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

Youssef Benzarti*
University of California, Berkeley

March 25, 2015

Abstract

I use a quasi-experimental design to estimate the burden of complying with the tax code. Employing a sample of US income tax returns, I observe the preferences of taxpayers when choosing between itemizing deductions and claiming the standard deduction. Taxpayers forego tax savings to avoid the hassle cost of itemizing, resulting in an average burden of itemizing of $617, with substantial heterogeneity. A revealed preference argument implies that itemizing deductions is as painful as working 19 hours. The burden of compliance is larger for richer households, consistent with the fact that the value of time increases with income. I explore two explanations for the result. First, it could be due to an extreme aversion to filing taxes. Such aversion implies that itemizing deductions imposes an aggregate compliance cost of 0.20% of GDP and back-of-the-envelope extrapolations to filing federal taxes yields an overall compliance cost of 1.25% of GDP. Second, if taxpayers are time inconsistent the axiom of revealed preferences fails, introducing a wedge between foregone benefits and compliance costs. Being present-biased leads taxpayers to forego large benefits even when compliance costs are relatively small. I provide evidence of taxpayers being present-biased. Both explanations – whether driven by preferences or mistakes – suggest that the burden imposed by tax compliance is significantly larger than previously estimated. I discuss policy implications of the result.

JEL Classification: D03, H24, H31, H83, M41
Keywords: Behavioral, Preferences, Subsidies, Household Economics, Record Keeping

*benzarti@econ.berkeley.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 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, Colin Raymond, Antonio Rosato, Jesse Rothstein, Josh Schwartzstein, Dan Silverman, Joel Slemrod, Charles Sprenger, Justin Sydnor, Alisa Tazhitdinova, Danny Yagan, Erick Zwick and seminar participants at UC Berkeley, the Tax Systems Conference (Oxford and U. of Michigan), Paris II and CREST. Funding from the Robert D. Burch Center for Tax Policy and Public Finance is gratefully acknowledged.

"Death, taxes and childbirth! There's never any convenient time for any of them."
Margaret Mitchell, *Gone with the Wind*

1 Introduction

While there is a long tradition in public finance of assessing the magnitude of the efficiency cost of taxation, very little attention has been given to the burden of complying with taxes. Every year, more than 140 million taxpayers have to file taxes in the US. With the tax code becoming increasingly complex, taxpayers have to spend a significant amount of time filling out the 1040 form, various schedules and keeping records of their transactions. Some individuals fail to take up benefits such as the Earned Income Tax Credit (EITC) or Unemployment Insurance[1] (UI). Some explanations of this phenomenon were investigated – such as stigma or lack of information – but hassle costs were rarely addressed. How large is the burden of compliance and are taxpayers foregoing benefits because of it? I answer this question by observing the choice of individuals over two tasks offering a trade-off between compliance costs and benefits. Itemizing deductions requires some effort cost but can provide large tax savings. Claiming the standard deduction saves time and effort but results in more taxes due. I show that taxpayers could save money by itemizing and are aware of it but still claim the standard deduction. I then estimate the burden of compliance by using the observed preferences of taxpayers over these two tasks.

With no compliance costs, taxpayers should itemize if the benefit of itemizing is greater than zero. With compliance costs, itemizing is only beneficial 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 is greater than the standard deduction amount. The main identification challenge is to differentiate between individuals who fail to itemize deductions because of compliance costs from individuals who claim the standard deduction because their total deductions are smaller than the standard deduction amount. This is particularly difficult because taxpayers who claim the standard deduction are not required to report their deductions, implying that their true level of deductions is not observable in tax data.

If individuals are truly foregoing tax benefits because of compliance costs, there should be a

[1] See for example Blank and Card (1991).

missing mass in the density of deductions immediately to the right of the standard deduction threshold. I test this hypothesis by graphing the density of deductions for years ranging from 1980 to 2003 using a representative sample of US tax returns. The shape of the density function suggests the presence of a missing mass in the neighborhood of the standard deduction. To confirm that this shape is due to taxpayers claiming the standard deduction when they could benefit from itemizing, I turn to a quasi-experimental design. Following an increase in the standard deduction amount in 1971 and 1988, I observe a drop in the mass of itemizers in the neighborhood of the post-reform standard deduction threshold. The post-reform density is systematically lower than the pre-reform one in the neighborhood of the post-reform standard deduction threshold and the two densities overlap further away from the standard deduction. I ensure that no other reforms are affecting the densities of itemized deductions.[2] I also explore alternative explanations such as lack of information about the possibility of itemizing, rational inattention, fear of being audited by the IRS, behavioral responses to a concave kink point and conclude that they cannot explain the result.

I use the missing mass to construct the distribution of foregone benefits. I find significant heterogeneity among taxpayers. Some taxpayers still itemize even when savings are modest and some forego large tax benefits, resulting in an average burden of itemizing of $617 per person.

If individuals switch to the standard deduction because they value their time more than the benefits they could potentially derive from itemizing, richer households should forego more tax benefits than poorer ones. To test this hypothesis, I break down individuals by income deciles and repeat the same identification strategy outlined above. The results show an increasing relationship between foregone tax benefits and income, consistent with the hypothesis that tax filing imposes a higher burden on richer individuals because they have a higher marginal value of time. Using a revealed preference argument, I estimate that itemizing is perceived to be as painful as working more than 19 hours at one's regular job.

I then proceed with providing an explanation for the heterogeneity in foregone benefits. I focus on factors that increase the likelihood of switching from itemizing to claiming the standard deduction. I find that taxpayers with a large proportion of deductions that are easy to keep track of are less likely to switch to the standard deduction and that the use of tax

[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. The full set of details is given in section 5.5.

preparation software or tax prepares does not eliminate the burden of itemizing. Similarly, I find that having a newborn in year t increases the chances of claiming the standard deduction in tax year t. Two effects are at play. On the one hand, newborns are likely to increase the deductions one can claim such as medical deductions. On the other hand, newborns increase the value of time making itemizing especially painful. My findings show that even though the benefits of itemizing are likely to be higher when having a newborn, the increase in the burden of itemizing prevails.

The cost of itemizing is the sum of two separate costs: the cost of record keeping and the cost of filling out Schedule A. Which one of the two is higher and drives the result? To answer this question, I consider the outside option of using a tax preparer. Tax preparers can provide assistance with filling out forms but they cannot perform any record keeping. The fee charged by tax preparers to file Schedule A is therefore an upper bound on the cost of filling out Schedule A. I find that the fee charged by tax preparers for filing *both* the 1040 form and Schedule A is less than \$220, implying that most of the cost is driven by record keeping.

Reasonable calibrations of the cost of itemizing suggest that it is unlikely that such a simple task requires so much time. Schedule A is one of the easiest forms to fill out as it does not require any calculations or any tax tables. The taxpayer only needs to copy numbers from receipts and then sum them up, which is unlikely to require more than an hour of work. It would imply that record keeping would require more than 18 hours, which is hard to believe.

Such high cost estimates could be consistent with an extreme aversion to filing taxes. Using a revealed preference argument and survey estimates of the time required to file federal taxes, I estimate that taxpayers dislike working on taxes 4.2 times more than they dislike working at their regular jobs. If this is the case, back-of-the-envelope calculations suggest that the overall burden of complying with income taxes is 1.25% of GDP. The policy implication of this result are that the time spent filing taxes should be reduced. This can be achieved by requiring less receipts, shortening forms and more generally simplifying the tax code.

However, there is compelling evidence that individuals are time inconsistent when saving for retirement,[3] searching for a job,[4] starting to smoke[5] and in many other situations. Time

[3]Madrian and Shea (2001).
[4]DellaVigna and Paserman (2005).
[5]Gruber and Köszegi (2001).

inconsistency leads to a failure of the axiom of revealed preferences[6] and introduces a wedge between compliance costs and foregone benefits. A model of time inconsistency based on present-bias shows that taxpayers forego large benefits even when compliance costs are modest. Present-bias forces individuals to delay filing until the deadline of April 15th making them more likely to face a high idiosyncratic cost. The policy implications of this model are more subtle than a mere reduction in time spent filing but they are also easier to implement and less costly. The model makes predictions about filing time and the behavior of taxpayers close to the deadline that are consistent with evidence that I gather from tax returns.

Overall, both aversion to tax filing and naive present-bias suggest that the burden of compliance is significantly larger than previously estimated whether due to compliance costs per se or psychological biases.

The results of this paper have implications in several dimensions. First, this is – to my best knowledge – the only paper to provide estimates of the burden of compliance by directly observing the behavior of taxpayers and using administrative tax data. Two other papers address this question using tax data: Pitt and Slemrod (1989) and Slemrod (1989). Because they cannot observe the preferences of taxpayers, they estimate a structural model. They find smaller compliance costs.[7] There is also a literature that uses survey evidence to estimate compliance costs. Although informative of the time spent filing taxes, it does not capture the preferences of taxpayers and in particular any aversion to filing taxes or any behavioral biases. The magnitude of my estimates emphasizes the policy relevance of reducing the compliance burden and advocates for a simplification of the tax code.

There is an extensive literature that documents low take up rates of government provided benefits.[8] Three explanations are generally offered: lack of information about the program, stigma costs and compliance costs. This paper is the first to disentangle compliance costs from lack of information and stigma costs and to show that they have a significant effect on benefit take up. The literature has mostly focused on the role of information. Bhargava and Manoli (2011) for example show that failure to claim the EITC can be explained by lack of information about the program but do not address compliance costs. My findings also provide a

[6]More generally, inferring preferences from choice behavior is discussed in Koszegi and Rabin (2007), Koszegi and Rabin (2008a) and Koszegi and Rabin (2008b).

[7]Possibly because of the structural assumptions they have to make.

[8]See Currie (2004) for a survey of the literature.

plausible and additional explanation for other phenomena reported by the literature. Jones (2012) shows that taxpayers fail to adjust their tax withholding resulting in foregone interest payments. He explains his results with inertia. An additional explanation could be the cost of filling out form W4 and sending it to the IRS. Engström et al. (2013) and Rees-Jones (2013)[9] show that taxpayers who have a balance due are more likely to reduce their balance to zero by claiming additional deductions. They provide compelling evidence that this behavior is driven by loss aversion. My estimates show that the cost of sending a cheque to the tax authorities could be an additional channel for their result.

This paper is also related to a literature in Marketing[10] and Behavioral Economics[11] documenting instances in which consumers fail to claim rebates. Some estimates suggest that only 1% of coupons are eventually redeemed.[12] Explanations of this findings are scarce. My results show that transaction costs (mailing the coupon etc.) are a plausible channel for this phenomenon.

Finally, this paper adds to a long tradition in public economics emphasizing the need to screen out applicants for welfare benefits by imposing high hassle costs[13] such as waiting in line, filling out forms etc. Poorer individuals value their time less – possibly because they are unemployed – and such policies can successfully target them by screening out richer individuals. My results show that this effect is indeed true because richer individuals tend to forego more benefits than poorer ones. However, given how large the costs are, such policies could be screening out too many individuals. In addition, time inconsistency could lead to unwanted distortions such as screening out naive individuals versus rational ones rather than rich ones versus poor ones.

2 Data and Institutional Background

2.1 The Decision to Itemize

Taxpayers can reduce their taxable income by claiming deductions. The most common deductions claimed by taxpayers are for mortgage interest payments, state and local income

[9] See also Feenberg and Skinner (1989).
[10] Silk and Janiszewski (2008).
[11] Ericson (2011) and Letzler and Tasoff (2014).
[12] Inmar (2012).
[13] Nichols et al. (1971), Duclos (1995) and Kleven and Kopczuk (2011).

taxes, charitable contributions and real estate taxes.

Consider, for example, a single person with an income of $150,000. In 1989 her marginal tax rate is 28%. If the person spends a total of $10,000 on different expenses that she is allowed to deduct from her income, her tax liability is reduced by $2,800. If instead she decides to claim the standard deduction – which in 1989 was $3,100 – her tax liability is reduced only by $868.

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

A rational taxpayer should account for these costs: if her total itemized deduction exceeds the standard deduction by an amount smaller than the cost of itemizing, she should claim the standard deduction even though it would result in a larger tax liability.

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

2.2 The Cost of Itemizing

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

The majority of taxpayers itemize four types of deductions:

- State and local income taxes: these are taxes paid in year t to the state or to the locality. They are reported on form W2 received in January of year $t + 1$. On average they represent 17% of total deductions.
- Mortgage interest: this is the interest paid to finance the main or second home of the taxpayer. It is reported on form 1098[14] which is received in January of year $t + 1$. On

[14]Mortgage interest on the purchase of a Recreational Vehicle (RV) or a boat used as a primary or secondary residence is not reported on the 1098. This is unlikely to bias my results given that very few people live in RVs or boats.

average they represent 40% of total deductions.

- Real estate taxes: these are taxes paid on real estate owned by the taxpayer. They can be found on the 1098 form, in financial records or by calling the county tax assessor. On average they represent 14% of total deductions.
- Charitable donations: any payment made for charitable purposes including to religious institutions. These payments are not subject to third-party reporting. The taxpayer has to keep records of her own receipts. On average they represent 12% of total deductions.

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

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

2.3 Data

The dataset used to carry this analysis consists of annual cross sections of individual tax returns. It is constructed by the IRS and called the Individual Public Use Tax Files. They are commonly referred to as the Statistics of Income (SOI) files. The data is available annually for the periods that I am analyzing. The number of observation per year ranges from 80,000 to 200,000. The repeated cross sections are stratified random samples where the randomization occurs over the Social Security Number. The data oversamples high-income taxpayers as well as taxpayers with business income but weights are provided by the IRS allowing my analysis to reflect population averages.

In addition, I use a panel of tax returns known as the University of Michigan tax panel. The panel covers 1979 to 1990 and contains the same variables as the SOI files but has a smaller sample size (less than 40,000 observations per year).

In rare cases individuals are forced to itemize deductions even though the standard deduction amount is larger than their deductions (details in section A.2). I drop these individuals from the sample. I focus on joint filers as they represent more than 50% of the sample yielding more precise estimates than if I were to use single filers or head of households.

3 Results

In this section I reconstruct the counterfactual density of itemizers and estimate the burden of itemizing deductions. First, if taxpayers are claiming the standard deduction even though they could profit from itemizing, there should be a missing mass in the neighborhood of the standard deduction. This missing mass is observed for any year (figure 1). Second, to show the causal relationship between the standard deduction and the missing mass in the neighborhood of the standard deduction, I use two reforms that increase the standard deduction amount, in 1971 and 1989 (see table C.6). I observe that the missing mass follows precisely the standard deduction threshold (figures 2a and 2b). Third, I develop a method to recover the counterfactual density of deductions. Finally, I use this counterfactual density to estimate the distribution of the burden of itemizing deductions in the population.

3.1 Missing Mass In the Neighborhood of the Standard Deduction

If some taxpayers are claiming the standard deduction even though the sum of their itemized deductions is greater than the standard deduction there should be a missing mass in the neighborhood of the standard deduction threshold.

I graph the density of deductions for all years ranging from 1980 to 2006 by bin sizes of \$2,000[15] in figure 1. The bin closest to the standard deduction only includes itemizers whose deductions are strictly larger than the standard deduction amount. Notice that the density is systematically low in the neighborhood of the standard deduction and then increases and peaks 2 to 3 bins away. This is consistent across years and across filing status.

The shape of the density of itemizers in the neighborhood of the standard deduction suggests that the density is discontinuous at the standard deduction. However, I cannot observe the density of itemizers below the standard deduction threshold because taxpayers who claim the standard deduction are not required to list their true deductions. In figure B.9, I consider the

[15] all dollar amounts are in 2014 dollars in the rest of the paper.

three different scenarios for the counterfactual density of deductions of taxpayers who claim the standard deduction. Approximately two-thirds of taxpayers claim the standard deduction which means that the density below the standard deduction threshold cannot be increasing from zero onwards and then connect with the density on the right-hand side of the standard deduction (graph (a)), as this would fail to account for a large portion of the population. If the density is smoothly decreasing on the left-hand side of the standard deduction threshold and it is single-peaked, then it is likely that the density is discontinuous at the standard deduction threshold (graph (c)). But I cannot rule out double peaked distributions (graph (b)) without knowing what the true distribution of total deductions is below the standard deduction threshold. This is why I turn to a quasi-experimental approach.

3.2 Identifying the Missing Distribution

The government adjusted the standard deduction in multiple instances. This resulted in large increases in the standard deduction amount. These natural experiments allow me to analyze the effect of the standard deduction on itemizers. The largest of these changes happened in 1971 and 1988. Table C.6 reports that the standard deduction is increased respectively by 50% and 33%.

I compare the pre-reform year to the post-reform year to account for lagged behavioral responses. The taxpayer has to perform two separate tasks when itemizing. Recall that first she has to keep a record of all the expenses she incurred in year t and keep track of her receipts. Second, she has to fill out Schedule A in year $t+1$. Filling out the form and record keeping happen in two consecutive years. Assume I am considering the 1988 reform. The taxpayer files her taxes in 1989 for 1988, which means that she is likely to have already collected the documents in 1988 and only has to work through Schedule A. Once done, she realizes that her itemized deductions are only slightly above the standard deduction amount and that it would have been more cost efficient for her to claim the standard deduction. However, documents have been collected already and forms filled out so she follows through with the decision to itemize. She promises herself however that next year she will save herself the trouble of dealing with all this administrative hassle and will claim the standard deduction instead. Accordingly, the burden estimated in 1988 should only reflect that of working through Schedule A. In 1989 however, the burden should reflect the entire hassle of

itemizing deductions. Comparing 1987 to 1988 shows that there is a lagged response.[16]

Figures 2a and 2b graph the density of deductions in pre and post-reform years for the 1971[17] and 1988 reforms. Notice that the shape of the distribution in year t+1 mirrors that of year t-1 and that the missing mass precisely follows the new standard deduction threshold. This shows that some itemizers switch to the standard deduction once it is increased even though their deductions are larger than the standard deduction.

The fact that the missing mass closely follows the standard deduction establishes that there is a discontinuity at the standard deduction due to the effort cost of itemizing. If this missing mass was a feature of the distribution and not due to the standard deduction, it should not follow the standard deduction once it is increased.

3.3 Recovering the Counterfactual Distribution Using the Reform Years

To calculate the burden of itemizing deductions I create bins of $2000.[18] I calculate the weighted frequency of individuals located in those bins. I subtract the mass of the 1989 bin from the mass of the corresponding bin in 1987 after adjusting the amounts to account for inflation.

To calculate the standard errors of the difference between the bins in the 1987 and 1989 densities and 1970 and 1971 densities, I use a bootstrap procedure with 100 replications. The results are reported in table C.7 and table C.8. The difference between the first and second bins is statistically significant with large z statistics (6.55 and 3.47). The rest of the bins are all overlapping with differences that are not significant even at the 10% level, at the exception of bin 10, 11 and 13 that are statistically significantly different at the 5 and 10% level, with differences of very small magnitude (less than 10 times that of the first or second bins).

This approach allows me to measure the percentage of individuals that claim the standard

[16] This is consistent for example with Gelber et al. (2013) who show that individuals have a lagged response to changes in Social Security benefits.

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

[18] I also consider $1,000 bin sizes in table 3, which yields similar results.

deduction even though their total itemized deductions exceed the standard deduction amount by multiples of $2,000.

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

To recover the counterfactual distribution of deductions I use the fact that the distribution of costs should be the same in 1987 and 1989.[19] I consider the first bins for which the 1987 and 1989 densities are overlapping. Figure 2a shows that these are bin number 3 in 1989 and bin 4 in 1987. In 1989, taxpayers located in bin number 3 can save from $4,000 to $6,000 worth of deductions. Given that the 1987 and 1989 densities are overlapping in the 1989 bin 3, no taxpayer is willing to forego more than $4,000 worth of deductions. In 1987, taxpayers who can save from $4,000 to $6,000 of deductions are located in bin 3. Since I assume that the cost is the same across years, this means that 1987 is the undistorted counterfactual for the 1989 density in bin 2. Bin 1 of the 1989 density is compared to bin 2 of the 1987 density. However, from observing the difference between the 1987 bin 3 and 1989 bin 2, I know that some taxpayers will forego deductions, biasing the 1987 bin 2 downward. By using the difference between the 1987 bin 3 and 1989 bin 2, I can calculate this bias and correct bin 2 in 1987 to get the true counterfactual for the 1989 bin 1.

