Youssef Benzarti

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

Online Appendix

APPENDIX 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.

APPENDIX B Cost Estimation

A 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 Simulated Example

The following example illustrates the estimation process outlined in Section III.A. 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 III.A, 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 III.A, 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.

APPENDIX C Sample Restrictions

A *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.

B *Figures 2a, 2b*

In Figure 2a and 2b, I focus on taxpayers who are married filing jointly for the reasons outlined in section C.A. 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 *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.

D *Figures 4c and 4b*

To generate Figures 4c and 4b, I consider joint filers as explained in section C.A. 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.

E 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.

F 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.

APPENDIX 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.

APPENDIX E The 1971, 1975 and 2003 reforms

A 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).

B 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.

APPENDIX 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)

APPENDIX 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:

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

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

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

of deductions with a 28% marginal tax rate). This means that the taxpayer has a range of uncertainty of deductions of more than \$6,479. This implies very high uncertainty in the beliefs of the benefits that the taxpayer can save from itemizing which is unlikely given that deductions are relatively stable from year to year as they are mostly constituted of mortgage payments and state taxes and are the results of active decisions. If a taxpayer's total deductions were to increase or decrease dramatically, she would most likely know about it because it would be due to for example to large income variations, the take up of a mortgage etc. which are salient.

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

APPENDIX 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:

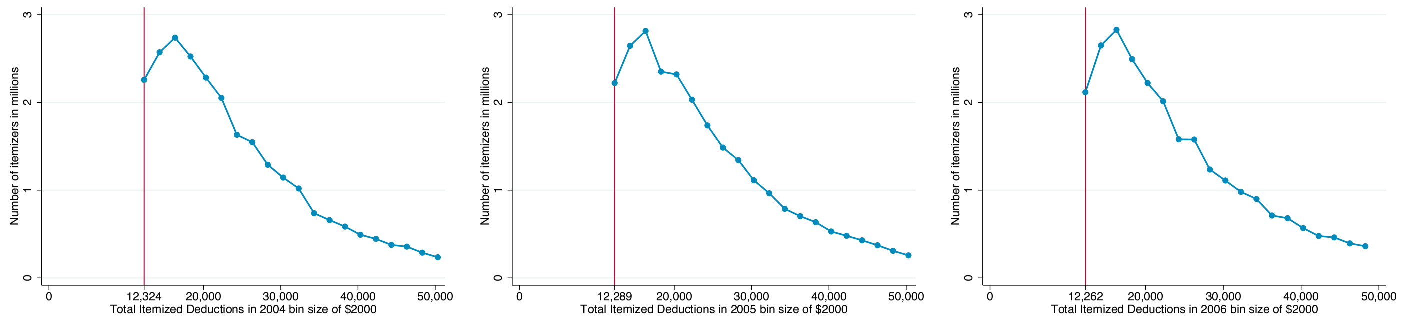
$$(1+r)\beta < 1. \tag{10}$$

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.

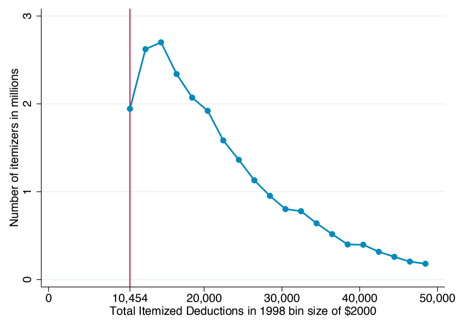
Testable Prediction 2: The cost of record keeping for naive present-biased taxpayers is greater than for rational ones. This predicts that taxpayers who file close to the deadline are likely to forgo more deductions.



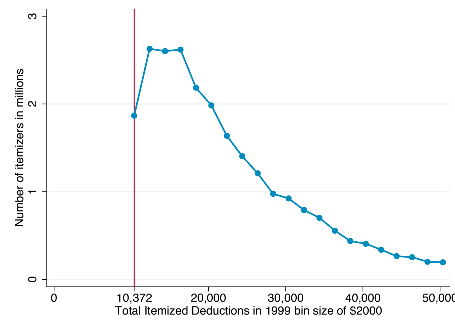
(a) 2004 (b) 2005 (c) 2006

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

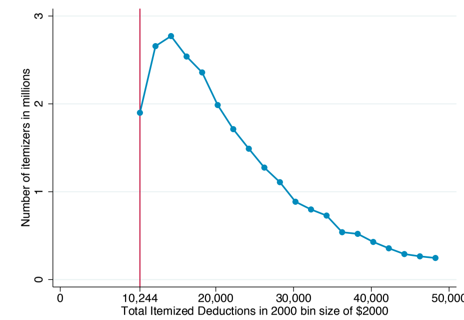
Note: The Figures above plot the density of deductions for itemizers filing jointly. The bin size is $2,000 and the vertical line represents the standard deduction threshold for each year.



