The Role of Firms in Workers' Earnings Responses to Taxes: Evidence From Pakistan*

Michael Carlos Best†
London School of Economics

May 2014

Abstract

This paper exploits employee-employer matched administrative tax data on firms and salaried workers in Pakistan to explore the underappreciated role of firms in determining how workers' taxable earnings respond to taxation. I present evidence on three ways in which firms affect workers' earnings responses. First, third-party reporting of salaries by employers makes underreporting taxable income more costly for workers and reduces evasion of the income tax. Second, firms' equilibrium salary-hours offers respond endogenously to the presence of adjustment costs in the labour market by tailoring offers to aggregate worker preferences. Third, workers learn about the tax schedule from firms' salary offers, making them more responsive to taxation both contemporaneously (by 130%) and in subsequent years (by 100%). However, while third-party reporting makes misreporting more costly, it does not eliminate it in a low tax-capacity setting: 19% of workers still underreport their salaries, leading to a loss of about 5% of tax revenue, and indicating high returns to investments in improving enforcement capacity. The large role played by firms in determining workers' earnings implies that firms need to play a central role in our analysis of income taxation in lower income countries.

*I would like to thank my supervisors, Henrik Kleven and Oriana Bandiera for their generous support during this project. I am grateful to Camille Landais, Gerard Padró i Miquel and Johannes Spinnewijn for numerous helpful discussions, to Miguel Almunia, Giuseppe Berlingieri, Steve Bond, Michael Devereux, Ulrich Doraszelski, Greg Fischer, Anders Jensen, Torsten Persson and Mazhar Waseem and numerous seminar participants for useful suggestions, and to Ali Arshad Hakeem, Samad Khurram, Jawad Abbasi and Ijaz Nabi for their help with the data and their enthusiasm. Financial support from the International Growth Centre, Pakistan Programme made this project possible. The plethora of remaining errors is mine alone.

†STICERD & Department of Economics. Houghton Street, London WC2A 2AE, UK. email: m.c.best@lse.ac.uk, web: http://personal.lse.ac.uk/bestm/research.htm

1 Introduction

The development of the capacity of the state to raise taxes is central to the process of economic development, but the public finance literature has been largely silent on the issue, either tending to assume that taxes can be perfectly and costlessly enforced or taking evasion and administrative costs as given (Besley & Persson, 2011, 2013). A recent literature has suggested that third-party reporting of tax bases to the tax authority, particularly by firms, is key to understanding the government's capacity to enforce taxes (Kleven *et al.*, 2009; Pomeranz, 2013). In the case of the personal income tax, third-party reporting of salaries and withholding of income tax by employers form the bedrock of the enforcement regime in modern tax administrations. Historically, the first successful modern income taxes in both the UK and the US featured withholding of income tax on the salaries of civil servants (Slemrod, 2008). Today, all 34 OECD countries require employers to report their employees' salaries and all except France require employers to withhold income tax on their employees, collecting over 75% of personal income tax revenues through withholding (OECD, 2013).

More generally, the role of firms in the study of taxation has been underplayed. Kopczuk & Slemrod (2006) appeal for firms to be central to models of taxation with imperfect enforcement, noting firms' key roll in collecting and enforcing taxes.[1] Firms may also play a broader role in determining the way that workers' reported taxable incomes respond to taxes. In the presence of adjustment costs in the labour market (such as costly search), workers are not simply paid their marginal product (obviating a role for firms). Instead, salary earnings are the outcome of a matching process to which firms are central (see, for example, Rogerson *et al.*, 2005 and Manning, 2011 for surveys, and Chetty *et al.*, 2011 for an application to taxation). Similarly, if workers face information frictions preventing them from perceiving the tax schedule they face accurately, firms' behaviour during the salary determination process can convey useful information on the tax schedule and this can influence how workers respond to taxation, both in their salary and non-salary earnings.

This paper provides evidence on both issues in the context of the taxation of salaried workers in Pakistan. I am able to leverage four key advantages of my data and setting in order to bring evasion under third-party reporting and withholding, and the role of firms in determining workers' earnings responses to taxation into sharp relief. First, I have benefited from being able to exploit large administrative datasets of tax records from the Federal Board of Revenue (FBR) on the universe of tax-registered firms and workers in Pakistan.[2] The large sample sizes on the universe of taxpayers (roughly a million records per year) and the existence of administrative identifiers permitting the linkage of individuals and firms across datasets and years allow me to overcome the problems of attrition, non-response, and measurement error that typically plague studies using survey data (Card *et al.*, 2010), especially in developing countries (De Mel *et al.*, 2009). Second, to my knowledge, this paper is the first to be able to directly study third party reporting of wages by firms and workers

[1] For example, in the United States, 82% of federal tax revenues are remitted by firms (Slemrod, 2008)

[2] By its very nature, informal economic activity is not captured in government records and so does not form a part of this study.

in an employer-employee matched dataset from a developing country.[3] Kumler *et al.* (2013) and Niehaus & Sukhtankar (2013) study similar issues of misreporting of wages, but are unable to match individual-level observations of salaries and so study discrepancies between the distributions of wages reported by firms and by workers.[4]

Third, the tax schedule for the salaried employees I study features a large number of tax brackets (between 17 and 20 during the period I study) and the kinks in firms' and workers' choice sets induced by the jumps in the marginal tax rate at the bracket thresholds provide multiple compelling sources of quasi-experimental variation in tax incentives. These discontinuities in marginal tax rates generate incentives for incomes to cluster around the bracket thresholds, allowing me to non-parametrically identify behavioural responses to marginal tax rates using a bunching approach (Saez, 2010; Chetty *et al.*, 2011; Kleven & Waseem, 2013). Fourth, in Pakistan, as in other low income countries, it is common for individuals to diversify their income sources. This means that my data contains a considerable number of observations on individuals with significant amounts of non-salary income. In combination with the large number of tax brackets, this provides a unique opportunity as worker-firm matches are likely to be responding to multiple kinks simultaneously, allowing me to disentangle responses in salary and non-salary incomes, and responses by firms and by workers.

To guide the empirical analysis, I set out a model in which firms and workers interact to determine salaries, and workers independently set their non-salary earnings. The model has three key features. First, workers can underreport their earnings at a cost. Second, the salaried labour market features adjustment costs. Third, some workers face information frictions preventing them from responding optimally to the incentives generated by the tax schedule. In turn, this provides three channels through which firms can affect workers' taxable earnings. Third-party reporting of salaries raises the cost of misreporting salaries; firms' salary offers can respond to the presence of adjustment costs; and firms' salary offers can convey information about the tax schedule to prospective workers. The model generates predictions regarding the extent of bunching of salary and taxable incomes around kinks in the tax schedule, which I then take to the data.

I present 5 sets of empirical findings. First, I document the presence of sharp bunching of overall taxable incomes around kinks in the tax schedule, providing direct evidence of behavioural responses to taxation in a lower income country context. Conceptually, these taxable income responses can be comprised of evasion responses, real earnings responses, and earnings shifting responses, so the next findings provide evidence on each of those.

Second, unilateral salary underreporting by workers is widespread. 19% of workers underreport their salary, underreporting it by an average of 16%. This leads to 4% of salary income going untaxed, or at least 5% of the tax revenue from salaried employees being lost. Consistent with the

[3] Gerard & Gonzaga (2013) study labour informality and unemployment insurance in Brazil using an employee-employer matched dataset, but do not have independent reports of wages from workers and from firms. Carrillo *et al.* (2013) exploit experimental variation in cross-checks of third party reports of business to business transactions in Ecuador.

[4] Similarly, Fisman & Wei (2004) study discrepancies in the distribution across product categories of reported imports and exports to detect tax evasion.

model, the prevalence and level of misreporting is positively correlated with the marginal tax rate faced by the worker, and with the share of the worker's total income that is self-reported. This misreporting is orders of magnitude larger than the available evidence from high income countries indicates. For instance, Kleven *et al.* (2011) find misreporting by only 1.3% of workers, amounting to 0.2% of income in Denmark. Moreover, since I am unable to detect misreporting that workers and firms collude in, the results presented here are only a loose lower bound on misreporting.

Third, I document *firm bunching* of salary incomes around kinks in the tax schedule amongst workers who do not face a kink in their budget set at these statutory kinks. Consistent with the predictions of the model, this provides direct evidence of firm-level responses to aggregate worker preferences, and of the presence of significant adjustment costs on the part of workers. In contrast to the finding by Chetty *et al.* (2011) of "aggregate bunching" around a kink in Denmark, a highly unionized labour market, in Pakistan the role of unions is negligible, and so I provide the first direct evidence of firms (rather than unions) aggregating workers' preferences.

Fourth, I document the presence of significant *double bunching*—individuals with salaries at one kink, and taxable incomes at a different kink. The model predicts that taxable income bunching is reduced when individuals are constrained by adjustment costs to accept a suboptimal salary income, and hours spent on salaried and non-salaried work are imperfectly substitutable. As a result of firm bunching, individuals with salaries at a kink are disproportionately likely to be facing significant adjustment costs, so the presence of significant double bunching indicates that workers are able to shift earnings between salary and non-salary income relatively easily. This is evidence of shifting between salary and non-salary income within the personal income tax base, rather than across bases as has more traditionally been studied (Gordon & Slemrod, 2000).

Fifth, I pursue both cross-sectional and event study methodologies to provide evidence that workers face information frictions preventing them from responding to the kinks in the income tax schedule, and that firm offers convey information about the tax schedule, increasing workers' overall responsiveness. I find that firm bunching increases worker responsiveness by around 130% in the year a worker receives a salary at a kink, and by 100% in subsequent years. These large results are similar in magnitude to those in Chetty *et al.* (2013) who find that moving to a high information neighbourhood from a median information neighbourhood roughly doubles workers' propensity to bunch at the refund-maximising kink in the Earned Income Tax Credit (EITC) in the United States.

This paper contributes to 3 literatures. First, there is a large literature on the determinants of tax evasion (see Andreoni *et al.*, 1998 and Slemrod & Yitzhaki, 2002 for surveys) and on estimating the extent of tax evasion (see Slemrod, 2007 and Slemrod & Weber, 2012 for surveys). This literature has been plagued with methodological and measurement issues, and this paper contributes to an emerging literature using discrepancies between two reports on the same tax base to study evasion (Fisman & Wei, 2004; Kumler *et al.*, 2013; Niehaus & Sukhtankar, 2013; Zucman, 2013). There is also a small literature studying the effects of third-party reporting and withholding on tax evasion, either in rich-country contexts, or studying evasion of taxes on firms, rather than workers (Yaniv, 1988; Slemrod *et al.*, 2001; Kleven *et al.*, 2011; Carrillo *et al.*, 2012; Pomeranz, 2013).

Second, a recent public finance literature posits that optimization frictions can account for the large discrepancies between microeconometrically estimated labour supply (or more generally, taxable income) elasticities and macro estimates (Chetty *et al.*, 2011; Chetty, 2012). Jones (2012) and Gelber *et al.* (2013) also study the implications of optimization frictions for the dynamics of adjustment to policy, finding a large role for adjustment costs. Of particular note here, Kleven & Waseem (2013) find that elasticities unattenuated by optimisation frictions are between 5 and 10 times larger than those implied by observed bunching behaviour at notches in the tax schedule in Pakistan. A second literature focuses specifically on information frictions, finding substantial effects of tax salience on demand elasticities (Chetty *et al.*, 2009; Finkelstein, 2009) and even on political instability (Cabral & Hoxby, 2012). A number of papers in this literature also consider the effectiveness of programmes that aim to increase responsiveness through the provision of information (Duflo & Saez, 2003; Liebman & Luttmer, 2011; Chetty & Saez, 2013) with mixed findings. Liebman & Zeckhauser (2004); Bernheim & Rangel (2009); Chetty *et al.* (2009); Mullainathan *et al.* (2012); Spinnewijn (2013b,a) also study the theoretical implications of misperception of choice sets for welfare and optimal policy.

Third, this paper contributes to a burgeoning literature on public finance and development (see Besley & Persson, 2013 for a survey) and particularly to work using administrative microdata and quasi-experimental methods to evaluate tax policy in developing countries (Kleven & Waseem, 2013; Best *et al.*, 2013; Kumler *et al.*, 2013; Pomeranz, 2013).

This paper proceeds as follows. Section 2 presents a model of salary determination by firms and workers and of non-salary earnings choices by workers to guide the empirial analysis. Section 3 presents the Pakistani context and the data I use. Section 4 presents the results on overall taxable income bunching, and evidence on its three constituent parts: evasion (4.2), real responses (4.3) and income shifting (4.4). Section 5 presents evidence that workers learn about the tax schedule from their interactions with employers. Finally, section 6 concludes.

2 Conceptual Framework

This section develops a simple, stylized model of the determination of salaries by firms and workers and of workers' joint choice of salary and non-salary earnings. The model has three key features. First, workers can underreport their earnings at a cost. Second, the salaried labour market features adjustment costs. Third, some workers face information frictions preventing them from responding optimally to the incentives generated by the tax schedule. While the model is extremely stylized, it captures the relevant features of the environment and serves to guide the empirical analysis by generating predictions regarding the extent of bunching of salary and taxable incomes around kinks in the tax schedule which I then take to the data.

Workers. Individuals have quasilinear preferences over consumption c, a CES aggregate of hours spent on salaried work l and hours spent on non-salaried work q, and reported salary and non-salary earnings, $\hat{s}$ and $\hat{n}$:

$$U\left(c,l,q,\hat{s},\hat{n}\right)=c-\frac{\left(\alpha+\beta\right)^{1-\frac{1}{\sigma}\left(1+\frac{1}{\varepsilon}\right)}}{1+\frac{1}{\varepsilon}}\left[\alpha\left(\frac{l}{\alpha}\right)^{\sigma}+\beta\left(\frac{q}{\beta}\right)^{\sigma}\right]^{\frac{1}{\sigma}\left(1+\frac{1}{\varepsilon}\right)}-e_{0}I\left\{\hat{s}<s\right\}-e\left(s-\hat{s},n-\hat{n}\right) \tag{1}$$

Individuals have heterogeneous tastes for working parameterized by $\alpha > 0$ and $\beta \geq 0$ capturing heterogeneity in abilities and disutilities of labour supply. A fraction η of individuals have $\beta = 0$ implying that these individuals earn only salary income, while the remaining fraction with $\beta > 0$ have access to a linear production technology allowing them to produce the same output as firms.[5] However, in contrast to firms, they are able to costlessly adjust their labour input. Denote the cdf of α, β as $G\left(\alpha, \beta\right)$ with corresponding density $g\left(\alpha, \beta\right)$.

Misreporting. Workers declare their salary income $\hat{s}$ and their non-salary income $\hat{n}$ to the tax authority, and pay taxes on their taxable income $\hat{z} = \hat{s} + \hat{n}$ according to a piecewise linear tax schedule $T\left(\hat{z}\right)$ featuring two kinks at threshold taxable income levels K_1, K_2 at which the marginal tax rate jumps up from τ_{j-1} to $\tau_j > \tau_{j-1}$, $j = 1, 2$. The worker's true earnings are $z = s + n$, the sum of their salary earnings $s = wl$ at wage rate w and their non-salary earnings $n = pq$, where p is the price of final output.[6]

However, at a cost, the workers can misreport their incomes, reporting $\hat{s} < s$ and/or $\hat{n} < n$. The cost of misreporting has two parts. First, due to third-party reporting of salaries by employers, the tax authority sometimes cross-checks the employee's and the employer's reports, and so reporting a salary of $\hat{s} < s$ carries a fixed cost of $e_0 > 0$. Since the probability that the two salary reports are cross-checked, and the reliability of the employer's report vary, workers are heterogeneous in the fixed cost they face, with the fixed cost distributed according to $D\left(e_0\right)$. Second, the cost of misreporting depends on the misreported amounts, where I assume that $e\left(\cdot, \cdot\right)$ is increasing in both arguments and convex, with $\partial^2 e/\partial\left(n - \hat{n}\right)\partial\left(s - \hat{s}\right) > 0$, and that $e(0, 0) = 0$. I also assume that workers who are more porductive in self-employment (higher β) are also better able to convincingly misreport their income, so that $\partial e/\partial\beta < 0$, but that this effect has diminishing returns, so that $\partial^2 e/\partial\left(n - \hat{n}\right)\partial\beta > 0$.

Information Frictions. I remain agnostic about the precise mechanism through which some individuals have failed to learn about the full tax schedule, and model information frictions in a reduced form way (Mullainathan *et al.*, 2012). I simply assume that some individuals (denoted by $\lambda = 1$) are aware of the kinks in the tax schedule and respond optimally to the full tax schedule, while the remaining individuals are naïve and behave as if the tax schedule did not feature kinks.[7] Workers' optimal salary and non-salary labour supplies are given by $s^*, n^* = \arg\max_{s,n} U$ where sophisticated workers optimize using the appropriate budget constraint $c = z - T\left(\hat{z}\right)$ while naïve

[5] The assumption that η is independent of α simplifies the exposition, but can easily be relaxed.

[6] For simplicity, I assume individuals have no non-labour income. In the case of quasi-linear utility, this makes no diference to the results. With more general utility, this will introduce income effects, but the qualitative conclusions remain unchanged.

[7] This could be for a variety of reasons. For example, it may be too costly for individuals to process the necessary information (Sims, 2003; Schwartzstein, 2012), they may confuse average and marginal rates (Liebman & Zeckhauser, 2004; Chetty *et al.*, 2009), or individuals may forget where the kinks are from year to year (Mullainathan, 2002).

taxpayers fail to, instead using some other budget constraint that does not feature the kinks induced by $T(\hat{z})$.

Firms. As in Chetty *et al.* (2011), firms are modelled extremely simply as producing output according to a linear, one-factor production function employing only labour. Firms post offers consisting of a package of a number of hours worked l and a wage rate $w(l)$, and commit to these offers *before* matching with workers, giving rise to a distribution of of hours offers $F^o(l)$. Furthermore, firms are unable to condition offers on the non-salary income of workers, i.e. $F^o(l|n) = F^o(l)\ \forall n$. A firm that posts a job requiring s hours thus earns profits of

$$\Pi = pl - w(l)\,l$$

Once a firm hires a worker, the worker's salary $w(l)\,l$ is also reported (truthfully) to the tax authority.

Fixed Costs of Adjustment. Workers must engage in a costly search process to match with jobs. As with the information frictions, I will remain agnostic about the precise source of these adjustment costs. Following Chetty *et al.* (2011) I assume that workers randomly sample a job offer from $F^o(l)$, and choose either to accept this job, or to pay a fixed cost ϕ, in which case I assume they find a job paying their preferred salary $s^* = w(l^*)\,l^*$ with certainty.

Equilibrium. The search process will map the distribution of posted offers $F^o(l)$ and the wage schedule $w(l)$ to a distribution of accepted salaries $\mathcal{D}[F^o, w]$ which combines the distribution of offers and the distribution of preferred hours. In order for the labour market to clear, it must be the case that the distribution of posted offers equals the distribution of accepted offers, or that $F^o(l) = \mathcal{D}[F^o(l), w(l)]$. That is, labour market equilibrium is a fixed point of $\mathcal{D}(\cdot)$. Furthermore, assuming free entry into a competitive market for final output, profits are bid down to zero and $w(l) = p\,\forall l$.

I proceed to analyze this model through a series of special cases. With the exception of the case focussing on salary misreporting, in each case, the model's equilibrium is summarized in terms of what it predicts for the degree of bunching of salary income and total taxable income at the kinks in the tax schedule, and how taxable income bunching varies with salary income. In particular, each equilibrium gives rise to a distribution of salary income $H(s)$ and a distribution of taxable income $J(z)$. Bunching in the salary income distribution is then the excess mass at the kinks

$$B_s(K_j) = H(K_j) - \lim_{s \uparrow K_j} H(s)\ j = 1, 2$$

and similarly bunching in the taxable income distribution is

$$B_z(K_j) = J(K_j) - \lim_{z \uparrow K_j} J(K_j)\ j = 1, 2$$

Finally, the amount of bunching of taxable incomes at K_2 amongst individuals with salary income s

is

$$B_{K_2}(s) = J(K_2|s) - \lim_{z \uparrow K_2} J(K_2|s)$$

2.1 Special Case 1: Frictionless Benchmark

As a benchmark, this section studies a special case of the model in which all workers, indexed by i, are (i) unable to misreport their income: $\partial e(0,0)/\partial(s-\hat{s}) = \partial e(0,0)/\partial(n-\hat{n}) = \infty$; (ii) costlessly able to find their preferred job: $\phi_i = 0\,\forall i$; and (iii) fully sophisticated: $\lambda_i = 1\,\forall i$.