The following hypothetical example illustrates the adjustment process and formalizes the approach I use to recover the true counterfactual density. I generate an undistorted hypothetical density of deductions in figure 5. Each bin size is equal to $100. I assume that the distribution of the burden of itemizing in the population is given by the following:

- 40% have a burden lower than $100
- 70% have a burden lower than $200

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

- 85% have a burden lower than \$300
- 95% have a burden lower than \$400

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

Denote by b_i^t the bin density, where i is the distance (in bins) to the standard deduction and t is the year. Year t corresponds to the pre-reform year and year $t+1$ to the post-reform year. When overlapping the density of deductions for year t and year $t+1$, b_i^t will be at the same location as b_{i-2}^{t+1} because the standard deduction increases by 2 bins because of the reform. If $b_i^t - b_{i-2}^{t+1} = 0$ then $d_{i-2} = 0$. I then start with the first undistorted bin. In graph 5 it corresponds to bin 7 in year t:

- $b_7^t - b_5^{t+1} = 0$ implies that $d_5 = 0$. This means that for both year t and $t+1$, b_5, b_6, b_7 etc. are undistorted as $d_5 = 0$ means that nobody has a burden greater than 500. This also means that I can use b_6^t as the true counterfactual to calculate d_4.
- $b_6^t - b_4^{t+1} = 5\%$ implies that $d_4 = 5\%$. Given that b_6^t is the true density (from the previous bullet point), I can use $\frac{b_4^t}{1-d_4}$ as the *true* counterfactual to calculate d_2.
- $b_5^t - b_3^{t+1} = 15\%$ implies that $d_3 = 15\%$. Given that b_5^t is the true density (from the first bullet point), I can use $\frac{b_3^t}{1-d_3}$ as the *true* counterfactual to calculate d_1.
- To calculate d_2 I need to use b_4^t. But I know from above (second bullet point) that b_4^t is distorted. This implies that the counterfactual density that I need to use to calculate d_2 is $\frac{b_4^t}{1-d_4}$ rather than b_4^t. Hence, $d_2 = \frac{b_4^t}{1-d_4} - b_2^{t+1} = 30\%$.
- Similarly, to calculate d_1 I need to use b_3^t. But I know from the third bullet point that b_3^t is distorted. This implies that the counterfactual density that I need to use to calculate d_3 is $\frac{b_3^t}{1-d_3}$ rather than b_3^t. Hence, $d_1 = \frac{b_3^t}{1-d_3} - b_1^{t+1} = 60\%$.

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

3.4 Estimation of the Burden of Itemizing Deductions

3.4.1 Distribution of the Burden

The counterfactual density that I constructed using the method outlined in the previous section allows me to observe the number of taxpayers who itemize in every bin and contrast it with the number of taxpayers who should have been itemizing had there been no cost to itemizing.

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

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

In bin 1 for example we know that:

- $b_1 \in (0, 2000]$.
- By using the difference between the mass of itemizers and the missing mass, I observe that $p_1 = 53\%$.
- This implies that $F(2000) = 1 - 0.53 = 47\%$ i.e. 47% of the population has a burden that is lower than \$2000 and greater than \$0.[20]

By repeating the same procedure for the remaining bins, I can construct the CDF of c_i. I

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

need the probability density function (PDF) to calculate the average perceived burden of itemizing, which can be derived from the CDF by taking the difference of the proportion of itemizers between each subsequent bin. Denote by m_k the PDF in bin bin k, then:

$$m_k = p_k - p_{k-1}. \tag{2}$$

The PDF and CDF are shown in table 2. Table 3 reports the CDF and PDF for a smaller bin size ($1,000).[21]

3.4.2 The 1988 reform

Besides the standard deduction reform, the only reform happening in 1988 that could affect the amount of deductions is the personal interest deduction phase-out, which I control for (details in section 5.5.1). There were no other reforms affecting deductions in 1988 or 1989 and the reforms affecting the 1987 distribution do not have lagged effects (see section 5.5.1 for the full list of reforms and appendix section A.1 for the TRA'86 reforms). I restrict my sample to taxpayers with the same marginal tax rate (28%) and who are not subject to the Alternative Minimum Tax (AMT). There is a marginal tax rate decrease for married filing jointly with Adjusted Gross Income (AGI) above $45,000 (in 1987 dollars) in 1988. I control for this change by only considering taxpayers with AGI below $45,000.

Table 2 implies that taxpayers forego large amounts of deductions, resulting in a burden of itemizing of $617 (s.e. 54.1). The average net annual wage for households in the neighborhood of the standard deduction is equal to $92,743. I assume that these households work on average 40 hours a week and 50 weeks a year. Given that all these households fall in the 28% marginal tax bracket, this results in a net wage of $33 for the household. A revealed preference argument implies that taxpayers perceive the task of itemizing to be equivalent to 19 hours of work on average. This is a lower bound as I am assuming that there is only one earner in the household and that their hourly wage is equal to the $33. If there are two earners and they work the same amount of hours, their individual wage would be $16.5. Filing the tax return only requires one person, implying that if there are two earners, they would perceive the task of itemizing to be as costly as working 37 hours at their regular jobs.

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

But to be conservative, I will assume that there is only one earner, keeping in mind that is likely to be a lower bound.

Every year, the IRS provides cost estimates for each tax form including both the time required to fill out the form and to perform the record keeping. In 1989, the IRS estimates that the average taxpayer needs 1 hour and 1 minute to fill out Schedule A, 2 hours and 47 minutes for record keeping, 26 minutes to learn about the form and 20 minutes to copy and assemble the documents before sending them to the IRS. This totals 4 hours and 34 minutes.[22]

3.4.3 The 1971 reform

To calculate the average average burden of itemizing deduction using the 1971 reform I need to control for the change in the parallel system of standard deduction which increases from 10% to 13% of AGI. The details of the adjustment are in section 5.5.2. In addition, and contrary to the 1988 reform, I can only calculate the average burden of itemizing using the average marginal tax rate because there were 25 marginal tax brackets in 1970 and 1971. I also focus on married taxpayers filing jointly to simplify the average marginal tax rate calculations and use bins of $3,065 ($500 in 1970 dollars) to reduce the noise. I estimate that the average burden of itemizing deductions is $996 (s.e. 126).

3.5 The Burden of Itemizing Deductions Increases With Income

If rich taxpayers value their time more than poor ones because their hourly wage is higher, we should expect them to forego more benefits. I can verify this assertion using the income reported on tax returns.

I break down the sample by deciles of AGI. Because this would significantly reduce the sample size and might make the results less precise, I consider a distribution around each AGI decile threshold. For example, the lower AGI group consists of every individual with AGI below the second decile threshold. And the second group consists of taxpayers with an AGI above the first decile and below the third decile etc. Some individuals will simultaneously belong to two groups: for example individuals whose AGI falls in the second AGI decile will belong to both the first group (AGI below the second AGI threshold) and the second group (AGI greater than the first decile threshold but smaller than the third decile threshold). This overlap is not

[22] Guyton et al. (2003) describe the methods used by the IRS to calculate the cost of filing taxes based on surveys of taxpayers.

a concern because the goal of this breakdown is to graph the relationship between AGI and foregone benefits. The precise location of a point in the AGI/foregone benefit space is of no particular interest. Instead, I am interested in plotting the general trend of the relationship.

Once the groups are constructed and because I have less data points, I fit a polynomial of degree 3 through each deduction bin. I construct confidence intervals around each bin. Any bins for which the confidence intervals overlap are considered as overlapping bins. Using the predicted bins from this polynomial, I am able to calculate the foregone benefits for each group by repeating the procedure developed in the previous section: I compare the distribution in 1987 to that in 1989, reconstruct the counterfactual distribution of itemized deductions and calculate the distribution of the burden of itemizing by comparing the counterfactual distribution to the true one. I only report results for the first six groups because deductions and AGI are positively correlated implying that there are very few high income individuals close to the standard deduction threshold. In figure 6(a), the x-axis represents the average AGI and the y-axis the average burden of itemizing for each income group. The relationship is increasing: as income increases taxpayers forego more benefits consistent with the idea that they value their time relatively more.

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

Figure 6(b) shows the relationship between income and the perceived hours required to itemize deductions. I assume that taxpayers work on average forty hours a day and fifty weeks a year and I divide their wages by the number of hours worked per year. By dividing the estimated burden of compliance by this measure of their hourly wage, I get the perceived hours required to itemize deductions.

Although of a lower magnitude and significance I find an increasing relationship between the value of time and the AGI deciles. The estimates range between 20 hours for the lowest income group to 40 hours for the highest one. If the relationship had been constant, it would have meant that richer and poorer taxpayers perceive the decision to itemize to be equally time consuming, but here the relationship is increasing. I interpret this as being due to the fact that individuals have different preferences over filing their taxes relative to working an

extra hour at their regular jobs. In other words, what I am calculating is the marginal rate of substitution (MRS) between an hour of work and an hour of filing taxes. The MRS could be different for two individuals for two reasons: they might enjoy working at their regular job equally but dislike filing taxes differently or, alternatively they might enjoy working at their jobs differently but dislike filing taxes equally. The most plausible story is that better paying jobs are more fulfilling and that individuals equally dislike filing taxes, which explains that the revealed value of time for rich households is higher than for poor households.

It is true that rich individuals have higher dollar amounts of deductions but it is unlikely that it takes more time to itemize them. The burden of itemizing is mostly fixed and does not generally increase in the amount of a given deduction. If a taxpayer has $10,000 worth of mortgage interest, she will spend the same amount of time itemizing them as a taxpayer who has $100,000 since they only have to keep track of one form and enter one number on Schedule A.

This is – to my best knowledge – the first piece of empirical evidence documenting that richer individuals value their time more than poorer ones.

3.6 Tax Preparers and Electronic Filing

Electronic filing and the use of tax preparers significantly reduce the cost of filling forms. However, it does not affect the cost of record keeping.

To test for whether electronic filing or using a tax preparer eliminates the burden of itemizing, I graph the density of itemizers who use a tax preparer and those who use electronic filing in graph 7 and look for whether there is still a missing mass close to the standard deduction threshold. The missing mass is still present implying that tax preparers or electronic filing does not eliminate the burden of itemizing.

Although surprising at first glance, this result is rather intuitive. Using tax preparers or electronic filing reduces the cost of filling out Schedule A significantly but does not affect the cost of record keeping: taxpayers still have to keep track of receipts and collect them at the time of filing. This is additional evidence that the cost of record keeping is indeed driving the result.

3.7 Who Is More Likely to Switch to the Standard Deduction?

3.7.1 Identification Strategy

I use the panel dataset to identify the reasons taxpayers switch to the standard deduction. I focus on taxpayers who itemize deductions in year t and observe their decisions in year $t+1$. Therefore, the variable of interest is a dummy equal to 1 if the taxpayer switches to the standard deduction in year $t+1$ and 0 if she keeps itemizing. I drop individuals who have to file other Schedules (B, C, etc.) as they could bias the results.[23] I do not consider individuals who switch from claiming the standard deduction to itemizing because this decision is not as easily available as the opposite one. Indeed, a person with deductions in excess of the standard deduction threshold can easily decide between itemizing and not. But a person who is claiming the standard deduction is likely to have too few deductions in total to be able to itemize. All my results are clustered at the individual level.

I regress a variable that indicates that the individual is switching to the standard deduction on several variables of interest that I explain below. I also control for the level of deductions in year t, a polynomial of AGI, marital status, year and state fixed effects. The results are reported on table 1. The regression specification is the following:

$$y_{it} = constant + newborn_{it} + newborn_{it} * close_{i(t-1)} + easyded_{i(t-1)} + ...$$

$$... + easyded_{i(t-1)} * close_{i(t-1)} + nopreparer_{i(t-1)} + nopreparer_{i(t-1)} * close_{i(t-1)} + x_{it} + \epsilon_{it}$$

- $y_{it} = 1$ if the taxpayer itemizes in year $t-1$ and switches to SD in year t and 0 if itemizes in year $t-1$ and year t.
- $close_{i(t-1)} = 1$ if the taxpayer reported itemized deductions within \$6,000[24] of the standard deduction in year $t-1$ and 0 otherwise.
- $newborn_{it} = 1$ if the taxpayer has a newborn in year t and 0 otherwise.
- $easyded_{i(t-1)} = 1$ if the sum of state tax and mortgage deductions in $t-1$ is greater than standard deduction and 0 otherwise.

[23] Some of these taxpayers have to deal with much higher record keeping costs than those required for itemizing.

[24] I choose \$6,000 because the pre and post-reform densities overlap 3 bins away from the standard deduction. I ran specifications with \$4000, \$5,000 and \$7,000 and found results of similar magnitude to the ones reported here.

- $nopreparer_{i(t-1)} = 1$ if the person does not use a tax preparer in year $t-1$ and 0 otherwise.
- x_{it} controls for a polynomial of income, the level of deductions in year $t-1$, the marital status, the state, the marginal tax rate and the year.

3.7.2 Newborn

Childbirth drastically reduces the amount of time available. Parents with a newborn are likely to value their time more than parents with no children because they have less leisure time available. For this reason, it is sensible to expect that families with newborns are more likely to switch to the standard deduction in the year when their child is born.

The results of the regression are reported in table 1, column 1, 5 and 6. I find a significant and positive coefficient for the interaction term of being close to the standard deduction threshold and having a newborn. A taxpayer who is close to the standard deduction threshold and who has a newborn is 5% more likely to switch to the standard deduction.

An exogenous shock to the value of time such as child birth has significant effects over the decision to itemize providing additional evidence that taxpayers are trading off time and money when deciding to itemize.

3.7.3 High Ratio of Third Party Reported Deductions

The mortgage payment deduction and the state and local income tax deduction are both third-party reported implying that taxpayers receive a "statement" with the 1098 and W2 forms in January of year $t+1$, significantly reducing the record keeping cost.

The results of the regression are reported in table 1 columns 2, 5 and 6. A taxpayer with a high proportion of mortgage interest and state tax deductions is 12% less likely to switch to the standard deduction when her deductions are close to the standard deduction threshold. This is consistent with the overall compliance burden being smaller for these two types of tax deductions because they have a relatively lower record keeping cost since both form W2 and form 1098 are received in January of year $t+1$, closer to the tax filing season. The fact that the record keeping cost is smaller if the receipts are sent closer to the tax filing season suggests that forms are harder to find or more likely to get lost as time elapses, possibly because they are not properly archived.

3.7.4 Tax Preparers

Tax preparers are readily available and provide taxpayers with assistance to file their returns. They also provide help in choosing the best options when filing taxes and ensuring that the taxpayer is "optimizing". However, they do not make the task of record keeping any easier.

Who uses tax preparers? Three types of individuals: low-income households who can get their refund faster when using tax preparers, households with complicated tax returns and households whose value of time is larger than the fee that they have to pay to the tax preparers.

The taxpayers who itemize deductions are unlikely to have low-incomes simply because deductions are strongly correlated with income.

To control for individuals who are using tax preparers because of the complexity of their tax return, I drop any person who files any other schedules but Schedule A. Those include individuals who have capital gains or dividends, or individuals who have profit or losses from farming etc. These Schedules are significantly more complicated and a visit to tax preparers might be necessary even for the most tax-savvy taxpayers.

I find that using a tax preparer has no effect on the decision to itemize. The absence of effect is likely due to the fact that tax preparers cannot provide any assistance with record keeping which is more costly than filling out Schedule A.

4 Welfare Implications

The task of itemizing deductions imposes an average burden of $617 on taxpayers. This is a significant amount of money given the average income of the population of interest.

A revealed preference argument implies that $617 corresponds to the true compliance cost. In other words, this means that when a rational taxpayer is faced with the decision to itemize or claim the standard deduction, she will only itemize if she can save more than $617. The fact that taxpayers experience such a large disutility from itemizing could be due to an extreme aversion to filing taxes.

However, the behavioral economics literature has documented several instances in which the axiom of revealed preferences fails, in particular because of time inconsistency. A failure of

this axiom introduces a wedge between foregone benefits and compliance costs, reconciling the large magnitude of my estimates with the survey evidence.

In what follows, I discuss the welfare implications of my result in light of both perspectives and argue that a model based on time inconsistency is likely to explain this behavior better than one that assumes that taxpayers are rational.

4.1 Aversion to Filing Taxes

The axiom of revealed preferences implies that taxpayers perceive the task of itemizing to be as costly as $617. This figure is calculated for taxpayers in the neighborhood of the standard deduction. Is it reasonable to assume that it is representative of the entire population of taxpayers? If the amount of deductions was randomly assigned across taxpayers, this assumption would hold. However, there is a strong relationship between deductions and income as richer taxpayers have higher state taxes, larger mortgages, more expensive houses etc. And given that richer taxpayers tend to forego more deductions because they value their time more (as shown in graph 6), the average cost of $617 is likely to be a lower bound on the cost of itemizing for richer taxpayers. Taxpayers who have deductions close to the standard deduction threshold are among the poorest itemizers and therefore extrapolating the $617 of perceived costs to the entire population of itemizers is likely to understate the burden of itemizing deductions imposed on the population of itemizers. Therefore, by extrapolating this amount to the rest of the population, I am understating the estimates of the aggregate burden of compliance.

The IRS estimates that itemizing requires less than 4.5 hours. A revealed preference argument implies that taxpayers perceive the task of itemizing to be as costly as working 19 hours at their regular jobs. My estimates are 4.2 times larger than the ones provided by the IRS. Their estimates are based on surveys of the time spent filing taxes. However, they do not ask how much taxpayers dislike filing taxes. If my estimates are truly driven by aversion to filing taxes then – taking the IRS estimates as given – my results suggest that spending an hour preparing taxes is 4.2 times more painful than spending one hour working. Taking this estimate as given, back-of-the-envelope calculations can inform us on the overall burden of filing taxes.

If these are the true preferences of taxpayers then I can use this estimate and the survey

estimates of the time required to file the various income tax forms to calculate the aggregate cost of filing taxes. If the wedge between survey estimates and revealed preference estimates is due to the aversion to filing taxes which cannot be captured by surveys, then multiplying the survey estimates by 4.2 would account for the full burden of tax filing including the aversion taxpayers experience when filing taxes. Table 4 shows the results of these calculations. Overall, the cost of complying with the individual tax code amounts to 1.25% of GDP in 1989 ($188 b.). In comparison, Feldstein (1999) estimates that the efficiency cost of Personal Income Tax and the Payroll Tax ranges between 2 and 5% of GDP. These orders of magnitude emphasize how important compliance costs are especially given how relatively little attention the Public Finance literature has given them compared to the efficiency cost of taxation.

4.2 Time Inconsistency

There is extensive evidence[25] that individuals are time inconsistent in different settings: credit card debt,[26] retirement saving,[27] addiction,[28] job search,[29] food stamps,[30] exercise[31] and others.

In this section, I provide a model based on naive present-bias that leads to time inconsistency and show that even small compliance costs can lead to large foregone benefits rationalizing the magnitude of my findings and the discrepancy between estimates based on revealed preferences and surveys. I start by showing in section 4.2.1 that time inconsistency predicts that taxpayers will procrastinate on filing their taxes and bunch at the deadline of April 15th. Next, I explore three different assumptions about the cost structure of filing taxes and show that taxpayers will forego more benefits because of time inconsistency. First, I assume in section 4.2.2 that the record keeping cost increases if taxpayers procrastinate on archiving their receipts leading to large foregone benefits due to large record keeping costs. In section 4.2.3, I assume that effort cost is convex. Naive present-biased taxpayers fail to smooth effort leading to high filing costs. Finally on section 4.2.4, I assume that cost realizations are stochastic and show that naive present-biased taxpayers expose themselves to higher costs by

[25] See DellaVigna (2009) for a survey of the literature.
[26] See Ausubel (1999).
[27] See Madrian and Shea (2001).
[28] See Gruber and Köszegi (2001).
[29] See DellaVigna and Paserman (2005).
[30] See Shapiro (2005).
[31] See DellaVigna and Malmendier (2006).

procrastinating. Overall, naive present biased taxpayers are more likely to forego deductions, consistent with evidence presented in section 4.2.5.

4.2.1 Deadline Effects

In this section, I show that in the simplest setting, the present-bias model[32] predicts that taxpayers will wait until the deadline of April 15th to file their taxes.

Assume a given taxpayer can derive a benefit b in period $t+1$ by incurring a cost c_t in period t. b is the level of total deductions minus the standard deduction and multiplied by the marginal tax rate. The rational taxpayer discounts the future with a daily discount factor δ (close to 1). Claiming the standard deduction does not provide any benefit and has no cost. She itemizes when:

$$\delta b - c_t \geq 0. \tag{3}$$

The rational taxpayer picks t such that c_t is the smallest. The naive present-biased taxpayer discounts the future by δ and the next day by $0 < \beta < 1$. She wants to itemize when:

$$\delta b - c_t \geq 0. \tag{4}$$

But procrastinates if

$$\delta\beta b - c_t \leq \delta\beta(\delta b - c_{t+1}). \tag{5}$$

With a constant cost $c_t = c$ and $\delta \approx 1$, this inequality simplifies to

$$\beta b - c \geq \beta(b - c). \tag{6}$$

Conditional on $\delta b - c_t \geq 0$, the naive present-biased taxpayer will procrastinate on itemizing until she reaches the deadline of April 15th at which point she has a choice between itemizing and claiming the standard deduction. She will itemize if

$$\beta b - c \geq 0. \tag{7}$$

Therefore, the model predicts that naive present-biased taxpayers will bunch at the deadline when filing their taxes, consistent with the anecdotal evidence of long lines of individuals

[32] Phelps and Pollak (1968) and Laibson (1997).

waiting at the post office on April 15th to submit their taxes.[33] Figure 8a graphs the volume of Google search of the term *1040* by week. There is a clear spike in the weeks that include April 15th consistent with the prediction of the model and suggesting that taxpayers are naive present-biased and procrastinate on filing their taxes.

