Behavioral Responses to Notches: Evidence from Pakistani Tax Records*

Henrik Jacobsen Kleven, London School of Economics

Mazhar Waseem, London School of Economics

Revised Version: March 2012

Abstract

Using administrative tax records from Pakistan, we investigate behavioral responses to notches created by discontinuous jumps in income tax liability. Notches introduce very strong incentives for bunching on the low-tax side and density holes on the high-tax side of cutoff points. We develop a method for estimating structural earnings elasticities with respect to the marginal tax rate using moments of the density distribution around notch points, and show that notches offer an ideal opportunity to estimate structural elasticities in a world where optimization frictions are important. We provide evidence of large and sharp bunching below notch points along with missing mass above notch points. Moreover, observed bunching is strongly attenuated by frictions as a large share of individuals are unresponsive even in strictly dominated regions above notches. The evidence of bunching and frictions is used to estimate long-run responses not attenuated by frictions. The implied earnings responses to notches are very large, but the underlying structural elasticities driving those responses are small. This finding shows the strongly distortionary nature of notched incentive schemes. We also present evidence on income shifting between wage income and self-employment income using notches, and consider a tax reform that facilitates a comparison between notches and kinks.

*We would like to thank Rashid Amjad, Tony Atkinson, Oriana Bandiera, Tim Besley, Michael Best, Richard Blundell, Raj Chetty, Estelle Dauchy, Roger Gordon, Adnan Khan, Wojciech Kopczuk, Ben Olken, Torsten Persson, Thomas Piketty, Dina Pomeranz, Imran Rasul, Emmanuel Saez, Monica Singhal, Joel Slemrod, Johannes Spinnewijn, and numerous seminar participants for helpful comments and discussions. Financial support from the International Growth Centre (IGC), Pakistan Programme is gratefully acknowledged. Corresponding author: Henrik Kleven, Department of Economics, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, United Kingdom. E-mail: h.j.kleven@lse.ac.uk.

1 Introduction

The historical expansion of fiscal capacity and government size in developed countries was facilitated by the ability to implement and enforce modern broad-based taxes such as income taxes and value-added taxes (e.g. Kleven, Kreiner, and Saez 2009). It is often argued that such taxes are central to the economic growth process, because they allow governments to provide basic public goods without destroying incentives and productive efficiency (e.g. Newbery and Stern 1987; Burgess and Stern 1993). Indeed, developing countries all over the world have started to replace traditional revenue sources by modern broad-based taxes (e.g. Keen and Lockwood 2010). Yet there is virtually no evidence, and in particular micro-based evidence, on the efficacy of such taxes within a less-developed country context. This paper takes an initial step to fill this gap by investigating behavioral responses to the income tax in Pakistan.

The context which we consider offers two key methodological advantages. First, the Pakistani income tax offers an unusual and compelling source of identifying variation. The system consists of a large number of tax brackets, each of which imposes its own *proportional* tax rate and therefore produces discontinuous jumps in tax liability—*notches*—at bracket cutoffs. Such notches create very strong incentives for bunching below cutoffs and density holes above cutoffs. By contrast, standard piecewise linear tax schedules are associated with discontinuous jumps in the *marginal* tax rate—*kinks*—which provide much weaker bunching incentives and do not produce holes. A recent literature using bunching around kink points to identify behavioral responses (Saez 2010; Chetty et al. 2011) finds that only very sharp and salient kinks create any bunching at all, and even there the response is modest. The likely explanation is the presence of optimization frictions such as switching and attention costs combined with the fact that the utility gain of bunching in response to jumps in marginal tax rates is not very large (Chetty 2011). Notches are different from kinks in two respects: (*i*) by creating much stronger variation than kinks, they are more likely to overcome frictions other things equal. (*ii*) Even if behavioral responses to notches are attenuated by frictions, the type of variation they produce offer an opportunity to directly estimate the degree of this attenuation bias. This allows for the separate estimation of short-run responses attenuated by frictions and long-run responses that overcome frictions.

Second, for this study, we have gained access to administrative records covering the universe of personal income tax filers in Pakistan over the period 2006-2008. While the use of large administrative tax records is emerging as the norm for public finance research on developed countries, such data have so far been unavailable for research on developing countries. For the study of bunching, it is crucial to use administrative data because survey data often feature too much measurement error to detect bunching (Hausman 1983; Saez 2010).

Our paper offers the following main contributions. First, we develop a method for estimating earnings elasticities using moments of the density distribution around notch points. Compared to the kink-based approach developed by Saez (2010), notches allow us to exploit two moments of the density distribution (bunching below and holes above notch points) instead of just one moment (bunching around kink points).[1] The advantage of this can be understood as follows. In the case of kinks, a situation with large structural elasticities and large frictions may produce the same amount of bunching as a situation with small structural elasticities and no frictions. Those two situations are then observationally equivalent, but have very different implications for long-run behavior and welfare. In the case of notches, on the other hand, the two situations may be associated with the same amount of bunching, but they would have different implications for the hole. We consider two approaches that exploit this difference to bound the structural elasticity. The first approach uses that notches create regions of strictly dominated choice above cutoffs where agents can increase both consumption and leisure by moving to the cutoff. Such regions should be completely empty in a frictionless world under any preferences, which implies that the observed density mass in strictly dominated regions can be used to measure attenuation bias from frictions. In particular, by scaling observed bunching using the hole in the dominated range, it is possible to obtain a lower bound on the true structural elasticity. The second approach is based on estimating the point of convergence between the observed and counterfactual densities above the notch point, and we show that this method provides an upper bound on the structural elasticity. Our paper therefore provides an alternative approach to bounding the long-run elasticity than the one developed by Chetty (2011).

Second, we provide clear graphical evidence of behavioral responses to notches. We consider wage earners and self-employed individuals separately as the two types are treated according to completely separate schedules. For the self-employed, we find large and sharp bunching below every notch combined with missing mass above every notch. For wage earners, we also find clear bunching and missing mass on each side of notches, but in smaller amounts than for the

[1] Only convex kinks create bunching around cutoffs; non-convex kinks create holes around cutoffs. Hence, kinks create either bunching or a hole, but not both at the same time.

self-employed. The smaller responses for wage earners are consistent with the notion that they have less flexibility to change their taxable income both because of less opportunity to evade and because of larger optimization frictions such as adjustment costs in labor supply.[2] We show that behavioral responses are strongly attenuated by optimization frictions as about 90 percent of wage earners in strictly dominated regions are unresponsive to notches, while 50 to 80 percent of self-employed individuals in such regions are unresponsive. This implies that, absent frictions, bunching would be 10 times larger than what we observe for wage earners and 2-5 times larger than what we observe for the self-employed. Interestingly, most of the difference in observed bunching between wage earners and self-employed individuals can be explained by differences in frictions, implying that differences in structural elasticities are relatively small.

Third, based on the estimates of bunching and missing mass around notches, we estimate the structural elasticity of earnings with respect to the marginal tax rate at different notches in the income distribution. Despite the finding of large and sharp bunching, the implied elasticities with respect to the marginal tax rate are modest (about 0.1-0.3 for the self-employed, smaller for wage earners). The combined findings of large bunching and small structural elasticities are not inconsistent: notches create extremely strong distortions and therefore induce large behavioral responses even under small elasticities. From a mechanism design perspective, this is indeed the key problem with notches. At the same time, the smallness of elasticities is striking because these are structural elasticity estimates not biased down by optimization frictions, and because these are taxable income elasticities in a context of weak enforcement. A couple of points are worth noting about this: (*i*) it is difficult to fully explain the result by the fact that these are elasticities for registered taxpayers, which may be a selected sample of relatively honest individuals in Pakistan. In particular, the findings by Kleven et al. (2011) show that tax evasion is high whenever detection probabilities are low even in a context of high tax morale. (*ii*) It is easier to explain the result by the fact that, according to the literature on tax evasion and tax enforcement (e.g. Andreoni et al. 1998; Slemrod and Yitzhaki 2002; Kleven et al. 2011), lax

[2] In general, bunching responses to discrete jumps in income tax rates (notches or kinks) facilitate estimations of the elasticity of taxable income. This elasticity concept has become the conventional measure of behavioral response to taxation in the public finance literature (see Saez, Slemrod, and Giertz 2010 for a recent survey) and encompasses the sum total of behavioral responses to taxation, including real responses and tax evasion. For wage earners, we distinguish between real responses and reporting responses by comparing bunching in the distributions of taxable income and third-party reported wage income, using the insight from the tax enforcement literature that income subject to third-party reporting tends to feature very little evasion (Kleven et al. 2011). This evidence suggests that about two-thirds of the response by wage earners is a real response.

enforcement leads to a large evasion *level* but not necessarily a large evasion *response* to the marginal tax rate. Hence, there is no *a priori* reason that the Pakistani tax notches would have to create large evasion responses.

Fourth, we provide evidence on shifting between wage income and self-employment income by considering a different type of notch. Each taxpayer is treated either as a wage earner or a self-employed person depending on whether the share of self-employment income in total income is below or above 50%, with tax rates on wage earners being substantially lower than on the self-employed. This income-composition notch creates a strong incentive to shift income between the two sources. We show that there is excess mass in the self-employment income share below this notch and missing mass above it, which provides compelling evidence of an income shifting response. While income shifting has been much discussed in the public finance literature (e.g. Slemrod 1995, 1998; Saez, Slemrod, and Giertz 2010), there is relatively little direct evidence on this type of response.

Fifth, we consider the effects of an unusual tax reform for wage earners in 2008. This reform combined the existing notched tax schedule with an alternative kinked schedule, and allowed taxpayers to choose between the two in order to minimize taxes. For a tax-minimizing individual, the reform effectively replaced each notch by two kinks, a convex kink and a non-convex kink, located at different income levels than the pre-existing notch. We find the following: (*i*) removing a notch eliminates bunching and missing mass on each side of the cutoff, (*ii*) introducing a convex kink sometimes creates small bunching, but this is not statistically significant, and (*iii*) introducing a non-convex kink creates no missing mass around the cutoff. These findings show the sharp difference between notches and kinks in generating behavioral responses.

Finally, we explore the nature of optimization frictions by investigating the determinants of strictly dominated choice. We find the following: (*i*) dominated choice is negatively associated with having prior experience as a tax filer, pointing to the importance of tax literacy and learning. (*ii*) Dominated choice is positively associated with dominated choice in the past, suggesting that some taxpayers consistently make mistakes while other consistently do not. We argue that this is likely to be explained by heterogeneity in factors such as misperception and inattention. (*iii*) Dominated choice is negatively associated with a number of indicators of taxpayer sophistication. Overall, these findings suggest that aspects of tax illiteracy such as misperception and unawareness of tax incentives are very important.

The conceptual and empirical approach to notches developed in this paper has the potential to be used in other contexts. Although notches have received relatively little attention from economists so far, they are not uncommon in the real world. Slemrod (2010) argues that notches are ubiquitous in tax systems around the world, and there are also many examples of notches in welfare programs, social insurance and social security. Some empirical studies have considered behavioral responses to notches in some of these other contexts, including the US Medicaid notch (Yelowitz 1995), social security notches (Gruber and Wise 1999), car taxation notches (Sallee and Slemrod 2010), and the UK in-work benefit notch (Blundell and Hoynes 2004; Blundell and Shephard 2011). Our paper takes a different approach than those papers by demonstrating how moments of the density distribution around notch points allows for the identification of the structural long-run elasticity in the presence of optimization frictions.

The paper is organized as follows. Section 2 sets out the theoretical framework and empirical methodology, section 3 describes the Pakistani income tax context and data, section 4 presents the empirical results, and section 5 concludes.

2 Theory and Empirical Methodology

2.1 A Model of Behavioral Responses to Notches

A recent literature in public finance (e.g. Saez 2010; Chetty et al. 2011) estimates behavioral elasticities using the variation created by discontinuities in marginal tax rates (kinks). Building on this idea, this paper develops a method for estimating behavioral elasticities using the variation created by discontinuities in tax liability (notches). We argue that notches may be better suited than kinks for identifying structural elasticities in a world where optimization frictions are important. This section sets out a theoretical model of behavioral responses to notches that will guide the empirical analysis. In the empirical section, we present both reduced-form evidence and a structural estimation of behavioral elasticities using the parametrized model presented here.

We start by analyzing the effect of notches on earnings choices at the intensive margin, assuming that the structural earnings elasticity is homogeneous around the notch and that there are no optimization frictions. We subsequently consider generalizations that allow for heterogeneous elasticities, optimization frictions, and extensive responses. Following the recent literature on bunching at kink points, consider individuals with quasi-linear and iso-elastic utility of the form

$$u = z - T(z) - \frac{n}{1+1/e} \cdot \left(\frac{z}{n}\right)^{1+1/e}, \tag{1}$$

where z is before-tax income, $T(z)$ is tax liability, and n is earnings capacity (ability). The quasi-linearity of utility rules out income effects of taxation such that the compensated and uncompensated elasticities are identical. The iso-elasticity of utility simplifies the exposition, but can easily be relaxed. As a baseline, we start by considering a linear tax system, $T(z) = t \cdot z$, where t is a proportional (average and marginal) tax rate. In this case, the maximization of utility with respect to before-tax income z yields the following income supply function

$$z = n(1-t)^e, \tag{2}$$

where e is the elasticity of earnings with respect to the marginal net-of-tax rate $1-t$. This is the structural parameter of interest as it serves as a sufficient statistic for tax revenue, Laffer rates, welfare, and optimal taxation. In the case of zero taxation, $t = 0$, equation (2) implies $z = n$ and therefore the ability parameter can be interpreted as potential earnings. The imposition of positive taxes depress actual earnings below potential earnings, with the strength of the effect determined by the elasticity e.

There is a smooth distribution of ability n in the population captured by a distribution function $F(n)$ and a density function $f(n)$. The combination of the ability distribution $F(n)$, $f(n)$ and the income supply function (2) yields a before-tax earnings distribution associated with the baseline linear tax system. We denote by $H_0(z)$, $h_0(z)$ the distribution and density functions for earnings associated with this baseline. Using (2), we obtain $H_0(z) = F\left(\frac{z}{(1-t)^e}\right)$ and hence $h_0(z) = H_0'(z) = f\left(\frac{z}{(1-t)^e}\right) / (1-t)^e$. Therefore, given a smooth tax system (no notches and no kinks), the smooth ability distribution converts into a smooth income distribution.

Suppose that a notch is introduced at the income cutoff z^*. A notch may be introduced either as a discrete change in tax liability at the cutoff with no change in the marginal tax rate on either side (a "pure notch") or as a discrete change in the proportional tax rate at the cutoff (a "proportional tax notch"). The latter form combines a pure notch with a discrete change in the marginal tax rate (a kink). The Pakistani tax schedule that we consider in the empirical application consists of jumps in the proportional tax rate, but in this conceptual analysis we allow for pure notches as well. The notched tax schedule can be written as $T(z) = t \cdot z + [\Delta T + \Delta t \cdot z] \cdot 1(z > z^*)$ where ΔT is a pure notch, Δt is a proportional tax notch, and

$1(.)$ is an indicator for being above the cutoff.

Figure 1 illustrates the implications of a proportional tax notch ($\Delta t > 0$, $\Delta T = 0$) in a budget set diagram (Panel A) and a density distribution diagram (Panel B). The notch creates a region of strictly dominated choice $\left(z^*, z^* + \Delta z^D\right]$ in which it is possible to increase both consumption and leisure by moving to the notch point z^*. There will be bunching at the notch point by all individuals who had incomes in an interval $(z^*, z^* + \Delta z^*]$ before the introduction of the notch, where the bunching interval is larger than the region of strictly dominated choice ($\Delta z^* > \Delta z^D$). Individual L has the lowest pre-notch income (lowest ability) among those who locate at the notch point; this individual chooses earnings z^* both before and after the tax change. Individual H has the highest pre-notch income (highest ability) among those who locate at the notch point; this individual chooses earnings $z^* + \Delta z^*$ before the tax change and is exactly indifferent between the notch point z^* and the best interior point in the upper bracket (point I) after the tax change. Every individual between L and H locates at the notch point. There is a hole in the post-notch density distribution as no individual is willing to locate between the notch point z^* and point I.

Under the notched tax schedule, we may distinguish between three groups of individuals. (i) Individuals with ability $n \leq z^*/(1-t)^e \equiv n^*$ are not affected by the notch and choose earnings according to (2). (ii) Individuals with ability $n \in (n^*, n^* + \Delta n^*]$ bunch at the notch point z^*. The bunching segment in the ability distribution corresponds to a bunching segment in the baseline earnings distribution $(z^*, z^* + \Delta z^*]$ according to the relationships $z^* = n^*(1-t)^e$ and $z^* + \Delta z^* = (n^* + \Delta n^*)(1-t)^e$. (iii) Individuals with ability $n > n^* + \Delta n^*$ locate in the interior of the upper bracket according to the income supply function $z = n(1 - t - \Delta t)^e$. This characterization encompasses situations with a pure notch ΔT due to the absence of income effects.

The basic idea in the empirical approach is that the width of the bunching segment Δz^* (corresponding to the earnings response of the marginal bunching individual) is determined by parameters of the tax notch and the structural elasticity e. Conversely, given knowledge of notch parameters and an estimate of the earnings response Δz^*, it is possible to uncover the elasticity e. To see how the method works, consider the marginal bunching individual at $n = n^* + \Delta n^*$, who is indifferent between the notch point and the best interior point in the upper bracket. At the notch point ($z = z^*$), this individual obtains utility

$$u_N = (1-t)\,z^* - \frac{n^* + \Delta n^*}{1+1/e}\left(\frac{z^*}{n^* + \Delta n^*}\right)^{1+1/e}. \tag{3}$$

At the best interior point ($z = (n^* + \Delta n^*)\,(1-t-\Delta t)^e$), the individual obtains utility

$$u_I = \left(\frac{1}{1+e}\right)(n^* + \Delta n^*)\,(1-t-\Delta t)^{1+e} - \Delta T \tag{4}$$

From the condition $u_N = u_I$ and using the the relationship $n^* + \Delta n^* = \frac{z^* + \Delta z^*}{(1-t)^e}$, we can rearrange terms so as to obtain

$$\frac{1}{1+\Delta z^*/z^*}\left[1 + \frac{\Delta T/z^*}{1-t}\right] - \frac{1}{1+1/e}\left[\frac{1}{1+\Delta z^*/z^*}\right]^{1+1/e} - \frac{1}{1+e}\left[1 - \frac{\Delta t}{1-t}\right]^{1+e} = 0. \tag{5}$$

This condition characterizes the relationship between the percentage earnings response $\frac{\Delta z^*}{z^*}$, the percentage jump in the average tax rate created by each type of notch $\frac{\Delta T/z^*}{1-t}$, $\frac{\Delta t}{1-t}$, and the elasticity parameter e. As we will estimate the earnings response directly using the observed earnings density around the notch, it is useful to view the relationship (5) as defining the elasticity e as an implicit function of $\frac{\Delta z^*}{z^*}$, $\frac{\Delta T/z^*}{1-t}$, and $\frac{\Delta t}{1-t}$. We denote this function by $e = g\left(\frac{\Delta z^*}{z^*}, \frac{\Delta T/z^*}{1-t}, \frac{\Delta t}{1-t}\right)$. It is not possible to obtain an explicit analytical solution for e, but it can be solved numerically given an estimate of Δz^* and observed values of the other arguments. Our empirical application will exploit proportional tax notches for identification, but pure notches or combinations of the two types of notches can also be used to identifying the elasticity using the characterization in (5).