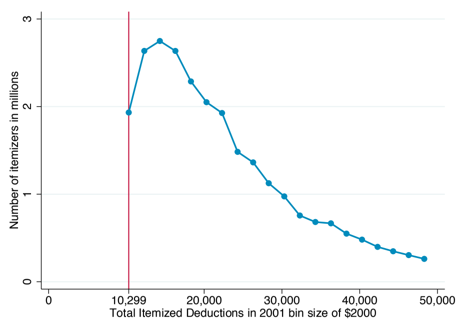
(a) 1998



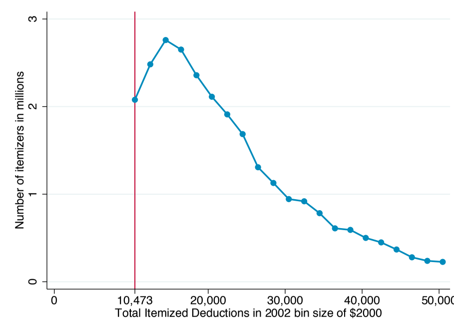
(b) 1999



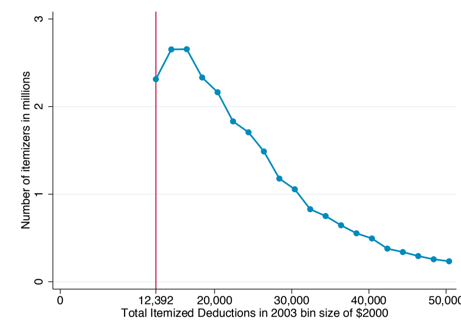
(c) 2000



(d) 2001



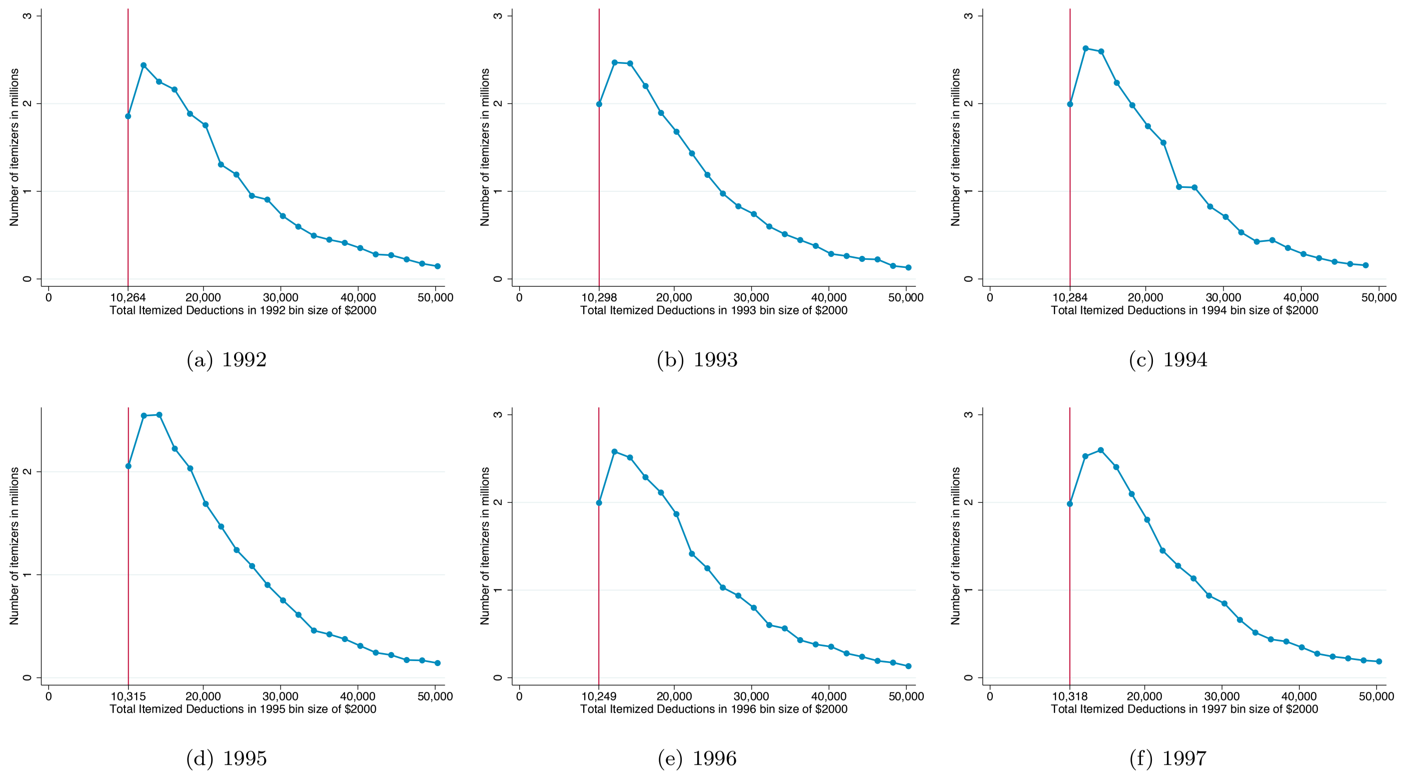
(e) 2002



(f) 2003

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

Note: 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.

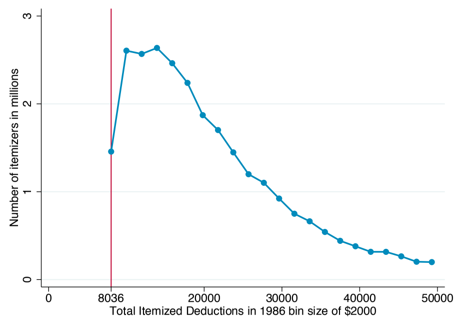


(a) 1992 (b) 1993 (c) 1994

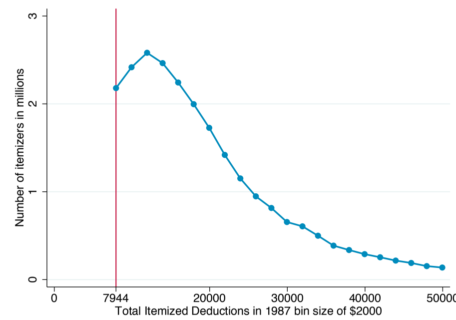
(d) 1995 (e) 1996 (f) 1997

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

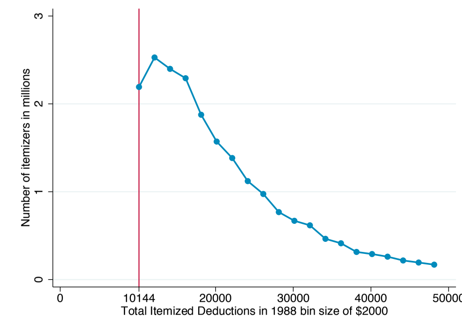
Note: The Figures above plot the density of deductions for itemizers filing jointly. The bin size is $2,000 and the vertical line represents the standard deduction threshold for each year.