Workers' optimal salary and taxable incomes are given by

$$\{s_i^*, z_i^*\} = \begin{cases} \alpha_i p^{1+\varepsilon}(1-\tau_0)^\varepsilon, (\alpha_i+\beta_i) p^{1+\varepsilon}(1-\tau_0)^\varepsilon & \text{if } \alpha_i+\beta_i < \underline{\delta}_1 \\ \frac{\alpha_i}{\alpha_i+\beta_i}K_1, K_1 & \text{if } \underline{\delta}_1 \le \alpha_i+\beta_i \le \bar{\delta}_1 \\ \alpha_i p^{1+\varepsilon}(1-\tau_1)^\varepsilon, (\alpha_i+\beta_i) p^{1+\varepsilon}(1-\tau_1)^\varepsilon & \text{if } \bar{\delta}_1 < \alpha_i+\beta_i < \underline{\delta}_2 \\ \frac{\alpha_i}{\alpha_i+\beta_i}K_2, K_2 & \text{if } \underline{\delta}_2 \le \alpha_i+\beta_i \le \bar{\delta}_2 \\ \alpha_i p^{1+\varepsilon}(1-\tau_2)^\varepsilon, (\alpha_i+\beta_i) p^{1+\varepsilon}(1-\tau_2)^\varepsilon & \text{if } \bar{\delta}_2 < \alpha_i+\beta_i \end{cases} \tag{2}$$

where $\underline{\delta}_j = K_j / \left[p^{1+\varepsilon}(1-\tau_{j-1})^\varepsilon\right]$ and $\bar{\delta}_j = K_j / \left[p^{1+\varepsilon}(1-\tau_j)^\varepsilon\right]$ for $j = 1, 2$. Since the labour market is frictionless and misreporting is infinitely costly, these are also workers' equilibrium outcomes and reported incomes. The following lemmas summarize the predictions for bunching in the frictionless benchmark model.

Lemma 1 (Frictionless Taxable Income Bunching)**.** *The distribution of taxable incomes, $J^*(z)$ features excess bunching at the kinks K_1, K_2: $B_z(K_j) > 0$, $j = 1, 2$*

Proof. See appendix A.1 □

Individuals' marginal incentive to accrue taxable income $1 - T'(z_i)$ jumps down as taxable incomes crosses a kink, so since the distribution of tastes $g(\alpha, \beta)$ is smooth, a mass of individuals choose to locate themselves at a kink. Turning to salary incomes,

Lemma 2 (Frictionless Salary Bunching)**.** *The distribution of the reported salary incomes for individuals with no non-salary income, $H^*(s^*|n^*=0)$ features excess bunching at the kinks K_1, K_2: $B_s(K_j|n^*>0) > 0$, $j = 1, 2$. However, the distribution of the preferred salaries of individuals with non-salary income, $H^*(s^*|n^*>0)$ does not feature excess bunching at the kinks: $B_s(K_i|n^*=0) = 0$, $j = 1, 2$.*

Proof. See appendix A.2 □

All individuals face incentives to have taxable incomes that bunch at the kinks. For individuals without non-salary income, this implies placing their salary income at a kink. However, for individuals with non-salary income, this is not the case. For these individuals, placing their taxable income at a kink implies placing their salary at a range of income levels in the interior of the tax brackets. Put differently, marginal incentives to accrue income $1 - T'(z_i)$ change as salary income crosses a kink for individuals without non-salary income, but not for individuals with non-salary income.

Finally, turning to how taxable income bunching varies with salary income,

Lemma 3 (Frictionless TI Bunching Conditional on Salary)**.** *The amount of excess bunching in taxable incomes at K_2 conditional on salary earnings s^*: $B^*_{K_2}(s^*)$ varies smoothly around K_1: $\lim_{s^* \uparrow K_1} B^*_{K_2}(s^*) = \lim_{s^* \downarrow K_1^*} B^*_{K_2}(s^*)$.*

Proof. See appendix A.3 □

As shown by equation (2), workers whose taxable income bunches at K_2 are those with $\underline{\delta}_2 \leq \alpha_i + \beta_i \leq \bar{\delta}_2$. The measure of this set of workers varies smoothly with α_i and hence with s_i^* since $g(\alpha, \beta)$ is smooth by assumption.

Having established these three properties of the frictionless equilibrium, I now turn to the equilibrium with frictions and study how these properties are affected by the presence of real adjustment costs and information frictions.

2.2 Special Case 2: Salary Misreporting

The first empirical predictions come from introducing the possibility of misreporting of incomes. I maintain the assumptions of costless labour market adjustment and full information, but allow individuals to misreport their incomes. Under the parameterization of evasion costs in 1, real decisions are undistorted by the possibility of evasion, so $s = s^*$, and $n = n^*$. Individuals must then choose whether to underreport their incomes and if so, by how much. If an individual misreports both her salary and non-salary income, her reports satisfy the first order conditions$e_s(s - \hat{s}^*, n - \hat{n}_s^*) = e_n(s - \hat{s}^*, n - \hat{n}_s^*) = \tau$, where e_s and e_n denote the partial derivatives of $e(s - \hat{s}, n - \hat{n})$ with respect to $s - \hat{s}$ and $n - \hat{n}$ respectively. By contrast, if she chooses only to misreport her non-salary income and avoid the fixed cost e_0 her choice of non-salary income report satisfies$e_n(0, n - \hat{n}_0^*) = \tau$. The worker then misreports her salary income iff

$$e_{0i} < \tau\left[s - \hat{s}^* + \hat{n}_0^* - \hat{n}_s^*\right] + e\left(0, n - \hat{n}_0^*\right) - e\left(s - \hat{s}^*, n - \hat{n}_s^*\right) = e_0^* \tag{3}$$

and so the fraction of workers who misreport their salary is $D(e_0^*)$. Intuitively, if misreporting salary income reduces the cost of misreporting non-salary income sufficiently (i.e. if e is sufficiently convex), then individuals will prefer to underreport both their salary and non-salary incomes rather than only non-salary income. That is, the cost savings from using a convex combination of salary and non-salary underreporting rather than only non-salary misreporting outweight the fixed cost of misreporting salary.

Prediction 1 (Misreporting and Marginal Tax Rates)**.** *Individuals facing higher marginal tax rates τ, are more likely to misreport their salary: $de_0^*/d\tau > 0$. Those that misreport their salary also misreport it by more: $d(s - \hat{s})/d\tau > 0$*

Proof. The first part follows immediately from application of the implicit function theorem to (3). The second part follows from inspection of the first order condition for $s - \hat{s}$. □

Intuitively, the bigger the marginal tax rate, the greater the returns to underrporting income, and so the more likely individuals are to be willing to do so. Furthermore, individuals with more non-salary income are more likely to misreport their salary:

Prediction 2 (Misreporting and Self-Reported Income)**.** *Individuals with larger non-salary (self-reported) incomes are more likely to misreport their salary: $de_0^*/d\beta > 0$. Those that misreport their salary also misreport it by more: $d\left(s - \hat{s}\right)/d\beta > 0$.*

Proof. See appendix A.4. □

Since individuals with higher non-salary incomes would like to underreport their incomes by more, they face a stronger incentive to also misreport their salary income, meaning that more people with higher non-salary incomes will also misreport their salaries. We can also summarize the implications of the presence of misreporting for bunching at the kinks as follows

Prediction 3 (Misreporting and Taxable Income Bunching)**.** *Bunching of reported taxable incomes at kinks is stronger in the presence of evasion than without evasion:*

$$B_z\left(K_j|\hat{z} \leq z\right) > B_z\left(K_j|\hat{z} = z\right)$$

Proof. See appendix A.5. □

The ability to misreport incomes makes taxable income more responsive as individuals have an additional margin along which to adjust. As a result, reported income is more sensitive to the tax rate, and so bunching is stronger.

2.3 Special Case 3: Firm Responses to Adjustment Costs

The second set of empirical predictions comes from a special case focusing on the role of adjustment cost in the labour market. For this, I assume that (i) workers are unable to misreport their income $\partial e\left(0,0\right)/\partial\left(s - \hat{s}\right) = \partial e\left(0,0\right)/\partial\left(n - \hat{n}\right) = \infty$; (ii) a proportion δ of individuals faces no search costs ($\phi_i = 0$) while the remaining workers have infinite search costs; and (iii) all workers are fully sophisticated: $\lambda_i = 1\,\forall i$.

With these assumptions the labour market equilibrium is very simply characterized. Workers have preferred salaries chosen as in (2), giving rise to a distribution of preferred hours $F^*\left(l\right)$. Workers who face no adjustment costs choose their preferred salaries, and have hours distributed according to the aggregate distribution of worker preferences $F^*\left(l\right)$, while workers with adjustment costs have salaries distributed according to the offer distribution $F^o\left(h\right)$. Therefore, the search process maps the distribution of offers and the distribution of worker preferences to a distribution of accepted offers according to

$$\mathcal{D}\left[F^{o}\right]=\delta F^{*}\left(l\right)+\left(1-\delta\right)F^{o}\left(l\right) \tag{4}$$

In equilibrium, the distribution of offers must equal the distribution of accepted salaries (a fixed point of $\mathcal{D}$) which here means that the distribution of offers matches the aggregate distribution of workers' preferences, $F^{o}\left(l\right)=F^{*}\left(l\right)$. Since workers without non-salary income have salary preferences featuring bunching at the kinks (by lemma 2) the aggregate distribution of workers' preferences will also feature bunching at the kinks, and as a result the distribution of accepted salaries of *all* workers, including those with non-salary income, will feature this *firm bunching* at the kinks (Chetty *et al.*, 2011). In particular,

Prediction 4 (Firm Bunching)**.** *The equilibrium distributions of salaries for workers both with and without non-salary income features excess bunching at the kinks in the tax schedule due to firm offers featuring excess bunching at the kinks – Firm Bunching:* $B_{s}\left(K_{j}|\hat{n}>0\right)>0$ *and* $B_{s}\left(K_{j}|\hat{n}=0\right)>0$ $j=1,2$.

Proof. The aggregate distribution of preferred hours is given by

$$F^{*}\left(l\right)=\eta F^{*}\left(l|n^{*}=0\right)+\left(1-\eta\right)F^{*}\left(l|n^{*}>0\right)$$

which features bunching at the kinks since by lemma 2 $F^{*}\left(l|n^{*}=0\right)$ features bunching at the kinks. For workers without non-salary income, the equilibrium distribution of offers is given by

$$\begin{aligned}F^{e}\left(l|n^{*}=0\right) &= \delta F^{*}\left(l|n^{*}=0\right)+\left(1-\delta\right)F^{*}\left(l\right)\\ &= \left[\delta+\left(1-\delta\right)\eta\right]F^{*}\left(l|n^{*}=0\right)+\left(1-\delta\right)\left(1-\eta\right)F^{*}\left(l|n^{*}>0\right)\end{aligned}$$

which again features bunching at the kinks due to the first term, coming from workers without non-salary income. Similarly for workers with non-salary income, the equilibrium distribution of accepted salaries is

$$\begin{aligned}F^{e}\left(l|n^{*}>0\right) &= \delta F^{*}\left(l|n^{*}>0\right)+\left(1-\delta\right)F^{*}\left(l\right)\\ &= \left(1-\delta\right)\eta F^{*}\left(l|n^{*}=0\right)+\left[\left(1-\delta\right)\left(1-\eta\right)+\delta\right]F^{*}\left(l|n^{*}>0\right)\end{aligned}$$

and this distribution also features bunching due to the preferences of workers without non-salary income. □

Despite the fact that workers with non-salary income do not face a kink in their budget set when their salary is at a kink in the tax schedule, the distribution of salaries they accept will feature bunching around the kinks. This bunching arises because firms, being unable to condition offers on worker characteristics such as sophistication and non-salary earnings, tailor offers to aggregate preferences, which feature bunching around the kinks. This firm bunching is absent in the frictionless model in section 2.1 because workers can costlessly adjust their salaries to find their preferred salary-hours packages (even though the overall distribution of salaries is the same). This means that looking for excess bunching in the distribution of salary incomes of workers with non-salary income provides a test for the presence of firm bunching, and consequently for the presence of fixed adjustment costs.

2.4 Special Case 4: Income Shifting Responses

Firm bunching attenuates the impact of adjustment costs by tailoring salary hours packages to the average preferences of workers. For workers with representative preferences – workers without non-salary income, this helps to mitigate the negative effects of the existence of adjustment costs. By contrast, for workers with unrepresentative preferences – workers *with* non-salary income, firm bunching makes it more difficult for workers to find a salary hours package fitting their preferences. However, these workers potentially have another means of responding - they can adjust the amount of their non-salary earnings in response to the salary income offers they receive.

This section characterizes the extent to which workers are able to respond by adjusting non-salary earnings. I focus on how much taxable incomes are able to respond to kinks in the tax schedule – the amount of bunching at K_2 by workers with nonsalary income – in the presence of firm bunching. Consider an equilibrium in which (i) workers are unable to misreport their income $\partial e\left(0,0\right)/\partial\left(s-\hat{s}\right)=\partial e\left(0,0\right)/\partial\left(n-\hat{n}\right)=\infty$; (ii) all workers have non-salary income ($\eta=0$); (iii) a proportion δ of individuals faces no search costs ($\phi_i=0$) while the remaining workers have infinite search costs; and (iv) all workers are fully sophisticated: $\lambda_i=1\,\forall i$. In this case,

Prediction 5 (Income Shifting and Double Bunching)**.** *Workers respond to adjustment costs in salary earnings determination by shifting income between salary and non-salary earnings. The strength of this response is governed by the substitutability between l and q in the utility function, σ. In particular, we have that (i) $B_z\left(K_2|\phi_i=0\right)\geq B_z\left(K_2|\phi_i=\infty\right)$; (ii) $B_z\left(K_2|\phi_i=0\right)=B_z\left(K_2|\phi_i=\infty\right)\iff\sigma=1$; and (iii) $B_z\left(K_2|\phi_i=0\right)-B_z\left(K_2|\phi_i=\infty\right)$ is increasing in σ.*

Proof. See appendix A.6 □

Adjustment costs sometimes force workers to accept suboptimal jobs. If salary and non-salary earnings are perfect substitutes ($\sigma=1$), then this doesn't affect the worker's taxable income, as she simply shifts from salary to non-salary income. However, if she is unable to perfectly substitute between salary and non-salary income ($\sigma>1$), then being constrained in her choice of salary hours will impact on her taxable income, and make it more difficult for her to bunch her taxable income at a kink. As a result, if we still observe strong bunching of taxable incomes amongst individuals facing adjustment costs, this implies that salary and non-salary incomes are readily substitutable, so that σ is modest.

2.5 Special Case 5: Learning by Bunching

The final empirical prediction comes from introducing information frictions. I first abstract from adjustment costs and show that if receiving a salary at K_1 makes neighbouring kinks more salient, we should expect more bunching at K_2 from workers with $s = K_1$ than from workers with nearby salaries — an *information effect*. I then introduce adjustment costs and show that when firm bunching pushes individuals to accept salaries at kinks, this causes an additional *mismatch effect* and reduces taxable income bunching at K_2. Finally, I combine these two effects and characterize the total change in the amount of excess bunching at K_2 expected for individuals with salary $s = K_1$ compared to nearby salaries.

To see the information effect, consider an equilibrium without evasion or adjustment costs: $\partial e\left(0,0\right)/\partial\left(s-\hat{s}\right) = \partial e\left(0,0\right)/\partial\left(n-\hat{n}\right) = \infty$, and $\phi_i = 0\,\forall i$ However, some individuals do not perceive the kinks ($\lambda_i = 0$). I assume that before searching for a job, all workers are equally likely to have perceived the kinks, but that when workers receive a firm offer at a kink, this makes the kinks salient and this increases the probability that a worker is sophisticated by $\Delta\gamma$. Denoting the probability that a worker is sophisticated conditional on his/her salary as $\gamma\left(s\right) \equiv \mathsf{P}\left(\lambda = 1|s\right)$, this amounts to assuming that

$$\gamma\left(s\right) = \begin{cases} \bar{\gamma} & \text{if } s \notin \{K_1, K_2\} \\ \bar{\gamma} + \Delta\gamma & \text{if } s \in \{K_1, K_2\} \end{cases} \tag{5}$$

For my empirical strategy, what is important here is not that γ is constant away from the kinks, but that it is continuous everywhere except at K_1, where it jumps up due to the kinks becoming salient. Since naïve workers do not perceive the kinks in the tax schedule, I assume that amongst these workers there is no excess bunching of taxable incomes around K_2.

Lemma 4 (Information Effect)**.** *Excess bunching in taxable income as a function of salary income jumps up discretely at $s = K_1$.*

$$\begin{aligned} B_{K_2}\left(s\right) &= \gamma\left(s\right) B^*_{K_2}\left(s|\lambda_i = 1\right) \\ &= \left[\bar{\gamma} + \Delta\gamma I\left\{s = K_1\right\}\right] B^*_{K_2}\left(s|\lambda_i = 1\right) \end{aligned}$$

where $B^*_{K_2}\left(s\right)$ is the amount of taxable income bunching at K_2 by workers who receive salary s and accurately perceive the tax schedule.

Proof. By assumption, the distribution of taxable incomes amongst individuals who do not perceive the tax schedule properly ($\lambda_i = 0$) does not feature excess bunching around the kinks ($B_{K_2}\left(s|\lambda_i = 0\right) = 0$), so the amount of taxable income bunching at K_2 conditional on salary level s is $B_{K_2}\left(s\right) = \gamma\left(s\right) B_{K_2}\left(s|\lambda_i = 1\right) + \left[1 - \gamma\left(s\right)\right] \times 0$. The result then follows by noting that by lemma 3 $B_{K_2}\left(s|\lambda_i = 1\right)$ is smooth everywhere, and $\gamma\left(s\right)$ is smooth everywhere except at the kinks. □

To see the mismatch effect, reintroduce adjustment costs so that a proportion δ of workers have $\phi_i = 0$ while the remainder have $\phi_i = \infty$. Now consider individuals who accurately perceive the tax schedule ($\lambda_i = 1$), for these workers we have,

Lemma 5 (Mismatch Effect)**.** *The probability that an individual with salary s has taxable income bunching at K_2 jumps down discretely at $s = K_1$. The amount of excess bunching at K_2 per worker with salary s is given by*

$$p(s) = \tilde{\delta}(s) B_{K_2}(s|\phi_i = 0) + \left(1 - \tilde{\delta}\right) B_{K_2}(s|\phi_i = \infty) \tag{6}$$

where

$$\tilde{\delta}(s) = \frac{\delta}{\delta + (1-\delta)\left[f^o(s) / f^*(s|n > 0)\right]}$$

and $f^*(s|n \neq 0)$ is the density of preferred salary incomes for individuals with non-salary income, $f^o(s)$ is the density of firm salary offers.

Proof. At any salary level there are $\delta f^*(s|n > 0)$ individuals with $\phi_i = 0$ who have chosen salary s because it is their preferred salary. $B_{K_2}(s|\phi_i = 0)$ of these individuals have a taxable income that bunches at K_2. The remaining $(1-\delta) f^o(s)$ individuals have $\phi_i = \infty$ and have accepted a salary offer at s despite it not being their preferred salary. $B_{K_2}(s|\phi_i = \infty)$ of these individuals bunch at K_2. Therefore, by Bayes' rule, the probability that an individual with salary s also has $\phi_i = 0$ is $\tilde{\delta}(s)$. By prediction 4 $f^o(s)$ features bunching at $s = K_1$, while by lemma 2 $f^*(s|n > 0)$ does not. Therefore, $\tilde{\delta}(s)$ jumps down discretely at $s = K_1$ and $p(s)$ assigns greater weight to bunching amongst constrained individuals. Finally, note that by prediction 5 $B_{K_2}(s|\phi_i = \infty) < B_{K_2}(s|\phi_i = 0)$ and the result follows. □

Having characterized the effects of both adjustment costs and information frictions separately, on taxable income bunching at K_2, I can now combine them to characterize the implications of both effects together on taxable income bunching at K_2.

Prediction 6 (Learning By Bunching)**.** *The probability that an individual with salary income s has taxable income bunching at K_2 jumps discretely at $s = K_1$, and the proportional jump is*

$$\frac{p(K_1)}{\lim_{x \to K_1} p(x)} = \left(1 + \frac{\Delta\gamma}{\bar{\gamma}}\right) \frac{1 + b_S \tilde{B}}{1 + b_S} \tag{7}$$

where $b_S \equiv B_s(K_1) / \lim_{s \to K_1} h(s)$ is normalized excess firm bunching of salaries at K_1, and
$\tilde{B} \equiv B_{K_2}(K_1|\phi_i = 0) / \left[\tilde{\delta} B_{K_2}(K_1|\phi_i = 0) + \left(1 - \tilde{\delta}\right) B_{K_2}(K_1|\phi_i = \infty)\right] \in [0, 1]$.

Proof. The full proof is in appendix A.7. Intuitively though, the first term captures the information effect as characterized in lemma 4, which increases taxable income bunching at K_1, while the second term captures the mismatch effect as characterized in lemma 5, which reduces taxable income bunching at K_1 □

Equation 7 characterizes the effect of adjustment costs and information frictions on workers' propensity to be taxable income bunchers as a function of their salary. In section 5 I will estimate $\frac{p(K_1)}{p(x)}$, while in section 4.3 I estimate firm bunching b_S. While this still leaves $\Delta\gamma/\bar{\gamma}$ underidentified, note that it can be bounded as

$$\frac{p(K_1)}{\lim_{x\to K_1} p(x)} - 1 \leq \frac{\Delta\gamma}{\bar{\gamma}} \leq \frac{p(K_1)}{\lim_{x\to K_1} p(x)}(1+b_S) - 1 \tag{8}$$

since $\tilde{B}$ is bounded between 0 (mismatch effect completely eliminates TI bunching) and 1 (mismatch effect is 0).

