

TAX ENFORCEMENT USING A HYBRID BETWEEN SELF- AND THIRD-PARTY REPORTING

Sarah Clifford and Panos Mavrokonstantis[†]

September 2021

Abstract

We study a tax enforcement policy combining elements of self- and third-party reporting. Taxpayers self-report to the authority but must file documentation issued by a third-party to corroborate claims. Exploiting salary-dependent cutoffs governing documentation requirements when claiming deductions for charitable contributions in Cyprus, we estimate that deductions increase by £0.7 when taxpayers can claim £1 more without documentation. Second, using a retroactive reform we find that at least 64% of the response is purely a reporting adjustment representing mainly over-reporting of deductions. Finally, reporting rules drive the behaviour of many taxpayers who display little responsiveness to financial incentives for giving.

Keywords: *Tax enforcement · Tax compliance · Charitable giving · Tax design*

[†]Clifford: University of Oxford, Department of Economics, Nuffield College and The Centre for Business Taxation, 1 New Road, Oxford OX1 1NF, United Kingdom (e-mail: sarah.clifford@economics.ox.ac.uk), Mavrokonstantis: Department of Economics, London School of Economics, 32 Lincoln's Inn Fields, London, WC2A 3PH, United Kingdom (e-mail: p.mavrokonstantis@lse.ac.uk). We thank the editor, Keith Marzilli Ericson, and three anonymous referees for very helpful suggestions. We are also grateful to Frank Cowell, Artur Doshchyn, Niels Johannesen, Claus T. Kreiner, Camille Landais, Alexandra Charalambides, Andreas Peichl, Viktorie Sevchenko, and seminar and conference participants at the University of Oxford, University of Copenhagen, University of Reading and the 35th annual congress of the EEA for helpful comments and suggestions. We would like to thank the Tax Department of the Republic of Cyprus for providing access to their data, and especially Elina Symeonidou for preparing and explaining the data. We also thank the Ministry of Labour, Welfare and Social Insurance for providing access to their archive of trade unions' collective bargaining agreements. The views expressed are those of the authors and do not necessarily reflect the views of the Tax Department or Ministry of Labour, Welfare and Social Insurance of the Republic of Cyprus. All remaining errors are those of the authors.

1 Introduction

Tax enforcement is an essential part of all tax systems, and a central aspect of the enforcement mechanism is how tax information is reported and verified.¹ This paper studies the effect on tax compliance of a simple documentation requirement when taxpayers report their information to the tax authorities. We study this policy in the context of a semi-formal economy where enforcement is of key importance.²

Different methods of reporting are implemented globally, depending on the level of development and the administrative capacity of each country. Third-party reporting is commonly viewed as the gold-standard in the context of advanced economies (Kleven et al., 2011).³ In such a system, both the tax filer and a third-party provide information regarding a given claim directly to the fiscal authority. At the other extreme is pure self-reporting, where the sole provider of information is the tax filer. In this paper, we study a widely-used hybrid policy, which combines elements of both of these systems. The policy prescribes that only the filers themselves report their information to the tax authority, but also requires that they attach documentation, issued by another involved party, to prove their claim. It therefore falls short of third-party reporting as the third party does not itself provide any corroborating information directly to the authorities. However by requiring some documentation, it partially maintains the corroborative role of the third party. At the same time the policy does not require the existence of sophisticated institutions and allows the government to avoid the large investment in data infrastructure necessary to implement direct third-party reporting. It is therefore an alternative in many situations where more elaborate policies are infeasible or too costly. This simple policy is observed in many different countries and contexts. In countries such as Germany and Austria it is used in the context of deductions for work-related expenses, while Russia uses the policy for enforcement of the tax relief for foreign earned income. In Ghana donations made with the purpose of fighting the Covid-19 pandemic are a tax deductible expense subject to this simple documentation requirement.⁴

In this paper, we analyse behavioural responses to the hybrid reporting policy by exploiting the context of charitable contributions in the Republic of Cyprus, where the reporting setup is ideal for

¹For a recent review of the literature on tax enforcement see Slemrod (2019).

²An economy is characterised as semi-formal when economic activity and exchange is often largely unregulated due to a high prevalence of cash transactions. Other semi-formal economies include Greece, Turkey, Russia and several countries in Latin America such as Argentina and Brazil (Artavanis et al., 2016).

³In the context of less developed economies the evidence is more mixed (Kumler et al., 2020; Carrillo et al., 2017).

⁴In France and the US the policy is used as a tax enforcement mechanism for charitable contributions claims. With charitable giving levels in the US exceeding \$420 billion in 2018 (Giving USA, 2019), this tax deduction represents a significant tax expense for the US government each year.

our purposes.⁵ Charitable giving in Cyprus is subsidised through a tax deduction, and this deduction is subject to the simple documentation requirement described above. Importantly for our empirical methodology, the requirement to file third-party issued documentation only activates if taxpayers report aggregate charitable contributions above a pre-determined threshold. Below this threshold no documentation is required and taxpayers self-report. During the time period we observe, this threshold varies across taxpayers, and it moves at least once for all taxpayers, offering multiple sources of quasi-experimental variation to identify the effects of this policy.

Within this setup, we present four sets of results. First, we precisely identify the effect on claimed deductions, and hence the tax base, from the hybrid reporting policy. Exploiting salary-dependence of the thresholds that govern documentation requirements, we start by presenting compelling graphical evidence of discontinuities at the salary cutoffs. We then employ a regression discontinuity approach, and find that individuals increase reporting by 0.7 pounds when 1 pound more can be claimed without providing documentation from a third party.

Second, we use a unique reform that retroactively shifted the location of the reporting threshold to separate real charitable giving from pure reporting responses to the change in enforcement environment. Exploiting the time-profile of responses using bunching techniques, we find that at least 64 percent of the large responses to this enforcement policy are purely changes in reporting behaviour. This separation of reporting and real responses is crucial in a setup where we expect positive externalities from real behaviour, such as expenditures on charitable contributions, investments in education, professional training, retirement savings, etc. In such cases, the goal of the fiscal authority is not only limited to raising tax revenue, but also to encourage particular behaviours through tax incentives. Third, exploiting the panel-dimension of the data, we look into the underlying micro-level patterns of behaviour to assess whether reporting responses are due to compliance costs or tax evasion. We find that tax evasion is the main driver of behaviour, with compliance costs playing only a minor role in this context.

Lastly, we show that for a large part of the population the key driver of behaviour is the reporting environment rather than the size of the subsidy for charitable giving. To illustrate this, we analyse the elasticity of reported charitable contributions with respect to the price of giving, using quasi-experimental variation in tax prices generated by reforms to the income tax schedule. We show that taxpayers who react to the reporting policy seem largely unresponsive to changes in the price of giving, while the remaining part of the population displays a large and significant elasticity. This difference in elasticities is driven by sticky behaviour around reporting thresholds. Our results tentatively suggest that the reporting environment can affect the success of policies providing financial

⁵Henceforth, we simply use the term Cyprus to refer to the Republic of Cyprus.

incentives, because it separates the real decision from the reporting decision for a subset of people. While we analyse the hybrid policy in the context of deductions for charitable giving, our results are readily generalisable to all deduction types, or types of income where a third party is involved.

The setting for our analysis is Cyprus, a semi-formal economy with an informal sector estimated to account for about one-fourth of economic activity (Schneider and Enste, 2000; Hassan and Schneider, 2016), and with a high reliance on cash transactions (Esselink and Hernández, 2017). Tax enforcement is an important dimension of all tax systems but is of particular concern in economies with a high level of informality where compliance levels tend to be low (Artavanis et al., 2016; Besley and Persson, 2014). Furthermore, it is essential to take the level of informality into account when thinking about the design of tax policies. For instance, literature shows that the level of taxes can affect the decision of economic agents to transition between the formal and informal sector in countries with substantial informality (Rocha et al., 2018; Waseem, 2018). We show that the simple hybrid reporting policy we study can lead to a substantial increase in tax collection in the context of a semi-formal economy. Given its low implementation cost and the lack of reliance on sophisticated institutions and data infrastructure, our findings provide a blueprint for effective enforcement tools in lower income economies, characterised by high levels of informality and low institutional quality. Ensuring tax compliance in developing countries is a central concern, since the ability to collect taxes is of key importance for the process of development (Besley and Persson, 2014).

This paper makes several contributions to the existing literature. Most importantly it adds to a growing literature on the effectiveness of tax enforcement initiatives and in particular policies regarding information reporting. While a great deal of attention has been devoted to the effectiveness of third-party reporting, much less focus has been directed at alternative reporting policies. This is surprising given the prevalence of these policies across both developed and developing countries, and given the institutional sophistication and large investments in data infrastructure necessary for a system of direct third-party reporting. This paper rigorously analyses a simple and low-cost reporting policy. While the literature has presented initial evidence of a response to this type of policy, we contribute on various fronts by studying a fundamentally different semi-formal context as well as providing a detailed analysis of the effects on various margins of behaviour using a unique natural experiment. This experiment allows us to directly observe and separate a pure reporting component from the behavioural response to an enforcement reform, and thereby set bounds on real behaviour affecting the true tax liability.

Within this literature the most related paper is Fack and Landaïs (2016), who consider a reform introducing a similar hybrid enforcement policy in the context of charitable contributions in France.

While the main focus of the paper is the effect of enforcement strictness on key elasticities and hence a different question than we address here, it also shows initial evidence of a drop in claimed deductions around the introduction of this reporting policy. However, the design and nationwide scope of the reform in France means that general fluctuations in donations over time cannot be controlled for. A separation of the components of the effect is also not feasible in this setting.⁶ We contribute by considering this enforcement policy within a substantially different context characterised by a large informal sector and a heavy reliance on cash transactions. Our policy variation and data allow us to provide a precise estimate of the magnitude of the effect on the tax base, as well as a direct separation of pure reporting adjustments from any potential real responses affecting true tax liability. Further, we document micro-level patterns in behaviour underlying the response using a unique population-wide panel dataset from Cyprus. This allows us to analyse the motives behind responses and whether pure reporting effects are driven by the hassle cost of compliance or tax evasion. The policy implications of these separate motives are widely different.

Among other papers analysing tax reporting policies, a large portion deals with the impact of direct third-party reporting on compliance. Examples include Kleven et al. (2011), who conduct an audit experiment in Denmark and show that the evasion rate on third-party reported income is very low compared to self-reported income. Phillips (2014) finds a similar result analysing audit data from the US, while Alm et al. (2009) use a lab experiment to show that compliance rates increase with the proportion of income subject to direct third-party reporting. Gillitzer and Skov (2018) look at a Danish reform which concurrently introduced automatic pre-filled tax return information and third-party reporting on deductions for charitable contributions. Contrary to other studies (Fack and Landaïs, 2016; Ackerman and Auten, 2011), they find little evidence of over-reporting of contributions before the introduction of third-party reporting. This also contradicts the findings of Kleven et al. (2011) for other income and deduction types in Denmark.

A small set of papers analyse alternative reporting policies. These policies are typically very specific to a particular type of income or deduction. For instance, LaLumia and Sallee (2013) find that introducing a requirement to report children’s full social security numbers leads to a reduction in the number of dependants claimed on the tax return. Ackerman and Auten (2011) examine tax deductions for donated vehicles and similarly find that tightening the vehicles’ valuation requirements leads to a significant drop in reported valuations. Tazhitdinova (2018) also uses the context

⁶Fack and Landaïs (2016) cannot observe whether the response to the French reform is pure reporting or real giving, but they do show evidence that a large French charity did not experience a material change in received contributions around the timing of the reform. If the charity is representative, this finding is suggestive that at least some of the response is not reflected in real giving.

of charitable contributions, but analyses an enforcement policy imposing the requirement to provide fully self-reported details of non-cash donations to charities. She finds a significant effect on evasion from this policy.⁷

The paper also relates to a recent literature on tax enforcement and information reporting in less formal economies. Important contributions to this strand of literature include Pomeranz (2015) and Naritomi (2019) who study different aspects of third-party reporting in the VAT systems of Chile and Brazil respectively. Both find a positive effect on compliance and tax collection. Carrillo et al. (2017) analyse the limitations of third-party reporting in the context of corporate taxes in Ecuador while Kumler et al. (2020) study the accuracy of employer-provided wage reports in Mexico and find that especially small firms under-report. Jensen (2019) argues that the increase in information trails coming from the long-run move from self-employment to employee-jobs underpins tax collection and is key for development.⁸ We contribute to this literature by showing that a simple self-reported documentation requirement can substantially increase tax collection in a semi-formal setting.

Lastly, the paper contributes to a small literature on the importance of tax system design for behavioural responses to taxes and subsidies. We observe individual level behaviour in response to both the reporting environment and the subsidy level on charitable contributions. Using this unique feature of our context and data, we attempt to answer whether the strong responses to the reporting setup impede behavioural responses to the generosity of tax subsidies. The concept that features of the tax system matter for the size of behavioural responses has been addressed theoretically by Slemrod and Kopczuk (2002) and Slemrod (1994), while a few other papers provide supporting empirical evidence. Kopczuk (2005) for instance finds that the elasticity of reported income with respect to tax rates depends on the level of deductions in the tax system, while Fack and Landaïs (2016) show that both the elasticity of reported income with respect to the tax rate and the elasticity of reported charitable contributions with respect to price are sensitive to the strictness of enforcement.⁹ Slemrod (1989) uses tax returns subject to audits to investigate the misreporting and real components of the elasticity of giving.

The rest of the paper is structured as follows: Section 2 describes the institutional environment

⁷For a theoretical contribution deriving the allocation of resources between audits and information reporting in an optimal tax enforcement policy setting see Kuchumova (2017). For literature focusing instead on firms' responses to information reporting see for instance Almunia and Lopez-Rodriguez (2018), Slemrod et al. (2017), Agostini and Martínez A (2014).

⁸An important theoretical contribution here is Gordon and Li (2009).

⁹A related paper in this context is Doerrenberg et al. (2017), which documents responsiveness of total deductions to tax changes in Germany, as well as a significant difference between the elasticity of gross and taxable income.

and the data we use in the empirical analysis. Section 3 presents a brief conceptual framework, before section 4 goes through our results from the regression discontinuity analysis investigating behavioural responses to the reporting policy. In section 5 we use bunching techniques to separate real and reporting responses where-after we in section 6 analyse the components of the reporting response. Section 7 uses our framework and results to discuss policy implications for the optimal placement of the reporting threshold. Section 8 presents our results on the importance of reporting environment and thresholds for the tax price elasticity of charitable contributions, before section 9 concludes.

2 Institutional context and data

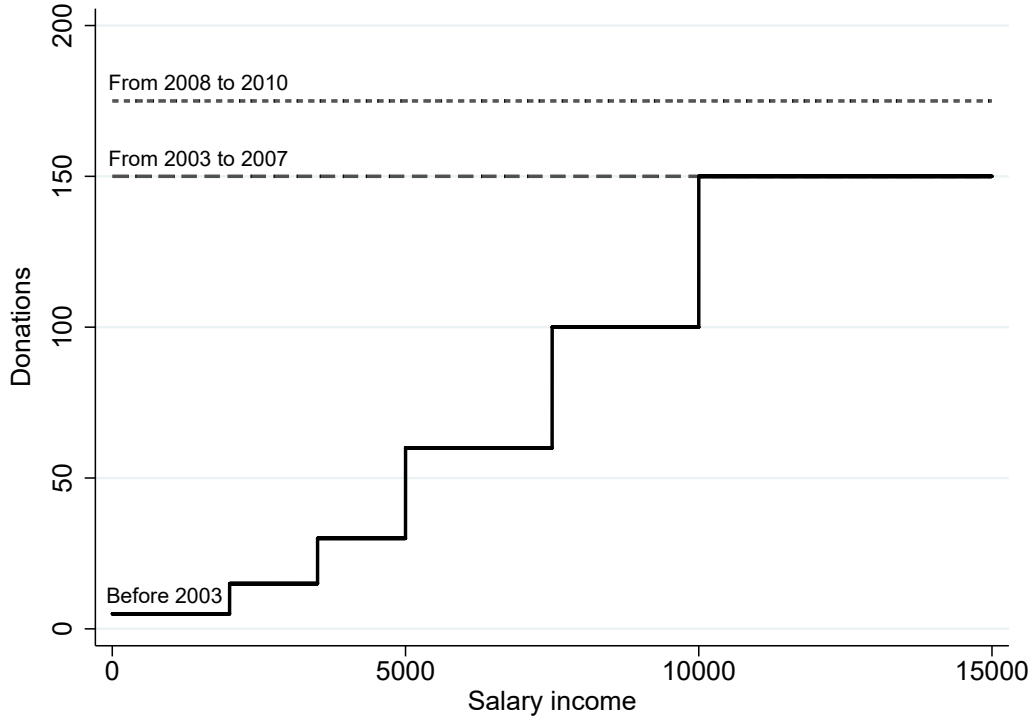
In this section, we describe the institutional context of charitable giving in Cyprus, the associated reforms we exploit, and the administrative dataset we use to analyse this enforcement policy.

2.1 Institutional details

As in most countries, charitable giving in Cyprus is subsidised through tax incentives. Specifically, the amount donated is deducted from taxable income, reducing the effective price of giving to $(1 - t)$, where t is the marginal tax rate. Due to administrative constraints, there is no automatic third-party reporting by charities to the tax authority. Instead, tax filers are required to provide receipts of donations. To reduce both hassle and administrative costs, a threshold for aggregate donations has been set, up to which no receipts are necessary. For any amounts claimed beyond this threshold, documentation must be provided.

We exploit several sources of exogenous variation in these reporting thresholds and in marginal income tax rates, which allow us to examine how contributions respond to the filing environment. We first explain the reforms associated with the filing environment. The reporting threshold schedule, determining at which aggregate donation level receipts are necessary, features several discontinuities and reforms between 1999-2010. This is illustrated in Figure 1. Prior to 2003 the maximum amount one could declare without providing receipts was a function of salary income. For salary earnings above CYP 10,000, this threshold was CYP 150, for earnings between CYP 7,500-10,000 it was CYP 100, etc. These salary-dependent cutoffs, introduced in 1989, were abolished by the Regulatory Administrative Act No. 823 of 2003. No new law or regulatory act set any new thresholds; rather, the tax authority created a de-facto threshold at CYP 150 for everyone by clearly stating the following on the 2003 tax form: “*For donations above £150 please attach receipts*” (shown in Appendix Figure 10b). This is the first tax form that denotes a specific threshold; up to 2002 the tax form simply stated: “*attach relevant receipts*” (Appendix Figure 10a). The 2003 wording was kept the same up to 2007.

Figure 1:
Reporting thresholds over the sample period



Notes: The figure illustrates the thresholds up to which people could claim deductions for charitable contributions without providing receipts. In the years 1999-2002 this threshold was dependent on salary income with 5 different notches in the schedule. From 2003 the threshold became independent of income and was set at CYP 150. In 2008 this threshold was changed again to CYP 175.

This threshold was changed again when Cyprus switched currency and adopted the Euro. The Euro was phased in during 2008, and the tax return for the 2008 fiscal year (which coincides with the calendar year) had to be filed in Euros. The tax return now stated “*attach receipts only for donations above €300*” (Appendix Figure 10c). Given the locked exchange rate¹⁰ of CYP0.585274 = €1, this was equivalent to CYP 175. Tax returns are published after the end of the fiscal year and need to be submitted by the end of April. Therefore, this new threshold was published *after* the end of the 2008 fiscal year, precluding any real responses in contributions during 2008.

Besides salary specific discontinuities and reforms, we are also able to exploit exogenous variation in the tax price of giving generated by marginal tax rate reforms. The Appendix Figure 11 shows the

¹⁰This became legally binding by the Regulatory Administrative Act No. 311/2007.

income tax schedule in Cyprus between 1999-2010, where marginal tax rates were changed six times in total, affecting all parts of the income distribution.

Our empirical strategy draws on three sources of variation generated by the institutional setting. We start by focusing on the pre-2003 salary-based discontinuities in the filing threshold to establish that donations respond strongly to this reporting policy. We then exploit the unique timing of the 2008 reform of the reporting threshold to set bounds on the real and pure reporting components of the response. We combine this with a deeper analysis of the micro-patterns of reporting behaviour present in our panel data to characterise the underlying driver of the pure reporting responses. Lastly we draw on the variation in marginal tax rates to estimate the tax price elasticity of reported donations and examine whether this is sensitive to the design of the filing environment i.e. the presence of reporting thresholds.

2.2 Data

The data come from first-time access to the administrative records of the Tax Department of the Republic of Cyprus. It covers the universe of tax filers between 1999-2010, and includes information from the main fields of the I.R. 1A tax return, as well as basic demographic and firm-related characteristics. All employees are required to file taxes, unless they earn a gross amount below some tax free level. The self-employed are required to file regardless of the amount earned. Besides individuals with earnings from labour, the dataset also includes pensioners and individuals out of the labour force who may be filing because it is a requirement for accessing government welfare programmes. To create our working dataset, we impose the following restrictions. First, we consider only individuals with a single employer, who report at least some positive salary income and are aged between 25-54.¹¹ Second, we drop individuals in the top 0.1% of donations. Our working dataset contains about 1.5 million observations and 225,000 unique individuals. Appendix Table 3 shows summary statistics for our sample.

It is important to note that due to the way the tax administration provides the tax data, we only observe a composite variable measuring both deductions for donations, trade union subscription fees, and a so-called "professional" tax. These tax deductions all appear on the same section of the tax return. The "professional" tax was in place in years prior to 2003, and was a lump-sum tax calcu-

¹¹As is explained below, we need to know workers' sector and salary to determine their potential union membership fees. In our data, we can observe individual salaries, but not salaries per employer. We therefore drop the 3% of our sample having more than one job, to ensure we can do this accurately. Further, the standard retirement age during our sample period is 60 years of age. Some sectors have the option of early retirement up to 5 years before the standard retirement age and some jobs have an option to work less hours in the years leading up to retirement. To make sure we are looking at full-time employees, and our results are not biased by such labour-market considerations we only include workers up to the age of 54.

lated as a step-function of earnings (Figure 12 in the Appendix shows the exact schedule). Since we observe individual earnings we can calculate the amount of the tax and subtract this from the variable where appropriate.

In the case of union subscription fees we cannot directly observe the size of this fee or whether an individual is a member of a union. However, using detailed sectoral information available in our dataset combined with information on union fee rates, we can back it out with some noise. Union fees are a sector specific fixed proportion of salaries, deducted every month from employers through the PAYE system. We collect information on union fee rates for each sector directly from the trade unions and from the Ministry of Labour and Social Insurance. Combining union rates, salaries and sectors we can residualise our donation measure from union fees in fully (or highly) unionised sectors. To our advantage, union membership in Cyprus, which has about 50% coverage in our study period (Ioannou and Sonan, 2014) is highly concentrated in just a few sectors. We elaborate further on this in Section 8.

Note that union fee rates are fixed across salaries and did not change following any tax or threshold reforms. Therefore this data issue does not affect the empirical strategies used in the first part of our analysis. These methods exploit variation in the data in such a way that adding noise to the measure should not affect any results. In the case of the regression discontinuity design the noise will shift the level on both sides of the cutoff, but not affect the size of the discontinuity. Similarly for the bunching estimates, fees should affect the entire distribution equally leaving the estimate of excess mass unchanged. However, for the last part of the analysis, where we consider elasticities, these union fees play a role and we need to correct for them as explained above. For each of our empirical strategies, we run a battery of robustness checks to show that our results are not affected by the way we deal with union fees.

3 Conceptual framework

To clarify the incentives created by the reporting environment in the Cypriot context, we present a stylised theoretical framework based on the model developed in Tazhitdinova (2018). While the framework is simple, it encompasses the main features of our empirical setup. Assume individual i solves the following problem:

$$\max_{c_i, d_i, r_i} u_i(c_i, d_i) \quad s.t. \quad c_i = y_i - d_i + t \cdot r_i - e_i(r_i) \cdot \mathbb{1}_{(r_i > d_i)} - h_i \mathbb{1}_{(r_i > \alpha)}$$

choosing consumption c_i , real charitable donations d_i , and reported donations r_i . Utility is increasing in both consumption and donations, with decreasing marginal utility in both. The individual takes as fixed his income y_i , his tax rate t , and a threshold α , above which documentation for real giving

is required to get a deduction.