There are three conceptual points to make about the elasticity formula (5). First, the formula applies only to the case of upward tax notches (increase in tax liability at the cutoff) where bunching is created by individuals coming from above the notch. The same formula cannot be applied to the case of downward tax notches (fall in tax liability at the cutoff) where bunching would be created by individuals coming from below. Using an approach analogous to the one above, it is possible to derive an elasticity formula for the case of downward notches that depends on the same basic variables as equation (5).

Second, as the compensated elasticity e converges to zero (Leontieff preferences), equation (5) implies

$$\lim_{e\to 0} \Delta z^* = \frac{\Delta T + \Delta t \cdot z^*}{1-t-\Delta t} \equiv \Delta z^D. \tag{6}$$

Hence, under Leontieff preferences, the bunching interval Δz^* converges to the strictly dominated range Δz^D in which taxpayers can increase both consumption and leisure by lowering earnings to the notch point.[3] The dominated range therefore represents a lower bound on the earnings response to notches under any compensated elasticity in this frictionless model.[4] The fact that notches create bunching even with a zero compensated elasticity represents a fundamental difference from kinks where an elasticity of zero means zero bunching.

Third, the above analysis encompasses the standard kink analysis as a special case. This can be seen by considering a combination of a pure notch ΔT and a proportional tax notch Δt that neutralizes the jump in tax liability at z^* and leaves only a jump in the marginal tax rate. This corresponds to the case where $\Delta T = -\Delta t \cdot z^*$ in which case the elasticity formula (5) collapses to the standard kink-based formula by Saez (2010) and the zero-elasticity result (6) becomes $\Delta z^* = 0$.

The determination of the structural parameter e using the characterization in (5) requires us to estimate the earnings response Δz^*. To facilitate this estimation, the model provides a simple relationship between the earnings response and estimable entities. Denoting excess bunching at the notch by B, we have

$$B = \int_{z^*}^{z^*+\Delta z^*} h_0(z)\, dz \approx h_0(z^*)\, \Delta z^*, \tag{7}$$

where the approximation assumes that the bunching segment $(z^*, z^* + \Delta z^*)$ is sufficiently small that the counterfactual density $h_0(z)$ will be roughly constant on this segment. A similar approximation is made in studies of bunching at kink points, but it is stronger in the context of notches because they create larger tax variation and therefore may create bunching responses over larger segments. We will estimate the counterfactual distribution in a flexible way that allows for curvature and therefore do not have to rely on the approximation, but it turns out that such curvature is not strong enough to have a significant impact on the results.

We now consider three extensions of the model that will be important in empirical applications: heterogeneity in elasticities, optimization frictions, and extensive responses. Figure 2 illustrates the effect of a notch on the density distribution under the baseline case (Panel A) as well as the various generalizations (Panels B-D). To simplify the exposition, the figure assumes that the notch is associated with a small jump in the *marginal* tax rate, so that intensive responses by those who stay above the notch are negligible. In this case, the pre-notch and

[3] The width of the dominated range Δz^D is defined such that the earnings level $z^* + \Delta z^D$ ensures the same consumption as the notch point z^*, i.e. $(1-t-\Delta t)\left(z^* + \Delta z^D\right) - \Delta T = (1-t)\, z^*$.

[4] This result is specific to upward tax notches. In the case of downward tax notches, a zero elasticity is associated with a zero earnings response.

post-notch densities coincide above the bunching segment $(z^*, z^* + \Delta z^*)$.[5] All of the notches we consider in the empirical application are associated with small marginal tax rate changes (but large tax liability changes) and therefore satisfy this approximation.[6]

First, we consider heterogeneity in the elasticity parameter. In this case, there is a joint distribution of abilities and elasticities described by the density function $\tilde{f}(n, e)$ on the domain $(0, \infty) \times (0, \bar{e})$. At each elasticity level, behavioral responses to the notch can be characterized as in the baseline case. The bunching segment for those with elasticity e is given by $(z^*, z^* + \Delta z_e^*)$, where Δz_e^* is increasing in e and takes the value Δz^D for $e = 0$. The post-notch earnings density in the full population will look like the solid blue curve in Panel B. The density is empty in the strictly dominated range and then increases gradually until it converges with the pre-notch density at $z^* + \Delta z_{\bar{e}}^*$. The grey shaded area in the post-notch density consists of those whose elasticity is too low for bunching given their location in the baseline earnings distribution.

In the case of heterogeneity, bunching can be used to estimate the average earnings response $E[\Delta z_e^*]$. Denoting by $\tilde{h}_0(z, e)$ the joint earnings-elasticity distribution in the baseline and by $h_0(z) \equiv \int_e \tilde{h}_0(z, e)\, de$ the unconditional earnings distribution, we have

$$B = \int_e \int_{z^*}^{z^* + \Delta z_e^*} \tilde{h}_0(z, e)\, dz de \approx h_0(z^*)\, E[\Delta z_e^*], \tag{8}$$

where the approximation is again based on the assumption of a constant counterfactual density on the bunching segment, $\tilde{h}_0(z, e) \approx \tilde{h}_0(z^*, e)$ for $z \in (z^*, z^* + \Delta z_e^*)$ and all e. Using equation (8), estimates of excess bunching and the counterfactual earnings density reveal the average earnings response.[7] The response estimated this way will be smaller than the response estimated from the point of convergence between the observed and counterfactual densities, which would reflect the response by the highest-elasticity individuals in the population.

Second, we consider the implications of optimization frictions driven by aspects such as switching and attention costs. In general, there are two possible implications of frictions. One is that some of the individuals who would move to the notch point in the absence of frictions

[5] We always refer to $(z^*, z^* + \Delta z^*)$ as the "bunching segment" (the inverval of the pre-notch earnings distribution over which individuals move to the cutoff), but under the approximation in Figure 2 it corresponds exactly to the "hole segment" (the interval of the post-notch distribution featuring a hole).

[6] It is not a contradiction to say that a proportional tax notch is large in terms of the jump in tax liability (the pure notch component) and at the same time small in terms of the jump in the marginal tax rate (the kink component). This is because, as a taxpayer crosses the cutoff z^*, the notch component changes tax liability by $\Delta t \cdot z^*$ while the kink component changes tax liability only by $\Delta t \cdot (z - z^*)$.

[7] Applying the the average earnings response $E[\Delta z_e^*]$ to the elasticity formula (5) provides only an approximation of the average elasticity in the population due to the nonlinearity of the elasticity formula.

stay above the notch. The other is that individuals who do respond may not be able to target the cutoff precisely, so that excess bunching is diffuse below the cutoff rather than a point mass exactly at the cutoff. In the empirical application, the first aspect turns out to be important (for example, there is significant density mass in strictly dominated ranges) while the second aspect is much less important (bunching is always very sharp). This suggests a model where responding to the notch is associated with a fixed adjustment cost, but conditional on incurring the adjustment cost individuals are able to control income quite precisely. This is the situation depicted in Panel C where adjustment costs create additional mass on the bunching segment $(z^*, z^* + \Delta z^*_{\bar{e}})$ compared to the frictionless model, but bunching still manifests itself as a sharp spike at the cutoff z^*. There is heterogeneity in adjustment costs (besides heterogeneity in abilities and elasticities), so that at each earnings-elasticity level some individuals respond and some do not. The light-grey area in the figure consists of those who do not respond because of low structural elasticities, while the dark-grey area consists of those who do not respond because of high adjustment costs.

A key distinction in this model is between the earnings response conditional on bunching, Δz^*_e at elasticity e, and the actual earnings response given frictions. We may interpret the first one as a long-run response that overcomes frictions and the second one as the short-run response attenuated by frictions. Existing micro studies of behavioral responses to taxation capture mostly short-run responses, but recent work tries to estimate long-run responses based on large tax variation and assumptions about the size of adjustment costs (Chetty 2011). A key advantage of notches is that they allow for an explicit estimation of both short-run and long-run responses using estimates of frictions that come directly from the data. We now describe two alternative approaches to deal with frictions, which provide lower and upper bounds on the long-run elasticity.

For the first approach, we denote by $a(z, e)$ the share of individuals at earnings level z and elasticity e with sufficiently high adjustment costs that they are unresponsive to the notch. We then have

$$B = \int_e \int_{z^*}^{z^* + \Delta z^*_e} (1 - a(z, e)) \tilde{h}_0(z, e) \, dz de \approx h_0(z^*) (1 - a^*) E[\Delta z^*_e], \tag{9}$$

where the approximation assumes a locally constant counterfactual density (as above) and a locally constant share of individuals with large adjustment costs, $a(z, e) = a^*$ for $z \in (z^*, z^* + \Delta z^*_e)$ and all e. In equation (9), $E[\Delta z^*_e]$ is the average long-run response not affected by frictions while $(1-a^*)\,E\left[\Delta z_e^*\right]$ is the average short-run response attenuated by frictions. Without knowledge of a^*, estimates of $B, h_0(z^*)$ can reveal only the the short-run response as in the previous literature using kink points. With an estimate of a^*, we can separately estimate short-run and long-run responses. Given the assumption that a locally constant share a^* are unresponsive due to frictions, this share can be estimated from the strictly dominated range where any remaining mass must be the result of frictions. Denoting by $h(z)$ the observed earnings density in the presence of the notch, we have $a^* = \frac{\int_{z^*}^{z^*+\Delta z^D} h(z)dz}{\int_{z^*}^{z^*+\Delta z^D} h_0(z)dz} \approx \frac{E[h(z)|D]}{h_0(z^*)}$ where $E\left[h(z)\,|D\right]$ is the average post-notch density in the dominated range.

This approach arguably provides a lower bound on the true structural response. To see why, notice that $a(z,e)$ is an endogenous variable that depends on the utility gain of moving to the notch point and the distribution of adjustment costs. If the distribution of adjustment costs is smooth, $a(z,e)$ would tend to be increasing over the bunching segment $(z^*, z^*+\Delta z_e^*)$ as the utility gain of moving to the notch point becomes smaller and so the minimum adjustment cost sufficient to prevent a response becomes smaller.[8] In this case, the approach based on estimating a^* from the strictly dominated range under-corrects for frictions and therefore under-estimates the long-run response. On the other hand, in a situation with dichotomous adjustment costs (either zero or prohibitively high) such that fixed shares of the population either do or do not respond, the approach would yield unbiased estimates of frictions and long-run responses.

The main innovation in the above approach compared to standard bunching approaches is to combine two moments of the distribution—bunching B and the hole in the dominated range $1-a^*$—to obtain a structural elasticity not attenuated by frictions. The method scales observed bunching by the inverse of the hole $1/(1-a^*)$ in order to measure structural bunching and the structural elasticity. This implies that the larger is observed bunching *and* the smaller is the hole, the larger is the structural elasticity. Moreover, comparing two situations with the same amount of observed bunching, the one with a smaller hole (more friction) will have a larger structural elasticity.

We also consider a second approach that provides an upper bound on frictions and the long-run response. Notice that to gauge the exact importance of frictions in attenuating bunching,

[8] For example, $a(z,e)$ will certainly be increasing over the bunching segment if the distribution of adjustment costs is independent of the earnings-elasticity distribution. Without such independence, it is theoretically possible to turn this around, but it would require implausible assumptions on the joint earnings-elasticity-friction distribution.

we would need to know how much of the observed mass on the segment $(z^*, z^* + \Delta z^*_{\bar{e}})$ can be explained by low elasticities in a frictionless world, corresponding to the light-grey area in Panel C. An upper-bound approach would be to assume that all of the mass on the segment $(z^*, z^* + \Delta z^*_{\bar{e}})$ is driven by frictions, not low elasticities. This corresponds to an assumption of homogeneous structural elasticities at $e = \bar{e}$. In this case, the structural long-run response can be determined as the point of convergence between the observed and counterfactual distributions. If there is heterogeneity in elasticities, this approach estimates the long-run response by the highest-elasticity individuals and is therefore an upper bound on the average long-run response. In the empirical application, we consider both the lower-bound approach (bunching scaled by $1/(1-a^*)$) and the upper-bound approach (convergence between $h(z)$ and $h_0(z)$) to estimating the structural elasticity, but in general the a^*-scaled bunching method is empirically much more robust than the convergence method.[9]

Finally, we consider a model with extensive responses. A conceptual difference between notches and kinks is that the former, by introducing a discrete jump in tax liability, may create both intensive and extensive responses. We consider a model with fixed costs of participation that are smoothly distributed in the population, and assume for simplicity that adjustment costs are the same for intensive and extensive responses. Earnings choices conditional on participation can be characterized as above, and all those with fixed costs of work below a critical value (that depends on tax liability) participates. In this model, the introduction of a notch induces some of the individuals who are close to the indifference point between working and not working to drop out of the labor market entirely rather than bunching at the cutoff. This pulls down the density distribution above the notch point as shown in Panel D of Figure 2. For those located slightly above the cutoff z^*, extensive responses will always be negligible because, conditional on incurring the adjustment cost, the notch point offers a utility almost as high as the utility obtained before the introduction of the notch, which by revealed preference dominates non-participation. As we move up along the segment $(z^*, z^* + \Delta z^*_{\bar{e}})$, the notch point becomes gradually less attractive compared to the situation before the notch in which case individuals who were originally close to the indifference margin between participation and non-participation prefer dropping out of the labor market to bunching at the notch point. Above the earnings

[9] The lack of robustness of the convergence method is not surprising, because it does not explicitly use excess bunching (which is the sharpest evidence of behavioral response in the data) and instead relies on estimating the counterfactual distribution over a large range above the notch point.

level $z^* + \Delta z^*_{\bar{e}}$, there are no longer any bunching responses, but extensive responses continue throughout the distribution.

It should be noted that notches always provide much stronger incentives for bunching (which depend on the implicit marginal tax rate created by the average tax rate jump) than for participation responses (which depend on the average tax rate jump per se). Hence, while extensive responses are theoretically possible, notches have to be very large to create clear extensive responses in empirical applications. In our application, extensive responses to notches are unlikely to be important for two reasons: (i) the average tax rate jumps are mostly between 0.5 and 2.5 percentage points and therefore provide weak variation in participation incentives, (ii) this variation occurs at the top tail of the income distribution (from around the 80th percentile and up) where we expect participation elasticities to be very small.[10]

2.2 A Reduced-Form Approximation of the Earnings Elasticity

The approach set out in the previous section relies on inferring the structural elasticity from an estimate of the earnings response (bunching segment) using a specific functional form for utility. Since the functional form assumption cannot be tested, it is also of interest to develop a reduced-form approach that does not make specific parametric assumptions about preferences. In the case of kinks, a reduced-form approach is straightforward as kinks create exogenous variation directly in the marginal tax rate of relevance to the structural elasticity. In the case of notches, a reduced-form approach is less straightforward as notches create exogenous variation in the average tax rate, which is not the parameter directly relevant for the structural parameter of interest. Here we set out a reduced-form approach for notches, which provides an approximation (upper bound) for the true structural earnings elasticity.

The basic idea in the reduced-form approach is to relate the earnings response Δz^* to the change in the implicit marginal tax rate between z^* and $z^* + \Delta z^*$ created by the notch. Considering a proportional tax notch, the implicit marginal tax rate t^* is given by

$$t^* \equiv \frac{T\left(z^* + \Delta z^*\right) - T\left(z^*\right)}{\Delta z^*} = t + \frac{\Delta t \cdot \left(z^* + \Delta z^*\right)}{\Delta z^*} \approx t + \frac{\Delta t \cdot z^*}{\Delta z^*}, \tag{10}$$

where the approximation requires that Δt is small (this approximation is not necessary, but simplifies slightly the elasticity formula below). The reduced-form elasticity of earnings with

[10] For example, if an individual is willing to participate in the labor market at the 90th percentile given a tax rate of 10%, this individual is unlikely to completely drop out of the labor market at a tax rate of 11%, especially when he has the option of avoiding the extra tax by moving to the notch point.

respect to the implicit net-of-tax rate is then defined as

$$e_R \equiv \frac{\Delta z^*/z^*}{\Delta t^*/\left(1-t^*\right)} \approx \frac{\left(\Delta z^*/z^*\right)^2}{\Delta t/\left(1-t\right)}. \tag{11}$$

This simple quadratic formula provides an alternative to the structural approach in the previous section. This formula essentially treats the notch as a hypothetical kink creating a jump in the marginal tax rate from t to t^*.[11]

Figure 3 illustrates the relationship between the reduced-form and structural approaches using a budget set diagram. The reduced-form formula (11) treats the response to the notch Δz^* as if it were generated by the kink shown by the intersection of the lower budget segment with the solid green line. As shown in the figure, this kink schedule includes interior points that are strictly preferred to the cutoff by the individual initially located at $z^* + \Delta z^*$, who would therefore not become a buncher if faced with this kink. In this case, the bunching response to the notch Δz^* overstates the bunching response that would be created by the kink Δt^*, implying that the reduced-form elasticity e_R constitutes an upper bound. The key reason why this is true in the figure is that the best interior point I is located on the left—or at least not too far on the right—of $z^* + \Delta z^*$ in which case the marginal bunching individual under the notch would not be willing to bunch under the hypothetical kink. This corresponds to an assumption that the uncompensated earnings elasticity is not too strongly negative, which is a fairly weak assumption. Since the assumption is satisfied for the quasi-linear case considered in the structural approach, the reduced-form elasticity (11) always provides an upper bound for the structural elasticity based on (5).

2.3 Estimating the Counterfactual Density and Bunching

The framework presented above allows us to evaluate the structural elasticity using an estimate of the earnings response Δz^*, which can be inferred from excess bunching B, the counterfactual density $h_0\left(z\right)$, and the actual density $h\left(z\right)$. This section describes how $B, h_0\left(z\right)$ can be estimated from the observed density $h\left(z\right)$.