In the following sections, I derive a prediction that shows taxpayers who are assumed to be naive present-biased will forego more deductions than rational ones and then I provide evidence of this prediction by comparing taxpayers who file their returns in April to those who file them in March.

4.2.2 Increasing Cost of Record Keeping

In what follows, 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.

Assume that the taxpayer has T periods to perform the two tasks, that they have to be performed on two separate days and that record keeping has to be done before filling out Schedule A. This means that the last period in which record keeping can be performed is $T-1$ whereas filling out Schedule A can be performed no later than in period T.

If the taxpayer succeeds in performing the two tasks she receives a one time benefit b. 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.

[33] Redelmeier and Yarnell (2012) show that there are more road crash fatalities precisely on April 15th.

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

In what follows, I use two definitions:

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

Similarly:

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

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:

$$\max_{t,u} \delta^t(-c(1+r)^{t-1} + \delta^u(b-k)), \tag{8}$$

conditional on itemizing being δ-worthwhile.

Cost c is incurred as soon as the taxpayer starts the record keeping. If she waits an additional u periods before filling out Schedule A the cost of record keeping is multiplied by $(1+r)$ for every period.

Assuming that $b-k>0$ i.e. that the benefit is large enough to justify filling out Schedule A, the taxpayer would want to itemize as soon as possible, which means having $u=1$.

The taxpayer is left with choosing t such that:

$$\max_{t} \delta^t(-c(1+r)^{t-1} + \delta(b-k)) \tag{9}$$

Assume the taxpayer is contemplating the decision to perform the record keeping task in the first period yielding utility: $-c+\delta(b-k)$. She will only perform it if $-c+\delta(b-k)>0$. And if she waits an additional period she will receive $\delta(-c(1+r)+\delta(b-k))$, 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 present-biased taxpayer maximizes the following utility function:

$$\max_{t,u} \delta^t(-c(1+r)^{t-1} + \beta\delta^u(b-k)). \tag{10}$$

Since benefit b and cost k are both incurred in the same period, the taxpayer will not procrastinate on performing the second task. This means that $u = 1$. The present-biased taxpayer is maximizing

$$\max_{t} \delta^t(-c(1+r)^{t-1} + \beta\delta(b-k)). \tag{11}$$

First, itemizing should be β-worthwhile, i.e. it should be profitable for the present-biased taxpayer to itemize now. This happens when the following condition is satisfied

$$-c + \beta\delta(b-k) > 0. \tag{12}$$

This condition simply states that the cost today should be smaller than the discounted net benefit tomorrow.

Second, she can perform the record keeping now or she can wait and perform it next period. She will prefer performing it next period if the following inequality is satisfied

$$\delta^t[-c(1+r)^{t-1} + \beta\delta(b-k)] < \delta^{t+1}\beta[-c(1+r)^{t-1} + \delta(b-k)]. \tag{13}$$

For δ close to one, this inequality simplifies to:

$$\beta < \frac{1}{1+r}. \tag{14}$$

Condition 14 implies that the taxpayer will procrastinate on archiving the charitable donation receipt if the benefit of waiting an additional period is greater than the increase in archiving cost.

Provided that condition 14 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 on it until the deadline.

Standard present-bias models show that with a strict deadline, naives are likely to procras-

tinate on completing a task but will perform it with certainty in the very last period (as illustrated in the previous section). Filing taxes has a strict deadline (April 15th) but when record keeping costs are increasing, naives will never perform the task simply because the expected benefit b has become too low for the task to be β-worthwhile. The intuition is the following:

1. In the first stage, itemizing deductions is both β and δ-worthwhile. The present-biased taxpayer would profit from itemizing now, but believes she would profit more from doing it later and keeps on postponing the task to the next period.
2. In the second stage, the expected value of itemizing has decreased enough to make itemizing not β-worthwhile anymore but is still δ-worthwhile. The taxpayer would not want to itemize now but still believes that she will do it tomorrow.
3. In the third stage, itemizing is neither β nor δ-worthwhile. The taxpayer does not want to perform it neither now, nor later.

For this to happen, we need the task to stop being β-worthwhile and still be δ-worthwhile in the next period. The naive present-biased taxpayer does not foresee this because everyday she believes that the task will be performed the next day as long as it is β-worthwhile. This arises when the two following inequalities hold:

$$-c(1+r)^{t-1} + \beta(b-k) < 0, \tag{15}$$

$$-c(1+r)^{t} + b - k > 0. \tag{16}$$

Which can be rewritten as follows:

$$\frac{\beta(b-k)}{c}^{\frac{1}{t-1}} - 1 < r < \frac{(b-k)}{c}^{\frac{1}{t}} - 1. \tag{17}$$

Under this condition, the taxpayer keeps procrastinating on archiving the receipt until the cost of doing so is too large to justify itemizing altogether. This result holds even for a relatively small c and predicts that large benefits are foregone by naive present-biased taxpayers.[34]

[34]The partially naive (also called sophisticated) taxpayer has self-control problems but is aware of them. She is able to look forward, solve the problem backwards and realize that at some point, itemizing will stop being β-worthwhile prompting her to itemize before it is not worthwhile anymore. Therefore her behavior is similar to that of the rational taxpayer.

This prediction is consistent with the results reported in table 1 columns 2, 5 and 6: taxpayers with a high proportion of deductions that have a low record keeping costs (the state and local tax and mortgage interest deductions) are less likely to switch from itemizing to claiming the standard deduction.

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

I can identify this fee in the dataset: individuals who itemize their deductions are allowed to deduct the tax preparer fee from their income. The average tax preparer fee for individuals who file the 1040 and Schedule A but not Schedule B, C, D etc. is $220. This is the fee for filling out both the 1040 form and Schedule A, submitting the documents and helping with audits if the need arises. This means that $220 is a generous upper bound. If the burden of itemizing deductions is driven by the cost of filling out Schedule A then taxpayers have the outside option of paying someone to perform this task and - for some of them - save large sums of money. This suggests that any cost in excess of $220 should be attributed to record keeping. Since the estimated cost is equal to $617, the record keeping cost accounts for more than 64% of the burden of itemizing.[35] This is consistent with taxpayers having large record keeping costs.

4.2.3 Increasing Marginal Disutility of Labor

In this section, I assume that the effort cost of itemizing is convex and show that it leads naive present-biased taxpayers to forego large deductions because they fail to smooth effort over time. I assume that taxpayers have an increasing marginal disutility of labor: it is more painful to work on taxes after 4 hours of work than after one hour. Knowing that, the rational taxpayer smoothes effort of time: she files the 1040 form, state taxes and Schedule A on three separate days. The naive present-biased taxpayer procrastinates on filing her taxes and ends up "pulling an all-nighter" on the very last day experiencing more disutility than the rational taxpayer. Given that state taxes and the 1040 form are both compulsory she has

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

no choice but to complete them. Itemizing however is optional and after working for several hours on her state taxes and 1040 form her marginal disutility of labor is so high that she is likely to turn down large sums of money to avoid spending more time looking for receipts and filling out Schedule A.

Assume the taxpayer needs to complete a task that requires H hours and provides a benefit b after it is completed. She can spread out the effort cost over T days. Her effort cost is given by a function $e(\cdot)$ with $e'(\cdot) > 0$ and $e''(\cdot) > 0$ implying that the effort cost increases in the number of hours spent working on taxes and the marginal disutility of effort is increasing.

The rational taxpayer will itemize provided that it is δ-worthwhile, and given that her marginal disutility of effort is increasing, she will smooth out the effort over the T days. Denote by h_i the number of hours spent working on taxes on day i. Formally, she is maximizing

$$\max_{h_1,h_2,...,h_T} -\sum_{i=1}^{T} \delta^i e(h_i) + \delta^{T+1} b, \tag{18}$$

subject to

$$-\sum_{i=1}^{T} \delta^i e(h_i) + \delta^{T+1} b > 0, \tag{19}$$

and

$$-\sum_{i=1}^{T} h_i = H. \tag{20}$$

Since δ is a daily discount factor, I set it equal to 1.[36] Equation (19) ensures that itemizing is δ-worthwhile. The taxpayer will smooth effort over time by choosing $h_i = \frac{H}{T}$ for any period i, provided that condition (19) holds. The equilibrium path for the rational taxpayer is $(h_1^R, h_2^R, ..., h_T^R)$, such that for any i, $h_i^R = \frac{H}{T}$.

The naive-present-biased taxpayer has a preference for instant gratification that results in her discounting anything that happens on the next day by $\beta < 1$. Her naivete implies that she believes that in the next day she will not overvalue the present. This leads her to be time inconsistent. Formally, every period t she believes that she will smooth the effort cost over the remaining $T - t$ periods but fails to do so every day.

[36] The main results are invariant to this assumption.

Because of her naivete, At time $t = 0$ she believes that she will be solving the following optimization problem:

$$\max_{h_1,h_2,\ldots,h_T} -\sum_{i=1}^{T} \delta^i e(h_i) + \delta^{T+1} b, \tag{21}$$

subject to

$$-\sum_{i=1}^{T} \delta^i e(h_i) + \delta^{T+1} b > 0, \tag{22}$$

and

$$-\sum_{i=1}^{T} h_i = H. \tag{23}$$

These conditions are the same as the one for the rational taxpayer: the naive taxpayer believes she will behave as if $\beta = 1$. But at time $t = 1$ she has a preference for instant gratification, and solves:

$$\max_{h_1,h_2,\ldots,h_N} -e(h_1) + \beta \left(-\sum_{i=2}^{N} \delta^i e(h_i) + \delta^{T+1} b \right), \tag{24}$$

subject to

$$-\sum_{i=1}^{N} \delta^i e(h_i) + \delta^{T+1} b > 0, \tag{25}$$

and

$$-\sum_{i=1}^{T} h_i = H. \tag{26}$$

Equation (24) is different from equation (18) in that everything except from period 1's cost is discounted by β.

Solving this problem in period t gives the following condition that describes the path of costs of the naive-present-biased taxpayer for $t < T$

$$e'(h_t) = \beta \sum_{i=t+1}^{T} \frac{1}{T-t} \delta^{i+1} e' \left(\frac{H - \sum_{i=1}^{t} h_i}{T-t} \right), \tag{27}$$

and for $t = T$

$$h_T = C - \sum_{i=1}^{T-1} h_i. \tag{28}$$

She equates the marginal disutility of effort today to the marginal disutility of effort in subsequent periods. She has wrong beliefs about β in the future and therefore thinks that she will smooth effort starting from tomorrow. Hence, for $i > t+1$, she believes that $h_i = \frac{H - \sum_{i=1}^{t} h_i}{T-t}$.

To calibrate this model, I assume that $\delta = 1$ and $e(h) = \frac{h^{1+\sigma}}{1+\sigma}$. From equation (27) it follows that

$$h_t = \frac{\beta^{\frac{1}{\sigma}}(H - \sum_{i=1}^{t-1} h_i)}{T - t + \beta^{\frac{1}{\sigma}}} \tag{29}$$

I further assume that $\sigma = 4$, $\beta = 0.5$ and that taxpayers can start filing their taxes as early as February 1st and as late as April 15 (corresponds to $T = 75$). In graph B.10(a), I plot the calibrated per-period disutility experienced by each type of taxpayer. In graph B.10(b), I plot the calibrated per-period number of hours spent working on taxes. The naive taxpayer works less than optimal the first days and works more the last days. This results in her experiencing a large disutility of effort in the last days because of the convexity of the effort function.

The taxpayer is required to file the 1040 form as well as any state income tax forms by April 15th. Itemizing however is optional. Given that she postpones most of her work to the very last days, itemizing can end up being very costly as her marginal disutility of effort in those days is high. It can become optimal at that time to forego large amounts of deductions even if it only requires 5 hours of work given how large of a disutility it would imply. The rational taxpayer on the other hand smoothes effort over time resulting in a relatively low marginal disutility of effort and would not forego the benefits of itemizing.

4.2.4 Idiosyncratic Shocks

In what follows, I assume that the cost of itemizing is stochastic: on some days, taxpayers have a high value of time (because they are busy) and on other days they are free and willing to itemize at a low cost. Rational taxpayers are aware of this variation in cost, have an option value of waiting and will do so to wait for a low cost realization. Naive present-biased taxpayers procrastinate on the task and fail to itemize even when costs are relatively low. This leads them to itemize on the last day exposing them to the full distribution of costs and leading them to forego large benefits when cost realizations are high.

Assume that the taxpayer has a choice between a costly task (itemizing deductions) and a

cost-free task (claiming the standard deduction). Assume that the cost of itemizing in period t, is given by $C = (1 + \alpha_t)c$ where c is the number of hours required to file Schedule A and α_t is stochastic and follows a distribution $F(\cdot)$. α_t represents the taxpayer's disutility from filing taxes on day t. Itemizing provides a deterministic benefit b. Claiming the standard deduction has a cost of zero and provides no benefit.

I build upon the search model developed by O'Donoghue and Rabin (1999b).[37]

Assume that the taxpayer has T periods to itemize. She is solving a Bellman equation with finite horizon. In the last period, she itemizes if $b - c(1 + \alpha_T) > 0$, which happens with a probability $F(\frac{b}{c} - 1)$. Denote by γ_t^R the threshold for α below which the task is performed by the rational taxpayer in period t. Then $\gamma_T^R = \frac{b}{c} - 1$. The utility derived from itemizing is given by $V_T = F(\gamma_T^R)(b - c(1 + \mathbb{E}(\alpha_T | \alpha_T < \gamma_T^R)))$.

In the period before, the task is performed when $b - c(1 + \alpha_{T-1}) > V_T$ i.e. when the benefit today is greater than the expected benefit tomorrow. This means that the cutoff is given by $\gamma_{T-1}^R = \frac{b - V_T}{c} - 1$ and $V_{T-1} = F(\gamma_{T-1}^R)[b - c(1 + \mathbb{E}(\alpha_{T-1} | \alpha_{T-1} < \gamma_{T-1}^R))] + [1 - F(\gamma_{T-1}^R)](V_T)$.

By induction, the cutoff and the continuation utility are given by:

$$\gamma_t^R = \frac{b - V_{t+1}}{c} - 1, \tag{30}$$

$$V_{t+1} = F(\gamma_{t+1}^R)[b - c(1 + \mathbb{E}(\alpha_{t+1} | \alpha_{T-1} < \gamma_{t+1}^R))] + [1 - F(\gamma_{t+1}^R)](V_{t+2}). \tag{31}$$

The naive present-biased taxpayer discounts the future by $0 < \beta < 1$. She mistakenly believes she will behave similarly to the rational taxpayer in the subsequent periods and thinks her cutoff vector will be γ^r. Let's denote her true cutoff vector by γ^n and compute it backwards. In the last period, γ_T^n is given by $\beta b - c(1 + \alpha_T) > 0$ i.e. $\gamma_T^n = \frac{\beta b}{c} - 1$. Notice that $\gamma_T^n < \gamma_T^r$. The naive present-biased taxpayer believes that she has the same cutoffs as the rational taxpayer, but her true cutoffs are given by $\gamma_t^n = \beta \gamma_t^R$. This implies that in every period, she is less likely to complete the task and more likely to delay it.

Denote by $Q(i, \beta, b)$ the probability that the taxpayer itemizes in a given period between $t = 1$ and $t = i$ given a present-bias parameter β and benefit from itemizing b:

[37] And more generally on Laibson (1997), O'Donoghue and Rabin (1999a), O'Donoghue and Rabin (2001) and O'Donoghue and Rabin (2008).

- $Q(1,\beta,b) = F(\beta\gamma_1^R)$,
- $Q(2,\beta,b) = F(\beta\gamma_1^R) + (1 - F(\beta\gamma_1^R))F(\beta\gamma_2^R) = Q(1) + (1 - F(\beta\gamma_1^R))F(\beta\gamma_2^R)$,
- $Q(3,\beta,b) = F(\beta\gamma_1^R) + (1 - F(\beta\gamma_1^R))[F(\beta\gamma_2^R) + (1 - F(\beta\gamma_2^R))F(\beta\gamma_3^R)] = Q(2) + \prod_{i=1}^{2}[1 - F(\beta\gamma_i)]F(\beta\gamma_3)$,
- And by induction:

$$Q(t,\beta,b) = Q(t-1) + \prod_{i=1}^{t-1}[1 - F(\beta\gamma_i)]F(\beta\gamma_t). \tag{32}$$

$Q(T,\beta,b)$ is the probability that the task is completed at time T. Given the benefit that a taxpayer can derive from itemizing, her present-bias parameter etc. this model predicts the likelihood that she will itemize and can be matched to the estimates that I provided in the previous sections. $Q(T,\beta,b)$ decreases in β: as taxpayers are more present-biased the cutoff γ is decreased implying more delaying of the task. Once the deadline is reached, the naive present-biased taxpayers cannot delay filing anymore and face the full range of idiosyncratic cost realizations. Taxpayers who face a high cost realization will forego large amounts of deductions.

4.2.5 Empirical Evidence

In sections 4.2.4, I showed that naive present-biased taxpayers are more likely than rational ones to forego large deductions even when the cost of itemizing is small. In particular, they are likely to procrastinate on archiving receipts leading to large record keeping costs, fail to smooth effort over time and expose themselves to higher idiosyncratic shocks. Naive present-biased taxpayers are likely to file their tax returns closer to the deadline of April 15th because of their procrastination. I can therefore use late filing as a tag for being naive present-biased.

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.

Consistent with the prediction that naive present-biased taxpayers are more likely to file close to April 15th and more likely to forego deductions as shown in the previous paragraphs, I find that the density of itemizers filing in April has a larger missing mass close to the standard deduction threshold than those who file in March. Figure 8b shows that the density of March itemizers matches that of April itemizers except in the neighborhood of the standard deduction where there are relatively fewer April itemizers. Notice that the two densities overlap everywhere except in bins that are close to the standard deduction threshold.

Similarly, regressions (4), (5) and (6) of table 1 show that – even controlling for observables – taxpayers who file late in year t are 6% more likely to switch to the standard deduction in year $t+1$.

Overall, this shows that taxpayers who file late are more likely to forego deductions, consistent with them being naive present-biased.

Notice 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 forego benefits. If that is the case, then we should observe that taxpayers who file late in year t are more likely to file earlier in year $t+1$. To test for this, I graph the average week in which returns are processed in year $t+1$ by week of processing in year t, in figure B.11. If taxpayers who bunch at the deadline are doing so for rational reasons, the relationship should be constant as we should observe mean reversion. If they are doing so because of a systematic bias, the relationship should be increasing as year t week of processing should predict year $t+1$ week of processing. Figure B.11 shows an increasing relationship between processing week in year t and year $t+1$ consistent with the explanation that late filing is due to a systematic bias.

4.3 Cost of Compliance When Taxpayers Are Naive Present-Biased

Time inconsistency implies a failure of the axiom of revealed preferences introducing a wedge between foregone benefits and compliance costs. In what follows, I estimate the compliance costs of taxation when taxpayers are naive present-biased. The estimated compliance cost

under this model is smaller than when assuming that taxpayers are fully rational, because naive present-biased taxpayers forego benefits both because of the cost of filing taxes and because of their bias. For simplicity, I assume away the increasing cost of record keeping and the increasing marginal disutility of effort and only assume that the cost of itemizing is stochastic (as in section 4.2.4). I then use the Generalized Method of Moments (GMM) to estimate the model.

The counterfactual distribution that I reconstructed in section 3 allows me to calculate the proportion of taxpayers who claim the standard deduction when they could benefit from itemizing (table 2 and 3). I use these proportions to estimate the parameters of the model: the cost distribution $F(\cdot)$ and the bias for the present parameter β. I match equation 32 that determines the probability of itemizing with the observed probabilities of itemizing in table 2 using the Generalized Method of Moments (GMM).

In section 4.2.4, I assume that the cost is stochastic and follows a distribution $F(\cdot)$. To estimate the model, I assume that $F(\cdot)$ is the CDF of a uniform distribution with support $[0, \alpha]$ and that $T = 20$. I estimate α and β using GMM:

	β	$(1+\alpha)$
Rational	1	12.3
		(0.01)
Naive PB	0.35	7.00
	(0.01)	(0.21)

Notice that the estimated average compliance cost of the rational taxpayer is larger than that of the present-biased taxpayer. To explain such large foregone benefits and without assuming time inconsistency, one has to assume very large aversion to filing taxes. The naive present-bias model can explain the empirical findings without assuming high costs of filing taxes.

The difference between the estimated compliance costs when assuming full rationality and when assuming time inconsistency emphasizes the importance of accurate behavioral modeling when drawing welfare implications. If taxpayers are truly rational then the estimated cost distribution is the true cost of compliance. On the other hand, if taxpayers are naive present-biased, then a portion of the burden of itemizing deductions is not due to compliance

costs per se but due to the time inconsistency of taxpayers.

This is important in two ways. First, it draws different conclusions about the magnitude of compliance costs. Second, it calls for different policy interventions. If taxpayers are rational, then the only possible intervention is to reduce true compliance costs (less record keeping, less forms etc.). If they are time inconsistent, interventions that specifically target the bias itself should also be considered.

