-S) = \frac{g_S(d)}{f(d)}. \tag{2}$$

In other words, the cost of itemizing is related to the missing mass $\frac{g_S(d)}{f(d)}$. However, because $f(\cdot)$ cannot be observed directly, it needs to be reconstructed using $g_0(\cdot)$ and $g_\delta(\cdot)$. Two assumptions are necessary:

- **A1**: The cost is similar pre- and post-reform.
- **A2**: The cost is independent of the level of deductions.

Assumptions A1 and A2 imply that $C_0(\cdot) = C_\delta(\cdot)$, and from equation (2), it follows that:

$$C_0(d) = \frac{g_0(d)}{f(d)} = \frac{g_\delta(d+\delta)}{f(d+\delta)} = C_\delta(d), \tag{3}$$

[10] An illustrative simulated example is provided in Appendix Section B.2.

which implies that the same proportion of individuals is missing d deductions above the pre-reform standard deduction and $d + \delta$ deductions above the post-reform standard deduction.

Assumption A1 can be verified by graphing two densities in years with no reforms and ensuring that they are overlapping. This assumption is verified on all years from 1980 to 2006.[11] A failure of A2 introduces a bias: in Appendix section B.1, I provide an upper bound on the size of this bias and show that it is small. Intuitively, if A2 fails, $g_\delta(\cdot)$ can be used instead of $f(\cdot)$. This will necessarily yield lower bounds, since $g_\delta(\cdot) < f(\cdot)$. For joint filers in the 28% bracket, for example, the estimated cost would lie between \$519 and \$591 if A2 fails, instead of \$591. In addition, if these two assumptions hold, then the missing masses in the pre- and post-reform distributions should be proportional. Appendix Figure H.18 shows that this proportionality assumption indeed holds.

To estimate $C_0(\cdot)$ and reconstruct $f(\cdot)$, I proceed in three steps. First, if $d \in [\delta + c_{max}; +\infty]$, then the benefit of itemizing is greater than its cost both pre- and post-reform, and taxpayers will not forgo deductions by claiming the standard deduction. This corresponds to the rightmost area in Figure 3. Formally, if $d \in [\delta + c_{max}; +\infty]$, then $C_0(d) = 1$ and $g_\delta(d) = g_0(d) = f(d)$, i.e., the pre- and post-reform distributions of itemizers overlap for ranges of deductions exceeding the post-reform standard deduction δ by more than the largest possible cost c_{max}. And for any $d \in [\delta + c_{max}; +\infty]$, $f(d) = g_0(d)$, i.e., the pre-reform observed distribution of itemizers $g_0(\cdot)$ corresponds to the undistorted distribution $f(\cdot)$.

Second, if $d \in [c_{max}; \delta + c_{max}]$, then, over this range, the pre-reform taxpayers do not forgo any deductions, but the post-reform ones do. This corresponds to the middle area in Figure 3. As a consequence, the pre-reform distribution is not affected by its proximity to the standard deduction and is equal to the undistorted distribution, i.e., $g_0(d) = f(d)$, but the post-reform distribution is distorted, i.e., $g_\delta(d) < f(d)$. From equation (3), it follows that $\forall d \in [c_{max}; \delta + c_{max}]$:

$$C_0(d - \delta) = \frac{g_\delta(d)}{f(d)} = \frac{g_\delta(d)}{g_0(d)}, \tag{4}$$

which allows me to estimate $C_0(\cdot)$ over $[c_{max} - \delta; c_{max}]$.

[11] See Appendix Figure H.15.

Third, if $d \in [c_{max} - \delta, c_{max}]$, then both the pre-reform and post-reform itemizers are forgoing deductions. This corresponds to the leftmost area in Figure 3. In this case, both the pre- and post-reform distributions are distorted by their proximity to the standard deduction and $g_0(\cdot)$ is now different from $f(\cdot)$. To reconstruct $f(\cdot)$, I use the estimate of $C_0(\cdot)$ over $[c_{max} - \delta; c_{max}]$ from equation 4 to correct the pre-reform distribution by using the definition of $g_0(\cdot)$ from equation (2): $f(d) = \frac{g_0(d)}{C_0(d)}$. From equation 3, it follows that $\forall d \in [c_{max} - \delta; c_{max}]$

$$C_0(d - \delta) = \frac{g_\delta(d)}{f(d)} = \frac{g_\delta(d)}{g_0(d)} C_0(d), \tag{5}$$

which allows me to estimate $C_0(\cdot)$ over $[c_{max} - 2\delta; c_{max} - \delta]$. By repeating this procedure over $[c_{max} - 3\delta; c_{max} - 2\delta]$, $[c_{max} - 4\delta; c_{max} - 3\delta]$, etc., I can recover $C_0(\cdot)$ and $f(\cdot)$ over $[0, c_{max}]$.

3.2 Cost Estimates

I apply the methodology outlined above to the 1988 reform, which increased the standard deduction from $2,540 to $3,000 for single filers, from $3,760 to $5,000 for joint filers and from $2,540 to $4,400 for head-of-households filers. Besides the standard deduction reform, the only other 1988 reform that could have affected the amount of deductions was the phase-out of the personal interest deduction, which I control for (details in Section 5.4).[12] Each cost estimate is performed on individuals with the same marginal tax rate and who were not subject to the AMT. There was a marginal tax rate decrease for married filing jointly with income above $45,000 (in 1987 dollars) in 1988. I address this by estimating the cost separately for individuals above and below this cutoff.

I use 1989 rather than 1988 as the post-reform year because the reform occurred in 1988. If taxpayers learned about the increase in the standard deduction when filing their taxes, we should observe the full response in 1989. Figure 2b confirms that the effect was smaller during the reform year.

Table 1 shows the estimated costs for single, joint and head-of-households filers in the 15% and 28% marginal tax brackets. Costs range from 0.57% to 0.85% of AGI. In dollar amounts, they vary from $175 for single filers in the 15% bracket to $591 for joint filers in the 28% bracket. Costs expressed in dollars

[12]See Appendix Section D for the TRA'86 reforms.

are systematically lower for individuals in lower tax brackets. They are, however, more homogenous when expressed as a percent of AGI. This suggests that income matters in determining the cost, as shown in Section 4.1.

To calculate the standard errors of the difference between the bins in the 1987 and 1989 densities, I use a bootstrap procedure. The results are reported in Table I.3. 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, with the exception of bins 10, 11 and 13, which are statistically significantly different at the 5% and 10% level, with differences of a very small magnitude (less than ten times that of the first or second bins).

4 Anatomy of the Missing Mass

4.1 Costs Increase 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 test this using the income reported on tax returns. I break down the sample used above by deciles of income. This reduces the power, which I deal with in two ways. First, I use a moving average of income deciles. Second, I focus on joint filers in the 28% marginal tax bracket, as they represent by far the largest group of taxpayers.

I then fit a flexible polynomial through each deduction bin and calculate the difference in density for each bin. When this difference is not statistically significant, I consider that the bins are overlapping, and, therefore, no deductions are forgone in that specific bin. Using the predicted bins from this polynomial, I calculate the forgone benefits for each group by repeating the procedure developed in Section 3.[13] The results are plotted in Figure 4a: as income increases, taxpayers forgo more deductions, consistent with the idea that they value their time more. Notice, first, that all taxpayers in Figure 4a 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. Second,

[13] I only report results for the first seven groups because deductions and income are positively correlated and very few high-income individuals are close to the standard deduction threshold. However, the increasing relationship is robust to including those bins, see Appendix Figure H.19.

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

4.2 Record Keeping

Electronic filing (e-filing) and tax preparers may reduce the cost of filling out forms. However, they do not affect the cost of record keeping. Therefore, e-filing and tax preparers are unlikely to eliminate the observed missing mass if it is due to the hassle of filling out forms. Survey estimates of the cost of filing taxes have consistently documented that record keeping is the main driver of the cost of itemizing.[14]

To test for the effect of e-filing and tax preparers, I plot the density of itemizers who use a tax preparer or e-file in Figures 4c and 4b. In both cases, the distribution of itemizers exhibits a clear missing mass, implying that e-filing or tax preparers do not eliminate the cost of itemizing.

Figure 4c compares the density of taxpayers who use e-filing to those who do not. It shows a slightly smaller missing mass for taxpayers who e-file.[15] However, e-filing only slightly reduces the cost of itemizing and does not eliminate the missing mass, which is consistent with record keeping being the main driver of compliance costs. I cannot perform a similar test for taxpayers who use tax preparers, as the two densities do not overlap away from the standard deduction, making a direct comparison of the missing mass impossible without additional parametric assumptions.[16]

Given that costs appear to be driven by record keeping rather than filling out schedule A, it is likely that costs increase with the number of receipts one has to claim, but not necessarily with the dollar amount claimed.[17] Consider the

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

[15]The difference is statistically significant: bootstrapped standard errors are reported in Table I.6

[16]One could also compare states with and without state income taxes. Since state income taxes are reported on W2s they are unlikely to impose a large cost of record keeping. Appendix Figure H.20 shows a missing mass for both states with and without income taxes, suggesting that the state income tax deduction is not a large driver of costs.

[17]A direct test of this assumption is not possible with the data I use, because the number of receipts is not reported in the data.

mortgage interest deduction or the state income tax deduction: taxpayers receive a statement, the 1098 for mortgage interest and W2 for state income taxes, close to the filing deadline, summing up all mortgage interest or state income taxes paid. For this reason, the record keeping cost of claiming $1,000 or $10,000 worth of mortgage interest are likely to be similar. However, this is not true anymore for non-third-party reported items, and in particular for charitable donations. In this case, since no statement is provided, each additional donation is likely to be costly in terms of record keeping. This suggests that the main driver of record keeping costs is not necessarily the level of donations but instead the number of separate donations: a $1,000 donation imposes less record keeping costs than 20 different $50 donations, even though they add up to the same total amount. The fact that costs are likely to be variable for certain deductions is consistent with the findings from Rees-Jones (2018) who shows that taxpayers can manipulate the balance due on their taxes – presumably by adjusting their deductions – to precisely target a balance of $0. Overall, if record keeping costs are partly variable, partly fixed, it is likely that the variable portion of these costs is pivotal to the decision to itemize. This would imply that the deductions that impose a variable cost on taxpayers, such as charitable donations, are likely to be the main driver of the missing mass.

5 Alternative Explanations to Compliance Costs

In this Section, I consider alternative explanations for the missing mass. Note that information is unlikely to explain it, since I focus on taxpayers who switch from itemizing to claiming the standard deduction, they should be aware of the decision to itemize. In addition, taxpayers are reminded on the 1040 form that they can make the choice between itemizing and claiming the standard deduction.

5.1 Concave Kink Points

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, possibly inducing some taxpayers to be indifferent between these two points. 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. This argument is illustrated in Figure 5.

However, in this case, the size of the missing mass should not respond to variations in income when controlling for marginal tax rates. The only reason taxpayers would adjust their deductions in response to a concave kink point is variation in marginal tax rates, while changes in income should not matter. On the other hand, a behavioral response due to compliance costs predicts that richer taxpayers will forgo more money because they have a higher opportunity cost of time, even controlling for marginal tax rates. Consistent with the compliance cost interpretation, Figure 4a graphs the relationship between forgone benefits and income - *controlling for marginal tax rates* - and finds an increasing relationship.

In addition, behavioral responses to concave kink points lead individuals to locate away from the concave kink point. If behavioral responses to concave kink points led to the observed missing mass, as the standard deduction increases, the bi-modal distribution should track the new standard deduction threshold and the pre- and post-distribution peaks should not overlap. The observed pre- and post- distributions in Figures 2a and 2b contradict this prediction: the pre- and post-distribution peaks are overlapping, once again rejecting the hypothesis that the missing mass is caused by behavioral responses to concave kink points. A graphical illustration of this argument is provided in Figure 5.

Overall, both of these empirical tests rule out responses to concave kink points. This is consistent with the previous literature. Saez (2010), Kleven and Waseem (2013) and Tazhitdinova (2017) directly test the predictions of a behavioral response to both concave and convex kink points, and find responses to convex kink points but not to concave kink points. Kleven (2016), in a survey of the bunching literature confirms that there is no evidence of bunching at concave kink points.

5.2 Evasion

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

However, since audit probabilities are very low, 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 conduct a survey of 195 individuals.[18] The survey allows me to elicit their beliefs about both the audit probabilities and the perceived costs of audits.

Surveyed individuals have levels of income similar to those of joint filers in the 1988 28% marginal tax bracket. On average, they believe that audits occur with a probability of 8.72%, which is 7.9 times the true audit probability.[19] This accounts for, at most, 25% of the $591 estimated forgone benefits for joint filers in the 28% marginal tax bracket.[20]

5.3 Rational Inattention

Can uncertainty about the level of deductions lead taxpayers to switch to the standard deduction and explain the observed missing mass? Table I.4 shows the results of the calibration of a model illustrative of this type of behavior with varying levels of risk aversion.[21] Taxpayers would need an uncertainty range of at least $\pm\$14,000$ in order to forgo amounts of money as those found in this paper when their true deductions are $10,000. This uncertainty range is large and unlikely for two reasons.

First, I focus on taxpayers who were itemizing in the previous year. Second, total deductions are highly serially correlated across years for a given individual, since 71% of total deductions are mortgage interest, state taxes and property taxes, which are relatively stable for a given person year after year. Indeed, Appendix Figure H.21 plots the distribution of year-on-year percent change in the level of deductions by taxpayer, using the 1979 to 1990 panel. It shows that, on average, deductions tend to increase nominally. Therefore, it is likely that taxpayers who were previously itemizing will believe that their deductions will be nominally higher this year and therefore will hold relatively precise beliefs over the lower bound on the tax savings they can derive from itemizing.[22]

[18] Appendix section F details the survey instrument.

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

[20] On average, expected audit costs are $147, with a 95% confidence interval of $[126, 169]$.

[21] The model is outlined in Appendix section G.