Based on the model allowing for heterogeneous elasticities and frictions, an empirical density distribution around a notch may look like the one depicted in Panel A of Figure 4. Indeed, the

[11] Compared to the kink formula proposed by Saez (2010), the reduced-form notch formula differs in two respects: (i) the denominator is specified in terms of the average rather than marginal tax rate and (ii) the numerator is quadratic rather than linear in the percentage earnings response.

density distributions considered in the empirical application below have the same qualitative properties as the hypothetical density in Figure 4. This density features excess bunching right below the cutoff, missing mass right above the cutoff, but no completely empty hole in the strictly dominated range reflecting the presence of optimization frictions. In order to estimate the counterfactual (no-notch) density from the observed (notched) density, we start by defining an earnings range that is clearly affected by intensive responses to the notch and will be excluded in the estimation. The excluded range is the area between the dashed vertical lines at z_L, z_U and includes the bunching part below the cutoff and the missing-mass part above the cutoff. We then estimate the counterfactual by fitting a flexible polynomial to the observed distribution, using only observations outside the excluded range. The resulting counterfactual is illustrated in Panel B. A comparison between the observed and counterfactual distributions $h(z)$ and $h_0(z)$ provides measures of excess bunching B, missing mass M, and the share of individuals in the dominated range who are unresponsive because of frictions a^*. These estimates allow us to infer the earnings response and the structural elasticity using the characterization above. Notice that, because we use observations to the right of z_U in the estimation of the counterfactual density, it will incorporate potential extensive responses and our estimates therefore capture only intensive responses. However, as explained above, the notches we consider are not large enough and occur at too high income levels to induce non-negligible extensive responses.

We now describe the estimation of the counterfactual distribution in more detail. We distinguish between two cases: the standard case shown in Figure 4 where the empirical distribution features excess bunching only at notches and a case where the observed distribution features excess bunching both at notches and salient round numbers. The case of round-number bunching will be relevant in the empirical application for Pakistan.

Standard Case

The procedure used to estimate the counterfactual distribution is an extension of the approach by Chetty et al. (2011). We start by grouping individuals into small earnings bins (100 Pakistani rupee bins in the application below) and denote by c_j the number of individuals and by z_j the earnings level in bin j. We fit a polynomial of order q to the bin counts in the empirical distribution, excluding bins in the range (z_L, z_U), by estimating the following regression:

$$c_j = \sum_{i=0}^{q} \beta_i \cdot (z_j)^i + \sum_{i=z_L}^{z_U} \gamma_i \cdot \mathbf{1}[z_j = i] + \nu_j, \quad (12)$$

where γ_i is a bin fixed effect for each bin in the excluded range (so that the regression gives a perfect fit in that range). The counterfactual distribution is then estimated as the predicted values from (12) omitting the contribution of the dummies in the excluded range, i.e. $\hat{c}_j = \sum_{i=0}^{q} \hat{\beta}_i \cdot (z_j)^i$. Excess bunching is the difference between the observed and counterfactual bin counts in the part of the excluded range that falls below the notch point, i.e. $\hat{B} = \sum_{j=z_L}^{z^*} (c_j - \hat{c}_j)$. The share of individuals in the dominated range who are unresponsive due to frictions is estimated as $\hat{a}^* = E[c_j|D] / E[\hat{c}_j|D]$. Finally, the estimates of $\hat{c}_j, \hat{B}, \hat{a}^*$ are used to obtain estimates of the earnings response and structural elasticity using the characterization in sections 2.1-2.2.

Three points are worth noting about the implementation of this method. First, the parameters q, z_L, z_U are determined such that the counterfactual distribution looks credible given the observed distribution and a careful analysis of robustness to alternative parameter values will be carried out. Second, the application considered below involves many notches and we therefore carry out an integrated estimation across several notches simultaneously as this gives the most convincing and robust estimates. This implies that the specification (12) is implemented with many excluded ranges. Third, standard errors are calculated based on a bootstrap procedure, which generates N sets of counts $c_j^1, ..., c_j^N$ by drawing from the estimated vector of errors $\hat{\nu}_j$ in (12) with replacement and then applies the technique above to each set of counts so as to obtain a distribution of estimates for the different variables of interest. The standard error of each variable is the standard deviation of the distribution of estimates from the bootstrap.

Notch Bunching vs. Round-Number Bunching

We find that taxpayers have a tendency to report taxable income in round numbers, which creates mass points at round numbers in the empirical distribution. There are two observations to make about the anatomy of round-number bunching in the data. First, some round numbers are rounder than others: for example, while there is excess mass at any income level that is a multiple of 1k, there is stronger excess mass at multiples of 5k, 10k, 25k and 50k. Second, there is evidence of both annual and monthly rounding, the latter being a situation in which annual taxable income divided by 12 is a multiple of a round number. For example, an annual income of 240k is a round number in both annual an monthly terms while an annual income of 230k is round only in annual terms. These two points together implies that round-number bunching is strongest at income levels that can be represented as multiples of many salient round numbers (1k, 5k, 10k, 25k, 50k,...) in both monthly and annual terms.

There are three conceptual points to make about round-number bunching. First, there are no direct tax incentives to report income in round numbers: tax liability obviously does not depend on roundness and audit strategies also do not depend on this. We observe much more rounding among taxpayers who have mostly self-reported income (e.g. self-employed individuals) than among taxpayers with mostly third-party reported income (e.g. employees in formal firms), suggesting that rounding is a side-effect of poor record keeping. In the absence reliable records such as payslips from employers, it is difficult for the individual to reconstruct exact income in which case it is natural to report a round number. While this point holds for both honest and dishonest tax filers, the fact that rounding is correlated with absence of accurate income records implies that it may be correlated with tax evasion.

Second, since notches are themselves located at salient round numbers, implementing the specification (12) without controlling for rounding would confound true notch bunching with round-number bunching and therefore overstate behavioral responses to the notch. It does not matter what is the underlying explanation for rounding (and whether it is correlated with tax evasion or other behavioral margins), except that it is not driven by the notch per se. If a notch located at a salient round number were removed, some bunching would remain due to rounding.

Third, it is possible to control for round-number bunching at notches by using excess bunching at "similar round numbers" that are not notches as counterfactuals. In order to construct such round-number counterfactuals convincingly, we account for the underlying anatomy of rounding described above by estimating a rich set of round-number fixed effects that depend on the degree of roundness in both the annual and monthly dimension. This allows us to estimate round-number bunching at each notch point based on the underlying anatomy of the cutoff number.

As the estimation of the counterfactual density is based on 100 rupee bins, we need to worry only about rounding above the 100 rupee level. The regression specification we consider is the following

$$c_j = \sum_{i=0}^{q} \beta_i \cdot (z_j)^i + \sum_{k \in K} \sum_{r \in R, 12 \cdot R} \rho_{rk} \cdot \mathbf{1}\left[z_j \in k\right] \cdot \mathbf{1}\left[\frac{z_j}{r} \in \mathbb{N}\right] + \sum_{i=z_L}^{z_U} \gamma_i \cdot \mathbf{1}\left[z_j = i\right] + \nu_j, \qquad (13)$$

where $\mathbb{N}$ is the set of natural numbers, $R = \{1k, 5k, 10k, 25k, 50k\}$ is a vector of round number multiples that capture annual rounding, $12 \cdot R$ is a vector of round number multiples that capture monthly rounding (as z_j is defined as annual income), and K is a set of income brackets allowing us control for the fact that the tendency for rounding may change through the income distribution (there tends to be more rounding at the bottom than at the top). The parameter ρ_{rk} is the fixed effect associated with round number multiple r (e.g. $10k$) in income bracket k. The estimate of the counterfactual distribution is defined as the predicted values from the regression (13) omitting the contribution of the dummies around the notch, but not omitting the contribution of round-number dummies. As we will show, the estimated counterfactual distribution predicts round-number bunching in the observed distribution very precisely, suggesting that we have controlled for round numbers in a sufficiently rich way.

3 Context and Data

3.1 Pakistani Income Tax and Enforcement System

The personal income tax in Pakistan currently raises revenue of 1.1 percent of GDP, or 11 percent of total tax revenue, and the share of registered taxpayers in the working-age population is about 2 percent.[12] This level of income tax is obviously very small compared to advanced countries, but it is broadly consistent with the level in neighboring India (Piketty and Qian 2009) and in developing countries more generally (Gordon and Li 2009). Individuals not registered for income tax fall in two categories: (*i*) those who are *legally* unregistered either because their income is below the exemption threshold or because of other types of exemptions (the most important of which is the exemption of agriculture income), (*ii*) those who are *illegally* unregistered and operate in the informal sector. Although informality is an important issue in a country like Pakistan, the income exemption threshold (which is above the 70th percentile of the income distribution) and the exemption of agriculture (which constitutes almost half of the economy) can in fact explain the bulk of non-registrations. Outside of the exemptions, the personal income tax applies to all wage earners, self-employed individuals and unincorporated firms. The tax treatment of married couples is fully individual-based and the tax schedule differs slightly between men and women.

Two unusual features of the Pakistani income tax are exploited in the empirical analysis. First, the tax system consists of two separate schedules, one for wage earners and another one

[12] World Bank (2009) and authors' calculations.

for self-employed individuals and unincorporated firms, with substantially higher tax rates in the latter schedule than in the former. Each individual is classified either as a wage earner or a self-employed individual depending on whether self-employment income is below or above 50% of total income, and is then taxed according to the assigned schedule on the entire income. Because average tax rates are substantially different in the two schedules, this creates a strong discontinuity in tax liability at a self-employment income share equal to 50%. This *income-composition notch* allows us to estimate income shifting between wage income and self-employment income.

Second, the two income tax schedules consist of a large number brackets each of which is associated with a given proportional tax rate. Each bracket threshold therefore creates a discontinuity in the average tax rate. Figure 3 shows the average tax rate as a function of taxable income in Pakistani Rupees (PKR) in each schedule.[13] These are the schedules for male taxpayers and apply to years 2006-08 in the case of self-employed individuals and to years 2006-07 in the case of wage earners.[14] The tax rate on self-employed individuals increases from 0 to 25 percent over thirteen notches and the size of notches is increasing in income. The tax rate on wage earners increases from 0 to 20 percent over twenty notches, but we only show the first thirteen notches in the figure as the top-seven notches occur at extremely high incomes and affect very few people. We use these *income-level notches* to estimate the elasticity of earnings with respect to the marginal net-of-tax rate.[15]

The tax system in Pakistan is not indexed for inflation despite the fact that inflation is high in Pakistan (8-20% annually over the period we consider). A nominal tax system in a high-inflation economy is associated with substantial "bracket creep" as nominal wage growth pushes taxpayers into higher tax brackets over time.[16] To deal with this issue, the Pakistani income tax schedule has been subject to frequent tax reforms (approximately every third or fourth year) that increase bracket thresholds to compensate for inflation. While there was no such reform for the self-employed over the period considered here (2006-08), a tax reform was implemented for wage earners in 2008.

[13] The PKR-USD exchange rate is about 90 as of March 2012.

[14] The only difference in the tax schedule for females is that the first bracket threshold (exemption threshold) is located at a higher income level.

[15] The notches described above (income-level and income-composition notches) were introduced in Pakistan in 2004. Before that time, the income tax schedule featured standard kinks and the same schedule applied to wage earners and self-employed individuals. It is not feasible to directly analyze the impact of the 2004 reform, because no individual tax return data are available from that period.

[16] Such bracket creep is in itself a potentially useful source of identifying variation (Saez 2003), but we leave this aspect aside in the present paper.

Besides increasing bracket thresholds, the 2008 reform involved a more fundamental change in the tax treatment of wage earners. In particular, it combined the notch-based schedule with an alternative kink-based schedule and allowed wage earners to choose between the two in order to minimize taxes. As we describe later, the reform effectively replaced each notch by two kinks, a convex kink and a non-convex kink, located at a higher income level than the pre-existing notch. This unusual tax reform is very useful for two reasons: (*i*) it allows us to see what happens when a notch is eliminated, thereby providing a check on the empirical strategy used to identify behavioral responses, (*ii*) it allows us to explore the difference between notches and different types of kinks in creating responses.

Registered taxpayers are required to file income tax returns declaring taxable income, tax liability, and taxes owed net of withholding unless they meet certain filing exemption requirements.[17] The system operates on a pure self-assessment basis whereby the tax return is considered final unless selected for audit. Moreover, the enforcement system involves some third-party information reporting and withholding, the extent and form of which vary across taxpayer types. For most wage earners, there is third-party reporting and withholding by employers. As shown by Kleven et al. (2011), the crucial element in ensuring effective income tax enforcement is precisely third-party reporting, suggesting that tax enforcement is quite strong for the population of wage earners in Pakistan. By contrast, self-employed individuals and firms face no third-party information reporting, although they are subject to various withholding schemes. These are schemes that withhold taxes in connection with certain transactions (e.g. electricity bills, phone bills, and cash withdrawals), which are credited against income tax liability at the time of filing. This type of withholding comes with no third-party information on the tax base itself (taxable income), and is therefore not as powerful for enforcement as the more traditional system of third-party reporting and withholding in place for wage earners. The bottom line is that, for self-employed individuals and firms, tax evasion is deterred primarily by the threat of audits and penalties. As tax audits in Pakistan are rare, ineffective, and associated with frequent allegations of corruption, tax enforcement is very weak for this population.

[17] In particular, the following filing exemption rule will play a role in the empirical analysis: a wage earner is exempt from filing if (*i*) wage income is below 500,000 rupees, (*ii*) the employer has filed a tax return (third-party report), and (*iii*) the taxpayer has no other non-wage income and does not claim any deductions. For wage earners who do not file under this exemption, the third-party reported wage income declared by the employer is taken as the taxable income of the taxpayer. Notice that this creates a different form of notch at 500,000 rupees, as wage earners below this threshold do not have to incur personal filing costs. We investigate the effects of this filing notch in the empirical analysis.

3.2 Data

Our study uses administrative data from the Federal Board of Revenue (FBR) in Pakistan. The dataset includes the universe of personal income tax returns filed for tax years 2006/07, 2007/08 and 2008/09 (referred to as 2006-08), amounting to about 1 million tax returns per year.[18] For this period, returns were filed either electronically through the FBR website or by hard copy at designated bank branches and fed to computers using an IT firm distinct from FBR. This data collection process, introduced in 2006, implies that the data has much less measurement error than what is typically the case for an under-developed country. Variables in the data include taxpayer type (wage earner, self-employed, rentier, and unincorporated firm), a number of income variables, tax liability, taxes withheld, geographical area, and gender.

There are several points worth mentioning about the nature of the sample we consider. First, the universe of tax filers is not perfectly overlapping with the universe of registered taxpayers (national tax number holders) for two main reasons: (*i*) some registered taxpayers do not file returns either because of noncompliance or because of a filing exemption and (*ii*) some national tax numbers are invalid or inactive (World Bank 2009). Second, the population of tax filers is a higher-income subsample than the general population due to the high income exemption threshold: average annual income in our data is 175,000 PKR for self-employed individuals and about 560,000 PKR for wage earners, both substantially higher than the current GDP per capita of 85,000 PKR. Third, the population of individual tax filers in Pakistan is predominantly male (more than 99 percent of all filers). This is an implication of the individual-based tax system combined with the high exemption threshold as there are very few women in Pakistan with individual incomes above the threshold of the personal income tax. Fourth, a little less than half of all personal income tax filers are wage earners, about half of them are self-employed, and 2 percent are unincorporated firms.[19] Finally, the sample of tax filers may be a relatively tax-compliant subsample of the population. This means that the tax evasion level is smaller in this sample than in the general population, but it does not necessarily mean that the tax

[18] As described in section 3.1, some wage earners are exempt from filing provided that the employer has filed a third-party information report (and some other conditions are met), in which cases gross wage income reported by the employer is the taxable income of the individual. For taxpayers who opt for non-filing under these rules, the data contains income information from the third-party reports. Hence, the dataset contains all taxpayers for whom returns have been provided either by the individual or by the employer.

[19] These three categories do not quite add up to the full population of personal filers as a small share of taxpayers are listed as rentiers receiving most of their income from property, capital gains and other sources.

evasion response to the marginal tax rate is smaller. We come back to this point later.

4 Empirical Results

4.1 Self-Employed Individuals and Firms: Responses to Notches

We start by presenting bunching evidence for (male) self-employed individuals and unincorporated firms.[20,21] As discussed earlier, self-employed individuals show a strong tendency to report taxable income in round numbers, which leads to the presence of round-number bunching in the data. In fact, almost 60 percent of self-employed individuals report income in even thousands, and a disproportionate amount of those round-filers show a preference for the more salient round numbers such as multiples of 10 or 50 thousand. Since notch points are themselves located at salient round numbers, a failure to control for round-number bunching will lead to upward bias. Section 2 described an empirical strategy that controls for round-number bunching. In this section, we take a two-pronged approach to deal with rounding. First, we split the sample by those who report income in even thousands ("rounders") and those who do not ("non-rounders"). We separately analyze the non-rounder subsample, where the issue of round-number bunching is non-existent and we can implement the standard empirical specification (12) described above. Second, we consider the full sample of rounders and non-rounders, where we control for round-number bunching at notches using excess bunching at counterfactual round numbers that are not notches, using the empirical specification (13).

Figure 6 presents the empirical density distribution of taxable income for non-rounders around the six bottom notches in Panel A and around the four middle notches in Panel B.[22] The density graphs plot the number of taxpayers in different income bins, using a bin width of 1000 rupees in Panel A and 2500 rupees in Panel B. Each dot represents the upper bound of a given bin, so that a dot located at a notch point represents the number of taxpayers right below

[20] As described above, the tax schedule differs slightly between male and female taxpayers, and the two groups should therefore be analyzed separately. Since the population of personal income tax filers in Pakistan contains very few women, we are able to study bunching responses among female filers only around the mode of the earnings distribution. In this earnings range, we find that bunching responses by self-employed women are roughly similar (but somewhat smaller in magnitude) to the bunching responses by self-employed men shown in this section.

[21] We always consider self-employed individuals and firms together as results for the two groups separately are qualitatively very similar.

[22] Our findings are qualitiatively similar for the three top notches, which are located in the far upper tail of the income distribution. But we focus on the bottom and middle notches where the density mass is much larger and estimates are therefore more precise.

the cutoff. Each notch point is shown by a black vertical line and is itself part of the tax-favored side of the notch. Three main findings emerge from the figure. First, every notch is associated with large and sharp bunching immediately below the cutoff combined with a clear drop in the density immediately above the cutoff. This provides strong evidence of a response to the tax structure. Second, the degree of bunching is increasing in the size of the notch, which can be seen by comparing adjacent notches associated with different tax rate jumps: bunching at the 125k notch is stronger than at the preceding 110k notch (tax rate jumps of 1 and 0.5 percentage points, respectively) and bunching at the 300k notch is stronger than at the preceding 200k notch (tax rate jumps of 2.5 and 1 percentage points, respectively). Third, although the density always falls discretely after a notch and therefore features missing mass, there are no discernible holes in the distribution. In particular, there is substantial mass in strictly dominated regions where taxpayers could increase consumption and leisure by reducing earnings to the cutoff. The empirical distribution provides direct evidence of the importance of optimization frictions in attenuating bunching responses, and is consistent with the general theoretical model with frictions and heterogeneous elasticities presented in section 2.