In table 5, I calculate the aggregate cost of filing taxes assuming taxpayers are naive present-biased using the parameters derived from the GMM estimation. The rational taxpayer has an upper bound on the cost distribution equal to 12.3 and the naive present biased equal to 7. This means that assuming that taxpayers are rational implies 1.75 times larger compliance cost. This implies an aversion to tax filing coefficient of 2.23 for the naive present biased individual. The cost of compliance when the taxpayer is assumed to be naive present biased amounts to 0.66% of GDP which corresponds to 53% of the aggregate compliance costs estimated when assuming that taxpayers are not time inconsistent.

5 Alternative Explanations

5.1 Lack of Information

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

5.2 Fear of Audit and Evasion

Could it be that taxpayers believe that itemizers are more likely to be audited than individuals claiming the standard deduction? The probabilities of audit for this portion of the population are lower than 1%[38] and are virtually the same for itemizers in the neighborhood of the standard deduction and individuals who claim the standard deduction. Assume that the taxpayer has Von Neumann-Morgenstern (VnM) preferences with a Constant Relative Risk

[38]See Miller et al. (2012) and Slemrod and Gillitzer (2013).

Aversion (CRRA) utility function $U(x) = \frac{1}{1-\theta}x^{1-\theta}$. Denote by p the probability of audit, S the after tax benefit of the standard deduction, T the after tax benefit of itemized deductions, c the cost of itemizing deductions[39] and k the cost imposed by an audit on the taxpayer, which includes both a fixed cost of being audited (collecting receipts and dealing with the IRS) and the penalty that the taxpayer may have to pay. Assume that the entire portion of deductions above the standard deduction (i.e. $T - S$) is not real: taxpayers evade taxes by reporting $T - S$ fake deductions.[40] Therefore, if a taxpayer is audited, her deduction level will be brought back to S from T and she will incur cost k.

The taxpayer will itemize deductions if the expected benefit of itemizing given a probability p of facing an audit is greater than the benefit of claiming the standard deduction

$$p\left[\frac{1}{1-\theta}(S-k-c)^{1-\theta}\right] + (1-p)\left[\frac{1}{1-\theta}(T-c)^{1-\theta}\right] \geq \frac{1}{1-\theta}S^{1-\theta}. \tag{33}$$

The first term of the is the benefit derived if the taxpayer is audited: she can only deduct the standard deduction (S) and incurs the cost of itemizing (c) and the cost of evasion (k). It is multiplied by the probability of audit p. The second term is the benefit derived from itemizing: it is equal to the level of deductions T minus the cost of itemizing c and is multiplied by the probability of not being audited $(1 - p)$. Overall, the sum of these two terms is equal to the expected benefit of itemizing. The right hand side of the inequality is the benefit from claiming the standard deduction. To perform the calibration, I vary p, k, θ, assume that $c = 149$ and calculate the benefits T that a taxpayer would forego. I present the results of the calibration of this model in table C.10 and C.11 and show that for reasonable parameters, audit probabilities cannot explain the magnitude of the estimated foregone benefits.

In addition, if evasion was truly driving the result we should observe that taxpayers who itemize - even when they are close to the standard deduction threshold - have a low proportion of deductions that are easy to evade. Mortgage interest, state and local tax deductions are hard to evade because they are third party reported to the IRS. Charitable donations however are easy to evade because they are not third party reported.[41] Figure 9 shows the proportion

[39] I use the cost estimated by the IRS, which is equal to \$149 given an hourly wage of \$33.

[40] Assuming that only part of the deductions are due to evasion makes this model even less credible.

[41] Kleven et al. (2011) show that taxpayers understand that third-party-reported deductions are harder to evade and behave accordingly.

of charitable donations for taxpayers who are close to the standard deduction threshold and rejects the assumption that taxpayers switch to the standard deduction by reducing their deductions because of a fear of audit.

5.3 Bunching at Concave Kink Points

When claiming the standard deduction, taxpayers are paying the full cost of the charitable donation or mortgage payment. When they itemize however, they only pay a portion of it. The standard deduction acts as a concave kink point: the price of charitable donations is lower when itemizing. The indifference curve of a given taxpayer can be tangent at two points of the kinked budget set possibly inducing some taxpayers to switch from itemizing to claiming the standard deduction. Could the missing mass in the neighborhood of the standard deduction be due to the presence of this concave kink point? It is unlikely for the following reasons. First, other papers looked at the behavior of individuals at concave kink points and found no evidence of a behavioral response.[42] Second, I can verify that there is no behavioral response at another concave kink point for deductions by looking at the distribution of deductions for taxpayers with an AGI close to the a marginal tax rate as the change in marginal tax rate also corresponds to a concave kink point. I consider tax year 1989 when there were two tax brackets and look at the deduction that is the most likely to be responsive: charitable donations. The tax rate increases from 15% to 28% for taxpayers with an AGI that exceeds $30,950 effectively changing the price of charitable donations in a similar way to the standard deduction. If taxpayers were responding to this concave kink point, we should observe a missing mass to the left and to the right of the marginal tax rate threshold for charitable donations. There is no such missing mass in figure B.13. Third, figure 9 graphs the ratio of charitable donations to total deductions for taxpayers by their distance to the standard deduction from 1980 to 2006. If my results were driven by a behavioral response to concave kink points, we would expect a decrease in the proportion of charitable donations as a percentage of total deductions close to the standard deduction threshold. Figure 9 shows that the ratio is constant rejecting a behavioral response to concave kink points.

[42] See Saez (2010) and Kleven and Waseem (2013).

5.4 Rational Inattention

Could taxpayers forego large amounts of deductions because they are uncertain of whether their total deductions are larger than the standard deductions threshold?[43]

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

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

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

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

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

$$\mathbb{E}\left[\frac{1}{1-\theta}(\tau - c)^{1-\theta}\right] \geq \frac{1}{1-\theta}S^{1-\theta}. \tag{34}$$

This equation does not have a closed form solution, so I use a Taylor expansion of second

[43] Although in a different setting, Abeler and Jäger (2013) show that taxpayers do not seem to be rationally inattentive when responding to taxes.

degree around the mean of $\tau - c$, as follows:

$$\frac{1}{1-\theta}(\mu - c)^{1-\theta} - \frac{1}{2}\theta(\mu - c)^{-1-\theta}\sigma^2 \geq \frac{1}{1-\theta}S^{1-\theta}. \tag{35}$$

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

$$\log(\mu - c) - \frac{\sigma^2}{2(\mu - c)^2} \geq \log(S). \tag{36}$$

The first term in equation 36 is the expected benefit that the taxpayer derives from itemizing. The second term is a correction for the risk aversion of the taxpayer: she will itemize deductions if the benefit of itemizing corrected for her risk aversion is greater than the benefit she derives from itemizing. Holt and Laury (2002) find a θ that ranges between -0.95 and 1.37. I assume here that $\theta = 1$ (similar to Chetty (2006)) but also consider $0 < \theta \leq 2$[44] in table C.9. 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 $c = 149$. I can calculate a lower bound on the standard deviation of the taxpayer's beliefs over τ (σ). Using these parameters, I find that for rational inattention to explain the magnitude of the foregone benefits, the standard deviation of after tax deductions σ has to be greater than \$1,814 (which corresponds to \$6,479 worth of deductions with a 28% marginal tax rate). This means that the taxpayer has a range of uncertainty of deductions of more than \$6,479. This implies very high uncertainty in the beliefs of the benefits that the taxpayer can save from itemizing which is unlikely given that deductions are relatively stable from year to year as they are mostly constituted of mortgage payments and state taxes and are the results of active decisions. If a taxpayer's total deductions were to increase or decrease dramatically, she would most likely know about it because it would be due to for example to large income variations, the take up of a mortgage etc. which are salient.

If I assume a standard deviation of $\sigma = 200$ – which corresponds to a standard deviation of deductions of \$714 – then rational inattention with $\theta = 1$ predicts that taxpayers would claim the standard deduction up to total deductions of \$10,557 and forego an average of \$557 worth of deductions, i.e. 156 of after tax dollars given a cost c=\$149. With reasonable parameters, rational inattention does not explain the magnitude of the result.

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

5.5 Other Reforms Affecting the Distribution of Deductions?

5.5.1 The 1988 reform

A few other changes happened in 1989. In this section, I describe these changes and explain how I adjust for the ones that are likely to affect my estimates. The estimates derived in section 3.4.2 already accounted for these adjustments. The fact that the pre and post-reform densities overlap away from the standard deduction threshold shows that the pre-reform density is a relevant counterfactual for the post-reform in figure 2a density and that – after adjusting for these changes – the missing mass estimates are not affected by these changes.

The personal interest deduction was phased out starting from 1986. In 1987, taxpayers could only deduct 65% of their personal interest, 40% in 1988 and 20% in 1989. This is likely to affect the distribution of deductions from 1987 to 1989. To control for this effect, I adjust the 1987 distribution - which is the counterfactual for 1989 - by recalculating the personal interest deduction as if only 20% of it could be deducted. This leads some taxpayers to have deductions below the standard deduction whom I drop. To ensure that there is no behavioral effect associated with the phasing out of the personal interest deduction, I compare the distribution of deductions for individuals below the 28% marginal tax rate bracket and above. If there was a behavioral effect, we should observe more deductions for individuals above the 28% marginal tax bracket. Graph B.12 shows that there is no discontinuity at the marginal tax rate change at $30,950 in 1989. This is rather intuitive because the majority of the personal interest deduction is claimed for interest on student loans which are hard to adjust once they are contracted. In addition, after making this correction, I can compare the overlap between the pre and post-reform densities. Away from the standard deduction, the two graphs overlap implying that the post-reform density is an appropriate counterfactual for the 1989 density.

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

5.5.2 The 1971 reform

In 1970 taxpayers could claim as a standard deduction the smaller of $6130 or 10% of their AGI. In 1971, both thresholds were increased to $8809 or 13% of the AGI if the AGI is greater

than $46,983, and the larger of $6166 or 13% of AGI for taxpayers with an AGI smaller than $46,983.

If I were to only look at the density of itemizers above $6130 in 1970 and compare it to the density of itemizers above $8809 in 1971, my estimates would be biased because some taxpayers who have deductions greater than $8809 in 1971 are likely to stop itemizing – not because of compliance costs – but only because their deductions are now smaller than 13% of their AGI. To control for this, I only consider taxpayers whose deductions exceed 13% of their AGI and $6166 in 1970. This provides an accurate counterfactual for 1971.

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

6 Policy Implications

6.1 Cost of Compliance

Policy makers have no precise estimates of the burden of complying with the tax code. Most of the literature on the compliance cost is based on survey evidence (Slemrod and Sorum (1984) and Blumenthal and Slemrod (1992)). Taxpayers are only surveyed about the time they spend working on taxes but not on the disutility it causes them. To my knowledge, this is the first paper to use a non-parametric approach along with administrative data to reveal the preferences of taxpayers over the compliance cost. The burden of compliance is large, suggesting that the welfare lost because of compliance is substantial and of policy importance. The cost is also distortionary as it impacts individuals differently: it varies with income, type of deductions etc. which can raise equity concerns.

6.1.1 Addressing Compliance Costs

If taxpayers are truly averse to filing taxes and if preferences are indeed driving the result, the only available policy instrument is a direct reduction of the cost. This can be achieved by reducing the complexity of the tax system, having less deductions, less credits etc.

6.1.2 Addressing Time Inconsistency

With naive present-biased taxpayers, the model derived in section 4.2 shows that the policy intervention should not necessarily target the collection process itself but rather the behavior of the individual. In light of my evidence, a policy that would aim at reducing the cost of filling out forms seems misguided since the majority of the cost is precisely due to record keeping. One approach could be to require less evidence of expenses when the taxpayer itemizes. This would prove out to be efficient in reducing the compliance cost but is likely to result in more evasion. The policy maker has to trade off the cost that evasion imposes on society and the cost that compliance imposes on individuals. Tazhitdinova (2014) explores this tradeoff in the case of charitable donations.

The model derived in section 4.2 also shows that there are relatively inexpensive policy interventions that can significantly reduce the cost of compliance. Advocates of pre-populated forms argue that they are likely to reduce evasion and mistakes by taxpayers. My results show that pre-populated forms are also likely to improve welfare by reducing compliance costs. Two of the three most common deductions are mortgage interest payments and state and local income taxes. Both are third-party reported implying that the IRS knows the amount of deductions that the taxpayer qualifies for. Gillitzer and Skov (2013) show that the introduction of pre-populated forms in Denmark increases claimed deductions consistent with the fact that taxpayers fail to claim deductions for expenses they incurred because of compliance costs. Kotakorpi and Laamanen (2013) show that pre-populated deductions increase but non pre-populated deductions decrease in Finland with the introduction of pre-populated forms.

The use of electronic receipts is another channel through which record keeping costs can be further reduced.[45] Some employers issue form W2 online and some banks provide an electronic 1098. Keeping track of an electronic document can be much easier than a paper one. However, this would only benefit taxpayers who have access to the Internet possibly creating further inequalities.

Given that taxpayers tend to procrastinate on filing their taxes and wait until April 15th, potentially facing a large cost of itemizing on that day, the policymaker can ensure that the

[45] Kopczuk and Pop-Eleches (2007) show, for example, that electronic filing increases the EITC take-up rate.

deadline for filing taxes falls on a day when people are likely to be less busy such as the weekend. The IRS actually has the opposite policy: if April 15th is a weekend day, the deadline is postponed to the next Monday. This was probably relevant when e-filing was not available and taxpayers had to visit the post-office to send their returns, but less relevant now with the prevalence of e-filing.

To address the issue of the increasing marginal disutility of labor, the IRS can have two deadlines: one on April 15th for the 1040 form and one a week later for Schedule A. Having only one deadline can result in the taxpayer filing all her taxes on the same day, but having two allows the naive present-biased taxpayer to smooth effort over time. This policy would help naive present-biased taxpayers but would not hurt rational ones as they would still be able to file their taxes on the same day if they want to. Similarly, the IRS could also coordinate with local governments to ensure that the deadlines for state and federal taxes do not overlap.

Naive present-bias justifies shifting the burden of compliance from individuals to firms. Firms are less likely to be subject to psychological biases. For example, in the case of charitable donations, requiring charitable organizations to report donations and then having the IRS send a statement to taxpayers (or pre-populate Schedule A) will result in less compliance burden on aggregate because firms are less likely to be time inconsistent and more likely to have a system of information that deals with receipts in a systematic way.

6.1.3 Should the Compliance Cost Be Reduced?

Kaplow (1998) argues that some compliance cost can be efficient when designing a tax system. The social gains of deductions are unclear and some advocate that they are relatively small.[46] If political economy concerns prevent the government from repelling these deductions, one way of ensuring that taxpayers do not claim them is to impose significant hassle costs.

6.2 Screening Literature

There is a long tradition in public economics that emphasizes the benefits of conditioning transfers on fixed characteristics and more particularly imposing transaction costs when providing welfare to screen out richer households. To my knowledge there was no empirical

[46] See Slemrod and Bakija (2008) for example.

evidence confirming that hassle costs are larger for richer households. The results in this paper show that it is the case and that such policy can be efficient. It warns however that transaction costs need to be chosen with care as they can be relatively large and can end up screening out more income groups than optimal. They can also screen out naive present-biased taxpayers versus rational ones rather poor taxpayers versus poorer ones.

6.3 Details Matter

Similarly to Saez (2009) this paper also shows that details matter in designing tax systems. Details that in theory should not have any impact can result in significant behavioral distortions. In this case, the time at which a receipt is sent to a taxpayer should not matter much if it is archived as soon it is received. But if taxpayers procrastinate on archiving them, it can lead to significant differences.

7 Conclusion

Research on the compliance cost of taxation has struggled with estimating the burden of compliance. And the literature on the failure to take up government benefits has not shown that hassle costs can lead individuals to leave benefits on the table.

Using a quasi-experimental design and a novel method to recover the counterfactual density of deductions, I find that taxpayers who fail to itemize forego large amounts of deductions, resulting in an average burden of itemizing of $617. This implies a compliance burden of a much larger magnitude than previously estimated.

If one is to believe that taxpayers are truly behaving rationally when making the decision to forego deductions, then aggregate compliance costs are large: back-of-the-envelope calculations suggest an order of magnitude of the total cost of complying with the federal tax code of 1.25% of GDP.

If taxpayers are instead present-biased, the axiom of revealed preferences fails, introducing a wedge between compliance cost and foregone benefits. I provide evidence to support this explanation.

Both explanations however suggest that the burden of compliance imposed on taxpayers is high. I offer policy recommendations to reduce it.

References

Abeler, J. and S. Jäger (2013): "Complex Tax Incentives-An Experimental Investigation," *American Economic Journal: Applied Microeconomics*, Forthcoming.

Ausubel, L. M. (1999): "Adverse Selection In the Credit Card Market," Working paper, University of Maryland.

Bhargava, S. and D. Manoli (2011): "Why Are Benefits Left On the Table? Assessing the Role Of Information, Complexity, and Stigma On Take-Up With an IRS Field Experiment," Working paper.

Blank, R. M. and D. E. Card (1991): "Recent Trends in Insured and Uninsured Unemployment: Is There an Explanation?" *The Quarterly Journal of Economics*, 106, 1157 – 1189.

Blumenthal, M. and J. Slemrod (1992): "The Compliance Cost of the US Individual Income Tax System: A Second Look After Tax Reform," *National Tax Journal*, 185 –202.

Chetty, R. (2006): "A new method of estimating risk aversion," *The American Economic Review*, 1821–1834.

Currie, J. (2004): "The Take Up of Social Benefits," *NBER working paper 10488*.

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

DellaVigna, S. and U. Malmendier (2006): "Paying Not to Go to the Gym," *The American Economic Review*, 694 –719.

DellaVigna, S. and M. D. Paserman (2005): "Job Search and Impatience," *Journal of Labor Economics*, 23.

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

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

Ericson, K. M. M. (2011): "Forgetting We Forget: Overconfidence and Memory," *Journal of the European Economic Association*, 9, 43 –60.

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

Feldstein, M. (1999): "Tax Avoidance and the Deadweight Loss of the Income Tax," *Review of Economics and Statistics*, 81, 674 – 680.

Gelber, A., D. Jones, and D. Sacks (2013): "Earnings Adjustment to Policy: Evidence From the Social Security Earnings Test," *NBER Working Paper 19491*.

Gillitzer, C. and P. E. Skov (2013): "Evidence On Unclaimed Charitable Contributions From the Introduction of Third-Party Information Reporting in Denmark," Working paper.

Gruber, J. and B. Köszegi (2001): "Is Addiction Rational? Theory and Evidence," *The Quarterly Journal of Economics*, 116, 1261–1303.

Guyton, J. L., J. F. O'Hare, M. P. Stavrianos, and E. J. Toder (2003): "Estimating the Compliance Cost of the US Individual Income Tax," *National Tax Journal*, 673 –688.

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

Inmar (2012): "2013 Coupon Trends," Year end report.

Jones, D. (2012): "Inertia and Overwithholding: Explaining the Prevalence of Income Tax Refunds," *American Economic Journal: Economic Policy*, 4, 158 –185.

Kaplow, L. (1998): "Accuracy, Complexity, and the Income Tax," *Journal of Law, Economics, and Organization*, 14, 61–83.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011): "Unwilling Or Unable to Cheat? Evidence From a Tax Audit Experiment In Denmark," *Econometrica*, 651–692.

Kleven, H. J. and W. Kopczuk (2011): "Transfer program complexity and the take-up of social benefits," *American Economic Journal: Economic Policy*, 3, 54–90.

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.

Kopczuk, W. and C. Pop-Eleches (2007): "Electronic filing, tax preparers and participation in the Earned Income Tax Credit," *Journal of Public Economics*, 91, 1351 –1367.

Koszegi, B. and M. Rabin (2007): "Mistakes in Choice-Based Welfare Analysis," *The American Economic Review*, 97, 477 – 481.

——— (2008a): "Choices, Situations and Happiness," *Journal of Public Economics*, 92, 1821 – 1832.

——— (2008b): "Revealed Mistakes and Revealed Preferences," in *The Foundations of Positive and Normative Economics: a Handbook*, 193 –209.

Kotakorpi, K. and J.-P. Laamanen (2013): "Complexity, Salience and Income Tax Filing Behavior: Evidence From a Natural Experiment," Manuscript, Working Paper, University of Turku.

Laibson, D. (1997): "Golden Eggs and Hyperbolic Discounting," *The Quarterly Journal of Economics*, 443–477.

Letzler, R. and J. Tasoff (2014): "Everyone Believes in Redemption: Overoptimism and Nudges," *Journal of Economic Behavior and Organization*, 107–122.

Madrian, B. C. and D. F. Shea (2001): "The Power of Suggestion: Inertia in 401 (k) Participation and Savings Behavior," *The Quarterly Journal of Economics*, 116, 1149–1187.

Miller, S., R. Marcuss, S. Boehmer, B. Johnson, and W. Kei (2012): "Internal Revenue Service Data Book," Annual report.

Nichols, D., E. Smolensky, and T. N. Tideman (1971): "Discrimination By Waiting Time In Merit Goods," *The American Economic Review*, 312 –323.