As in our empirical setting the taxpayer gets a deduction from taxable income equal to reported donations, and hence the subsidy rate is equal to the tax rate. If individual i over-reports donations ($r_i > d_i$) he faces a fixed (individual specific) evasion cost below the threshold, $e_i(r_i) = e_i \forall r_i \leq \alpha$. For simplicity we assume that the evasion cost is infinitely high above the threshold, since receipts are inspected and hence evasion detected with certainty $e_i(r_i) = \infty \forall r_i > \alpha$.¹² For taxpayers with real donations above the threshold, the process of collecting and attaching receipts might represent a compliance cost. The compliance cost of collecting documentation for real donations is fixed and individual specific, h_i . This captures the notion that an individual with a preference for giving to multiple charities has a larger compliance cost than a person giving all donations to a single charity. Note that only people with a weak inherent preference for charitable giving have the opportunity to evade, since evasion is only possible if real giving is below the threshold α . People below the threshold either report truthfully $r_i = d_i$ if evasion cost is larger than evasion benefit ($e_i > t(\alpha - d_i)$) or they evade as much as possible and report $r_i = \alpha$ if evasion benefit is larger than evasion cost ($t(\alpha - d_i) > e_i$).

People with a strong inherent preference for giving cannot evade, but will choose between reporting truthfully $r_i = d_i > \alpha$ and hence incurring the compliance cost, or reporting $r_i = \alpha$ to avoid the hassle of collecting receipts. Those that choose to report $r_i = \alpha$ might also decrease their true giving since they do not get the subsidy for their marginal giving when they do not incur the compliance cost.¹³

4 Behavioural responses to a hybrid reporting policy

We begin our empirical analysis by presenting motivating evidence showing that reported contributions seem to respond strongly to changes in the reporting environment. As explained above, our period of analysis includes two reforms, implemented in 2003 and 2008, both of which increased the

¹²This is a simplifying assumption and is not needed for our general conclusions. Here we focus on a fixed cost of evasion, guided by our empirical findings, however one could easily consider a variable component of the cost. A cost which depends on the amount evaded is typically justified by considerations such as increasing probability of detection. Such considerations do not, however, seem relevant in this setting.

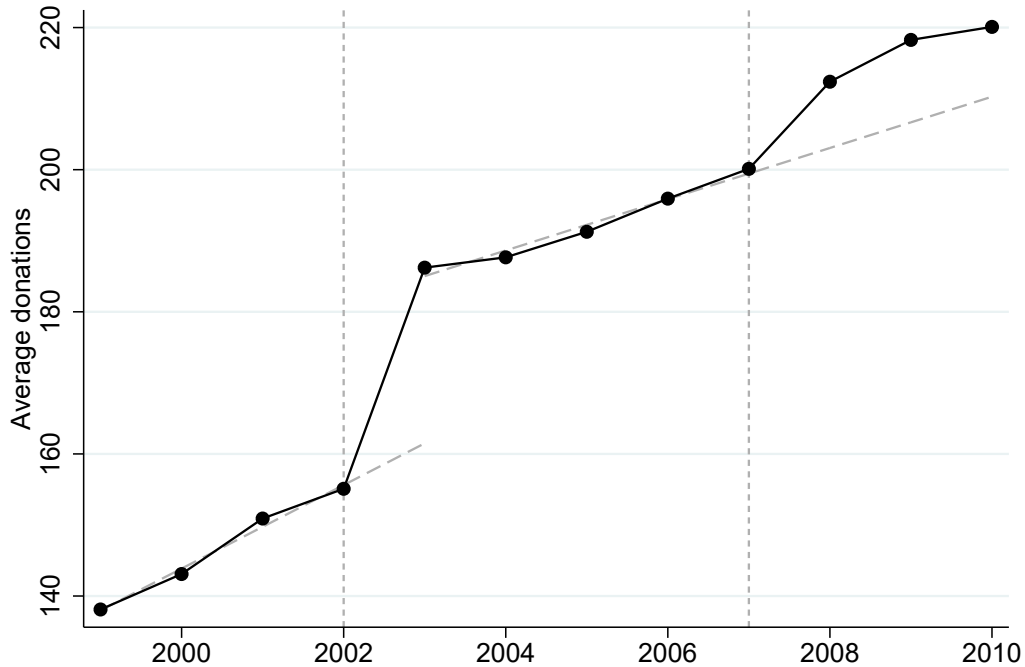
¹³A possible extension to our model is to introduce loss aversion and treat the thresholds as reference points. We do not however adopt this because it is inconsistent with the empirical patterns we observe in our data. Such a model predicts not only bunching but also an asymmetry in the distribution around the threshold, due to a drop in the density of donations above it. Our extensive bunching analysis, which we elaborate on in Section 5, reveals no such patterns. If anything, we observe a tendency for the density to tend slightly upwards in the region just above our thresholds. For a detailed explanation of the predictions for the shape of the empirical densities in a reference-point model, see Seibold (2020).

levels of the reporting threshold. In Figure 2, we check whether they relate to reported donations by plotting average reported donations over time, and marking the years before the reforms with vertical dashed lines. Apart from an increasing time trend, the raw timeseries clearly reveals two sharp jumps exactly in the two reform years. This initial time-profile of reported donations suggests that these reforms, which uniformly relaxed the enforcement environment, caused a substantial increase in reported donations in Cyprus.

Given this evidence, the following section aims to identify the causal effect of this hybrid enforce-

Figure 2:

Average reported donations among donors



Notes: The figure shows the yearly average of reported donations using only observations where donations are positive. We remove the top 0.01% of donations within each year. The sample includes all tax filers in the age range 25-54, with some positive salary income and only one job within a given year.

ment policy on reported donations. We do this by exploiting our first source of quasi-experimental variation: the salary-based discontinuities in the amount of aggregate donations tax filers can report without receipts before 2003.

4.1 Regression discontinuity estimates

As shown in Figure 1, the reporting threshold was a function of gross salaries before 2003. This setup lends itself to a regression discontinuity design. For our main RD estimates, we use years 1999-2001 and restrict our sample to those with only salary income. We exclude 2002 because a reform in that year shifted the first income tax threshold to CYP 9,000, meaning that individuals with salaries below this level cease to be a reliable sample as they had no obligation to file a tax return and no tax incentive to claim deductions. We also exclude individuals with non-salary income because the threshold we want to exploit is a function of salary income only, and we want to preserve the income trend in donations.¹⁴ We focus on two discontinuities: the jump from CYP 100 to CYP 150 at the CYP 10,000 salary cutoff, and the jump from CYP 60 to CYP 100 at the CYP 7,500 salary cutoff. We do not consider lower cutoffs because they are located at income levels where individuals have no tax filing obligation in part or all of the period considered.

Figure 3 plots the average reported donation by salary bins of width 50 between 1999-2001. As is clearly seen, donations jump at exactly the salary cutoffs associated with different reporting thresholds, but otherwise evolve smoothly. Note that our measure here also includes professional taxes and union fees. We do not remove these, since neither involve any discontinuities at our salary cutoffs of interest. This is evident in Appendix Figure 12, which shows that the professional tax indeed evolves smoothly across these cutoffs. Likewise, union fees are always set at a fixed percentage of salary, and hence do not jump at different income levels.

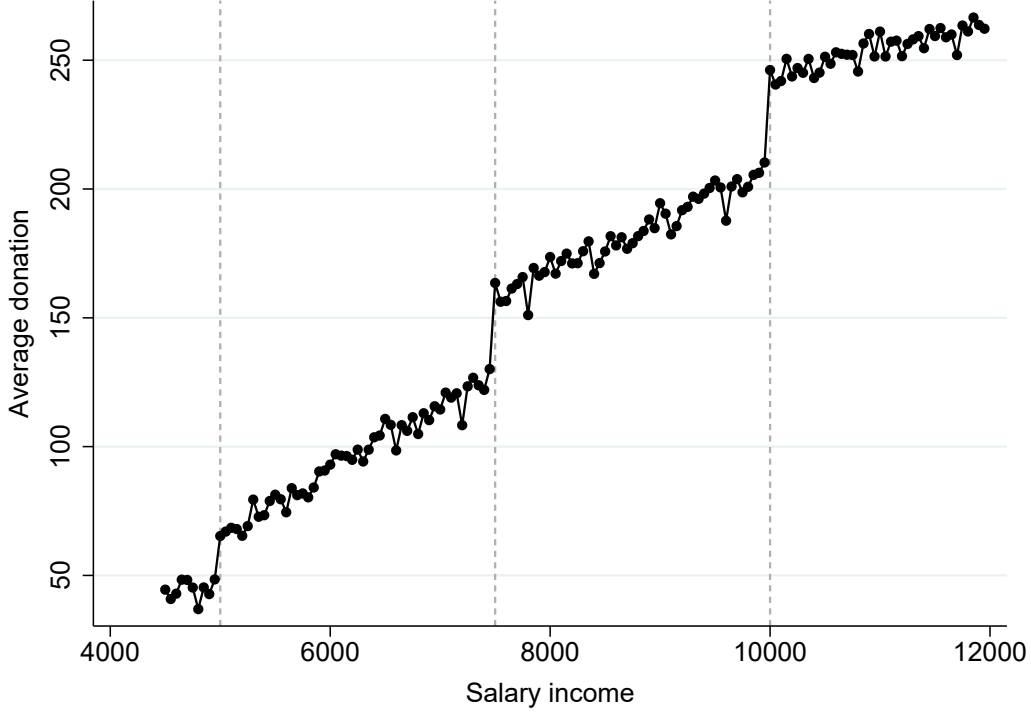
Our aim is to estimate the size of the discontinuous jumps in reported donations using a regression discontinuity design with individual salaries s_i as the assignment variable. The identifying assumption is absence of precise manipulation of the assignment variable, i.e. workers cannot precisely choose their salaries in order to manipulate the different thresholds. If workers just to the right of a cutoff strategically placed themselves there in order to be able to report more, then workers with salaries just below the cutoff would not provide a valid counterfactual. The possibility that workers specifically search for wage-hours packages in order to respond to the cutoffs associated with charitable giving is however very unlikely.¹⁵ Further, we do not observe any signs of manipulation in the density of salaries or any discontinuities in covariates around the cutoffs (see Appendix A). Salaries

¹⁴For robustness, we also run our main specification including 2002, and including individuals with non-salary income, and find very similar results.

¹⁵There are significant labour market frictions associated with searching for wage-hours packages. Indeed, a public finance literature on taxable income bunching (Chetty et al., 2011; Kleven and Waseem, 2013; Gelber et al., 2017) and work hours constraints (Dickens and Lundberg, 1993; Blundell et al., 2008) shows that there are significant frictions associated with precisely choosing earnings. Mavrokonstantis and Seibold (2020) also find frictions in the case of Cyprus.

Figure 3:

Average reported donations by income 1999-2001



Notes: The figure shows average reported donations by income pooled in the years 1999-2001. We use income bins of 50 including the left-hand value. We include people with only salary income and only one job within a single year and remove people with an income at an exact round number (multiples of 500).

in general have a high propensity to be set at round numbers, and hence we see some bunching in the density of salary income around these values. We proceed by dropping the rounders from our estimation sample, but show as a robustness check that their inclusion does not change our results. To estimate the size of the discontinuities, we treat individual salary s_i as our assignment variable, and run regressions of the form:

$$r_i = \alpha_0 + \alpha_1 Treated_i + f(s_i, \beta) + Treated_i \times f(s_i, \gamma) + X_i' \delta + \epsilon_i \quad (1)$$

where we define, for each threshold $T \in \{7500, 10000\}$ separately, $Treated_i = \mathbb{1}\{s_i \geq T\}$. Here r_i measures reported donations of individual i while $f(s_i, \cdot)$ is a polynomial function with parameter vectors β and γ that controls for the salary trend and for the interaction between the salary trend and treatment status respectively. The parameter α_1 measures the jump in donations due to the change

in the reporting threshold. Some specifications also include a vector of controls X (sex, year and sector fixed effects). Lastly, we only consider bandwidths of up to 2000 to ensure that no estimation sample includes more than one cutoff.

Table 1 shows our results split into two panels: A for the CYP 10,000 cutoff and B for the CYP

Table 1:

Regression discontinuity estimates of reported donations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A:								
Above 10k	36.64*** (0.95)	33.80*** (1.41)	34.40*** (0.97)	32.51*** (1.45)	35.08*** (1.33)	35.12*** (2.04)	33.79*** (1.38)	34.15*** (2.12)
Scaled response	0.73	0.68	0.69	0.65	0.70	0.70	0.69	0.69
Observations	84 255	84 255	68 697	68 697	43 095	43 095	35 404	35 404
R^2	0.21	0.21	0.28	0.28	0.13	0.13	0.20	0.20
Panel B:								
Above 7.5k	29.64*** (0.74)	31.77*** (1.08)	30.48*** (0.79)	30.91*** (1.14)	30.78*** (1.01)	26.50*** (1.51)	30.21*** (1.06)	26.30*** (1.58)
Scaled response	0.74	0.79	0.76	0.77	0.77	0.66	0.76	0.66
Observations	92 433	92 433	71 095	71 095	51 850	51 850	40 123	40 123
R^2	0.29	0.29	0.37	0.37	0.17	0.18	0.27	0.27
Polynomial	p_1	p_2	p_1	p_2	p_1	p_2	p_1	p_2
Controls	-	-	✓	✓	-	-	✓	✓
Bandwidth	2 000	2 000	2 000	2 000	1 000	1 000	1 000	1 000

Notes: This table shows the results from estimating specification (1) on our main sample pooled over 1999-2001. The *scaled response* is calculated as the parameter estimate divided by the notch size in the donation schedule (i.e. in panel A the parameter estimate is divided by 50 while in panel B the parameter estimate is divided by 40). p_s indicates that we fit a polynomial of order s on each side of the notch, while controls include sex, year and sector fixed effects. Robust standard errors clustered at the individual level in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

7,500 cutoff. To assure robustness of our estimates, we present results from different specifications of equation (1): with a first and second order polynomial of the assignment variable,¹⁶ with and

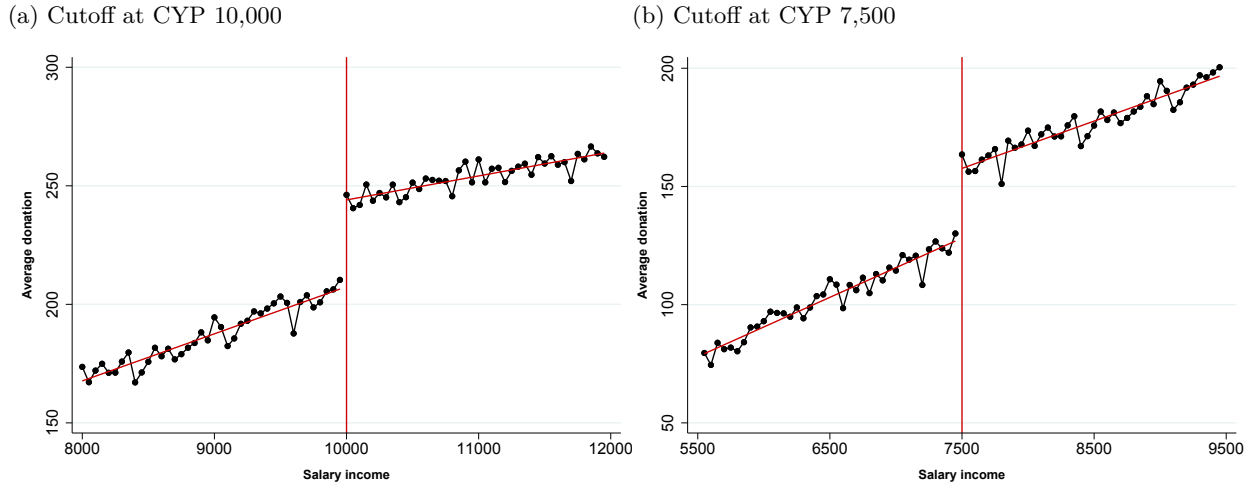
¹⁶As is clear from Figure 3, a linear specification allowing for different slopes on each side of each threshold should suffice. For robustness we nevertheless also report results using second-order polynomials. We have also estimated specifications with higher order polynomials, and found that this does not change any results (not reported).

without controls, and with different bandwidths. Each panel reports two estimates: (a) the size of the discontinuity, and (b) the scaled response, which divides our estimate by the size of the notch in the donation schedule.

Starting with the 10,000 cutoff, column (1) of Panel A shows that the effect of increasing the threshold from CYP 100 to CYP 150 is a CYP 36.64 increase in reported donations, with this effect estimated with very high precision. This estimate implies a scaled response of 73%, i.e. that workers increase their reported donations by 0.73 for every unit increase in the amount of donations that can be reported without providing receipts. The result is highly robust to the choice of polynomial order, inclusion of controls, and bandwidth. As columns (2) - (8) show, the estimated effect is on average CYP 34.5 and the scaled response is hence about 70%. We illustrate the RD estimation in Figure 4 panel (a), showing the discontinuity in average reported donations at CYP 10,000 and the linear regression lines estimated on either side of the cutoff.

Very similar results are found when we consider the second cutoff at CYP 7,500 (Table 1 panel Figure 4:

Regression discontinuity estimates of reported donations



Notes: The figure shows average reported donations by salary bin, pooled for the years 1999-2001, with a linear fitted line on either side of the specific salary cutoff. In each sub-figure, we use a bandwidth of 2000 and income bins of 50 including the left-hand value. We include people with only salary income and only one job within a single year and remove people with an income at an exact round number (multiples of 500).

B, Figure 4 panel (b)). The increase in reported donations is estimated close to CYP 30 across all specifications. This is as expected lower than the effect estimated at the higher cutoff given that the discontinuity in the reporting threshold is also smaller in magnitude. When we scale the effect by the size of the notch, we find a very similar magnitude of the response, estimated at 74%, which suggests

that the behavioural responses are highly comparable across the two cutoffs. Again, estimates are extremely robust to the alternative specifications across (1)-(8). The discontinuity is illustrated in Figure 4 panel (b) which shows the clear jump in reported donations around CYP 7,500.

To test the robustness of our estimation sample we first include individuals with round-number salaries. Second, we include the year 2002 and third we include individuals who also have non-salary income (conditional on having some salary income). Fourth, we use our variable cleaned from professional taxes and union fees.¹⁷ Lastly, we run separate regressions for individuals working in highly and not highly unionised sectors. The results are shown in Appendix Tables 4-10. Our estimates are extremely robust to every variation we consider.¹⁸

5 Separating real and reporting responses

Having established that reported donations respond strongly to this enforcement policy, we now focus on characterising the composition of this response by exploiting the timing of the 2008 reform. As explained, the threshold up to which no receipts are necessary was moved from CYP 150 to CYP 175 (300 Euros) in 2008, but this change was only announced after the end of the 2008 fiscal year. Hence, at the time of the announcement it was no longer possible to adjust real giving behaviour and therefore any response to the new threshold in 2008 can only be a pure reporting adjustment. To exploit this unique characteristic of the reform we examine bunching patterns around the reporting thresholds and how these patterns change in the year of the reform. To do this, we implement standard bunching techniques (Saez, 2010; Chetty et al., 2011; Kleven, 2016) and estimate the excess mass of individuals located at each threshold between 2003-2010. Our bunching results then allow us to estimate what proportion of observed responses are pure reporting effects by comparing the bunching (excess mass) in 2008 at the CYP 175 new threshold (B_{175}^{2008}), which can only be driven by a pure reporting response, to the bunching in 2007 at the CYP 150 threshold (B_{150}^{2007}):

$$L_R = \frac{B_{175}^{2008}}{B_{150}^{2007}} \quad (2)$$

In words, L_R reports the fraction of the excess mass at the previous threshold that moves to the new threshold before real responses are feasible. L_R provides a lower bound on the pure reporting response, since responses in subsequent years can include both a real and a reporting dimension. Note that we here use non-normalised bunching estimates (denoted by uppercase B), rather than

¹⁷In this case, we drop the public sector from Table 7. Due to the existence of two different rates for the union membership fee in the public sector we cannot correct the donation measure with high precision.

¹⁸Besides testing robustness, we also tested for heterogeneity in responses by sex and age but did not find notable differences.

normalised estimates (denoted by lowercase b), to ensure that differences in the counterfactual distribution around the two thresholds do not affect the size of L_R .

5.1 Bunching responses to reporting thresholds

We start by showing the dynamic bunching patterns in the data using our main sample, defined in the same way as in the RD section. We do not remove union fees because we want to preserve the raw patterns in the data. This of course means that we are identifying our effects from the non-unionised sample. We show in the next section that our results are highly robust to accounting for union fees. We group reported donations in bins of width 5 and fit an 11th order polynomial to estimate the counterfactual mass of filers in the absence of these thresholds. The difference between the actual and counterfactual count is therefore the excess mass ascribed to the discontinuous change in reporting requirements at the threshold. In our estimation, we also control for round number bunching in multiples of 50 and 100 (thereby flexibly allowing for different levels of roundedness for each).¹⁹ As highlighted in the theoretical framework presented in section 3 it is a priori unclear whether the bunching is coming from above or below. We therefore do not impose the integration constraint, i.e. an upward shift of the counterfactual density in bins to the right of the threshold. We discuss our methodology in more detail in Appendix B, but results are not sensitive to this choice.²⁰ In a later section we also analyse in detail whether bunching originates from above or below.

Figure 5 shows the empirical density of reported donations between 2003-2010 for our main sample, in bins of CYP 5. To get a sense of the magnitude of the bunching in each year, each sub-figure reports the normalised excess mass at each threshold, b_{150} and b_{175} , with bootstrapped standard errors shown in parentheses.²¹ To highlight how the bunching moves across the two thresholds, each sub-figure also demarcates the threshold in place in a given year by a solid vertical line, and the other threshold by a dashed vertical line. We do not include the estimated counterfactuals here to avoid cluttering, but show these separately for each threshold in Appendix Figures 13 and 14.

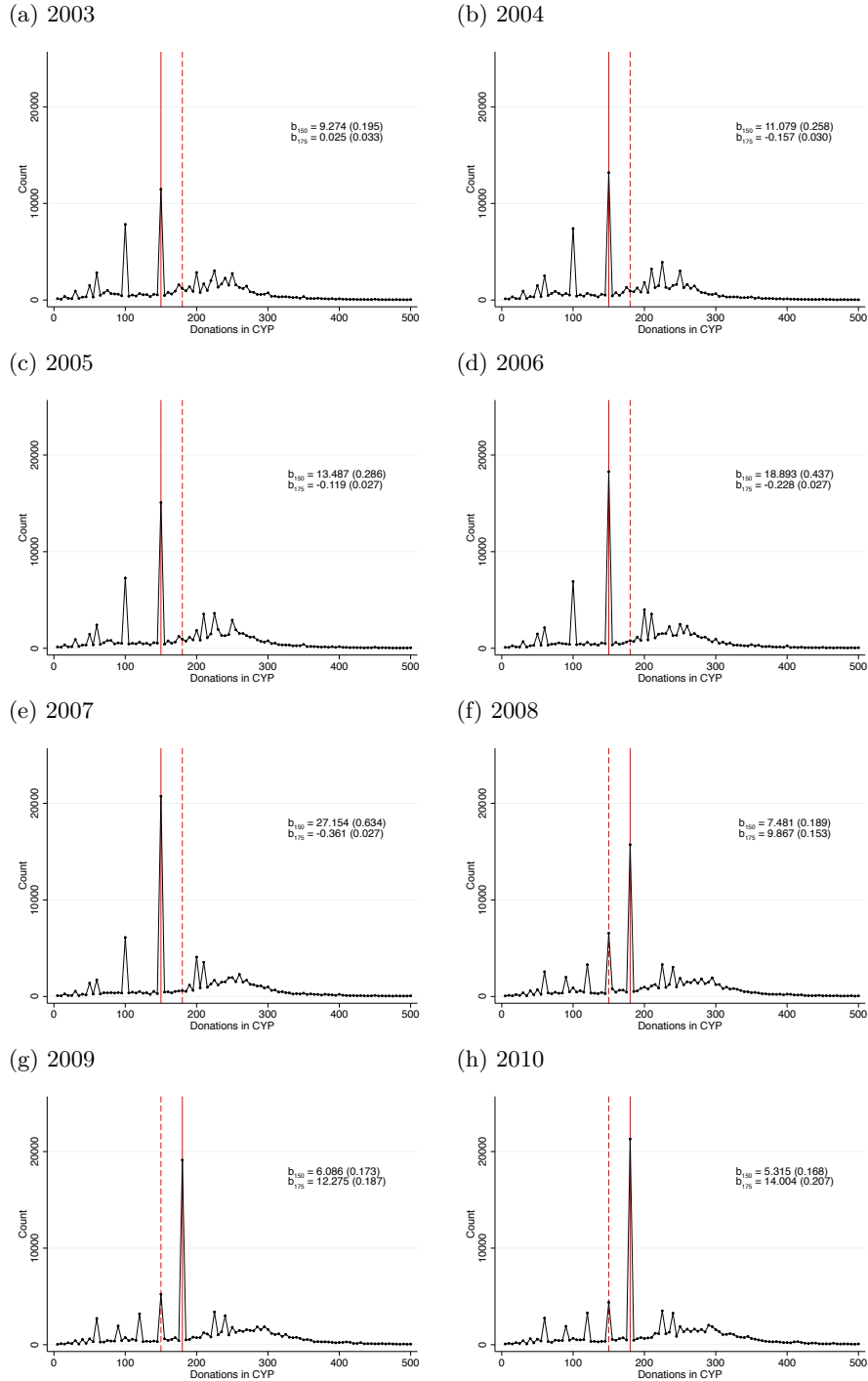
¹⁹For years 2008-2010, we have converted the currency from Euros to CYP using the official exchange rate. In this case, we control for round numbers by using the CYP converted amounts of the round numbers in Euros, since that was the actual currency used to file the tax return.