3 Context & Data

Pakistan is a large developing country with a population of around 190 million. Tax revenues represent only 9% of GDP, a small amount even by lower income country (LIC) standards (Gordon & Li, 2009). Of total revenues, around 60% is collected through various withholding regimes, including the income tax on salaried employees that I study here, one of the bedrocks of the Pakistani tax system. For the perwonal income tax, an indvidual's taxable income is the sum of an individual's salary income, business income, capital income, foreign-source income and "other" income minus charitable deductions.[8] There are two tax schedules for the taxation of individual income, depending on the composition of an individual's taxable income. If an individual's salary income is less than half of their taxable income, they are taxed as self-employed individuals, using a tax schedule featuring notches (discrete jumps in the average tax rate) at threshold incomes (these notches form the basis of the paper by Kleven & Waseem, 2013). By contrast, in this paper I focus on individuals whose salary income represents more than half of their taxable income, who are taxed using the schedule for salaried employees (roughly a quarter of income tax returns are filed by individuals taxed as salaried employees). These individuals face a complicated tax schedule featuring between 17 and 20 income brackets (or "slabs" as they are known in Pakistan).

At the thresholds between brackets, the average tax rate on income jumps up creating a discrete jump up in the individual's tax liability—a notch. Panel A of figure 1 shows an example of a budget constraint in consumption-earnings space affected by a notch at K at which the average tax rate jumps from τ_1 up to τ_2. However, a complex system known as marginal relief was introduced in 2008 to smooth these notches, creating a pair of kinks at each bracket threshold, one convex and one non-convex. The marginal relief system allows taxpayers with incomes above the bracket threshold to opt to pay a high marginal tax rate τ_M on the income they earn above the threshold. Panel B of figure 1 shows the marginal relief schedule as the blue, dashed line in the budget set diagram. This smooths the discontinuous jump in tax liability at the notch, replacing it with a sharp convex kink where the marginal tax rate jumps from τ_1 to τ_M. At some point, however, the marginal relief no longer minimizes an individual's tax liability and she optimally opts to pay τ_2 on her entire income, creating a concave kink where the marginal tax rate jumps down from τ_M to τ_2. Panel C of figure 1 shows the tax-minimizing tax schedule around K. Table 1 shows the full schedule for the year 2009-10, giving a sense of how complicated the tax schedule is.

[8] Agricultural income is taxed by the provincial governments separately

Two additional features of the Pakistani setting are important to note. First, as is common accross the world, employers of salaried workers are required to withold income tax on their employees (Slemrod, 2008) treating their salary as if it was their total taxable income, and to remit the tax to the government on their employees' behalf. In addition, employers are required to declare all their employees, their gross salaries, and the tax withheld on them to the tax authorities. However, apart from withholding income tax, firms have no other tax or benefit obligations linked to the level of the salaries they report. In particular, there is no payroll tax, and there are no social security contributions linked to workers' salaries.[9] This means that both firms and workers have incentives to underreport salaries, in contrast to the Mexican setting studied by Kumler *et al.* (2013) in which payroll taxes and benefits linked to reported salaries generate opposite incentives for firms and workers to misreport salaries. Furthermore, since employers are able to deduct their entire wage bill (salaried plus non-salaried employees) from their corporate income tax liability, underreporting salaries need not affect their corporate tax liability as it can be accompanied by overreporting non-salaried wages.

Second, trade unions that could be determining wages at a collective level are almost completely absent in Pakistan, particularly amongst the salaried, private sector workers I study. Fewer than 3% of formal sector workers in Pakistan are unionized, only 1% are under collective bargaining agreements, and unions mostly represent public sector workers (Mehmood, 2012), while the data used here covers only the private sector. The two biggest unions in Pakistan are the railway workers' union and the airline union, and all results are robust to excluding them. In addition, a number of textile firms have unionized workers, but these unions mostly represent contract workers, not salaried workers, and so do not appear in my data.

3.1 Data

This paper is one of the first to have access to administrative data on the universe of taxpayers in a low income country.[10] I use data from income tax returns from Pakistan covering the fiscal years 2007/08–2011/12 (though I focus mostly on 2008/09–2011/12 as this is the period during which the kinked schedule described above is in place)[11] from the Federal Board of Revenue, Pakistan (FBR). I also use third-party reports on salaries from Employer Statements in which employers declare their employees' salaries and income tax witheld (the equivalent of the W-2 form in the United States).

I merge the employer statements with the income tax returns to have both salary and taxable income data for workers, and both employer and employee reports of the workers' salary. As shown in table B.1, which outlines the merging procedure, the match rate is just over 50%. This rate is pulled down by two factors. First, the Employer Statement data only covers the private sector, and a large part of the salaried workforce is employed in the public sector. Second, the

[9] Some employers do make pension contributions linked to workers' pay, but these contributions are not reported to FBR, and are, in any case independent of the salary reported to FBR.

[10] predecessors include Kleven & Waseem, 2013; Kumler *et al.*, 2013; and Best *et al.*, 2013.

[11] The fiscal year in Pakistan runs from July 1 to June 30.

employer statements are not automatically checked for internal consistency, and so many records have missing or inaccurate identifiers, preventing a match with the income tax returns.

The salary data features strong bunching at round-number multiples of Rs. 5,000 in monthly terms (Rs. 60,000 in annual terms). In order to avoid conflating this heuristic bunching at round numbers with responses to kinks in the tax schedule, I drop the roughly 7.5% of jobs with round number salaries. Since 2 of the kinks in the tax schedule—at Rs. 900,000 and at Rs. 1,200,000—are at round numbers, I also exlude these two kinks from my analysis, though the results are robust to including them. After these steps, the dataset consists of 314,994 employee-employer-year observations.

The main variables I use are a worker's salary from the employer statements, and the worker's salary income, total income and taxable income (total income minus deductions) from the tax returns. In addition, I use a number of observable characteristics of the firms and workers as control variables and to estimate heterogeneity in bunching. Table 2 shows summary statistics of the matched dataset, and the subsample of workers whose taxable income differs from their salary (as reported by their employer) by at least 2%.

4 Taxable Income Responses: Kinks, Imperfect Enforcement, and Adjustment Costs

4.1 Taxable Income Responses: Bunching Around Kinks

The model in section 2 predicts that taxable incomes z will bunch around the kinks in the tax schedule. Conceptually, these taxable income responses combine real changes in earning behaviour by firms and workers; shifting responses as workers shift earnings across tax bases and/or between salary and non-salary earnings; and evasion responses. The following sections provide evidence on the presence of all three classes of responses, but I begin by establishing that there is clear evidence of sharp behavioural responses to the tax schedule by studying bunching of overall taxable incomes.

Figure 2 shows the distribution of taxable incomes around kinks in the tax schedule amongst all workers in the merged sample. Each individual's taxable income is scaled by the kink it is closest to, permitting me to pool all the kinks and years into a single figure.[12] The blue dots show the observed distribution of scaled taxable incomes, while the red line is the estimated counterfactual distribution in the absence of kinks in the tax schedule.[13] The figure also shows an estimate of the

[12]Bunching around the kinks is stable across the years in the sample, and as shown below, tends to be stronger at the lower kinks in the tax schedule, though it is present at all kinks.

[13]Grouping the data into bins 0.1% wide, the counterfactual is estimated as

$$c_j = \sum_{m=0}^{q} \beta_m (d_j)^m + \sum_{r=\bar{k}^-}^{\bar{k}^+} \gamma_r \mathbf{1}\{j = r\} + \mu_j \tag{9}$$

where c_j is the number of observations in bin j, d_j is the distance of bin j from a kink, $(j-100)/0.1$, and q is the order of the polynomial ($q = 7$ in figure 2). The second term excludes bins in a region $[\bar{k}^-, \bar{k}^+]$ around the kinks, and μ_j is an error reflecting misspecification of the estimating equation. Standard errors are obtained by bootstrapping as in Chetty excess bunching mass in the distribution normalized by the counterfactual density at the kink, b: a statistic which permits comparison across figures, and which is proportional to the magnitude of the earnings response to the tax rate (Saez, 2010).[14] Bunching is significant and extremely sharp, $b = 1.08\ (0.127)$ demonstrating clearly that behaviour is responding to the tax schedule.[15]

While previous findings from high income countries have tended to find very diffuse bunching around kinks (Saez, 2010; Chetty *et al.*, 2011, 2013), the sharp bunching found here suggests that behavioural responses are strong and precise. This is despite the fact that this is a group of workers who we expect to have trouble adjusting reported earnings to taxes as they are likely to face rigid hours constraints (Rosen, 1976; Altonji & Paxson, 1988) and search costs (Rogerson *et al.*, 2005; Manning, 2011) in determining their salary earnings, which make it difficult for workers to target their earnings at kinks precisely. This is also among the first compellingly identified evidence of behavioural responses to individual income taxation from a developing country context. Kleven & Waseem (2013) provide evidence of behavioural responses from Pakistan, focusing mostly on an earlier time period during which the notched tax schedule was in place, and Kumler *et al.* (2013) study employer compliance with payroll taxation in Mexico.

4.2 Evasion Responses: Unilateral Salary Misreporting

The model presented in section 2 predicts that all workers will misreport their non-salary income, and that some workers will choose to misreport their salary income (predictions 1 & 2). In order to asess these predictions, I exploit the fact that I have independent reports of a worker's salary from the employee and the employer, and look for discrepancies between the two.

Despite employers withholding income tax on their employees, workers may wish to underreport their salary for a number of reasons. Most importantly, workers who also report non-salary income must remit the difference between the tax on their total taxable income and the tax that their employer has already withheld on their salary, so underreporting their salary reduces the amount they must pay. Even workers who only report salary income have incentives to underreport their salary. All individuals face a number of direct taxes on consumption items (effectively excise taxes) that are classed as income taxes for administrative purposes and reported together with the tax on taxable income, and so underreporting their salary will reduce their net tax liability. They may also wish to claim that their employers have overwithheld income tax on them and claim a tax refund.

Firms also have incentives to underreport salaries. Underreporting a salary reduces the amount

et al. (2011).

[14] Total excess mass around the kinks in the distribution is given by $B = \sum_{r=\bar{k}^-}^{\bar{k}^+} (c_r - \hat{c}_r)$ where $\hat{c}_r$ is the counterfactual mass in bin r predicted by estimating (9) omitting the contribution of the dummies for the excluded range around the kink, $\hat{c}_r = \sum_{m=0}^{q} \hat{\beta}_m (d_r)^m$. To permit comparison across distributions, the figures show $b = B/c_0$, the excess mass normalized by the average counterfactual density in the excluded range, $c_0 = \left[\frac{\bar{k}^+ - \bar{k}^-}{0.1}\right]^{-1} \sum_{r=\bar{k}^-}^{\bar{k}^+} \hat{c}_r$.

[15] This finding is estimated amongst workers who can be matched to their employer's salary report. Taxable income bunching is stronger in the full population of income tax filers $b = 1.66\ (0.115)$. However, this is mostly driven by the rate at which workers can be matched to their employers being higher at the higher kinks. Reweighting the income tax filers to match the composition of the merged sample, bunching becomes statistically indistinguishable in the two samples $b = 1.32\ (0.095)$.

that the employer has to withhold and remit to the government, relieving liquidity constraints and allowing employers to offer higher net of reported tax wages to their employees. Since both employers and employees have incentives to underreport salaries, they may collude in what they report. In this case, the two reports will be the same, but will both be smaller than the true salary. Since I am only focusing on unilateral underreporting by workers, the misreporting reported here should be interpreted as a very loose lower bound on the true extent of salary misreporting.[16]

Table 5 shows the extent of salary misreporting. 19.3% of workers unilaterally underreport their salary.[17] Furthermore, these workers underreport by a significant amount. Taking the firms' reports as the truth, workers who underreport understate their income by 15.6% leading to 3.6% of overall income going untaxed. The tax losses are larger still. To estimate the amount of tax lost, I calaculate the tax liability implied by each report, assuming that the reported salary is the worker's total taxable income. Since most workers do not have any non-salary income, this assumption is accurate in most cases. However, this assumption will tend to underestimate the effect of underreporting salary on the tax liability for workers with non-salary income due to the convexity of the tax schedule. As shown in panel C, evaders understate their tax liability by at least 21.3%, leading to a loss of at least 5.1% of tax revenue.

Figure 3 tests whether predictions 1 and 2 are borne out in the data. As predicted by the model, panel A shows clearly that individuals facing higher marginal tax rates are indeed more likely to underreport their salary with individuals facing the highest marginal rates almost 5 times more likely to underreport their salary. Panel B shows that there is also evidence that individuals facing higher marginal tax rates misreport their salary by more, though this effect is mainly concentrated in individuals in the upper tax brackets above 10%. Furthermore, consistent with prediction 2, panel C shows that individuals with a higher share of self-reported income are more likely to misreport their salary income, and panel D shows that individuals with greater non-salary income also report larger discrepancies, though this relationship is not statistically significant.

Prediction 3 predicts that taxable income bunching should be stronger amongst individuals with more evasion opportunities. Since direct measures of evasion opportunities are not available, I rely on proxies for evasion opportunities that the previous literature has identified.[18] Table 4 show the results of estimating taxable income bunching separately in various sub-samples. Bunching is significantly stronger around the lower kinks in the tax schedule, for workers employed by individually owned firms (as opposed to incorporated businesses or partnerships), for firms that

[16]In principle, firms also have incentives to unilaterally underreport salaries. A particularly stark example is given in Yaniv (1988) who studies employers' incentives to underwithhold by reporting a lower salary to the tax authorities than to the worker, allowing firms to withhold more income tax on the worker than they remit to the authorities. However, workers only report their total salary earnings on their tax returns. This means that in cases where the employer reports less than the worker, I cannot distinguish between workers with a job that is not in the employer reports and firms that are underreporting, so this data does not permit an analysis of firm underrporting. I drop all workers that I observe in more than one employer statement, dropping around 9% of observations in the process.

[17]To avoid conflating underreporting with marginal differences due to rounding errors, I restrict attention to discrepancies of at least 0.25%

[18]In principle, salary misreporting is a direct measure of evasion opportunities, but since this directly affects the bunching behaviour, estimating taxable income bunching separately on misreporters vs. non-misreporters would not be a suitable test of prediction 3.

are not registered for the Value Added Tax (VAT) and firms that are not under the purview of a Large Taxpayers Unit (LTU). Bunching is also stronger for workers at smaller firms and firms in the trading, construction and services sectors. This evidence is consistent with a significant part of taxable income responses being driven by evasion, as previous work has suggested that evasion should be negatively correlated with firm size (Kleven *et al.*, 2009; Bigio & Zilberman, 2011), with the increased papertrail from being in the VAT net (Pomeranz, 2013), and with the increased scrutiny from a LTU (Almunia & Lopez-Rodriguez, 2013). It is also consistent with the patterns of heterogeneity in *corporate* income tax evasion in Pakistan shown in Best *et al.* (2013).

The evidence on unilateral misreporting presented here is is in sharp contrast to the (limited) available evidence from rich countries. For example, Kleven *et al.* (2011) find that only 1.3% of of workers in Denmark underreport their third-party reported personal income (their salary), and that the underreported income is only 0.2%. Similarly, IRS (2012) estimates that the net tax gap for salaries in the United States is under 1%. Since the evidence in Kleven *et al.* (2011) is based on pre-post audit comparisons, it also includes any underreporting in which firms and workers collude, and any jobs that are completely unreported that are detected by the auditors, which are not included in the estimates presented here.

Overall, this suggests that evasion of third-party reported salary income in Pakistan is orders of magnitude larger than in high income countries. What is more, the sample studied here – those for whom both firm and worker reports of salary are available – is likely to be the most compliant segment of the workforce. Workers whose employers fail to report their salary do not face the possibility that their report is cross-checked with the employer's and so the risk of detection is smaller, presumably increasing the amount of evasion. Conversely, it suggests that greater use of cross-checking firm and worker reports can lead to large increases in compliance and revenues.

4.3 Real Responses and Adjustment Costs: Firm Bunching

As shown by prediction 4 of the model in section 2, in the presence of adjustment costs on the part of workers and hours constraints on the part of firms, firm salary-hours offers will bunch at kinks. In order to establish whether firms are driving any of the bunching of workers' *taxable* incomes, I focus on the subset of workers who report significant non-salary earnings (the 2% sample) and investigate bunching in their *salary* incomes. As the tax schedule is a function of taxable income (the sum of salary and non-salary earnings, net of deductions), these individuals do not face a kink in their budget set when their salary is at a kink in the tax schedule and so should not have salaries bunched around kinks in the absence of adjustment costs and hours constraints (lemma 2). Therefore, if we find bunching in their salary incomes it is direct evidence that bunching is being partly driven by firms.

Figure 4 shows the findings. It shows the distribution of salary incomes (scaled by their closest kink) for individuals in the 2% sample. There is clear, sharp bunching around kinks: the normalized excess mass is $b = 2.14\ (0.211)$ indicating that firms are placing salary-hours offers around kinks, even for workers whom this does not benefit, consistent with the conceptual framework in section

2.[19] The only similar finding in previous work is Chetty *et al.* (2011), who find bunching of salaries at statutory kinks for workers with significant deductions in Denmark. However, collective wage bargaining is highly prevalent in Denmark, making it impossible to distinguish aggregation of worker preferences by firms or by trade unions. As discussed in section 3, the role of unions in Pakistan is insignificant, particularly for the salaried, private sector workers I am able to observe here.[20] The findings here are therefore the first to provide direct evidence of *firm* responses to worker incentives.

A potential concern with the interpreting this finding as evidence that firms are responding to worker-level incentives is that workers with uncertain non-salary income may prefer to accept a salary near a kink before their uncertain non-salary income is realized so that their taxable income is near a kink in expectation. In this case, the uncertainty in these workers' non-salary income would cause their taxable incomes to bunch diffusely around kinks in the tax schedule. However, figure 5 shows that this is not the case. Taxable income bunching is stronger amongst workers with non-salary income than workers with only salary income ($b = 2.03\ (0.183)$ vs $b = 0.98\ (0.137)$). More importantly, taxable income bunching is just as sharp, suggesting that salary bunching is indeed driven by firms' offers.

Firm bunching is not uniform across firms and workers. Bunching of salary incomes is weaker amongst workers who only have salary income, $b = 1.02, (0.142)$ suggesting that it is particular types of firms that respond, and that they employ particular types of workers. Propensity score reweighting observations without non-salary income to account for all observable firm and worker characteristics (DiNardo *et al.*, 1996) raises the estimated degree of bunching in salaries to $b = 1.51\ (0.191)$, suggesting only half of the discrepancy between firm-worker pairs with non-salary income and without is accounted for by observable differences between these groups, with the remainder accounted for by unobservable characteristics of the firm and the worker such as the cost of misreporting salaries and firms' ability to substitute between labour and other inputs.[21]

4.4 Income Shifting Responses: Double Bunching

Prediction 5 of the model in section 2 is that taxable income bunching should be smaller amongst individuals who face large adjustment costs in their salary determination. Direct measures of adjustment costs are unavailable, but as shown in lemma 5, firm bunchers—individuals with salaries at a kink, are more likely to face large adjustment costs than other workers. This implies

[19]The 2% sample uses a 2% threshold to classify individuals as having significant non-salary income. Table B.2 shows that this result is robust to using alternative thresholds.

[20]Dropping workers in the two big unionized sectors - airline and railway workers does not change any of the findings.

[21]Workers are weighted by $\rho_i / (1 - \rho_i)$ where ρ_i is their estimated propensity to have non-salary income. The propensity scores are estimated using a year-specific cubic spline in salary income with knots at the kinks in the tax schedule; year-specific dummies for reporting other income sources (business, property, capital, other) or deductions (charitable, contributions to worker's welfare funds); year-industry dummies; year-region dummies; year-specific cubics in firm age, the firm's number of workers, firm sales, and the proportion of the firm's workers with salaries at a kink; cubics in firm age, worker age, and how long an individual has been registered for; dummies for gender, registration for VAT, and the tax year.

that double bunching—having salary bunching at one kink and taxable income bunching at another kink— should be small if workers are unable to shift income between salary and nonsalary earnings.

Figure 6 shows the distribution of taxable income for workers with salaries very near a kink (their employer's report of their salary is within 0.5% of a kink) but with taxable incomes away from that kink.[22] There is clearly strong and sharp bunching at the kinks—$b = 5.46\ (1.24)$. From this I can conclude that workers respond to adjustment costs in salary determination by adjusting their non-salary income, shifting earnings between salary and non-salary income. Furthermore, the strength of the bunching at the kinks suggests that this shifting is relatively easy for workers, i.e. that σ, the complementarity between salaried and nonsalaried hours worked, is modest.