O'Donoghue, T. and M. Rabin (1999a): "Doing It Now Or Later," *American Economic Review*, 103 –124.

——— (1999b): "Incentives For Procrastinators," *Quarterly Journal of Economics*, 769 –816.

——— (2001): "Choice and Procrastination," *Quarterly Journal of Economics*, 121 –160.

——— (2008): "Procrastination On Long-Term Projects," *Journal of Economic Behavior & Organization*, 66, 161 –175.

Phelps, E. S. and R. A. Pollak (1968): "On second-best national saving and game-equilibrium growth," *The Review of Economic Studies*, 185–199.

Pitt, M. 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, 1486–1488.

Rees-Jones, A. (2013): "Loss Aversion Motivates Tax Sheltering: Evidence From US Tax Returns," Working paper.

Saez, E. (2009): "Details Matter: The Impact Of Presentation and Information On the Take-Up of Financial Incentives For Retirement Saving," *American Economic Journal: Economic Policy*, 1, 204 –228.

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

Shapiro, J. M. (2005): "Is there a daily discount rate? Evidence from the food stamp nutrition cycle," *Journal of Public Economics*, 89, 303–325.

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

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

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

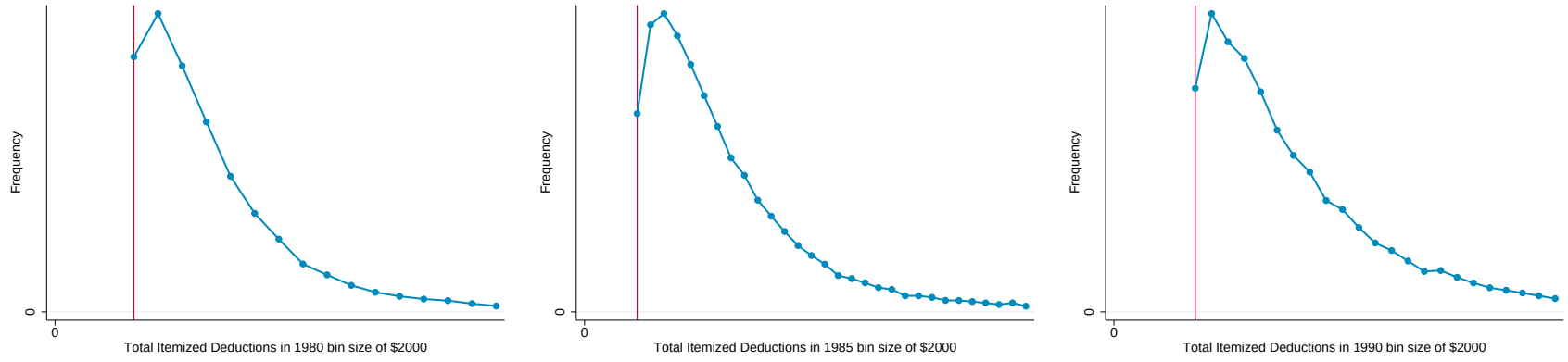
Slemrod, J., C. Christian, R. London, and J. A. Parker (1997): "April 15 syndrome," *Economic Inquiry*, 35, 695–709.

Slemrod, J. and C. Gillitzer (2013): *Tax Systems*, MIT Press.

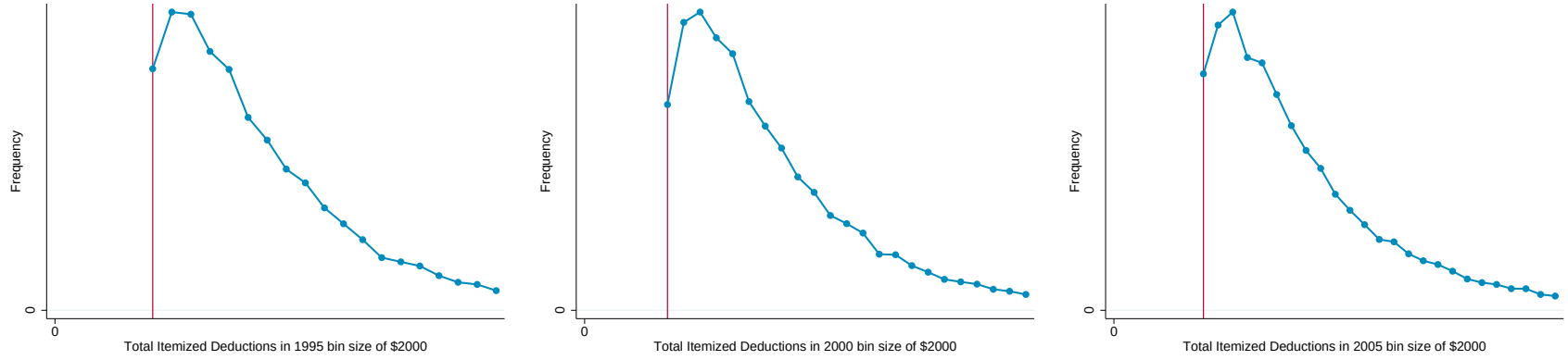
Slemrod, J. and N. Sorum (1984): "The Compliance Cost of the US Individual Income Tax System," *National Tax Journal*, 461 –474.

Tazhitdinova, A. (2014): "Optimal Reporting Threshold: Theory and Evidence From Charitable Contributions," Manuscript, University of California, Berkeley.

Figure 1: Missing Mass In the Neighborhood of the Standard Deduction



(a) 1980 **(b) 1985** **(c) 1990**

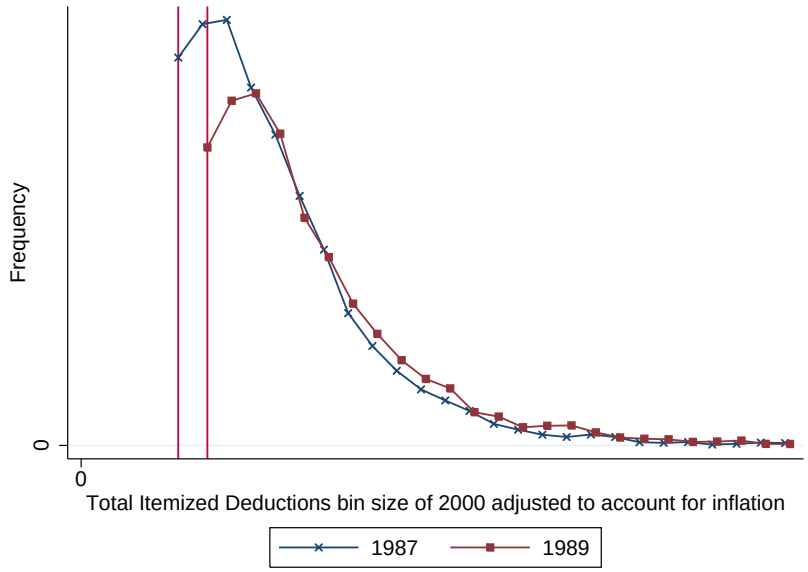


(d) 1995 **(e) 2000** **(f) 2005**

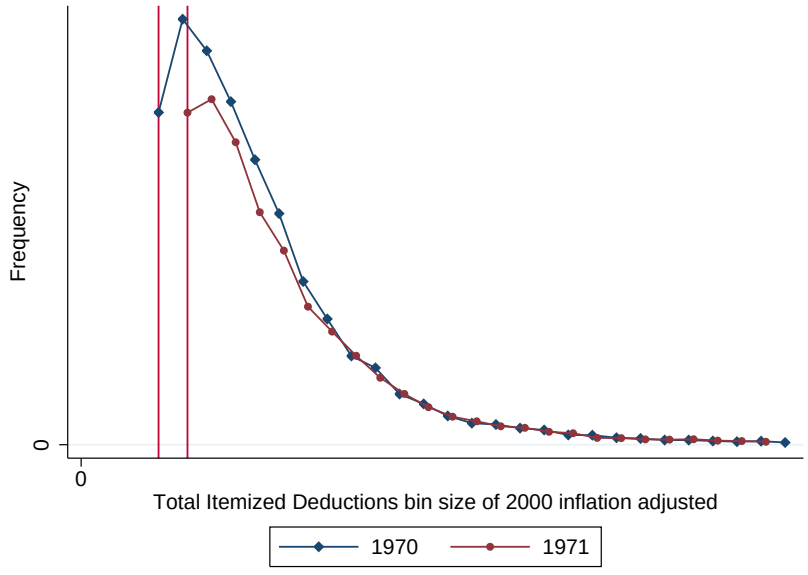
Notes: The figures above plot the density of deductions for itemizers filing jointly. The bin size is $2,000 and the vertical line represents the standard deduction threshold for each year. Notice the missing mass in the neighborhood of the standard deduction threshold. Additional years are reported in appendix figures B.2, B.3, B.4 and B.5.

Figure 2: Density of Deductions for Itemizers Filing Jointly Before and After the Standard Deduction Is Increased

(a) The 1988 Reform

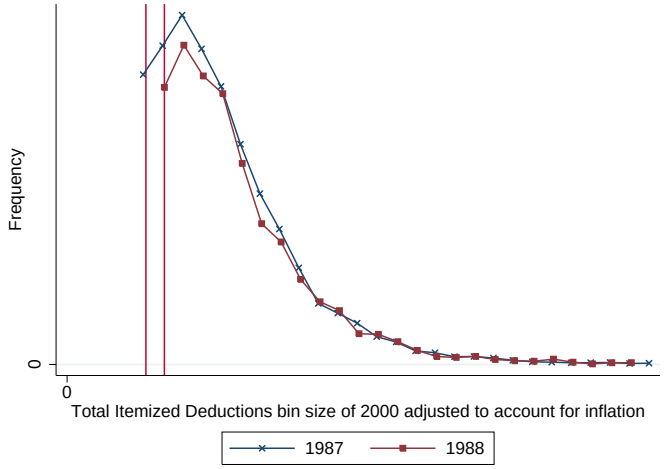


(b) The 1971 Reform



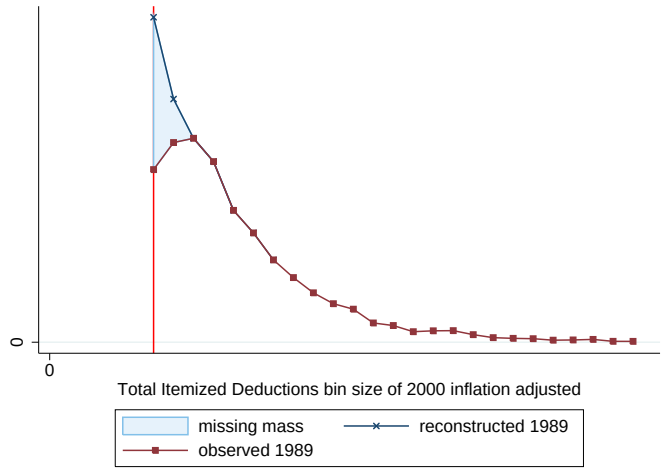
Notes: The first graph plots the density of deductions for the 1988 reform and the second one for the 1971 reform. Notice that the pre-reform density is higher than the post-reform density specifically in the neighborhood of the standard deduction, whereas the two densities are very similar when comparing them further away from the standard deduction. The statistical difference between the two densities is reported on appendix tables C.7 for 1988 and C.8 for 1971.

Figure 3: Lagged Response: Small Effect During Reform Year (1987-1988)



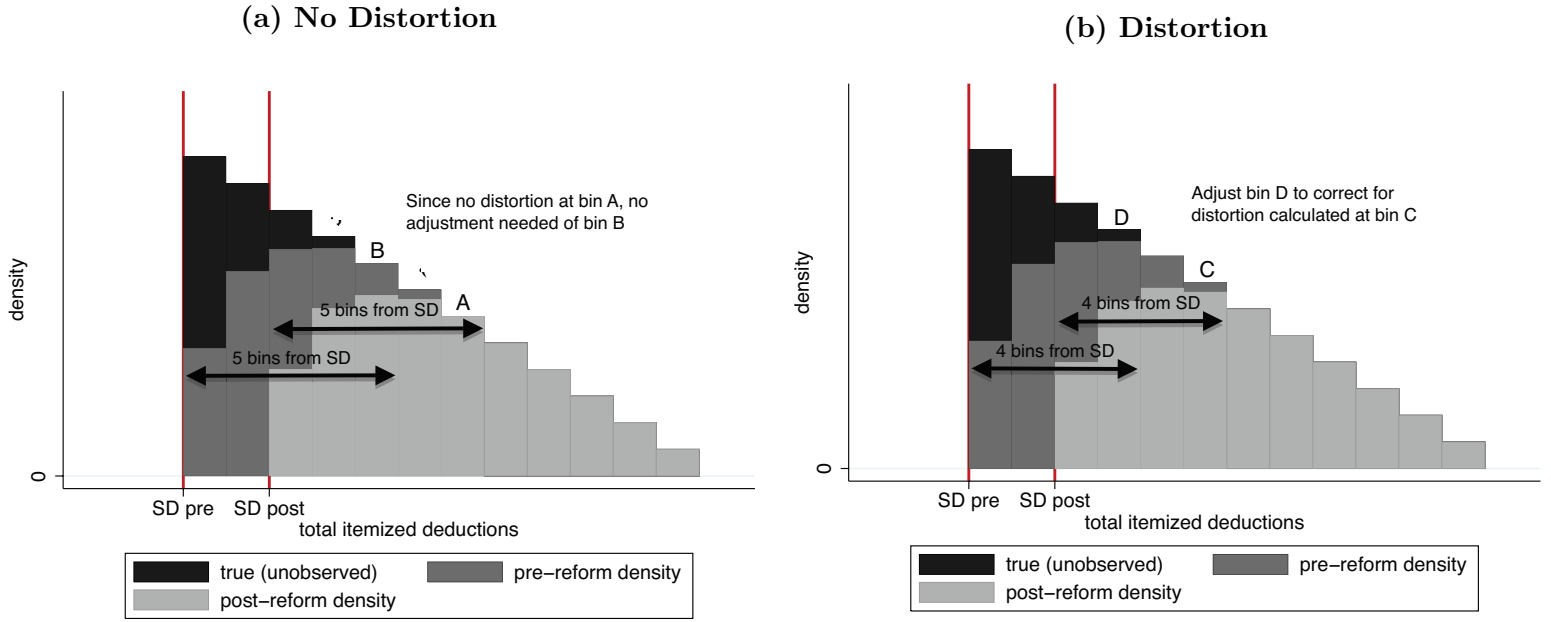
Notes: This graph plots the distribution of deductions for itemizers filing jointly in 1987 and 1988. Notice that the missing mass is smaller than in figure 2a showing that there is a lagged response to the reform.

Figure 4: Reconstructed Density and Missing Mass in 1989



Notes: This graph plots the reconstructed density in 1989 using the method that I outline in section 3.3 and the observed density for 1989. The missing mass that allows me to estimate the burden of itemizing is given by the area lying between the two curves. The distribution of the burden of itemizing is provided in table 2 for a bin size of $2,000 and table 3 for a bin size of $500.

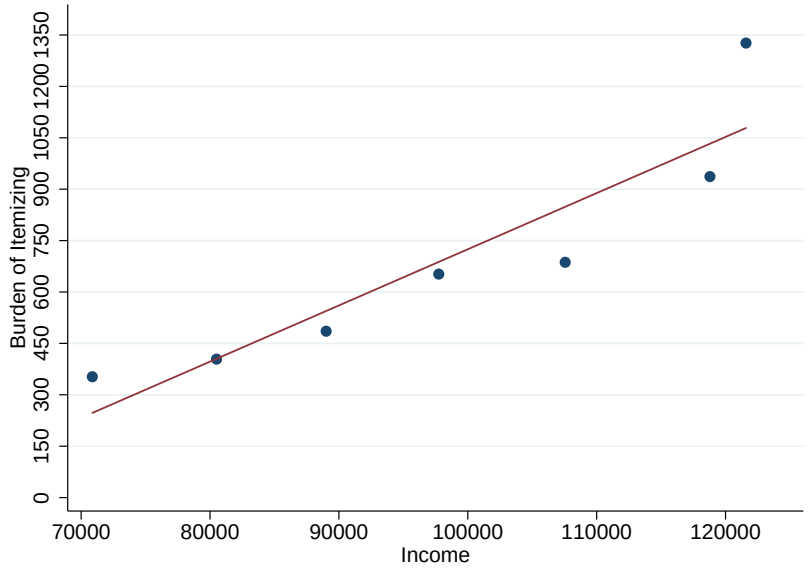
Figure 5: Reconstructing the Counterfactual Density



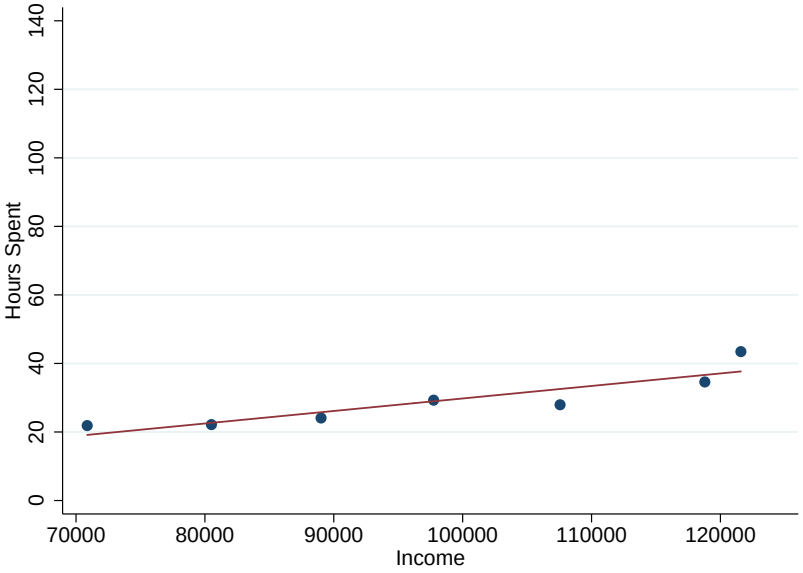
Notes: These graphs illustrate the method that I use to reconstruct the counterfactual density. The darkest histograms correspond to the true density of deductions when assuming that there is no cost. The next shade corresponds to the pre-reform year and the lightest one to the post-reform year. The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

Figure 6: Relationship Between Income and the Burden of Itemizing Deductions

(a) Burden of Itemizing and Income



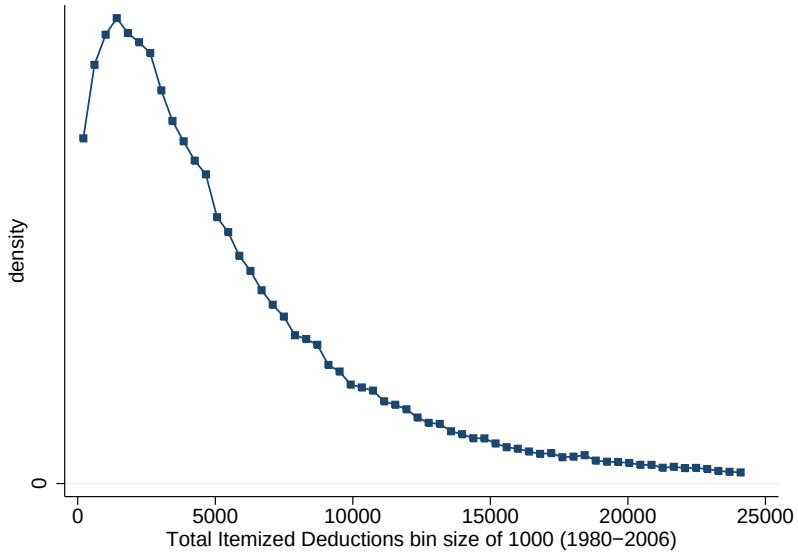
(b) Hours Spent Itemizing and Income



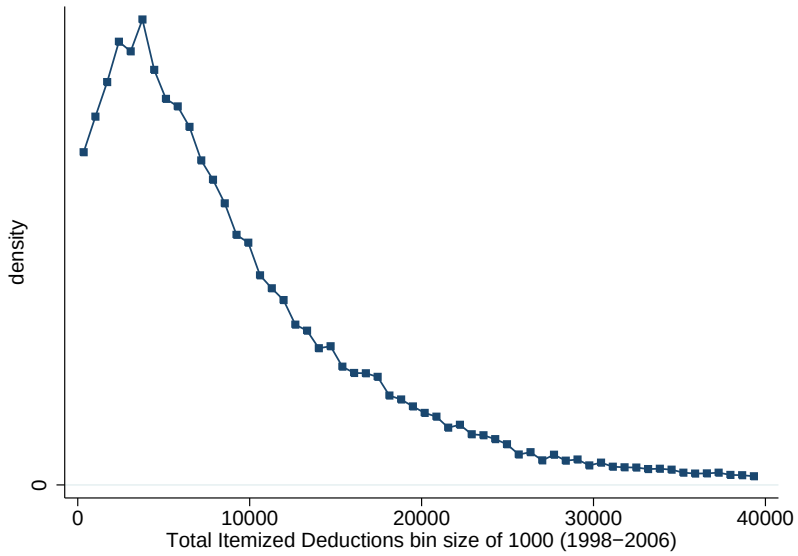
Notes: (a) The first graph shows the increasing relationship between income and the burden of itemizing: richer households are more likely to forego deductions (b) The second graph divides the burden of itemizing by the hourly wage of each household and shows the hours spent itemizing by each income group.