²⁰Estimates in this section were obtained using the R package **bunching** (Mavrokonstantis, 2019), available at <https://CRAN.R-project.org/package=bunching>.

²¹Note that we here estimate the normalised excess mass b by scaling the excess mass by the height of the counterfactual. We denote normalised bunching estimates by lowercase b (and non-normalised estimates by uppercase B).

Figure 5:

Bunching around reporting thresholds



Notes: This figure shows the bunching dynamics of reported donations among salary earners between 2003-2010, by plotting the yearly empirical distributions in bins of width CYP 5. Each sub-figure reports the normalised excess bunching mass b around the CYP 150 and CYP 175 thresholds, with bootstrapped standard errors in parentheses. Vertical solid lines mark the relevant threshold that is in place in a given year (CYP 150 during 2003-2007 and CYP 175 during 2008-2010), while dashed lines mark the other threshold that has either been eliminated, or has not been yet introduced.

We find that bunching at the CYP 150 threshold is very large in magnitude and sharp (i.e. there is no diffuse bunching around the threshold). The normalised excess mass steadily increases between 2003 and 2007, starting from a level of 9.3 and peaking at 27.2 in the last year this threshold is effective. By 2007 therefore, there are 27 times as many individuals at CYP 150 compared to what there would be absent the filing discontinuity. At the same time, there is no excess mass at CYP 175 throughout this period (marked with a dashed vertical line). These patterns are followed by a dramatic change in 2008. The bunching at CYP 150 stops growing and instead exhibits a large drop to 7.5, and continues decreasing thereafter. Bunching at CYP 175 now appears, producing a normalised excess mass of 9.9 in 2008. In a symmetrically opposite way to the bunching at CYP 150, bunching at the new threshold exhibits further growth in years 2009-2010. What is striking is that while there is no bunching in any year before 2007 at CYP 175, a very large spike appears in 2008, even though there was no knowledge of this new threshold, and hence no real response possible during the 2008 fiscal year. The bunching dynamics also suggests learning, as it seems to take time for individuals to understand the incentives created by the thresholds and respond to them over the years.

5.2 Estimating a lower bound for pure reporting responses

We next turn to our estimate of L_R . To obtain this, we restrict our sample to a balanced panel of tax-filers present in our data in all years between 2003-2010.²² This is important when we directly compare patterns across time as we may otherwise bias our estimates due to entry and exit from the sample. To generate our estimate of L_R we use the non-normalised excess mass (B_{150}^{2007} and B_{175}^{2008}) such that if the entire excess mass from the CYP 150 threshold moves to the new threshold in 2008 we get $L_R = 1$.

To visualise the dynamics, Figure 6 plots the non-normalised bunching estimates at each threshold across time, with the shaded areas demarcating our 95% confidence intervals. The patterns show how tax-filers in our panel shift around the two thresholds across time, and elucidate how striking the reversal in the bunching mass between the two thresholds is around the time of the reform.

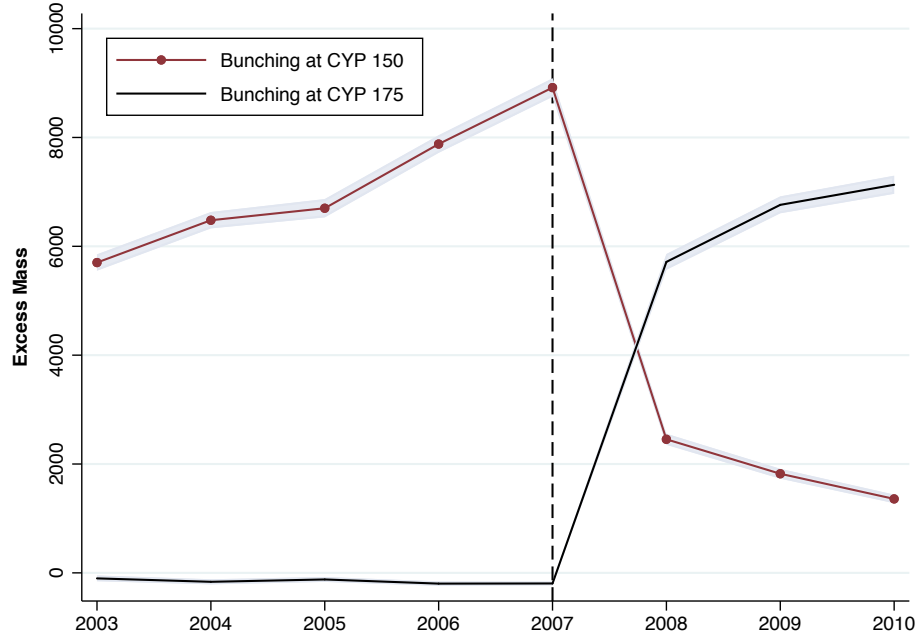
Using our estimates of B_{150}^{2007} and B_{175}^{2008} we get $L_R = 0.64$. This implies that at least 64% of the response to an increase in the reporting threshold is due to pure changes in reporting, rather than real changes in contributions.²³ In relation to our results from the previous section, this means that taxpayers increase reported donations by £0.7 when they can claim £1 more without documentation, and at least £0.45 of this increase is a pure reporting adjustment unrelated to real giving. To the

²²The bunching patterns presented in the previous section remain identical when we impose this sample restriction.

²³If we impose a less strict sample restriction of requiring tax-filers only to be present in 2007 and 2008, this estimate increases to 67%.

Figure 6:

Bunching (excess mass) estimates over time



Notes: This figure shows for the balanced panel of our main sample of salary earners, the estimates of the excess mass around both the CYP 150 and 175 thresholds, between 2003-2010. The shaded areas demarcate 95% confidence intervals.

extent that individuals take time to learn about, and understand, the changes in the reporting environment, the responses in 2009 and 2010 may also capture reporting responses and hence L_R is a lower bound on the pure reporting component.

Next, we further exploit the panel dimension of our data to check whether the patterns we observe are indeed driven by individuals moving from the old threshold to the new one. In Appendix Figure 15, we plot the empirical density of reported donations between 2003-2010, for the sample of taxpayers bunching at CYP 150 in 2007. We find that the majority of the 2007 bunchers are "repeat" bunchers, locating at CYP 150 in earlier years as well. They also overwhelmingly shift to the new threshold in 2008, while some take a further one or two years to complete the shift. Appendix Figure 16 repeats the analysis for those bunching at the new threshold in 2008. Again, we find the same patterns; these are individuals who were previously bunching at the old threshold for several years, with the shift being nearly complete by 2010.

Figure 18 in the Appendix plots the fraction of individuals with reported donations at (1) the 150 threshold, (2) the 175 threshold, (3) above 150 and (4) above 175 throughout the entire sample period. The trends in the fraction of individuals filing 150 and 175 is in line with our previous results. The

fraction of individuals at 150 is increasing from 2003 and peaks at 2007, before dropping sharply in 2008. Conversely, there are very few at 175 until 2008, when it increases sharply. What is more interesting is the trend in the proportion of people filing more than 150 and more than 175. While they exhibit parallel trends up to 2008, there is a sharp increase in the fraction of individuals filing above 150, but no change in the fraction of those filing above 175. This confirms that the movement is purely between these two thresholds, and emphasises how important the reporting environment is for taxpayer behaviour.

5.3 Robustness

We next discuss the robustness of these results and conduct a battery of checks on our main findings. Figure 6 clearly shows an increasing trend in the level of bunching over the years from 2003 to 2007. Each year more people bunch at the threshold, absent any changes in the reporting environment. This increasing trend is likely evidence of learning, since other optimisation frictions are less likely in a setup where the taxpayer himself notes the amount of donations on the tax form. By measuring the fraction of excess mass in 2007 that moves in 2008, the fraction L_R does not take yearly changes in bunching into account. This approach to measuring the pure reporting component embeds the underlying assumption that overall bunching remains constant from 2007 to 2008. To make sure our conclusions are not affected by yearly changes in the overall level of bunching we produce two alternative estimates. The first relates the bunching in 2008 at CYP 175, to the aggregate bunching at both thresholds in 2008. This approach assumes that the sum of the excess mass at CYP 175 and CYP 150 in 2008, represents what the excess mass at CYP 150 would have been in 2008 if the threshold had not been moved. Our second alternative estimate uses a linear fitted trend in the overall level of bunching at CYP 150 between 2003 and 2007, to predict what the level of bunching at CYP 150 would have been in 2008 absent the reform. We then relate the level of bunching at CYP 175 in 2008, to the predicted bunching at CYP 150 in 2008. Appendix Figure 28 illustrates the linear prediction. Hence,

$$L_R^1 = \frac{B_{175}^{2008}}{B_{150}^{2008} + B_{175}^{2008}} \quad \text{and} \quad L_R^2 = \frac{B_{175}^{2008}}{B_{150}^{\text{Predicted}, 2008}}$$

Using the balanced panel of our main sample we get $L_R^1 = 0.70$ and $L_R^2 = 0.60$. Both alternative estimates are very close to our main estimate of 0.64.

A second concern relates to potential anticipation of the 2008 reform. It is very important for the validity of the lower bound that the 2008 threshold change was not anticipated or somehow made public before the end of the fiscal year. Before the introduction of the Euro there was a large

government campaign informing citizens that during the transition they should simply use the official locked exchange rate to convert prices, salaries etc. Following this, filers should have expected the threshold to remain unchanged at a converted value of €250, not €300. Further, tax returns are not published before the end of the fiscal year. Even if it was published early through unofficial channels, it is highly unlikely that filers would be so keen to obtain their tax return before the end of the fiscal year that they would search for it.²⁴

Next, we consider whether our bunching patterns could be affected by the existence of union fees. This would be implied by two, extremely unlikely, scenarios. The first is that we are picking up bunching at thresholds that is driven purely by fees which coincidentally place individuals at the thresholds. The second is that we are picking up the sum of union fees and donations, which again happen to consistently sum to these thresholds. Both are implausible, especially given that union fees are a fixed percentage of salary.²⁵ Nevertheless, we repeat our analysis first restricting the sample to workers in non-unionised sectors and second removing union fees for workers in highly unionised sectors. Estimates in both cases are very similar to our main results.²⁶

Lastly, we check whether our main results are sensitive to the choice of the polynomial order used to estimate the counterfactual density. This affects our estimate of the bunching mass and can thereby influence our estimate of the reporting response, L_R . Figure 29 plots estimates of L_R for every polynomial order in unit intervals from 3 to 12, and shows that our results are extremely stable around the estimated value of 64%.²⁷

6 Tax evasion or compliance cost?

The excess mass that we observe at the reporting threshold can be due to either over-reporting, under-reporting or a combination of both. Taxpayers whose true donations are below the threshold can over-report up to the threshold, and thereby evade taxes. Taxpayers with true donations above the threshold can under-report down to the threshold and thereby avoid the hassle of collecting receipts. Consequently, the movement between thresholds that we interpret as a pure reporting response, can be either an increase in over-reporting or a decrease in under-reporting. While both

²⁴The fiscal year ends four months in advance of the tax return submission deadline. Looking at the data on tax return submission dates, we see that the vast majority of tax filers procrastinate, submitting their return just before the deadline of April 30th of the following year (see Figure 17 in the Appendix). Most file in the last week of the deadline. This behaviour is clearly at odds with active tax return search and filing behaviour.

²⁵This is confirmed if we look at the salary growth rate of the 2007 bunchers around the reform, see Appendix Figure 19.

²⁶The full set of results are shown in Appendix Figures 20 - 27. In the case of highly unionised sectors we exclude the public sector since we cannot correct the donation measure for union fees in this sector.

²⁷We have also checked for heterogeneity across sex and age groups and did not find any substantial differences.

movements represent pure reporting behaviour, the policy implications are very different and hence it is important that we can separate the two sources. In this section we exploit the richness of the data to disentangle the underlying source of the changes in reporting.

A simple way of evaluating the relative importance of over- versus under-reporting, is to look at the distributional patterns of donations around the threshold. A potential hassle cost would create a downward notch in the filer's budget set, and thereby a dominated region. Hence, if under-reporting due to the hassle of collecting receipts plays a significant role in our context, we should observe signs of a missing mass above the threshold. In Figure 7 we plot the full distribution of donations after removing those individuals bunching at exactly CYP150 to highlight the patterns surrounding the threshold.²⁸ From this figure it is difficult to detect any missing mass above the threshold - if anything the level above the threshold appears higher than the level of the distribution below. This would suggest that the bunching mass predominantly originates from below rather than above the threshold and hence is the result of tax evasion rather than a hassle cost.

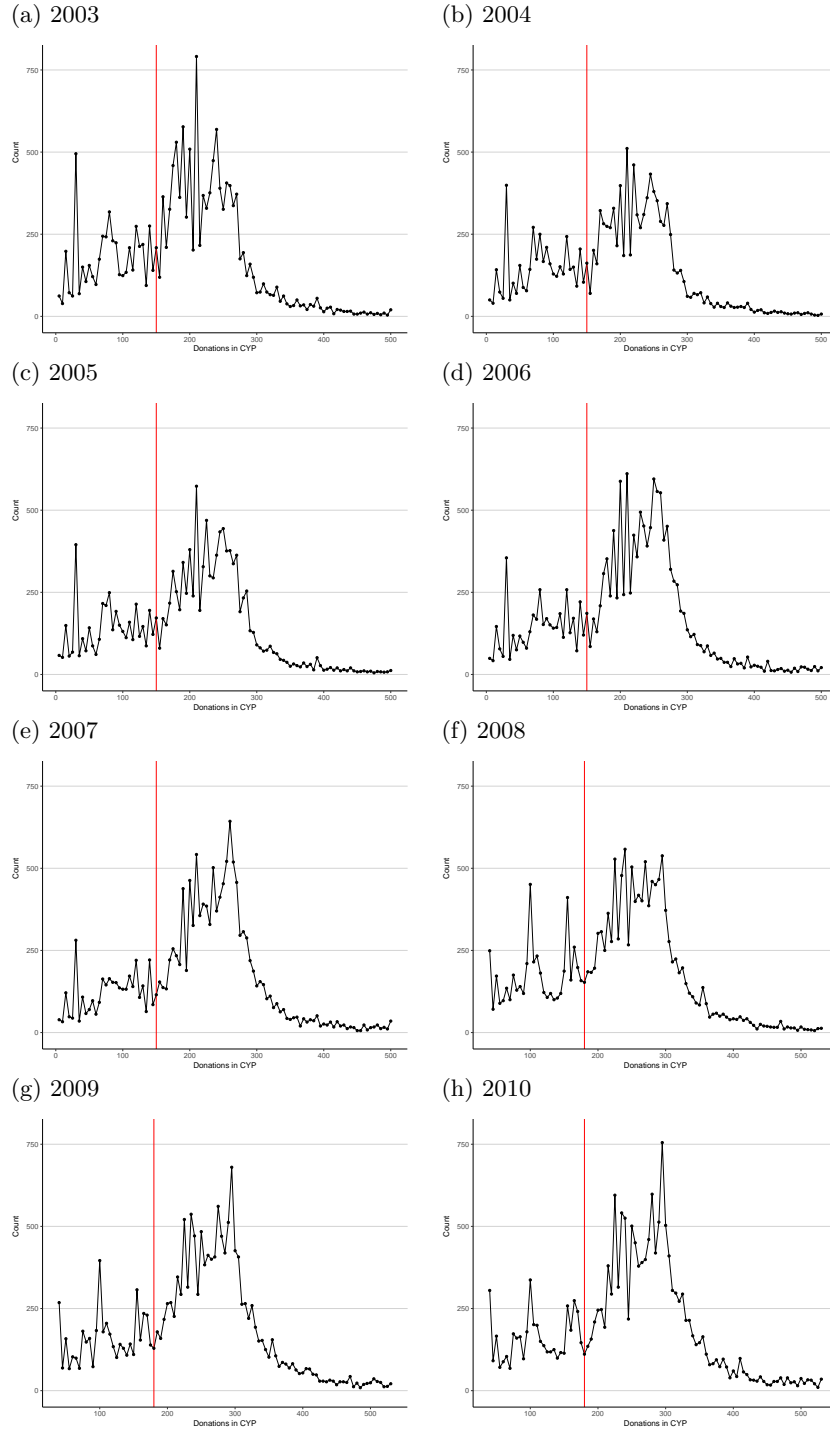
A standard justification for the lack of missing mass in other bunching studies has been the presence of frictions preventing taxpayers from accurately hitting a desired target (Chetty et al., 2011; Kleven and Waseem, 2013; Gelber et al., 2017). In our setting this is very unlikely to be the case since hitting a target is as easy as writing the correct number on a tax return form. A counter-argument could be that the dominated region could be too small to be visible. This however also appears unlikely as even a negligible hassle cost of CYP 5, would predict a dominated region spanning at least 5 bins above CYP150, depending on the tax filer's marginal tax rate.²⁹

²⁸The figure uses the sample of salary earners in non-unionised sectors to allow precise measurement of the reported donation variable for each individual over time. All results in this section are robust to using the full sample of salary earners in all sectors corrected for union fees (excluding the public sector for which we cannot separate union fees and donations as discussed earlier).

²⁹To estimate this, we consider a taxpayer with real donations above the threshold, $d_i > \alpha$. From the theoretical framework presented in section 3 consumption is given by: $c_i = y_i - d_i + t \cdot r_i - h_i \mathbb{1}_{(r_i > \alpha)}$. The consumption gain from collecting receipts and reporting true donations d_i rather than under-reporting down to the threshold is then, $\Delta c_i = t \cdot (d_i - \alpha) - h_i$. Hence, the taxpayer will collect receipts if $\Delta c_i > 0$, or $d_i > \alpha + h_i/t$. This gives us the upper bound of the dominated region for taxpayer i .

Figure 7:

Distribution of reported donations excluding bunchers



Notes: The figure shows the yearly empirical distributions of reported donations in bins of width CYP 5. The sample contains salary-earners in non-unionised sectors. We remove bunching at the threshold to clearly see the patterns of the distribution around the threshold. We also remove potential bunching at round numbers (multiples of 50). Vertical solid lines mark the relevant threshold that is in place in a given year (CYP 150 during 2003-2007 and CYP 175 during 2008-2010).

An alternative approach to evaluating the motive of the reporting response is to look at taxpayers' behaviour before they bunch at the threshold. Each year new taxpayers move to the threshold creating a gradual build-up of mass. This could be due to some delay in learning about the rules and the reporting setup or it could be caused by fluctuations in real giving from year to year. Appendix Figure 30 shows reported donations of taxpayers in the year before they move to the CYP 150 threshold. For each year between 2003 and 2006 we look at the sample of taxpayers who start bunching at the reporting threshold in the subsequent year. For this sample we display the fraction of taxpayers with reported donations in each of the bins above and below the threshold.³⁰ The figure hence illustrates where taxpayers are moving from, when they decide to move to the threshold. Naturally there will be some movements from the entire span of the distribution, however, we see a very clear pattern of additional mass coming from very low levels of donations. This pattern is stable, with only little variation across years. Interestingly, very few move from a position closely below the threshold, which seems to rule out any prominent role for bunching due to inflation or other similar upward time trends.

These findings are largely consistent with two scenarios. If we assume that true giving is relatively stable over time, then this is firstly consistent with a tax evasion story. Taxpayers with very low levels of real donations have most to gain from over-reporting up to the threshold. If we assume a fixed cost of evasion, we should exactly expect this group to be most likely to evade. At the same time, we observe no pattern of additional mass from the area immediately above the threshold. If we consider a compliance cost motive for bunching, then we should expect taxpayers with real donations just above the threshold to be the most likely to bunch, as this group has the least to gain from collecting receipts. Secondly, these patterns are also consistent with a scenario where peoples' real donations fluctuate a lot from year to year and tend to be either very high (above the threshold) or very low (close to zero). If taxpayers report truthfully when their real giving is low, while they report exactly at the threshold when their real giving is high to avoid a hassle cost, then we would also observe these patterns. While we cannot directly distinguish between these two scenarios, the tax evasion scenario perhaps appears most relevant if we factor in the distributional patterns around the threshold illustrated above in Figure 7.

In a final approach to analysing the source of reporting responses, we investigate the group of taxpayers most likely to move across the threshold as a result of the 2008 reform. In 2008 the threshold moves up unexpectedly, and hence taxpayers with real donations closely above the threshold find themselves below the new threshold. Before the reform this group has a strong incentive to bunch by

³⁰About 3000 new taxpayers move to the CYP 150 threshold per year between 2003-2006.

the hassle cost motive while after the reform they can report truthfully without collecting receipts.³¹ Following these considerations we should expect a number of previous bunchers at CYP 150 to move into the range between the old and the new threshold in 2008 if bunching comes from above and reporting is truthful below the threshold.³² If this is not the case, then either the bunching is coming from below and taxpayers start evading more after the reform, or taxpayers moving across the threshold go directly from under-reporting to over-reporting. Both of these latter scenarios would suggest a substantial role for over-reporting in this context.

Looking at the range between the two thresholds in 2008, we observe very few taxpayers re-locating from the earlier threshold into this range. To compare this number to what we would expect to observe in this range, we generate a prediction using a 2007 counterfactual distribution. Using standard bunching methodology we shift the 2007 distribution above the CYP 150 threshold upwards using a fraction of the bunching mass. Consequently, this counterfactual is based on the assumption that a specific percentage of the 2007 bunching mass originates from above. Even if we assume this percentage to be 1 and hence shift the counterfactual upwards by a negligible amount, our counterfactual still predicts more people in this range than what we observe in 2008. This again suggests a small role for the hassle cost motive unless this group of taxpayers moves directly from being under-reporters to being over-reporters.

Considering the patterns presented in this section, they all point towards tax evasion as a likely underlying source of the reporting responses we observe in previous sections. This does not rule out the logical notion that a hassle cost exists, but rather suggests that this cost is likely small and is not the driver of the results we find in the Cypriot context.

7 Optimal location of the reporting threshold

As mentioned earlier, the presence of a reporting threshold is a common feature of this enforcement policy, observed across many different countries and settings. In fact thresholds are widely used in the context of other reporting policies as well. Using our conceptual framework and the empirical results presented above, we can now examine the key trade-offs associated with the optimal place-

³¹All taxpayers with true donations above the threshold can choose either to collect receipts and receive a deduction equal to true donations, or to disregard receipts and forego part of the deduction. The forgone part of the deduction is equal to the distance between true donations and the threshold amount. Consequently filers whose true donations are closer to the threshold will forego a smaller deduction than those further from the threshold by choosing not to collect receipts. If a substantial part of the bunching mass at CYP 150 represents under-reporting, we would thus expect many bunchers to have real donations closely above the threshold.

³²Here we assume heterogeneous tastes for true giving and hence a continuous distribution of true donations absent any incentives created by the reporting environment.

ment of reporting thresholds, as well as the implications for optimal policy in the Cypriot context.

From the government's point of view the optimal reporting threshold trades off multiple effects. Firstly, there is a budget effect on the overall cost of providing the subsidy. This effect consists of two behavioural components - more evasion makes it costlier to provide the subsidy, while under-reporting due to compliance costs makes it less costly. Secondly, the compliance burden imposed on everyone above the threshold constitutes a utility loss,³³ while the gain from evasion represents a utility gain for evaders. Further, any decrease in real giving, caused by the compliance cost above the threshold, leads to a loss in the overall externality from giving. Lastly, the placement of the threshold affects the administrative burden on the government since it determines the amount of documentation that needs to be validated.