Figure 7 compares the empirical distribution to an estimated counterfactual distribution that would apply in the absence of notches. We focus on the part of the income distribution that is above the three bottom notches (100k, 110k, and 125k) as this makes the analysis cleaner for two reasons: (*i*) the three bottom notches are very closely spaced, making it possible that taxpayers jump more than one notch at a time, (*ii*) the bottom notches are located around the mode of the distribution, which makes it difficult to obtain robust estimates of missing mass above cutoffs. By focusing on the part of the distribution that is monotonically declining and where notches are relatively far apart, the estimation of the counterfactual distribution, bunching, and structural elasticities is much more robust. The estimated counterfactual shown in the figure is based on fitting a ninth-order polynomial to the observed density, excluding the data around notch points, as specified in (12). Panel A in the figure shows the observed and counterfactual densities over the full income range in the estimation, with excluded ranges around notch points demarcated by vertical dashed-blue lines and notch points demarcated by vertical solid-black lines. The subsequent panels focus on smaller income ranges and specific notches. In those panels, bunching b is defined as excess mass in the excluded range below the notch in proportion to the average counterfactual frequency in that range, a^* is the proportion of taxpayers in the strictly dominated range who do not respond to the notch, and the vertical dashed-red line demarcates the upper bound of the dominated region.

The main findings in Figure 7 are the following. First, excess bunching is between 1.6 to 4.4 times the height of the counterfactual distribution and precisely estimated everywhere except at 600k. Second, there is missing mass over large earnings ranges above notch points, resulting from the discrete drops in the observed density on the right side of cutoffs. Third, despite the evidence of large bunching and missing mass, behavioral responses are strongly attenuated by optimization frictions: the share of individuals in strictly dominated ranges who are unresponsive to notches vary between 50 and 80 percent and is precisely estimated. Using this amount of friction to scale observed bunching (i.e., scaling by the inverse of missing mass in the dominated range, $1/(1-a^*)$) implies that bunching would be 2 to 5 times larger in the absence of frictions. As argued in section 2, this method is likely to provide a lower bound on true structural responsiveness.

Figure 8 turns to the full population of rounders and non-rounders, focusing on the income range above 200k. Although bunching in the full sample is extremely strong at lower earnings levels, it is difficult to control for round-number bunching at the bottom where all the most salient round numbers are notch points. Hence, there are no good counterfactual round numbers in the vicinity of the lower notches, and so implementing the specification (13) in that range is likely to under-control for rounding and lead to upward bias. Panel A in Figure 8 shows the empirical density distribution for the full sample over the income range used in the estimation, Panel B adds the estimated counterfactual density with round-number fixed effects, and Panels C-F zoom in on specific notches.

The findings for the full population are the following. First, the raw distribution features very large excess mass at notch points, but it also features smaller mass points at other places in the distribution. All of the non-notch mass points occur exactly at round numbers (for example, the strongest one in Panel A is at 250k). The counterfactual density accurately captures mass points at round numbers that are not notches, and it also features mass points at notches as they are always located at round numbers. Second, excess bunching b is a bit smaller in the full sample than in the round-number sample, while frictions a^* are a bit larger in the full sample. Hence, while observed bunching is smaller for rounders than for non-rounders, long-run bunching adjusted for frictions is about the same for the two groups. Notice that there were no *a priori* reason to expect either smaller or larger responses among rounders because of two offsetting effects: (*i*) As discussed earlier, rounding is likely to be a side-effect of the absence of accurate business records, which is associated with larger evasion opportunities and potentially larger evasion responses to notches. (*ii*) Rounding is a signal of lower taxpayer sophistication and therefore larger frictions (from misperception, inattention, etc.), which is associated with smaller observed bunching responses. Our findings suggest that the second effect dominates and create smaller observed bunching among rounders, but that the structural responsiveness is about the same for rounders.

We now turn to the estimation of earnings elasticities with respect to the marginal net-of-tax rate, combining the evidence above with the theoretical framework in section 2. As shown in that section, it is possible to estimate such earnings elasticities from notches by estimating the earnings response Δz^* of the marginal buncher (or the average earnings response $E[\Delta z^*]$ across all preference types) and applying the structural relationship (5) or the reduced-form approximation (11). Our objective is to estimate the structural elasticity parameter e, which governs behavioral responses absent optimization frictions. We put bounds on this parameter using the two approaches described in section 2: a lower bound is obtained by scaling observed bunching based on frictions in the dominated range (as measured by the inverse of the hole $1/(1-a^*)$) and an upper bound is obtained by estimating the earnings response as the point of convergence between the counterfactual and observed distributions. The results are presented in Table 1 for the non-rounder sample in Panel A and the full sample in Panel B. The table shows the notch point in column (1), the average tax rate jump in column (2), the dominated range in column (3), earnings responses in columns (4)-(5), elasticities based on the structural model in columns (6)-(7), and elasticities based on the reduced-form approximation in columns (8)-(9).

The main findings in Table 1 are the following. First, estimated earnings responses are very large at all notches, which reflects the combined findings of large bunching *and* a small hole (large frictions). The response varies between 5 and 15 percent of earnings across most notches, which is an extremely large response to tax incentives. This reflects the strong distortions introduced by notches. Second, the structural elasticities driving those large earnings responses are in general modest. The lower-bound elasticities fall mostly in the interval 0.05-0.20 while the upper-bound elasticities fall mostly in the interval 0.10-0.30, except at 200k where the upper bound is very large (but also has a large standard error).[23] The smallness of elasticities obviously cannot be explained by frictions as the methodology explicitly controls for frictions. The combined findings of large bunching responses and small structural elasticities crystallizes the key mechanism design problem with notches, and suggests that alternative tax structures could do much better. Third, estimates of short-run effects that include the effect of frictions can be obtained by multiplying the estimates in the table by the share of responders $1-a^*$ (equal to 0.2-0.5 across the different notches as shown in the previous figures). This exercise implies average short-run elasticities extremely close to zero. Fourth, the reduced-form estimates in the last two columns (which are always upper bounds on the structural estimates as shown in section 2.2) are not very far from the structural estimates, showing that the much simpler reduced-form approximation is useful.

The smallness of elasticities in Pakistan is *a priori* surprising for two reasons: (*i*) the elasticities are identified from very large tax variation and control for frictions, (*ii*) these are elasticities of taxable income (including real and evasion responses) among self-employed individuals and small firms in a context of weak tax enforcement. Let is briefly discuss two possible explanations. First, the population of tax filers in Pakistan is likely to be a selected subsample of taxpayers with high tax morale and therefore low evasion. A weakness of this interpretation is that the findings of Kleven et al. (2011) from a context with high tax morale (Denmark) shows that, whenever detection probabilities for evasion are low, taxpayers engage in large tax evasion. Second, it is important to distinguish clearly between the evasion *level* and the evasion *response* to a higher marginal tax rate. A weak enforcement system will be associated with large tax evasion, but not necessarily a large response of tax evasion to the marginal tax rate. For example, in a benchmark model with a linear income tax, a linear penalty in the evaded tax and an exogenous audit probability, the substitution effect of the marginal tax rate on evasion is always zero (e.g. Andreoni et al. 1998; Slemrod and Yitzhaki 2002). In the presence of nonlinear penalties (including side payments to inspectors), nonlinear taxes and endogenous audit probabilities, the substitution effect is non-zero but the sign is in general ambiguous (e.g. Kleven et al. 2011). Hence, there is no reason to believe that a weak enforcement system is necessarily associated with large evasion responses to notches. Kleven et al. (2011) study this

[23] The convergence approach (upper bound) in Table 1 defines convergence as the first earnings point after the notch where the difference between the observed and counterfactual densities is less than 5% over a 5k interval above the point.

question empirically using Danish audit data and a large kink in the income tax schedule, and finds that for self-employed individuals (whose evasion level is high) the evasion response to the marginal tax rate is very small. Hence, while small elasticities of taxable income in Pakistan are somewhat surprising, they are not inconsistent with the presence of weak enforcement.

4.2 Wage Earners: Responses to Notches and Kinks

As described earlier, the tax treatment of wage earners was changed fundamentally in 2008 from the notch-based schedule shown in Figure 3 to a complicated kink-based schedule. We start by presenting evidence on behavioral responses to the notched schedule in years 2006-07, and then turn to the effects of the 2008 tax reform.

We first make two points about the sample and notches we consider. First, the issue of round-number bunching is much less important for wage earners than for self-employed individuals, as discussed in section 2. Only about 8 percent of wage earners report income in even thousands as opposed to 60 percent of self-employed individuals. The limited amount of rounding that we do observe for wage earners mostly takes the form of monthly rounding (as job contracts are negotiated in monthly rather than annual terms) and is therefore a concern only at a few notch points. Given that rounding affects a small fraction of wage earners, we focus on the non-rounder sample throughout this section as the graphs and estimation are more straightforward for this group. Considering the full sample does not change the qualitative conclusions. Second, the tax schedule for wage earners has twenty notches, but we do not consider all of them here. In particular, the density distribution is too noisy around the bottom notches (in part because of filing exemptions at the bottom) and around the top notches (because of very few observations at the extreme top) to estimate bunching responses precisely, and we therefore focus on a set of middle notches where the density distribution is sufficiently clean to obtain robust estimates. As those middle notches apply to both male and female taxpayers, we show evidence for the two groups together.

Figure 9 shows density distributions of taxable income (green graph with dots) and gross wage income (grey graph with triangles) for all wage earners in 2006-07 around six middle notches demarcated by black vertical lines. Taxable income equals gross wage income plus non-wage income (self-employment income, capital and property income) minus deductions. As wage income is third-party reported for this sample and therefore difficult to successfully under-report

(Kleven et al. 2009, 2011), tax evasion by wage earners will occur mostly in the non-wage and deduction component of taxable income.[24] Hence, following Chetty et al. (2011), we consider bunching in gross wage income around statutory cutoffs (defined with respect to taxable income) to detect real responses. A bunching individual with any non-wage income or deductions will be mechanically displaced from the notch when considering the distribution of wage income alone, and hence bunching in this distribution is likely to reflect mostly real responses.[25] We can compare this to bunching in taxable income, which reflects the total behavioral response including evasion.

The following findings emerge from the figure. First, the distribution of taxable income features sharp bunching below every notch along with missing mass above every notch. Unlike the findings for self-employed individuals, there is evidence of clear holes above some of the notches (especially at 700k). Second, as one would expect, bunching is not as strong for wage earners as it is for self-employed individuals, although it should be noted that the notches for wage earners are associated with smaller tax rate jumps than for the self-employed. Third, there is clear evidence of bunching even in the distribution of third-party reported wage income, which suggests that there is a real labor supply response to the tax structure. Bunching in wage income is about two-thirds of bunching in taxable income. Finally, the evidence of bunching at the 500k notch should be interpreted with some caution. As described in section 3, wage earners are exempt from the filing requirement if wage income is below 500k *and* a third-party information report has been submitted by the employer. For taxpayers who opt for non-filing under these rules, taxable income is given by third-party reported gross wage income. Since income tax filing is not costless to the taxpayer (due to for example administrative hassle), these rules introduce a *filing notch* and create an additional incentive for bunching below 500k. Hence, bunching at this particular cutoff possibly combines responses to the tax rate notch and the filing notch. We will be able to explore this below, taking advantage of the fact that the 2008 tax reform removed the tax rate notch—but not the filing notch—at 500k.

Figure 10 compares the observed density to an estimated counterfactual density obtained

[24] We cannot completely rule out that there is some evasion in gross wage income even though it is third-party reported, for example because of collusion between the taxpayer and the third party (see Kleven et al. 2009 for a study of this question).

[25] Arguably, this is a lower bound on real responses as individuals with non-wage income and deductions who are bunching in taxable income (but not in third-party wage income) could have achieved bunching using partly real wage income.

from specification (12) assuming a ninth-order polynomial. Panel A shows the empirical distribution along with excluded ranges in the estimation of the counterfactual, Panel B adds the estimated counterfactual to the picture, while Panels C-F zoom in on specific notches and show estimates of excess bunching b and frictions in the dominated range a^*. For the notches in the figure, excess bunching is between 50% and 80% of the height of the counterfactual density and precisely estimated. Moreover, observed bunching is strongly attenuated by frictions as about 90% of individuals in strictly dominated ranges are unresponsive to notches. The fact that frictions are larger for wage earners than for self-employed individuals is consistent with the idea that adjustment costs in labor supply matter more for wage earners. Scaling bunching based on the amount of friction in the dominated range implies that, absent frictions, bunching by wage earners would be 10 times larger than what we observe.

Table 2 presents estimates of earnings responses to notches and the implied structural elasticities, using the method of scaling bunching by $1/(1-a^*)$.[26] The key findings in the table are the following. First, earnings responses are between 2 and 5 percent of earnings across all notches and strongly significant. Second, those earnings responses are driven by very small structural elasticities, around 0.05 or lower. Although the structural elasticities are smaller for wage earners than for self-employed individuals, the difference is small compared to the difference on observed bunching between the two groups. This suggests that the difference in short-run behavioral responses by wage earners and the self-employed has more to do with frictions than with preferences. Third, the estimates provide a clear illustration of the implications of using two moments of the density distribution to estimating the structural elasticity. Notice that the structural elasticity is particularly tiny at 700k, although this notch features large bunching compared to the other notches. The key reason is that this notch is associated with a larger hole in the dominated range and therefore with less friction. Because of the clear hole above this notch, observed bunching can be almost fully explained by people moving from within the dominated range (the earnings response is about 12k compared to a dominated range of about 11k) and therefore the compensated elasticity is extremely close to zero.

We now turn to the analysis of an unusual tax reform in 2008, which allows us to compare the effects of notches and kinks. Rather than replacing the notched schedule by a kinked schedule, the reform introduced a kink option in addition to the pre-existing notched schedule and left it

[26] We do not consider the convergence method here, because this method does not yield very robust results for wage earners.

to the taxpayer to self-assess tax liability in a tax-minimizing fashion. The effect of the reform on the budget set is illustrated if Figure 11. In Panel A of this figure, the solid blue curve shows the pre-reform budget set featuring a notch at z^*. Following the reform, a kink option (shown by the dashed red curve) with a marginal tax rate t_k above the cutoff was introduced. A kink alternative of this sort was introduced at every notch, with the marginal tax rate t_k varying across different ranges of the income distribution (from 20% at the bottom to 60% at the top). Because the marginal tax rate above the cutoff is much higher under the kink than under the notch ($t_k >> t + \Delta t$), it eventually becomes optimal to switch back to the notch schedule as income increases sufficiently above the cutoff (where the solid and dashed curves intersect in Panel A). Therefore, and as shown in Panel B, the post-reform budget set for a tax-minimizing individual features two kinks for every notch: a convex kink at z^* and a non-convex kink at z^{**}. Hence, the tax reform effectively replaced the twenty notches by forty kinks, half of which are convex (and should produce bunching) and half of which are non-convex (and should produce holes). At the same time, bracket cutoffs were increased such that each pair of kinks were located above the pre-existing notch.

Two aspects of this reform are worth emphasizing. First, the post-reform tax system is very complex, and it requires a high degree of taxpayer sophistication to take advantage of the incentives created by the system. While the reform replaces notches by kinks and hence eliminates strictly dominated regions *in principle*, it leaves this entirely to taxpayers' self-assessments. As those self-assessments are corrected only in the case of audits, it is possible for taxpayers to be on the notch schedule when they should be on the kink schedule, and vice versa. We find evidence of such behavior in the data: among taxpayers with incomes in between the convex and non-convex kinks (all of whom should pay tax according to the kink schedule), about 20% self-assess their tax liability according to the notch schedule. The presence of such misperception suggests that potential bunching at the newly established convex kinks could reflect a response to a perceived notch. Second, under the new tax schedule, the convex kink and non-convex kink are very different in terms of salience. The convex kinks are located at statutory cutoffs in the tax system (all of which are round numbers as in the pre-existing notch schedule) and are therefore relatively salient. The non-convex kinks, on the other hand, are located at intercepts between the two underlying schedules, which is endogenous to all the different tax parameters $t, \Delta t, t_k, z^*$. Hence, the non-convex kinks are much less salient than the convex kinks.

We present evidence on the effects of the reform in Figure 12, which shows observed and counterfactual density distributions for wage earners after the reform around notches that were eliminated and kinks that were introduced by the reform. Before-reform notch points are shown by solid vertical lines and after-reform *convex* kink points are shown by dashed lines. The figure does not demarcate the non-convex kink points as they never produce any discernible response (i.e., no density holes). Our findings are the following. First, there is no longer bunching at notch points that were eliminated by the reform, except possibly at 500k where a very small (but statistically significant) amount of bunching remains. As discussed above, this specific cutoff represents both a tax rate notch (before the reform) and a filing exemption notch (before and after the reform), and hence post-reform bunching at 500k identifies a small behavioral response to the filing notch.[27] For all other cutoffs, the evidence in the figure confirms the identifying assumption that, in the absence of discontinuities in tax rates, the income distribution would be smooth around the cutoff. Second, the introduction of convex kinks seems to create small bunching, although this is never statistically insignificant. Third, the form of bunching at convex kinks suggests that some taxpayers respond to those kinks as if they were still notches: there is excess bunching below the cutoff but not above, whereas in principle we would expect kinks to create bunching *around* the cutoff. This points to misperception of incentives even among those who do respond to the tax system. Overall, the figure confirms the identification strategy used elsewhere in the paper and also shows that kinks create little or no bunching, despite the fact that all of the kinks are associated with very large marginal tax rate changes.

4.3 Shifting Between Self-Employment and Wage Income

Having studied behavioral responses among wage earners and self-employed individuals separately, we now turn the question of shifting between the two. While income shifting has been much discussed in the literature, there is relatively little direct evidence on this type of behavioral response. As identifying variation, we use the income-composition notch described earlier: each taxpayer is treated either as a wage earner or a self-employed individual depending on a cutoff rule applying to the composition of income, with tax rates on the latter being much higher than on the former. In particular, if the share of self-employment income in total income is greater than or equal to 50%, the individual is treated as self-employed. Otherwise, the indi-

[27] For this particular cutoff, the elasticity estimates in Table 2 include the filing exemption response and are therefore (slightly) upward biased.

vidual is treated as a wage earner. This creates a large notch at an income share of 50% and provides very strong incentives to shift income from self-employment to wages in order to be treated as a wage earner.

We provide evidence on income shifting responses in Figure 13, which shows the density distribution of the self-employment income share between zero and one in bins of 0.025. Taxpayers are treated as wage earners and face a low tax rate on the left side of the vertical line at 0.50, and are treated as self-employed and face a high tax rate on the right side of the vertical line. Importantly, unlike the density diagrams shown earlier, the notch point itself belongs to the high-tax region and we would therefore expect to see bunching strictly below the notch rather than at the notch. To evaluate this, each bin excludes the upper bound of the interval such that the bin right below the notch spans $[0.475; 0.50)$ and the bin right above spans $[0.50; 0.525)$.