[22] Taxpayers may also hold biased beliefs over tax rates, leading them to mis-perceive the savings they can derive from itemizing. Kuziemko et al. (2015), for example, shows that taxpayers are inattentive to certain exemption thresholds, which could lead them to over-estimate aver-

5.4 Other Reforms Affecting Deductions?

Other changes took place in 1988. In this section, I describe these changes and explain how I adjust for those that are likely to affect my estimates. The estimates derived in Section 3 have 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 in 1986. Taxpayers could deduct only 65% of their personal interest in 1987, 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, and I drop them from my sample. 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 above and below the 28% marginal tax rate bracket. If there had been a behavioral effect, we should observe more deductions for individuals above the 28% marginal tax bracket. I find no significant behavioral response of personal interest deductions.[23] 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.

age tax rates. On the other hand, De Bartolome (1995) and Rees-Jones and Taubinsky (2019) show that taxpayers tend to "iron" tax schedules, leading to an underestimate of marginal tax rates. Abeler and Jäger (2015) show that complexity plays a large role in these misperceptions. Since the 1987, 1988 and 1989 tax schedules were only comprised of two tax brackets, these biases, while possible, are unlikely to be meaningfully large.

[23] See Appendix Figure H.23.

5.5 Contemporary Relevance

There are three main reasons why the tax filing costs estimated in this paper are likely to be relevant for recent years. First, while electronic filing has been rising rapidly over the past two decades, it does not appear to have drastically reduced filing costs. Figure 6a shows that the distribution of itemized deductions still exhibits a missing mass just above the standard deduction, even in the most recently available year (2009). Figure 6b similarly shows the missing mass for electronic filers in 2009.[24] These Figures imply that taxpayers are still failing to itemize in spite of the advent of electronic filing. This result is consistent with filing costs being mostly driven by record keeping.

In addition, it is likely that overall filing costs have increased over time, as we observe a steady increase in the *number* of schedules filed over time. Figure 4d plots the number of forms filed by schedule over time. While one additional Schedule was filed for each 1040 form in the mid-1980's, this proportion is almost two-to-one in recent years, suggesting that costs may have been increasing over time, as taxpayers have to file more and more schedules.

6 Compliance Costs or Behavioral Costs?

There is extensive evidence that individuals are time-inconsistent and tend to procrastinate.[25] If taxpayers procrastinate on filing their taxes, one should observe a large proportion of taxpayers filing on April 15th and procrastinators forgoing more deductions.[26]

First, consistent with individuals procrastinating on filing their taxes, I find that taxpayers bunch at the April 15th deadline. Figure 7a graphs the volume of Google searches for the term *1040* by week, and Figure 7b uses data from irs.gov[27] and graphs the number of tax returns filed by week. Both exhibit a clear spike in the weeks that include April 15th. This is consistent with Hoopes et al. (2015), who show that taxpayers search more actively on Google and Wikipedia for capital-gains-tax-related information close to April 15th.

[24] Similar patterns are observed in 2009 for taxpayers who use a tax preparer (Figure 6c) and who self file their returns (Figure 6d).

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

[26] This argument is formalized in Appendix section H, which shows that procrastination can lead to high record-keeping costs, resulting in individuals failing to itemize.

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

Second, I also find that taxpayers who file close to the deadline tend to forgo more deductions, consistent with procrastination accounting for a portion of the estimated forgone deductions. Figure 7c[28] 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: (1) by delaying filing, they forgo interest on their refunds; and (2) they expose themselves to higher filing costs. Indeed, the sample I use to generate Figure 7c includes only those taxpayers who are owed a refund by the IRS and, therefore, have an incentive to file as early as possible to save on interest.[29] Second, filing costs are substantially higher closer to the deadline because lines at the post office are longer; appointments with tax preparers are scarcer; and it is harder to get tax help from the IRS because their phone lines are busier than usual.[30]

Note 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 $t + 1$. To test for this, in Figure 7d, I graph the average week in which returns are processed in year $t + 1$ by week of processing in year t. 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 the year t week of processing should predict the year $t + 1$ week of processing. Figure 7d shows a clear increasing relationship between the processing weeks in year t and year $t + 1$, consistent with the explanation that late filing is due to a systematic bias.

[28]Appendix section C.5 explains how the graph is constructed.

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

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

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 tax benefits, suggesting large tax filing costs. The identification strategy used in this paper can be exported to estimate other compliance costs when individuals have a choice between a low-cost/low-benefit option versus a high-cost/high-benefit one. It can also be used when identifying responses from a censored distribution above or below a certain threshold.

References

Abeler, J. and S. Jäger (2015). Complex Tax Incentives. *American Economic Journal: Economic Policy 7*(3), 1–28.

Aghion, P., U. Akcigit, M. Lequien, and S. Stantcheva (2017). Tax Simplicity and Heterogeneous Learning.

Bertrand, M., S. Mullainathan, and E. Shafir (2006). Behavioral Economics and Marketing in Aid of Decision Making Among the Poor. *Journal of Public Policy & Marketing 25*(1), 8–23.

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.

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.

De Bartolome, C. A. (1995). Which Tax Rate do People Use: Average or Marginal? *Journal of public Economics 56*(1), 79–96.

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

Gronau, R. (1973). The Effect of Children on the Housewife's Value of Time. *The Journal of Political Economy*, S168–S199.

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.

Harju, J., T. Matikka, and T. Rauhanen (2019). Compliance Costs vs. Tax Incentives: Why Do Entrepreneurs Respond to Size-Based Regulations? *Journal of Public Economics 173*, 139–164.

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. Slemrod (2015). Taxpayer search for information: Implications for rational attention. *American Economic Journal: Economic Policy 7*(3), 177–208.

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.

Kuziemko, I., M. I. Norton, E. Saez, and S. Stantcheva (2015). How Elastic Are Preferences For Redistribution? Evidence From Randomized Survey Experiments. *American Economic Review 105*(4), 1478–1508.

Little, A. D. (1988). Development of Methodology for Estimating the Taxpayer Paperwork Burden. *Working Paper*.

Nelson, F. D. (1977). Censored Regression Models with Unobserved, Stochastic Censoring Thresholds. *Journal of Econometrics 6*(3), 309–327.

Nichols, A. L. and R. J. Zeckhauser (1982). Targeting Transfers Through Restrictions on Recipients. *The American Economic Review 72*(2), 372–377.

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. (2018). Quantifying loss-averse tax manipulation. *The Review of Economic Studies 85*(2), 1251–1278.

Rees-Jones, A. and D. Taubinsky (2019). Measuring "Schmeduling". *NBER Working Paper*.

Saez, E. (2010). Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy*, 180–212.

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.

Tazhitdinova, A. (2016). Income Shifting and the Cost of Incorporation. *Working Paper*.

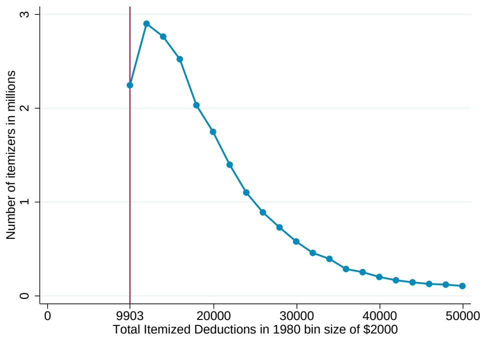
Tazhitdinova, A. (2017). Adjust Me if I Can't: The Effect of Firm Incentives on Labor Supply Responses to Taxes. *Working Paper*.

Tazhitdinova, A. (2018). Reducing Evasion Through Self-Reporting: Evidence From Charitable Contributions. *Journal of Public Economics 165*, 31–47.

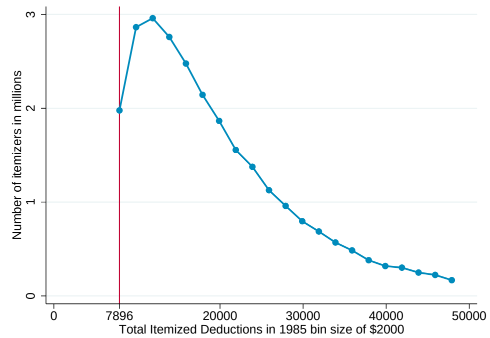
Wicks, J. H. (1965). Taxpayer Compliance Costs From the Montana Personal Income Tax. *Montana Business Quarterly 3*, 36–42.

Wicks, J. H. and M. N. Killworth (1967). Administrative and Compliance Costs of State and Local Taxes. *National Tax Journal*, 309–315.

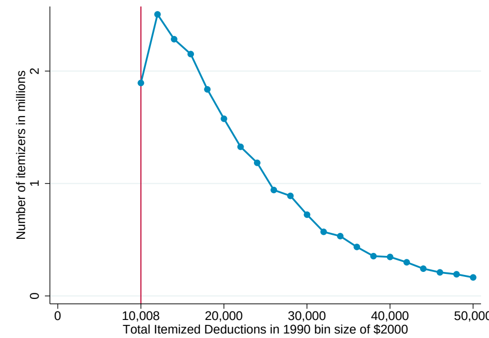
Figure 1: Missing Mass Just Above the Standard Deduction



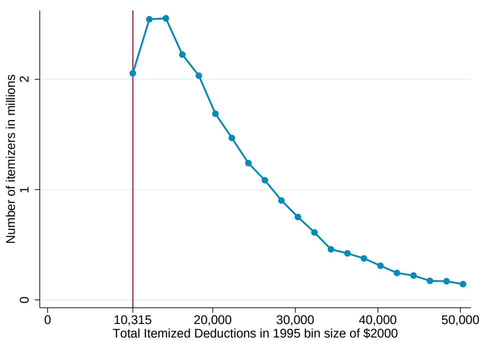
(a) 1980



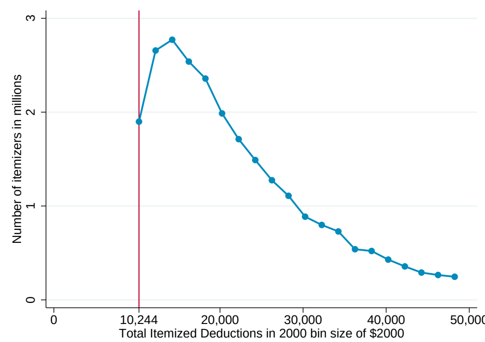
(b) 1985



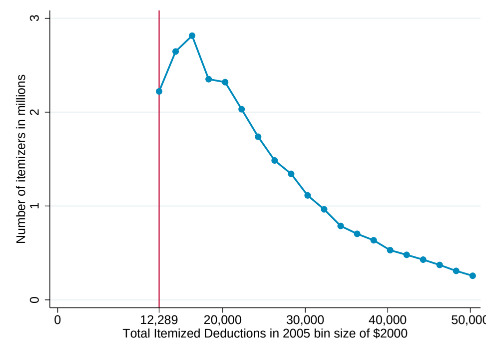
(c) 1990



(d) 1995



(e) 2000

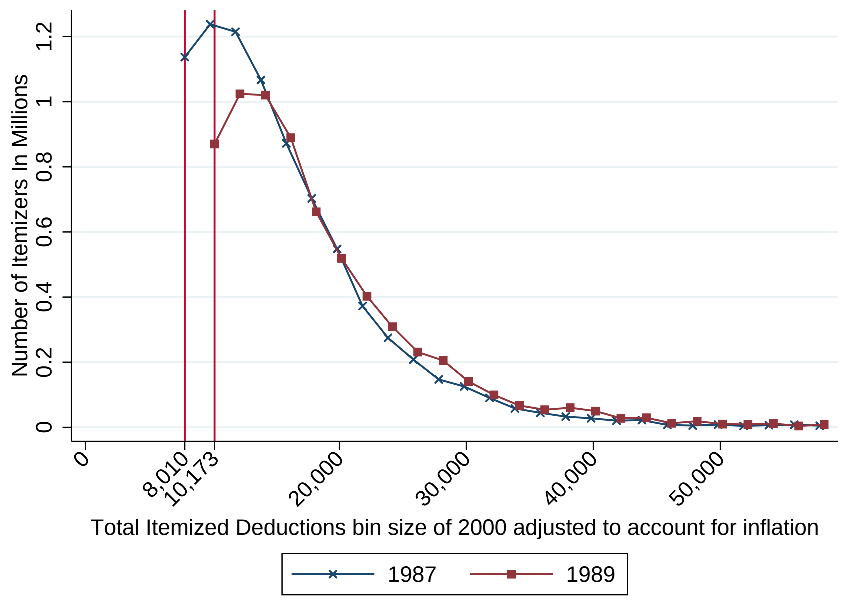


(f) 2005

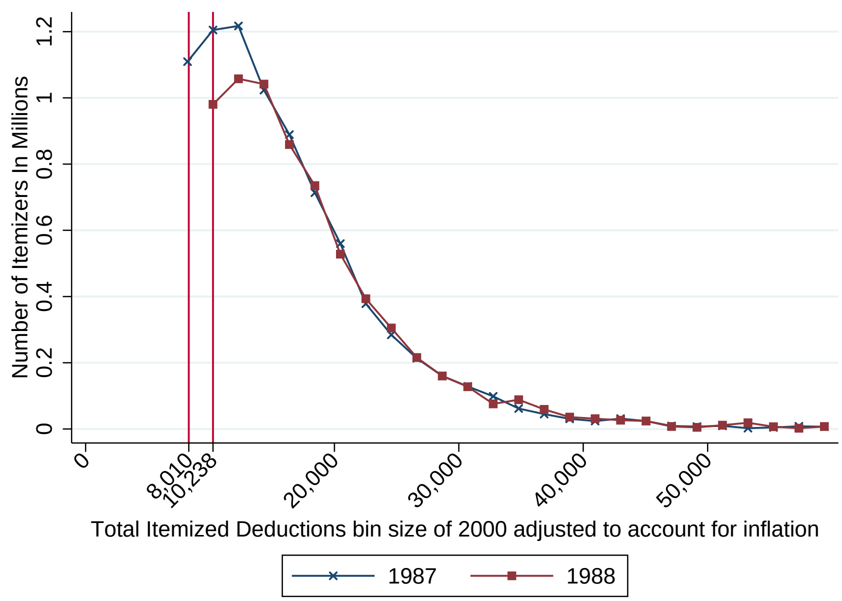
Notes: These Figures plot the density of deductions for itemizers filing jointly. The bin size is $2,000 and the vertical line represents the standard deduction threshold for each year. Additional years are reported in Appendix Figures H.9, H.10, H.11 and H.12 and Figure H.13 for single filers.