Furthermore, this is not purely driven by workers misreporting their earnings. As figure 7 shows, there is still strong evidence of double-bunching amongst workers who don't underreport their salaries ($b = 6.69\ (1.81)$ in panel A) and who self-report a salary at a kink ($b = 6.28\ (1.38)$ in panel B). To see more clearly that workers are adjusting their non-salary earnings to the presence of adjustment costs, figure 8 shows part of the distribution of non-salary income for individuals with salaries within 0.5% of a kink (in blue circles), and for individuals with salaries more than 0.5% away from a kink, but within 2.5% of a kink (in orange diamonds). The distributions are clearly similar with the exception of the presence of clear bunching in the blue distribution at Rs. 50,000, Rs. 150,000, and to a lesser extent Rs. 250,000. A Kolmogorov-Smirnov test rejects the equality of the two distributions with $p = 0.0069$. These amounts coincide with the distances between various kinks in the tax schedule, confirming that individuals are adjusting their non-salary income.[23]

The shifting behaviour identified here is distinct from what the previous literature has usually focused on as it occurs within the personal income tax base, whereas previous work has tended to study shifting across bases (for example, between the corporate and personal income tax base) in response to differences in tax rates across bases (Gordon & Slemrod, 2000; Kleven & Schultz, 2013).[24] However, these spillovers from the taxation of salary income onto non-salary income are important for two reasons. First, individuals in lower income countries are much more likely to have both salary and non-salary income than people in high income countries, so these spillovers have a real relevance for taxable income elasticities and tax policy. Second, as shown in section 4.2, individuals with significant amounts of non-salary earnings are more likely to evade their tax liabilities by misreporting their salary income so shifting responses will affect the overall level of evasion of the income tax.

[22] I define a worker as having taxable income "away" from that kink if the closest kink to his/her salary income is not the same as the closest kink to his/her taxable income.

[23] These numbers are also salient, round numbers and individuals may have a tendency to report non-salary income at these round numbers, but this does not invalidate these findings as Rs. 50,000 is no rounder an amount for individuals with salaries at kinks than for individuals with salaries near kinks.

[24] Kleven & Waseem (2013) is a notable exception here that studies the same issue using different variation for identification. They study shifting between salary and non-salary income using the notch by the shift from the salaried tax schedule to the non-salaried tax schedule when salary income falls below half of taxable income described in section 3 whereas I study shifting between salary and non-salary income within the salaried tax schedule.

5 Learning by Bunching

This section presents results arguing that workers learn about the tax schedule through their interactions with employers to determine their salaries. In particular, receiving a salary at a kink teaches the workers about the importance of kinks, making them more likely to have taxable income bunching at a different kink. Section 5.1 takes a cross-sectional approach to demonstrate this, while section 5.2 uses an event study methodology to control for potential selection effects coming from assortative matching between sophisticated firms and workers.

5.1 Cross-Sectional Results

Consistent with prediction 6 of the model in section 2, this section shows that the probability that workers have taxable income bunching at a kink is discretely higher when their salary income is at a kink due to firm bunching, consistent with the presence of a large information effect. To do this, I first develop a methodology to estimate the counterfactual level of taxable income bunching that would have been observed were it not for firm bunching. To do this, I borrow from the bunching literature (Saez, 2010; Chetty *et al.*, 2011; Kleven & Waseem, 2013) and fit a flexible, high-order polynomial to binned data on the observed outcome using data near the kinks, but excluding data very near the kinks. I then use the predicted values from this estimate as my counterfactual at the kinks.

Under the identifying assumption that all firm and worker characteristics determining worker responsiveness covary smoothly with salary around the kinks, this method will identify a valid counterfactual for responsiveness in the absence of firm bunching. I pool all the kinks together by scaling salaries by the closest kink to them, K_S. Then, within bins of scaled salaries, I calculate the fraction of individuals who have taxable incomes near (defined as being within 0.5%) a kink K_{TI}, but that kink is not the same kink that their salary is near, $K_S \neq K_{TI}$. I denote this conditional probability by p_j, the probability of taxable income bunching in a bin centered at j%. Grouping the data into bins of scaled salary 0.2% wide, I estimate the counterfactual conditional probability with the following polynomial:

$$p_j = \sum_{m=0}^{q} \beta_m \left(d_j\right)^m + \sum_{r=\bar{k}^-}^{\bar{k}^+} \gamma_r \mathbf{1}\left\{j = r\right\} + \mu_j \tag{10}$$

where d_j is the distance between bin j and a kink, $(j - 100)/0.2$, and q is the order of the polynomial ($q = 7$ in figures 9 & 10). The second term excludes bins in a region $[\bar{k}^-, \bar{k}^+]$ around the kinks reflecting the possibility that firm bunching may be slightly diffuse (though as shown in figure 4 firm bunching is extremely sharp). Finally, μ_j is a residual reflecting misspecification of the conditional probability equation (10).

From the estimates of the coefficients in equation (10), I calculate the counterfactual conditional probability in the excluded region $[\bar{k}^-, \bar{k}^+]$ as $\hat{p}_j = \sum_{m=0}^{q} \hat{\beta}_m \left(d_j\right)^m$, and my estimate of the discrete change in the conditional probability at the kinks is the proportional difference between the observed probability at the kink and the average counterfactual probability in the excluded region around the kink

$$\Delta p = \frac{p_{100}}{\left(\frac{0.2}{\bar{k}^{+}-\bar{k}^{-}+1}\right)\sum_{j=\bar{k}^{-}}^{\bar{k}^{+}}\hat{p}_j} - 1$$

I estimate the standard error of this estimate using by bootstrapping as in Chetty *et al.* (2011).

Figure 9 shows the baseline results. The blue circles show the observed conditional probabilities p_j, while the orange line is the estimated counterfactual conditional probability function estimated excluding bins ±0.2% from a kink. The figure also shows the estimate of Δp, how much the probability changes at the kink due to firm bunching, along with its bootstrapped standard error. Figure 9 clearly shows a sharp spike in taxable income responsiveness for workers with salaries at a kink. The estimate of $\Delta p = 1.283\ (0.265)$ indicates that the responsiveness of workers affected by firm bunching is more than double that of workers with salaries near, but not at a kink. Using equation (8) and the estimate of firm bunching $b_S = 2.14\ (0.219)$, we can bound the information effect $\Delta\gamma/\bar{\gamma}$ as lying between $\Delta p = 1.283$ and $\Delta p \times (1 + b_S) = 4.029$, significantly larger than 0.

One potential concern with this could be that the findings are driven by misreporting of the worker's salary rather than responses by workers to their true salary. Figure 10 repeats the same exercise, but uses the worker's salary report instead of the employer's report. In this case, $\Delta p = 0.691\ (0.240)$, which is significantly smaller than the result in figure 9, and does not seem visually to be larger than the spikes at other salary levels due to noise. This implies that it is the true salary, not the salary that the worker (mis)reported that matters for the worker's information.

A more serious concern is with the identifying assumption that firms that offer salaries that bunch around the kinks are not differentially likely to employ workers who are more responsive to tax incentives, i.e. there is no assortative matching between firms and workers on tax responsiveness. To address this, I turn to an event study methodology to rule out time-invariant selection effects.

5.2 Event Study Results

This section pursues an event study methodology to allow me to control for fixed firm and worker characteristics such as sophistication and long-run responsiveness (Jacobson *et al.*, 1993; Hilger, 2013; Chetty *et al.*, 2013). The strategy consists of comparing the outcomes of individuals who experience a treatment event – receiving a salary at a kink in the tax schedule, to the outcomes of individuals who experience a control event – receiving a salary in the interior of a tax bracket.

Let $g \in \{K, I\}$ denote whether a worker experiences a salary at a kink, K, or in the interior, I in a year s. Let t denote the year an outcome y is observed, and define $q = t - s$ as event time, the number of years since the event g. The event study model for individual i's outcome y, allowing the effects of kink and interior events to vary by period is then

$$y_{i,g,t,s} = \alpha + \sum_{j=q^{-}}^{q^{+}} \mu_j^K \mathbf{1}\{g = K, q = j\} + \sum_{j=q^{-}}^{q^{+}} \mu_j^I \mathbf{1}\{g = I, q = j\} + \Gamma X_{i,g,t,s} + u_{i,g,t,s} \qquad (11)$$

where $q^- < 0$ and $q^+ > 0$ are the minimum and maximum values of q, respectively, $X_{i,g,t,s}$ is a vector of observable covariates,[25] and $u_{i,g,t,s}$ is an independently distributed error term. The key advantages of the event study model over a standard difference in difference (DD) model are that it allows the effects to vary arbitrarily over time and does not impose a fixed difference between treatment and control groups. This flexible specification then permits assessment of whether the assumptions necessary for identification in a traditional DD framework are reasonable.

To operationalize the event study I define salary bunching as receiving a salary within 1% of a kink, and taxable income bunching as having taxable income within 1% of a kink.[26] For the event study I construct two samples. The "Kink" sample consists of all workers who receive a salary at a kink in year s and also receive a salary in the interior of a tax bracket in year $s-1$. The restriction on year $s-1$ salary is intended to facilitate the interpretation of the results as the impacts of first-time exposure to the kinks in the tax schedule rather than repeated exposure. The "Interior" sample consists of workers who receive a salary in the interior of a tax bracket in year s but work at a firm where at least 1 worker received a salary at a kink. These workers must also have received a salary in the interior in year $s-1$ to match the restriction on workers in the Kink sample. I also include data from 2007/08 to increase the time dimension of the panel.[27] Table 3 shows summary statistics of the Kinks and Interior samples.

The outcomes I analyze are a variety of bunching behaviours combining taxable income and salary income bunching as summary indicators of sophisticated tax responsiveness. I will be interested in both the contemporaneous impact β_0 and medium-run impact $\beta_{>0}$ of experiencing a salary at a kink rather than a salary in the interior of a tax bracket, which I will estimate as DD treatment effects from equations of the form

$$\begin{aligned} y_{i,g,t,s} \quad = \quad & \delta + \lambda I\{g = K\} + \psi_{-1} I\{q = -1\} + \psi_0 I\{q = 0\} + \psi_{>0} I\{q > 0\} \\ & + \beta_{-1} I\{g = K, q = -1\} + \beta_0 I\{g = K, q = 0\} + \beta_{>0} I\{g = K, q > 0\} \\ & + \Gamma X_{i,g,t,s} + u_{i,g,t,s} \end{aligned} \tag{12}$$

where the ψ_{-1} and β_{-1} terms are included to account for the fact that as a result of the definition of the samples, salary bunching is mechanically 0 at event-time -1 in both samples.[28]

The identifying assumption required in order to interpret β_0 and $\beta_{>0}$ as the causal effects of

[25] $X_{i,g,t,s}$ contains fixed effects for event by observation years, firm sector, tax office, worker gender, having business income, property income, capital income, foreign income, other income, charitable deductions, and other deductions; year-specific cubics in firm size, firm age, the degree of firm bunching in the firm, worker age and time since worker registration for taxes; and a year-specific cubic spline in salary with knots at the kinks in the tax schedule.

[26] Results are similar though noisier when using 0.5% for taxable income bunching

[27] This data was not appropriate for use in the earlier analysis as it is a year in which the thresholds the workers faced were notches rather than kinks. See Kleven & Waseem (2013) for a detailed analysis of these notches.

[28] Note that the DD estimators β_0 and $\beta_{>0}$ are related to the event study dummies μ_j^g according to $\beta_0 = \left(\mu_0^K - \mu_0^I\right) - \left(-q^-\right)^{-1} \sum_{j=q^-}^{-1} \left(\mu_j^K - \mu_j^I\right)$ and $\beta_{>0} = \left(q^+\right)^{-1} \sum_{j=1}^{q^+} \left(\mu_j^K - \mu_j^I\right) - \left(-q^-\right)^{-1} \sum_{j=q^-}^{-1} \left(\mu_j^K - \mu_j^I\right)$.

receiving a salary at a kink is that

$$I\{g=K,q=0\},I\{g=K,q>0\}\perp u_{i,g,t_O,t_E}$$

which in economic terms requires that

1. There be parallel trends between the treatment and control groups before the event is experienced: $\mu_j^K-\mu_j^I=\rho\,\forall j<0$
2. Individuals do not anticipate receiving a salary at a kink and respond preemptively in period $q<0$
3. There are no time-varying unobserved worker or firm characteristics that are correlated both with event time q and with responsiveness to taxes y.

I verify the reasonableness of condition 1 visually through inspection of the μ_j^G dummies for $j<0$. Condition 2 is unlikely to be a major concern as inflation is high and volatile over the period I study in Pakistan, making it difficult to predict future wages with the precision necessary to target the kinks. Any violation of condition 2 will likely increase responsiveness in years before the event however, attenuating my estimates. Condition 3 is addressed through the addition of a rich set of individual-year and firm-year controls, reducing the scope for unobserved factors to affect the estimates.

Figure 11 shows the evolution of overall bunching in salary (in panel A) and taxable (in panel B) incomes in the Kink sample and the Interior sample. The blue circles show the estimated μ_j^K for the Kinks sample from equation (11), while the orange crosses show the μ_j^I from the Interior sample. Each panel also shows the estimated contemporaneous effect β_0 and medium-term effect $\beta_{>0}$ from estimating equation (12) alongside the mean pre-event level of bunching in the Kink sample. In both cases it is striking that the trends before the event in the Kinks and Interior samples are remarkably parallel. Furthermore, experiencing a salary at a kink has a large effect on future bunching behaviour. It increases future salary bunching by 1.7 percentage points, a 28% increase, and taxable income bunching by 0.9 percentage points, an increase of 36%.

Figure 12 shows event study results from decomposing taxable income bunching into its 3 constituent parts and is constructed in the same way as figure 11. It shows taxable income bunching rates when salary income bunches at a different kink (panel A), when salary income is not at a kink (panel B), and when salary income is at the same kink as taxable income (panel C). Though the results are considerably noisier as a result of the smaller samples, it is again striking that the pre trends in the two samples are remarkably parallel. Panels A and C show that there is a strong effect on taxable income bunching accompanied by salary bunching either at the same kink or at a neighbouring kink, with medium term bunching increasing by 100% for double bunchers in panel A, and 40% for those without non-salary income in panel C. However, panel B does not provide strong evidence of an effect on taxable income bunching when salary is not at a kink, suggesting that workers learn about the significance of the kinks in the tax schedule and seek out both salaries and taxable incomes at kinks in the medium term.

6 Conclusion

This paper has exploited unique access to employee-employer matched administrative tax data on firms and salaried workers in Pakistan to explore the underappreciated role of firms in determining how workers' taxable earnings respond to taxation. Consistent with the model presented in section 2, I present evidence on three ways in which firms affect workers' earnings responses. First, third-party reporting of salaries by employers makes underreporting taxable income more costly for workers. Second, firms' equilibrium salary-hours offers respond endogenously to the presence of adjustment costs in the labour market. Third, workers learn about the tax schedule from firms' salary offers, making them more responsive to taxation both in the same year and in subsequent years.

While third-party reporting of salaries raises the cost of misreporting, it has not eliminated misreporting, as 19% of workers still misreport their salaries. This casts doubt on the efficacy of the third-party reporting that recent work has suggested is central to tax enforcement (Kleven *et al.*, 2009; Pomeranz, 2013) in low-capacity environments where cross checks of multiple reports of the same tax base are absent or limited. Since salaried workers are generally the most compliant group of personal income tax payers, this suggests that the self-employed will be even more responsive to the tax schedule. Together these suggest that the returns to investment in fiscal capacity are large, particularly in cross-checking third-party reports of tax bases and increased scrutiny of individuals with non-salary income, both subjects for future work.

This paper has also shown that in addition to their central role in the collection of taxes, firms play a key role in mitigating the impact of adjustment costs and information frictions on workers' responsiveness to taxes. Firms reduce the impact of adjustment costs by aggregating the preferences of workers and this is manifested in bunching in their salary offers around kinks in the tax schedule. This equilibrium level response by firms to worker level incentives also indirectly increases the responsiveness of taxable earnings amongst individuals facing information frictions since firms respond to aggregate preferences, including those of workers who do not face information frictions.

Furthermore, firm offers at kinks in the tax schedule directly affect the information frictions attenuating worker responses to the tax schedule. The effects of this are large, workers are around 130% more responsive to taxation in years they receive salaries at a kink, and 100% more responsive in future years, suggesting that information frictions play a large role in attenuating responses to the tax schedule. This implies that policies that make the tax schedule more salient, or simplify the tax schedule can have large impacts on how firms and workers respond to the tax schedule.[29] Of course, whether increasing responsiveness to the tax schedule improves welfare depends on whether

[29] Indeed, in 2012, the tax schedule was simplified to a set of 10 standard kinks, partly in response to the perception that the tax schedule was overly complicated. Of course, 10 kinks is still more than are found in most OECD countries suggesting further scope for simplification.

individuals are suffering a utility loss from their attenuated responsiveness due to optimization frictions (see Liebman & Zeckhauser, 2004; Bernheim & Rangel, 2009 for e.g.).

Overall, this paper has shown that in lower income country contexts firms affect evasion decisions, real earnings decisions, the impact of adjustment costs, and the information workers use in their decisions. In light of this, the virtual absence of firms in the public finance literature on income taxation (Kopczuk & Slemrod, 2006) has come at a great cost. Firms must play a central role in our analysis of income taxation in lower income countries.

References

ALMUNIA, MIGUEL, & LOPEZ-RODRIGUEZ, DAVID. 2013. *Firms' Responses to Tax Enforcement Strategies: Evidence from Spain*. Mimeo: University of Warwick. 19

ALTONJI, JOSEPH G., & PAXSON, CHRISTINA H. 1988. Labor Supply Preferences, Hours Constraints, and Hours-Wage Trade-Offs. *Journal of Labor Economics*, **6**, 254–276. 17

ANDREONI, JAMES, ERARD, BRIAN, & FEINSTEIN, JONATHAN. 1998. Tax complicance. *Journal of Economic Literature*, **36**, 818–860. 3

BERNHEIM, B. DOUGLAS, & RANGEL, ANTONIO. 2009. Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics. *Quarterly Journal of Economics*, **124**, 51–104. 4, 27

BESLEY, TIMOTHY, & PERSSON, TORSTEN. 2011. *Pillars of Prosperity: The Political Economy of Development Clusters*. Princeton University Press. 1

BESLEY, TIMOTHY, & PERSSON, TORSTEN. 2013. Taxation and Development. *Chap. 2, pages 51–110 of:* AUERBACH, ALAN J., CHETTY, RAJ, FELDSTEIN, MARTIN, & SAEZ, EMMANUEL (eds), *Handbook of Public Economics, Volume 5*. Elsevier. 1, 4

BEST, MICHAEL CARLOS, BROCKMEYER, ANNE, KLEVEN, HENRIK JACOBSEN, SPINNEWIJN, JOHANNES, & WASEEM, MAZHAR. 2013. *Production vs Revenue Efficiency With Limited Tax Capacity: Theory and Evidence From Pakistan*. Mimeo: London School of Economics. 4, 15, 19

BIGIO, SAKI, & ZILBERMAN, EDUARDO. 2011. Optimal Self-Employment Income Tax Enforcement. *Journal of Public Economics*, **95**. 19

CABRAL, MARIKA, & HOXBY, CAROLINE. 2012. *The Hated Property Tax: Salience, Tax Rates, and Tax Revolts*. NBER Working Paper No. 18514. 4

CARD, DAVID, CHETTY, RAJ, FELDSTEIN, MARTIN, & SAEZ, EMMANUEL. 2010. *Expanding Access to Administrative Data for Research in the United States*. NSF White Paper. 1

CARRILLO, PAUL, SINGHAL, MONICA, & POMERANZ, DINA. 2013. *Tax Me if You Can: Third Party Cross-Checks and Evasion Substitution*. Mimeo: Harvard University. 2

CARRILLO, PAUL E., EMRAN, M. SHAHE, & RIVADENEIRA, ANITA. 2012. *Do Cheaters Bunch Together? Profit Taxes, Withholding Rates and Tax Evasion*. Mimeo: George Washington University. 3

CHETTY, RAJ. 2012. Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply. *Econometrica*, **80**, 969–1018. 4

CHETTY, RAJ, & SAEZ, EMMANUEL. 2013. Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients. *American Economic Journal: Applied Economics*, **5**, 1–31. 4

CHETTY, RAJ, LOONEY, ADAM, & KROFT, KORY. 2009. Salience and Taxation: Theory and Evidence. *American Economic Review*, **99**, 1145–1177. 4, 5

CHETTY, RAJ, FRIEDMAN, JOHN N, OLSEN, TORE, & PISTAFERRI, LUIGI. 2011. Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence From Danish Tax Records. *Quarterly Journal of Economics*, **126**, 749–804. 1, 2, 3, 4, 6, 10, 16, 17, 20, 22, 23, 34

CHETTY, RAJ, FRIEDMAN, JOHN, & SAEZ, EMMANUEL. 2013. *Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings*. forthcoming: American Economic Review. 3, 17, 23

DE MEL, SURESH, MCKENZIE, DAVID J., & WOODRUFF, CHRISTOPHER. 2009. Measuring Microenterprise Profits: Must we Ask How the Sausage is Made? *Journal of Development Economics*, **88**. 1

DINARDO, JOHN, FORTIN, NICOLE M., & LEMIEUX, THOMAS. 1996. Labor Market Institutions and the Distribution of Wages, 1973–1992: A Semiparametric Approach. *Econometrica*, **64**, 1001–1044. 20