Panel A shows the raw density distribution of the self-employment income share. While this panel shows a clear effect on the density distribution, it is puzzling that bunching seems to occur *around* the notch rather than below it. Interestingly, it turns out that almost all of the taxpayers in the first bin above the notch have a self-employment income share *exactly* equal to the cutoff value of 0.50. This points to two possible explanations. The first possibility is that this is a form of "round-number bunching" as analyzed earlier. These taxpayers may not know the exact composition of their income into wage income and self-employment income and therefore naturally report the same amount in the two cells on the self-assessment. A second possibility is that these taxpayers are actually trying to game the system by bunching at the notch point, but are unaware of the subtlety that the notch point itself (unlike all other notches in the tax system) belongs to the high-tax range.

Whatever the explanation for taxpayers locating at an income share of precisely 0.50, an empirical test of income shifting responses to the notch can be obtained by slicing out those specific taxpayers from the distribution. That is, if there were no response to the notch, the density distribution would be smooth around the notch even if we drop observations at one specific value. Hence, Panel B shows the density distribution when dropping taxpayers with a self-employment income share equal to 0.50, and this diagram shows a very clear income shifting response to the notch: there is bunching immediately below the notch along with a very sharp drop in the density immediately above the notch. The missing mass above the notch is much larger than excess bunching below the notch, suggesting that individuals move from the upper half of the distribution to the interior of the lower half, possibly because of optimization frictions. This makes it impossible to estimate the true size of the shifting response directly based on excess bunching. In this case, a better approach would be to estimate the shifting response based on the point of convergence between the observed density and a counterfactual density above the notch point. Such an approach would yield very large shifting effects as measured by the absolute change in the self-employment income share. However, the large absolute shifting effect does not convert into a large shifting elasticity, because the income-composition notch considered in this section is much larger than the income-level notches considered earlier.[28]

4.4 Strictly Dominated Choice and the Nature of Optimization Frictions

A central feature of (upward) tax notches is that they create strictly dominated regions where the taxpayer can increase both consumption and leisure by lowering earnings to the notch point. Since everyone should move out of such regions in a frictionless world, our finding that a large fraction of individuals in those regions (50-80% of self-employed individuals, 90% of wage earners) are unresponsive provides direct evidence of the importance of optimization frictions in attenuating bunching. We have so far been agnostic about the exact nature of optimization frictions, but in this section we present suggestive evidence on this question by investigating which taxpayer characteristics are associated with strictly dominated choice.

Table 3 reports the results of OLS regressions of a dummy for locating in a 500 rupee range above a notch on various dummy covariates. The 500 rupee range is smaller than the strictly dominated range at every notch. By considering a constant range above each cutoff rather than the full dominated range, we avoid mechanical effects of the dominated range becoming larger as the cutoff increases. The table shows results for self-employed individuals and firms in columns (1)-(2) for years 2007 and 2008, respectively, and for wage earners in columns (3)-(4) for the same years.[29] The dummy covariates we consider include (i) dummies for being a registered tax

[28] The tax rate jump at the income-composition notch is determined by the vertical distance between the tax schedules for wage earners and self-employed persons shown in Figure 3. This tax rate jump depends on income level and is therefore heterogeneous across taxpayers, but is in general very large except at the very bottom of the distribution. Among the sample of filers in Figure 13, the tax rate jump between wage-earner and self-employment status (taking the level of total taxable income of each filer as given) is above 6 percentage points on average. This is a truly massive notch and therefore the observed shifting response is associated with a small shifting elasticity.

[29] A caveat is in place with regards to the regression for wage earners in 2008 in column (4). As the tax reform in 2008 introduced a kink option in addition to the notch schedule, locating in a 500 rupee range above the cutoff

filer in prior years, (ii) dummies for being located in strictly dominated regions in prior years, (iii) income quantile dummies (percentiles 25-50, 50-75, 75-90, 90-95, 95-100), (iv) firm dummy, (v) capital and property income dummy, (vi) rural region dummy, and (vii) male gender dummy.

The following main findings emerge from the table. First, strictly dominated choice in a given year is negatively associated with being a registered tax filer in previous years. Compared to the unconditional probability of strictly dominated choice (shown at the bottom of the table), this effect is very large. Hence, taxpayers who have less experience in dealing with the tax code are much more likely to make mistakes, which suggests that tax illiteracy and learning are important for behavioral responses. Second, dominated choice in a given year is positively associated with dominated choice in previous years, and the effect is very strong compared to the baseline probability of dominated choice.[30] The fact that mistakes are strongly persistent over time also points to aspects of tax illiteracy such as misperception or inattention. In particular, switching costs in labor supply are less likely to explain this specific effect for two reasons: (*i*) such switching costs would not necessarily keep a taxpayer in a strictly dominated region year after year due to the combination of high wage inflation and a nominal tax system, (*ii*) the effect is stronger for self-employed individuals than the for wage earners. Of course, this does not rule out that switching costs in hours worked also contribute to optimization frictions, only that they are unlikely to explain the persistence in dominated choice. Third, taxpayers at the top of the income distribution are less likely to locate in strictly dominated ranges than taxpayers at the bottom, consistent with high-income individuals being more sophisticated and the incentives created by notches being stronger at the top. Fourth, the firm dummy has a negative coefficient presumably because firms are more sophisticated and have larger capacity to optimize than individuals. Finally, the presence of capital and property income has a negative effect (but is not always significant), which may be driven by the fact that taxpayers with such income forms are relatively sophisticated.

Overall, these findings suggest that optimization frictions may be driven to a large degree

is not necesarily strictly dominated. However, as analyzed in section 4.2, many taxpayers did not understand this complex reform and behave as under the notch schedule (e.g., a substantial fraction of taxpayers right above the cutoff self-assess tax liability according to the notched schedule rather than the kinked schedule). Hence, it makes sense to consider a regression for wage earners in 2008 that defines dominated choice as being located in a 500 rupee range above a cutoff.

[30] These effects are unlikely to be driven by serial correlation in income, because we are considering a very small earnings range above each notch point and inflation in Pakistan is very high (the 500 rupee range is less than 0.5% of income at every notch, while inflation in Pakistan is 10-20% annually over the period considered).

by aspects of tax illiteracy such as misperception or unawareness of tax incentives.

5 Conclusion

Notches tend to have a bad reputation among economists, although there is relatively little empirical evidence on their effects. Indeed, one of the key results in the canonical optimal income tax framework by Mirrlees (1971) is that the marginal tax rate is everywhere below one, which rules out discontinuous increases in tax liability as created by the notches analyzed in this paper.[31] Such notches create regions of strictly dominated choice where the implicit marginal tax rate is larger than one. In the standard economic model, because no taxpayer is willing to locate in a region of strictly dominated choice, the presence of such regions create strong distortions without collecting any revenue. This can never be socially optimal.[32]

Notches are interesting to study for two reasons. First, they are conceptually and practically interesting in their own right given how widespread they are in tax systems and other public policies around the world. It is valuable to learn if the negative view on notches implied by economic theory lines up with empirical evidence from different contexts in order to form policy recommendations to the many governments using notched schemes. Second, notches provide an ideal opportunity to identify structural elasticity parameters for the analysis of tax reform and optimal taxation, allowing us to move beyond notches and think about alternative tax structures. In particular, we have argued that notches may be better suited than kinks or tax reforms for the identification of structural parameters in a world where optimization frictions are important. Notches provide compelling exogenous variation just like kinks, but in a way that allows us to measure frictions and estimate long-run elasticities.

Our finding of large behavioral responses combined with small structural elasticities illustrates the key problem with notches: by creating extremely large implicit marginal tax rates around cutoffs, they induce very large behavioral responses and efficiency costs even when

[31] The Mirrlees model also rules out discontinuous drops in tax liability (downward notches). In general, if the fundamentals of the model are continuous (ability distribution and preferences), then the optimal tax schedule is also continuos, which rules out both downward and upward notches.

[32] Exceptions to this negative view on notches are provided by Blinder and Rosen (1985) and Slemrod (2010). Both of these papers advocate a view that notches may be optimal in the presence of restrictions on policy instruments (e.g. due to administrative constraints). In particular, Blinder and Rosen show that a notched schedule may dominate a fully linear schedule under some conditions. By contrast, the Mirrleesian literature on optimal nonlinear incentive schemes shows that there always exists a better schedule without notches if we allow for enough flexibility in instruments.

structural elasticities are small. This lends support to the normative conclusions based on the Mirrleesian framework described above. On the other hand, the fact that a substantial fraction of taxpayers in strictly dominated ranges do not respond to notches is clearly inconsistent with the frictionless Mirrleesian model. However, by itself this does not provide a compelling argument in favor of notches for two reasons. First, in the long run, behavior would be governed by the structural elasticities that overcome frictions. Second, under a short horizon, the amount of friction that we identify for notches would presumably be at least as large under alternative continuous tax schedules (although they would be empirically more difficult to measure under such schedules). Hence, even in the short run, it is not likely that notched schedules are more revenue-efficient than continuous schedules. This implies that, absent considerations having to do with administration or simplicity, it seems hard to justify the type of notched tax schedule observed in Pakistan.

References

Andreoni, J., B. Erard, and J. Feinstein (1998). "Tax Compliance." *Journal of Economic Literature* 36, 818-860.

Blinder, A. S. and H. S. Rosen (1985). "Notches." *American Economic Review* 75, 736-747.

Blundell, R. and H. W. Hoynes (2004). "Has "In-Work" Benefit Reform Helped the Labor Market?" in R. Blundell, D. Card, and R. B. Freeman (eds.), *Seeking a Premier Economy: The Economic Effects of British Economic Reforms, 1980-2000*, University Of Chicago Press, Chicago.

Blundell, R. and A. Shephard (2011). "Employment, Hours of Work and the Optimal Taxation of Low Income Families." Forthcoming *Review of Economic Studies.*

Burgess, R. and N. Stern (1993). "Taxation and Development." *Journal of Economic Literature* 31, 762-830.

Chetty, R. (2011). "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply." Forthcoming *Econometrica.*

Chetty, R., J. Friedman, T. Olsen, and L. Pistaferri (2011). "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records." *Quarterly Journal of Economics* 126, 749-804.

Gordon, R. and W. Li (2009). "Tax structures in developing countries: Many puzzles and a possible explanation." *Journal of Public Economics* 93, 855-866.

Gruber, J. and D. A. Wise (1999). *Social Security and Retirement Around the World*, National Bureau of Economic Research, Inc.

Hausman, J. A. (1983). "Stochastic Problems in the Simulation of Labor Supply." In M. Feldstein (ed.), *Behavioral Simulation Methods in Tax Policy Analysis*, University of Chicago Press, Chicago.

Keen, M. and B. Lockwood (2010). "The value added tax: Its causes and consequences." *Journal of Development Economics* 92, 138-151.

Kleven, H., C. Kreiner, and E. Saez (2009). "Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries." NBER Working Paper 15218.

Kleven, H. J, M. Knudsen, C. Kreiner, S. Pedersen, and E. Saez (2011). "Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark." *Econometrica* 79, 651-692.

Mirrlees, J. A. (1971). "An Exploration in the Theory of Optimal Income Taxation." *Review of Economic studies* 38, 175-208.

Newbery, D. and N. Stern (1987). *The Theory of Taxation for Developing Countries*, Oxford University Press, New York.

Piketty, T. and N. Qian (2009). "Income Inequality and Progressive Income Taxation in China and India, 1986-2010." *American Economic Journal: Applied Economics* 1, 53–63.

Saez, E. (2003). "The Effect of Marginal Tax Rates on Income: A Panel Study of Bracket Creep." *Journal of Public Economics* 87, 1231-1258.

Saez, E. (2010). "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy* 2, 180-212.

Saez, E., J. Slemrod, and S. Giertz (2010). "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." Forthcoming *Journal of Economic Literature*.

Sallee, J. M. and J. Slemrod (2010). "Car Notches: Strategic Automaker Responses to Fuel Economy Policy." NBER Working Paper 16604, December 2010.

Slemrod, J. (1995). "Income Creation or Income Shifting? Behavioral Responses to the Tax Reform Act of 1986." *American Economic Review Papers and Proceedings* 85, 175-180.

Slemrod, J. (1998). "Methodological Issues in Measuring and Interpreting Taxable Income Elasticities." *National Tax Journal* 51, 773-88.

Slemrod, J. (2010). "Buenas Notches: Lines and Notches in Tax System Design." Working Paper, September 2010.

Slemrod, J. and S. Yitzhaki (2002). "Tax avoidance, evasion and administration," in A. J. Auerbach and M. Feldstein (eds.), *Handbook of Public Economics* III, Elsevier, Amsterdam.

World Bank (2009). *Pakistan Tax Policy Report: Tapping Tax Bases for Development*, Volume II, July 2009.

Yelowitz, A. S. (1995). "The Medicaid Notch, Labor Supply, and Welfare Participation: Evidence from Eligibility Expansions." *Quarterly Journal of Economics* 110, 909-39.