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

(a) 1987-1989 Comparison

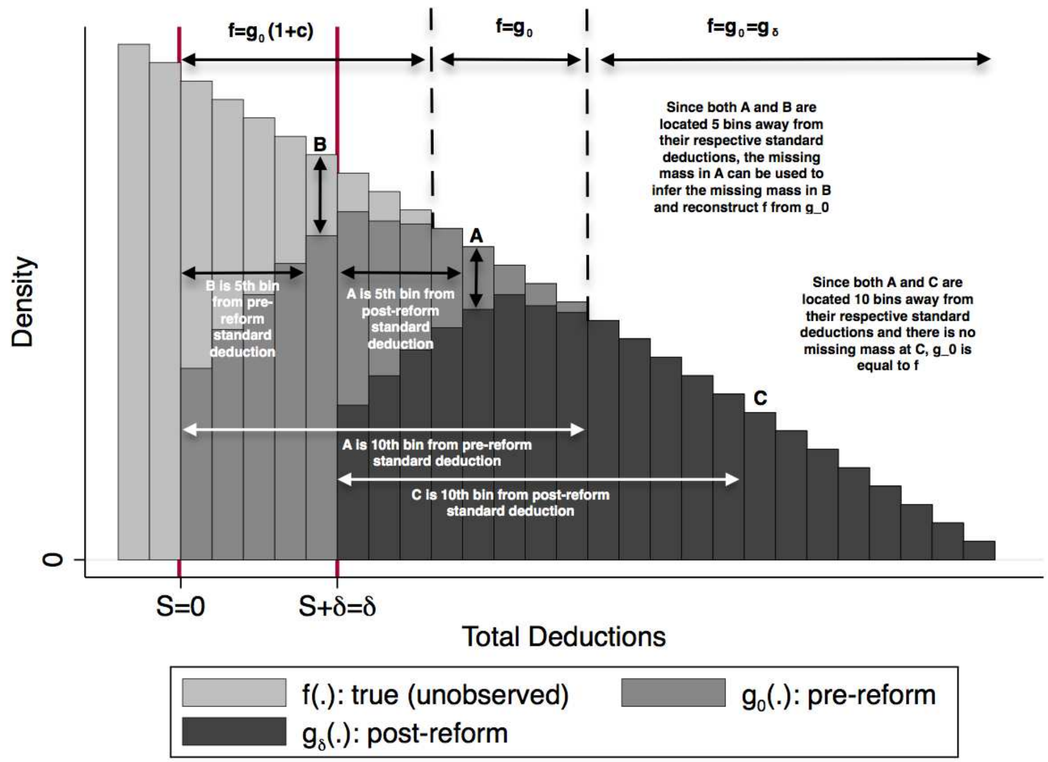


(b) 1987-1988 Comparison



Notes: The first graph plots the density of itemizers one year before and one year after the standard deduction reform, while the second one plots these densities one year before and during the reform.

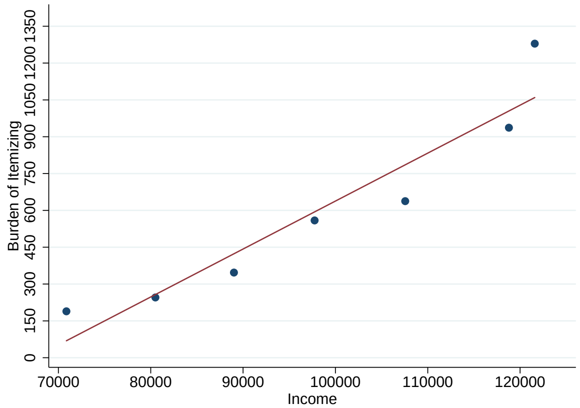
Figure 3: Recovering the Counterfactual Distribution of Deductions



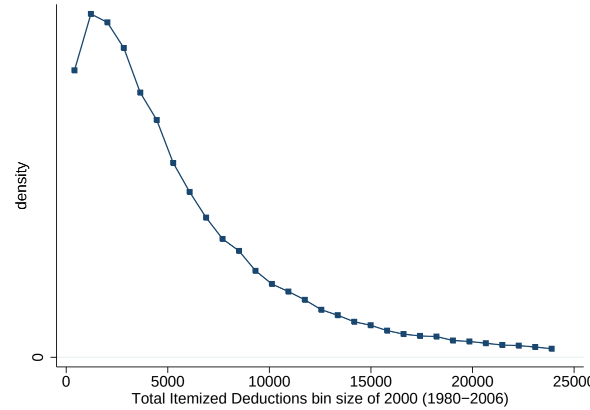
Notes: This graph illustrates the method used in Section 3.1 to reconstruct the counterfactual density of itemizers $f(.)$ using the pre- and post-reform densities $g_0(.)$ and $g_\delta(.)$.

Figure 4: Anatomy of the Missing Mass

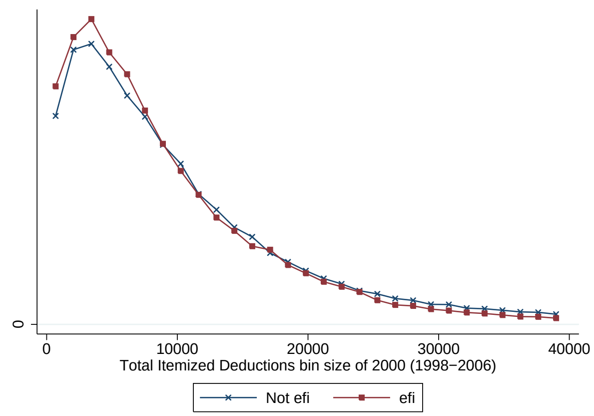
(a) Costs Increase with Income



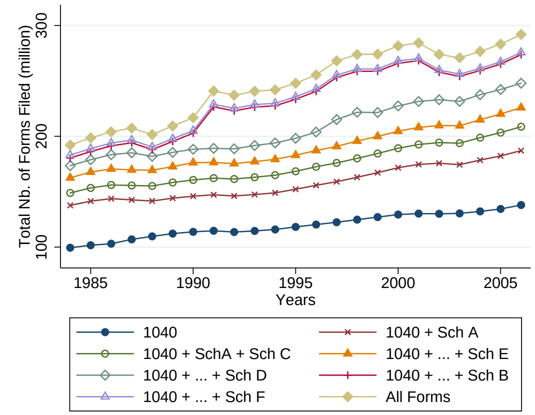
(b) Tax Preparers



(c) Electronic Filing



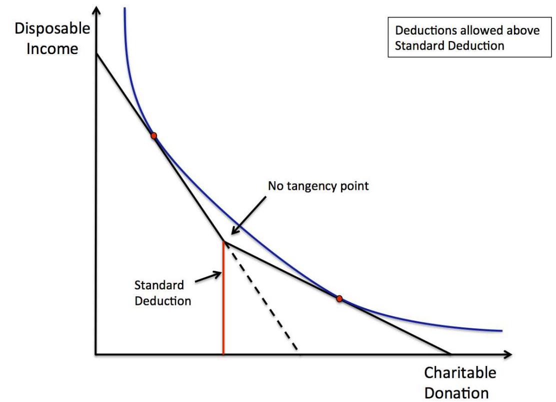
(d) Number of Forms Filed



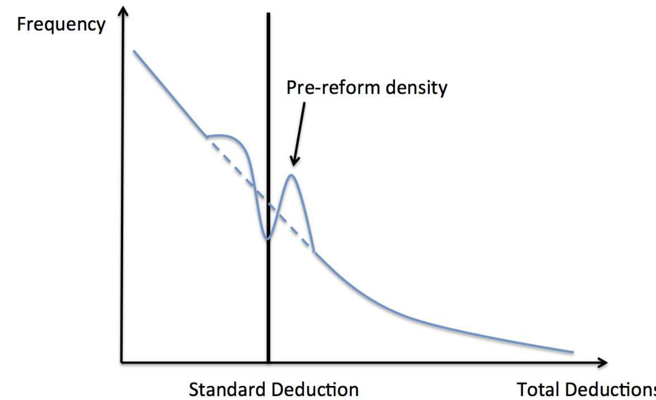
Notes: The first graph plots the relationship between income and the cost of itemizing. This relationship controls for the variation in MTR across the different income groups. The second and third Figures show the distribution of itemized deductions for taxpayers who use tax preparers and e-filing. The x-axis is normalized such that 0 corresponds to the standard deduction threshold. The third graph pools years 1980 to 2006 and the fourth one years 1998 (start of e-filing) to 2006. The fourth graph shows the evolution of the number of forms filed over time.

Figure 5: Concave Kink Point: Densities Following Reform Should Not Overlap

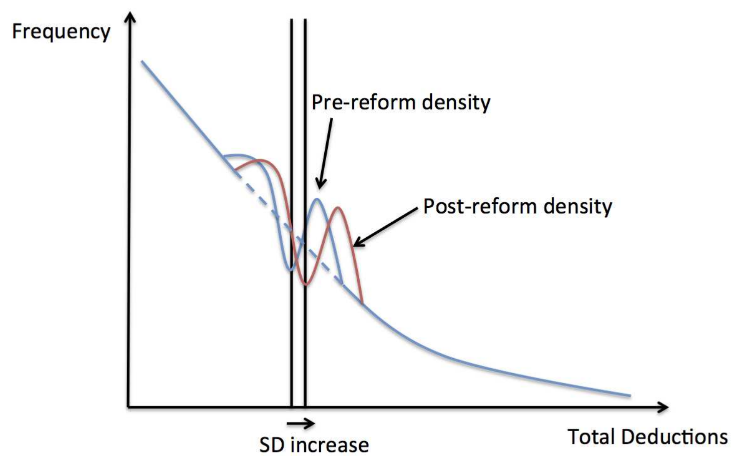
(a) Concave Kink Point



(b) Missing Mass Due to Concave Kink Point



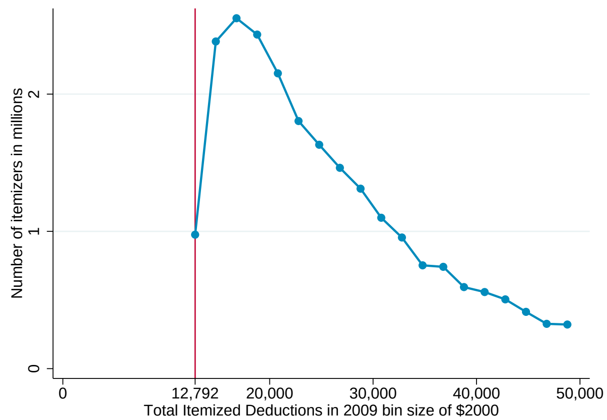
(c) Pre and Post Reform Densities



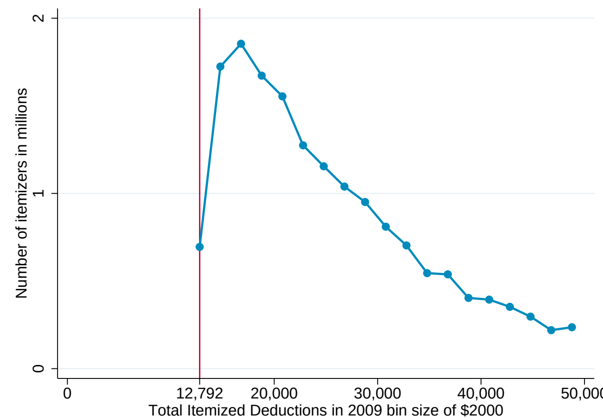
Notes: Panel (a) displays a budget set with a concave kink point. Panel (b) shows the effect that a concave kink point could in theory have on the density of itemizers. Panel (c) shows that if itemizers were responding to the concave kink point, we should observe that the pre and post reform densities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

Figure 6: Distributions in the Most Recently Available Year (2009)