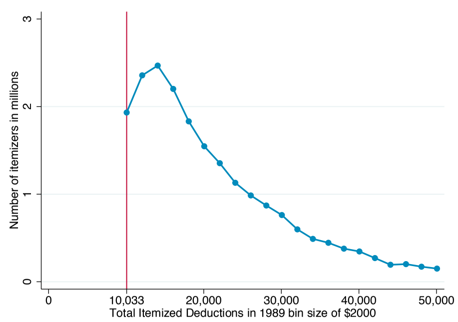
(a) 1986



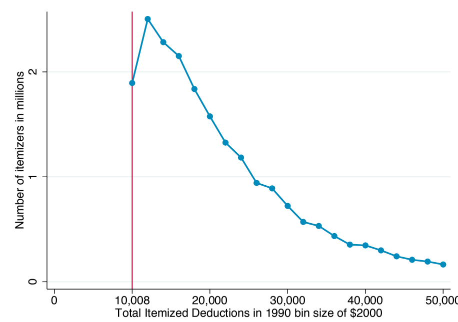
(b) 1987



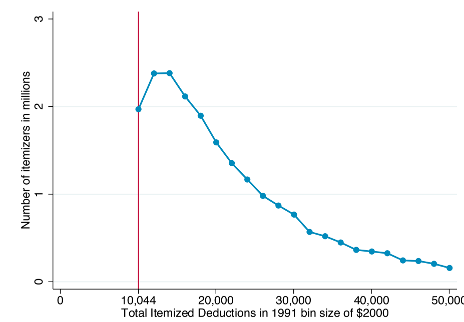
(c) 1988



(d) 1989



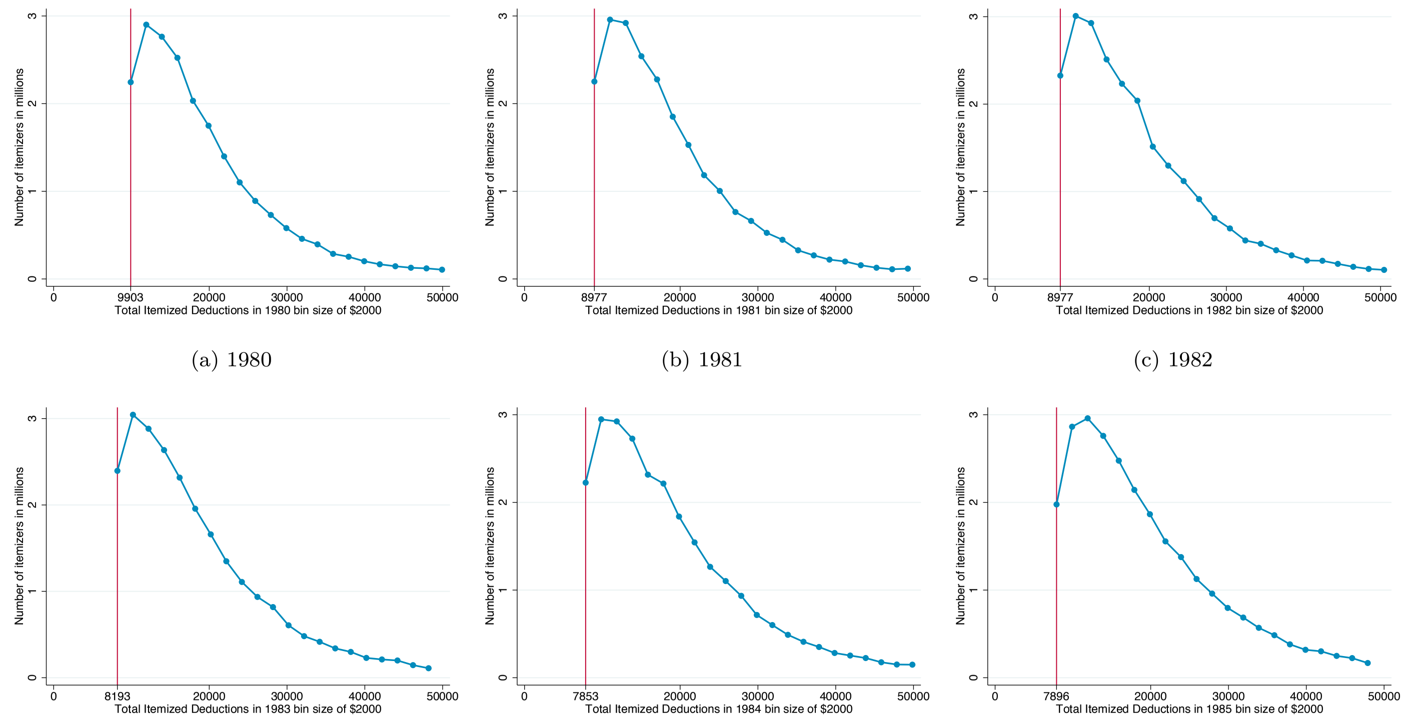
(e) 1990



(f) 1991

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

Note: 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.



(a) 1980 (b) 1981 (c) 1982

(d) 1983 (e) 1984 (f) 1985

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

Note: 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.