DUFLO, ESTHER, & SAEZ, EMMANUEL. 2003. The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence From a Randomized Experiment. *Quarterly Journal of Economics*, **118**, 815–842. 4

FINKELSTEIN, AMY. 2009. E-ztax: Tax Salience and Tax Rates. *Quarterly Journal of Economics*, **124**, 969–1010. 4

FISMAN, RAYMOND, & WEI, SHANG-JIN. 2004. Tax Rates and Tax Evasion: Evidence from "Missing Imports" in China. *Journal of Political Economy*, **112**, 471–500. 2, 3

GELBER, ALEX, JONES, DAMON, & SACKS, DAN. 2013. *Earnings Adjustment to Policy: Evidence From the Social Security Earnings Test*. NBER Working Paper No. 19491. 4

GERARD, FRANÇOIS, & GONZAGA, GUSTAVO. 2013. *Informal Labor and the Cost of Social Programs: Evidence from 15 Years of Unemployement Insurance in Brazil*. Mimeo: Columbia University. 2

GORDON, ROGER, & LI, WEI. 2009. Tax Structures in Developing Countries: Many Puzzles and a Possible Explanation. *Journal of Public Economics*, **93**, 855–866. 14

GORDON, ROGER H., & SLEMROD, JOEL. 2000. Are 'Real' Responses to Taxes Simply Incoe Shifting Between Corporate and Personal Tax Bases? *Pages 240–280 of:* SLEMROD, JOEL (ed), *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*. New York: Russell Sage Foundation; Cambridge and London: Harvard University Press. 3, 21

HILGER, NATE G. 2013. *How Does Family Income Affect College Attendance? Evidence From Timing of Parental Layoffs*. Mimeo: Brown University. 23

INTERNAL REVENUE SERVICE (IRS). 2012. *Overview of the Tax Year 2006 Tax Gap Estimates*. Washington, DC. 19

JACOBSON, LOUIS S., ALONDE, ROBERT J., & SULLIVAN, DANIEL G. 1993. Earnings Losses of Displaced Workers. *American Economic Review*, **83**, 685–709. 23

JONES, DAMON. 2012. Inertia and Overwithholding: Explaining the Prevalence of Income Tax Refunds. *American Economic Journal: Economic Policy*, **4**, 158–185. 4

KLEVEN, HENRIK, KNUDSEN, MARTIN, KREINER, CLAUS THUSTRUP, PEDERSEN, SOREN, & SAEZ, EMMANUEL. 2011. Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark. *Econometrica*, **79**, 651–692. 3, 19

KLEVEN, HENRIK JACOBSEN, & SCHULTZ, ESBEN. 2013. *Estimating Taxable Income Responses Using Danish Tax Reforms*. American Economic Journal: Economic Policy, forthcoming. 21

KLEVEN, HENRIK JACOBSEN, & WASEEM, MAZHAR. 2013. Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan. *Quarterly Journal of Economics*, **128**, 669–723. 2, 4, 14, 15, 17, 21, 22, 24

KLEVEN, HENRIK JACOBSEN, KREINER, CLAUS T., & SAEZ, EMMANUEL. 2009. *Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries*. NBER Working Paper No. 15218. 1, 19, 26

KOPCZUK, WOJCIECH, & SLEMROD, JOEL. 2006. Putting Firms into Optimal Tax Theory. *American Economic Review Papers and Proceedings*, **96**, 130–134. 1, 27

KUMLER, TODD, VERHOOGEN, ERIC, & FRÍAS, JUDITH. 2013. *Enlisting Employees in Improving Payroll-Tax Compliance: Evidence from Mexico*. Mimeo: Columbia University. 2, 3, 4, 15, 17

LIEBMAN, JEFFREY B., & LUTTMER, ERZO F.P. 2011. *Would People Behave Differently if They Better Understood Social Security? Evidence From a Field Experiment*. NBER Working Paper No. 17287. 4

LIEBMAN, JEFFREY B., & ZECKHAUSER, RICHARD J. 2004. *Schmeduling*. Mimeo: Harvard Kennedy School of Government. 4, 5, 27

MANNING, ALAN. 2011. Imperfect Competition in the Labor Market. *Chap. 11, pages 973–1041 of:* CARD, DAVID, & ASHENFELTER, ORLEY (eds), *Handbook of Labor Economics, Volume 4b*. Elsevier. 1, 17

MEHMOOD, AMJAD. 2012. *Annual Labour Report 2011*. Tech. rept. Labour Education Foundation, Pakistan. 15

MULLAINATHAN, SENDHIL. 2002. A Memory-Based Model of Bounded Rationality. *Quarterly Journal of Economics*, **117**, 735–774. 5

MULLAINATHAN, SENDHIL, SCHWARTZSTEIN, JOSHUA, & CONGDON, WILLIAM J. 2012. A Reduced-Form Approach to Behavioral Public Finance. *Annual Review of Economics*, **4**, 511–540. 4, 5

NIEHAUS, PAUL, & SUKHTANKAR, SANDIP. 2013. The Marginal Rate of Corruption in Public Programs. *Journal of Public Economics*, **104**, 52–64. 2, 3

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. 2013. *Tax Administration 2013: Comparative Information on OECD and Other Advanced and Emerging Economies*. OECD Publishing. 1

POMERANZ, DINA. 2013. *No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax*. Mimeo: Harvard Business School. 1, 3, 4, 19, 26

ROGERSON, RICHARD, SHIMER, ROBERT, & WRIGHT, RANDALL. 2005. Search-Theoretic Models of the Labor Market: A Survey. *Journal of Economic Literature*, **43**, 959–988. 1, 17

ROSEN, HARVEY S. 1976. Taxes in a Labor Supply Model with Joint Wage-Hours Determination. *Econometrica*, **44**, 485–507. 17

SAEZ, EMMANUEL. 2010. Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy*, **2**, 180–212. 2, 17, 22, 52

SCHWARTZSTEIN, JOSHUA. 2012. *Selective Attention and Learning*. Mimeo: Dartmouth College. 5

SIMS, CHRISOPHER. 2003. Implications of Rational Inattention. *Journal of Monetary Economics*, **50**, 665–690. 5

SLEMROD, JOEL. 2007. Cheating Ourselves: The Economics of Tax Evasion. *Journal of Economic Perspectives*, **21**, 25–48. 3

SLEMROD, JOEL. 2008. Does it Matter Who Writes the Check to the Government? The Economics of Tax Remittance. *National Tax Journal*, **61**, 251–275. 1, 15

SLEMROD, JOEL, & WEBER, CAROLINE. 2012. Evidence of the invisible: toward a credibility revolution in the empirical analysis of tax evasion and the informal economy. *International Tax and Public Finance*, **19**(1), 25–53. 3

SLEMROD, JOEL, & YITZHAKI, SHLOMO. 2002. Tax avoidance, evasion, and administration. *Chap. 22, pages 1423–1470 of:* AUERBACH, A. J., & FELDSTEIN, M. (eds), *Handbook of Public Economics*. Handbook of Public Economics, vol. 3. Elsevier. 3

SLEMROD, JOEL, BLUMENTHAL, MARSHA, & CHRISTIAN, CHARLES. 2001. Taxpayer response to an increased probability of audit: evidence from a controlled experiment in Minnesota. *Journal of Public Economics*, **79**(3), 455–483. 3

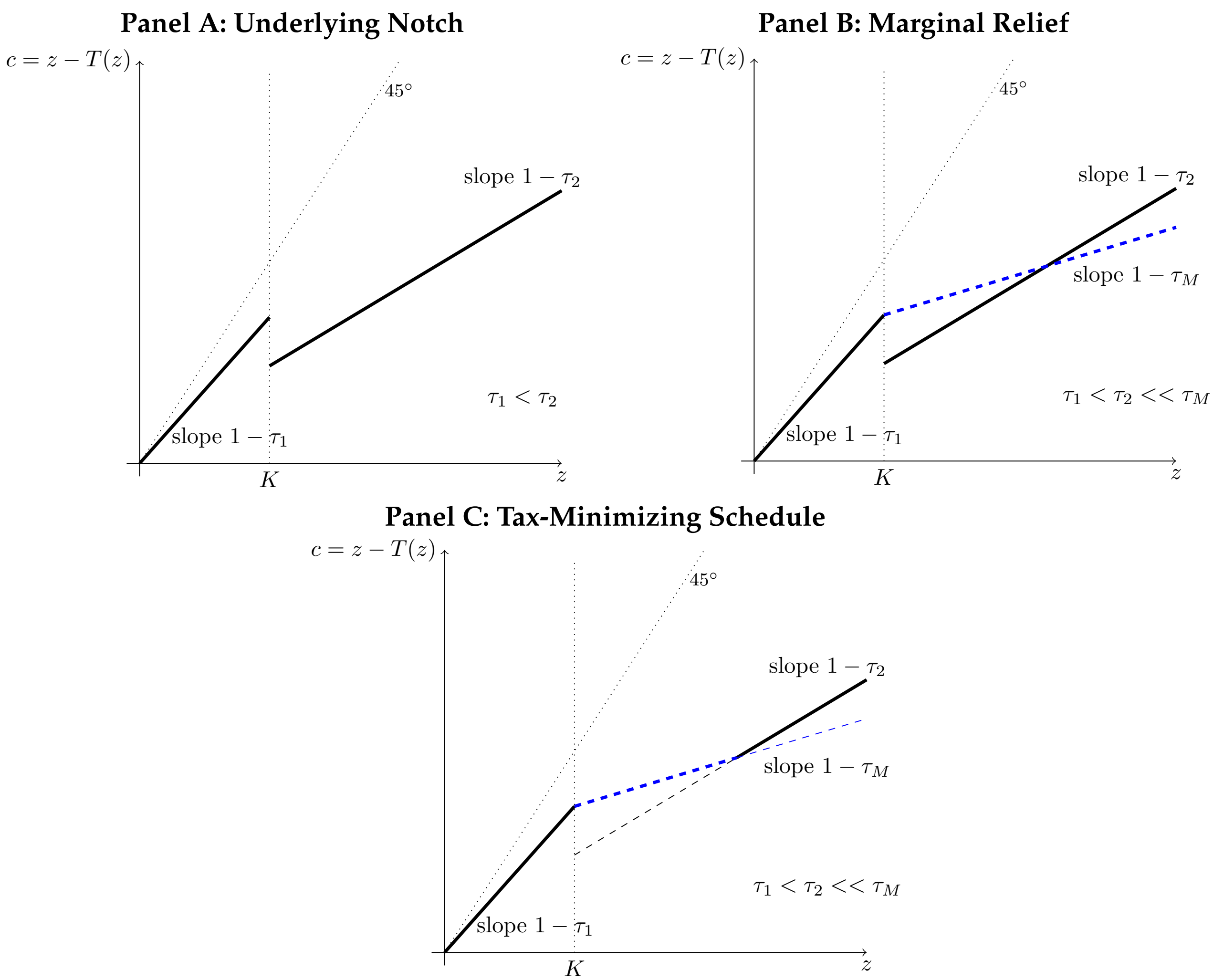
SPINNEWIJN, JOHANNES. 2013a. *Heterogeneity, Demand for Insurance and Adverse Selection*. Mimeo: London School of Economics. 4

SPINNEWIJN, JOHANNES. 2013b. *Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs*. Journal of the European Economic Association, forthcoming. 4

YANIV, GIDEON. 1988. Withholding and Non-Withheld Tax Evasion. *Journal of Public Economics*, **35**, 183–204. 3, 18

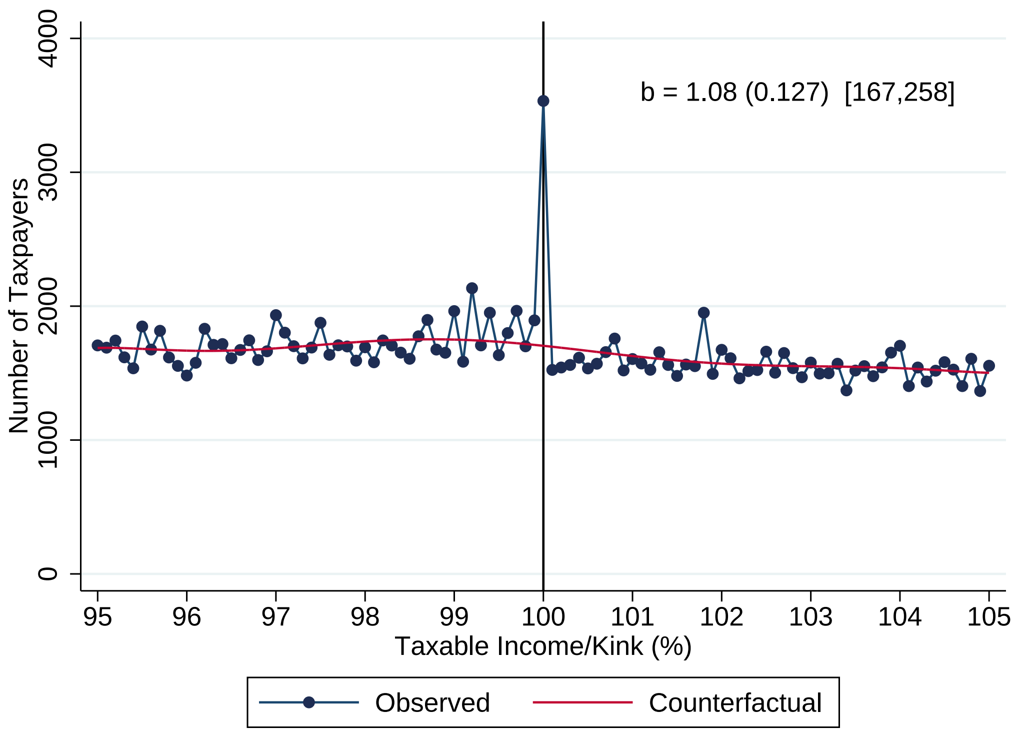
ZUCMAN, GABRIEL. 2013. The Missing Wealth of Nations: Are Europe and the U.S. net Debtors or net Creditors? *Quarterly Journal of Economics*, **128**, 1321–1364. 3

FIGURE 1: EFFECT OF TAX SCHEDULE ON BUDGET SET



Notes: The figure shows the effect of the tax schedule for salaried workers $T(z)$ on a worker's budget constraint around a bracket threshold K. The budget constraint shows the relationship between consumption $c = z - T(z)$ and taxable income z. Panel A shows the underlying notch at the threshold K where the average tax rate jumps up from τ_1 to τ_2 generating a discrete fall in consumption at the threshold. Panel B shows the effect of marginal relief (the blue, dashed line) allowing taxpayers to opt to pay a high marginal tax rate $\tau_M >> \tau_2$ on their income above the threshold generating a convex kink at K where the marginal tax rate jumps up from τ_1 to τ_M. At some point, the marginal relief ceases to minimize the taxpayer's tax liability, and so an optimizing taxpayer opts to pay the flat rate τ_2 on their entire income, generating a concave kink where the marginal tax rate jumps down from τ_M to τ_2. Panel C shows the tax minimizing schedule around the threshold K combining the convex and concave kinks.

FIGURE 2: TAXABLE INCOME BUNCHING

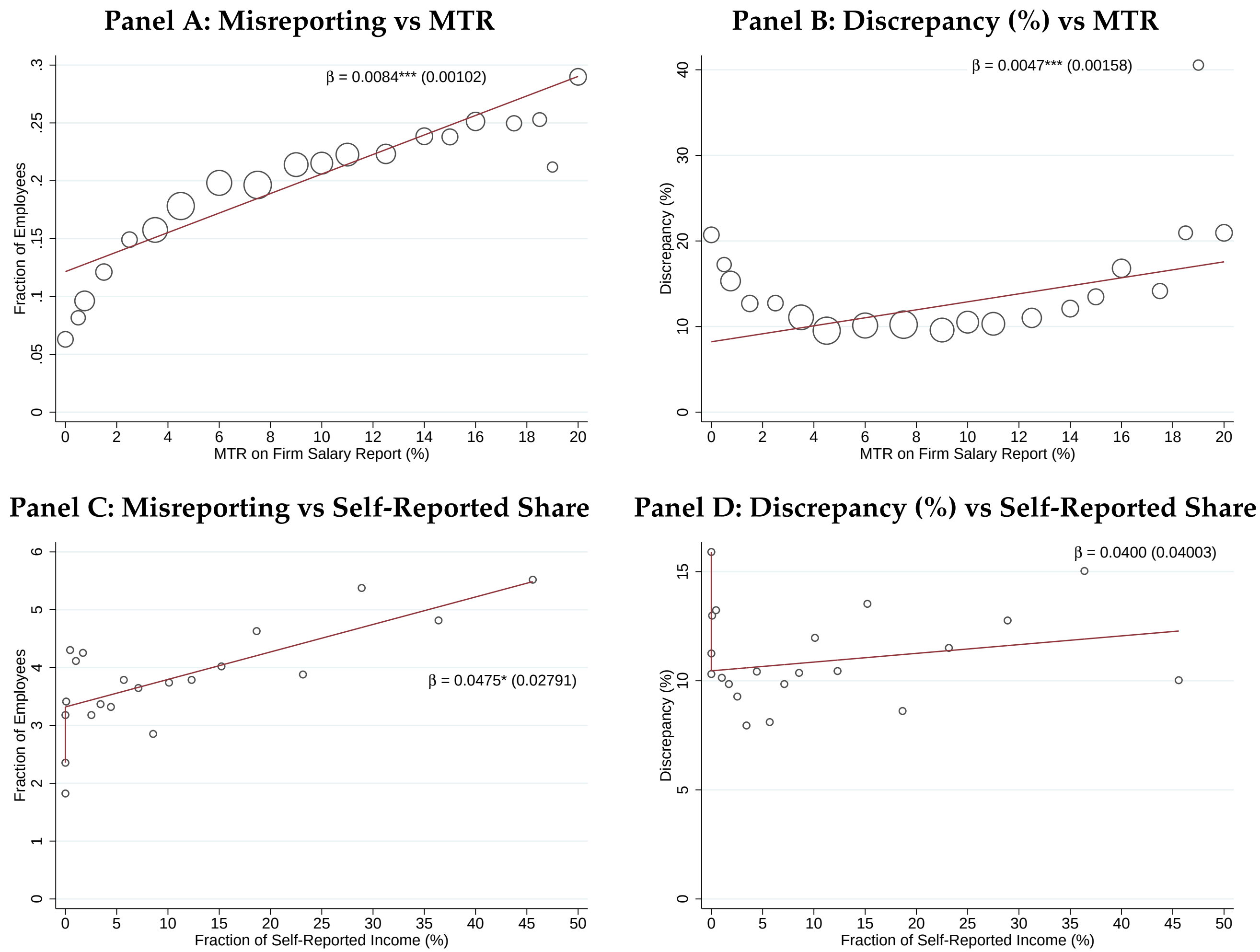


Notes: The figure shows the observed distribution of workers' taxable incomes (as a percentage of the nearest kink) in blue dots, alongside an estimate of the counterfactual density that would be observed if the tax schedule did not feature a kink at 100%. Grouping the data into bins 0.1% wide, the counterfactual is estimated as

$$c_j = \sum_{m=0}^{q} \beta_m \left(d_j\right)^m + \sum_{r=\bar{k}^-}^{\bar{k}^+} \gamma_r \mathbf{1}\left\{j = r\right\} + \mu_j \tag{13}$$

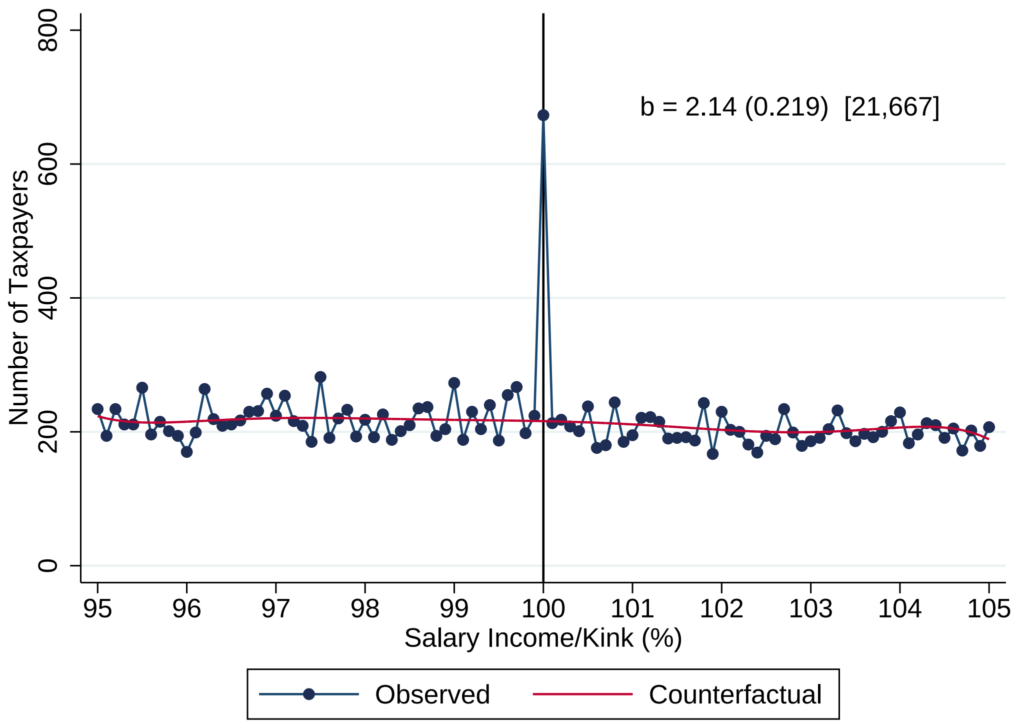
where c_j is the number of observations in bin j, d_j is the distance of bin j from a kink, $(j - 100)/0.1$, and q is the order of the polynomial ($q = 7$ in figure 2). The second term excludes bins in a region $\left[\bar{k}^-, \bar{k}^+\right]$ around the kinks, and μ_j is an error reflecting misspecification of the estimating equation. The figure also shows the normalized estimated excess mass in the observed distribution around the kinks. Total excess mass around the kinks in the distribution is given by $B = \sum_{r=\bar{k}^-}^{\bar{k}^+} \left(c_r - \hat{c}_r\right)$ where $\hat{c}_r$ is the counterfactual mass in bin r predicted by estimating (9) omitting the contribution of the dummies for the excluded range around the kink, $\hat{c}_r = \sum_{m=0}^{q} \hat{\beta}_m \left(d_r\right)^m$. The figure shows $b = B/c_0$, the excess mass normalized by the average counterfactual density in the excluded range, $c_0 = \left[\frac{\bar{k}^+ - \bar{k}^-}{0.1}\right]^{-1} \sum_{r=\bar{k}^-}^{\bar{k}^+} \hat{c}_r$. Standard errors are obtained by bootstrapping as in Chetty *et al.*, 2011. The number of observations used for the estimation is shown in square brackets.