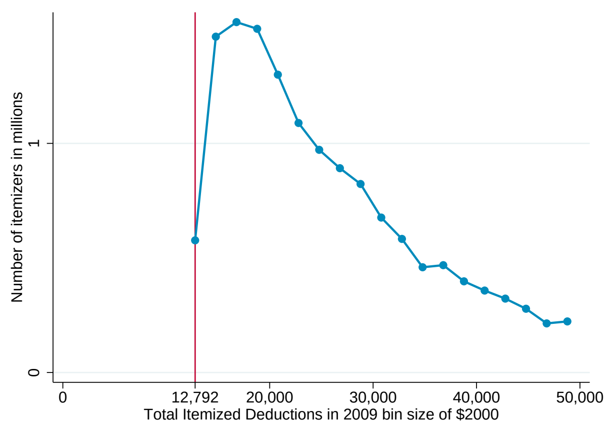
(a) All Taxpayers



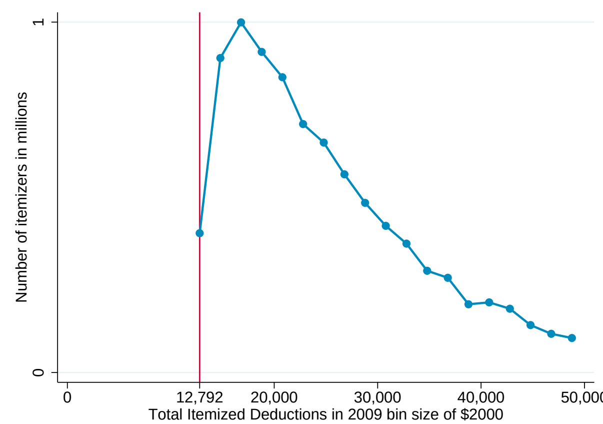
(b) Electronic Filers



(c) Tax-Preparer-Filed Returns

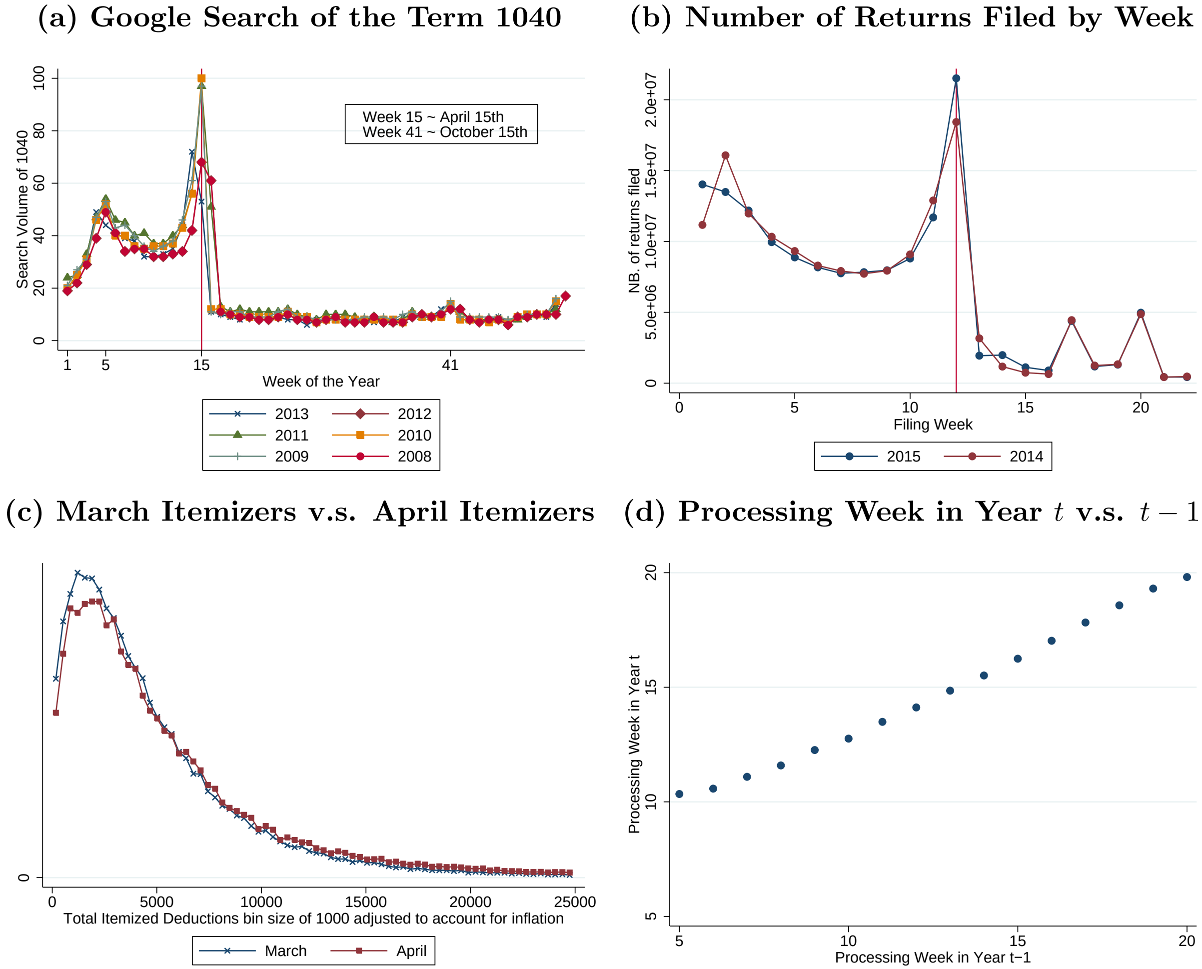


(d) Self-Filed Returns



Notes: These Figures plot the distribution of itemizers in 2009, which is the most recently available year in the SOI dataset. Figure (a) shows the distribution for all taxpayers, Figure (b) for taxpayers who file electronically, Figure (c) for taxpayers who use a tax preparer and Figure (d) for taxpayers who self file.

Figure 7: Deadline Effects



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.

Table 1: Cost Estimates

Filing Status	MTR	Cost as % of AGI	Implied Hours	Cost in $
Single	15%	0.83	14.7	$175
Single	28%	0.85	15.2	$369
Joint	15%	0.57	10.2	$243
Joint	28%	0.74	13.2	$591
Head	15%	0.76	13.6	$270
Head	28%	0.72	12.8	$458

Notes: This Table shows the cost of itemizing as estimated in section 3 for different brackets and filing status.

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 compliance 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.

B Cost Estimation

B.1 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.

Assumption A2 is important for equation 3. 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. 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 $g_\delta(d)$ as the true counterfactual instead of $f(d)$. This is a generous upper bound because it assumes that the pre-reform distribution is undistorted just above the standard deduction in spite of Figure 2a showing a clear distortion. In this case, the estimated cost would be $519 instead of $591. Therefore if A2 fails, the cost of itemizing would lie between $519 and $591.

B.2 Simulated Example

The following example illustrates the estimation process outlined in Section 3.1. Assume an undistorted and hypothetical distribution of deductions f(.) as represented by the lightest distribution in Figure 3. Assume for simplicity and without loss of generality that each bin size is equal to $100 of deductions. And assume that the cost distribution in the population is given by the following, where the cost is expressed in deductions rather than dollars (to get dollars, one simply needs to multiply the deduction amount by the marginal tax rate):

- 40% have a cost lower than 100
- 50% have a cost lower than 200
- 60% have a cost lower than 300
- 70% have a cost lower than 400
- 80% have a cost lower than 500
- 90% have a cost lower than 600
- 92% have a cost lower than 700
- 96% have a cost lower than 800
- 100% have a cost lower than 900

Introduce a standard deduction at the second bin in Figure 3 and apply the cost outlined above to the density. The distributions denoted by $g_0(.)$ is the pre-reform distribution, which is empirically observed and corresponds to the distributions plotted in Figure 1. In Figure 3, $g_0(.)$ is representedby the second lightest distribution. If $f(.)$ was observable, one would simply compare the percentage difference between $f(.)$ and $g_0(.)$ in order to calculate the cost distribution. However, $f(.)$ is not observable. This is why I rely on comparing the pre- and post-reform distributions $g_0(.)$ and $g_\delta(.)$. Figure 3 assumes that the cost distribution is the same across years and introduces a reform that increase the standard deduction amount by \$500 (5 bins). I denote by d_i the distortion introduced by the standard deduction in bin i. 40% of the population has a cost that is 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, $d_2 = 50\%$, $d_3 = 40\%$, $d_4 = 30\%$, $d_5 = 20\%$, $d_6 = 10\%$, $d_7 = 8\%$, $d_8 = 4\%$ and $d_i = 0$ for any $i > 8$. I use these values in order to generate $g_\delta(.)$. Using the method discussed in Section 3.1, I recover the true density of deductions $f(.)$ in the following way.

Denote by b_i a bin of size 100 that starts at i. For example, $b_{200} = [200, 300]$. Relatedly, $f(b_{200})$ corresponds to number of itemizers with total deductions that range between 200 to 300. Since $f(.)$ corresponds to the distribution of itemizers had there been no distortion imposed by the cost of itemizing, as described in Figure 3, $f(b_{200})$ is the number of itemizers one would observe with deductions between 200 and 300 had there been no distortion imposed by the cost of itemizing. Similarly, $g_0(b_{200})$ is the number of itemizers with deductions 200 to 300 in excess of the pre-reform standard deduction; and $g_\delta(b_{200})$ is the number of itemizers with deductions 200 to 300 in excess of the post-reform standard deduction once it increases by δ.

First, consider $g_\delta(b_{900})/g_0(b_{1400})$. In Figure 3, this ratio corresponds to the ratio of the darkest and second darkest distributions at the bin that is 14 bins away from the pre-reform distribution (and therefore 9 bins away from the post-reform distribution). At this bin, $g_0(.)$ and $g_\delta(.)$ are overlapping, implying that $g_\delta(b_{900})/g_0(b_{1400}) = 1$.

Recall, from Section 3.1, that $g_\delta b_{900}/g_0 b_{1400} = C(900)$, where $C(.)$ is the CDF of costs. This implies that $C(900) = 1$, i.e., no taxpayer has a cost in excess of 900. Intuitively, since $g_\delta b_9 00$ is undistorted, that must mean that there is no taxpayer with a cost exceeding 900. This has two implications. First, we now know that costs do not exceed \$900. Second, this tells us that both the pre- and post-reform distributions are undistorted when considering taxpayers with deductions \$900 in excess of the standard deduction. This means that for any deductions 900 in excess of the standard deduction, $f(.) = g_0(.)$.

Next, consider $g_\delta(b_{800})/g_0(b_{1300})$. In Figure 3, this ratio corresponds to the ratio of the darkest and second darkest distributions at the bin that is 13 bins away from the pre-reform distribution. Since $g_0 b_{1300}$ is undistorted, $g_\delta(b_{800})/g_0(b_{1300}) = g_\delta(b_{800})/f(b_{1300}) = C(800) = 96\%$.

The same procedure is applied to recover $C(700) = 92\%$, $C(600) = 90\%$, $C(500) = 80\%$ and $C(400) = 70\%$. These are all the bins in Figure 3 where $f(.) = g_0(.)$, and where $g_0(.)$ can be used as the true counterfactual.

This is not the case anymore for $C(300)$, since $f(.)$ and $g_0(.)$ do not overlap anymore. Therefore, we need to re-construct $f(.)$ from $g_0(.)$ and $g_\delta(.)$. Notice that the bin at which $f(.)$ and $g_0(.)$ diverge is 8 bins away from the pre-reform standard deduction. Notice also, that we know that $C(800) = 96\%$, which means that 4% of individuals will not itemize when their deductions are lower than 800. This implies that the distortion imposed by the cost of itemizing on $g_0(.)$ is 4%.

Therefore, $f(800) = g_0(800)/96\%$.

Now that we know $f(800)$, we can infer $C(300) = f(800) - g_\delta(300) = g_0(800)/C(300) - g_\delta(300)$, which is equal to 60%. Similarly, we can calculate $C(200) = f(700) - g_\delta(200) = g_0(700)/C(200) - g_\delta(200)$, which is equal to 50% and $C(100) = f(600) - g_\delta(100) = g_0(600)/C(100) - g_\delta(100)$, which is equal to 40% and therefore recover the cost distribution.

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. Figure H.13 shows the same patterns for single taxpayers.

C.2 Figures 2a, 2b

In Figure 2a and 2b, I focus on 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 change in marginal tax rates, 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.

C.3 Figure 4a

In Figure 4a, I use the same sample restrictions as in Figure 2a and 2b and break down the sample into deciles of income.

C.4 Figures 4c and 4b

To generate Figures 4c and 4b, 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 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.

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. These individuals are dropped from my sample. 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 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.

E The 1971, 1975 and 2003 reforms

E.1 The 1971 and 1975 reforms

In 1970 and 1975 taxpayers could claim as a standard deduction the smaller of the standard deduction 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. In 1975, a similar two tiered standard deduction existed with an AGI limit of 16% and a dollar limit of $74,431.

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 compliance costs – but only because their deductions are now smaller than 13% of their income. This is why using 1971 and 1975 will not yield accurate estimate of compliance costs (they tend to over-estimate them).

E.2 The 2003 reform

Two main changes occurred in 2003 that affect the post-reform standard deduction. The first one is that tax rates were reduced 2 to 3 percentage points (depending on the bracket), reducing the incentive to itemize. The second one is that electronic filing was rapidly expanding in the early 2000's complicating the comparison between the pre and post-reform standard deduction.

F 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 $1,000)

G 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? Figure H.21 shows that deductions tend to increase year-on-year. This means that taxpayers should have a precise signal of the lower bound of tax savings they can derive from itemizing. I formalize this argument below:
Assume that a given 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 this 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:

$$\mathbb{E}\left[\frac{1}{1-\theta}(\tau-c)^{1-\theta}\right] \geq \frac{1}{1-\theta}S^{1-\theta}. \tag{6}$$

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:

$$\frac{1}{1-\theta}(\mu-c)^{1-\theta} - \frac{1}{2}\theta(\mu-c)^{-1-\theta}\sigma^2 \geq \frac{1}{1-\theta}S^{1-\theta}. \tag{7}$$

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

$$\log(\mu-c) - \frac{\sigma^2}{2(\mu-c)^2} \geq \log(S). \tag{8}$$

The first term in equation 8 is the expected benefit that the taxpayer derives from itemizing. The second term is a correction for the risk aversion of the taxpayer who will itemize deductions if the benefit of itemizing corrected for risk aversion is greater than the benefit derived 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 $0 < \theta \leq 2$[31] in Table I.4. I fix the standard deduction at \$10,000 for joint filers. The cost estimated by the IRS of the time required to itemize deductions is $c = 149$. 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.

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

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.

H 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:

$$-c(1+r)^t - k < \beta[-c(1+r)^{t+1} - k]. \tag{9}$$

A sufficient condition for equation 9 to hold is:

$$(1 + r)\beta < 1. \tag{10}$$

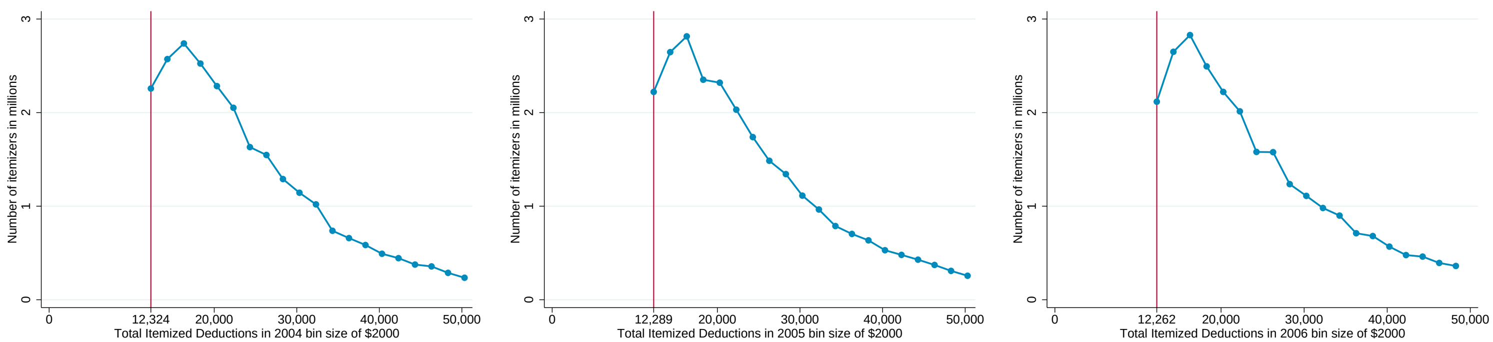
Intuitively, for the naive present-biased taxpayer to procrastinate on archiving her receipt, it is sufficient that the rate at which the record keeping cost increases be smaller than the rate at which she discounts the future.