Figure 7: Use of Tax Preparer and Electronic Filing

(a) Tax Preparer



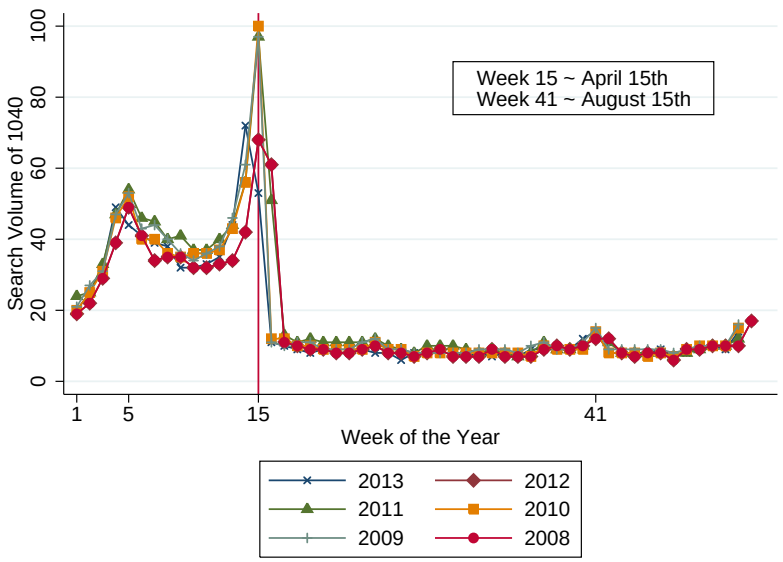
(b) Electronic Filing



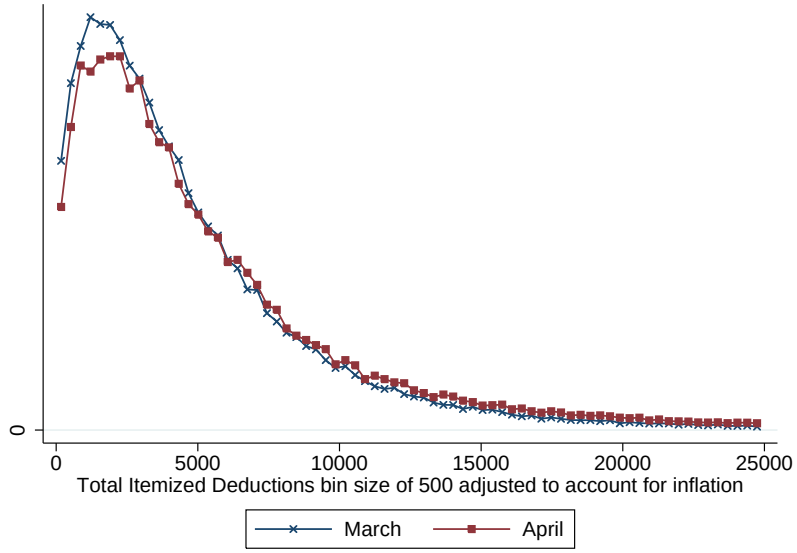
Notes: Graph (a) plots the density of total deductions for taxpayers who use tax preparers from 1980 to 2006 (excluding 1985 and 1990 because the variable is not available in those years) by bin size of $1000. Graph (b) plots the density of total deductions for taxpayers who file returns electronically from 1998 to 2006 (few taxpayers used electronic filing prior to 1998) by bin size of $1000. Both graphs exhibit a significant missing mass close to the standard deduction implying that neither tax preparers nor electronic filing eliminate the burden of itemizing.

Figure 8: Deadline Effects

(a) Google Search of the Term 1040

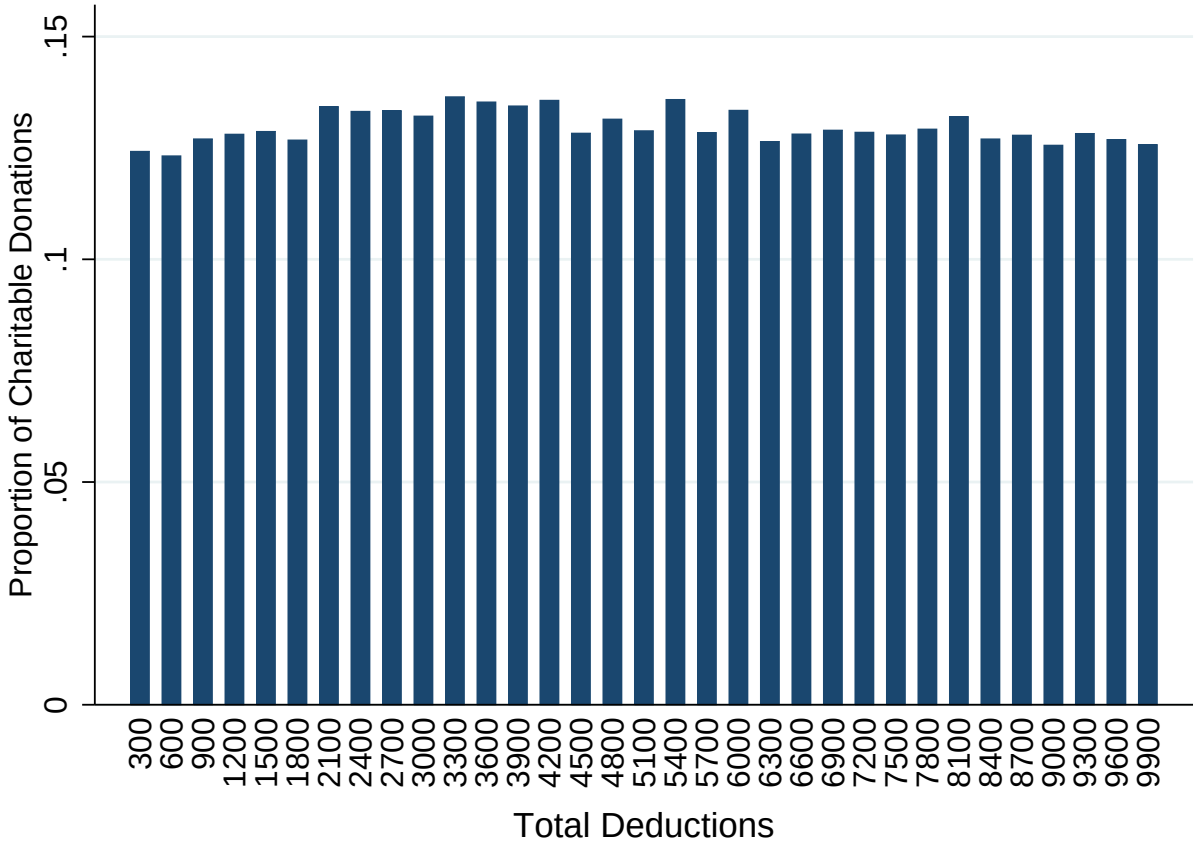


(b) March Itemizers v.s. April Itemizers



Notes: The first graph plots the volume of search of the term "*1040*" in Google. The x-axis is in weeks. April 15th typically falls in week 15, while August 15th – which corresponds to the filing extension deadline – falls in week 41. Notice the spike in search on week 15 and the spike in week 41 consistent with the prediction of the naive present-bias model that time inconsistency leads taxpayers to file at the last moment. The second graph plots the density of itemizers who file their returns in March and the density of those who file in April. The dataset is constituted of 20 repeated cross section (1980 to 1999: years in which the variable is available) pooled together. Consistent with the prediction of the naive present-bias model, April filers tend to forego more deductions in the neighborhood of the standard deduction than March filers.

Figure 9: Proportion of Charitable Donation Deductions By Distance to the Standard Deduction Threshold



Notes: This graph shows the proportion of deductions that are charitable donations for itemizers from 1980 to 2006 by their distance to the standard deduction. Deductions are adjusted for inflation and the standard deduction amount is subtracted from them to calculate the distance to the standard deduction. The proportion of charitable donations does not change close to the standard deduction threshold implying that taxpayers do not respond to the change in the standard deduction by reducing their charitable donations. This rules out the explanations of the missing mass based on the behavioral response to a concave kink point and evasion.

Table 1: Determinants of the Likelihood of Switching to the Standard Deduction

Outcome:	Switch to the standard deduction: {0,1}					
	(1)	(2)	(3)	(4)	(5)	(6)
newborn x close to SD	0.03				0.05**	0.05**
	(0.02)				(0.02)	(0.02)
newborn	-0.02*				0.01	-0.01
	(0.01)				(0.01)	(0.01)
easy ded. x close to SD		-0.12***			-0.12***	-0.12***
		(0.01)			(0.01)	(0.01)
easy ded.		-0.12***			-0.18***	-0.11***
		(0.01)			(0.01)	(0.01)
no preparer x close to SD			-0.00		0.00	-0.00
			(0.01)		(0.01)	(0.01)
no preparer			-0.00		-0.01	0.00
			(0.01)		(0.01)	(0.01)
late x close to SD				0.04***	0.03**	0.04***
				(0.01)	(0.01)	(0.01)
late				0.02***	0.01	0.01
				(0.01)	(0.01)	(0.01)
close	0.06***	0.06***	0.06***	0.05***	0.10***	0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Controls	Yes	Yes	Yes	Yes	No	Yes
R^2	0.264	0.307	0.264	0.267	0.160	0.309
N	11799	11799	11799	11799	11799	11799
Clusters (individual)	4659	4659	4659	4659	4659	4659

Notes: Each cell reports an estimate from a separate regression of the likelihood of switching from itemizing to claiming the standard deduction. **newborn** indicates whether the family experienced a birth during the previous tax year. **easy ded.** is equal to 1 if the sum of mortgage interest and state tax deduction is greater than the standard deduction. **late** indicates that the return has been processed by the IRS later than week 18. **no preparer** indicates that the individual did not use the help of a tax preparer during the previous filing season. **close to SD** indicates that the taxpayer's total deductions were close to the standard deduction threshold the previous year. I include controls for the number of children, the level of deductions in year t-1, income polynomial, marital status, state, marginal tax rate, year t-1 and t-2 medical expenses.

Table 2: Cumulative Distribution Function of the Burden of Itemizing (bin size of $2,000)

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

Table 3: Cumulative Distribution Function of the Burden of Itemizing (bin size of $1,000)

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

Notes: These two tables report the Cumulative Distribution Function (CDF) and Probability Density Function (PDF) of the perceived burden of itemizing deductions. Table 2 uses a bin size of $2,000 and table 3 uses a bin size of $1,000. The first column corresponds to deductions, and the second to the after tax deductions. For example, the second row of table 3 corresponds to taxpayers who can save $1,000 to $2,000 of deductions, which is on average $1,500 of tax deductions and corresponds to $420 with 28% marginal tax rate. The CDF is calculated by comparing the proportion of taxpayers who itemize and those who fail to itemize. In the first row for example, 43% of taxpayers itemize implying that their perceived burden of itemizing is less than $1,000 of deductions. The average burden of itemizing is a weighted average given by the product of the average benefit and the PDF. b_k, m_k and c_k refer to the notation used in section 3.3

Table 4: Aggregate Burden of Filing Taxes Assuming Taxpayer Is Rational

Form	Hours (from IRS)	Hourly Wage (in $)	Tax-Aversion Coefficient	Individual Burden (in $)	Nb. of Taxpayers (million)	Aggregate Burden (in $b.)	% of GDP
1040	9.4	17.7	4.22	702	112	78.6	0.72
Sch. A	4.53	34.7	4.22	663	32.0	21.2	0.20
Sch. B	1.28	22.3	4.22	120	12.4	1.49	0.01
Sch. C	9.63	22.1	4.22	899	14.0	12.6	0.12
Sch. D	3.75	37.1	4.22	586	8.52	5.00	0.05
Sch. E	5.83	35.7	4.22	877	14.2	12.5	0.11
Sch. F	16.1	21.4	4.22	1456	2.36	3.44	0.03
Sch. SE	1.13	17.0	4.22	81.1	11.6	0.94	0.01
Total						**135**	**1.25**

Table 5: Aggregate Burden of Filing Taxes Assuming Taxpayer Is Naive Present-Biased

Form	Hours (from IRS)	Hourly Wage (in $)	Tax-Aversion Coefficient	Individual Burden (in $)	Nb. of Taxpayers (million)	Aggregate Burden (in $b.)	% of GDP
1040	9.4	17.7	2.23	515	112	57.7	0.38
Sch. A	4.53	34.7	2.23	487	32.0	15.6	0.10
Sch. B	1.28	22.3	2.23	88.2	12.4	1.09	0.01
Sch. C	9.63	22.1	2.23	660	14.0	9.24	0.06
Sch. D	3.75	37.1	2.23	430	8.52	3.67	0.02
Sch. E	5.83	35.7	2.23	644	14.2	9.14	0.06
Sch. F	16.1	21.4	2.23	1068	2.36	2.52	0.02
Sch. SE	1.13	17.0	2.23	59.5	11.6	0.69	0.01
Total						**71.7**	**0.66**

Notes: This table reports the aggregate burden of filing each item of the income tax return. The hours are reported by the IRS in the documentation accompanying the 1040 form (the one used here is from 1989). The hourly wage is estimated from the reported wage in the SOI dataset by restricting the sample to individuals with positive wages and dividing the annual wage by 2000 hours. The number of taxpayers is reported in the SOI files. The aggregate burden is equal to the product of the hours required to file the form, the estimated aversion to filing coefficient, the average hourly wage and the number of taxpayers. The GDP in 1989 was 10.9 trillion dollars (in 2014 dollars). I derive the coefficient of aversion to filing taxes for the naive present-biased taxpayer from the model estimation.

A APPENDIX

A.1 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 truly 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.

A.2 Taxpayers Who Have To Claim the Standard Deduction

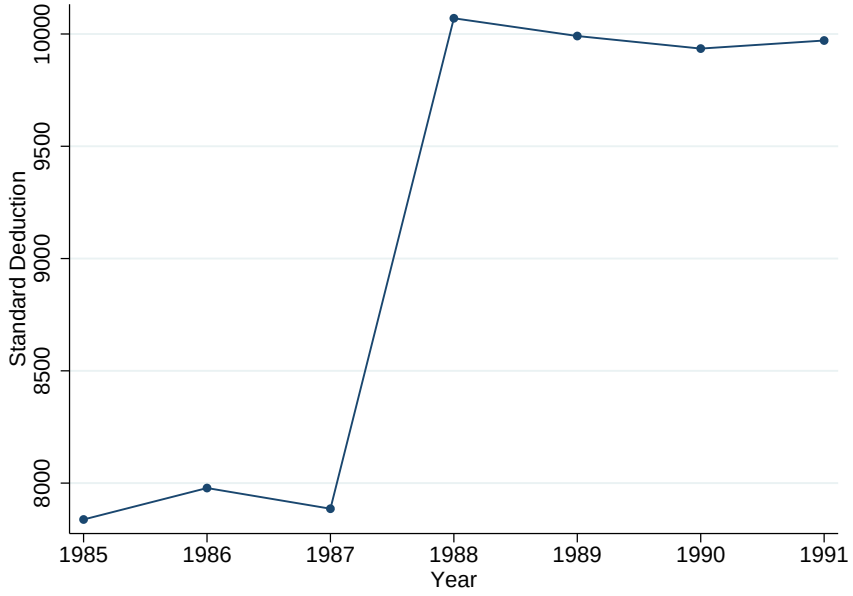
In rare cases, taxpayers have to claim the standard deduction even when their itemized deductions exceed the standard deduction. This happens in the following four cases:

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

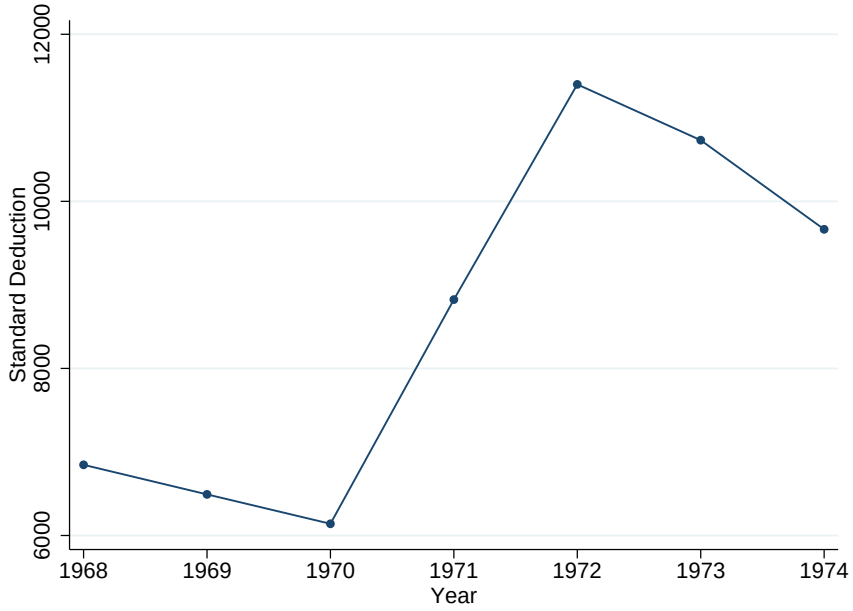
B APPENDIX FIGURES

Figure B.1: Reforms

(a) 1988 reform

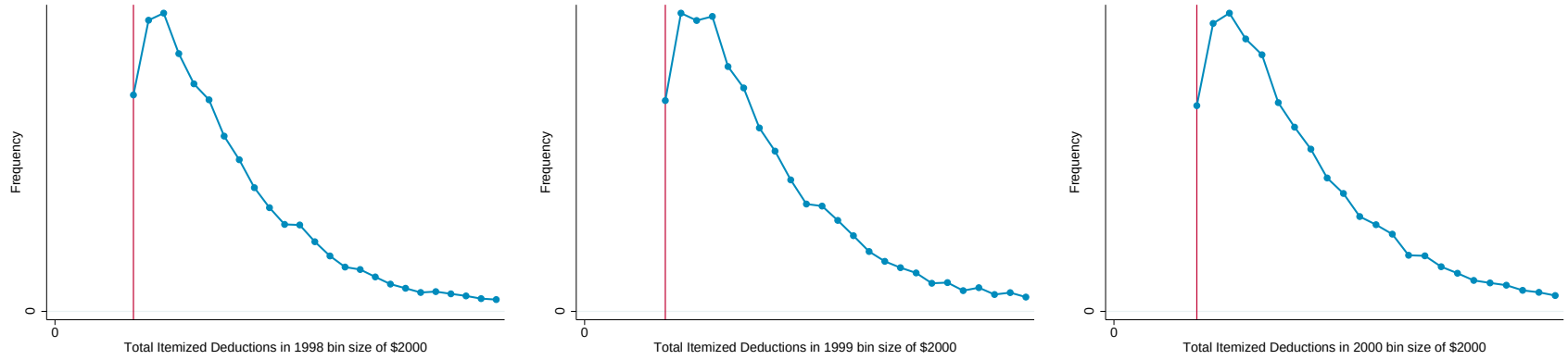


(b) 1971 reform

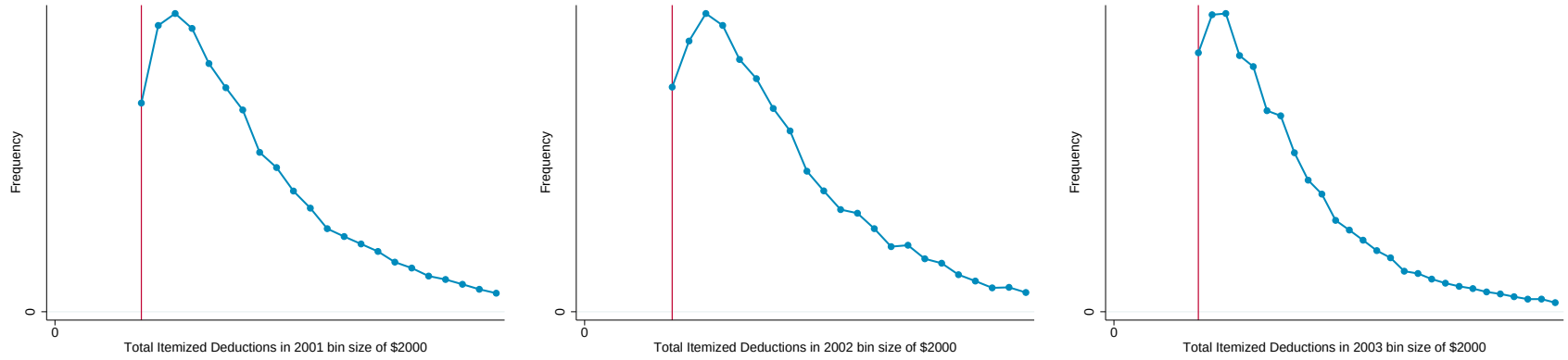


Notes: These graphs plot the level of the standard deduction by year. My identification strategy exploits the large increases in the standard deduction amount in 1988 and 1971. The amount of the standard deduction drops after 1972 because it is fixed in nominal terms. See appendix table C.6 for details.