FIGURE 3: CORRELATES OF MISREPORTING: PREDICTIONS 1 AND 2



Notes: Panels A and B show the correlation of the probability that a worker underreports his/her salary, and the size of the discrepancy amongst misreporters (as a percentage of the firm's salary report), with the marginal tax rate the worker faces, respectively. Panels C and D show the correlation between the probability that a worker underreports his/her salary, and the size of the discrepancy for misreporters and the marginal tax rate, and the share of his/her income that is self-reported, respectively. The fraction of self-reported income is calculated as the worker's reported non-salary income divided by the sum of the worker's reported non-salary income and the worker's employer's report of his/her salary. This measure is capped at 50% as above this the worker is no longer taxed as a salaried worker. In Panels A and B the grey circles show the averages within each tax rate, with the size of the circle proportional to the number of individuals facing each tax rate. The red line shows the fitted relationship from a linear OLS regression. In Panels C and D the grey circles show the averages within each vingtile of the distribution of the self-reported share. The red line shows the fitted relationship from a linear OLS regression of the outcome variable on a dummy for having zero self-reported income and the share of self-reported income. The figures also show the coefficient β from the regression along with its standard error clustered by tax office $\times$ employer type $\times$ year.

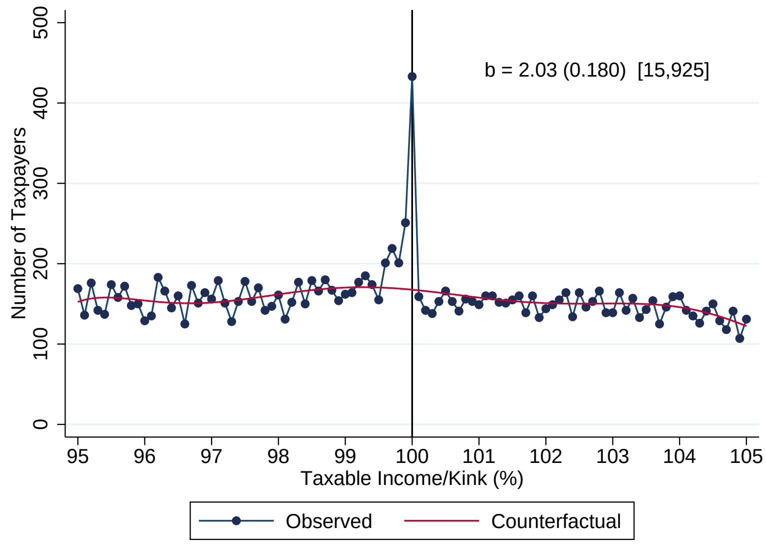
FIGURE 4: FIRM BUNCHING OF SALARIES



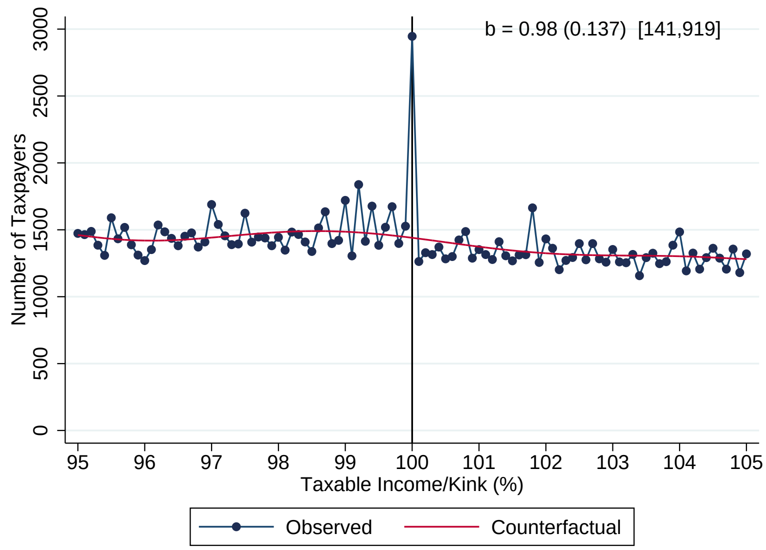
Notes: The figure shows the observed distribution of scaled salary incomes for workers with significant non-salary income (defined as having taxable income more than 2% different from salary income) in blue dots, alongside the estimated counterfactual distribution (red line) and the estimated normalized excess bunching mass b, its standard error in parentheses, and the number of observations used in square brackets (see notes to figure 2 for estimation details).

FIGURE 5: FIRM BUNCHING: FIRM RESPONSES OR UNCERTAIN INCOME?

Panel A: Taxable Income Distribution; Workers with Non-salary Income

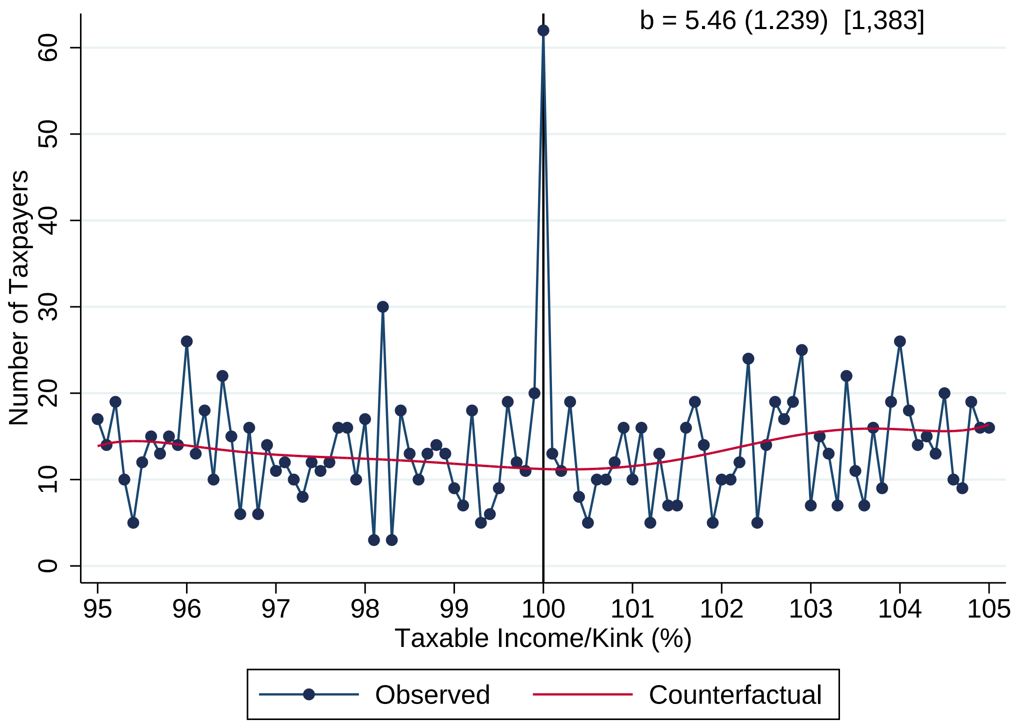


Panel B: Taxable Income Distribution; Workers without Non-salary Income



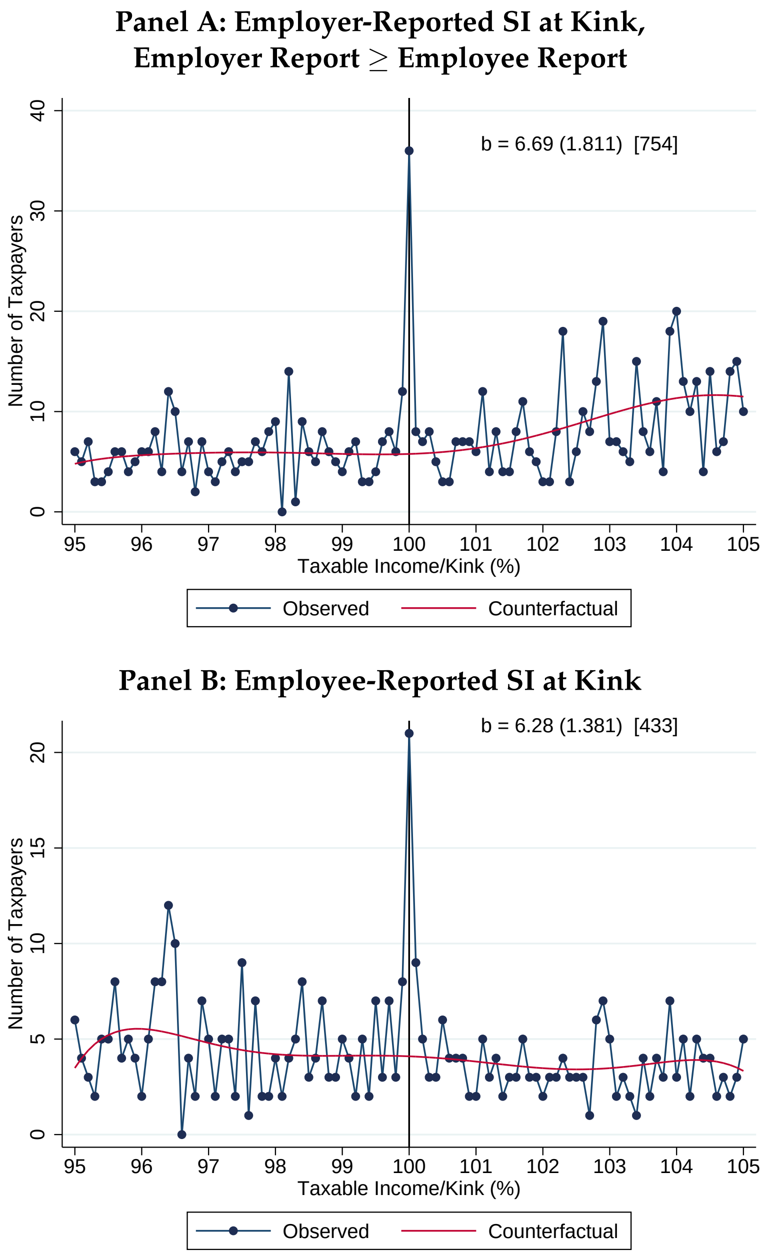
The figure shows bunching of scaled taxable incomes amongst workers with non-salary income (defined as having taxable income more than 2% different from salary income) in panel A, and without non-salary income in panel B. The figure shows the observed distributions in blue dots, alongside the estimated counterfactual distribution (red line) and the estimated normalized excess bunching mass b, its standard error in parentheses, and the number of observations used in square brackets (see notes to figure 2 for estimation details).

FIGURE 6: DOUBLE BUNCHING: TAXABLE INCOME DISTRIBUTION OF SALARY INCOME BUNCHERS



Notes: The figure shows taxable income distribution of workers with salary incomes near kinks, but with taxable incomes away from that kink. Workers are defined as having salary incomes near a kink if their employer reports a salary within 0.5% of a kink. Workers are defined as having taxable income away from that kink if the closest kink to their taxable income is not the same as the closest kink to their salary income. The blue dots show the observed distribution in 0.1% bins of scaled income, while the red line shows the estimated counterfactual distribution, and the panels also show estimates of the normalized excess bunching mass b. See the notes to figure 2 for details of the estimation methodology.

FIGURE 7: DOUBLE BUNCHING IS NOT DRIVEN BY SALARY MISREPORTING

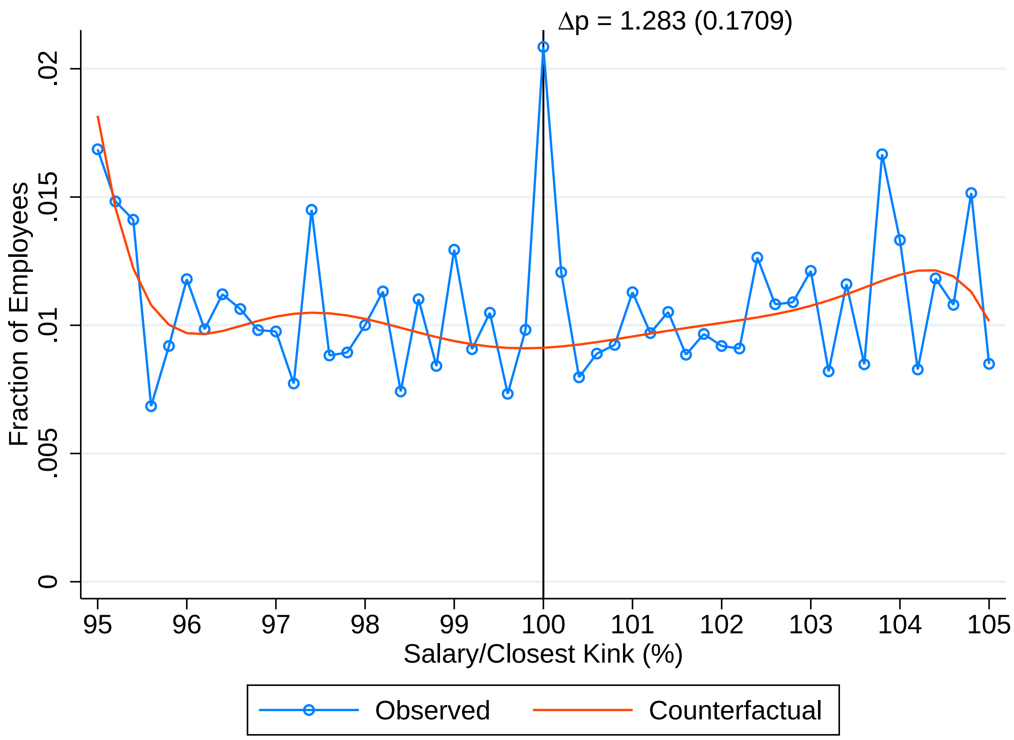


Notes: The figure shows taxable income distributions of workers with salary incomes near kinks. Workers are defined as having salary incomes near a kink if their employer reports a salary within 0.5% of a kink in panel A, or if the employee reports a salary within 0.5% of a kink in panel B. The panels show the distributions of taxable incomes for workers whose taxable income is away from the kink their salary is near. Specifically, if the closest kink to their taxable income is not the same as the closest kink to their salary income. The blue dots show the observed distribution in 0.1% bins of scaled income, while the red line shows the estimated counterfactual distribution, and the panels also show estimates of the normalized excess bunching mass b. See the notes to figure 2 for details of the estimation methodology.

FIGURE 8: DISTRIBUTIONS OF NON-SALARY INCOME

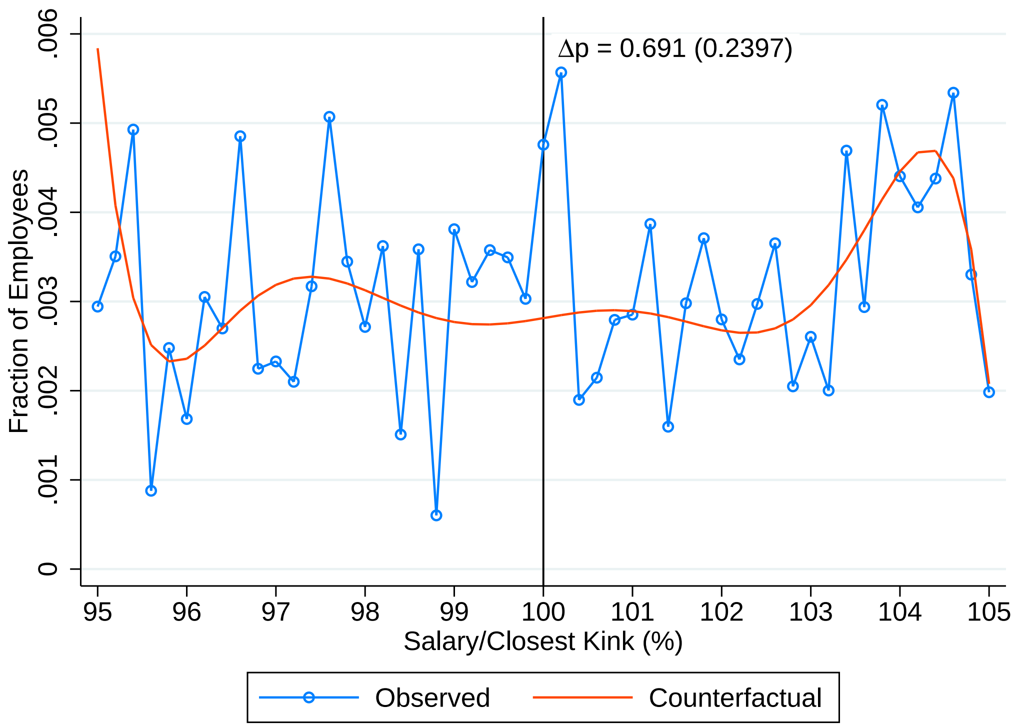
Notes: The figure shows the distributions of non-salary income (defined as the difference between taxable income and salary income) amongst workers with salaries at a kink (defined as being within 0.5% of a kink) in blue circles and workers with salaries near, but not at kinks (defined as being within 2.5% of a kink, but not within 0.5%) in orange diamonds. The figure also shows the *p*-value from a Kolmogorov-Smirnov test of the equality of the two distributions.

FIGURE 9: PROBABLITY OF TAXABLE INCOME BUNCHING AS DISTANCE OF SALARY FROM KINK VARIES



Notes: The figure shows how the probability that a worker has taxable income near a kink (defined as being within 0.5% of a kink) and that that kink is not the same kink as the closest kink to their salary, $K_{Si} \neq K_{TIi}$, changes as the distance between an individual's salary (as reported by his/her employer) and a kink varies. Each blue circle is the probability that a worker has taxable income near a kink within a bin of width 0.2%. The orange line is the estimated counterfactual probability estimated on the binned data using a 7th order polynomial as in equation (10) and excluding points in bins 0.2% above and below the kink. The figure also shows Δp, the observed increase in probability at the kink normalized by the average counterfactual probability in the excluded region around the kink, along with its standard error calculated by bootstrapping the procedure 200 times in brackets.