Consider a baseline scenario where the density of donors is roughly flat around the threshold and the individual level of donations is not correlated with other individual features such as the size of the hassle cost, h_i , or evasion cost, e_i . Now consider a small increase in the level of the reporting threshold. Evasion will increase because people below the threshold will evade more. Filers who were already evading will increase over-reporting as the individual evasion cost, e_i , is fixed. Further, for a subset of individuals the evasion benefit, $t(\alpha - d_i)$, will increase enough to outweigh the cost, e_i , and hence they will start evading. Lastly, some taxpayers with real donations just above the old threshold will now fall below and thereby get the opportunity to evade by a small amount. Those with a low cost of evasion will choose to do so when given the option.

As you increase the threshold fewer people are affected by the documentation requirement and hence fewer have to incur the compliance cost. This will lower the utility loss from the hassle imposed on everyone above the threshold. Further, when increasing the threshold the government has to validate fewer receipts and hence the administrative burden on the government decreases. Since the density is flat and the level of donations is uncorrelated with other individual features, the overall amount of under-reporting and the potential decrease in real giving above the threshold should stay roughly the same when moving the threshold.

Consequently, if the threshold increases, evasion will increase, but the collective compliance burden will decrease along with the administrative burden on the government. This is the basic trade-off faced by the government when setting the threshold.³⁴

³³Note that included in this utility loss is the potential loss from a suboptimal level of real giving if people above the threshold reduce genuine donations, as well as a utility loss from lost deductions from under-reporting.

³⁴For a formal description of the problem faced by the government and a theoretical derivation of the optimal threshold in a similar setup, see Tazhitdinova (2018).

Our empirical analysis above indicates very sizeable responses to the reporting threshold in the context of the hybrid reporting policy in Cyprus. This suggests that there might be sizeable gains from setting this threshold optimally and analogously considerable losses from setting it too high or too low. We find multiple indications that most of the massive bunching we observe originates from below rather than above the threshold. Accordingly, this would imply that the hassle cost plays a minor role, while the evasion component is substantial at the current level of the threshold. If the hassle cost is low, then the loss from an increased compliance burden when decreasing the threshold is low. On the other hand if evasion is high, then the gain from less evasion when lowering the threshold is high. Hence, under the assumption of a low or moderate administrative burden from validating documentation, our empirical evidence suggests that a decrease in the threshold is likely to be beneficial in this context. Of course this will depend on the objectives of the government, and how highly the utility of compliers is valued relative to the utility of evaders.

The baseline assumptions made here are more likely to hold when thinking about small local movements of the threshold. When considering large changes it may be important to factor in how the density and types of donors change. If the density of donors is decreasing then the number of people in the area immediately above the threshold decreases as the threshold is moved up. This means that fewer people will lower their reporting and potentially also their real giving in response to the threshold if it is placed higher in the distribution. If people who donate more for instance have lower compliance costs than people who donate less, then this could also mean less under-reporting above the threshold when the threshold is placed high. We abstract from such considerations in the discussion above since compliance costs appear to be low in our context and therefore these effects are likely of second order.

8 The elasticity of giving and the reporting environment

Taking a step back, the fundamental objective of the government when offering a tax deduction is to incentivise charitable giving by making it cheaper. As our analysis has so far established, the behaviour of many taxpayers is heavily influenced by the reporting environment surrounding this policy and specifically by the placement of thresholds. If taxpayers' behaviour is fundamentally guided by the enforcement setup, this raises the question of whether this precludes responses to price incentives. Put differently, does the design of reporting policies - in our case the presence of reporting thresholds - have implications for peoples' responsiveness to price changes, potentially undermining the objective of the government subsidy.

To look at these questions, we study the elasticity of reported donations with respect to price in the Cypriot population. For this analysis we exploit a further source of quasi-experimental variation -

changes in the price of giving generated by tax rate reforms. We look at the overall responsiveness in the population, but we also look separately at individuals who appear very responsive to the enforcement setup. Since the behaviour of this large group seems to be guided by enforcement concerns we might expect that they are less responsive to price incentives compared to the rest of the population. If this is the case then how large are the magnitudes of such differences and what does that mean for the interpretation of various elasticity estimates across the literature based on different samples of individuals subject to different reporting requirements?

The income tax rate reforms in our sample period are illustrated in the Appendix Figure 11. We can exploit these reforms as a source of variation in the price of giving, since the tax rate determines the size of the tax subsidy to charitable donations. The typical approach in the literature on the elasticity of the tax price of giving is to run log-specifications of the form:

$$\ln(r_{it}) = \beta_1 \ln(1 - \tau_{it}) + \beta_2 \ln(y_{it}) + \beta_3' X_{it} + \Gamma_i + \Gamma_t + \varepsilon_{it} \quad (3)$$

where r_{it} is the reported donation amount and y_{it} is disposable income before donations of individual i at time t . τ_{it} is the marginal tax rate and hence $1 - \tau_{it}$ is the price of giving. X_{it} is a vector of other controls. Specifications estimated using panel data can also include individual and time fixed effects, Γ_i and Γ_t . The price elasticity is then given by β_1 .

Estimating this equation using standard OLS leads to an endogeneity problem. Charitable donations can affect the price of giving because these may shift taxpayers to lower tax brackets, thereby reducing the tax price and causing an upward bias in the estimated elasticities. This is a well-known endogeneity problem in the literature, and has been typically dealt with by instrumenting the *last-pound* (observed) price of giving with the *first-pound* price of giving ($1 - \tau_{it|r_{it}=0} \equiv 1 - \tau_{it}^*$). This is the price a taxpayer faces for the first pound of charitable contribution. This removes any price variation due to charitable giving, and results in a very strong first-stage because the first- and last-pound prices are mechanically very highly correlated.

For the exclusion restriction to hold, the relationship between the first-pound price and the level of donations must solely go through the last-pound-price of giving. As argued by Almunia et al. (2020), this exclusion restriction is violated when using price variation from tax reforms because such reforms create a second source of endogeneity. Specifically, changes in marginal tax rates can also affect other choices such as individual labour supply and earnings more generally. Tax reforms therefore affect choices which both enter the donation decision and directly affect the first-pound price, because they determine which tax bracket a taxpayer is in, thereby violating the exclusion restriction.

Almunia et al. (2020) propose a solution that leverages the availability of panel data, based on the Gruber and Saez (2002) IV strategy that is widely used in the literature on the elasticity of taxable

income (for a review see Saez et al., 2012). Their instrument uses lagged income values to predict the reform-induced change in the price of giving. Specifically, they propose estimating the following differenced equation:

$$\Delta \ln(r_{it}) = \beta_1 \Delta \ln(1 - \tau_{it}^*) + \beta_2 \Delta \ln(y_{it}) + \beta_3' \Delta X_{it} + \Delta \varepsilon_{it} \quad (4)$$

where $\Delta \ln(x_{it}) = \ln\left(\frac{x_{it}}{x_{i,t-k}}\right)$ for $x_{it} = r_{it}, 1 - \tau_{it}^*, y_{it}$, and the log change in the first-pound price is instrumented by:

$$\ln\left(\frac{1 - \tau_{i,t}^*(y_{i,t-k}^*)}{1 - \tau_{i,t-k}^*(y_{i,t-k}^*)}\right) \quad (5)$$

The variable k determines the time horizon of the difference. In words, the instrument is the change in the price of giving from time $t - k$ to time t if taxable income at zero donations (y_{it}^*) remained unchanged. This instrument solves the endogeneity problem because it uses past (pre-tax) income which should not be affected by future reform-related choices. In our empirical application, we follow Almunia et al. (2020) and implement the differenced-IV specification (4). To compare our findings with existing practice, we also report results for the non-differenced IV version of specification (3) in the Appendix.³⁵

As already mentioned, our measure of donations includes a union membership fee for some subset of workers. We cannot directly observe the size of this fee or who is a member of a union. However, using detailed information on sector of employment we can back it out with some noise. Union membership in Cyprus is highly concentrated in a few sectors. One of these is the banking sector which is large and (almost) fully unionised due to industry-wide agreements and automatic enrolment upon employment. This means that all workers in this sector pay the same fraction of salary in membership fees and hence we can completely separate donations from fees.³⁶ A small number of other sectors, such as the construction and hotel services, are highly but not fully unionised.³⁷ We can therefore account for fee payments for these sectors with some small error. Lastly Cyprus has a number of sectors with very low unionisation. In these sectors the donation measure will be somewhat noisy due to the small number of unionised workers. Given these considerations, we use the sample of workers from the banking sector as our main sample in the analysis, but show that all results are robust to using either the full sample or the sample of highly unionised sectors instead.

³⁵Using this methodology we do not explicitly deal with censoring coming from people reporting zero donations. In general, this can potentially create bias because such observations are excluded due to the logarithmic specification. In our setting however this is not a substantial concern since close to 80 percent of tax filers report positive donations (after correcting for union fees in highly unionised sectors).

³⁶The banking sector makes up 7% of our sample.

³⁷Hotel and construction have unionisation rates of about 75%, and make up 3% and 7% of our sample respectively.

8.1 Elasticity estimates

We split our sample into two types of taxpayers - those who appear to respond strongly to the enforcement environment and those who do not. Concretely, the first group consists of taxpayers who at some point bunch, meaning that at some point in the sample period we observe them exactly at a threshold value. The second group consists of those taxpayers who never bunch meaning that we never observe them exactly at a reporting threshold. We then estimate elasticities for the whole population as well as separately for these two groups of taxpayers using the method introduced above. We need to clearly separate responses to prices from those caused by changes in reporting thresholds and hence related to the reporting setup. Therefore we focus on the period 2003-2007, since this is the longest period in the data for which the reporting environment remains constant.³⁸

Table 2 shows our results for these separate samples as well as for the full population. All columns

Table 2:

Elasticity of reported donations wrt price - bunchers vs non-bunchers

	(1) All	(2) All	(3) Non-B	(4) Non-B	(5) B	(6) B
$\Delta \ln(1 - \tau^*)$	-0.419*** (0.160)	-0.375*** (0.154)	-0.818*** (0.261)	-0.786*** (0.247)	-0.253 (0.203)	-0.207 (0.198)
Observations	15 957	15 957	4 586	4 586	11 371	11 371
R^2	0.20	0.21	0.27	0.28	0.17	0.18
Individual FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Controls	-	✓	-	✓	-	✓

Notes: The sample includes all workers from the banking sector in years 2003-2007. In all specifications we control for income. Additional controls include age squared and a dummy for changing employer. We drop people below the first tax bracket (i.e. people with no tax liability). *Non-B* denotes the sample of workers that are never observed at a reporting threshold, while *B* denotes the sample of workers that at some point in the sample period are observed exactly at a reporting threshold. We report robust standard errors clustered at the individual level in parenthesis, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

report estimates using the differenced IV strategy described above with a time horizon of one year and individual and year fixed effects.³⁹ Columns 1 and 2 show results from specifications with and without

³⁸In both threshold reform years 2003 and 2008 for instance, we observe changes in marginal tax rates and hence prices. Note that we cannot look exclusively at the last period 2008-2010, since we have no variation in tax rates in this period. Including the period before 2003 does not change our qualitative findings.

³⁹We use a time horizon of one year, since we are looking at a relatively short time period. However, if we use the full sample period and estimate the elasticity using different time horizons, overall results do not seem sensitive to the choice of horizon.

additional controls for the full sample including all taxpayers. Using our preferred specification (including controls in column 2) we find an elasticity of -0.38 . Columns 3 and 4 only consider taxpayers who never bunch at a reporting threshold, and show an elasticity estimate of -0.79 using our preferred specification. These results clearly show that the non-buncher group displays much more sensitivity to price changes than the collective group of taxpayers, with estimated elasticities doubling in magnitude when considering only non-bunchers. Columns 5 and 6 show estimated elasticities for taxpayers who at some point bunch at a threshold value. This group appears much less responsive to price changes than the rest of the population, with an estimated elasticity of only -0.21 . This estimate is further statistically insignificant at all conventional levels and the magnitude of the elasticity is around half of the estimate for the entire group and about a quarter of the estimate for the non-bunchers.⁴⁰

For comparison with existing literature we also report elasticities using the non-differenced first-pound price IV strategy. These results are reported in Appendix Table 11 and show larger estimated elasticities for all groups. However, importantly we find the exact same patterns in the relative magnitudes of elasticities between bunchers, non-bunchers and the full population. For robustness, we repeat the estimation using instead either the sample of all highly unionised sectors, or our full sample. These results reveal the same patterns.⁴¹

As one would expect these differences are mainly driven by stickiness in bunching behaviour around reporting thresholds. We observe bunchers staying at the reporting threshold for multiple consecutive years and hence bunchers are significantly less likely to react to price changes compared to the rest of the population. In the Appendix Table 14 we run a simple linear probability model using as outcome an indicator for changing reporting from one year to the next. We run this regression both conditional and unconditional on a change in price and find a large and statistically significant difference between those observed at a threshold value and those at other values, in the probability of changing reporting behaviour from one year to the next.

Given this reporting setup we find that a very large part of the population is fairly unresponsive to

⁴⁰To present a formal test of the difference in elasticities between bunchers and non-bunchers we run a pooled regression with full interactions. The difference in elasticities from our main specification with controls in Table 2 is $-0.207 + 0.786 = 0.579$. The standard error for this estimate from the pooled regression is 0.316, indicating marginal significance at the 10% level.

⁴¹See Appendix Tables 12 and 13 for the results of these regressions. The patterns are identical to our main results, but the estimated differences are slightly smaller in magnitude, consistent with the fact that we cannot precisely separate bunchers from non-bunchers in these groups given the noise in the donation measure. For the highly-unionised sample the difference in elasticities between bunchers and non-bunchers is 0.437 significant at the 10% level, while the difference is 0.297 for the full sample, significant at the 5% level.

price incentives for charitable giving and much less responsive than the rest of the population. The behaviour of this group is instead shaped by the enforcement setup around the subsidy. Consequently, the substantial elasticity of giving that we find in the population is almost entirely driven by the minority of taxpayers who do not appear responsive to the design of the reporting environment. This finding seems to suggest that reporting policies can affect the success of financial incentive policies by separating the real decision from the reporting decision for a subset of people. Since we do not directly observe taxpayer behaviour in the absence of reporting thresholds, we cannot conclude directly on how this feature affects elasticities. However, these results indicate that unless a large fraction of the population is inherently only weakly affected by price-levels in their donation decision, the elasticity of charitable giving would be substantially different under another reporting setup. Reporting thresholds are a common component of reporting policies across the world and especially in the context of deductions used to incentivise desired behaviours. Understanding how this feature impacts behaviour and interacts with the deduction is therefore important to be able to attain the desired objectives of the tax system. It is also paramount for the external validity and interpretation of various estimates of the elasticity of giving with respect to price as well as other similar elasticities reported in the literature. We show that such thresholds are a key driver for the behaviour of many taxpayers and therefore more focus is warranted to understand the direct effect on giving.

9 Conclusion

This paper studies behavioural responses to a widely-used tax enforcement policy that combines elements of self- and third-party reporting, using the context of charitable contributions in the Republic of Cyprus. We hence analyse a semi-formal economy, where we can exploit multiple sources of quasi-experimental variation in reporting requirements and tax-price subsidies to present several policy-relevant results.

First, we show evidence of substantial reactions to this hybrid reporting policy. Exploiting salary-dependent cutoffs that govern documentation requirements, we estimate that reported donations increase by 0.7 pounds when taxpayers can report 1 pound more without providing corroborating information from a third party. Second, by utilising a reform that retroactively shifted the location of the threshold activating the need for documentation we show that a very large part of this response is purely a reporting adjustment. Our bunching analysis reveals that at least 64 percent of the response is purely due to changes in reporting and not to changes in real giving. Further, looking closely into the time-patterns of individual reporting decisions reveals tax evasion as the main underlying source of the large reporting response. Finally, we show that the reporting environment appears to be im-

portant for taxpayers' responsiveness to the size of the subsidy for giving. We estimate the elasticity of reported giving with respect to price using quasi-experimental variation in tax prices generated by income tax reforms. We find that the reporting environment is the key driver of behaviour for a large part of the population who show little sensitivity to financial incentives.

Our findings have important implications, both for policy and tax theory. The very strong behavioural responses, observed around the reporting thresholds, imply that the policy strongly affects the level of misreporting of deductions and hence tax evasion. Consequently, hybrid reporting with a simple self-reported documentation requirement can potentially have a large effect on government revenue even in a semi-formal setting with high reliance on cash transactions. Further, our results suggest that taxpayers' responses to tax subsidies could potentially be very different in the absence of this reporting policy and the embedded thresholds determining the strictness of reporting requirements. Our findings thereby highlight how important the conjoint analysis of both the enforcement environment and tax prices is to optimal tax policy design. To the extent that the fiscal authority wants to incentivise certain forms of behaviours that generate positive externalities, it is crucial to understand the conditions under which tax subsidies cannot achieve this goal.

References

- Ackerman, D. and Auten, G. (2011). Tax expenditures for noncash charitable contributions. *National tax journal*, 64:651–687.
- Agostini, C. A. and Martínez A, C. (2014). Response of tax credit claims to tax enforcement: Evidence from a quasi-experiment in chile. *Fiscal Studies*, 35(1):41–65.
- Alm, J., Deskins, J., and McKee, M. (2009). Do individuals comply on income not reported by their employer? *Public Finance Review*, 37(2):120–141.
- Almunia, M., Guceri, I., Lockwood, B., and Scharf, K. (2020). More giving or more givers? the effects of tax incentives on charitable donations in the uk. *Journal of Public Economics*, 183:104114.
- Almunia, M. and Lopez-Rodriguez, D. (2018). Under the radar: The effects of monitoring firms on tax compliance. *American Economic Journal: Economic Policy*, 10(1):1–38.
- Artavanis, N., Morse, A., and Tsoutsoura, M. (2016). Measuring income tax evasion using bank credit: Evidence from greece. *The Quarterly Journal of Economics*, 131(2):739–798.
- Besley, T. and Persson, T. (2014). Why do developing countries tax so little? *Journal of Economic Perspectives*, 28(4):99–120.
- Blundell, R., Brewer, M., and Francesconi, M. (2008). Job changes and hours changes: understanding the path of labor supply adjustment. *Journal of Labor Economics*, 26(3):421–453.
- Carrillo, P., Pomeranz, D., and Singhal, M. (2017). Dodging the taxman: Firm misreporting and limits to tax enforcement. *American Economic Journal: Applied Economics*, 9(2):144–64.
- Chetty, R., Friedman, J. N., Olsen, T., and Pistaferri, L. (2011). Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from danish tax records. *The Quarterly Journal of Economics*, 126(2):749–804.
- Dickens, W. T. and Lundberg, S. (1993). Hours restrictions and labor supply. *International Economic Review*, 34(1):169–92.
- Doerrenberg, P., Peichl, A., and Siegloch, S. (2017). The elasticity of taxable income in the presence of deduction possibilities. *Journal of Public Economics*, 151:41–55.
- Esselink, H. and Hernández, L. (2017). The use of cash by households in the euro area. *ECB Occasional Paper*, (201).

-
- Fack, G. and Landais, C. (2016). The effect of tax enforcement on tax elasticities: Evidence from charitable contributions in france. *Journal of Public Economics*, 133:23–40.
- Gelber, A. M., Jones, D., and Sacks, D. W. (2017). Estimating earnings adjustment frictions: Method and evidence from the social security earnings test. *Goldman School of Public Policy Working Paper*.
- Gillitzer, C. and Skov, P. E. (2018). The use of third-party information reporting for tax deductions: evidence and implications from charitable deductions in denmark. *Oxford Economic Papers*, 70(3):892–916.
- Giving USA (2019). The annual report on philanthropy for the year 2018.
- Gordon, R. and Li, W. (2009). Tax structures in developing countries: Many puzzles and a possible explanation. *Journal of Public Economics*, 93(7-8):855–866.
- Gruber, J. and Saez, E. (2002). The elasticity of taxable income: evidence and implications. *Journal of Public Economics*, 84(1):1–32.
- Hassan, M. and Schneider, F. (2016). Size and development of the shadow economies of 157 worldwide countries: Updated and new measures from 1999 to 2013. *Journal of Global Economics*, 4(3):1–15.
- Ioannou, G. and Sonan, S. (2014). *Trade Unions in Cyprus: History of Division, Common Challenges Ahead*. Friedrich-Ebert-Stiftung.
- Jensen, A. (2019). Employment structure and the rise of the modern tax system. *NBER Working paper No. 25502*.
- Kleven, H. J. (2016). Bunching. *Annual Review of Economics*, 8:435–464.
- Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., and Saez, E. (2011). Unwilling or unable to cheat? evidence from a tax audit experiment in denmark. *Econometrica*, 79(3):651–692.
- Kleven, H. J. and Waseem, M. (2013). Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from pakistan. *The Quarterly Journal of Economics*, 128(2):669–723.
- Kopczuk, W. (2005). Tax bases, tax rates and the elasticity of reported income. *Journal of Public Economics*, 89(11-12):2093–2119.
- Kuchumova, Y. P. (2017). The optimal deterrence of tax evasion: The trade-off between information reporting and audits. *Journal of Public Economics*, 145:162–180.

-
- Kumler, T., Verhoogen, E., and Frías, J. (2020). Enlisting employees in improving payroll-tax compliance: Evidence from Mexico. *Review of Economics and Statistics*, pages 1–45.
- LaLumia, S. and Sallee, J. M. (2013). The value of honesty: empirical estimates from the case of the missing children. *International Tax and Public Finance*, 20(2):192–224.
- Mavrokonstantis, P. (2019). Bunching: Estimate bunching. *R package version 0.8.4*, URL: <https://CRAN.R-project.org/package=bunching>.
- Mavrokonstantis, P. and Seibold, A. (2020). Bunching and adjustment costs: Evidence from Cypriot tax reforms. *Working paper*.
- Naritomi, J. (2019). Consumers as tax auditors. *American Economic Review*, 109(9):3031–72.
- Phillips, M. D. (2014). Individual income tax compliance and information reporting: what do the US data show? *National Tax Journal*, 67(3):531–567.
- Pomeranz, D. (2015). No taxation without information: Deterrence and self-enforcement in the value added tax. *American Economic Review*, 105(8):2539–69.
- Rocha, R., Ulyssea, G., and Rachter, L. (2018). Do lower taxes reduce informality? evidence from Brazil. *Journal of Development Economics*, 134:28–49.
- Saez, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, 2(3):180–212.
- Saez, E., Slemrod, J., and Giertz, S. H. (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature*, 50(1):3–50.
- Schneider, F. and Enste, D. H. (2000). Shadow economies: Size, causes, and consequences. *Journal of Economic Literature*, 38(1):77–114.
- Seibold, A. (2020). Reference points for retirement behavior: Evidence from German pension discontinuities. *Working Paper*.
- Slemrod, J. (1989). Are estimated tax elasticities really just tax evasion elasticities? the case of charitable contributions. *The Review of Economics and Statistics*, pages 517–522.
- Slemrod, J. (1994). Fixing the leak in Okun’s bucket: optimal tax progressivity when avoidance can be controlled. *Journal of Public Economics*, 55(1):41–51.
- Slemrod, J. (2019). Tax compliance and enforcement. *Journal of Economic Literature*, 57(4):904–54.