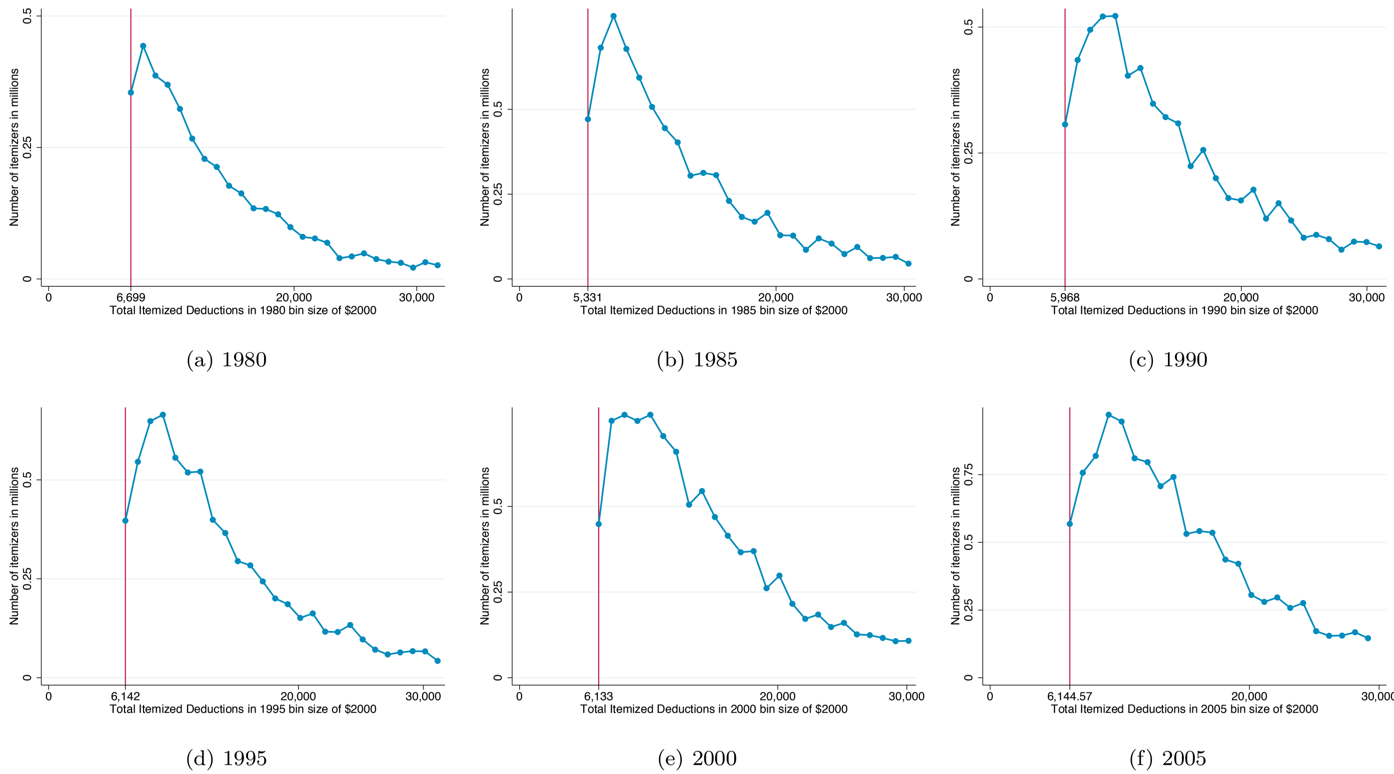
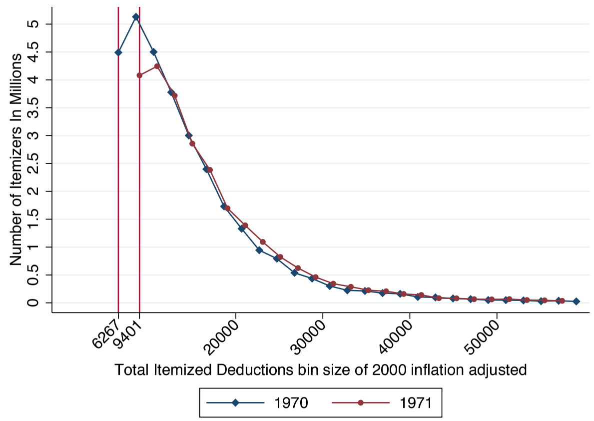
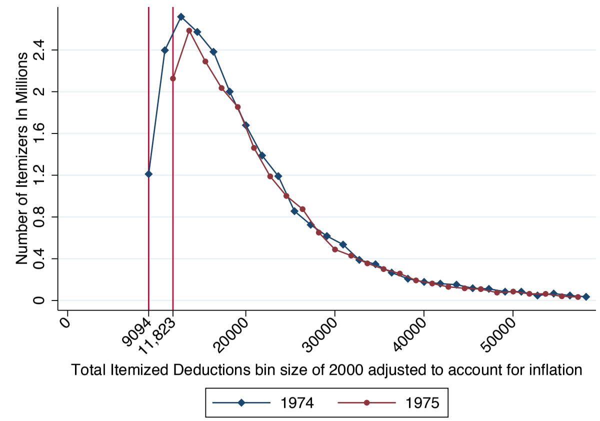


Figure H.13. : Missing Mass Just Above the Standard Deduction (Single Filers)

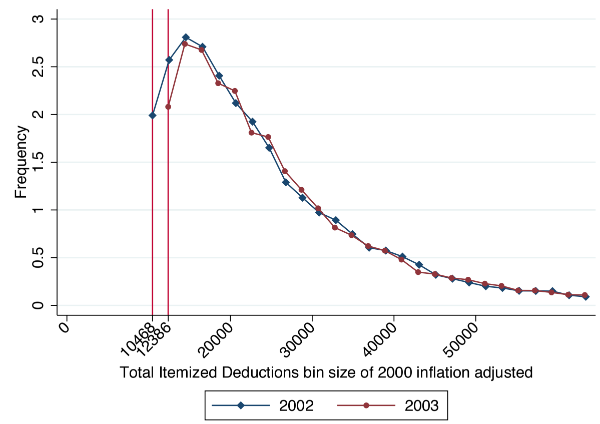
Note: The Figures above plot the density of deductions for single filers who itemize deductions. The bin size is $2,000 and the vertical line represents the standard deduction threshold for each year.



(a) 1970-1971 Comparison



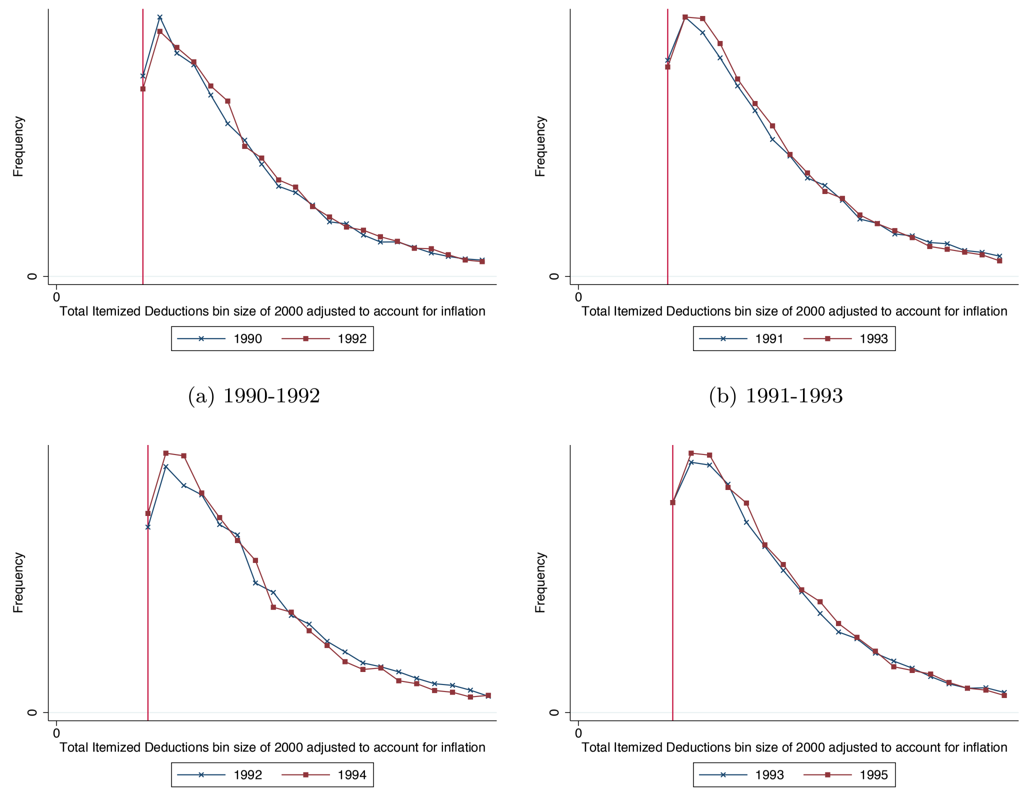
(b) 1974-1975 Comparison



(c) 2002-2003 Comparison

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

Note: 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.