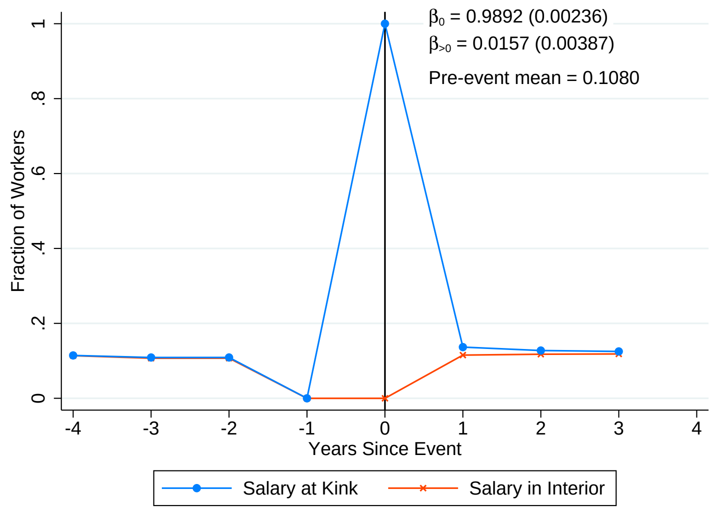
FIGURE 10: INFORMATION EFFECT IS NOT DRIVEN BY SALARY MISREPORTING



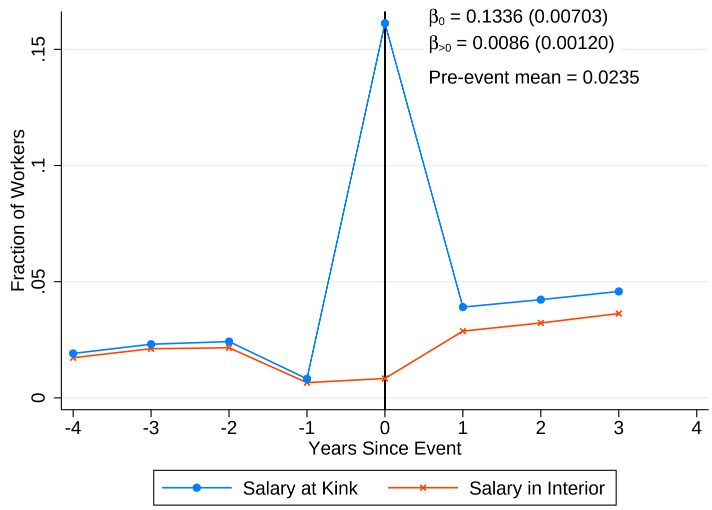
Notes: The figure shows how the probability that a worker has taxable income near a kink (defined as being within 0.5% of a kink) and that that kink is not the same kink as the closest kink to their salary, $K_{Si} \neq K_{TIi}$, changes as the distance between an individual's salary (as reported by the employee) and a kink varies. Each blue circle is the probability that a worker has taxable income near a kink within a bin of width 0.2%. The orange line is the estimated counterfactual probability estimated on the binned data using a 7th order polynomial as in equation (10) and excluding points in bins 0.2% above and below the kink. The figure also shows Δp, the observed increase in probability at the kink normalized by the average counterfactual probability in the excluded region around the kink, along with its standard error calculated by bootstrapping the procedure 200 times in brackets.

FIGURE 11: EVENT STUDY OF RECEIVING SALARY AT A KINK: OVERALL SALARY AND TAXABLE INCOME BUNCHING

Panel A: Salary Bunching



Panel B: Taxable Income Bunching

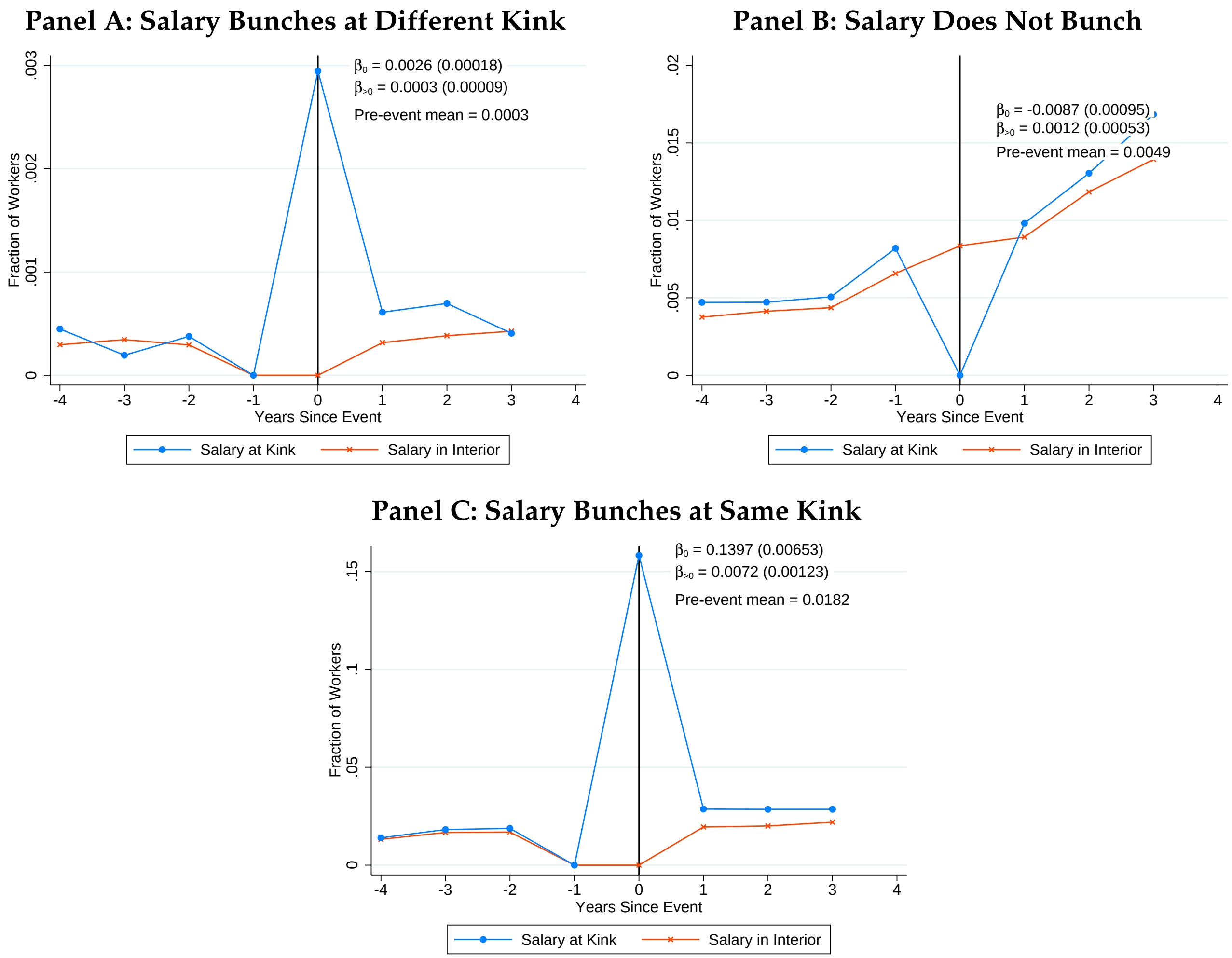


Notes: The figure shows the evolution of bunching behaviour in the Kinks Sample, who experience a salary at a kink in year 0 (defined as a salary within 1% of a kink), and in the Interior Sample, who experience a salary in the interior of a tax bracket in year 0. Panel A shows salary bunching in the two samples. By definition of the samples salary bunching is 0 in both samples in year -1, 0 in the Interior sample in year 0, and 1 in the Kink sample in year 0. Panel B shows taxable income bunching in the two samples defined as having taxable income within 1% of a kink. The panels also show the estimated contemporaneous β_0 and medium term $\beta_{>0}$ effects of receiving a salary at a kink estimated from equation (12):

$$\begin{aligned} y_{i,g,t,s} \quad = \quad & \delta + \lambda \mathbf{1}\{g = K\} + \psi_{-1}\mathbf{1}\{q = -1\} + \psi_0 \mathbf{1}\{q = 0\} + \psi_{>0}\mathbf{1}\{q > 0\} \\ & + \beta_{-1}\mathbf{1}\{g = K, q = -1\} + \beta_0 \mathbf{1}\{g = K, q = 0\} + \beta_{>0}\mathbf{1}\{g = K, q > 0\} + \Gamma X_{i,g,t,s} + u_{i,g,t,s} \end{aligned}$$

The standard errors shown are robust standard errors clustered at the g, t, s level. The figures also show the pre-event mean in the Kinks Sample.

FIGURE 12: EVENT STUDY OF RECEIVING SALARY AT A KINK: DECOMPOSITION OF TAXABLE INCOME BUNCHING



Notes: The figure shows the evolution of the 3 components of taxable income bunching (defined as having taxable income within 1% of a kink) behaviour in the Kinks Sample, who experience a salary at a kink in year 0 (defined as a salary within 1% of a kink), and in the Interior Sample, who experience a salary in the interior of a tax bracket in year 0. The panels show the probabilities that taxable incomes and salaries bunch at different kinks (panel A) that taxable incomes bunch while salaries do not (panel B) and that taxable incomes and salaries bunch at the same kink (panel C) in the two samples. By definition of the samples bunching is 0 in both samples in year -1, and 0 in the Interior sample in year 0 in panels A and C; and 0 in the Kinks sample in panel B. The panels also show the estimated contemporaneous β_0 and medium term $\beta_{>0}$ effects of receiving a salary at a kink estimated from equation (12):

$$y_{i,g,t,s} \quad = \quad \delta + \lambda \mathbf{1}\{g = K\} + \psi_{-1}\mathbf{1}\{q = -1\} + \psi_0 \mathbf{1}\{q = 0\} + \psi_{>0}\mathbf{1}\{q > 0\} + \beta_{-1}\mathbf{1}\{g = K, q = -1\} + \beta_0 \mathbf{1}\{g = K, q = 0\} + \beta_{>0}\mathbf{1}\{g = K, q > 0\} + \Gamma X_{i,g,t,s} + u_{i,g,t,s}$$

The standard errors shown are robust standard errors clustered at the g, t, s level. The figures also show the pre-event mean in the Kinks Sample.

TABLE 1: TAX SCHEDULE FOR SALARIED EMPLOYEES IN TAX YEAR 2009/10

From (Rs. 000s)	To (Rs. 000s)	Flat Rate (%)	Marginal Relief Rate (%)
0	200	0	
200	250	0.5	20
250	350	0.75	20
350	400	1.5	20
400	450	2.5	20
450	550	3.5	20
550	650	4.5	30
650	750	6	30
750	900	7.5	30
900	1,050	9	30
1,050	1,200	10	40
1,200	1,450	11	40
1,450	1,700	12.5	40
1,700	1,950	14	40
1,950	2,250	15	40
2,250	2,850	16	50
2,850	3,550	17.5	50
3,550	4,550	18.5	50
4,550	8,650	19	60
8,650	∞	20	60

Notes: The table shows the tax schedule for salaried employees in the tax year from 1 July 2009 to 30 June 2010. Each row represents a bracket of the tax schedule with its lower and upper bounds in the first two columns. The third column shows the flat average tax rate within the bracket, and the fourth column shows the marginal rate at which individuals can opt to be taxed on their income above the lower bound of the tax bracket.

TABLE 2: SUMMARY STATISTICS

Variable	Matched Sample Mean	Matched Sample s.d.	2% Sample Mean	2% Sample s.d.
Salary (Employer Report)	1,460,681	3,579,009	1,456,672	4,737,086
I{Salary ≈ Kink}	0.061	0.239	0.056	0.231
Salary (Employee Report)	1,712,665	73,329,352	2,139,826	119,638,823
I{Salary ≈ Kink}	0.059	0.236	0.055	0.227
Taxable Income	1,882,843	159,527,501	2,582,172	260,363,676
I{TI ≈ Kink}	0.060	0.238	0.055	0.228
Total Income	1,851,126	159,181,413	2,532,930	259,796,228
I{Business Income}	0.031	0.173	0.077	0.266
I{Capital Income}	0.001	0.036	0.003	0.051
I{Foreign Income}	0.000	0.015	0.000	0.022
I{Other Income}	0.014	0.118	0.029	0.168
I{Deductions}	0.063	0.244	0.095	0.293
I{Zakat Deductions}	0.047	0.212	0.072	0.259
I{WWF Deductions}	0.005	0.070	0.008	0.089
I{Charitable Deductions}	0.020	0.141	0.031	0.173
Age	43.0	11.75	43.1	12.46
I{Female}	0.035	0.183	0.037	0.189
Years Reg. For Tax	9.0	5.39	9.1	5.36
I{Reg for VAT}	0.049	0.216	0.093	0.291
Firm ♯ of Workers	1751.9	3295.37	1651.0	3225.64
Firm Sales (Rs. Millions)	13,202.82	60,794.638	12,845.64	64,760.363
Firm Salary Bunching	0.066	0.072	0.068	0.076
Firm Age	11.0	4.73	10.5	4.94
I{Firm Reg for VAT}	0.798	0.402	0.763	0.425
I{Firm Under LTU}	0.632	0.482	0.577	0.494
I{Corporate Employer}	0.943	0.232	0.936	0.245
I{Individual Employer}	0.013	0.113	0.014	0.116
I{Partnership Employer}	0.044	0.205	0.051	0.219
I{Agriculture}	0.011	0.105	0.007	0.085
I{Construction}	0.015	0.123	0.015	0.122
I{Finance}	0.177	0.381	0.174	0.379
I{Manufacturing}	0.338	0.473	0.298	0.457
I{Mining}	0.038	0.192	0.038	0.192
I{Services}	0.353	0.478	0.402	0.490
I{Trading}	0.024	0.153	0.026	0.160
I{Utilities}	0.033	0.179	0.028	0.165
I{Other}	0.010	0.099	0.011	0.102
2008/09 (♯ of obs)	78,070		26,671	
2009/10 (♯ of obs)	81,536		26,594	
2010/11 (♯ of obs)	74,254		29,055	
2011/12 (♯ of obs)	81,134		35,912	
Overall (♯ of obs)	314,994		118,232	

Notes: The table shows means and standard deviations of variables in the matched sample and the 2% sample whose taxable income differs from their employer-reported salary by more than 2%. Income being "≈ Kink" is defined as being within 0.5% of a kink. Zakat deductions are religious charitable giving, collected centrally by the state in Pakistan. WWF Deductions are employers' tax-deductible contributions to a workers' welfare fund. VAT is the Value Added Tax (called the generalised sales tax in Pakistan), and the LTU is the Large Taxpayers Unit.

TABLE 3: SUMMARY STATISTICS OF KINKS AND INTERIOR SAMPLES

Variable	Kinks Sample		Interior Sample	
	Mean	s.d.	Mean	s.d.
Salary (Employer Report)	617,137	3,058,758	619,935	1,443,355
I{Salary ≈ Kink}	0.306	0.461	0.053	0.224
I{TI ≈ Kink}	0.056	0.230	0.015	0.120
I{Business Income}	0.006	0.076	0.006	0.075
I{Capital Income}	0.000	0.015	0.000	0.016
I{Foreign Income}	0.000	0.008	0.000	0.007
I{Other Income}	0.003	0.050	0.002	0.049
I{Deductions}	0.076	0.265	0.077	0.267
I{Zakat Deductions}	0.007	0.085	0.007	0.084
I{WWF Deductions}	0.001	0.029	0.001	0.028
I{Charitable Deductions}	0.003	0.052	0.003	0.055
Age	9.5	18.33	8.7	17.70
I{Female}	0.030	0.170	0.033	0.178
Years Reg. For Tax	6.8	5.53	6.3	5.54
Firm ♯ of Workers	2961.8	4235.82	2808.2	4108.55
Firm Sales (Rs. Millions)	20,019.76	67,312.533	18,540.23	62,783.602
Firm Salary Bunching	0.069	0.070	0.064	0.044
Firm Age	11.0	4.45	11.0	4.44
I{Firm Under LTU}	0.702	0.457	0.719	0.449
I{Corporate Employer}	0.964	0.186	0.969	0.173
I{Individual Employer}	0.008	0.091	0.006	0.078
I{Partnership Employer}	0.027	0.163	0.024	0.155
I{Agriculture}	0.010	0.097	0.010	0.098
I{Construction}	0.014	0.118	0.014	0.116
I{Finance}	0.186	0.389	0.185	0.388
I{Manufacturing}	0.337	0.473	0.340	0.474
I{Mining}	0.060	0.237	0.059	0.235
I{Services}	0.346	0.476	0.349	0.477
I{Trading}	0.018	0.134	0.016	0.124
I{Utilities}	0.026	0.158	0.024	0.153
I{Other}	0.003	0.059	0.003	0.058
2007/08 (♯ of obs)	61,789		145,053	
2008/09 (♯ of obs)	90,756		221,942	
2009/10 (♯ of obs)	98,581		244,071	
2010/11 (♯ of obs)	93,526		228,142	
2011/12 (♯ of obs)	87,715		209,264	
Overall (♯ of obs)	432,367		1,048,472	
Overall (♯ of workers)	101,758		355,158	

Notes: The table shows means and standard deviations of variables in the Kinks Sample, who experience a salary at a kink in year 0 (defined as a salary within 1% of a kink), and in the Interior Sample, who experience a salary in the interior of a tax bracket in year 0. Income being "≈ Kink" is defined as being within 1% of a kink. Zakat deductions are religious charitable giving, collected centrally by the state in Pakistan. WWF Deductions are employers' tax-deductible contributions to a workers' welfare fund. The LTU is the Large Taxpayers Unit.

TABLE 4: TAXABLE INCOME BUNCHING AND PROXIES FOR EVASION OPPORTUNITIES

Worker Characteristics		Firm Characteristics				Sector	
TI ≤ Median	1.76 (0.205) [93,812]	**Corporate Employer**	1.30 (0.116) [156,957]	**# of Workers ≤ Median**	2.44 (0.237) [83,577]	**Agriculture**	0.70 (0.502) [1,927]
TI > Median	0.75 (0.092) [73,254]	**Individual Employer**	3.06 (0.642) [2,225]	**# of Workers > Median**	0.21 (0.069) [82,525]	**Mining**	0.25 (0.220) [5,685]
Age ≤ Median	1.28 (0.141) [57,570]	**Partnership Employer**	1.96 (0.326) [7,458]	**Sales ≤ Median**	1.81 (0.171) [83,876]	**Construction**	1.75 (0.472) [2,550]
Age > Median	1.38 (0.138) [52,697]	**Years Registered ≤ Median**	1.83 (0.169) [89,093]	**Sales > Median**	0.84 (0.096) [82,226]	**Manufacturing**	1.54 (0.176) [57,076]
Male	1.33 (0.123) [157,727]	**Years Registered > Median**	0.80 (0.100) [76,920]	**Size ≤ Median**	2.32 (0.218) [83,839]	**Trading**	2.65 (0.459) [3,985]
Female	2.63 (0.373) [5,884]	**Not VAT Registered**	1.87 (0.179) [33,180]	**Size > Median**	0.32 (0.073) [82,263]	**Services**	1.75 (0.162) [58,822]
Years Registered ≤ Median	1.33 (0.146) [84,987]	**VAT Registered**	1.20 (0.121) [130,827]			**Finance**	0.40 (0.115) [29,166]
Years Registered > Median	1.38 (0.122) [82,067]	**Not LTU**	3.07 (0.314) [61,406]			**Utilities**	0.50 (0.291) [5,796]
Not VAT Registered	1.17 (0.116) [75,801]	**LTU**	0.34 (0.065) [105,248]				
VAT Registered	5.58 (0.746) [2,587]						

Notes: The table shows estimated bunching of taxable income in various subsamples. For each subsample, the table shows the estimated normalized excess bunching mass b estimated as in section 4.1; the standard error of the estimate in round brackets, and the number of observations used for the estimation (those within 5% of a kink) in square brackets. A firm's (relative) size combines its sales and its number of employees by defining its size as the sum of its percentile in the distribution of number of workers and its percentile in the distribution of firm sales.

TABLE 5: PREVALENCE OF SALARY MISREPORTING

	Variable	Amount
Panel A: Underreporters (% of Workers)		
(1)	Employee $<$ Employer	19.3
Panel B: Underreported Salary Income (SI)		
(2)	Employee $<$ Employer (Rs. Bn)	15.6
(3)	Total Evaders' Employer Reported SI (Rs. Bn)	98.9
(4)	Total Employer Reported SI (Rs. Bn)	437.3
(5)	Employee Underreported SI (% of evaders' SI)	15.7
(6)	Employee Underreported SI (% of total SI)	3.6
Panel C: Underreported Tax Liability		
(7)	Employee $<$ Employer (Rs. Bn)	3.1
(8)	Total Evaders' Employer Reported Tax (Rs. Bn)	14.4
(9)	Total Employer Reported Tax (Rs. Bn)	60.6
(10)	Employee Underreported Tax (% of evaders' tax)	21.3
(11)	Employee Underreported Tax (% of total tax)	5.1

Notes: The table shows measures of underreporting of salaries based on discrepancies between employees' and employers' reports of workers' salaries. Panel A shows the remarkably high prevalence of discrepancies between the two reports. Row (1) shows the percentage of workers who report a salary at least 0.25% smaller than their employers using only individuals who have a single job in the employer statements. Panel B shows how much salary income is underreported. Row (2) sums the discrepancies, row (3) shows the total salary income reported by these individuals' employers, and row (4) shows the total salary income reported by all employers. Row (5) shows the extent of underreporting by evaders by dividing total underreported income (row (2)) by their employers' reported salary (row (3)). Row (6) shows the overall extent of underreporting by dividing underreported income (row (2)) by total reported salary income (row (4)). Rows (7)–(11) repeat this exercise converting the incomes into tax revenues assuming that the worker's salary is his/her taxable income and applying the tax schedule. Since most workers do not have any non-salary income, this approximation will be precise, and since most workers that *do* have non-salary income have positive non-salary income, this approximation will underestimate the effect due to the convexity of the tax schedule.