Provided that condition 9 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 9 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.

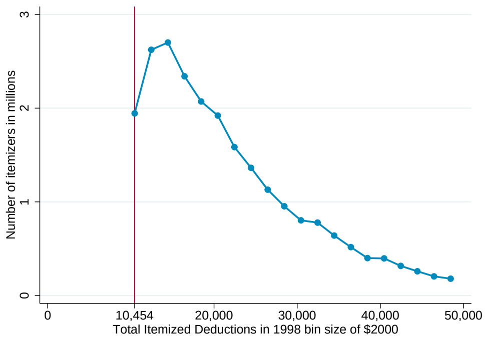
Figure H.8: Missing Mass Just Above the Standard Deduction 2004-2006 (Joint Filers)



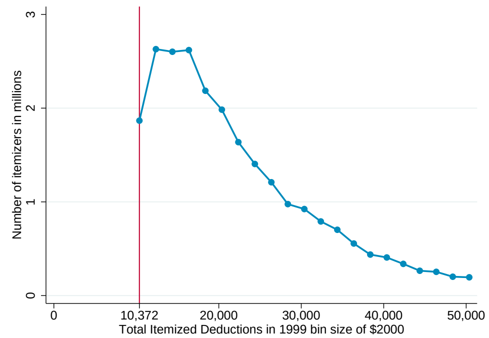
(a) 2004 **(b) 2005** **(c) 2006**

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.

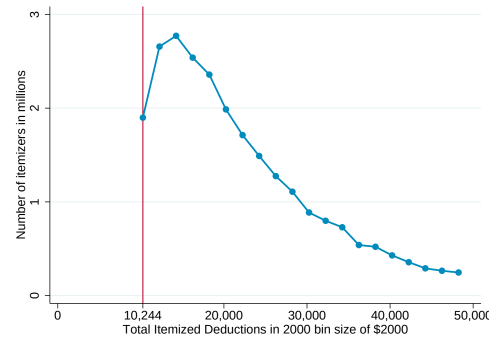
Figure H.9: Missing Mass Just Above the Standard Deduction 1998-2003 (Joint Filers)



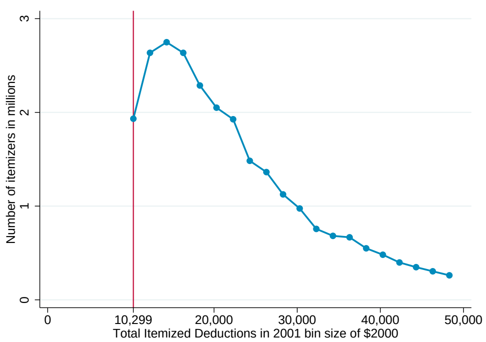
(a) 1998



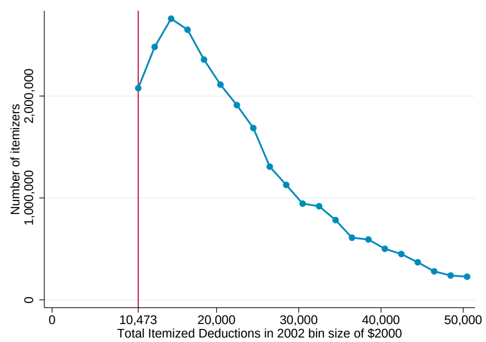
(b) 1999



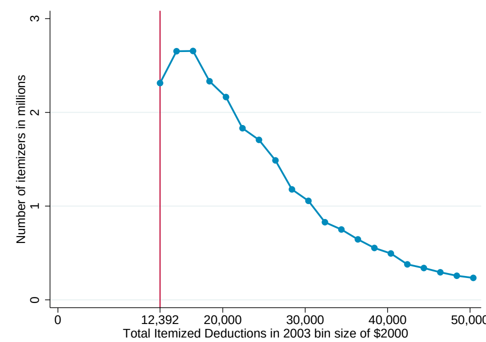
(c) 2000



(d) 2001



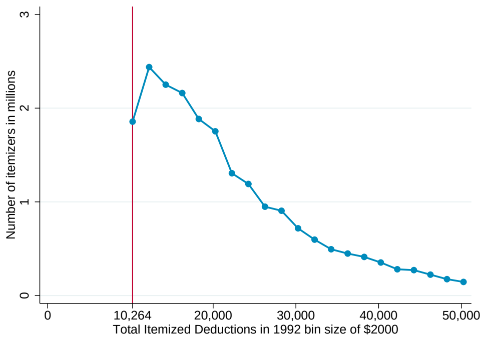
(e) 2002



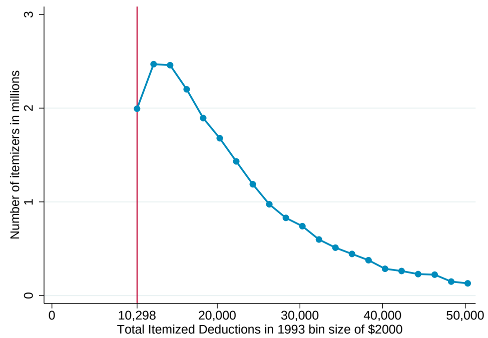
(f) 2003

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.

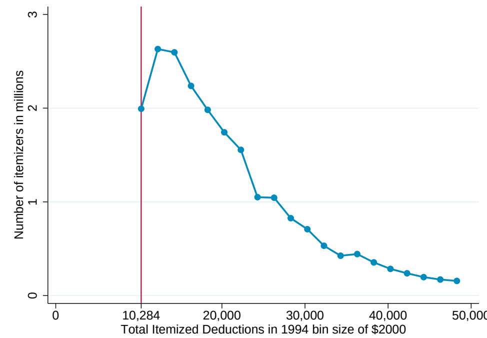
Figure H.10: Missing Mass Just Above the Standard Deduction 1992-1997 (Joint Filers)



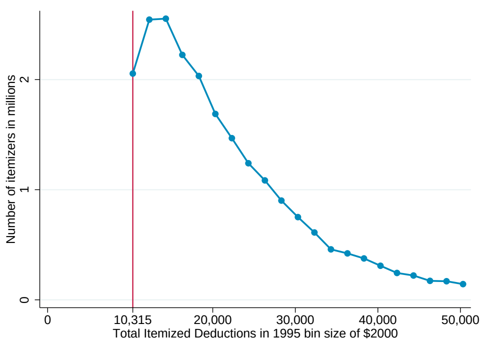
(a) 1992



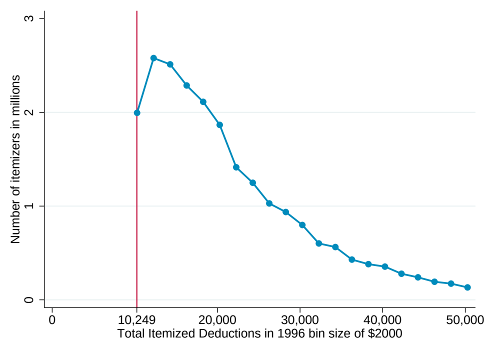
(b) 1993



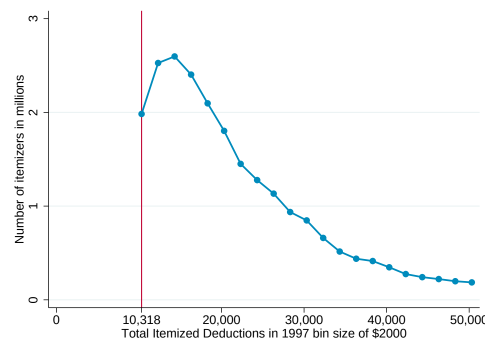
(c) 1994



(d) 1995



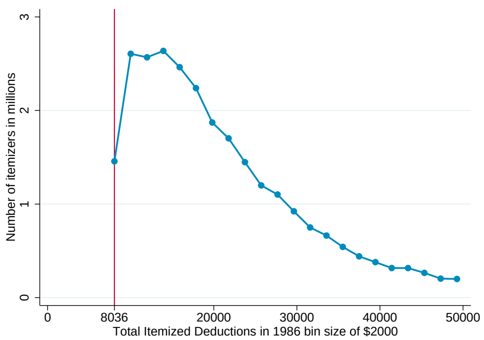
(e) 1996



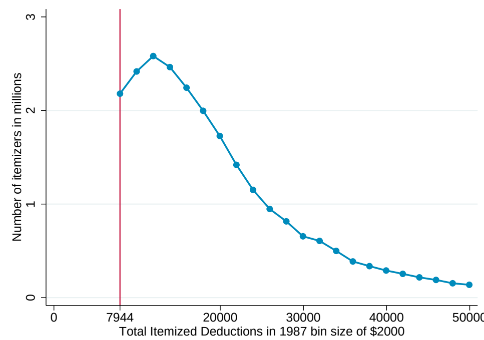
(f) 1997

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.

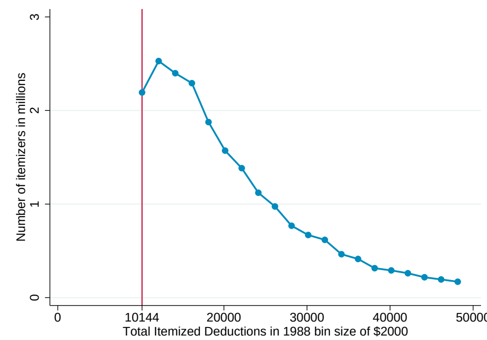
Figure H.11: Missing Mass Just Above the Standard Deduction 1986-1991 (Joint Filers)



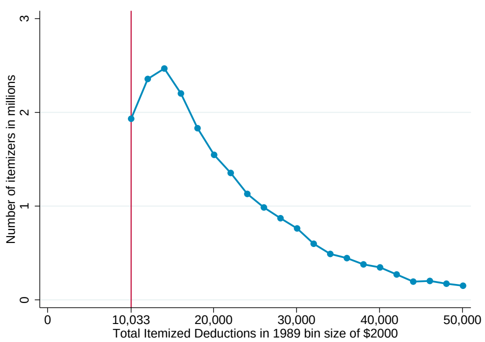
(a) 1986



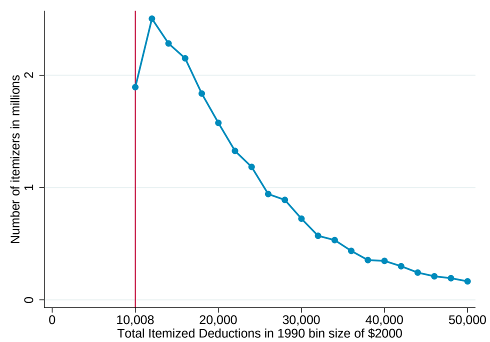
(b) 1987



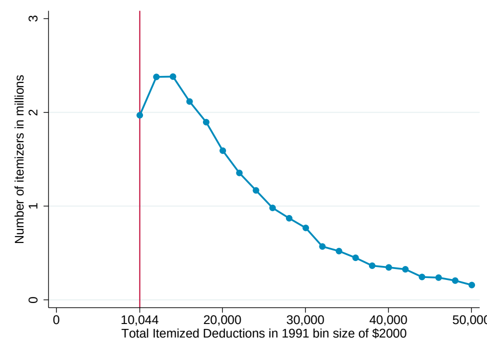
(c) 1988



(d) 1989



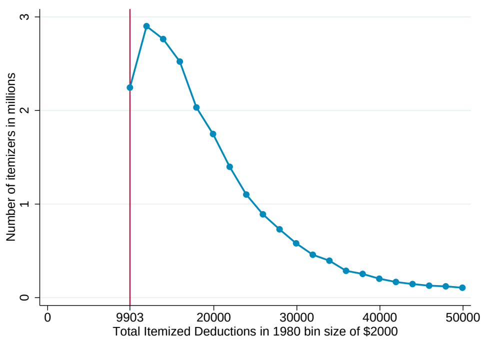
(e) 1990



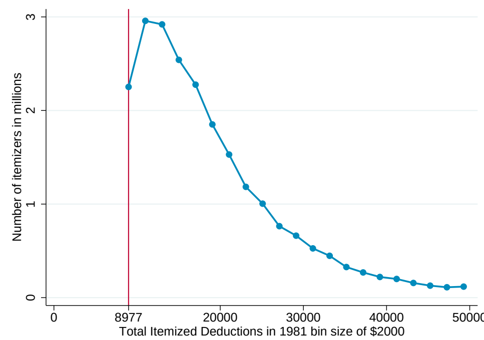
(f) 1991

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.