(a) 1990-1992

(b) 1991-1993

(c) 1992-1994

(d) 1993-1995

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

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

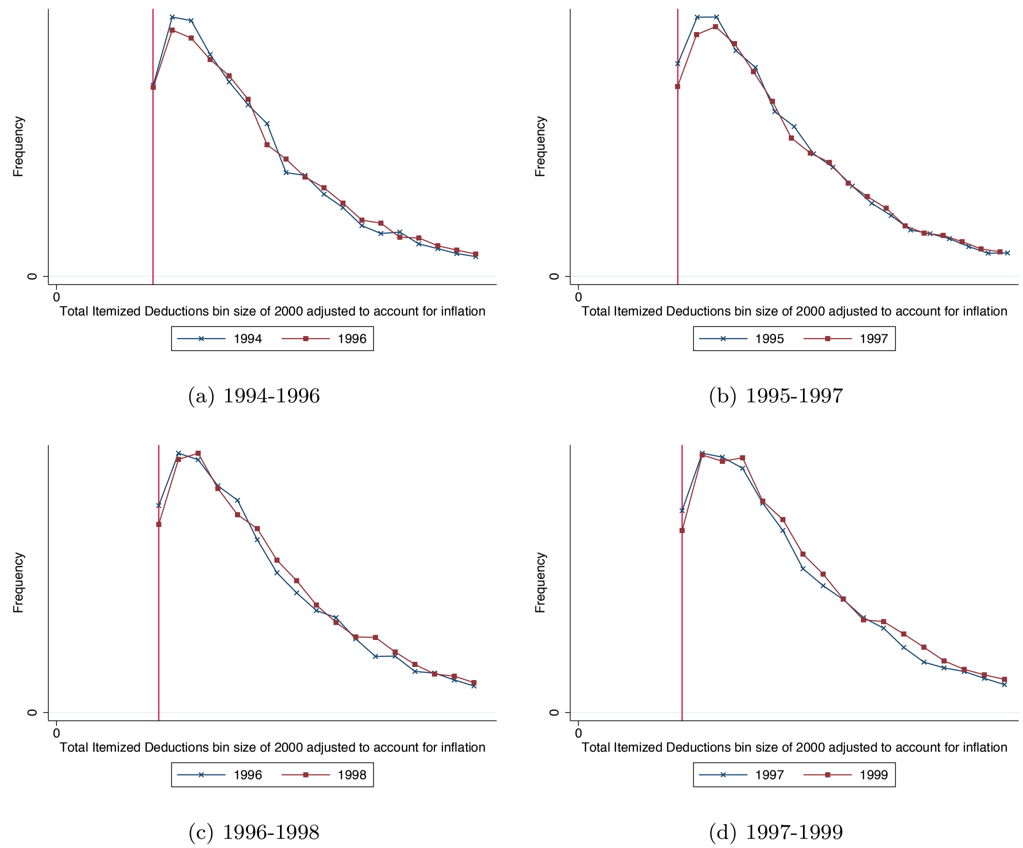
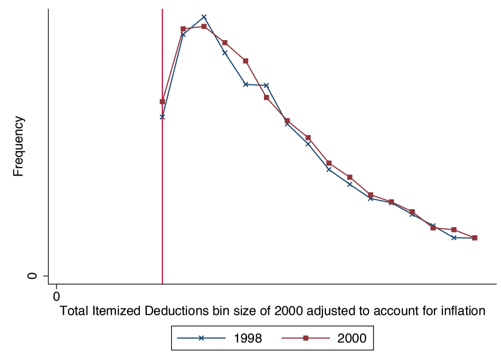
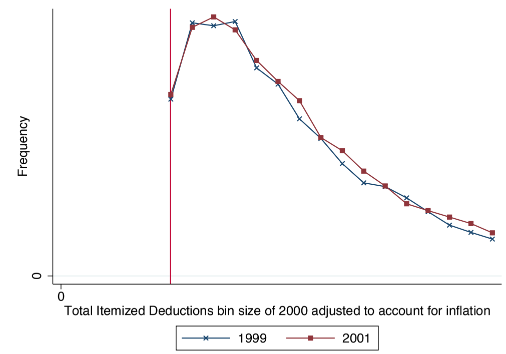


Figure H.16. : Placebo Test: Overlapping Densities In Years With No Reforms

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



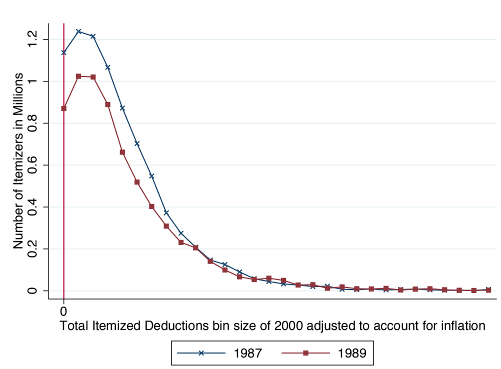
(a) 1998-2000



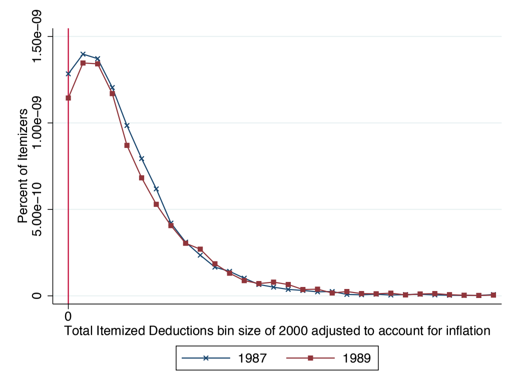
(b) 1999-2001

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

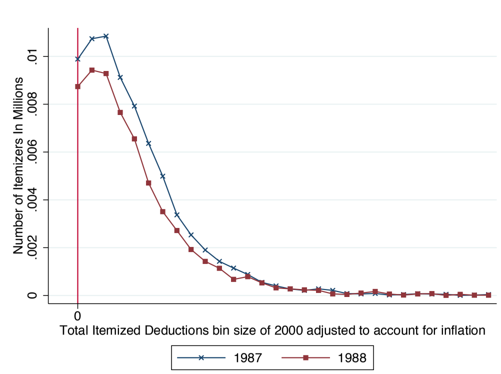
Note: Notes: The Figures above test assumption A1 which states that the cost of itemizing does not vary from year to year.