A Proofs

A.1 Proof of Lemma 1

Using (2), the distribution of taxable incomes is given by

$$J^*(z) = \begin{cases} \int_0^{z/\left[p^{1+\varepsilon}(1-\tau_0)^{\varepsilon}\right]} \int_0^{z/\left[p^{1+\varepsilon}(1-\tau_0)^{\varepsilon}\right]-\alpha} g(\alpha,\beta)\, d\beta d\alpha & \text{if } z < K_1 \\ \int_0^{z/\left[p^{1+\varepsilon}(1-\tau_1)^{\varepsilon}\right]} \int_0^{z/\left[p^{1+\varepsilon}(1-\tau_1)^{\varepsilon}\right]-\alpha} g(\alpha,\beta)\, d\beta d\alpha & \text{if } K_1 \leq z < K_2 \\ \int_0^{z/\left[p^{1+\varepsilon}(1-\tau_2)^{\varepsilon}\right]} \int_0^{z/\left[p^{1+\varepsilon}(1-\tau_2)^{\varepsilon}\right]-\alpha} g(\alpha,\beta)\, d\beta d\alpha & \text{if } K_2 \leq z \end{cases}$$

Which features bunching at the kinks. For example, at K_1, $J^*(K_1) = \int_0^{\bar{\delta}_1} \int_0^{\bar{\delta}_1-\alpha} g(\alpha,\beta)\, d\beta d\alpha$, while $\lim_{z\uparrow K_1} J^*(K_1) = \int_0^{\underline{\delta}_1} \int_0^{\underline{\delta}_1-\alpha} g(\alpha,\beta)\, d\beta d\alpha < J^*(K_1)$, and so $B_z(K_1) > 0$. A similar reasoning implies that $B_z(K_2) > 0$.

A.2 Proof of Lemma 2

Using (2), the salary distribution for workers without non-salary income is given by

$$H^*(s|n^*=0) = \begin{cases} G\left(\frac{s}{p^{1+\varepsilon}(1-\tau_0)^{\varepsilon}}, 0\right) & \text{if } s < K_1 \\ G\left(\frac{s}{p^{1+\varepsilon}(1-\tau_1)^{\varepsilon}}, 0\right) & \text{if } K_1 \leq s < K_2 \\ G\left(\frac{s}{p^{1+\varepsilon}(1-\tau_2)^{\varepsilon}}, 0\right) & \text{if } K_2 \leq s \end{cases}$$

which features excess bunching at the kinks as $\lim_{s\uparrow K_j} = G\left(\frac{K_j}{p^{1+\varepsilon}(1-\tau_{j-1})^{\varepsilon}}, 0\right) < G\left(\frac{K_j}{p^{1+\varepsilon}(1-\tau_j)^{\varepsilon}}, 0\right)$ for $j=1,2$. The excess bunching is given by $B_s(K_j|n^*=0) = G\left(\frac{K_j}{p^{1+\varepsilon}(1-\tau_i)^{\varepsilon}}, 0\right) - \left(\frac{K_j}{p^{1+\varepsilon}(1-\tau_{i-1})^{\varepsilon}}, 0\right)$.

Using (2) again, the salary distribution for workers with non-salary income is given by

$$H^*(s|n^*>0) = \int_0^s \int_0^\infty g(\alpha(s',\beta),\beta)\, d\beta ds'$$

where

$$\alpha(s,\beta) = \begin{cases} \frac{s}{p^{1+\varepsilon}(1-\tau_0)^{\varepsilon}} & \text{if } s < K_1 \,\&\, \beta \leq \underline{\delta}_1 - \frac{s}{p^{1+\varepsilon}(1-\tau_0)^{\varepsilon}} \\ \frac{\beta}{\frac{K_1}{s}-1} & \text{if } s < K_1 \,\&\, \underline{\delta}_1 - \frac{s}{p^{1+\varepsilon}(1-\tau_0)^{\varepsilon}} < \beta \leq \bar{\delta}_1 - \frac{s}{p^{1+\varepsilon}(1-\tau_1)^{\varepsilon}} \\ \frac{s}{(1-\tau_1)^{\varepsilon}} & \text{if } s < K_2 \,\&\, \bar{\delta}_1 - \frac{s}{p^{1+\varepsilon}(1-\tau_1)^{\varepsilon}} < \beta \leq \underline{\delta}_2 - \frac{s}{p^{1+\varepsilon}(1-\tau_1)^{\varepsilon}} \\ \frac{\beta}{\frac{K_2}{s}-1} & \text{if } s < K_2 \,\&\, \underline{\delta}_2 - \frac{s}{p^{1+\varepsilon}(1-\tau_1)^{\varepsilon}} < \beta \leq \bar{\delta}_2 - \frac{s}{p^{1+\varepsilon}(1-\tau_2)^{\varepsilon}} \\ \frac{s}{(1-\tau_2)^{\varepsilon}} & \text{if } \bar{\delta}_2 - \frac{s}{p^{1+\varepsilon}(1-\tau_2)^{\varepsilon}} < \beta \end{cases}$$

Since $\alpha(s,\beta)$ is continuous in both s and β, and since by assumption $g(\alpha,\beta)$ is continuous, $H^*(s|n^*>0)$ is continuous everywhere, including at K_1 and K_2.

A.3 Proof of Lemma 3

From (2), workers have taxable income K_2 whenever $\underline{\delta}_2 \leq \alpha_i + \beta_i \leq \bar{\delta}_2$. Therefore, for any $s < K_2$, the excess bunching mass of individuals with taxable income K_2 is

$$B^*_{K_2}(s) = \int_{\frac{K_2 - s}{p^{1+\varepsilon}(1-\tau_1)^{\varepsilon}}}^{\frac{K_2 - s}{p^{1+\varepsilon}(1-\tau_2)^{\varepsilon}}} g\left(\frac{\beta}{\frac{K_2}{s} - 1}, \beta\right) d\beta \tag{14}$$

Since the function $\beta / \left(\frac{K_2}{s} - 1\right)$ is continuous in β and s, and since $g(\alpha, \beta)$ is smooth by assumption, $B^*_{K_2}(s)$ is continuous in s at all s including K_1.

A.4 Proof of Prediction 2

First note that since real income choices are not distorted by the presence of evasion, and since n^* is increasing in β, individuals' non-salary income in equilibrium will be increasing in β and so we can erform comparative statics with respect to β. Applying the implicit function theorem to (3),

$$\frac{de^*_0}{d\beta} = \frac{\partial e\left(0, n - \hat{n}^*_0\right)}{\partial \beta} - \frac{\partial e\left(s - \hat{s}, n - \hat{n}^*_s\right)}{\partial \beta}$$

Since $e_n\left(0, n - \hat{n}^*_0\right) = e_n\left(s - \hat{s}^*, n - \hat{n}^*_s\right) = \tau$ and $\partial^2 e / \partial\left(s - \hat{s}\right)\partial\left(n - \hat{n}\right) > 0$, it must be the case that $n - \hat{n}^*_s < n - \hat{n}^*_0$. Then, since $\partial e / \partial \beta < 0$ and $\partial^2 e / \partial \beta \partial\left(n - \hat{n}\right) > 0$, the first part of the prediction follows. To see the second part, apply the implicit function theorem to the pair of first order conditions $e_s\left(s - \hat{s}^*, n - \hat{n}^*_s\right) = e_n\left(s - \hat{s}^*, n - \hat{n}^*_s\right) = \tau$ to see that

$$\frac{ds - \hat{s}^*}{d\beta} = \frac{e_{sn} e_{n\beta}}{e_{nn} e_{ss} - e^2_{ns}} > 0$$

where subscripts denote partial derivatives. The inequality follows from the convexity of e and the assumptions that $e_{sn} > 0$ and $e_{n\beta} > 0$.

A.5 Proof of Prediction 3

Equation (2) shows that in the absence of evasion, bunchers at kink $j = 1, 2$ are those for whom$\underline{\delta}_j \leq \alpha_i + \beta_i \leq \bar{\delta}_j$. For the case with evasion, assume for simplicity that $e_0 = 0$, and define

$$\begin{aligned} V\left(z, \hat{z}\right) &\equiv \max_{h,q,\hat{s},\hat{n}} U\left(c, l, q, \hat{s}, \hat{n}\right) \; s.t. \, p\left(l + q\right) = z \,\&\, \hat{s} + \hat{n} = \hat{z} \\ &= c - \frac{\left(\alpha + \beta\right)^{-1/\varepsilon}}{1 + 1/\varepsilon} z^{1+1/\varepsilon} - \tilde{e}\left(z - \hat{z}\right) \end{aligned}$$

as the maximal utility of earning z and reporting $\hat{z}$, where $\tilde{e}\left(z-\hat{z}\right)=e\left(\frac{\alpha}{\alpha+\beta}\left(z-\hat{z}\right),\frac{\beta}{\alpha+\beta}\left(z-\hat{z}\right)\right)$. Under a linear tax at rate τ, the optimal choices of $z\left(\tau\right)$ and $\hat{z}\left(\tau\right)$ satisfy

$$\begin{aligned}1-\left(\frac{z\left(\tau\right)}{\alpha+\beta}\right)^{\frac{1}{\varepsilon}}-\tilde{e}_{z}\left(z\left(\tau\right)-\hat{z}\left(\tau\right)\right) &= 0\\ -\tau+\tilde{e}_{z}\left(z\left(\tau\right)-\hat{z}\left(\tau\right)\right) &= 0\end{aligned}$$

where $\tilde{e}_z=\partial\tilde{e}/\partial\left(z-\hat{z}\right)$ implying that $z\left(\tau\right)=\left(\alpha+\beta\right)\left(1-\tau\right)^{\varepsilon}$ and that $\hat{z}\left(\tau\right)=z\left(\tau\right)-\tilde{e}_z^{-1}\left(\tau\right)$. Individuals who report taxable income at a kink are those for whom the optimal reported income under a linear tax at the lower rate below the kink is above the kink, and the optimal reported income under a linear tax at the higher rate is below the kink:[30]

$$\hat{z}\left(\tau_j\right)\leq K_j\leq\hat{z}\left(\tau_{j-1}\right)\ j=1,2$$

Solving this yields

$$\underline{\delta}_j+\frac{\tilde{e}_z^{-1}\left(\tau_{j-1}\right)}{\left(1-\tau_{j-1}\right)^{\varepsilon}}\leq\alpha_i+\beta_i\leq\bar{\delta}_j+\frac{\tilde{e}_z^{-1}\left(\tau_j\right)}{\left(1-\tau_j\right)^{\varepsilon}}$$

which is a larger range of $\alpha_i+\beta_i$ than in the case without evasion since $\tilde{e}$ is strictly convex. As long as the distribution of $\alpha_i+\beta_i$ is roughly uniform and/or the kink is small, this will mean a larger excess bunching mass at the kink. This continues to be the case when $e_0>0$, though the derivations are slightly more complicated.

A.6 Proof of Prediction 5

At salary level s there are $\delta f^*\left(s|n\neq0\right)$ individuals who have $\phi_i=0$ and have chosen salary s as their preferred salary. There are also $\left(1-\delta\right)f^o\left(s\right)$ individuals who have $\phi=\infty$ and happened to receive a salary offer at s. Among the unconstrained individuals, $B^*_{K_2}\left(s\right)$ choose taxable incomes at K_2 as defined in (14). Among the individuals who are constrained to accept a salary at s despite it not being their preferred salary, $\hat{B}_{K_2}\left(s\right)$ choose taxable incomes at K_2. Since these individuals are constrained in their choices by having a salary away from their preferred salary, fewer of them have taxable incomes that bunch at K_2. Among individuals with $\phi_i=0$, $B_z\left(K\right)$

Define the minimum disutility of achieving income z for individuals whose salary choice is unconstrained ($\phi_i=0$) as

$$\begin{aligned}V^*\left(z\right)\ \equiv\ &\min_{h,q}\frac{\left(\alpha+\beta\right)^{1-\frac{1}{\sigma}\left(1+\frac{1}{\varepsilon}\right)}}{1+\frac{1}{\varepsilon}}\left[\alpha\left(\frac{h}{\alpha}\right)^{\sigma}+\beta\left(\frac{q}{\beta}\right)^{\sigma}\right]^{\frac{1}{\sigma}\left(1+\frac{1}{\varepsilon}\right)}\\ &s.t.\,p\left(h+q\right)=z\end{aligned}$$

[30] This argument is made simpler by the absence of income effects under this parameterization of the utility function so that only the marginal tax rate matters. In the presence of income effects the comparison would still be of two linear taxes, but one would have to adjust the intercept of the tax schedule to account for the kink (see Saez, 2010 for further details).

which, solving, yields

$$V^{*}(z)=\frac{1}{1+\frac{1}{\varepsilon}}\left[\frac{z}{p(\alpha+\beta)}\right]^{\frac{1}{\varepsilon}} \tag{15}$$

Those who bunch at K_2 are those for whom $(1-\tau_2)\leq\frac{\partial V^{*}(z)}{\partial z}\leq(1-\tau_1)$, (solving this yields the fourth row of expression (2)). $\frac{\partial V^{*}}{\partial z}$ is strictly decreasing and continuous in α and β, so for each α there is an interval of values of β which lead the individual to bunch taxable income at K_2. In the unconstrained case, we have that

$$\frac{\partial^{2}V^{*}(z)}{\partial z\partial\beta}=-\frac{1}{\varepsilon}\frac{1}{\alpha+\beta}\frac{\partial V^{*}(z)}{\partial z} \tag{16}$$

For individuals constrained to earn a salary s ($\phi_i=\infty$), the minimum disutility of achieving taxable income $z\geq s$ is

$$\tilde{V}(z)=\frac{(\alpha+\beta)^{1-\frac{1}{\sigma}\left(1+\frac{1}{\varepsilon}\right)}}{1+\frac{1}{\varepsilon}}\left[\alpha\left(\frac{s}{\alpha}\right)^{\sigma}+\beta\left(\frac{z-s}{\beta}\right)^{\sigma}\right]^{\frac{1}{\sigma}\left(1+\frac{1}{\varepsilon}\right)}$$

and again, those who bunch are those for whom $(1-\tau_2)\leq\frac{\partial\tilde{V}(z)}{\partial z}\leq(1-\tau_1)$, i.e. those for whom

$$1-\tau_2\leq\left[\frac{\alpha}{\alpha+\beta}\left(\frac{s}{\alpha}\right)^{\sigma}+\frac{\beta}{\alpha+\beta}\left(\frac{z-s}{\beta}\right)^{\sigma}\right]^{\frac{1}{\sigma}\left(1+\frac{1}{\varepsilon}\right)-1}\left(\frac{z-s}{\beta}\right)^{\sigma-1}\leq 1-\tau_1$$

And differentiating we get that

$$\frac{\partial^{2}\tilde{V}(z)}{\partial z\partial\beta}=-\left\{\left[\frac{1}{\sigma}\left(1+\frac{1}{\varepsilon}\right)-1\right]\left[\frac{1}{\alpha+\beta}+\frac{\sigma-1}{\beta}\frac{\beta\left(\frac{z-s}{\beta}\right)^{\sigma}}{\alpha\left(\frac{s}{\alpha}\right)^{\sigma}+\beta\left(\frac{z-s}{\beta}\right)^{\sigma}}\right]+\frac{\sigma-1}{\beta}\right\}\frac{\partial\tilde{V}(z)}{\partial z} \tag{17}$$

When $\sigma=1$, this yields $\frac{\partial^{2}\tilde{V}(z)}{\partial z\partial\beta}=-\frac{1}{\alpha+\beta}\frac{1}{\varepsilon}\frac{\partial\tilde{V}(z)}{\partial z}=\frac{\partial^{2}V^{*}(z)}{\partial z\partial\beta}$, while when $\sigma=1+1/\varepsilon$, $\frac{\partial^{2}\tilde{V}(z)}{\partial z\partial\beta}=-\frac{1}{\beta}\frac{1}{\varepsilon}\frac{\partial\tilde{V}(z)}{\partial z}<\frac{\partial^{2}V^{*}(z)}{\partial z\partial\beta}$. It is straightforward to show that $\frac{\partial^{2}\tilde{V}(z)}{\partial z\partial\beta}$ is strictly decreasing in σ and continuous, so for each α the interval of β that leads to bunching is smaller in the constrained case than in the unconstrained case.

A.7 Proof of Prediction 6

In the presence of both information and mismatch effects, the probability of being a taxable income buncher at K_2 conditional on receiving a salary $x\neq K_1$ is

$$p(x)=\bar{\gamma}\frac{\delta f^{*}(x|n>0)\,B_{K_2}(x|\phi_i=0)+(1-\delta)\,f^{o}(x)\,B_{K_2}(x|\phi_i=\infty)}{\delta f^{*}(x|n>0)+(1-\delta)\,f^{o}(x)} \tag{18}$$

At K_1, there is firm bunching, so there are an additional $f^{0}(K_1)\,b_S$ workers with salaries at K_1 compared to x, $B_{K_2}(K_1|\phi_i=\infty)$ of whom have taxable incomes that bunch at K_2, where $f^{0}(K_1)$ is the counterfactual density at K_1 in the absence of firm bunching. This counterfactual density can be approximated by the observed density at x close to K_1, so that $f^{0}(K_1)\approx\delta f^{*}(x|n>0)+$

$(1-\delta) f^o(x)$. Combining these, and using the fact that x is close to K_1 so that $B_{K_2}(K_1|\phi_i=\infty) \approx B_{K_2}(x|\phi_i=\infty)$ and $B_{K_2}(K_1|\phi_i=0) \approx B_{K_2}(x|\phi_i=0)$,

$$\begin{aligned} p(K_1) &= \frac{(\bar{\gamma}+\Delta\gamma)}{(1+b)\left[\delta f^*(x|n>0)+(1-\delta) f^o(x)\right]}\Big\{\delta f^*(x|n>0) B_{K_2}(x|\phi_i=0) \\ &\quad + (1-\delta) f^o B_{K_2}(x|\phi_i=\infty) + b\left[\delta f^*(x|n>0)+(1-\delta) f^o(x)\right] B_{K_2}(x|\phi_i=\infty)\Big\} \end{aligned} \tag{19}$$

Combining (18) and (19) yields the result in equation (7)

B Appendix Figures and Tables

TABLE B.1: MERGING THE TAX RETURNS AND THE EMPLOYER STATEMENTS

Result	**2007/08**	**2008/09**	**2009/10**	**2010/11**	**2011/12**
Individual IT Returns	664,425	696,760	681,396	676,699	630,157
Returns with Salary >0	132,209	158,896	164,212	165,897	162,423
NTN & CNIC same on both	51,419	85,081	88,131	83,300	89,029
NTN Match; no CNIC on ES	11,027	1,492	957	665	110
CNIC Match; no NTN on ES	2,404	0	1	0	0
NTN Match; no CNIC on IT	274	13	103	24	11
Total Matched	65,124	86,586	89,192	83,989	89,150
Total Unmatched	67,085	72,310	75,020	81,908	73,273
Match Rate (%)	49.3	54.5	54.3	50.6	54.9

Notes: The table shows the outcome of the procedure used to merge the employer statements (ES) and the income tax returns (IT). Each IT record has a National Tax Number (NTN) identifier, and most also have a Computerised National Identity Card (CNIC) number. Most of the ES records also have at least one of these identifiers, though some have neither. Records are matched whenever the ES record and the IT records contain at least one matching identifier, and no conflicting identifiers. That is, a match occurs whenever a) both the NTN and the CNIC are the same; b) the NTNs are the same but either the ES or the IT record is missing a CNIC; or c) the CNICs are the same but the ES is missing the NTN. A match fails whenever a) the NTNs match, but the IT and ES records have conflicting CNICs; b) the CNICs match but the IT and ES records have conflicting NTNs, or c) when the ES record is missing both the NTN and the CNIC.

TABLE B.2: FIRM BUNCHING: ROBUSTNESS

Threshold	TI≠SI	TI≈SI; Unweighted	TI≈SI; PS Reweighted
	1.58	1.02	1.06
0%	(0.142)	(0.143)	(0.161)
	[67,457]	[92,086]	[91,320]
	1.62	1.02	1.32
1%	(0.139)	(0.143)	(0.176)
	[74,197]	[92,373]	[81,672]
	1.68	1.02	1.51
2%	(0.144)	(0.141)	(0.193)
	[75,828]	[92,378]	[80,019]
	1.71	1.02	1.58
3%	(0.145)	(0.140)	(0.196)
	[76,365]	[92,373]	[79,537]
	1.72	1.02	1.73
4%	(0.146)	(0.141)	(0.211)
	[76,601]	[92,345]	[79,958]
	1.72	1.03	1.88
5%	(0.146)	(0.139)	(0.227)
	[76,677]	[92,300]	[80,369]
	1.76	1.00	1.76
$\mathbf{K_{TI} \neq K_{TI}}$	(0.149)	(0.140)	(0.241)
	[76,864]	[91,667]	[78,553]

Notes: The table shows estimates of the normalized excess bunching in the distribution of scaled salaries, b, for different samples of workers, together with the estimate's bootstrapped standard error in round brackets, and the number of observations in the sample in square brackets. Each row uses a different threshold percentage difference between a worker's taxable income and his/her salary to define workers who have "significant" non-salary income or deductions. The first column shows the percentage used as a threshold. The second column shows the excess bunching amongst all workers who file a return. The third column shows the excess bunching amongst workers whose taxable income is approximately the same as their salary, and the final column shows aggregate bunching of salaries amongst workers whose taxable income is significantly different from their salary.