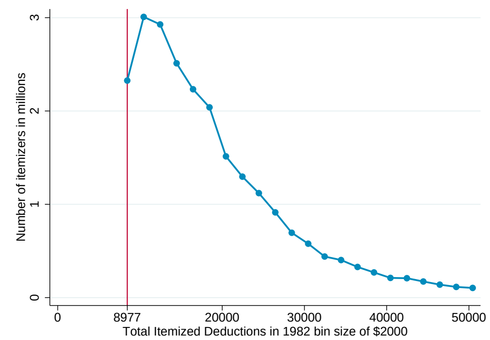
Figure H.12: Missing Mass Just Above the Standard Deduction 1980-1985 (Joint Filers)



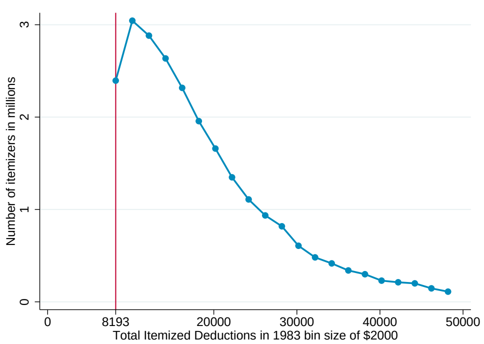
(a) 1980



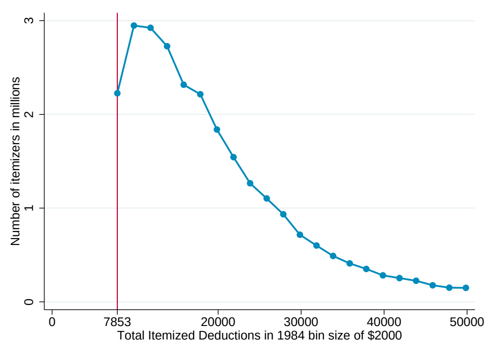
(b) 1981



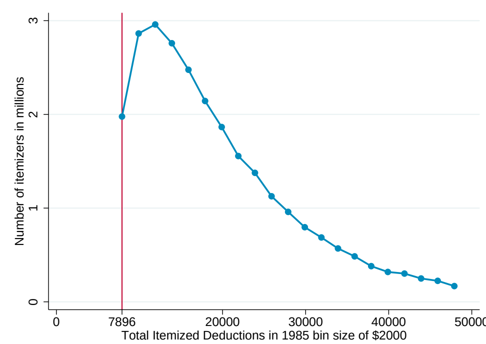
(c) 1982



(d) 1983



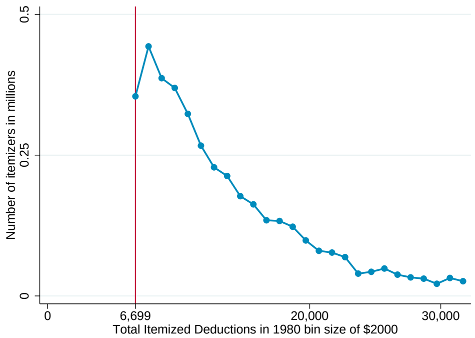
(e) 1984



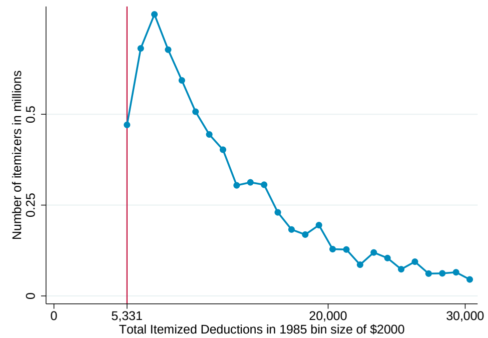
(f) 1985

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.

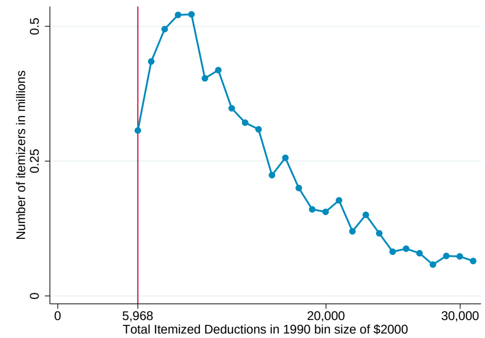
Figure H.13: Missing Mass Just Above the Standard Deduction (Single Filers)



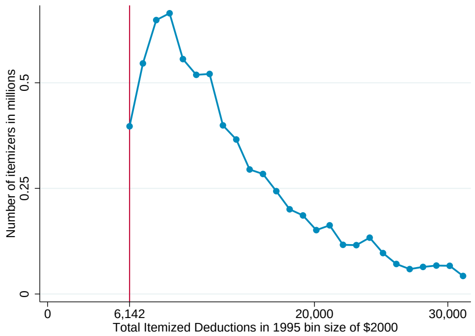
(a) 1980



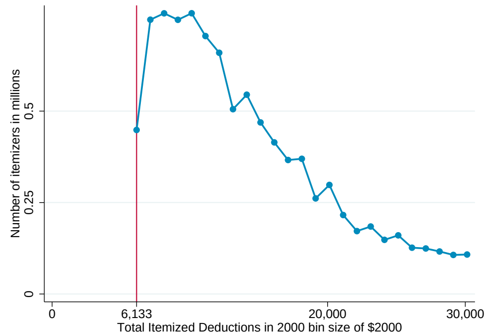
(b) 1985



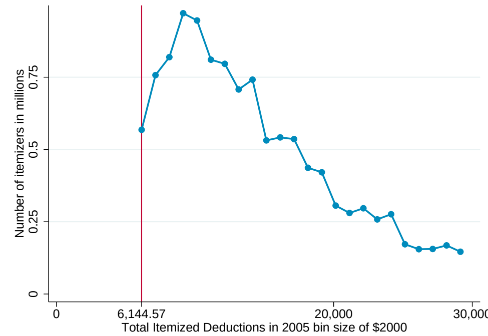
(c) 1990



(d) 1995



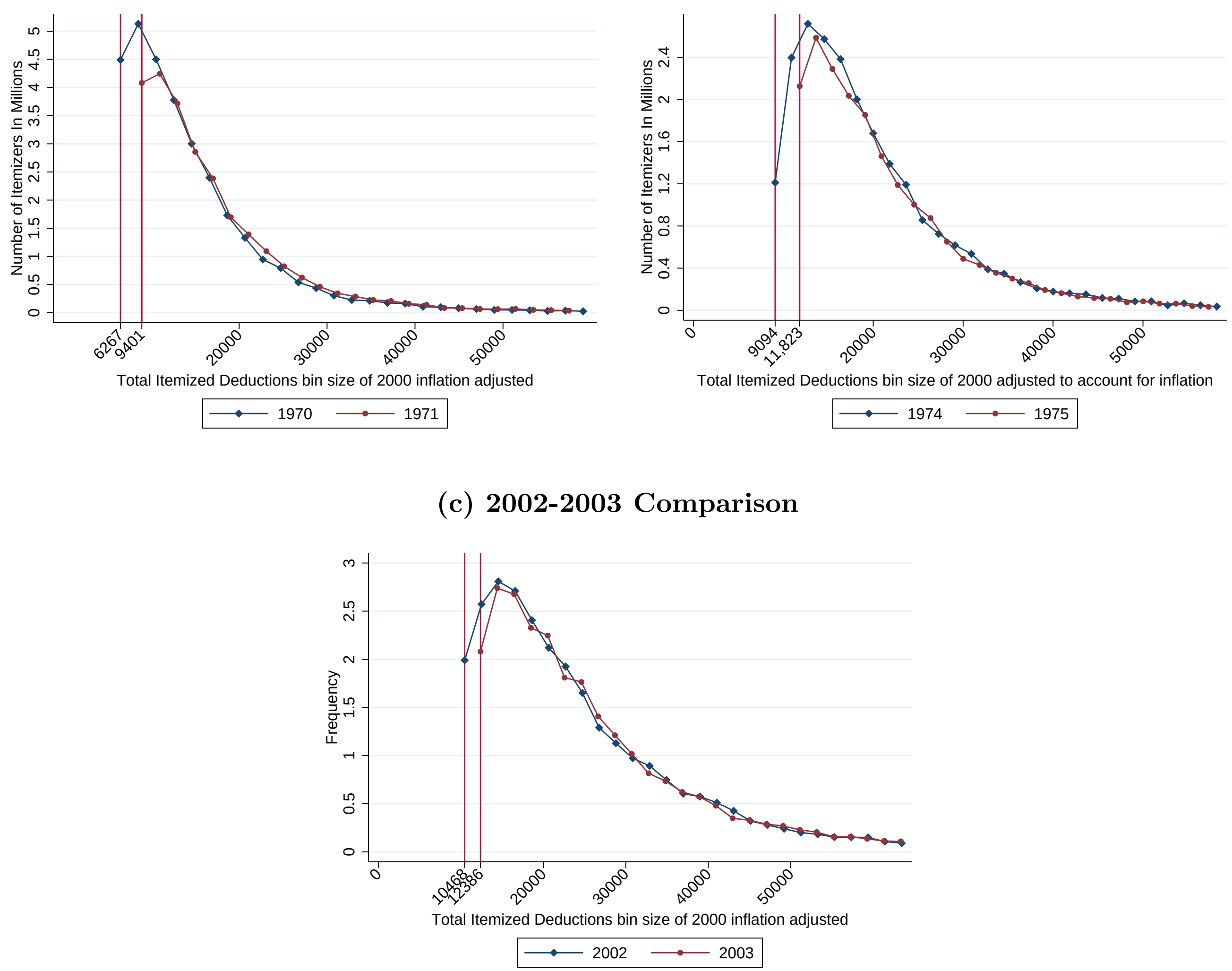
(e) 2000



(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.

Figure H.14: Reduced Form Evidence of the Existence of Compliance Costs



(a) 1970-1971 Comparison

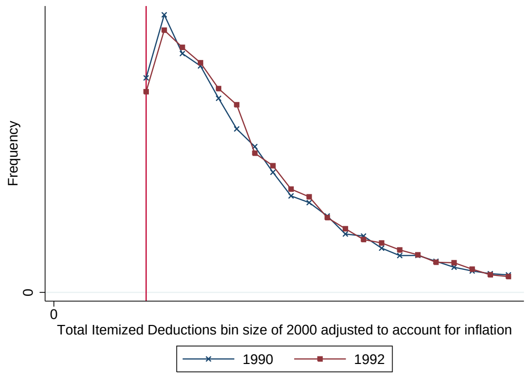
(b) 1974-1975 Comparison

(c) 2002-2003 Comparison

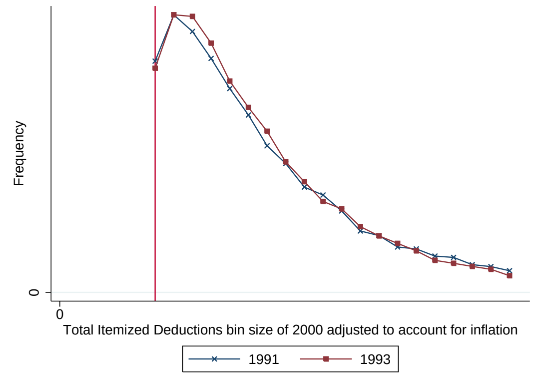
Notes: These graph plot the density of deductions before the 1971, 1975 and 2003 changes in the standard deduction amount. While these show reduced form evidence of the existence of compliance costs, they do not provide accurate estimates of these compliance costs because other changes occurred at the same time.

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

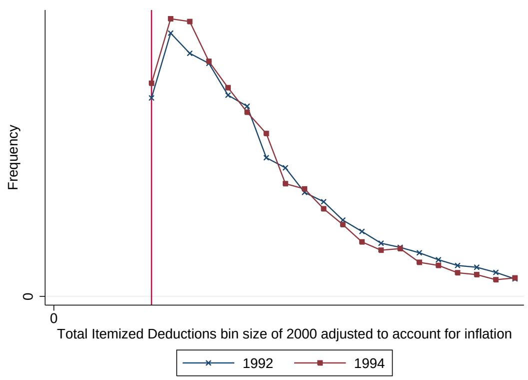
(a) 1990-1992



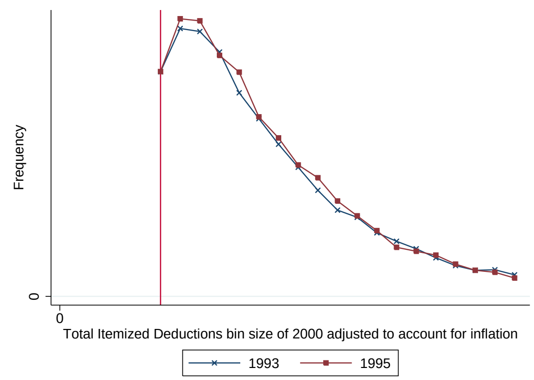
(b) 1991-1993



(c) 1992-1994



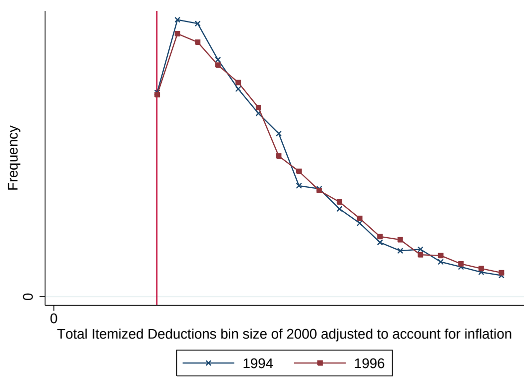
(d) 1993-1995



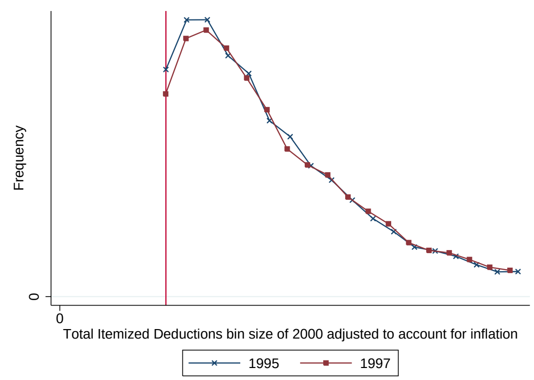
Notes: The Figures above test assumption A1 which states the cost of itemizing does not vary from year to year.