FIGURE 1
Behavioral Responses to a Tax Notch

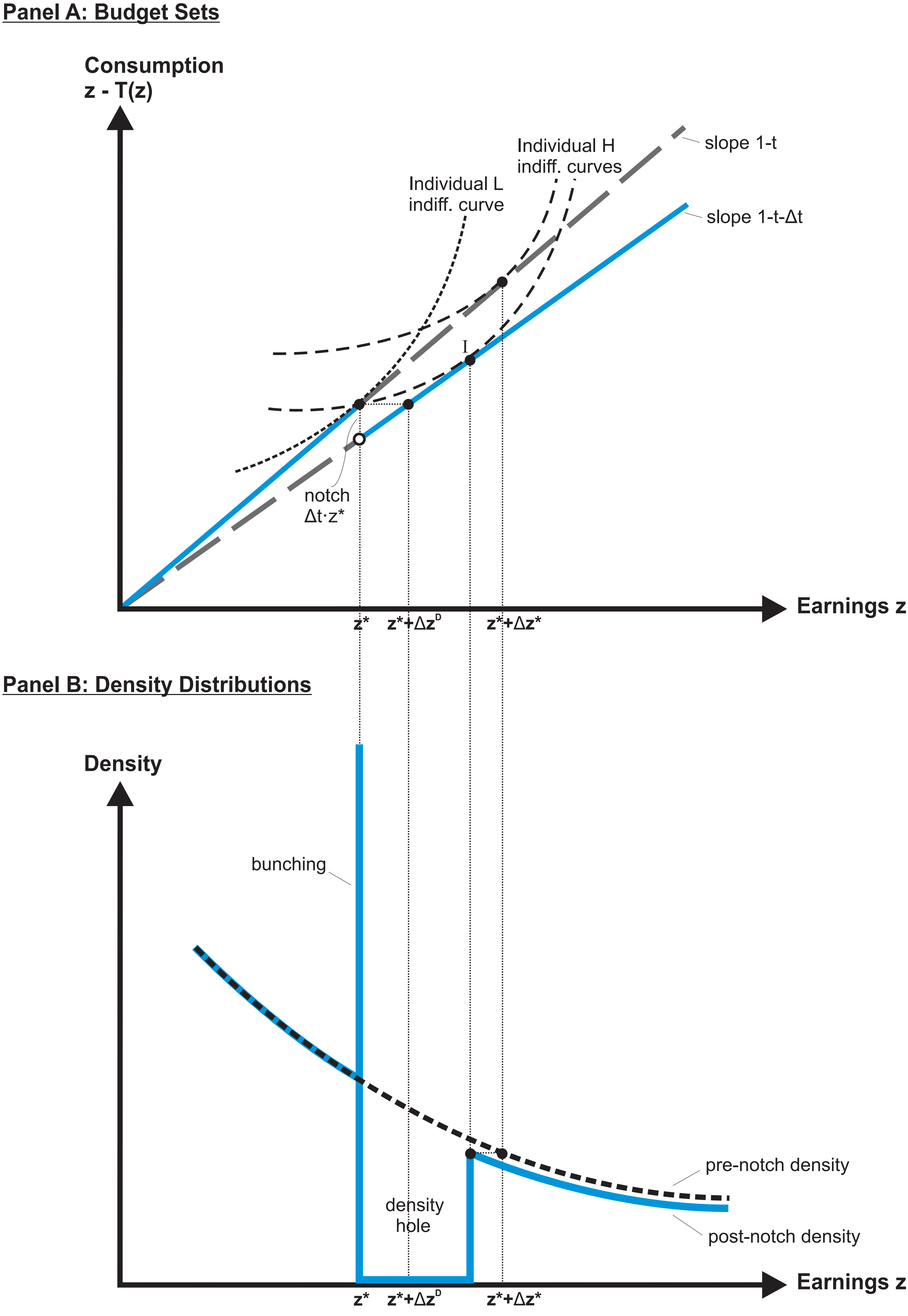


FIGURE 2
Density Distributions Under Different Model Extensions

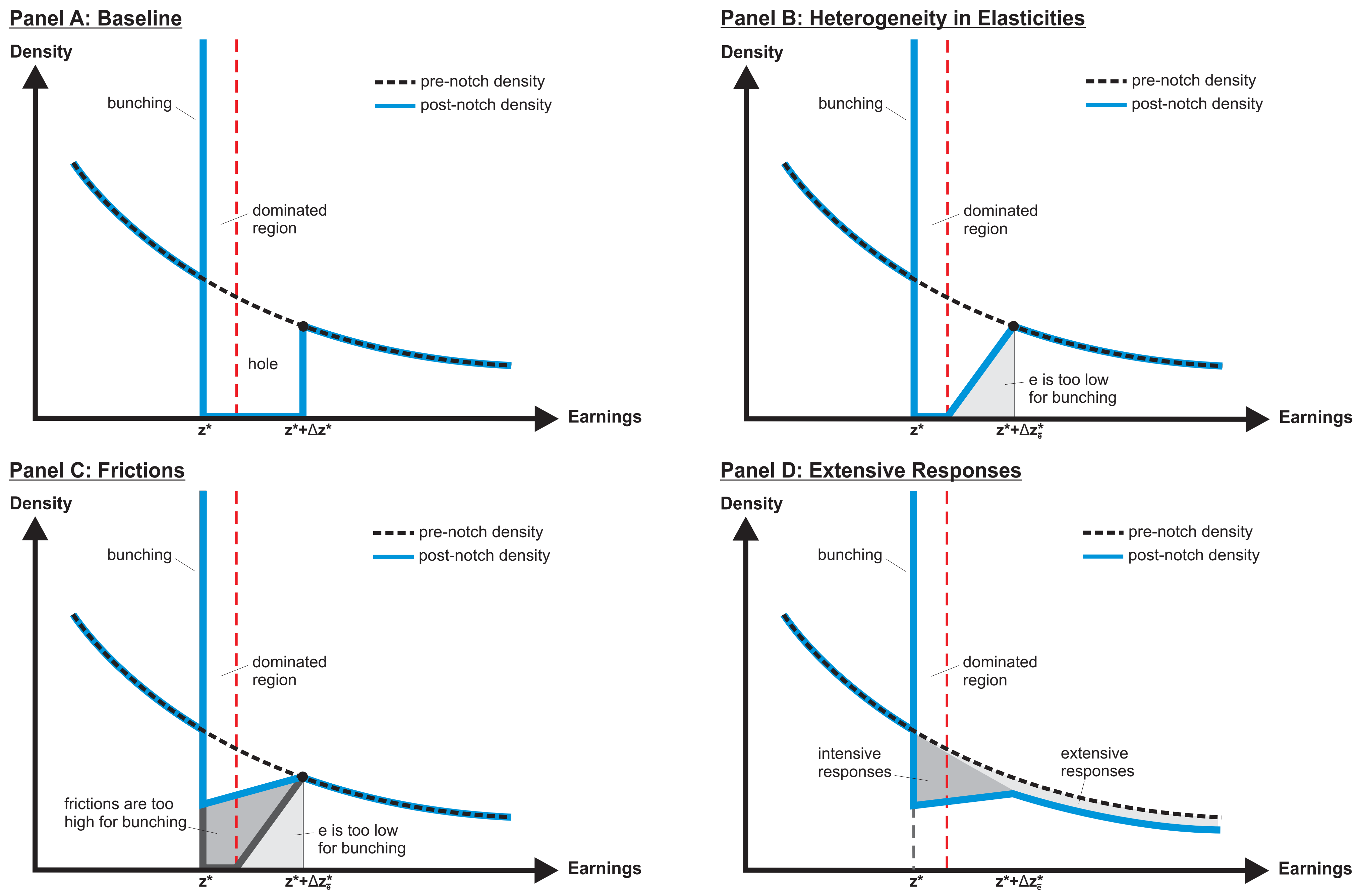


FIGURE 3

Reduced-Form Approximation of Earnings Elasticity

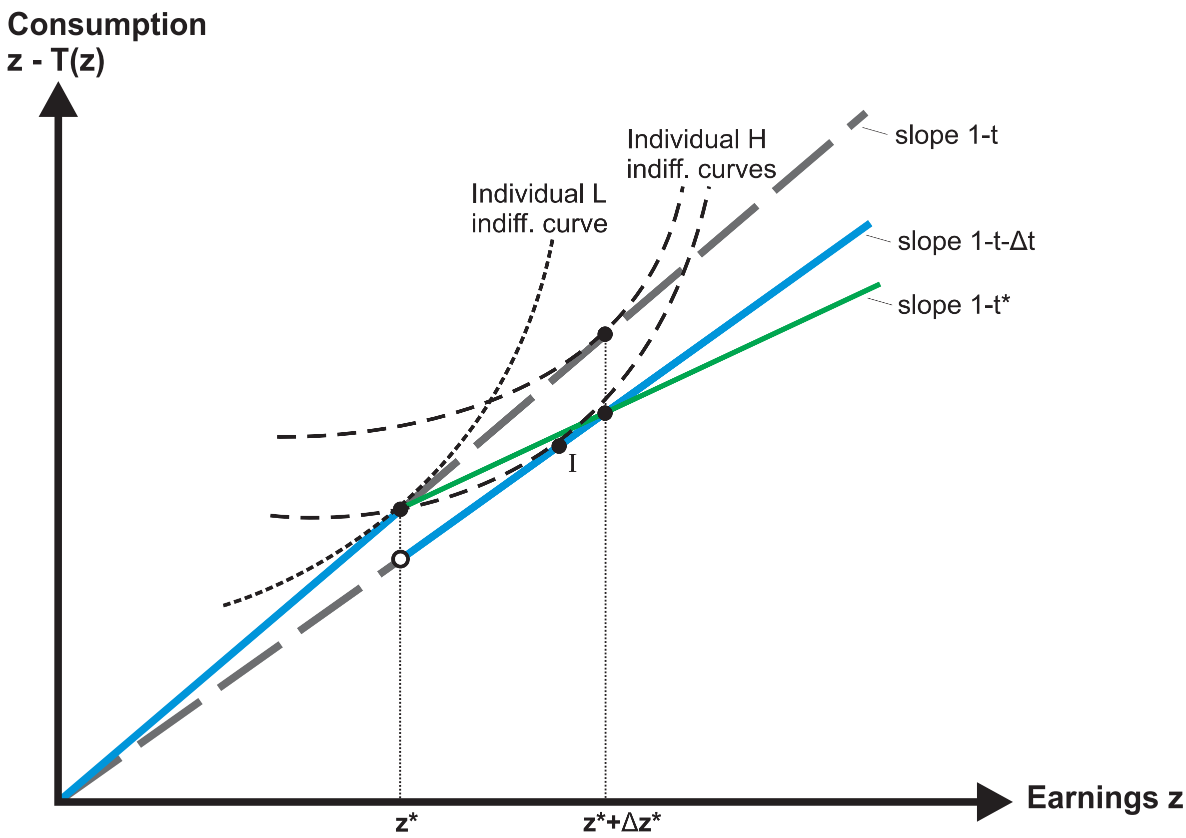


FIGURE 4

Estimating the Counterfactual Density Using an Observed Density

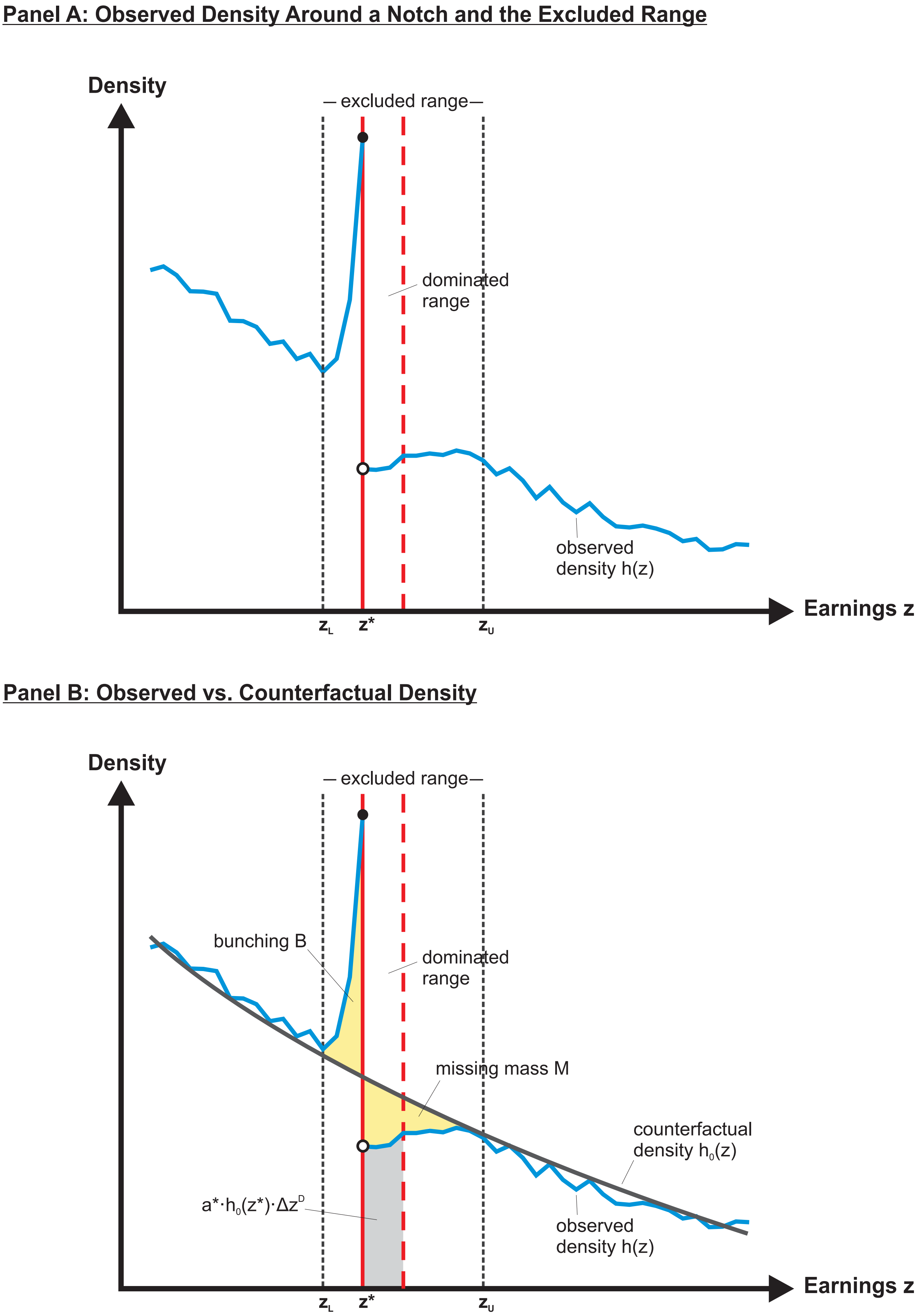
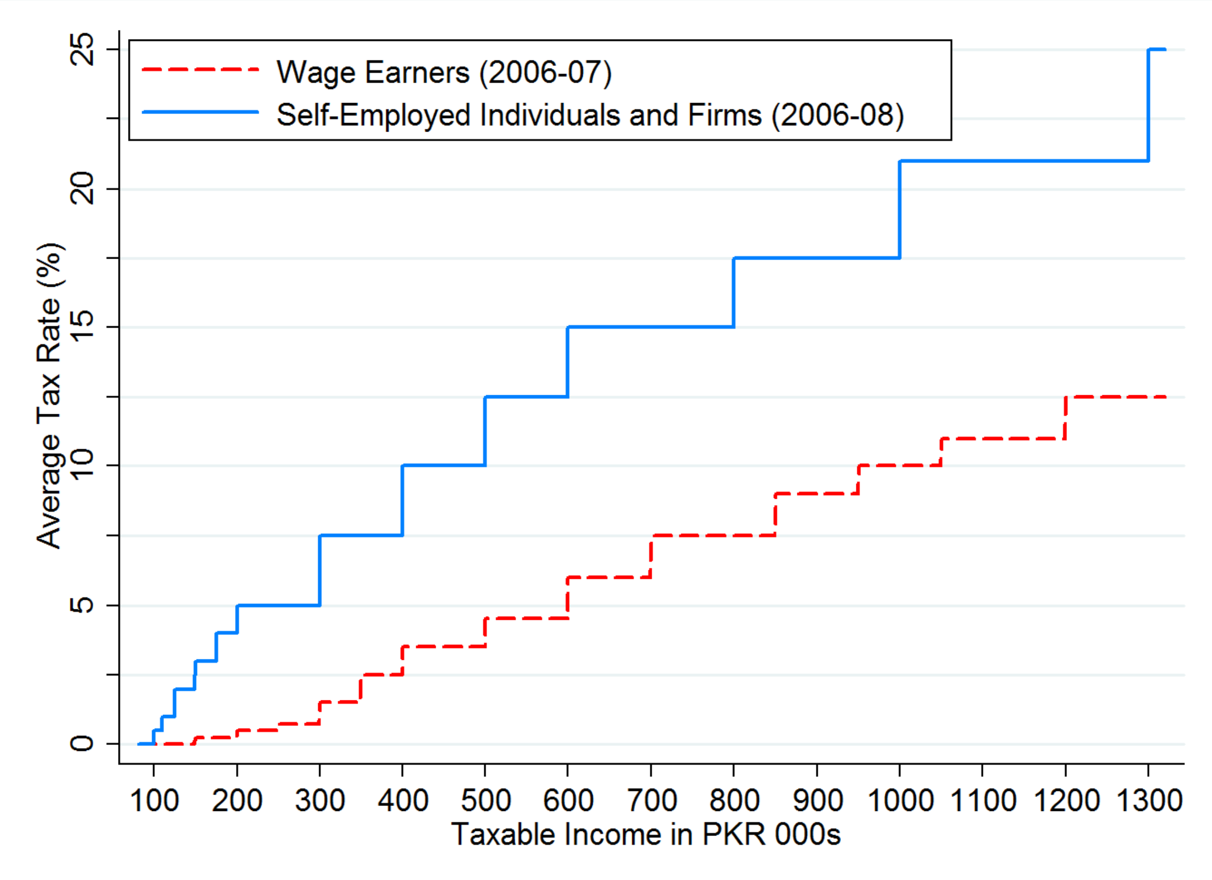


FIGURE 5

Personal Income Tax Schedules in Pakistan

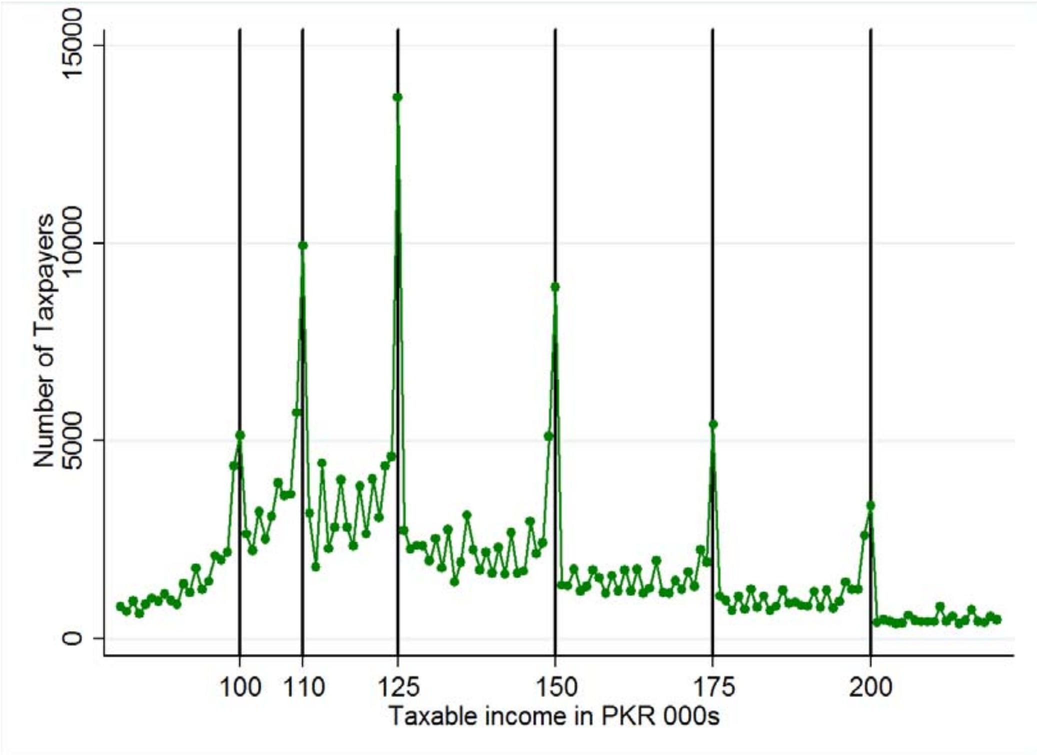


Notes: the figure shows the average tax rate as a function of annual taxable income in the schedules for wage earners (red dashed line) and for self-employed individuals and unincorporated firms (blue solid line). Taxable income is shown in thousands of Pakistani Rupees (PKR), and the PKR-USD exchange rate is about 90 as of March 2012. The tax schedule for self-employed individuals applies to years 2006-08, while the tax schedule for wage earners applies to years 2006-07. The tax system classifies individuals as either wage earners or self-employed based on whether income from wages or self-employment constitute the larger share of total income, and then taxes total income according to the assigned schedule. The tax schedule for self-employed individuals and firms consists of 14 brackets, while the tax schedule for wage earners consists of 21 brackets (the first 14 of which are shown in the figure). Each bracket cutoff is associated with a notch, and the cutoff itself belongs to the tax-favored side of the notch.

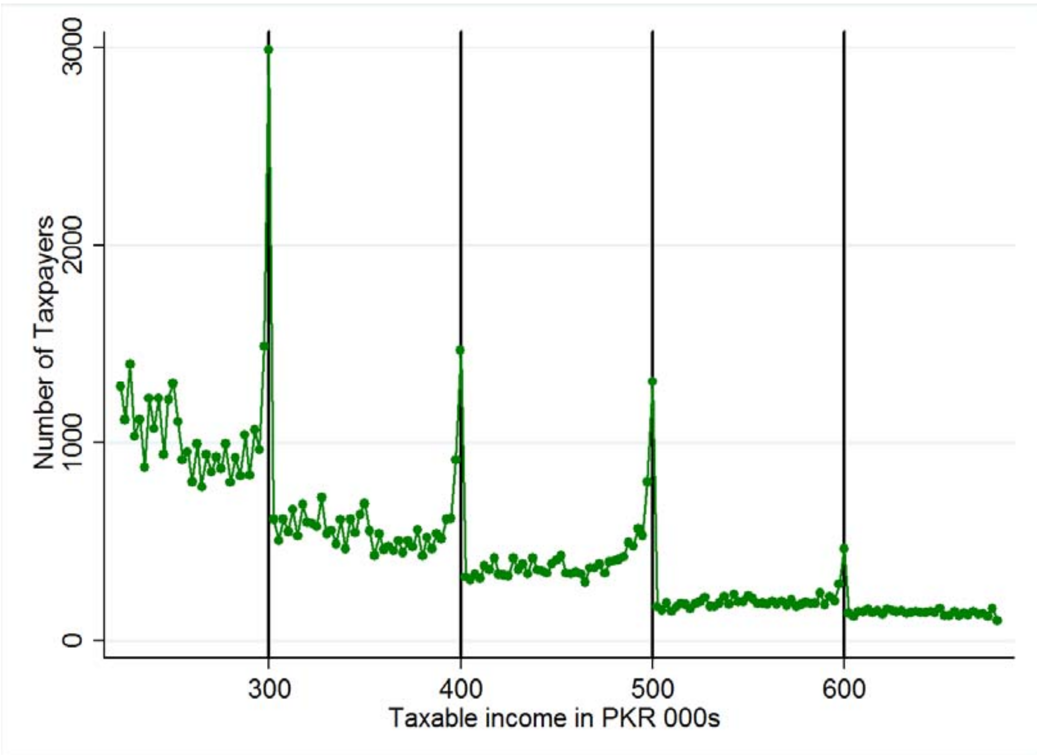
FIGURE 6

Density Distribution around First Ten Notches:
Self-Employed Individuals and Firms (Non-Rounders)