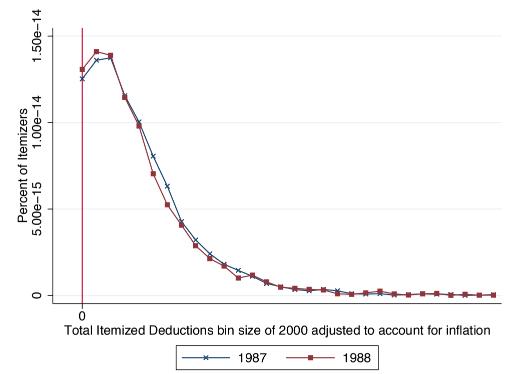
(a) 1987-1989 Raw



(b) 1987-1989 Normalized



(c) 1987-1988 Raw



(d) 1987-1988 Normalized

Figure H.18. : Test of the Proportionality Assumption

Note: 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.

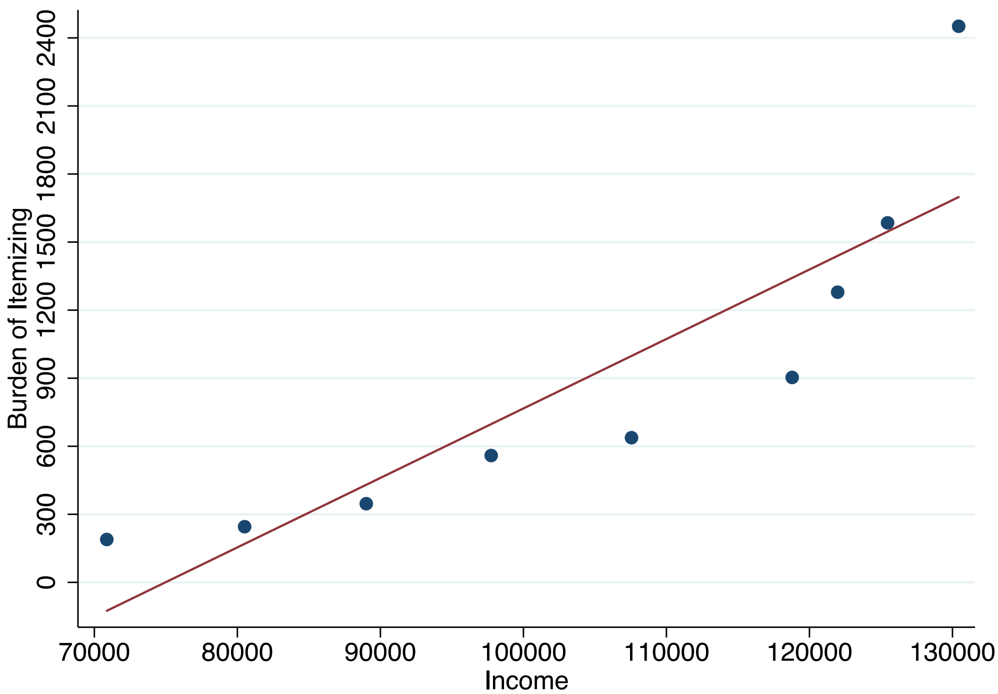
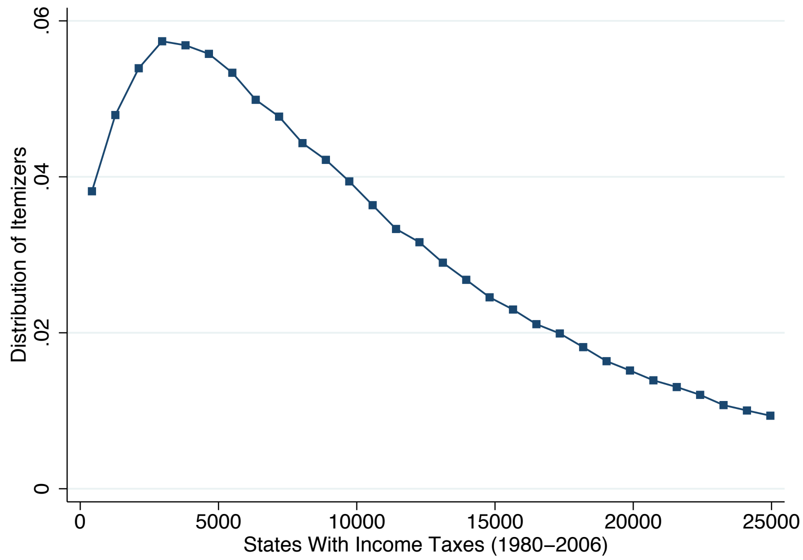
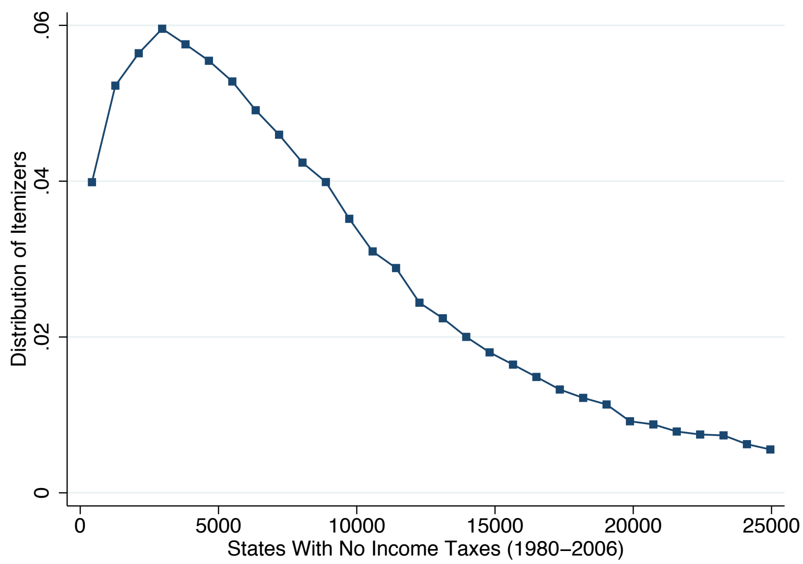