Figure H.16: Placebo Test: Overlapping Densities In Years With No Reforms

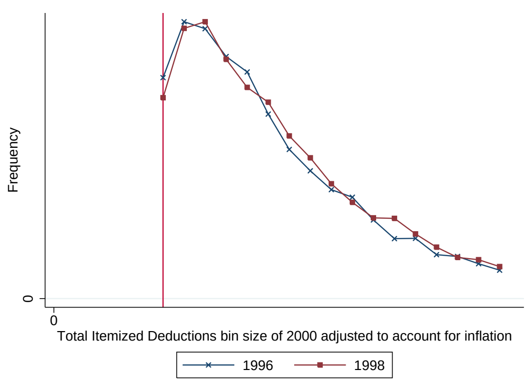
(a) 1994-1996



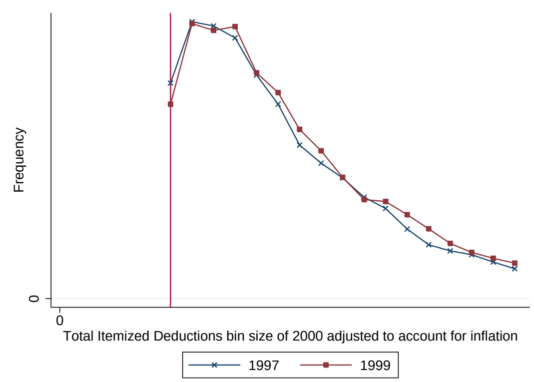
(b) 1995-1997



(c) 1996-1998



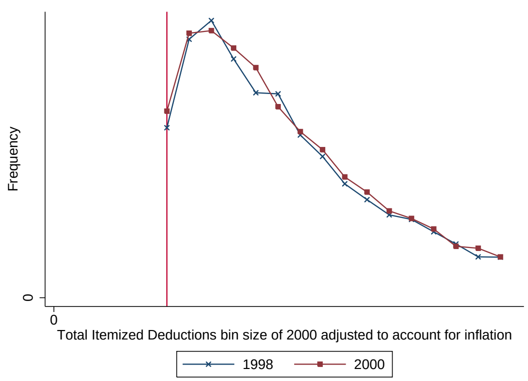
(d) 1997-1999



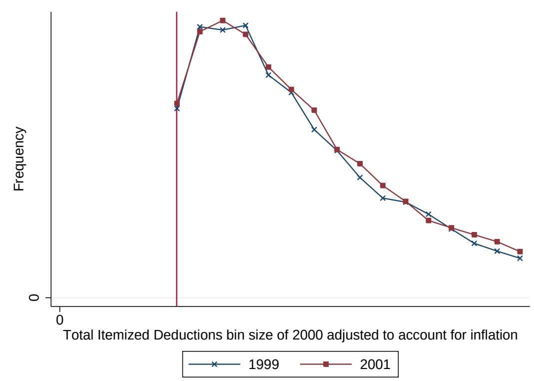
Notes: The Figures above test assumption A1 which states the cost of itemizing does not vary from year to year.

Figure H.17: Placebo Test: Overlapping Densities In Years With No Reforms

(a) 1998-2000

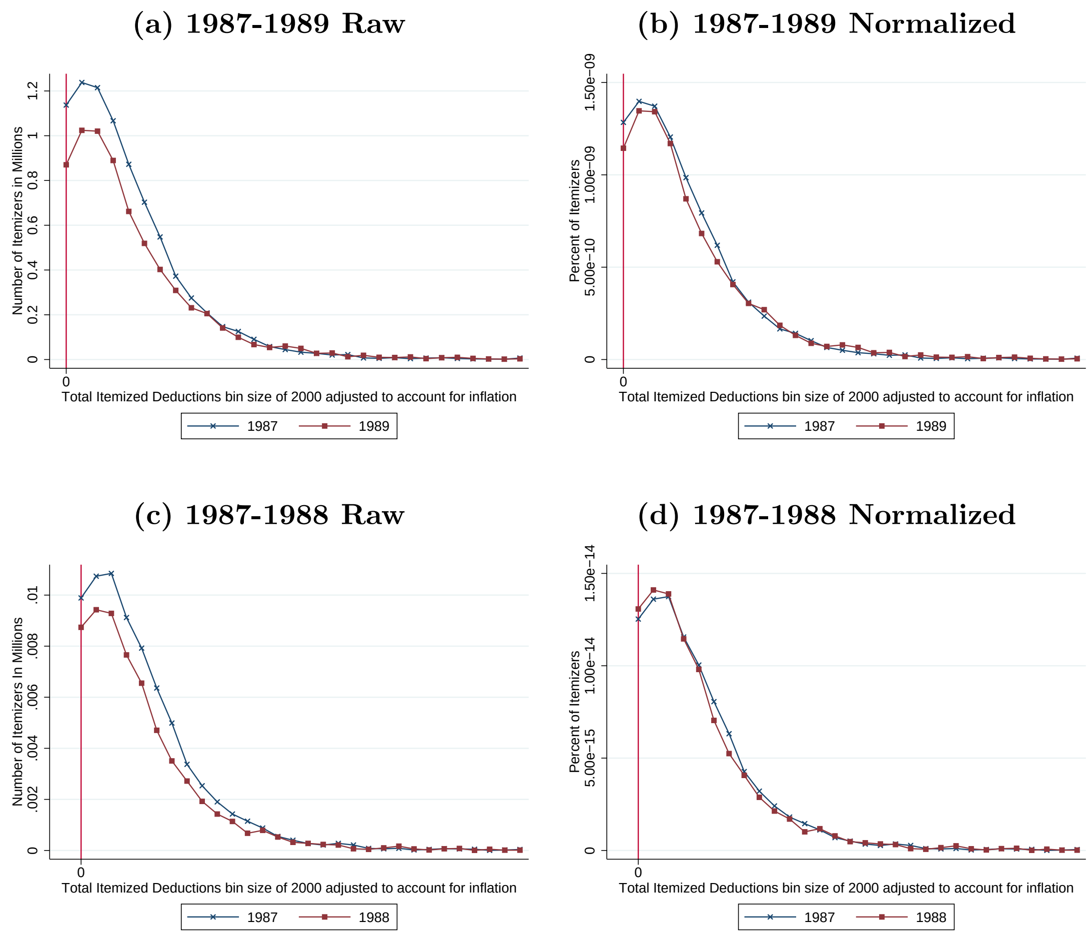


(b) 1999-2001



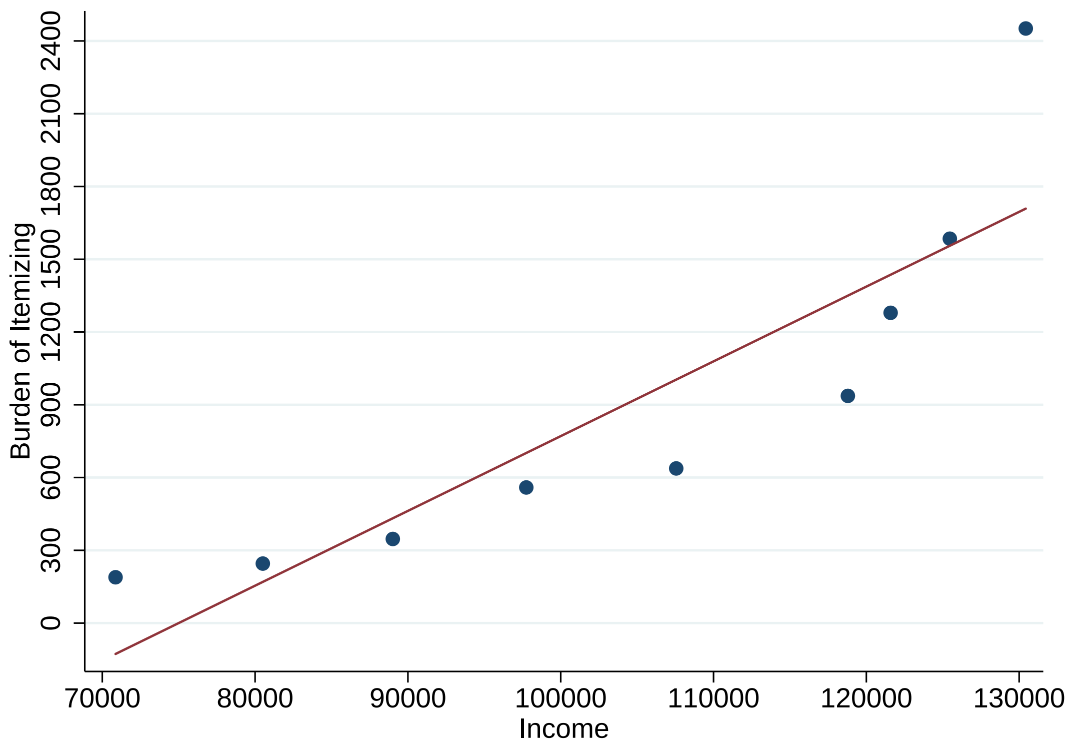
Notes: The Figures above test assumption A1 which states the cost of itemizing does not vary from year to year.

Figure H.18: Test of the Proportionality Assumption



Notes: Figures (a) and (c) plot the distribution of itemizers in 1987-1989 and 1987-1988, respectively, shifted to the left so as to have the same starting point as the pre-reform distribution. Figures (b) and (d) plot the the distribution of the number of itemizers per bin divided by the total number of itemizers with the same starting point for both 1987-1989 and 1987-1988, respectively.

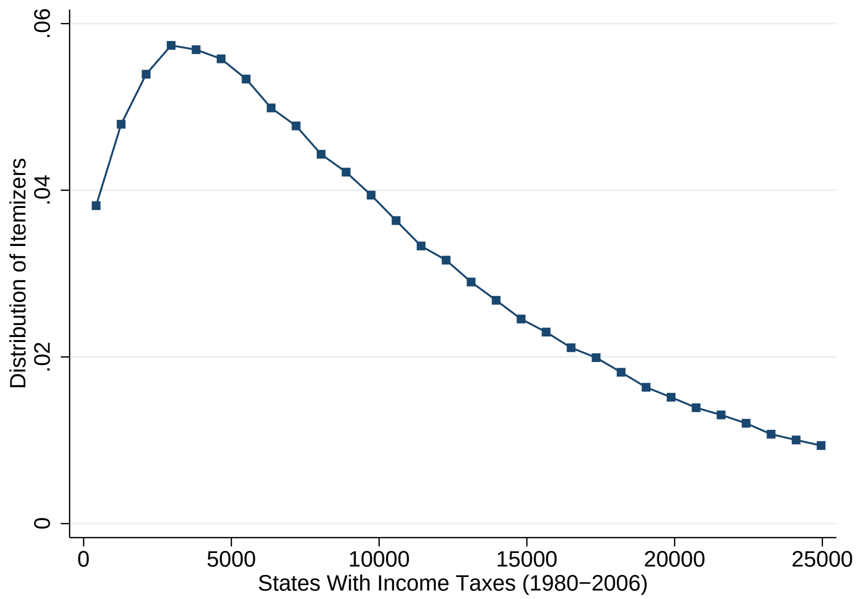
Figure H.19: Forgone Benefits Increase With Income



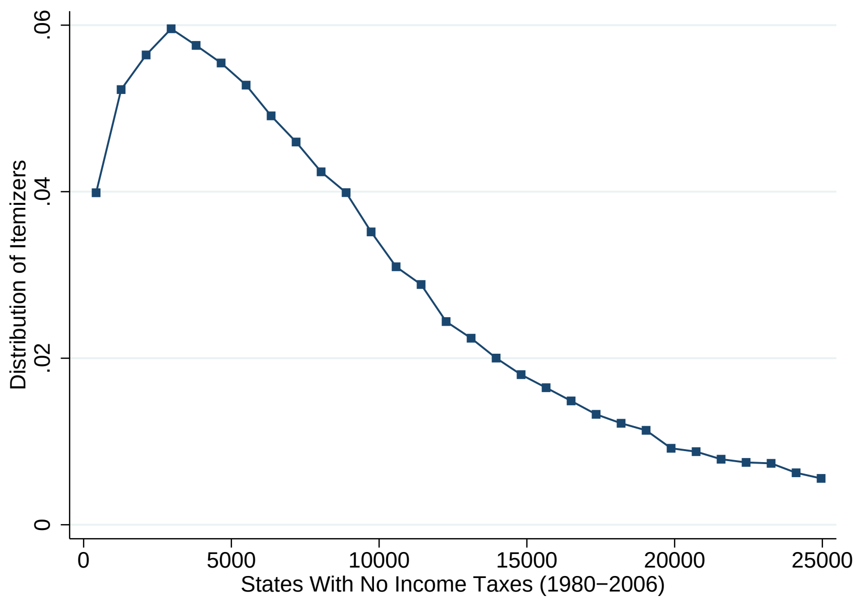
Notes: This graph plots the relationship between forgone benefits and income for all income deciles.