Figure B.2: Missing Mass In the Neighborhood of the Standard Deduction 1998-2003



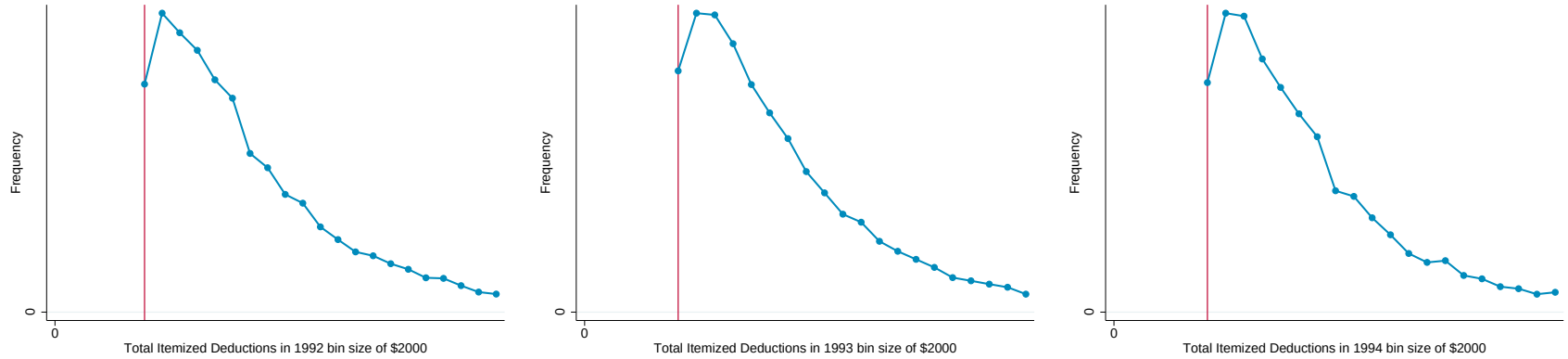
(a) 1998 **(b) 1999** **(c) 2000**



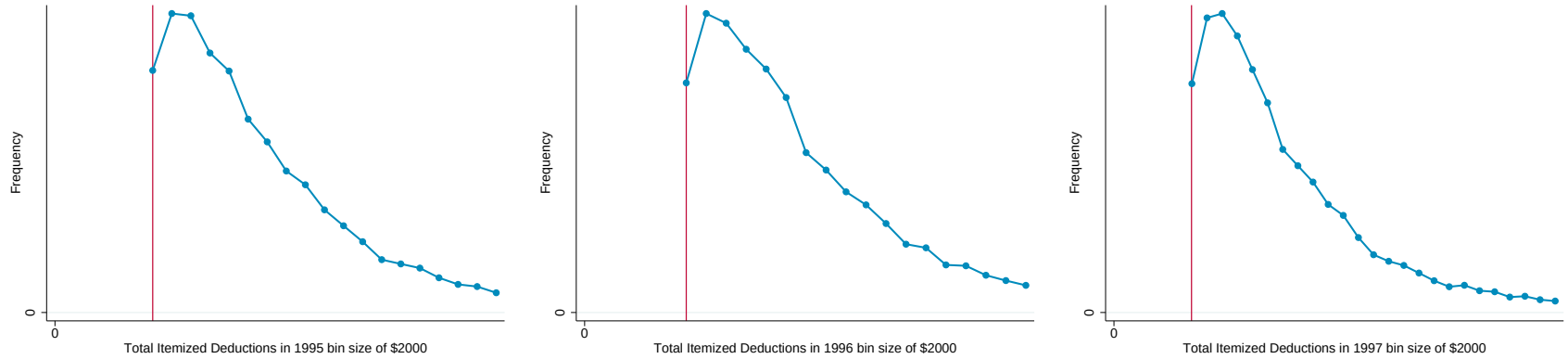
(d) 2001 **(e) 2002** **(f) 2003**

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

Figure B.3: Missing Mass In the Neighborhood of the Standard Deduction 1992-1997



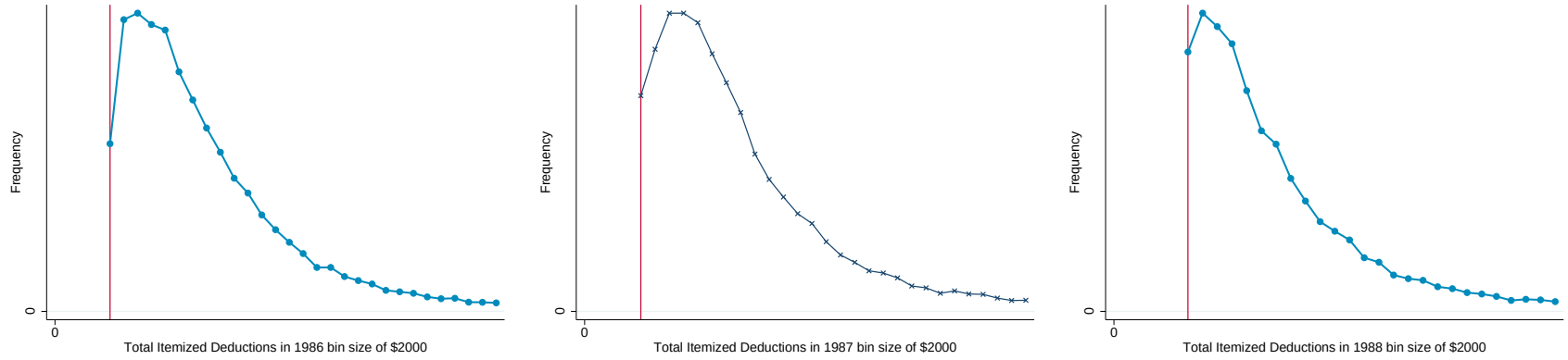
(a) 1992 **(b) 1993** **(c) 1994**



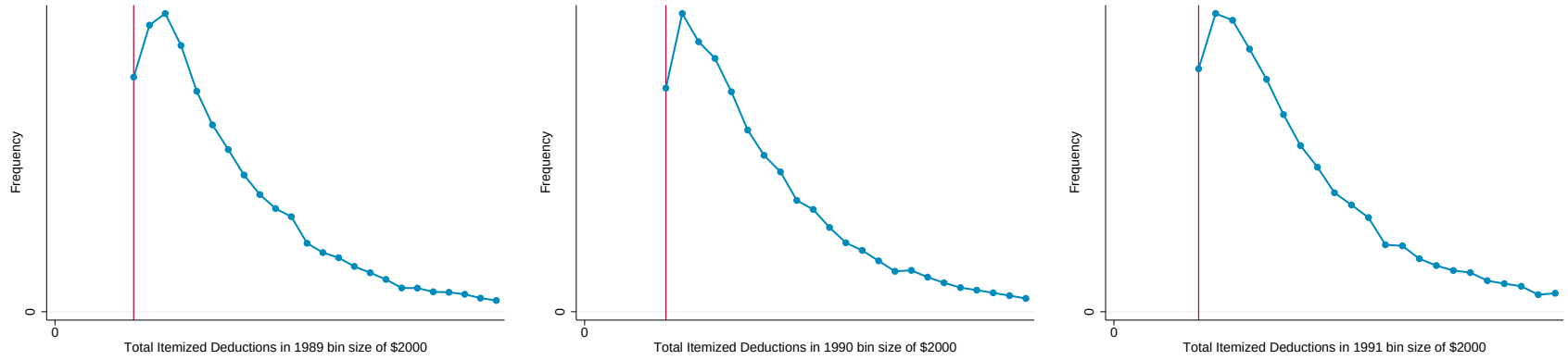
(d) 1995 **(e) 1996** **(f) 1997**

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

Figure B.4: Missing Mass In the Neighborhood of the Standard Deduction 1986-1991



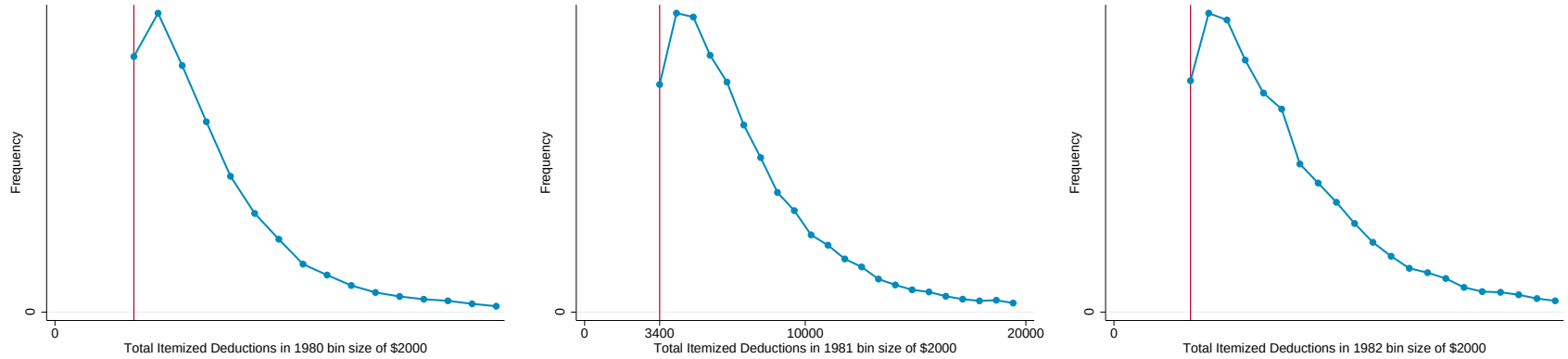
(a) 1986 **(b) 1987** **(c) 1988**



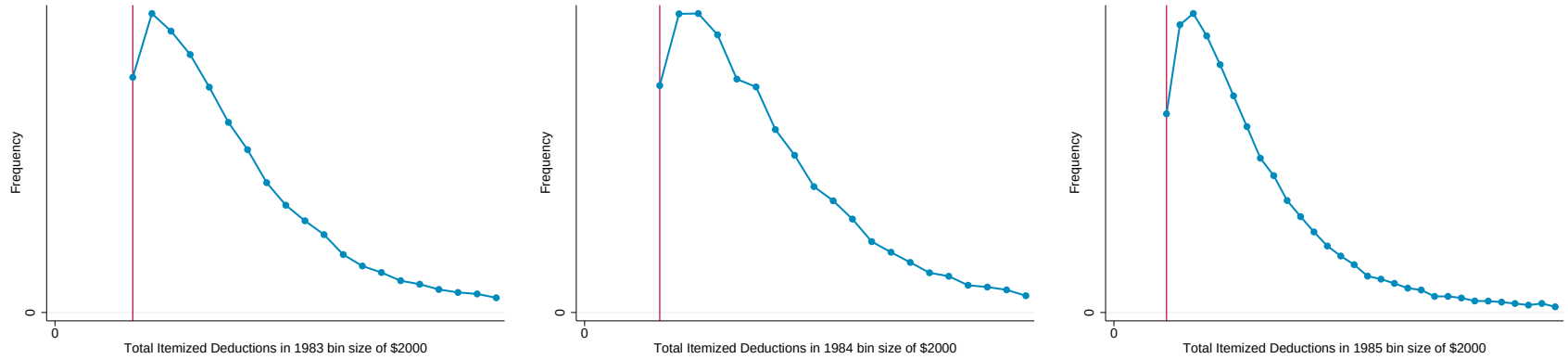
(d) 1989 **(e) 1990** **(f) 1991**

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

Figure B.5: Missing Mass In the Neighborhood of the Standard Deduction 1980-1985



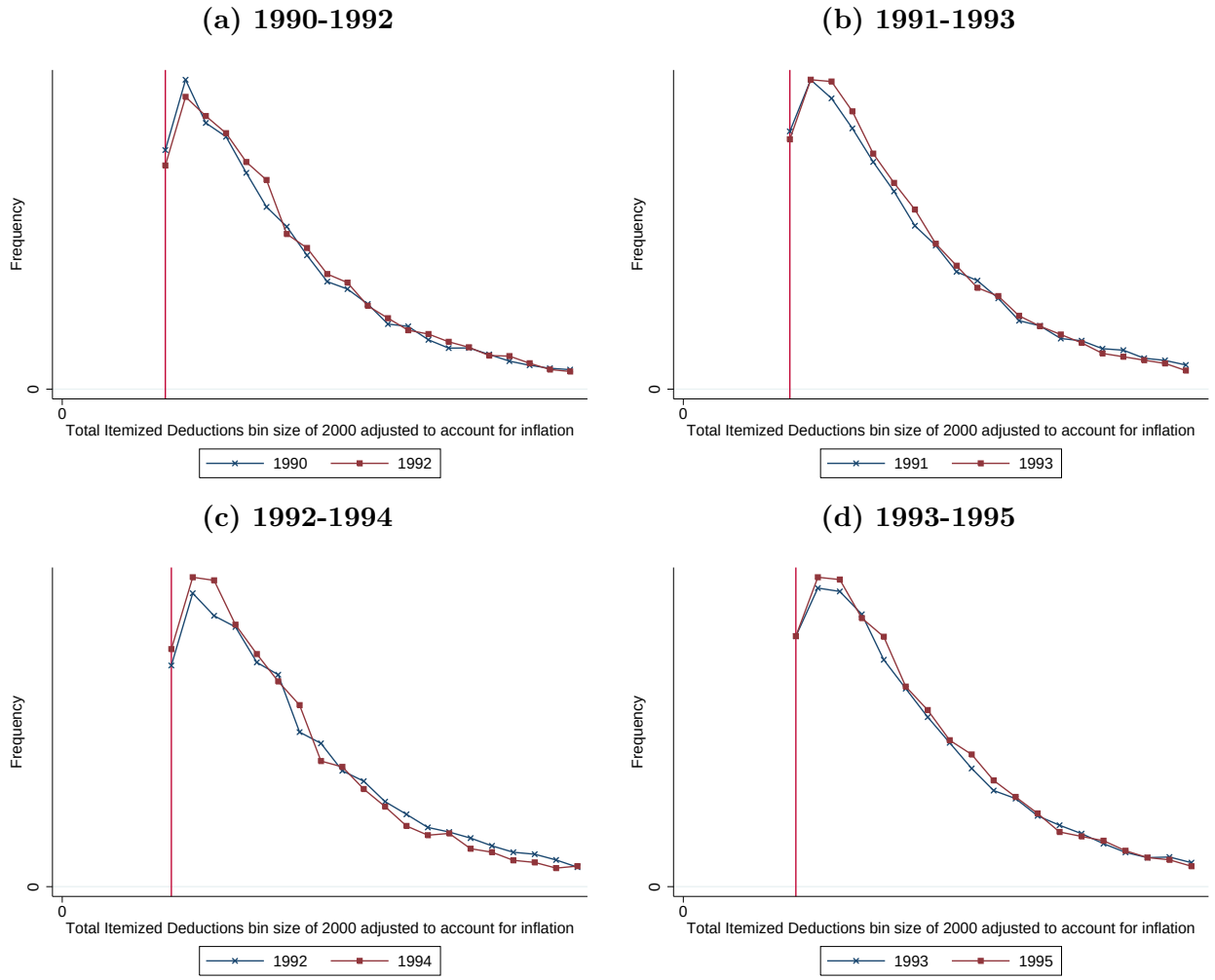
(a) 1980 **(b) 1981** **(c) 1982**



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

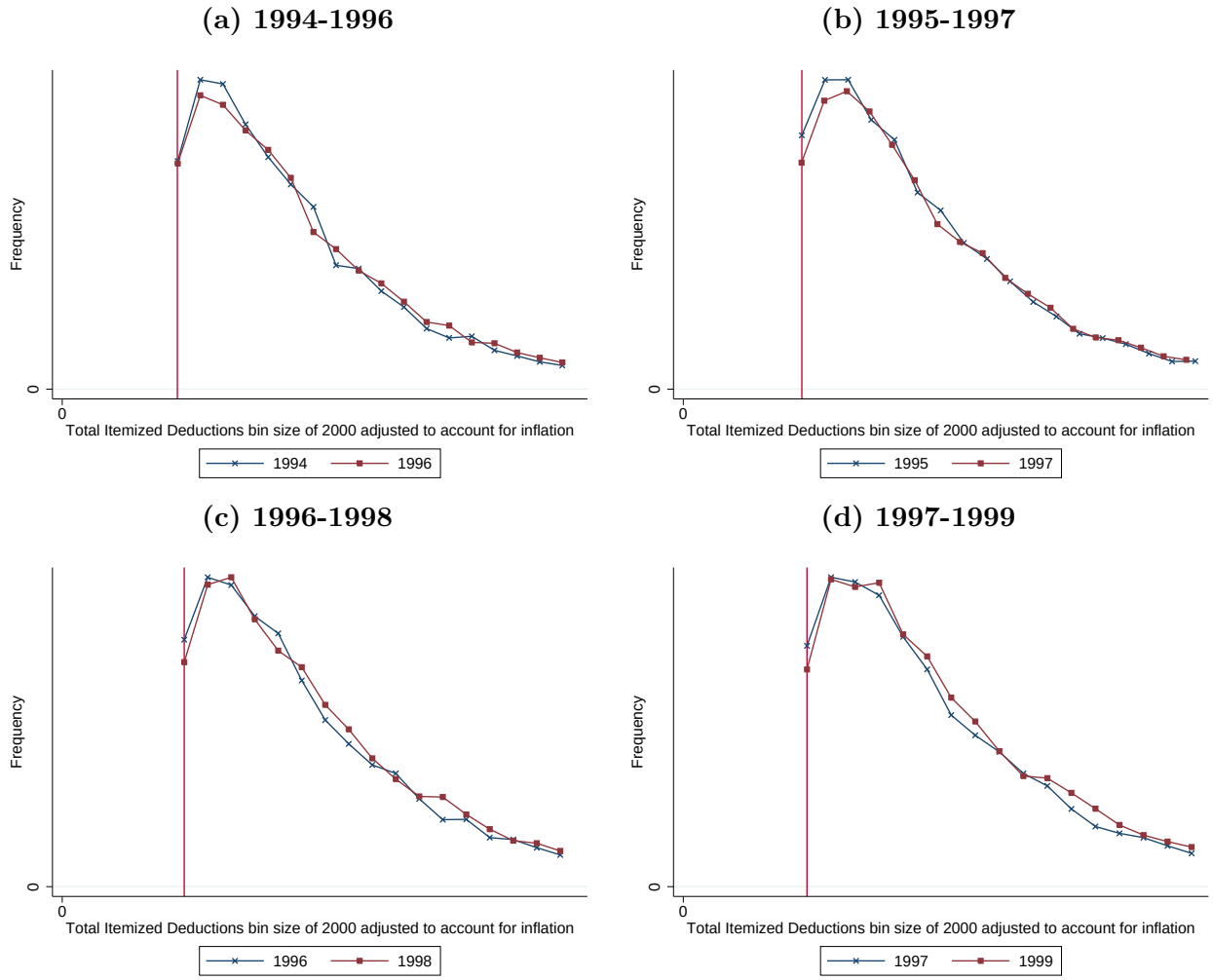
Notes: The figures above plot the density of deductions for itemizers filing jointly. The bin size is $2,000 and the vertical line represents the standard deduction threshold for each year. Notice the missing mass in the neighborhood of the standard deduction threshold.

Figure B.6: Placebo Test: Overlapping Densities In Years With No Reforms



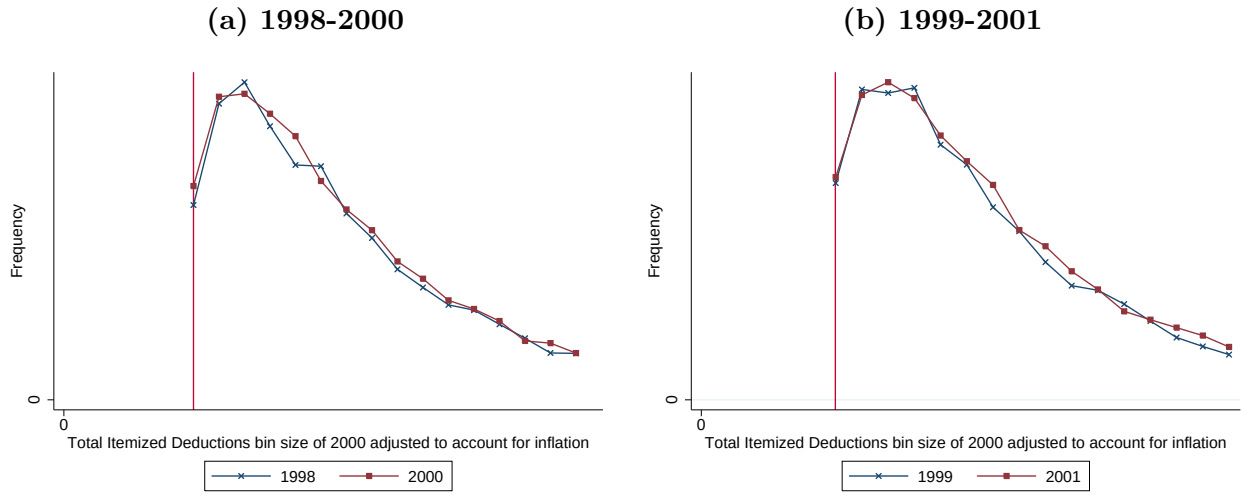
Notes: The figures above plot the density of deductions for itemizers filing jointly in years with no reforms of the standard deduction. Notice that there is no missing mass in the neighborhood of the standard deduction.