Figure H.19. : Forgone Benefits Increase With Income

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



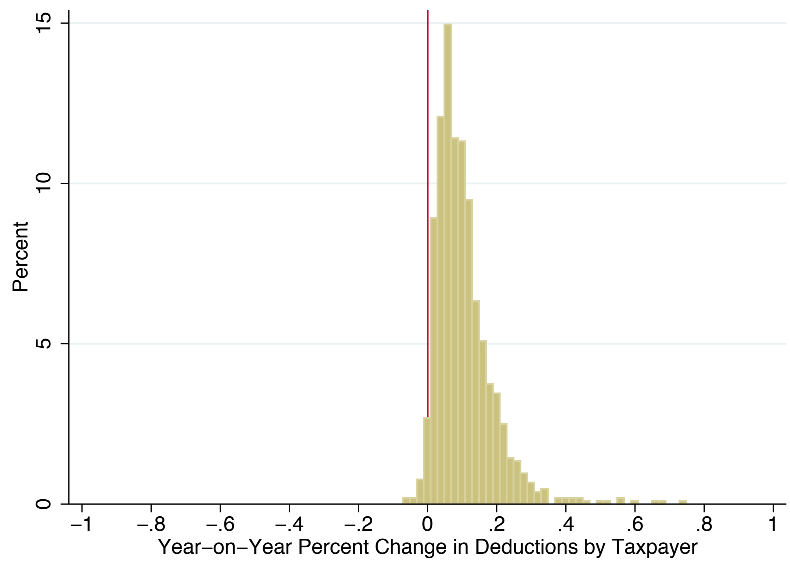
(a) States With Income Taxes



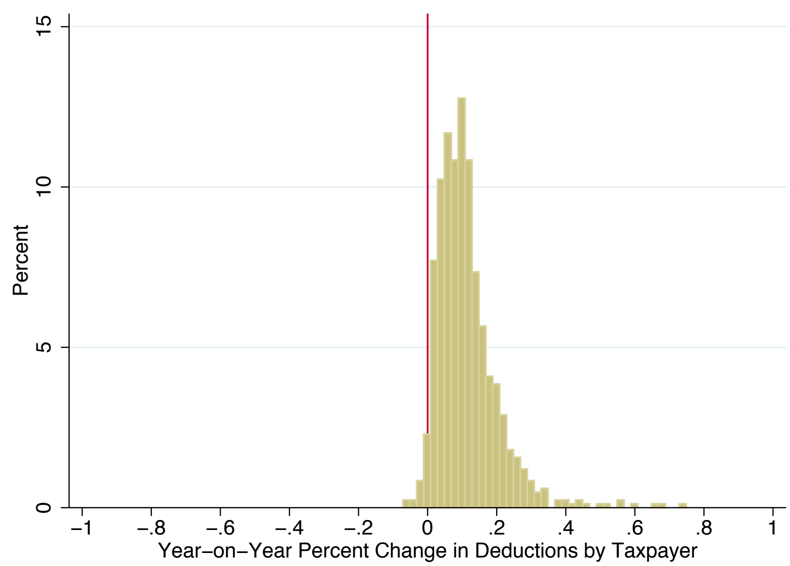
(b) States Without Income Taxes

Figure H.20. : States With and Without Income Taxes

Note: 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.



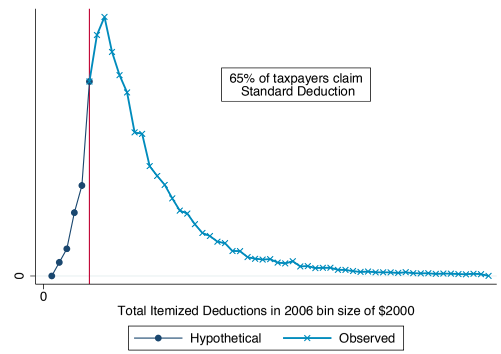
(a) All Itemizers



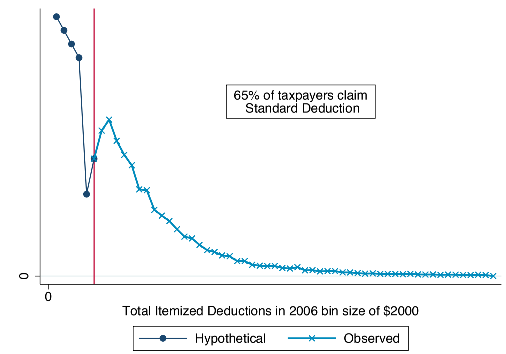
(b) Itemizers With Large Levels of Deductions

Figure H.21. : Stability of Deductions

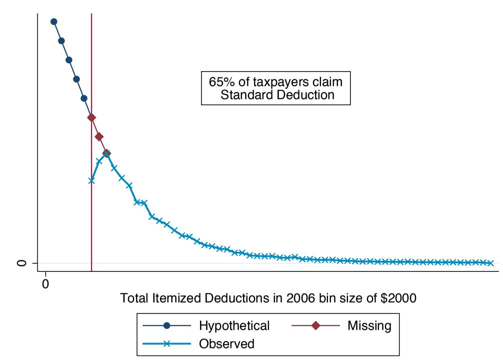
Note: These graphs plot the distribution of year-on-year total deduction changes by taxpayer. Panel (a) considers all itemizers and panel (b) is restricted to itemizers with total deduction levels exceeding \$15,000.



(a) Increasing: Impossible



(b) Double Peaked: Unlikely



(c) Missing Mass

Figure H.22. : Different Scenarios Below the Standard Deduction

Note: 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.