Panel A: First Six Notches



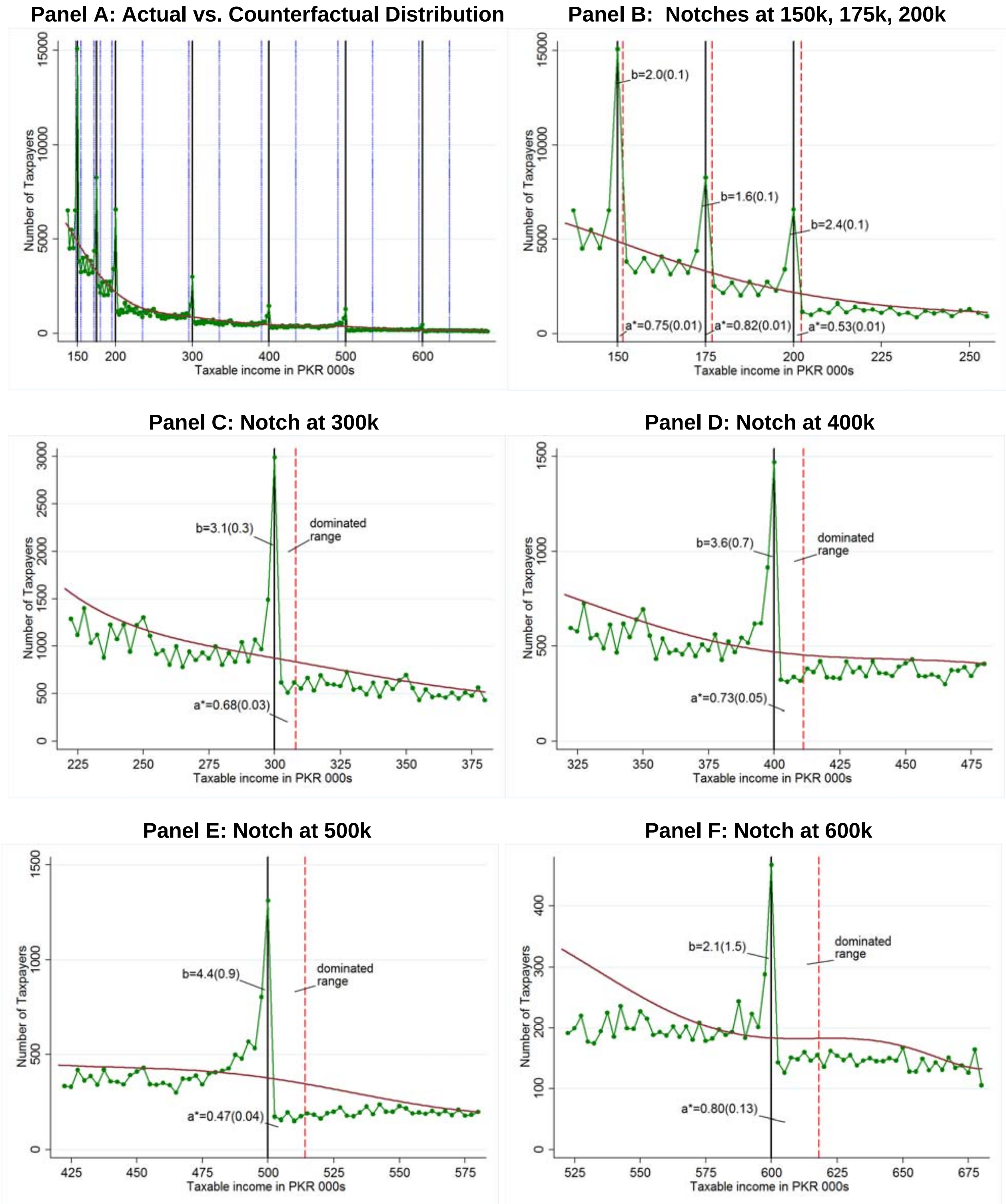
Panel B: Next Four Notches



Notes: the figure shows the taxable income distribution for male self-employed individuals and unincorporated firms in 2006-08. The distribution includes only non-rounders defined as those who do not report income in even thousands. The dots show the number of taxpayers in 1000 Rupee bins in Panel A and in 2500 Rupee bins in Panel B. Each dot is located at the upper bound of a given bin. Notch points are demarcated by black vertical lines.

FIGURE 7

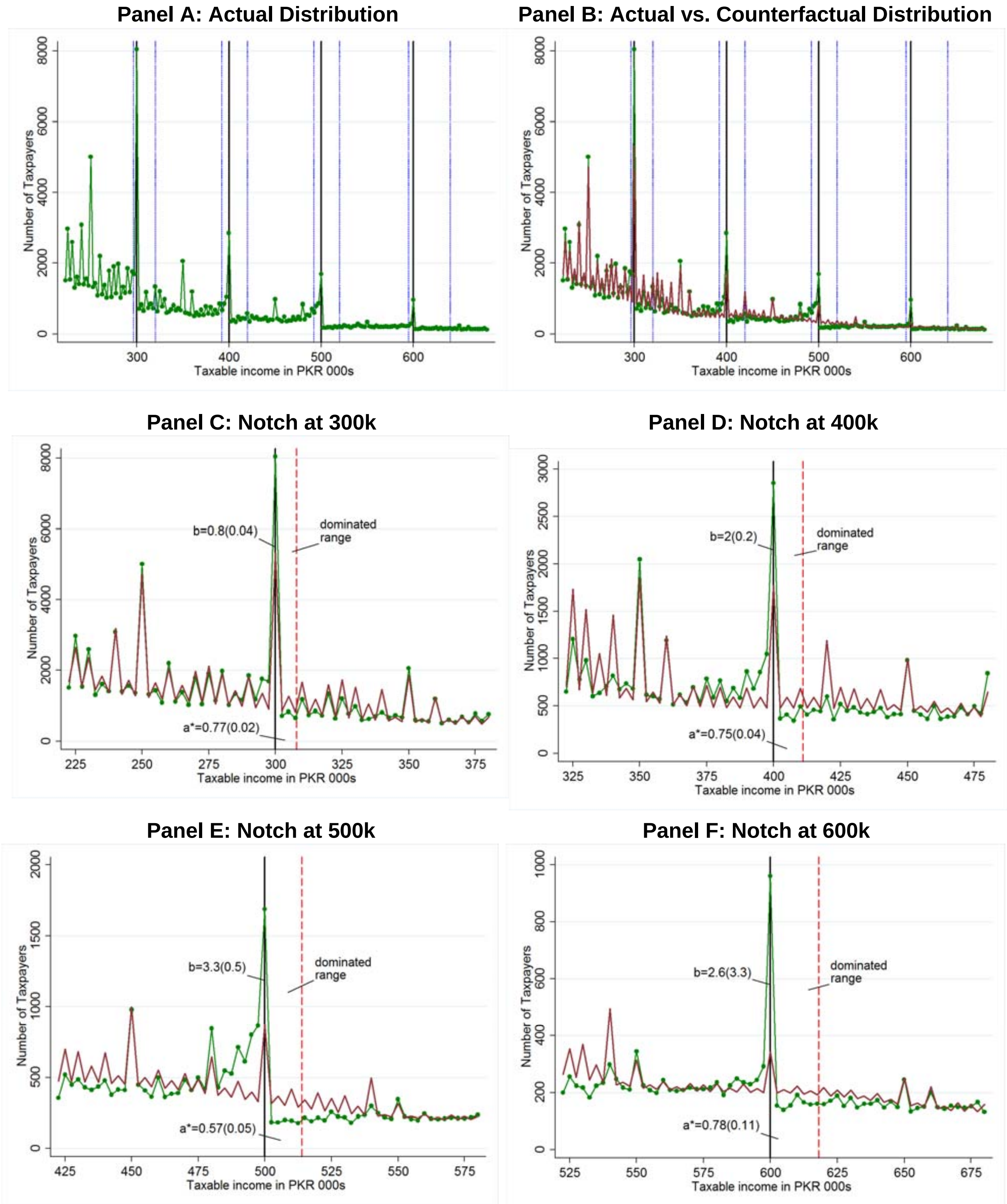
Actual and Counterfactual Distributions around Middle Notches:
Self-Employed Individuals and Firms (Non-Rounders)



Notes: the figure shows the observed distribution (dotted green graph) and the estimated counterfactual distribution (solid brown graph) of taxable income for male self-employed individuals and unincorporated firms in 2006-08. The distributions include only non-rounders defined as those who do not report income in even thousands. The counterfactual is a ninth-order polynomial estimated as in eq. (12), with excluded ranges around notch points demarcated by the dashed blue lines in Panel A. Notch points are shown by solid black lines and the upper bounds of dominated ranges are shown by dashed red lines. Bunching *b* is excess mass in the excluded range below the notch in proportion to the average counterfactual frequency in this range, and a* is the proportion of taxpayers in dominated ranges who are unresponsive. Standard errors are shown in parentheses.

FIGURE 8

Actual and Counterfactual Distributions around Middle Notches:
Self-Employed Individuals and Firms (Full Population)

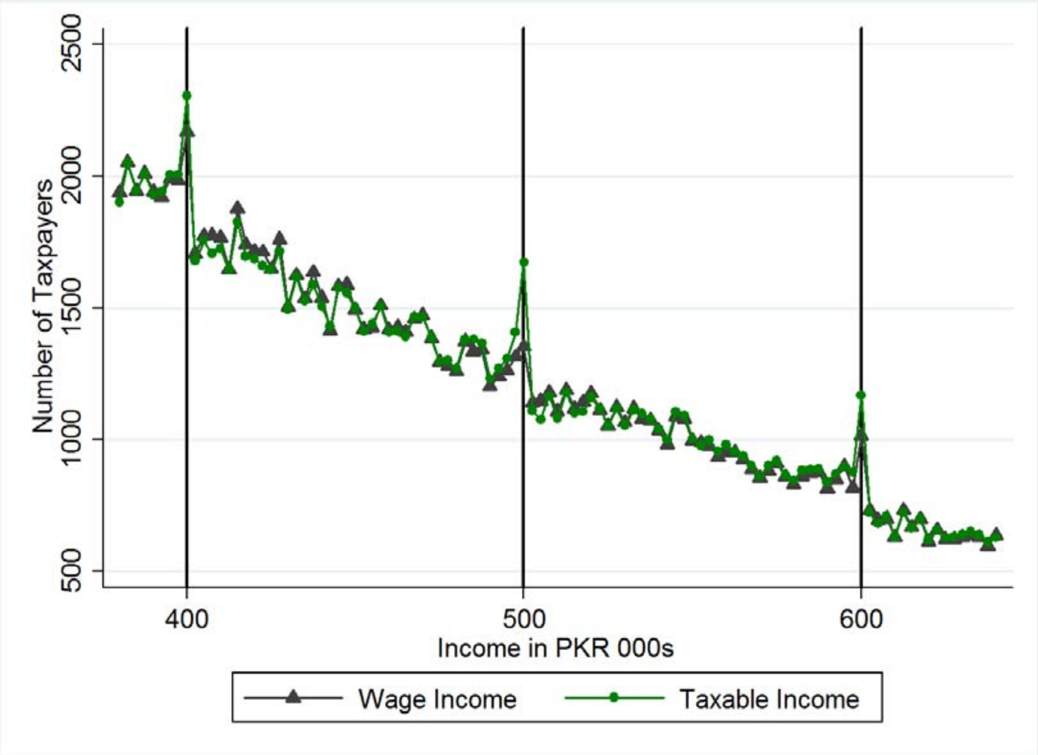


Notes: the figure shows the observed distribution (dotted green graph) and the estimated counterfactual distribution (solid brown graph) of taxable income for male self-employed individuals and unincorporated firms in 2006-08. The distributions are for the full sample and show spikes at round numbers. The counterfactual is a ninth-order polynomial estimated as in eq. (13), with excluded ranges around notch points demarcated by the dashed blue lines in Panel A-B. Notch points are shown by solid black lines and the upper bounds of dominated ranges are shown by dashed red lines. Bunching *b* is excess mass in the excluded range below the notch in proportion to the average counterfactual frequency in this range, and a* is the proportion of taxpayers in dominated ranges who are unresponsive. Standard errors are shown in parentheses.

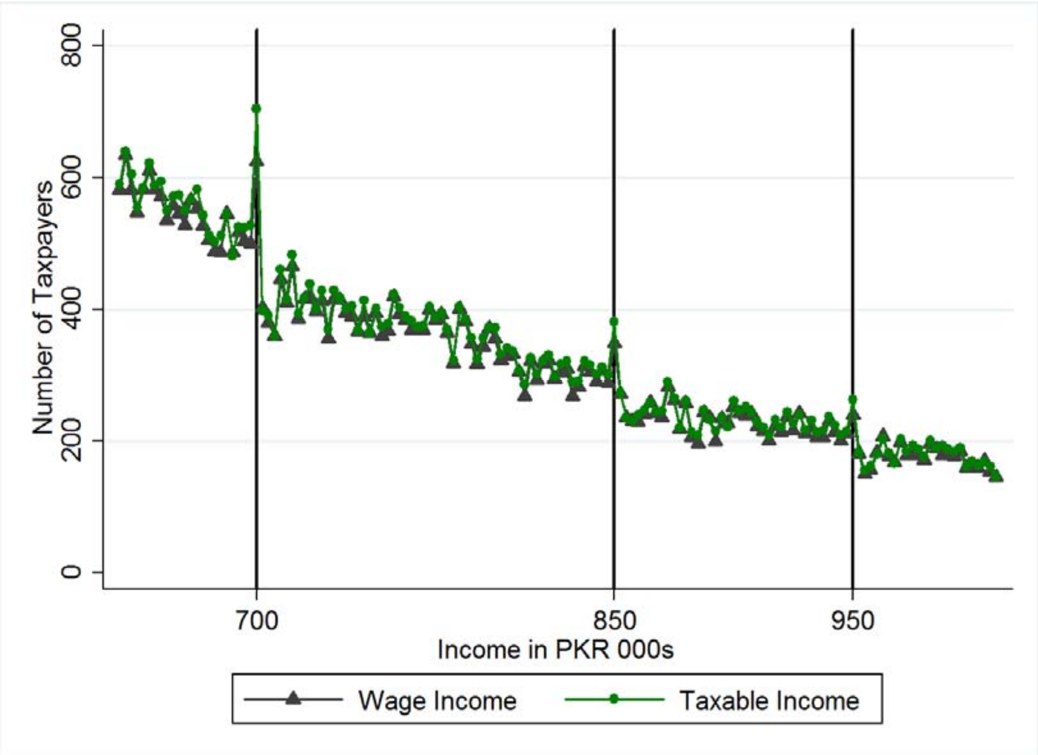
FIGURE 9

Density Distributions around Middle Notches:
Wage Earners (Non-Rounders)

Panel A: Notches at 400k, 500k, and 600k



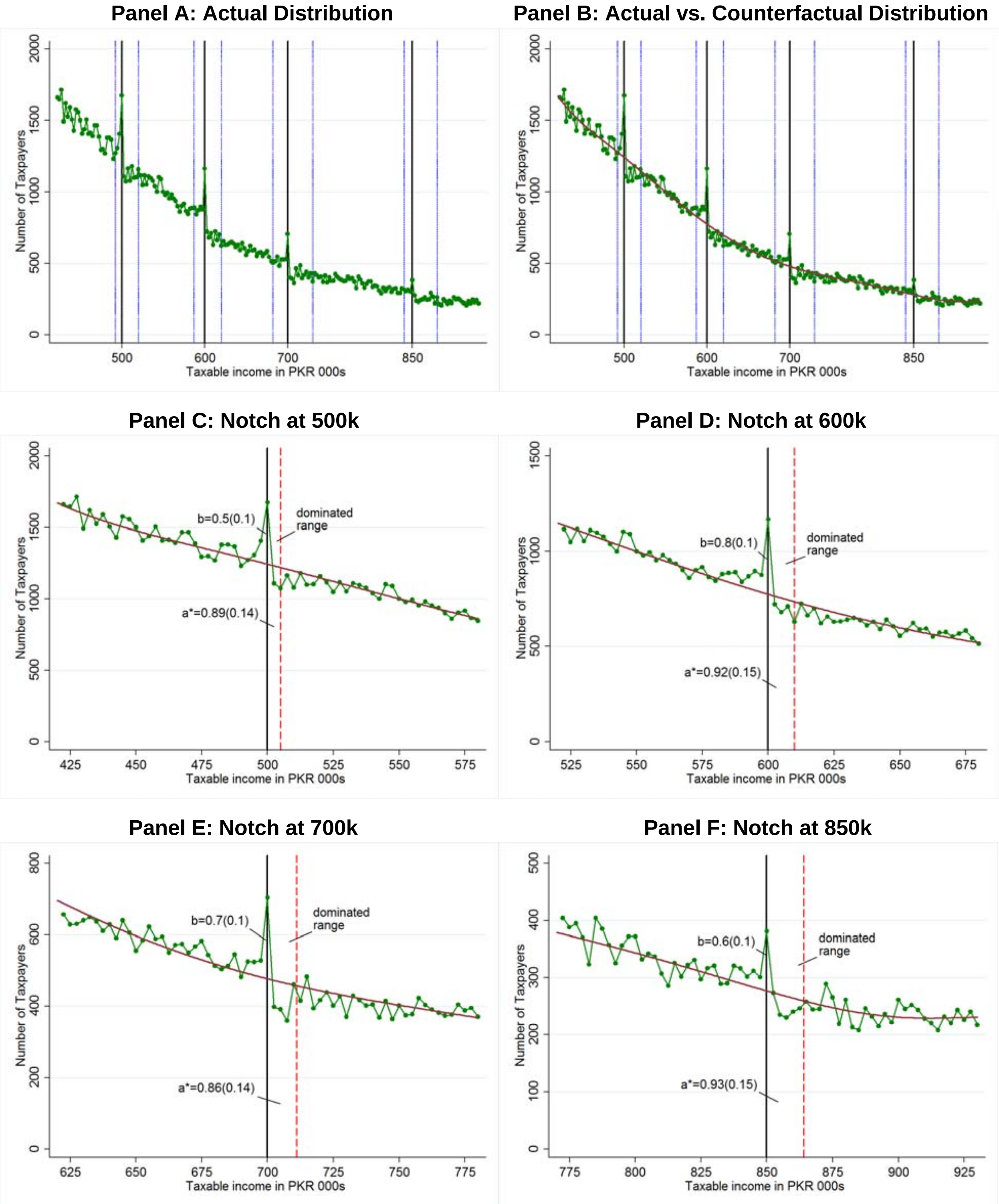
Panel B: Notches at 700k, 850k, and 950k



Notes: the figure shows the distributions of gross wage income and taxable income for all wage earners in 2006-07 around six middle notches in the wage earner tax schedule. The distributions exclude rounders defined as those who report income in even thousands. Each dot shows the number of observations in a 2500 Rupee bin below the dot. Notch points are demarcated by solid black lines.