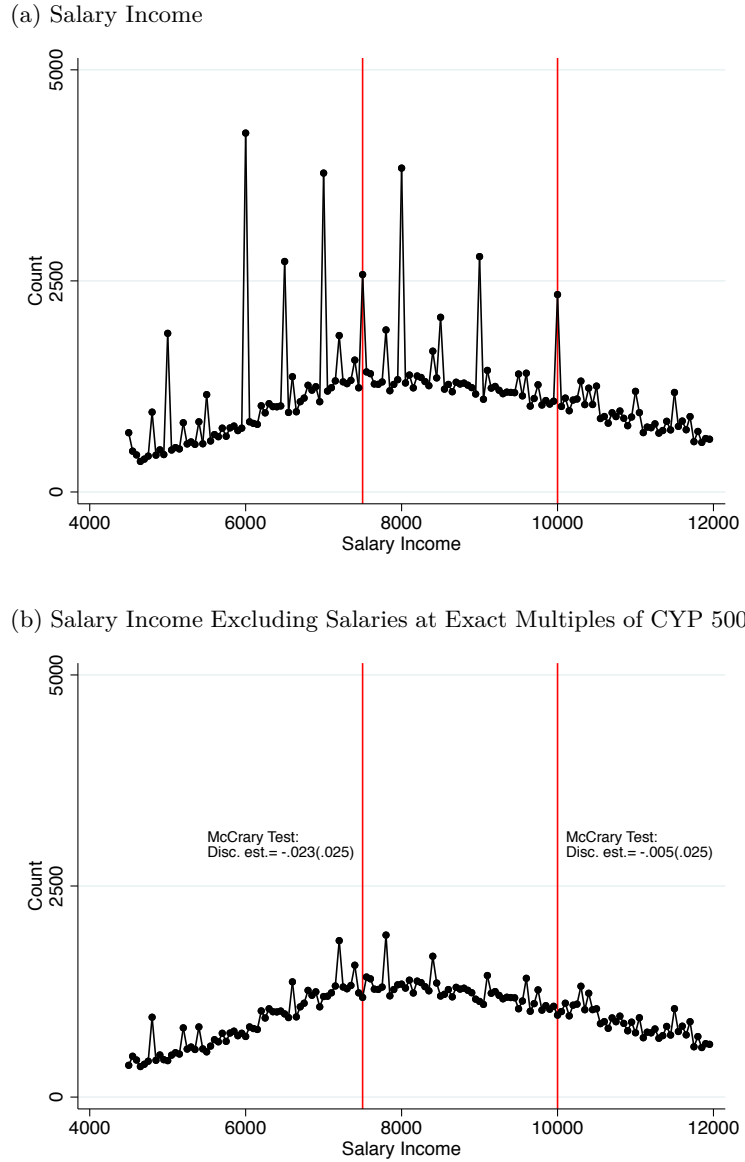
-
- Slemrod, J., Collins, B., Hoopes, J. L., Reck, D., and Sebastiani, M. (2017). Does credit-card information reporting improve small-business tax compliance? *Journal of Public Economics*, 149:1–19.
- Slemrod, J. and Kopczuk, W. (2002). The optimal elasticity of taxable income. *Journal of Public Economics*, 84(1):91–112.
- Tazhitdinova, A. (2018). Reducing evasion through self-reporting: Evidence from charitable contributions. *Journal of Public Economics*, 165:31–47.
- Waseem, M. (2018). Taxes, informality and income shifting: Evidence from a recent pakistani tax reform. *Journal of public economics*, 157:41–77.

10 Appendix A: Validity of RDD approach

In this appendix we check the validity of our regression discontinuity methodology used in Section 4. As mentioned the identifying assumption is absence of precise manipulation of the assignment variable, which in our context is individual salary income. We therefore investigate whether we observe any signs of such manipulation. Figure 8 shows the density of salary income around the salary cutoffs between 1999 and 2001. Figure 8a shows bunching in salary income at *all* multiples of CYP 500, which is characteristic of the fact that salaries have a high propensity to be set at round numbers. This is confirmed by Appendix Figure 8b where we drop individuals with a salary that is an *exact* multiple of CYP 500. This removes any signs of bunching in salary income around cutoffs. We use a McCrary test to formally test for the existence of any significant discontinuities in the density around each cutoff, the results of which are reported in the figure. The null of no discontinuity cannot be rejected, supporting our identifying assumption of no precise manipulation of s_i .

We also observe no discontinuities in covariates around the cutoffs. We consider multiple covariates: age, probability of being female, probability of working in a highly unionised sector, level of other deductions, the first and last price of giving (the prices before and after donations respectively). Figure 9 shows plots of each of these cases as a function of our assignment variable. All plots confirm smoothness around both cutoffs.

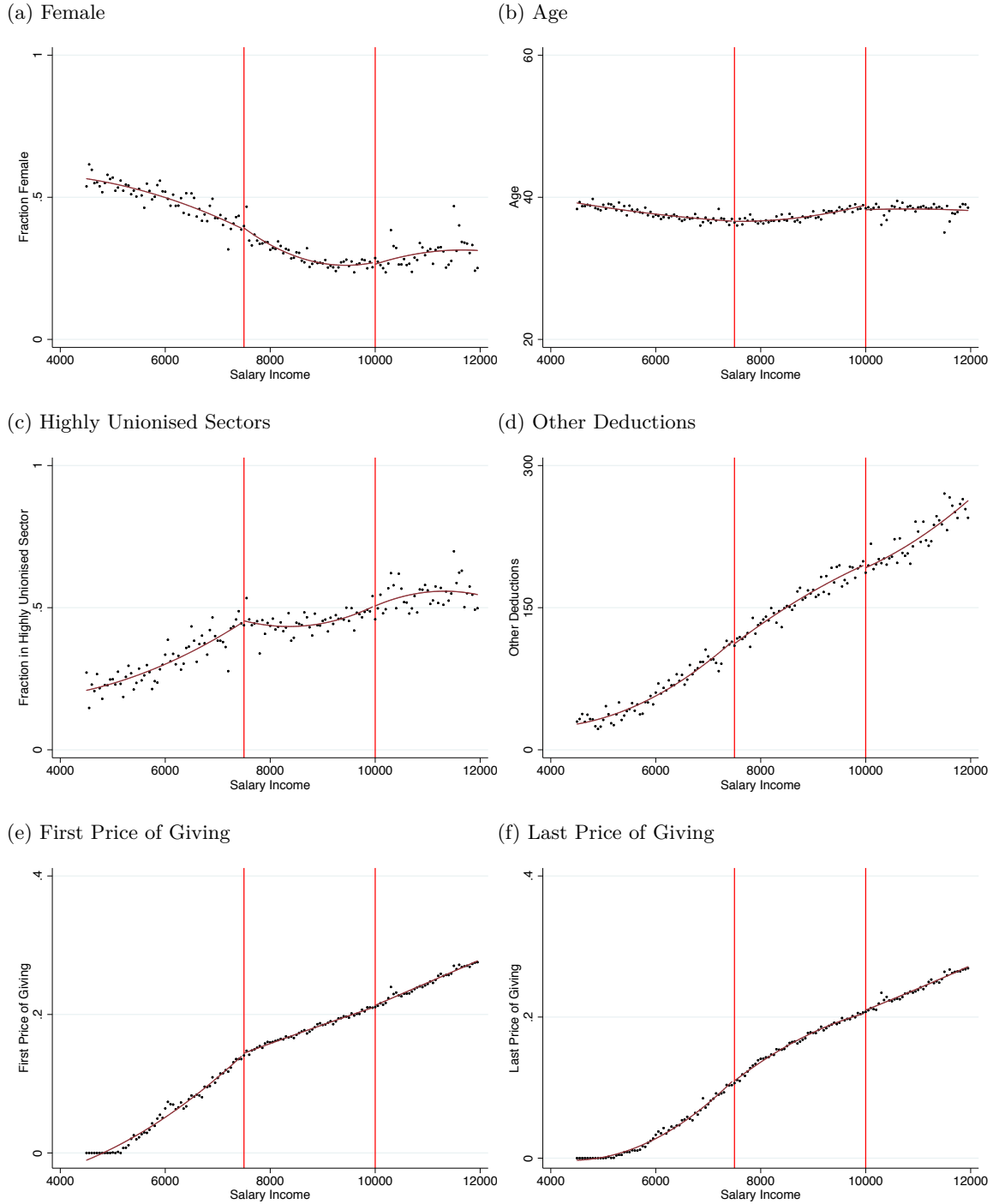
Figure 8:
Density of salary income between 1999-2001



Notes: This figure shows the density of salaries pooled over the years 1999-2001, for two samples: (a) all salary earners and (b) all salary earners excluding those earning at exact multiples of CYP500. In each case, the two earnings thresholds that we focus on (7,500 and 10,000) are marked with vertical lines. Panel (b) also reports the results from a McCrary test for discontinuities in the density of the assignment variable (the estimated log difference in height). The null of no discontinuity cannot be rejected at any of the two earnings cutoffs, in support of the assumptions of our RD design.

Figure 9:

Robustness check: Smoothness of covariates



Notes: This figure shows evidence in support of the RD identifying assumption. Each sub-figure shows the mean value of the given covariate in bins of width 50 of the assignment variable around each salary threshold. The sample is the same as in our main specification (pooled over 1999-2001).

11 Appendix B: Bunching methodology

This appendix provides further details of our bunching methodology. Following standard approaches we group our donation variable d_i into bins of width $\delta = \text{CYP5}$ and fit the following equation:

$$c_j = \sum_{i=0}^p \beta_i (d_j)^i + \sum_{i=d_L}^{d_U} \gamma_i \mathbb{1}[d_j = i] + \sum_{r \in R} \rho_r \mathbb{1}\left[\frac{d_j}{r} \in \mathbb{N}\right] + \sum_{k \in K} \theta_k \mathbb{1}\left[d_j \in K \wedge d_j \notin [d_L, d_U]\right] + v_j \quad (6)$$

where c_j is the observation count in donation bin j , p is the order of polynomial used to fit the counts (11th order), and d_L and d_U represent the lower and upper region that define the bunching region around the threshold d^* .

To control for round numbers, we include a set of dummy variables (R) for multiples of 50 and 100. We also control for further fixed effects (included in set K) that exhibit strong bunching mass outside our bunching region and which cannot be captured purely by our round number controls. For instance, this includes bin $j = 100$ for years up to 2007. For estimates of our bunching mass at CYP150 in 2008-2010, K includes bin $j = 175$, and likewise for estimates at CYP175, K includes bin $j = 150$.

The fitted counterfactual, \hat{c}_j is then given by:

$$\hat{c}_j = \sum_{i=0}^p \hat{\beta}_i (d_j)^i + \sum_{r \in R} \hat{\rho}_r \mathbb{1}\left[\frac{d_j}{r} \in \mathbb{N}\right] + \sum_{k \in K} \hat{\theta}_k \mathbb{1}\left[d_j \in K \wedge d_j \notin [d_L, d_U]\right] \quad (7)$$

The bunching mass is simply the difference between the observed and counterfactual density in our bunching region:⁴²

$$\hat{B} = \sum_{j=d_L}^{d_U} (c_j - \hat{c}_j) \quad (8)$$

In our main figures, we also report the normalized bunching mass, defined as:

$$\hat{b} = \frac{\hat{B}}{\hat{c}_0} \quad (9)$$

where \hat{c}_0 is the average counterfactual frequency in the excluded range.

We set $d_L = d^*$, since our bunching is very sharp, with no diffuse bunching below the threshold. We further set $d_U = d^*$ since it not possible for a tax filer to be a "diffuse buncher" above the threshold in our setting. This would require that, (a) he knows the placement of the threshold, but (2) he

⁴²Note that because $d_L = d_U = d^*$, our bunching mass estimate is the same regardless of whether we treat our threshold as a kink or notch, since $\hat{B} = \sum_{j=d_L}^{d_U} (c_j - \hat{c}_j)$ for kinks and $\hat{B} = \sum_{j=d_L}^{d^*} (c_j - \hat{c}_j)$ for notches.

instead writes a slightly larger figure than the threshold amount he wanted to report on her form. While this is unlikely to happen in the first place, reporting slightly above the threshold also requires collecting receipts. It seems infeasible that someone who is planning to bunch at the threshold would still go through the trouble of collecting receipts.

Finally, we follow Fack and Landaï (2016) and do not shift the counterfactual to the right of the threshold upwards, as it is a priori unclear whether the bunching is coming from above or below. Nevertheless, we have re-estimated our lower bound ratios after shifting the counterfactual upwards (i.e. assuming that all the bunching mass is coming from above), and find nearly identical results.

12 Appendix C - Tables

Table 3:
Summary Statistics

	Mean	Std. Dev.
Salary Earnings Only	0.845	0.362
Ratio of Salary to Total Earnings	0.960	0.141
Taxable Income	12646.498	11139.126
Job Switches	0.028	0.164
Marginal Tax Rate	0.182	0.133
Positive Donations	0.873	0.333
Donations (cond. positive)	171.711	118.089
Positive Donations (Net of Union Fees)	129.87	103.436
Donations (Net of Union Fees, cond. positive)	0.788	0.409
Age	40.437	8.070
Female	0.385	0.487
Agriculture	0.005	0.069
Mining	0.003	0.051
Manufacturing	0.086	0.281
Construction	0.074	0.262
Utilities	0.021	0.142
Trade	0.120	0.325
Hotel Services	0.033	0.180
Other Services	0.186	0.389
Commercial Banking	0.066	0.248
Other Financial Services	0.033	0.179
Public Sector	0.360	0.480
Other	0.011	0.102
Observations	1,462,409	

Notes: This table displays summary statistics for our sample. We distinguish between positive donations, and positive donations net of union fees, where we residualise our measure in the latter case from union fees. Both measures have professional taxes already removed.

Table 4:

RD estimates - including rounders

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A:								
Above 10k	38.43*** (0.93)	36.35*** (1.39)	35.81*** (0.95)	34.10*** (1.42)	37.06*** (1.28)	37.40*** (2.00)	34.85*** (1.33)	35.29*** (2.09)
Scaled response	0.77	0.73	0.72	0.68	0.74	0.75	0.70	0.71
Observations	91 768	91 768	74 064	74 064	46 676	46 676	38 012	38 012
R^2	0.26	0.26	0.32	0.32	0.20	0.20	0.26	0.26
Panel B:								
Above 7.5k	29.50*** (0.70)	31.79*** (1.02)	30.10*** (0.75)	30.82*** (1.09)	31.72*** (0.95)	27.66*** (1.42)	30.98*** (1.01)	26.56*** (1.50)
Scaled response	0.74	0.79	0.75	0.77	0.79	0.69	0.77	0.66
Observations	107 684	107 684	81 197	81 197	60 945	60 945	46 272	46 272
R^2	0.33	0.33	0.39	0.39	0.22	0.22	0.30	0.30
Polynomial	p_1	p_2	p_1	p_2	p_1	p_2	p_1	p_2
Controls	-	-	✓	✓	-	-	✓	✓
Bandwidth	2 000	2 000	2 000	2 000	1 000	1 000	1 000	1 000

Notes: This table shows robustness checks from estimating specification (1) when we include rounders to our main sample. In this case, the specification also includes round number fixed effects. The *scaled response* is calculated as the parameter estimate divided by the notch size in the donation schedule (i.e. in panel A the parameter estimate is divided by 50 while in panel B the parameter estimate is divided by 40). p_s indicates that we fit a polynomial of order s on each side of the notch, while controls include sex, year and sector fixed effects. Robust standard errors clustered at the individual level in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5:

RD estimates - years 1999-2002

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A:								
Above10k	38.18*** (0.83)	34.43*** (1.25)	36.28*** (0.84)	33.41*** (1.27)	35.89*** (1.17)	34.91*** (1.80)	34.74*** (1.19)	35.09*** (1.84)
Scaled response	0.76	0.69	0.73	0.69	0.72	0.70	0.69	0.70
Observations	114 588	114 588	93 830	93 830	59 023	59 023	48 459	48 459
R^2	0.20	0.20	0.29	0.29	0.13	0.13	0.22	0.22
Panel B:								
Above 7.5k	28.75*** (0.69)	30.70*** (1.00)	29.21*** (0.71)	29.50*** (1.04)	29.76*** (0.94)	26.76*** (1.41)	28.78*** (0.96)	26.37*** (1.44)
Scaled response	0.72	0.77	0.73	0.74	0.74	0.67	0.72	0.66
Observations	116 837	116 837	90 641	90 641	64 941	64 941	50 592	50 592
R^2	0.26	0.26	0.36	0.36	0.15	0.15	0.27	0.27
Polynomial	p_1	p_2	p_1	p_2	p_1	p_2	p_1	p_2
Controls	-	-	✓	✓	-	-	✓	✓
Bandwidth	2 000	2 000	2 000	2 000	1 000	1 000	1 000	1 000

Notes: This table shows robustness checks from estimating specification (1) when we also include year 2002 to our main sample. The *scaled response* is calculated as the parameter estimate divided by the notch size in the donation schedule (i.e. in panel A the parameter estimate is divided by 50 while in panel B the parameter estimate is divided by 40). p_s indicates that we fit a polynomial of order s on each side of the notch, while controls include sex, year and sector fixed effects. Robust standard errors clustered at the individual level in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 6:

RD estimates - including individuals with some non-salary income

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A:								
Above 10k	35.07*** (0.89)	31.84*** (1.35)	33.11*** (0.91)	31.07*** (1.39)	33.17*** (1.26)	32.36*** (1.95)	32.34*** (1.30)	31.78*** (2.03)
Scaled response	0.70	0.64	0.66	0.62	0.66	0.65	0.65	0.64
Observations	96 826	96 826	79 271	79 271	49 475	49 475	40 784	40 784
R^2	0.20	0.20	0.27	0.27	0.12	0.12	0.20	0.20
Panel B:								
Above 7.5k	27.63*** (0.72)	29.46*** (1.05)	28.66*** (0.76)	28.80*** (1.11)	28.65*** (0.98)	24.49*** (1.48)	28.30*** (1.03)	24.79*** (1.54)
Scaled response	0.69	0.74	0.72	0.72	0.72	0.61	0.71	0.62
Observations	105 388	105 388	81 401	81 401	59 031	59 031	45 849	45 849
R^2	0.27	0.27	0.35	0.35	0.16	0.16	0.25	0.25
Polynomial	p_1	p_2	p_1	p_2	p_1	p_2	p_1	p_2
Controls	-	-	✓	✓	-	-	✓	✓
Bandwidth	2 000	2 000	2 000	2 000	1 000	1 000	1 000	1 000

Notes: This table shows robustness checks from estimating specification (1) when we include individuals with income from both salary and non-salary earnings to our main sample. The *scaled response* is calculated as the parameter estimate divided by the notch size in the donation schedule (i.e. in panel A the parameter estimate is divided by 50 while in panel B the parameter estimate is divided by 40). p_s indicates that we fit a polynomial of order s on each side of the notch, while controls include sex, year and sector fixed effects. Robust standard errors clustered at the individual level in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7:

RD estimates - outcome variable corrected

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A:								
Above 10k	36.58*** (1.17)	35.19*** (1.72)	35.64*** (1.22)	33.45*** (1.79)	35.06*** (1.62)	35.51*** (2.48)	33.55*** (1.70)	32.96*** (2.58)
Scaled response	0.73	0.70	0.71	0.67	0.70	0.71	0.67	0.66
Observations	62 960	62 960	47 402	47 402	32 269	32 269	24 578	24 578
R^2	0.11	0.11	0.24	0.24	0.08	0.08	0.22	0.22
Panel B:								
Above 7.5k	28.28*** (0.86)	27.25*** (1.25)	27.98*** (0.92)	25.68*** (1.34)	26.25*** (1.18)	23.67*** (1.78)	25.11*** (1.25)	22.60*** (1.88)
Scaled response	0.71	0.68	0.70	0.64	0.66	0.59	0.63	0.57
Observations	75 557	75 557	54 219	54 219	41 914	41 914	30 187	30 187
R^2	0.14	0.14	0.26	0.26	0.10	0.10	0.23	0.23
Polynomial	p_1	p_2	p_1	p_2	p_1	p_2	p_1	p_2
Controls	-	-	✓	✓	-	-	✓	✓
Bandwidth	2 000	2 000	2 000	2 000	1 000	1 000	1 000	1 000

Notes: This table shows robustness checks from estimating specification (1) when we net out professional taxes and union payments from our outcome variable. In this case, we exclude individuals working in the public sector. The *scaled response* is calculated as the parameter estimate divided by the notch size in the donation schedule (i.e. in panel A the parameter estimate is divided by 50 while in panel B the parameter estimate is divided by 40). p_s indicates that we fit a polynomial of order s on each side of the notch, while controls include sex, year and sector fixed effects. Robust standard errors clustered at the individual level in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 8:

RD estimates - only highly unionised sectors with corrected outcome variable

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A:								
Above 10k	38.77*** (2.01)	33.67*** (2.81)	38.74*** (1.97)	33.88*** (2.75)	35.09*** (2.67)	34.97*** (3.94)	34.75*** (2.63)	34.43*** (3.87)
Scaled response	0.78	0.67	0.77	0.68	0.70	0.70	0.70	0.69
Observations	13 805	13 805	13 805	13 805	7 496	7 496	7 496	7 496
R^2	0.13	0.13	0.16	0.16	0.11	0.11	0.14	0.14
Panel B:								
Above 7.5k	25.80*** (1.76)	25.70*** (2.67)	27.26*** (1.67)	25.25*** (2.54)	25.63*** (2.45)	20.40*** (3.88)	25.38*** (2.34)	21.82*** (3.70)
Scaled response	0.65	0.64	0.68	0.63	0.64	0.51	0.63	0.55
Observations	13 531	13 531	13 531	13 531	7 638	7 638	7 638	7 638
R^2	0.11	0.11	0.19	0.19	0.09	0.09	0.17	0.17
Polynomial	p_1	p_2	p_1	p_2	p_1	p_2	p_1	p_2
Controls	-	-	✓	✓	-	-	✓	✓
Bandwidth	2 000	2 000	2 000	2 000	1 000	1 000	1 000	1 000

Notes: This table shows robustness checks from estimating specification (1) when we net out professional taxes and union payments from our outcome variable, and restrict our sample to workers in highly unionised sectors (excluding the public sector). The *scaled response* is calculated as the parameter estimate divided by the notch size in the donation schedule (i.e. in panel A the parameter estimate is divided by 50 while in panel B the parameter estimate is divided by 40). p_s indicates that we fit a polynomial of order s on each side of the notch, while controls include sex, year and sector fixed effects. Robust standard errors clustered at the individual level in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 9:

RD estimates - only highly unionised sectors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A:								
Above 10k	34.81*** (1.29)	31.21*** (1.92)	35.52*** (1.27)	31.38*** (1.89)	32.72*** (1.82)	34.62*** (2.81)	33.38*** (1.80)	33.92*** (2.75)
Scaled response	0.70	0.62	0.71	0.63	0.65	0.69	0.67	0.68
Observations	35 100	35 100	35 100	35 100	18 322	18 322	18 322	18 322
R^2	0.22	0.22	0.25	0.25	0.13	0.13	0.16	0.16
Panel B:								
Above 7.5k	31.52*** (1.17)	34.75*** (1.70)	32.44*** (1.15)	35.10*** (1.67)	33.07*** (1.57)	26.87*** (2.36)	33.19*** (1.53)	27.23*** (2.31)
Scaled response	0.79	0.87	0.81	0.88	0.83	0.67	0.83	0.68
Observations	30 407	30 407	30 407	30 407	17 574	17 574	17 574	17 574
R^2	0.29	0.29	0.33	0.33	0.19	0.19	0.23	0.24
Polynomial	p_1	p_2	p_1	p_2	p_1	p_2	p_1	p_2
Controls	-	-	✓	✓	-	-	✓	✓
Bandwidth	2 000	2 000	2 000	2 000	1 000	1 000	1 000	1 000

Notes: This table shows robustness checks from estimating specification (1) when we restrict our sample to those working in highly unionised sectors (including the public sector). The *scaled response* is calculated as the parameter estimate divided by the notch size in the donation schedule (i.e. in panel A the parameter estimate is divided by 50 while in panel B the parameter estimate is divided by 40). p_s indicates that we fit a polynomial of order s on each side of the notch, while controls include sex, year and sector fixed effects. Robust standard errors clustered at the individual level in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 10:

RD estimates - only sectors with low unionisation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A:								
Above 10k	33.08*** (1.54)	33.04*** (2.29)	33.55*** (1.49)	33.60*** (2.22)	33.88*** (2.17)	32.84*** (3.34)	34.57*** (2.11)	33.81*** (3.23)
Scaled response	0.66	0.66	0.67	0.67	0.68	0.66	0.69	0.68
Observations	33 700	33 700	33 597	33 597	17 134	17 134	17 082	17 082
R^2	0.19	0.19	0.25	0.25	0.11	0.11	0.18	0.18
Panel B:								
Above 7.5k	28.99*** (1.12)	27.89*** (1.61)	28.76*** (1.08)	27.60*** (1.55)	27.79*** (1.51)	24.08*** (2.25)	27.51*** (1.45)	25.19*** (2.15)
Scaled response	0.72	0.70	0.72	0.69	0.69	0.60	0.69	0.63
Observations	40 831	40 831	40 688	40 688	22 635	22 635	22 549	22 549
R^2	0.29	0.29	0.35	0.35	0.18	0.18	0.24	0.24
Polynomial	p_1	p_2	p_1	p_2	p_1	p_2	p_1	p_2
Controls	-	-	✓	✓	-	-	✓	✓
Bandwidth	2 000	2 000	2 000	2 000	1 000	1 000	1 000	1 000

Notes: This table shows robustness checks from estimating specification (1) when we restrict our sample to those not working in highly unionised sectors. The *scaled response* is calculated as the parameter estimate divided by the notch size in the donation schedule (i.e. in panel A the parameter estimate is divided by 50 while in panel B the parameter estimate is divided by 40). p_s indicates that we fit a polynomial of order s on each side of the notch, while controls include sex, year and sector fixed effects. Robust standard errors clustered at the individual level in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 11:

**Elasticity of donations wrt price - bunchers vs non-bunchers
(first-pound price IV)**

	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	Non-B	Non-B	B	B
$\ln(1 - \tau^*)$	-0.774 *** (0.142)	-0.781 *** (0.156)	-1.374 *** (0.265)	-1.411 *** (0.295)	-0.539 *** (0.160)	-0.539 *** (0.167)
Observations	22 074	22 074	6 515	6 515	15 559	15 559
R^2	0.57	0.57	0.69	0.69	0.46	0.46
Individual FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Controls	-	✓	-	✓	-	✓

Notes: The sample includes all workers from the banking sectors in years 2003-2007. In all specifications we control for income. Additional controls include age squared and a dummy for changing employer. We drop people below the first tax bracket (i.e. people with no tax liability). We report robust standard errors clustered at the individual level in parenthesis, *** p<0.01, ** p<0.05, * p<0.1.