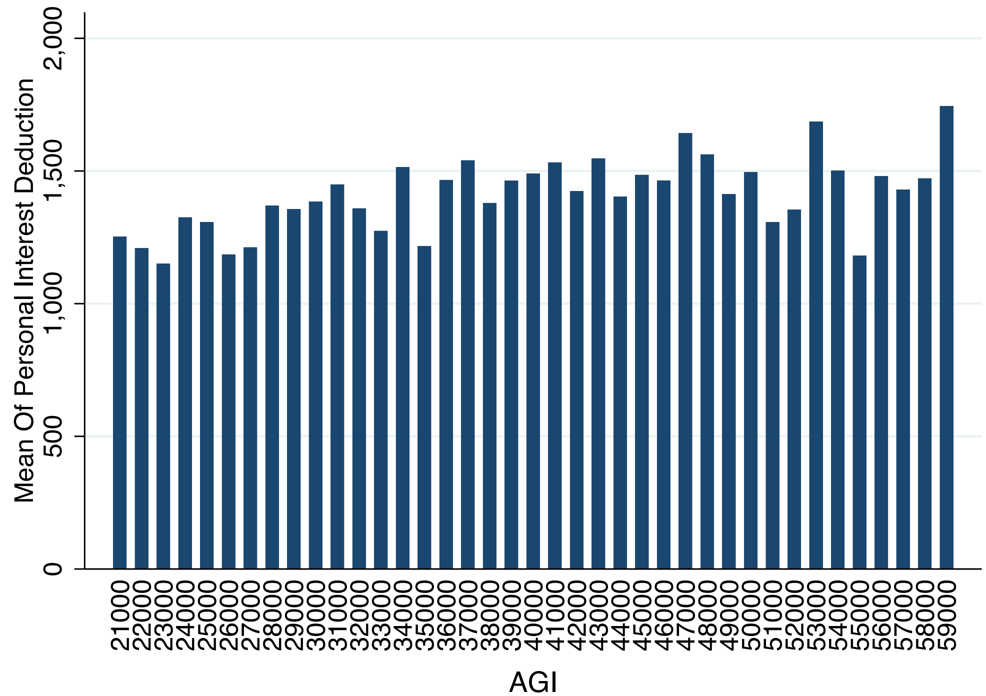


Figure H.23. : No Behavioral Response For Personal Interest Deduction

Note: 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.

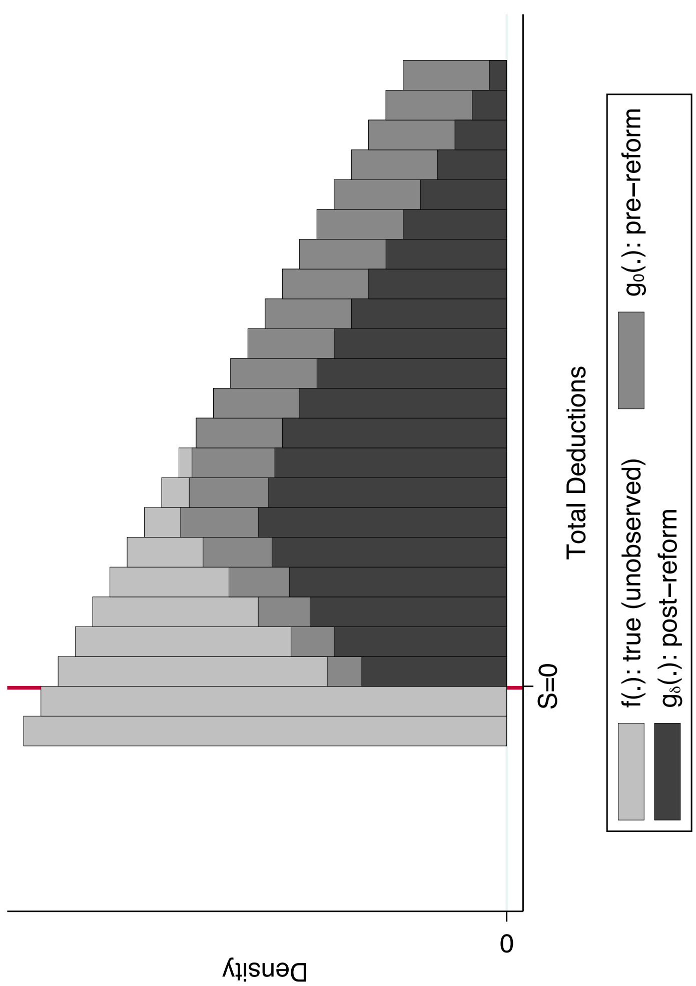


Figure H.24. : Shifted Pre- and Post-Reform Distributions

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

APPENDIX 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%

Note: 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]	0.0032744	0.0004801	6.82
2	(11991, 13991]	0.0015767	0.0005178	3.04
3	(13991, 15991]	0.0001276	0.0004963	0.26
4	(15991, 17991]	-0.0004995	0.0004008	-1.25
5	(17991, 19991]	0.0001943	0.0004058	0.48
6	(19991, 21991]	-4.63e-06	0.0003106	-0.01
7	(21991, 23991]	-0.0005501	.000266	-2.07
8	(23991, 25991]	-0.0003188	0.0002258	-1.41
9	(25991, 27991]	-0.0003171	0.0002144	-1.48
10	(27991, 29991]	-0.0005612	0.0001788	-3.14

Note: This Table shows the bootstrapped standard errors for the difference between bins in Figure 2a. 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

Note: This Table shows the results of a calibration of the rational inattention model derived in section V.C.

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]	0.0031746	0.0008349	3.80
2	(9140, 12140]	0.0019828	0.0007675	2.58
3	(12140, 15140]	0.0014309	0.0007114	2.01
4	(15140, 18140]	0.0024057	0.0005395	4.46
5	(18140, 21140]	0.0010975	0.0004993	2.20
6	(21140, 24140]	0.000593	0.0003337	1.78
7	(24140, 27140]	0.0002313	0.0003289	0.70
8	(27140, 30140]	0.000045	0.0002861	0.16
9	(30140, 33140]	0.0001287	0.000249	0.52
10	(33140, 36140]	-0.0001114	0.0001671	-0.67

Note: This Table shows the bootstrapped standard errors for the difference between bins in 1970 and 1971 for taxpayers with deductions below $30,000. 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

Note: 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.

Note: 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.