Figure B.7: Placebo Test: Overlapping Densities In Years With No Reforms



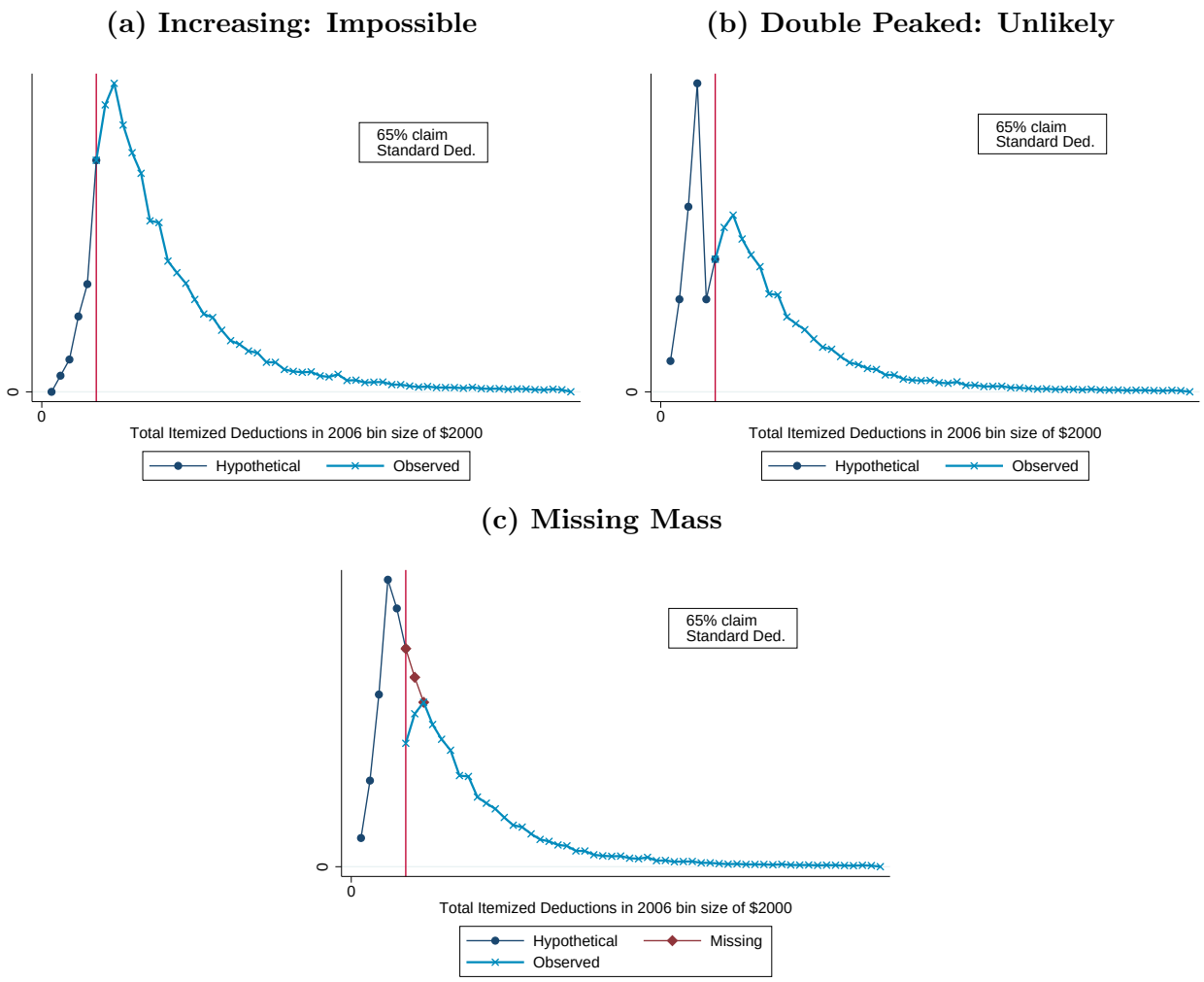
Notes: The figures above plot the density of deductions for itemizers filing jointly in years with no reforms of the standard deduction. Notice that there is no missing mass in the neighborhood of the standard deduction.

Figure B.8: Placebo Test: Overlapping Densities In Years With No Reforms



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

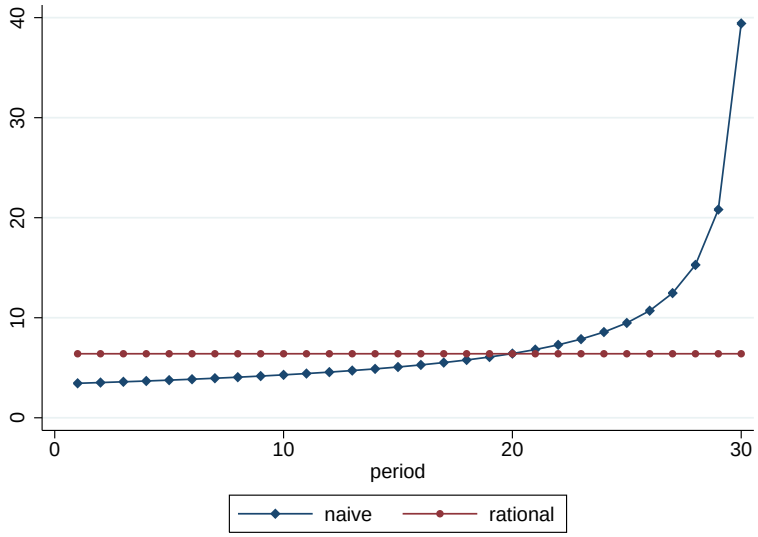
Figure B.9: Different Scenarios Below the Standard Deduction



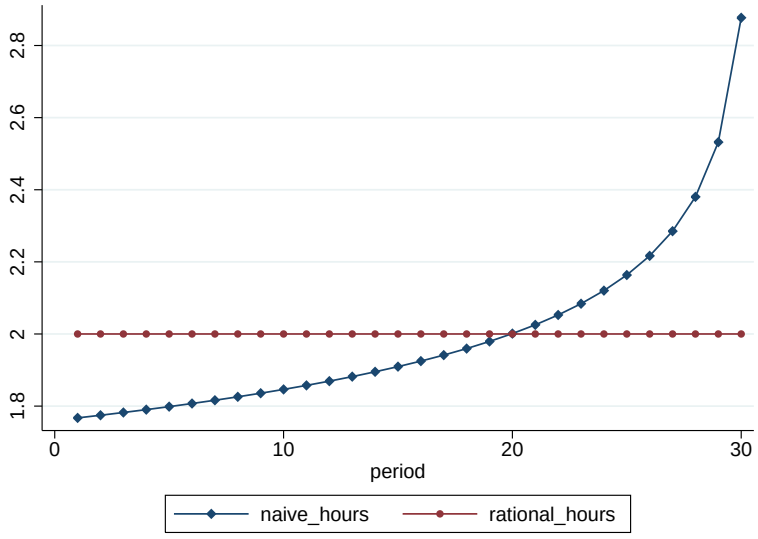
Notes: The graphs above plot the different scenarios that could be happening below the standard deduction. Graph (a) assumes that the density is strictly increasing, which is impossible given that 65% of taxpayers claim the standard deduction. This scenario would fail to account for most of the population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

Figure B.10: Calibration of Model With Convex Effort Costs

(a) Per-Period Disutility

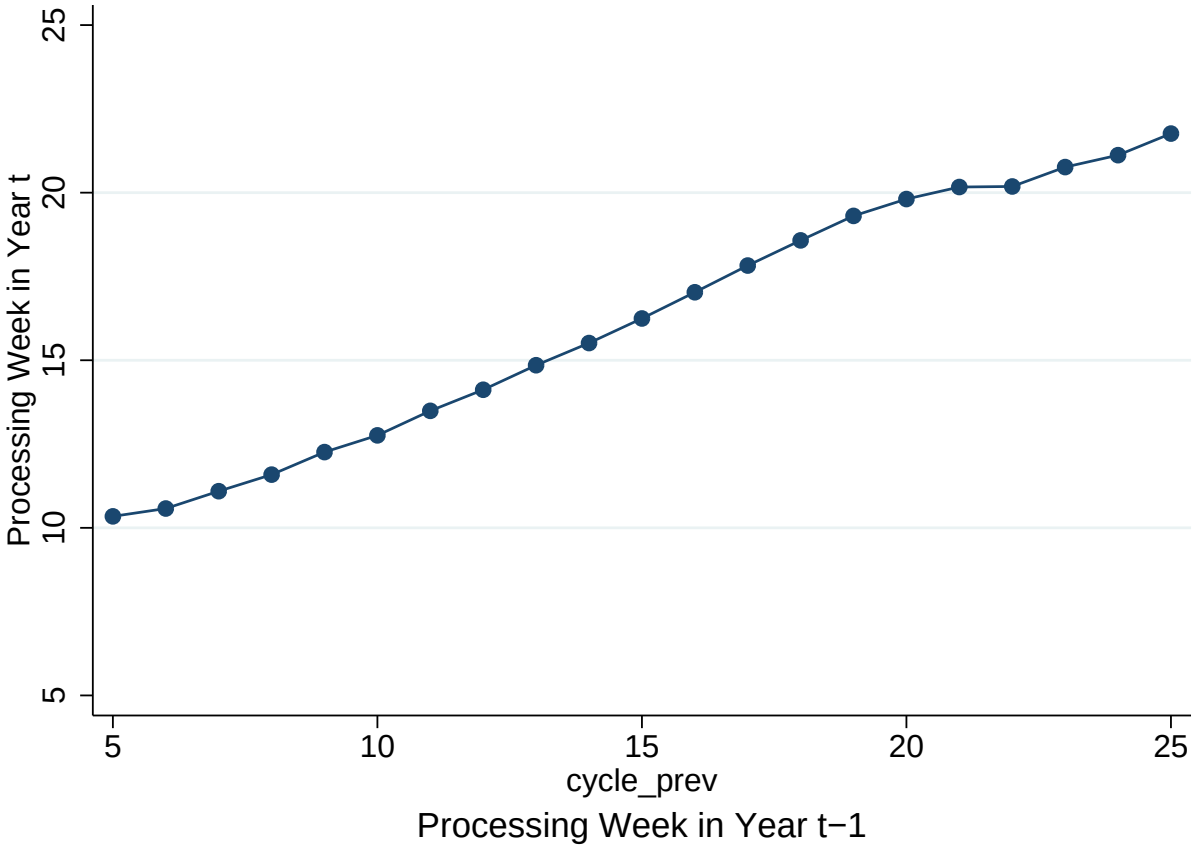


(b) Per-Period Effort



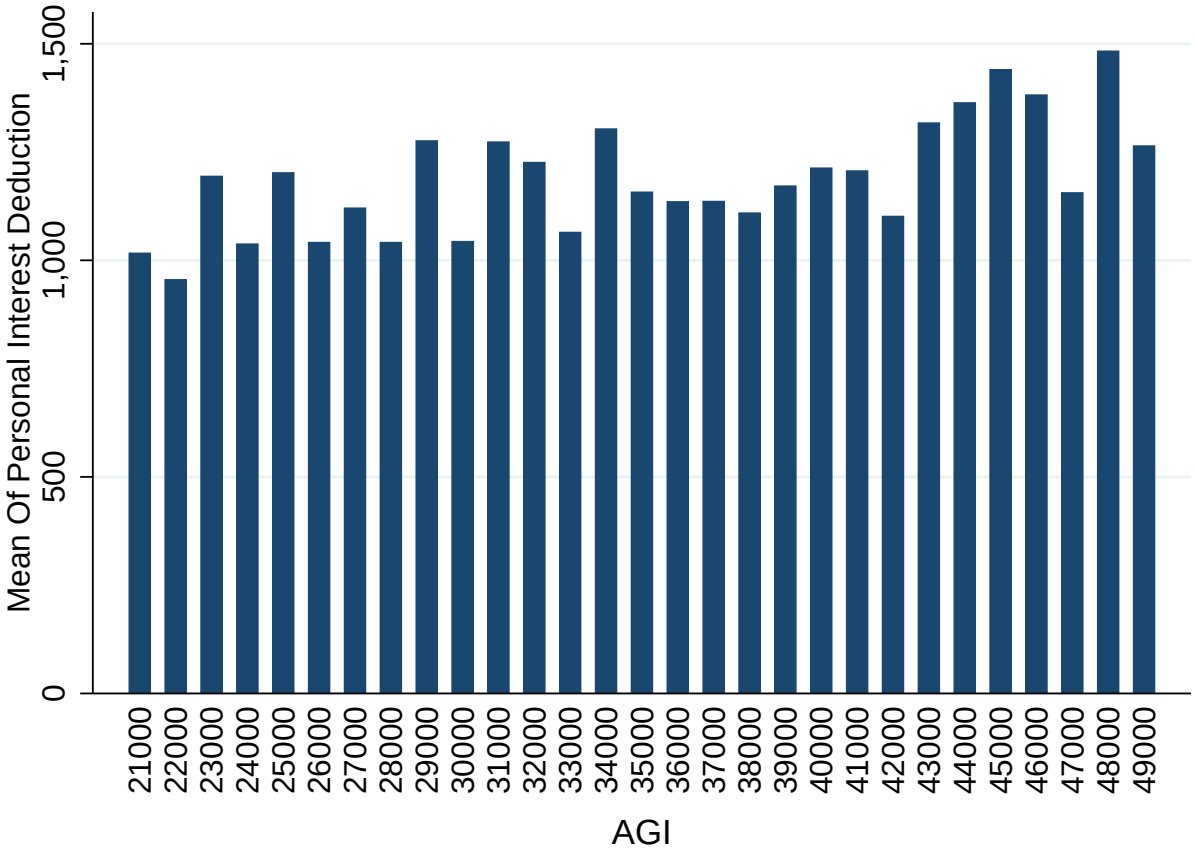
Notes: These two graphs are the result of a calibration of the model outlined in section 4.2.3. The first graph shows that the per-period disutility experienced by the naive present-biased taxpayer is higher than for the rational one. The second graph shows that the rational taxpayer smoothes effort over time whereas the naive present-biased one spends a lot of time filing taxes closer to the deadline.

Figure B.11: Processing Week in Year t v.s. Year $t-1$



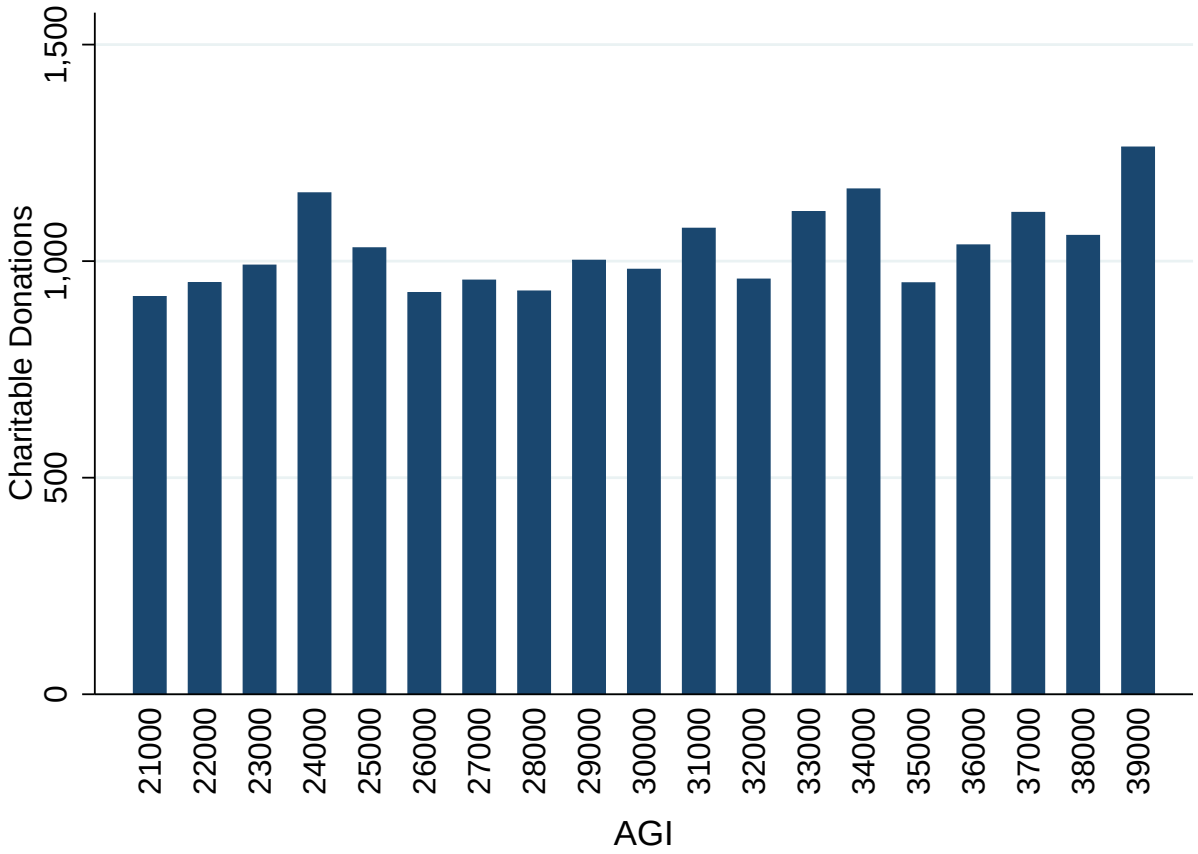
Notes: This graph plots the average week in which a return is processed in year t on the y-axis and the average week in which a return is processed in year $t-1$ on the x-axis. The relationship is increasing implying that taxpayers who file late in year $t-1$ are more likely to file late in year t consistent with the predictions of the naive present-bias model.

Figure B.12: 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 B.13: No Behavioral Response of Deductions to Concave Kink Points Formed by MTR Variation



Notes: The graph above plots the average charitable donation by AGI in 1989. The marginal tax rate is 15% for taxpayers with an AGI of less than $30,950 and 28% for taxpayers with an AGI above $30,950. Similarly to the standard deduction, an increase in marginal tax rate leads to a concave kink point in the budget set of taxpayers creating two tangency points with their indifference curve. Theoretically, this could lead to a missing mass close to the concave kink point. This graph shows that this prediction does not hold. Taxpayers do not reduce their deductions because of concave kink points. I focus on charitable donations because they are likely to be more responsive than state taxes or mortgage interest.

C APPENDIX TABLES

Table C.6: Standard Deduction By Year For Joint Filers

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

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

Table C.7: Standard Errors of the Difference Between the 1987 and 1989 Densities

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
11	(29991, 31991]	-0.00034**	0.00017	-2.00

Table C.8: Standard Errors of the Difference Between the 1970 and 1971 Densities

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

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

Table C.9: Calibration of Rational Inattention Model

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

Notes: This table shows the results of a calibration of the Rational Inattention model derived in section 5.4. Rational inattention cannot explain the magnitude of the foregone benefits unless one assumes that the standard deviation of the savings is greater than \$2000, which implies a standard deviation of deductions of \$7,143. Such high uncertainty is extremely unlikely given that deductions are stable and changes are usually due to active decisions on the part of the taxpayer (increase in income, take up of home mortgage etc.).

Table C.10: Calibration of Fear of Audit Model With True Audit Probabilities

	Cost k of audit in dollars						
	50	100	200	500	1000	1500	2000
CRRA coefficient							
0.5	151	152	153	156	162	169	178
1	151	152	153	157	164	174	191
1.5	151	152	153	157	166	181	211

Table C.11: Calibration of Fear of Audit Model With Inflated Audit Probabilities

	Cost k of audit in dollars						
	50	100	200	500	1000	1500	2000
CRRA coefficient							
0.5	200	213	240	325	484	674	920
1	201	215	244	340	544	846	1381
1.5	202	217	248	357	626	1134	2581

Notes: Table C.10 and C.11 calibrate a model based on evasion and fear of audit. The first table assumes the true audit probabilities (1%), the second table assumes 20 times the true audit probabilities. I assume that taxpayers are evading all their deductions in excess of the standard deduction. If they are audited by the IRS, their deductions are brought back to the standard deduction. In addition, they have to pay a cost k that includes both the hassle of an audit and the penalties they are charged. Note that penalties are usually waived by the IRS if the taxpayer fails to provide a receipt and argues that it was lost. Even assuming that one needs a full day of work to comply with an audit, this would imply a cost of \$264 which implies that any cost of being audited k above \$500 is hard to believe. The only way that audit probabilities can explain the magnitude of the foregone benefit is by assuming at the same time very large audit probabilities, high risk-aversion and high costs of audit.