Figure H.20: States With and Without Income Taxes

(a) States With Income Taxes



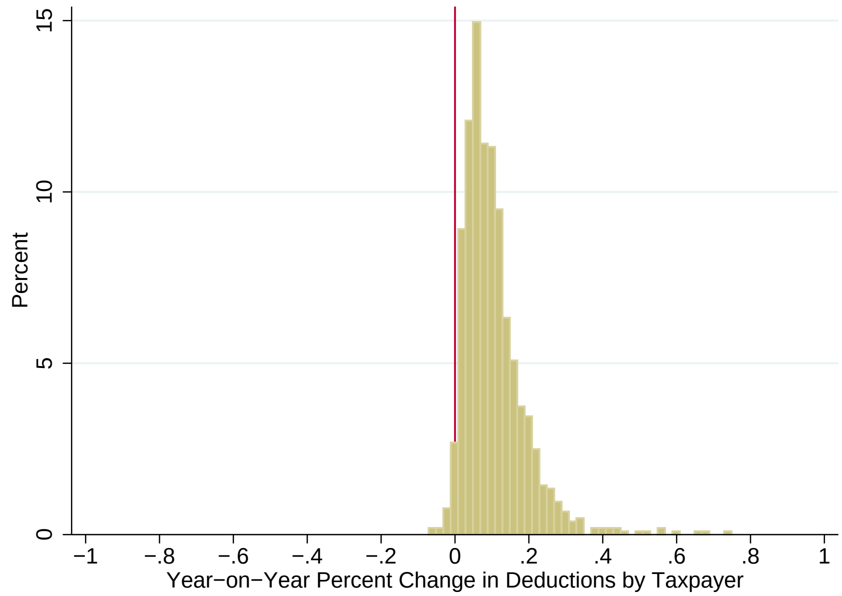
(b) States Without Income Taxes



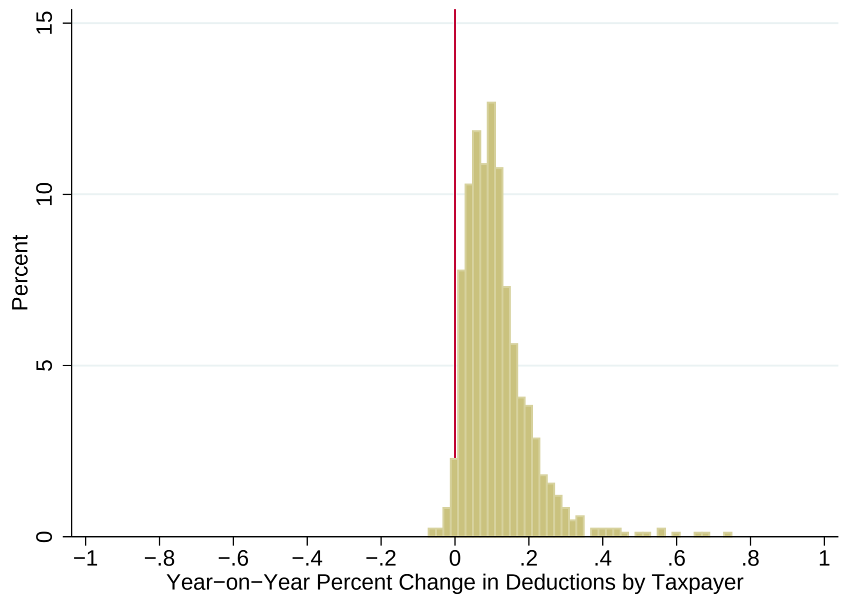
Notes: These Figures plot the distribution of itemizers in states with income taxes (panel a) and without income taxes (panel b). Both Figures pool years from 1980 to 2006.

Figure H.21: Stability of Deductions

(a) All Itemizers

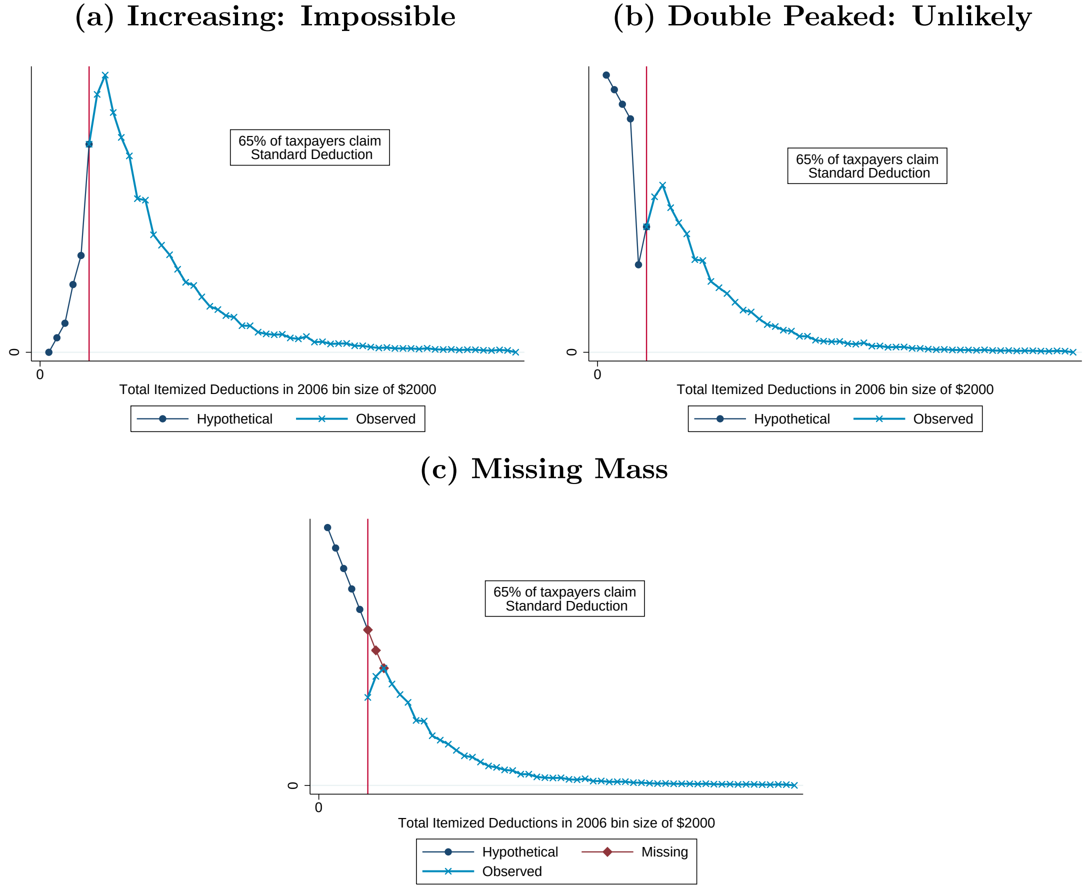


(b) Itemizers With Large Levels of Deductions



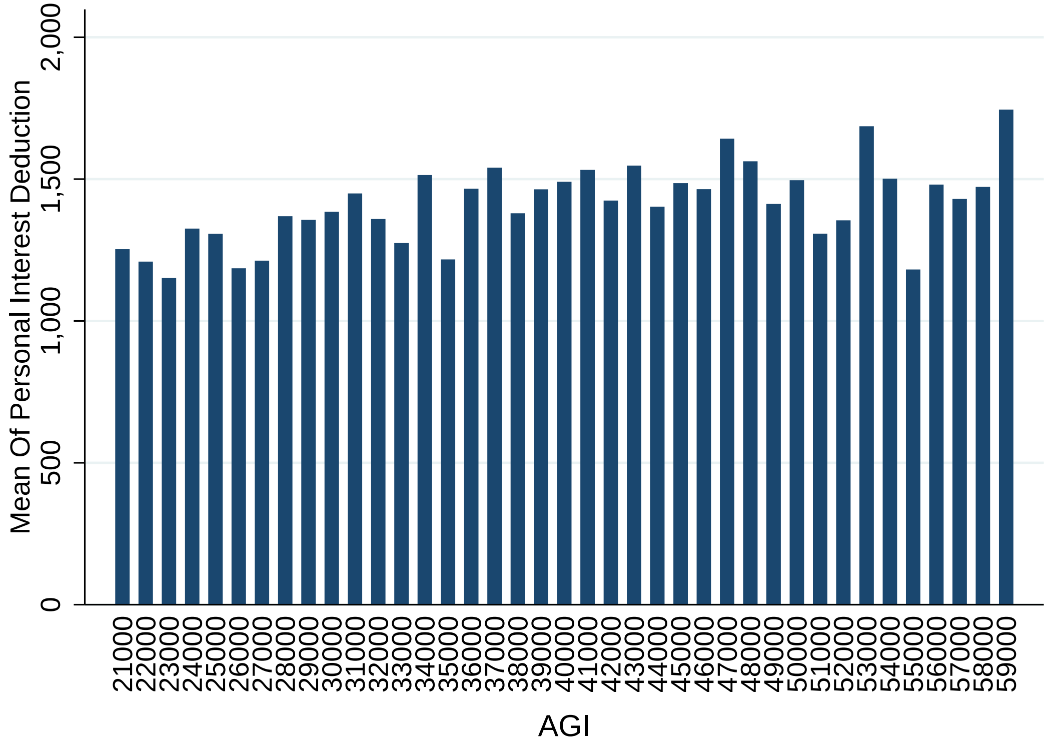
Notes: These graphs plot the distribution of year-on-year total deduction changes by taxpayer. Panel (a) considers all itemizers and panel (b) is restricted to itemizers with total deduction levels exceeding $15,000.

Figure H.22: Different Scenarios Below the Standard Deduction



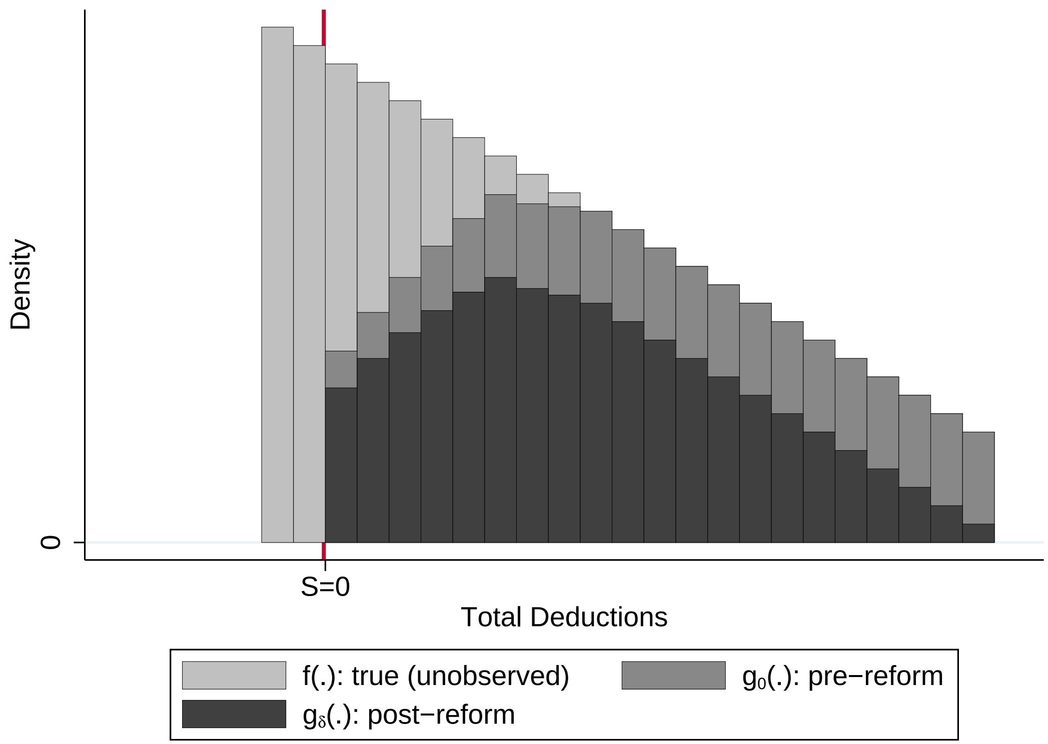
Notes: The graphs above plot the different scenarios that could be happening below the standard deduction. Graph (a) assumes that the density is strictly increasing, which is impossible given that 65% of taxpayers claim the standard deduction. This scenario would fail to account for most of the population 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 compliance costs creating a missing mass.

Figure H.23: 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.

Figure H.24: Shifted Pre- and Post-Reform Distributions



Notes: This graph shows hypothetical pre- and post-reform distributions that are shifted so as to have the same starting point.

I APPENDIX TABLES

Table I.2: 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	30.00%	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 cost of itemizing deductions are in bold.

Table I.3: 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

Notes: This Table shows the bootstrapped standard errors for the difference between bins in Figure 2a * denotes significance at the 10% level, ** at the 5% level and *** at the 1% level. I use 100 replications for the bootstrap estimation.

Table I.4: 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	0	0	0	1	5	28	70	152
0.25	0	0	0	2	11	44	167	203
0.5	0	0	1	4	22	86	64	625
0.8	0	0	1	5	35	134	462	880
1	0	1	2	7	44	164	547	1015
1.1	0	1	2	8	48	179	586	1074
1.25	0	1	2	9	54	200	640	1153
1.5	0	1	3	11	64	233	718	1262
1.8	0	1	3	13	76	270	799	1364
2	0	1	4	14	84	293	844	1417

Notes: This Table shows the results of a calibration of the rational inattention model derived in section 5.3.

Table I.5: Standard Errors of the Difference Between the 1970 and 1971 Densities (Figure H.14a)

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: This Table shows the bootstrapped standard errors for the difference between bins in 1970 and 1971 for taxpayers with deductions below $30,000. * denotes significance at the 10% level, ** at the 5% level and *** at the 1% level. I use 100 replications for the bootstrap estimation.

Table I.6: Standard Errors of the Difference Between the Density of Electronic Filers v.s. Paper Filers (Figure 4c)

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 I.7: Survey Based Estimates of the Compliance 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 I.8: IRS Hourly Cost Estimates

Form	Recordkeeping	Learning about the law or the form	Preparing the form	Copying, assembling and sending the form to the IRS	Total
1040	3 hrs., 7 min.	2 hrs., 32 min.	3 hrs., 10 min.	35 min.	9hrs., 24 min.
Sch. A	2 hrs., 47 min.	26 min.	1 hr., 1 min.	20 min.	4 hrs., 34 min.
Sch. B	33 min.	8 min.	16 min.	20 min.	1 hr., 17 min.
Sch. C	6 hrs., 13 min.	1 hr., 4 min.	1 hr., 56 min.	25 min.	9 hrs., 38 min.
Sch. D	1 hr., 2 min.	1 hr.	1 hr., 8 min.	35 min.	3 hrs., 45 min.
Sch. D-1	13 min.	1 min.	13 min.	35 min.	1 hr., 2 min.
Sch. E	2 hr., 52 min.	1 hr., 7 min.	1 hr., 16 min.	35 min.	5 hrs., 50 min.
Sch. F	9 hr., 41 min.	1 hr., 59 min.	3 hr., 52 min.	35 min.	16 hrs., 7 min.
Sch. R	20 min.	15 min.	22 min.	35 min.	1 hr., 32 min.
Sch. SE short	20 min.	11 min.	13 min.	14 min.	58 min.
Sch. SE long	26 min.	22 min.	37 min.	20 min.	1 hr., 45 min.

Notes: Each cell of this Table is an estimate of the time it takes to perform each task associated with each tax schedule. They are based on IRS surveys of taxpayers at the time of filing and are reported in the 1040 instructions (on page 3 in 1989). There is no information on Sch. R in the SOI public use files so its cost is not estimated in this paper.