This table shows the results from separate regressions on split samples. If we instead run a pooled regression with full interactions, the coefficient on the interaction between the price-instrument and the indicator for being a buncher is 0.872 with a standard error of 0.339 (using the specification with controls). Hence, the difference in elasticity between bunchers and non-bunchers is significant at the 5 percent level.

Table 12:

Elasticity of donations wrt price - bunchers vs non-bunchers

(Sample: All highly unionised sectors)

	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	Non-B	Non-B	B	B
$\Delta \ln(1 - \tau^*)$	-0.345 *** (0.126)	-0.363 *** (0.126)	-0.630 *** (0.199)	-0.669 *** (0.194)	-0.241 (0.162)	-0.232 (0.162)
Observations	22 644	22 644	7 708	7 708	14 936	14 936
R^2	0.26	0.27	0.32	0.32	0.23	0.24
Individual FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Controls	-	✓	-	✓	-	✓

Notes: The sample includes workers from all highly unionised sectors in years 2003-2007. In all specifications we control for income. Additional controls include age squared and a dummy for changing employer. We drop people below the first tax bracket (i.e. people with no tax liability). We report robust standard errors clustered at the individual level in parenthesis, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

This table shows the results from separate regressions on split samples. If we instead run a pooled regression with full interactions, the coefficient on the interaction between the price-instrument and the indicator for being a buncher is 0.437 with a standard error of 0.253 (using specification with controls). Hence, the difference in elasticity between bunchers and non-bunchers is marginally significant at the 10 percent level.

Table 13:

Elasticity of donations wrt price - bunchers vs non-bunchers

(Sample: All)

	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	Non-B	Non-B	B	B
$\Delta \ln(1 - \tau^*)$	-0.411 *** (0.068)	-0.394 *** (0.068)	-0.593 *** (0.101)	-0.580 *** (0.099)	-0.306 *** (0.094)	-0.282 *** (0.094)
Observations	47 574	47 574	20 729	20 729	26 845	26 845
R^2	0.27	0.27	0.30	0.30	0.25	0.25
Individual FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Controls	-	✓	-	✓	-	✓

Notes: The sample includes workers from all sectors in years 2003-2007. In all specifications we control for income. Additional controls include age squared and a dummy for changing employer. We drop people below the first tax bracket (i.e. people with no tax liability). We report robust standard errors clustered at the individual level in parenthesis, *** p<0.01, ** p<0.05, * p<0.1.

This table shows the results from separate regressions on split samples. If we instead run a pooled regression with full interactions, the coefficient on the interaction between the price-instrument and the indicator for being a buncher is 0.297 with a standard error of 0.137 (using the specification with controls). Hence, the difference in elasticity between bunchers and non-bunchers is significant at the 5 percent level.

Table 14:

Buncher stickiness - Linear probability model

	(1) All	(2) All	(3) $\Delta(1 - \tau^*) \neq 0$	(4) $\Delta(1 - \tau^*) \neq 0$
$\mathbb{1}_{\text{(buncher)}}$	-0.403 *** (0.007)	-0.391 *** (0.007)	-0.474 *** (0.013)	-0.458 *** (0.013)
Constant	0.785 *** (0.004)	0.813 *** (0.011)	0.913 *** (0.004)	0.895 *** (0.015)
Observations	20 621	20 621	6 567	6 567
R^2	0.16	0.20	0.27	0.28
Year FE	-	✓	-	✓
Controls	-	✓	-	✓

Notes: The sample includes workers from the banking sector in years 2003-2007. Controls include age squared, salary and a dummy for changing employer. We drop people below the first tax bracket (i.e. people with no tax liability). In columns (3) and (4) we restrict only to workers experiencing a change in the first-pound price. We report robust standard errors clustered at the individual level in parenthesis, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

13 Appendix D - Figures

Figure 10:

Information on tax returns regarding thresholds

(a) 2002

C MISCELLANEOUS DEDUCTIONS (attach necessary certificates)			
	1 DESCRIPTION	2 AMOUNT £	
1	Professional licence / Tax		4 Donations to approved Charities
2	Contributions to trade unions		5 Deposits under the specific savings scheme of the Housing Finance Corporation
3	Subscriptions		6 Any other deduction
TOTAL			

(b) 2003

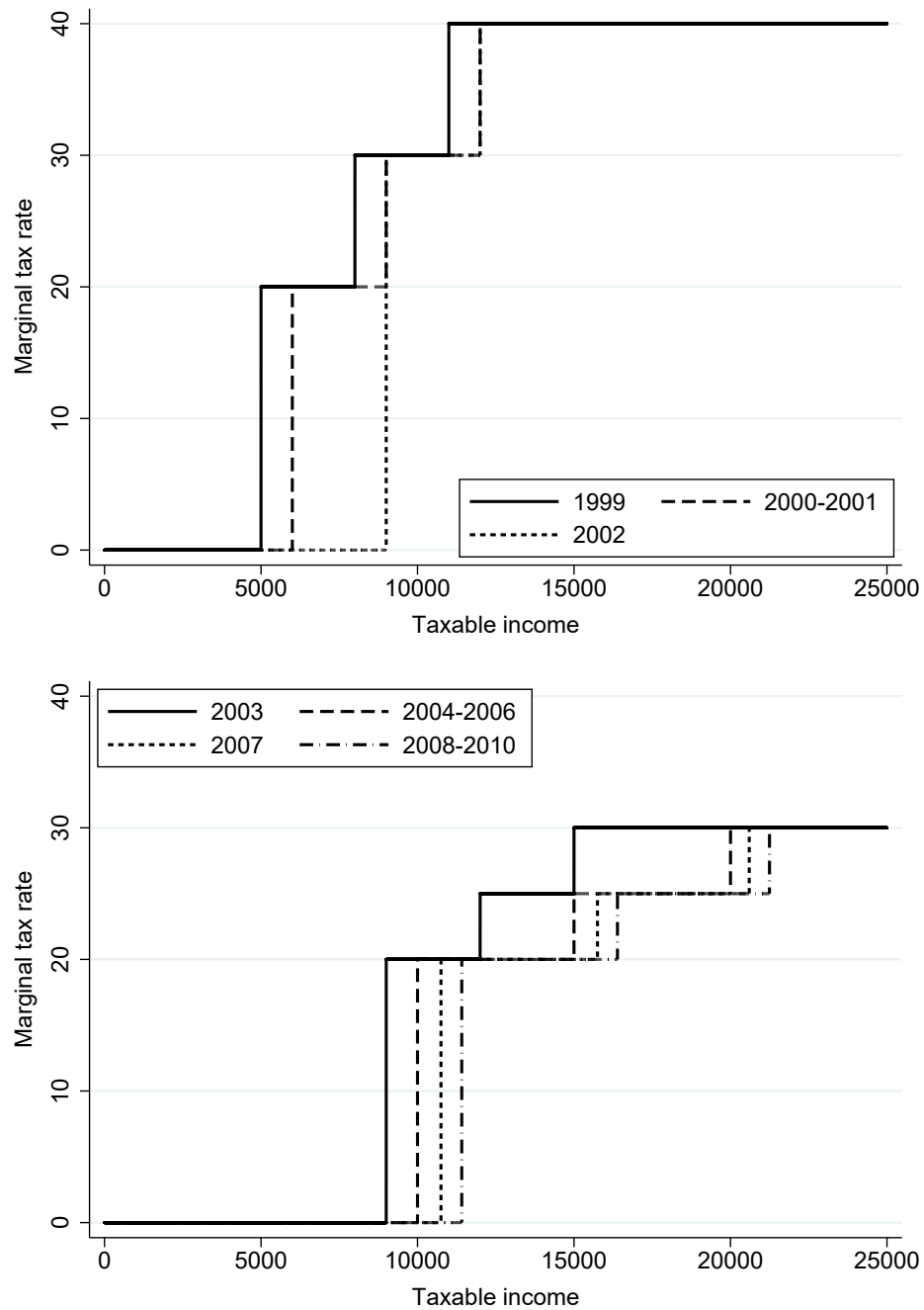
B	MISCELLANEOUS DEDUCTIONS (For donations over £150 please attach certificates / receipts. For donations of a lesser amount you should keep the certificates / receipts to submit when requested).	
	1 DESCRIPTION	2 AMOUNT £
1	TRADE UNION CONTRIBUTIONS	
2	PROFESSIONAL SUBSCRIPTIONS	
3	DONATIONS TO APPROVED CHARITABLE ORGANISATIONS	
4	ANY OTHER DEDUCTION	
5	TOTAL	

(c) 2008

PART 5 – DEDUCTIONS / ALLOWANCES		
A	MISCELLANEOUS DEDUCTIONS (Attach certificates / receipts only for donations over €300. For donations of a lesser amount you should keep the certificates / receipts to submit when requested).	
	1 DESCRIPTION	2 AMOUNT
1	TRADE UNION CONTRIBUTIONS	
2	PROFESSIONAL SUBSCRIPTIONS	
3	DONATIONS TO APPROVED CHARITABLE ORGANISATIONS	
4	ANY OTHER DEDUCTION	
5	TOTAL	

Notes: This set of figures shows the information provided on the tax return regarding filing thresholds for different years.

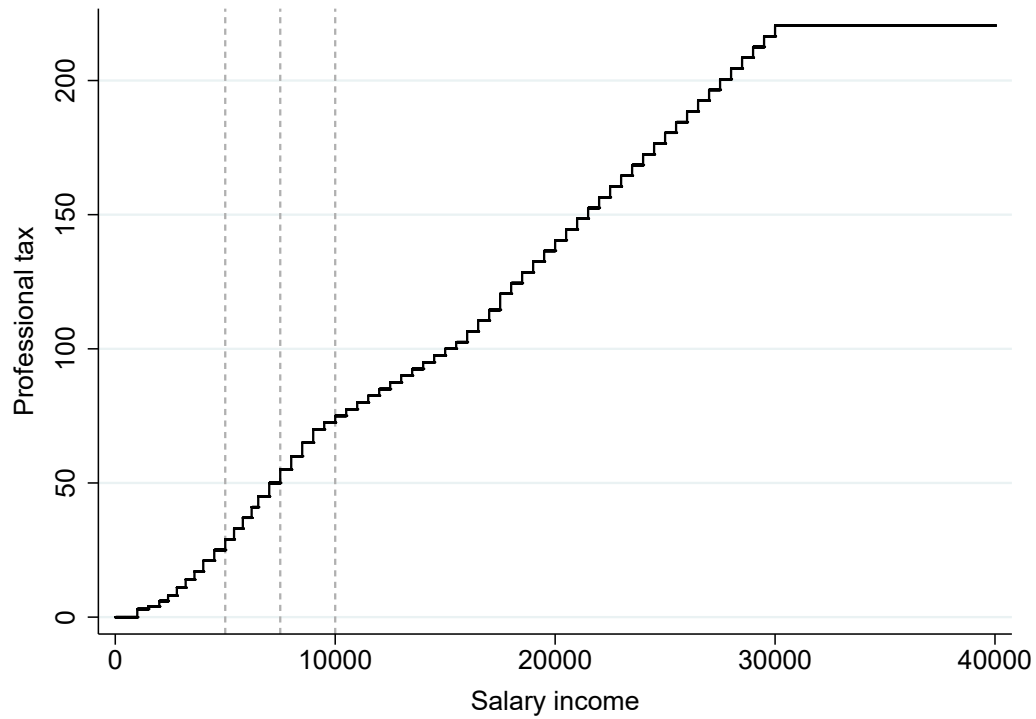
Figure 11:
Schedule of marginal tax rates (years 1999-2010)



Notes: The figure shows the schedule of marginal tax rates in place in the Republic Of Cyprus in the years 1999-2010.

Figure 12:

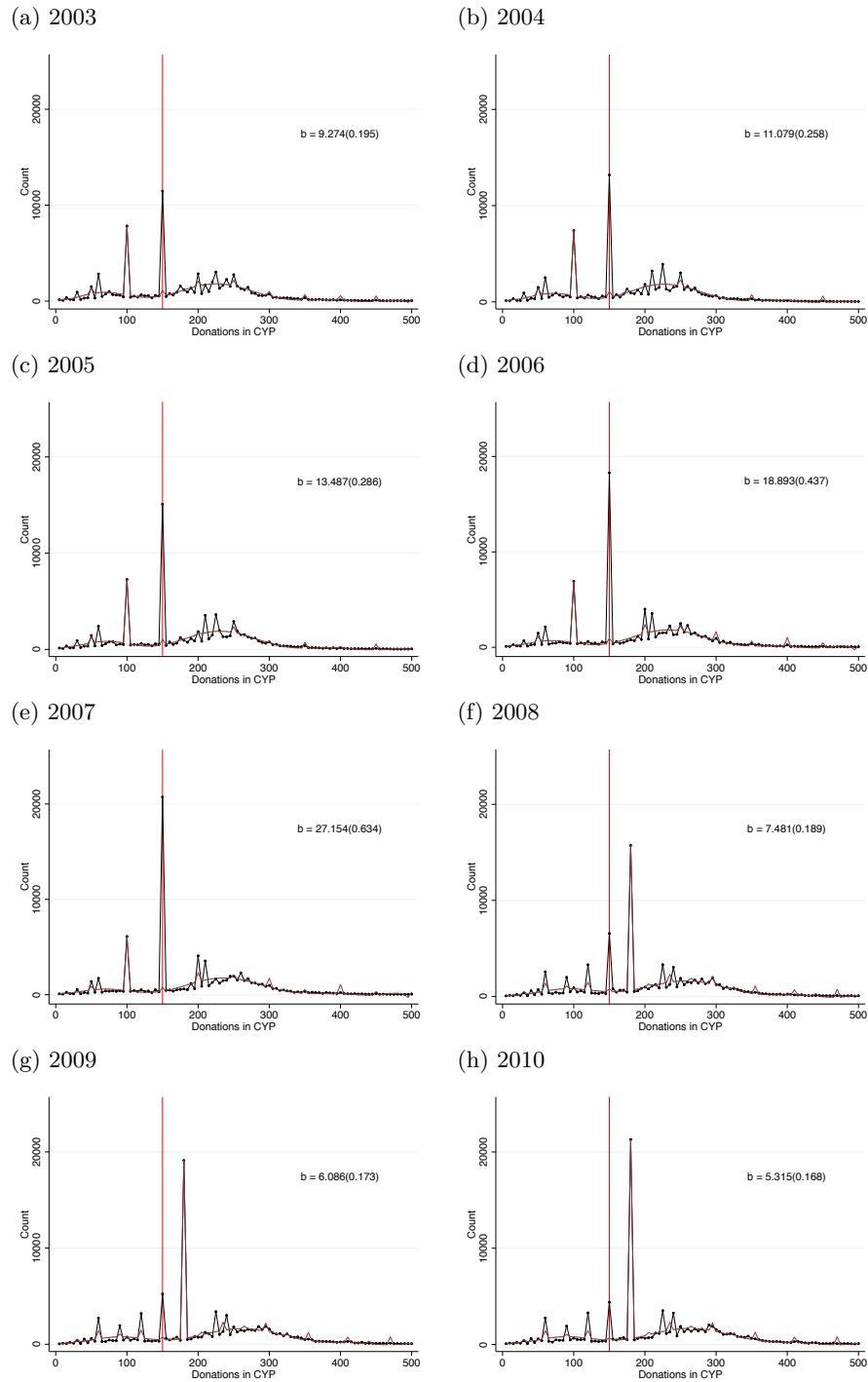
Schedule of professional taxes (before 2003)



Notes: The figure shows the schedule of professional taxes in place in the years prior to 2003. The tax was dependent on salary income, with incremental steps until CYP 30,001. Vertical lines indicate the notches in the schedule for deductible donations without the provision of receipts.

Figure 13:

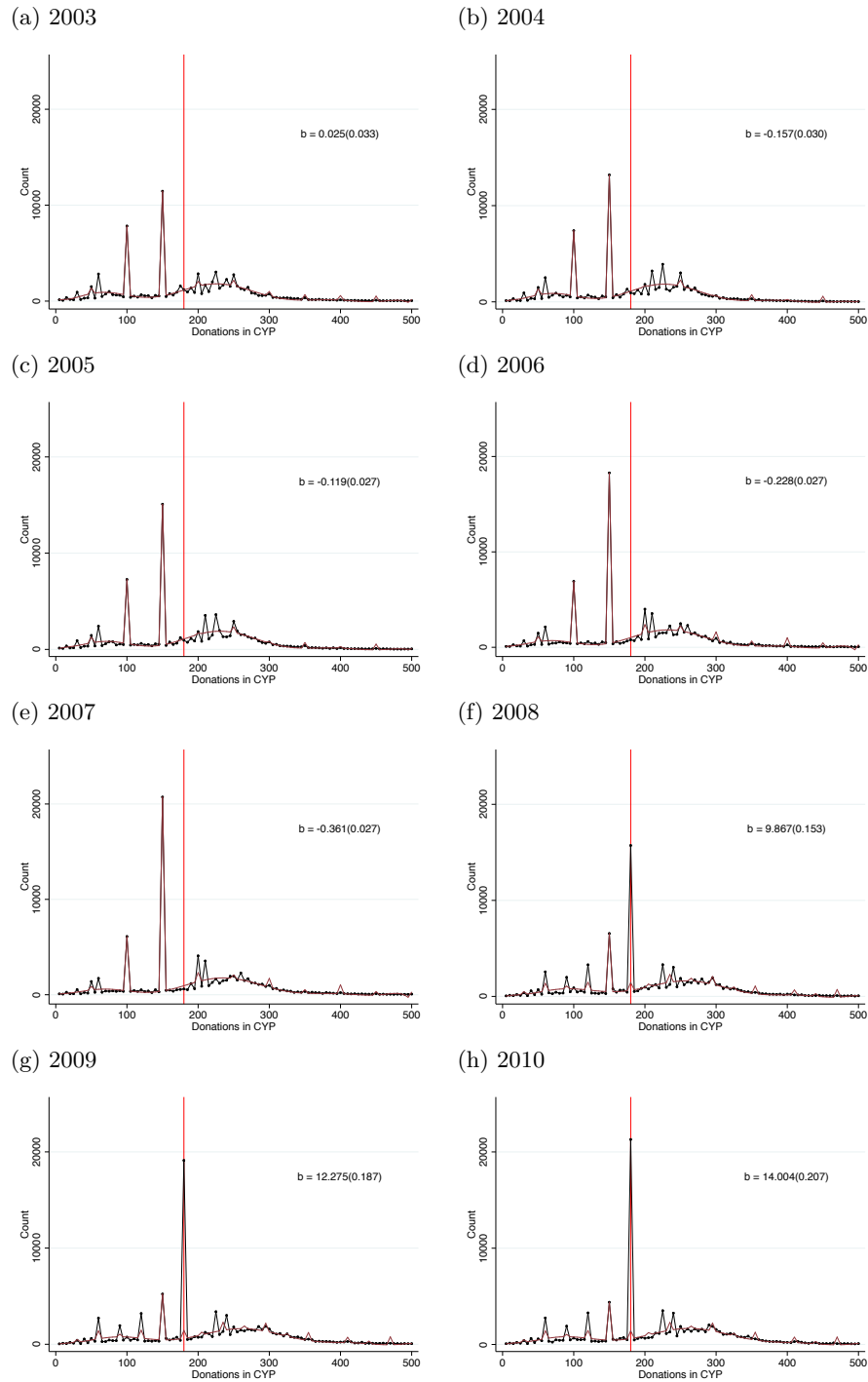
Bunching at CYP 150 with estimated counterfactual, main sample



Notes: This figure shows the bunching dynamics of positive donations among salary earners between 2003-2010. It plots the yearly empirical distribution in bins of width CYP 5, together with the estimated counterfactual. Each sub-figure reports the normalised excess bunching mass b around the CYP 150 threshold. Bootstrapped standard errors are in parentheses.

Figure 14:

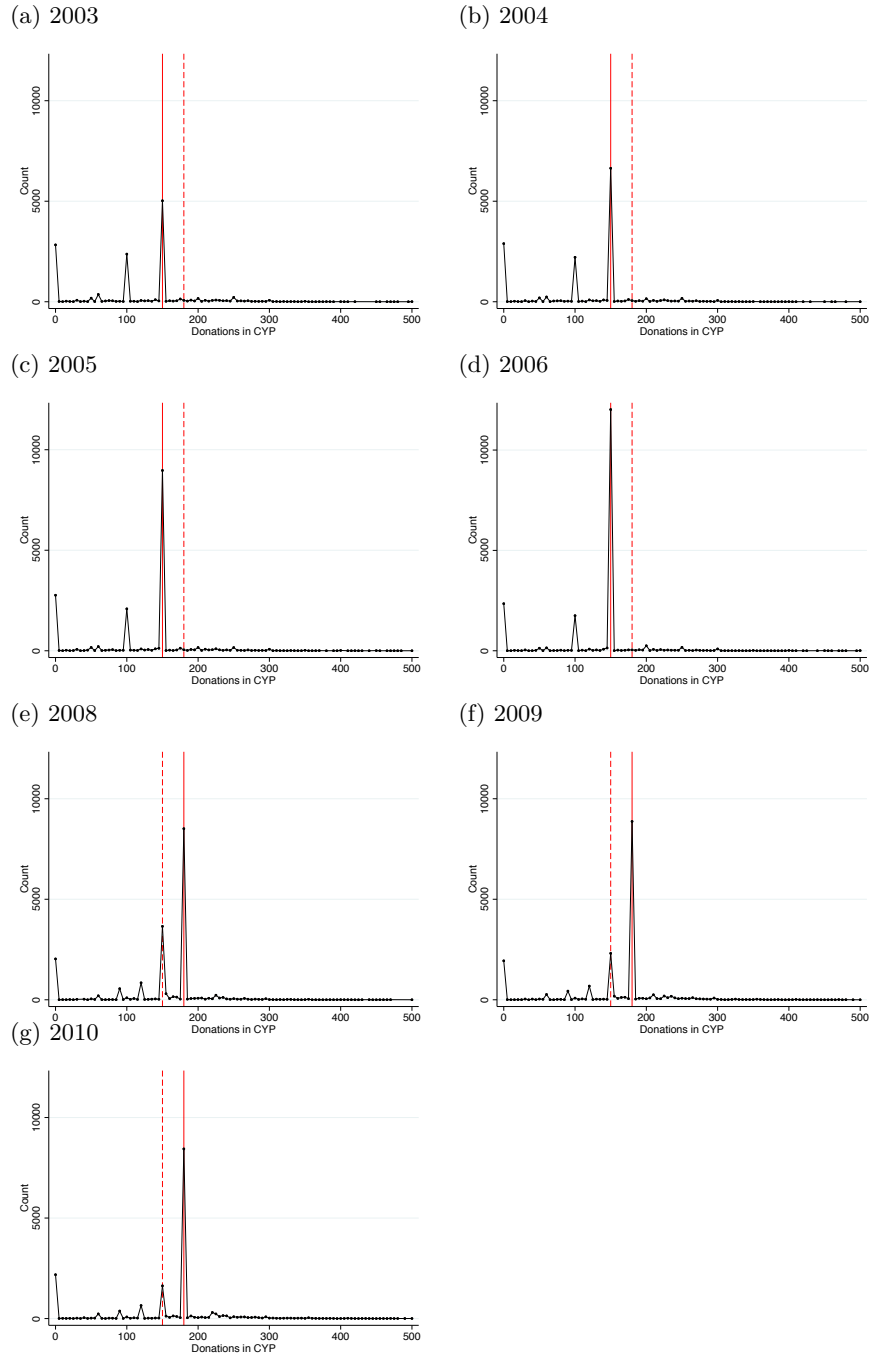
Bunching at CYP 175 with estimated counterfactual, main sample



Notes: This figure shows the bunching dynamics of positive donations among salary earners between 2003-2010. It plots the yearly empirical distribution in bins of width CYP 5, together with the estimated counterfactual. Each sub-figure reports the normalised excess bunching mass b around the CYP 175 threshold. Bootstrapped standard errors are in parentheses.

Figure 15:

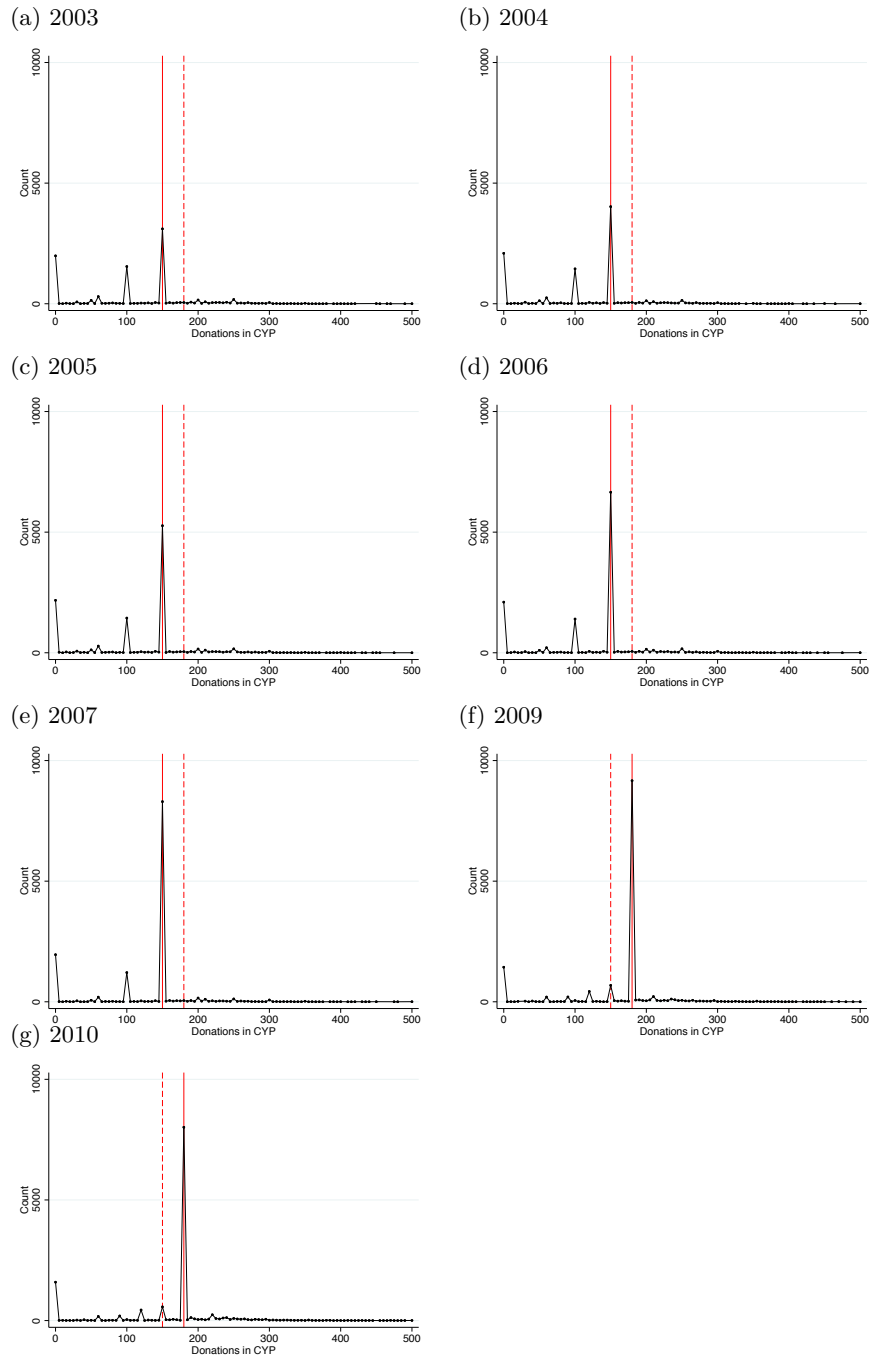
Donations between 2003-2010 of those bunching at CYP 150 in 2007



Notes: This figure shows the empirical distribution of donations before and after 2007, for the sample of salary earners who bunched at CYP 150 in 2007. Vertical solid lines mark the relevant threshold that is in place in a given year (CYP 150 during 2003-2006 and CYP 175 during 2008-2010), while dashed lines mark the other threshold that has either been eliminated, or has not been yet introduced.

Figure 16:

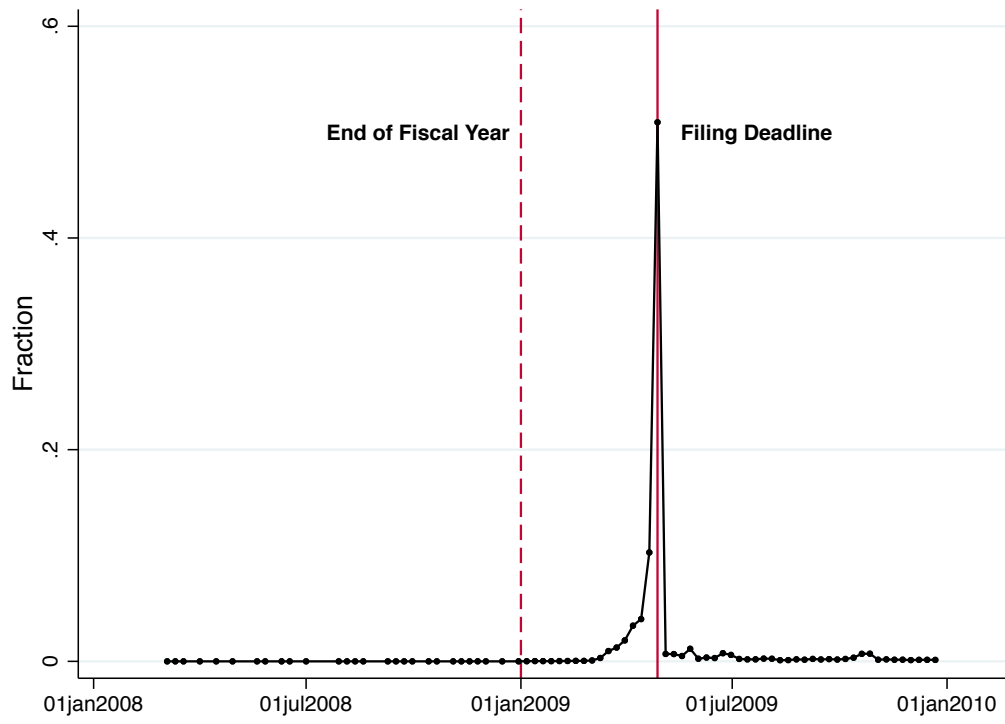
Donations between 2003-2010 of those bunching at CYP 175 in 2008



Notes: This figure shows the empirical distribution of donations before and after 2008, for the sample of salary earners who bunched at CYP 175 in 2008. Vertical solid lines mark the relevant threshold that is in place in a given year (CYP 150 during 2003-2007 and CYP 175 during 2009-2010), while dashed lines mark the other threshold that has either been eliminated, or has not been yet introduced.

Figure 17:

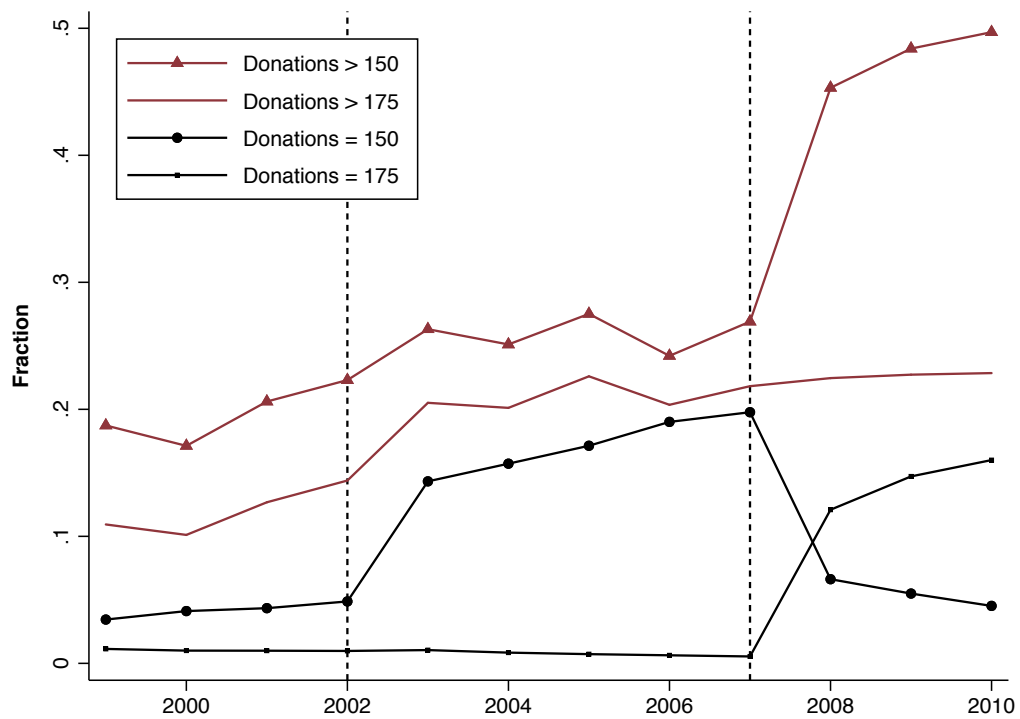
When did people file their taxes for 2008?



Notes: The figure shows, in weekly bins, the fraction of people filing their taxes for the fiscal year 2008. Vertical dashed and solid lines mark the end of the the fiscal year (31 December 2008) and filing deadline (30 April 2009) respectively.

Figure 18:

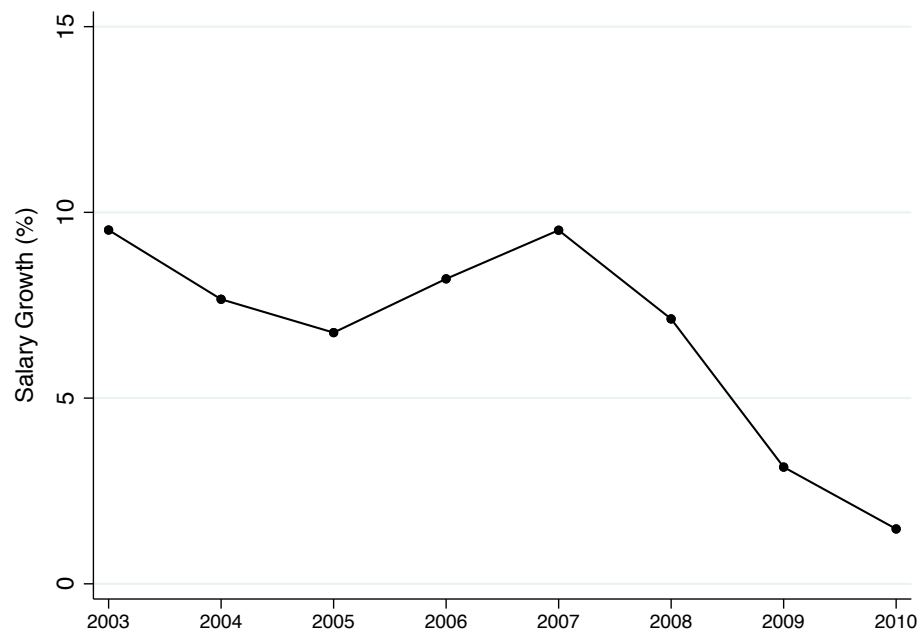
Fraction reporting specific amounts of donations over time



Notes: This figure shows the fraction of individuals reporting each of the following amounts of donations over time: over 175, over 150, 150 and 175 (all in CYP).

Figure 19:

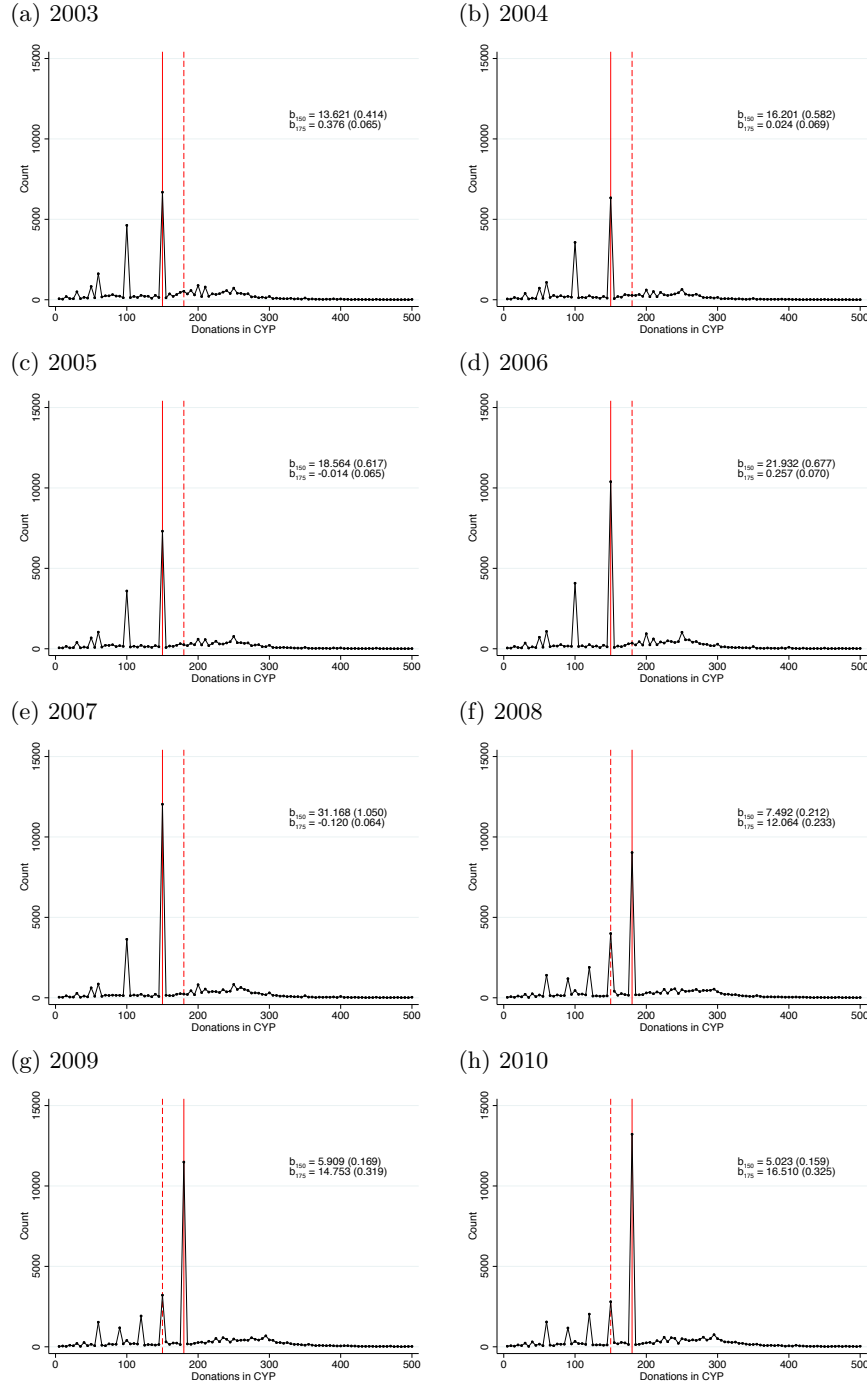
Salary growth rates of 2007 bunchers



Notes: This figure shows the yearly salary growth rate between 2003-2010 of salary earners bunching at the CYP 150 threshold in 2007.

Figure 20:

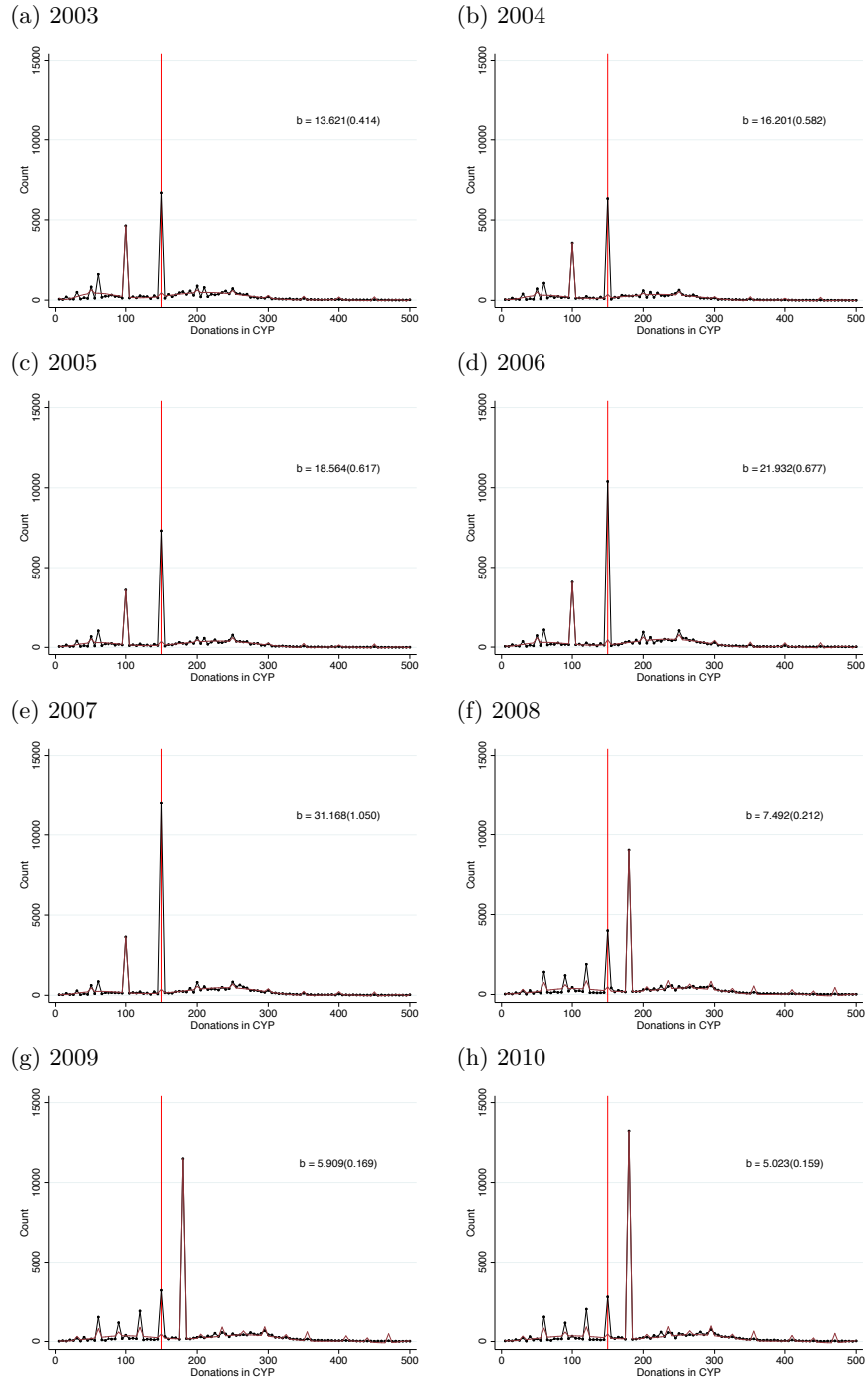
Bunching around reporting thresholds, excluding highly unionised sectors



Notes: This figure shows the bunching dynamics of positive donations among salary earners between 2003-2010 by plotting the yearly empirical distributions in bins of width CYP 5. The sample is restricted to those not in highly unionised sectors and drops those whose sector is not observed. Each sub-figure reports the normalised excess bunching mass b around the CYP 150 and CYP 175 thresholds, with bootstrapped standard errors in parentheses. Vertical solid lines mark the threshold that is in place in a given year, while dashed lines mark the other threshold that has either been eliminated, or has not been yet introduced.

Figure 21:

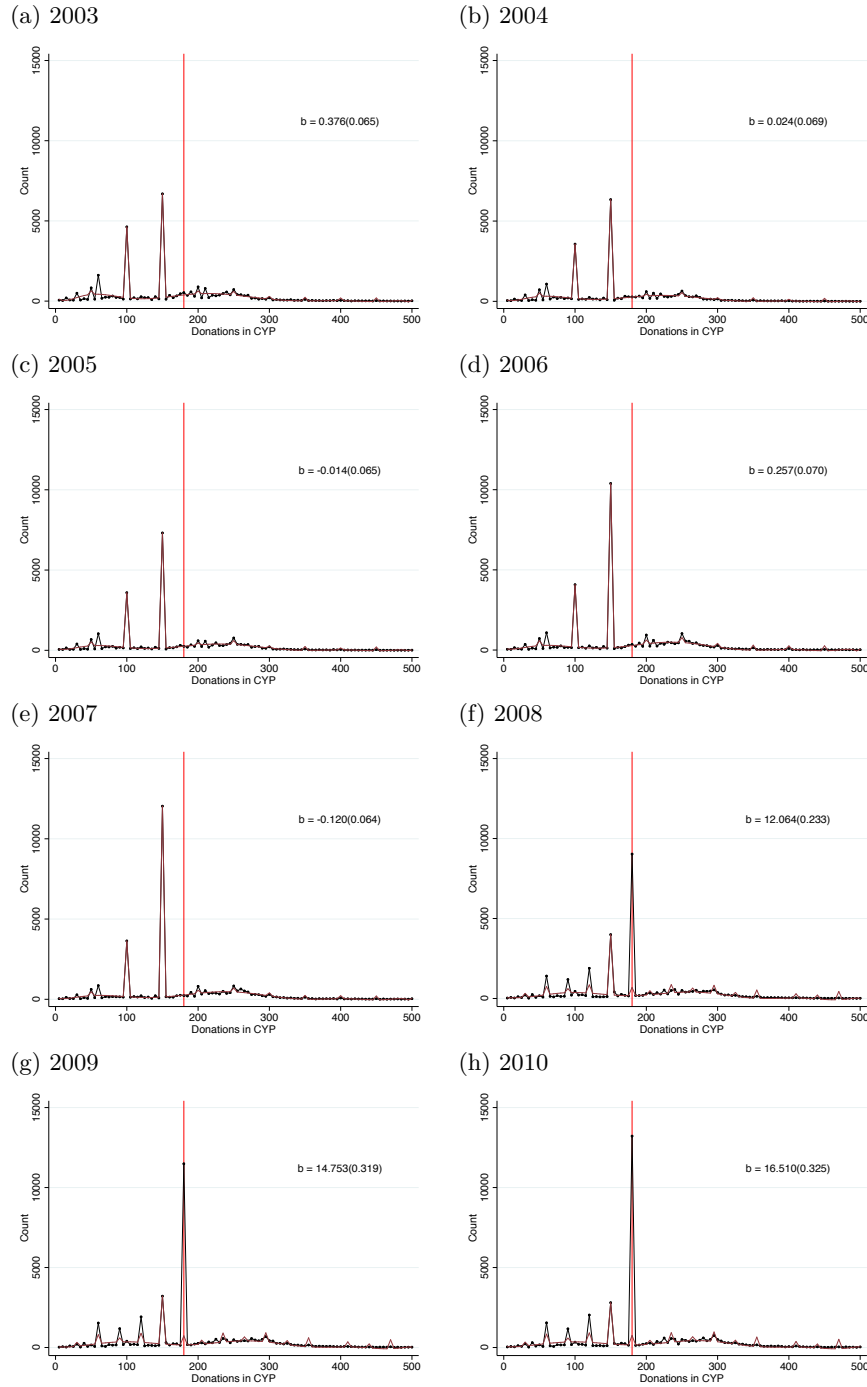
Bunching at CYP 150 with counterfactual, excl. highly unionised sectors



Notes: This figure shows the bunching dynamics of positive donations among salary earners between 2003-2010, restricting the sample to those not in highly unionised sectors and dropping those whose sector is not observed. It plots the yearly empirical distribution in bins of width CYP 5, together with the estimated counterfactual. Each sub-figure reports the normalised excess bunching mass b around the CYP 150 threshold. Bootstrapped standard errors are in parentheses.

Figure 22:

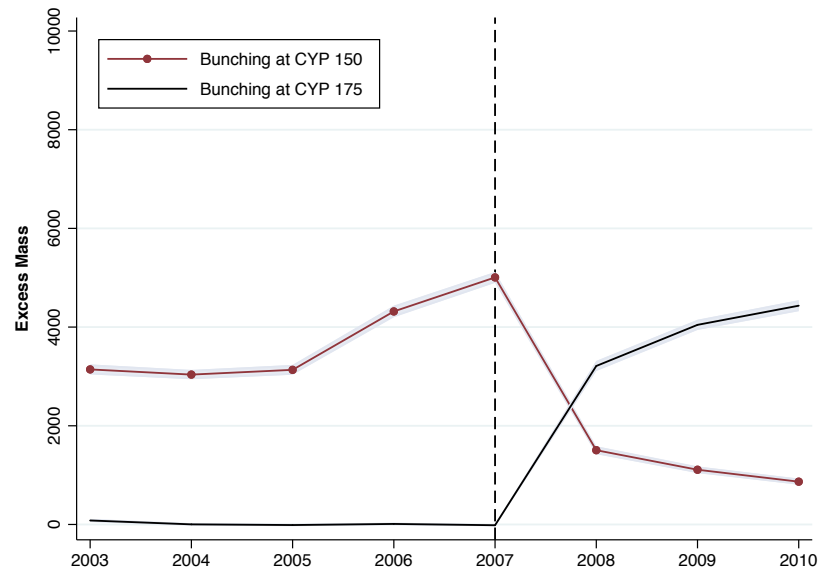
Bunching at CYP 175 with counterfactual, excl. highly unionised sectors



Notes: This figure shows the bunching dynamics of positive donations among salary earners between 2003-2010, restricting the sample to those not in highly unionised sectors and dropping those whose sector is not observed. It plots the yearly empirical distribution in bins of width CYP 5, together with the estimated counterfactual. Each sub-figure reports the normalised excess bunching mass b around the CYP 175 threshold. Bootstrapped standard errors are in parentheses.

Figure 23:

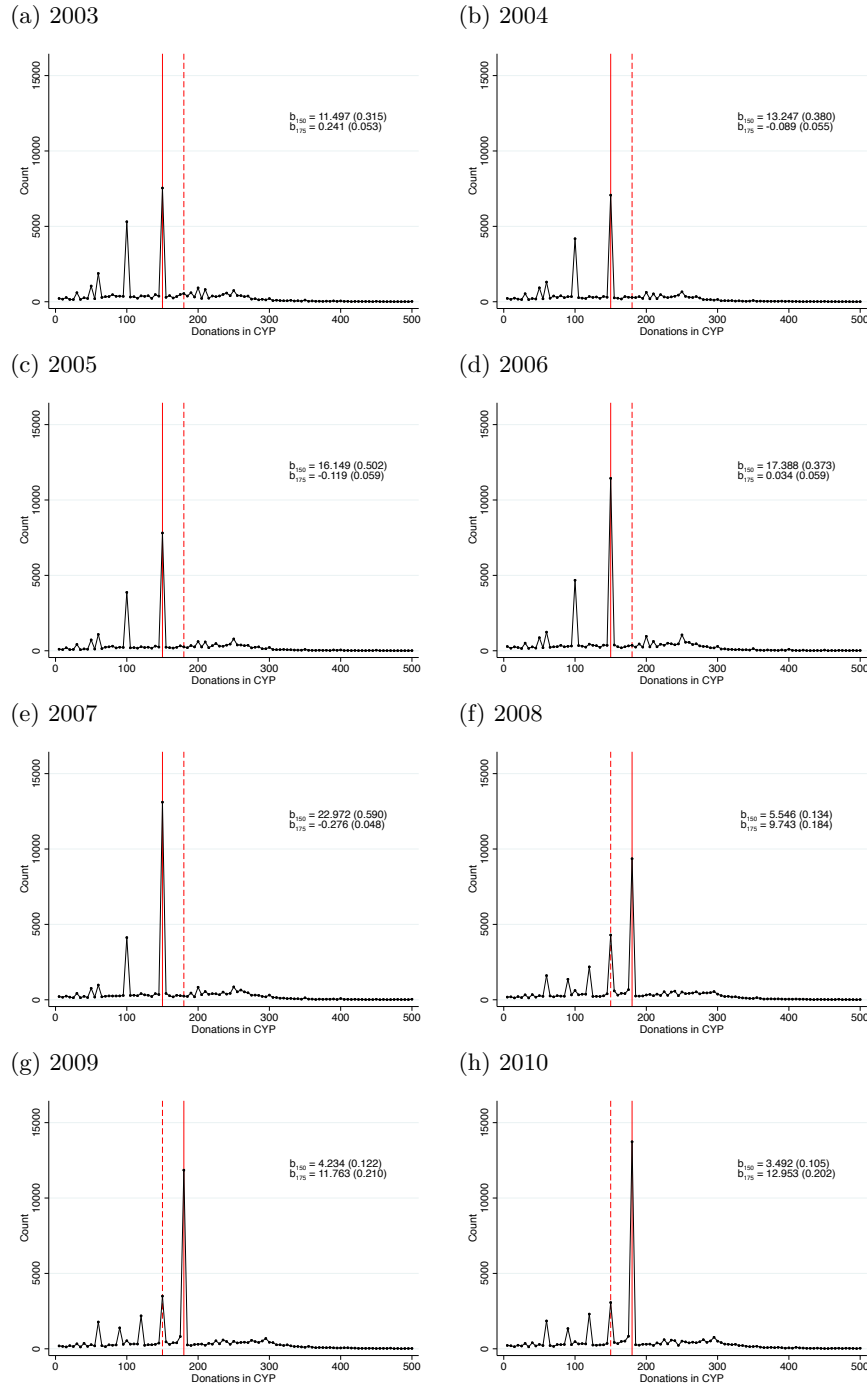
Bunching estimates over time, excluding highly unionised sectors



Notes: This figure shows the estimates of the excess mass around both the CYP 150 and 175 thresholds between 2003-2010, restricting the sample to those not in highly unionised sectors. The shaded areas demarcate 95% confidence intervals.

Figure 24:

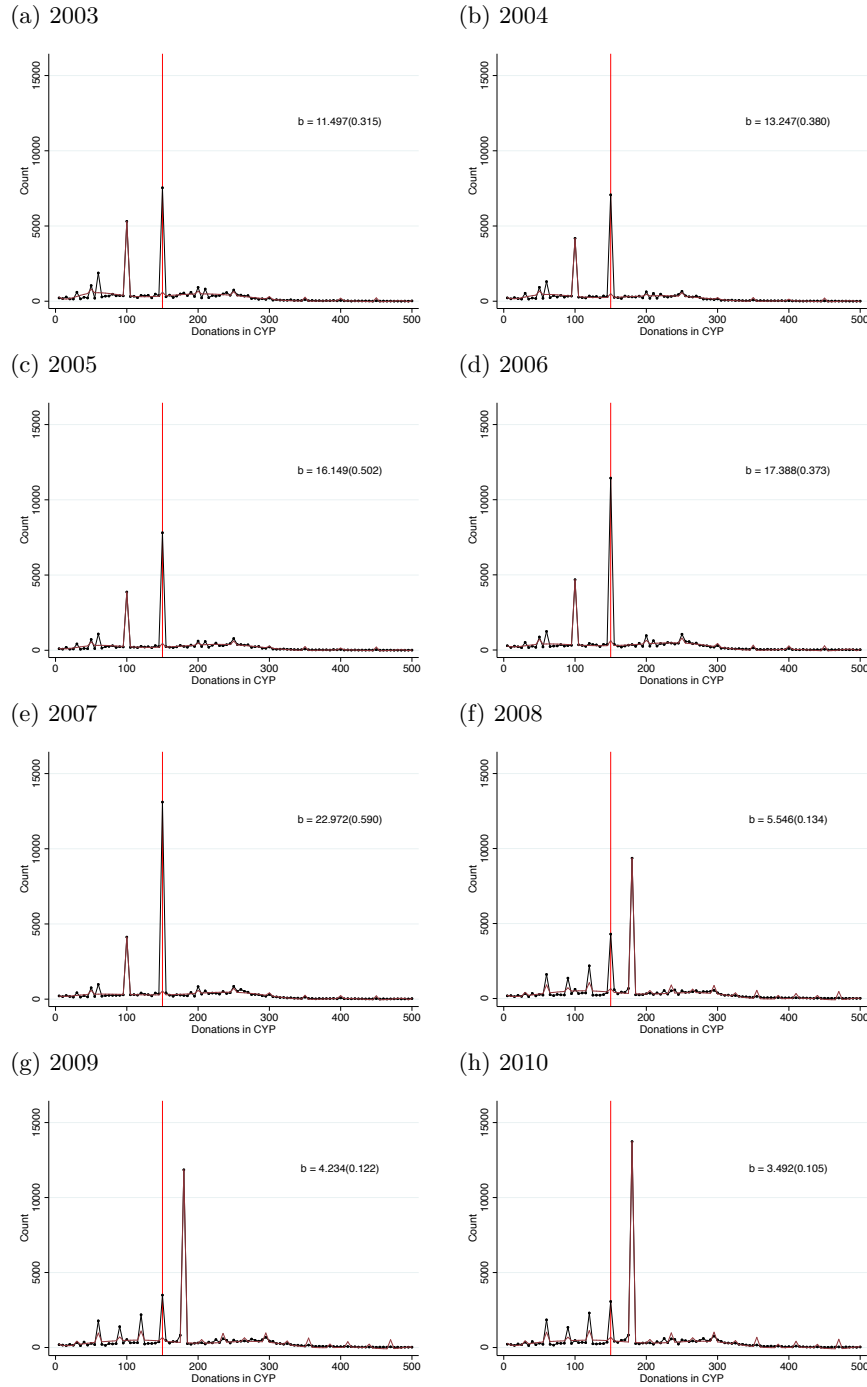
Bunching of donations, removing union fees



Notes: This figure shows the bunching dynamics of positive donations among salary earners between 2003-2010, by plotting the yearly empirical distributions in bins of width CYP 5. The sample is restricted to only those whose sector can be observed. For those in highly unionised sectors, the union fees have been removed. Each sub-figure reports the normalised excess bunching mass b around the CYP 150 and CYP 175 thresholds, with bootstrapped standard errors in parentheses. Vertical solid lines mark the threshold that is in place in a given year, while dashed lines mark the other threshold that has either been eliminated, or has not been yet introduced.

Figure 25:

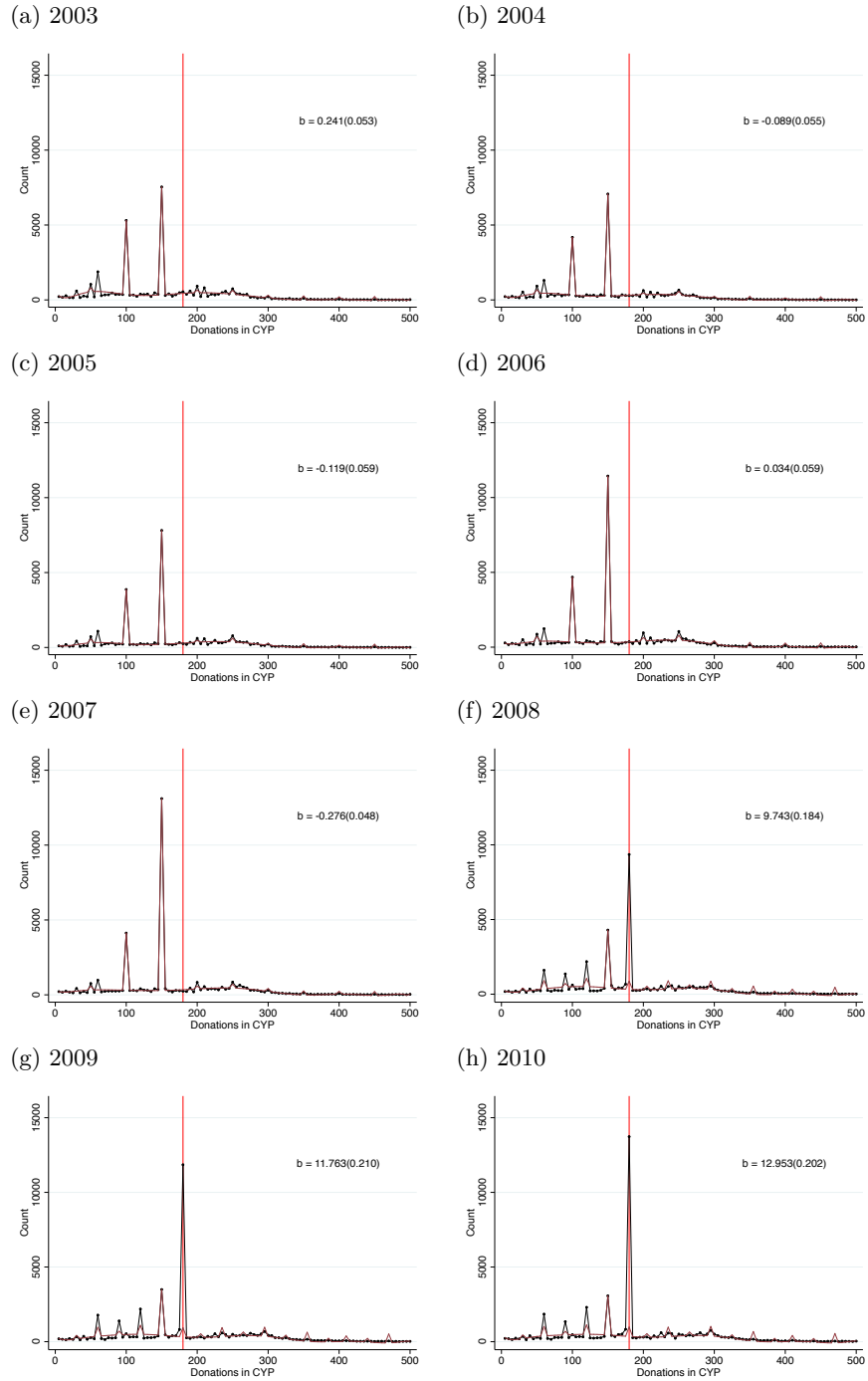
Bunching at CYP 150 with counterfactual, removing union fees



Notes: This figure shows the bunching dynamics of positive donations among salary earners between 2003-2010, for the main sample but restricted to only those whose sector can be observed. For those in highly unionised sectors, the union fees have been removed. It plots the yearly empirical distribution in bins of width CYP 5, together with the estimated counterfactual. Each sub-figure reports the normalised excess bunching mass b around the CYP 150 threshold. Bootstrapped standard errors are in parentheses.

Figure 26:

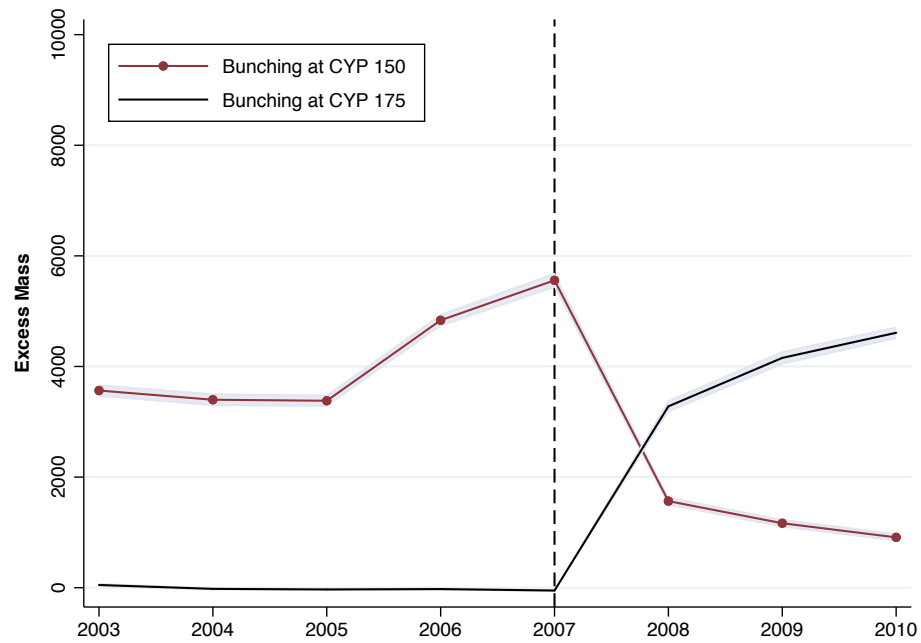
Bunching at CYP 175 with counterfactual, removing union fees



Notes: This figure shows the bunching dynamics of positive donations among salary earners between 2003-2010, for the main sample but restricted to only those whose sector can be observed. For those in highly unionised sectors, the union fees have been removed. It plots the yearly empirical distribution in bins of width CYP 5, together with the estimated counterfactual. Each sub-figure reports the normalised excess bunching mass b around the CYP 175 threshold. Bootstrapped standard errors are in parentheses.

Figure 27:

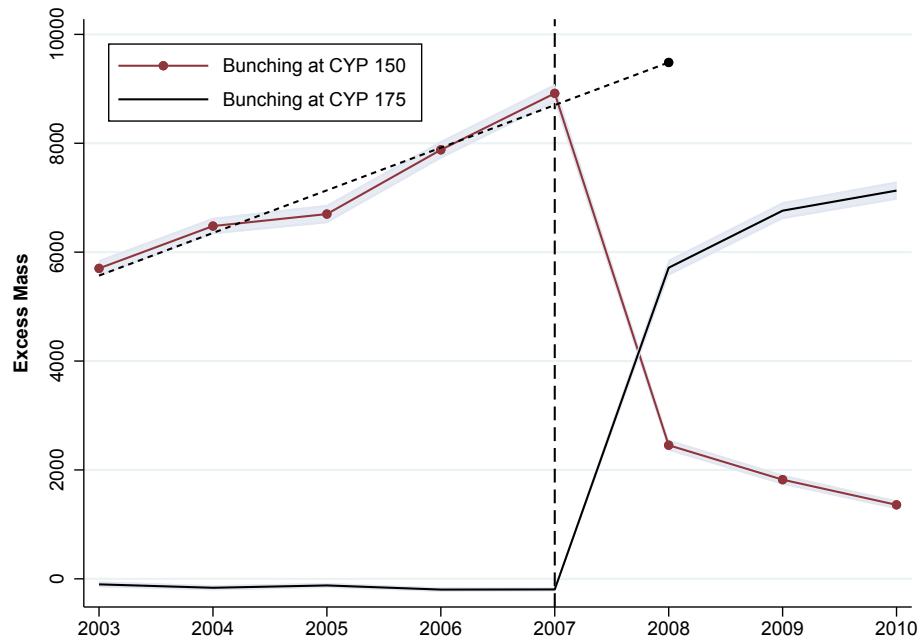
Bunching estimates over time, removing union fees



Notes: This figure shows the estimates of the excess mass around both the CYP 150 and 175 thresholds between 2003-2010. The shaded areas demarcate 95% confidence intervals. The sample is restricted to salary earners whose sector is observed (but excludes the public sector), and the outcome variable has been adjusted for union fees among those in highly unionised sectors.

Figure 28:

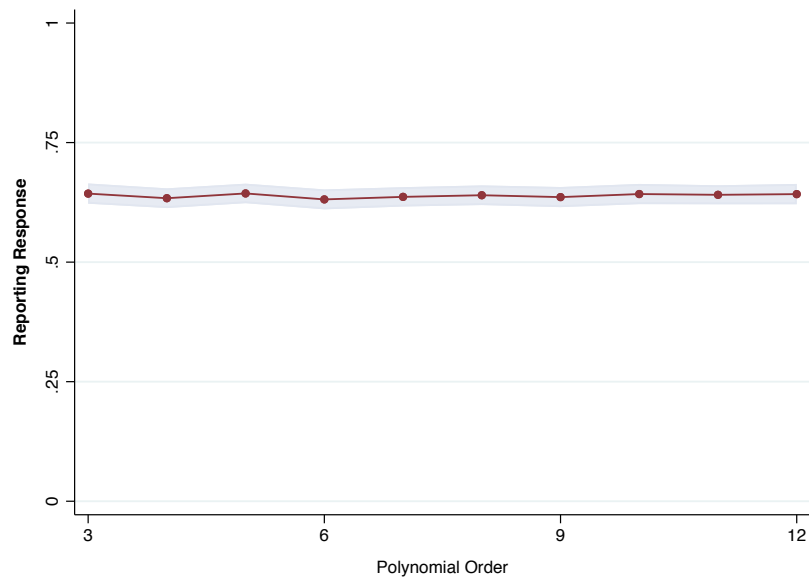
Bunching (excess mass) estimates over time



Notes: This figure shows for the balanced panel of our main sample of salary earners, the estimates of the excess mass around both the CYP 150 and 175 thresholds, between 2003-2010. The shaded areas demarcate 95% confidence intervals. The figure also shows the linear trend in bunching at CYP 150 estimated between 2003 and 2007. We use this linear trend to predict what the bunching at CYP150 would have been in 2008 absent the change in reporting threshold.

Figure 29:

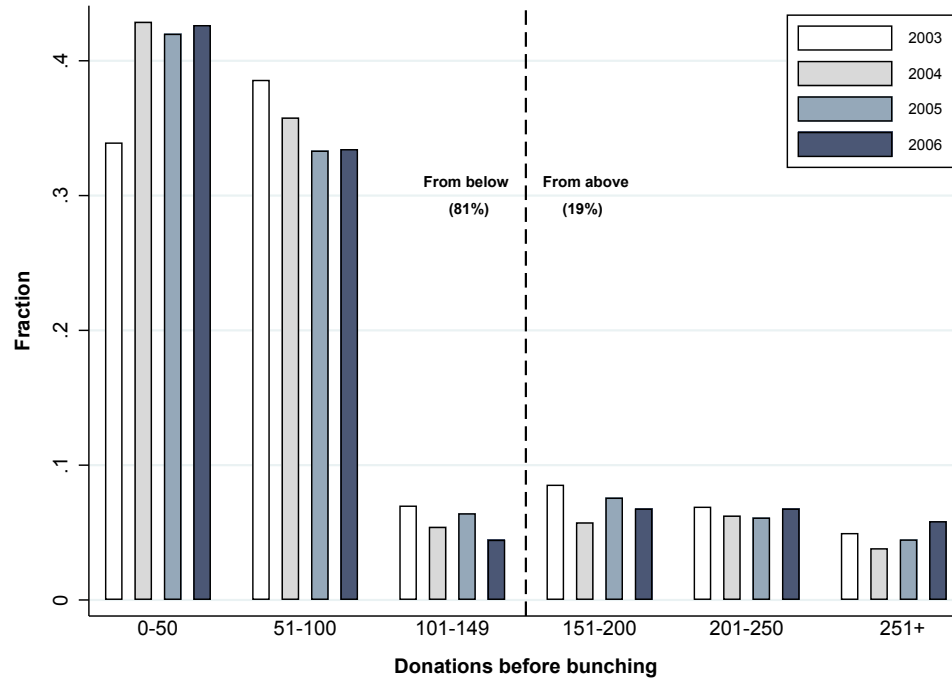
Robustness of reporting response to polynomial order of counterfactual fit



Notes: This figure plots estimates of (the lower bound of) the reporting response, L_R , for different values of the order of polynomial used to estimate the counterfactual density in our bunching analysis. The shaded areas demarcate 95% confidence intervals.

Figure 30:

Tracking bunchers



Notes: The figure uses the sample of salary-earners in non-unionised sectors. From this sample we then isolate taxpayers who move to the CYP 150 threshold between 2003 and 2007. The figure illustrates the donations of these taxpayers in the year before they start bunching at the CYP 150 threshold, divided into separate bins. Each colour represents a different year, i.e. 2003 shows the distribution of donations in 2003 for those taxpayers who start bunching in 2004. The vertical dashed line represents the placement of the reporting threshold at CYP 150. The percentage noted as coming from below (above) is calculated using all years as the fraction of taxpayers moving to the threshold between 2004 and 2007 who moved from a position below (above) the threshold.