FIGURE 10

Actual and Counterfactual Distributions around Middle Notches:
Wage Earners (Non-Rounders)

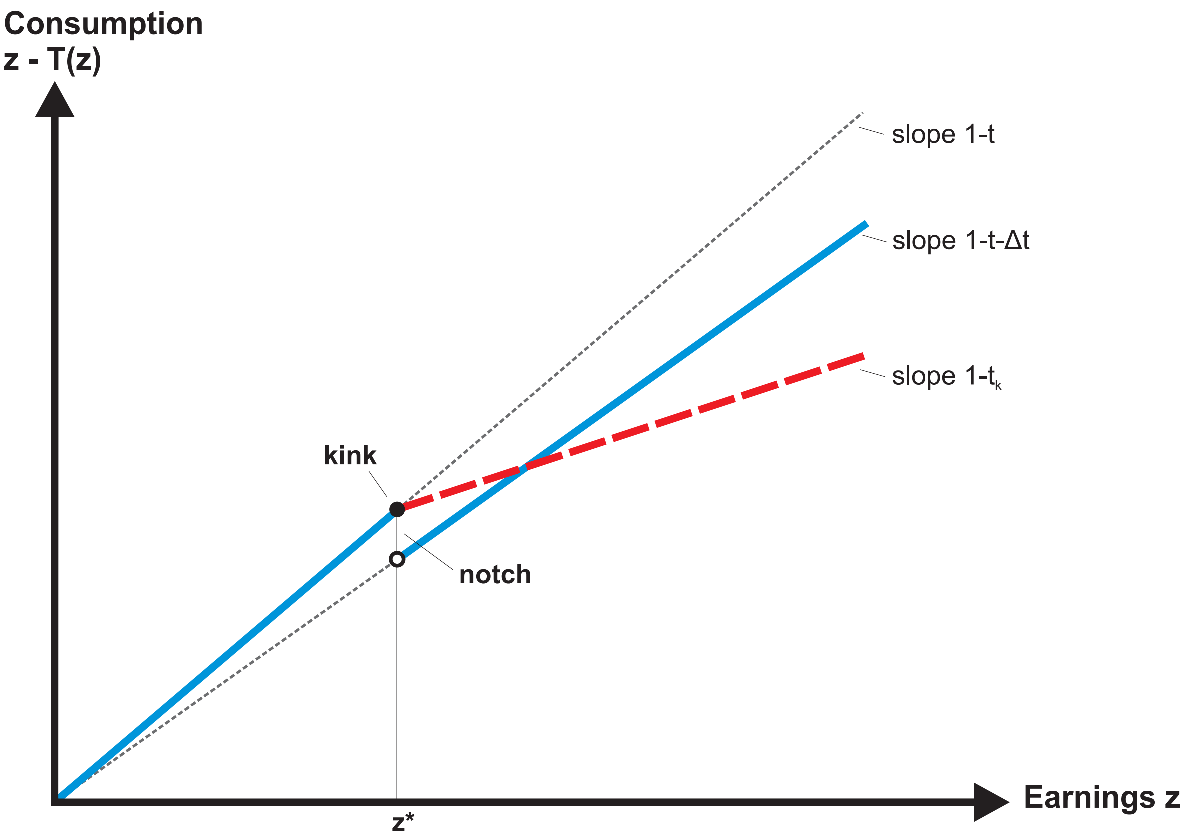


Notes: the figure shows the observed distribution (dotted green graph) and the estimated counterfactual distribution (solid brown graph) of taxable income for wage earners in 2006-07. The distributions exclude rounders defined as those who report income in even thousands. The counterfactual is a ninth-order polynomial estimated as in eq. (12), with excluded ranges around notch points demarcated by the dashed blue lines in Panel A-B. Notch points are shown by solid black lines and the upper bounds of dominated ranges are shown by dashed red lines. Bunching *b* is excess mass in the excluded range below the notch in proportion to the average counterfactual frequency in this range, and a* is the proportion of taxpayers in dominated ranges who are unresponsive. Standard errors are shown in parentheses.

FIGURE 11

Tax Reform for Wage Earners in 2008

Panel A: Budget Sets under the Notch and Kink Options



Panel B: Budget Set of a Tax-Minimizing Individual

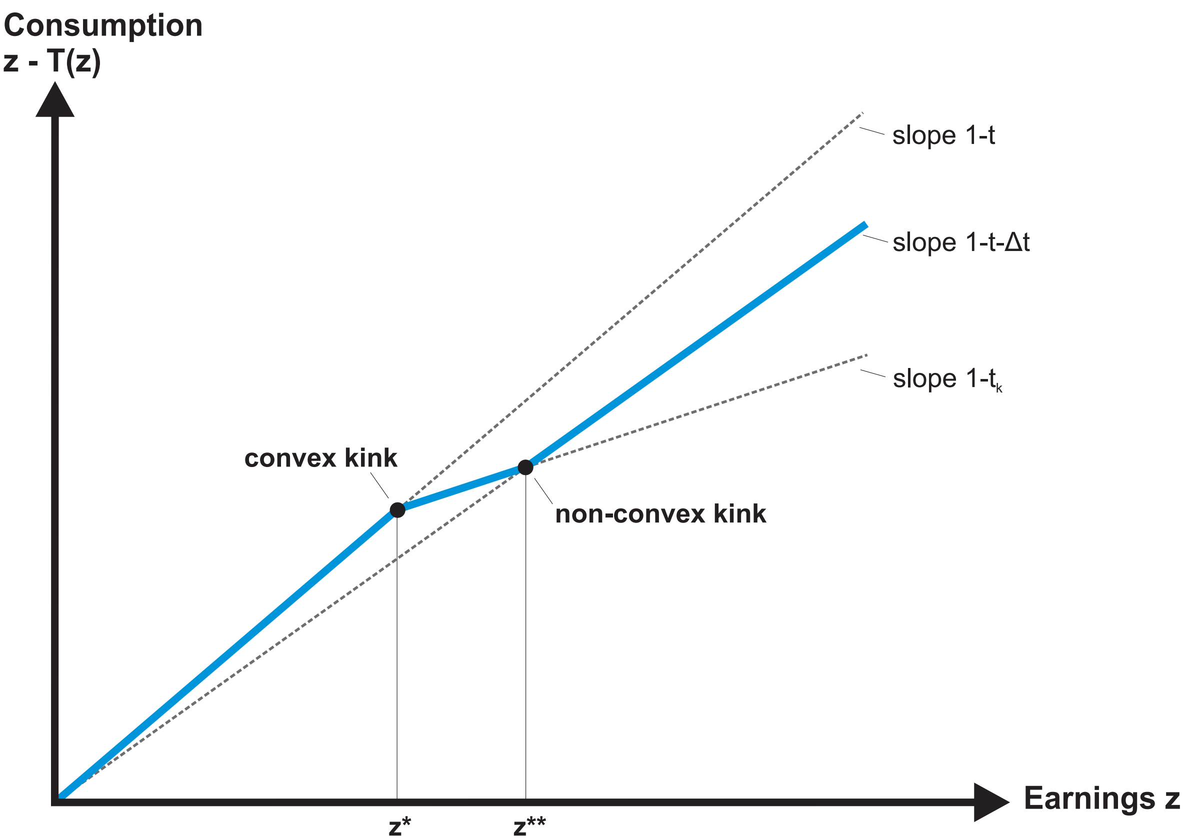
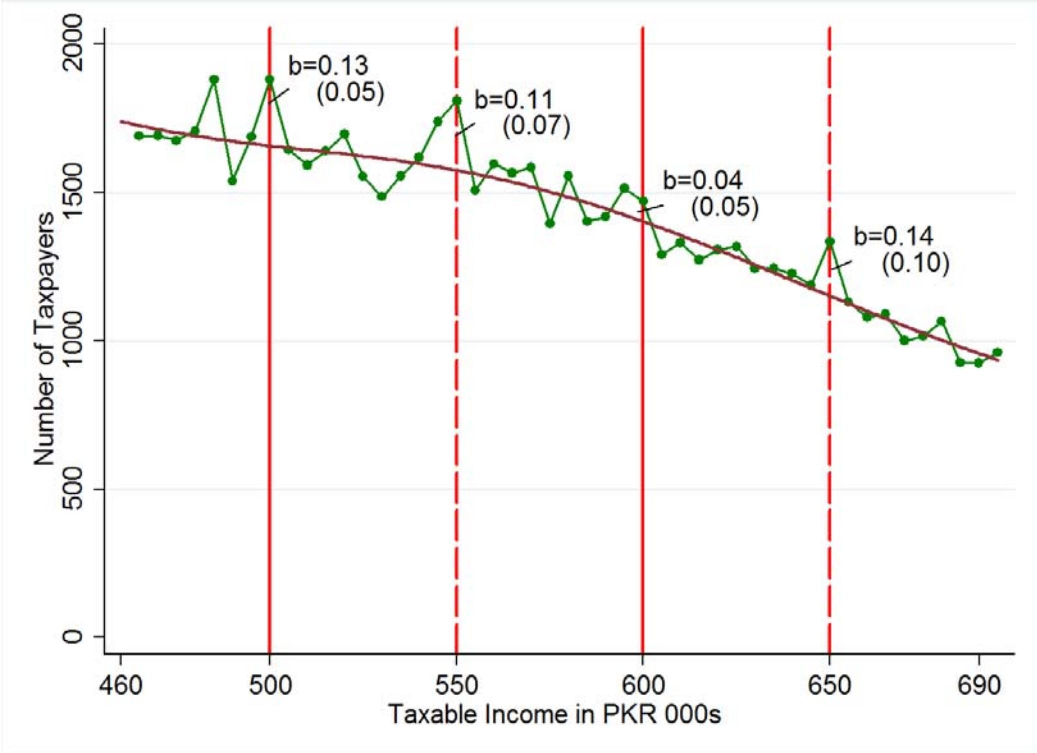


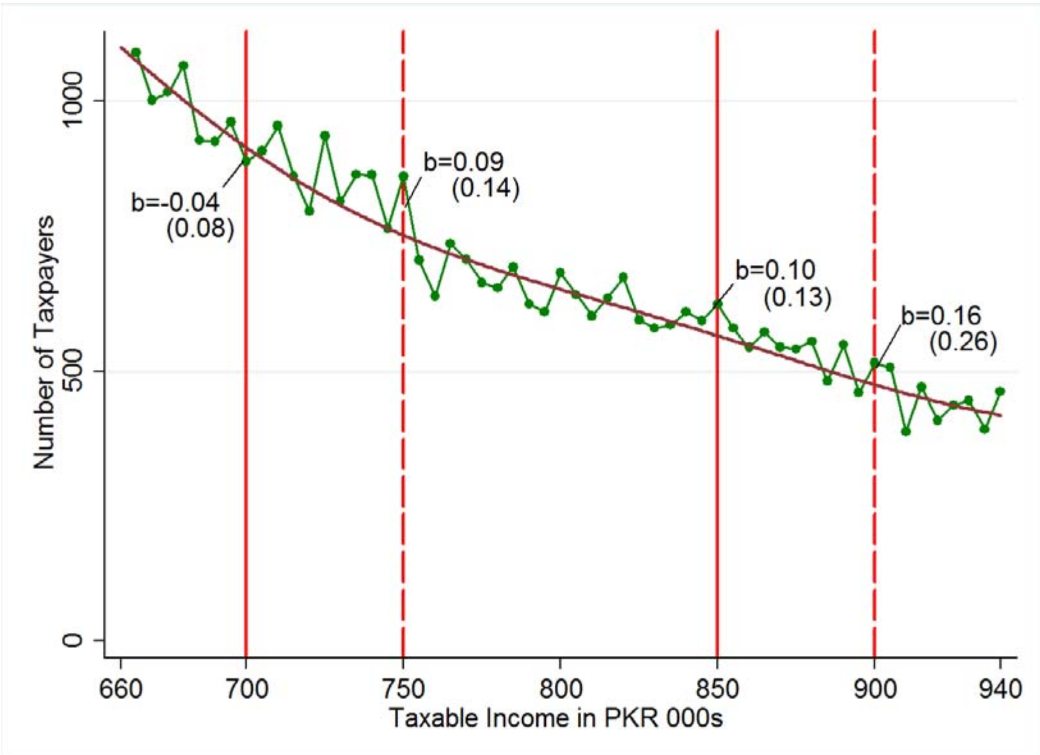
FIGURE 12

Actual and Counterfactual Distributions for Wage Earners after the 2008 Tax Reform: Elimination of Notches and Introduction of Kinks

Panel A: Old Notch Points at 500k & 600k, New Kink Points at 550k & 650k



Panel B: Old Notch Points at 700k & 850k, New Kink Points at 750k & 900k

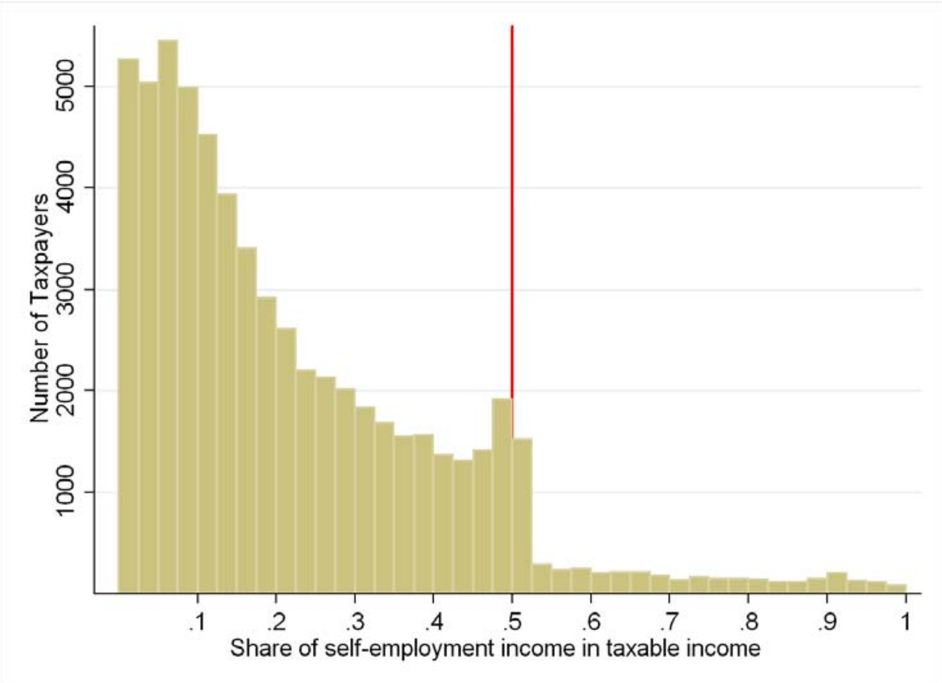


Notes: the figure shows the observed and counterfactual distributions of taxable income for wage earners (non-rounders only) in 2008. The counterfactual is a ninth-order polynomial estimated based on eq. (12). Notch points that were eliminated by the 2008 reform are shown by vertical solid lines, while convex kink points that were introduced by the 2008 reform are shown by vertical dashed lines. Bunching *b* measures excess mass (in proportion to the counterfactual frequency) in a 5k interval below the cutoff in the case of notches that were eliminated and in a 5k interval around the cutoff in the case of kinks that were introduced. Standard errors are shown in parentheses.

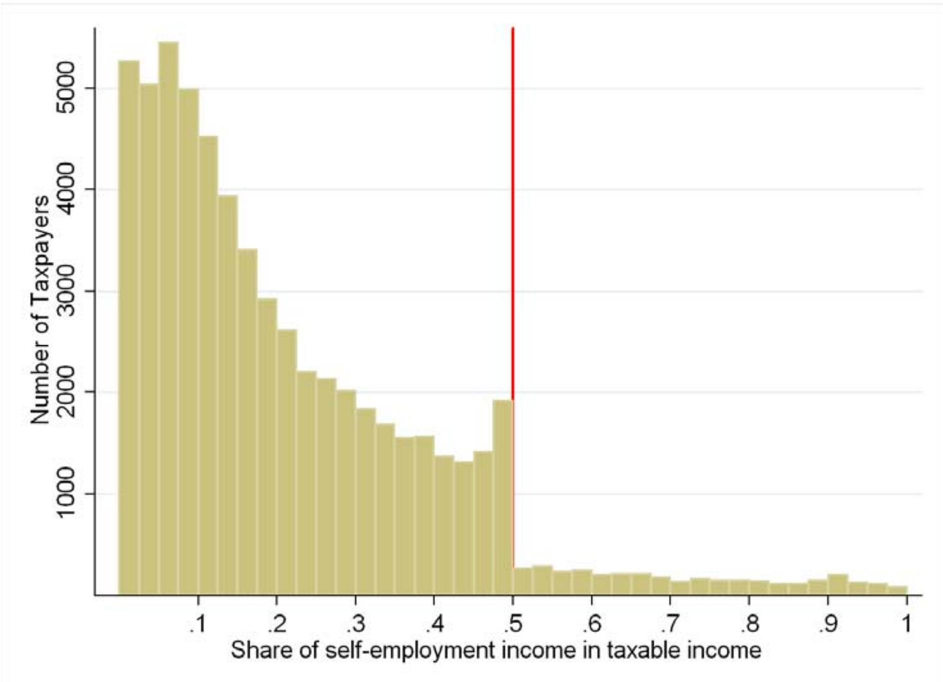
FIGURE 13

Distribution of Self-Employment Income Share around 50% Notch:
Shifting between Self-Employment Income and Wage Income

Panel A: All Taxpayers with Self-Employment Income Share in (0,1) Range



Panel B: Dropping Taxpayers at the Cutoff Value of 0.5



Notes: the figure shows the distribution of the share of self-employment income in total income in years 2006-08. Taxpayers are treated as wage earners (low tax rates) on the left side of the vertical line at 0.50 and as self-employed (high tax rates) on the right side of the vertical line. The notch point at 0.50 is itself part of the high-tax region. The density distribution is shown in bins of 0.025, where each bin excludes the upper bound of the interval such that the bin right below the notch spans [0.475;0.50) and the bin right above spans [0.50;0.525). Panel A considers all taxpayers who report strictly positive income amounts from both wages and self-employment. Panel B considers taxpayers who report strictly positive income amounts from both wages and self-employment, except those who report the exact same amount from the two sources (a self-employment income share exactly equal to the cutoff-value, which belongs to the high-tax region).

TABLE 1

Structural Earnings Elasticities for Self-Employed Individuals and Firms

Notch Point	ATR Jump (%-points)	Dominated Range	Earnings Response		Structural Elasticity		Reduced-Form Elasticity	
			Bunching/ Hole Method	**Conv. Method**	**Bunching/ Hole Method**	**Conv. Method**	**Bunching/ Hole Method**	**Conv. Method**
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Non-Rounder Sample								
200k	1.0	2,105	**14,800**	**44,600**	**0.210**	**1.651**	**0.253**	2.147
			(1,316)	(16,935)	(0.037)	(0.750)	(0.045)	(1.828)
300k	2.5	8,108	**26,700**	**45,000**	**0.101**	0.292	**0.144**	**0.398**
			(6,707)	(15,313)	(0.045)	(0.151)	(0.058)	(0.203)
400k	2.5	11,111	**35,300**	**46,700**	**0.095**	**0.172**	**0.138**	**0.238**
			(8,290)	(11,257)	(0.033)	(0.054)	(0.043)	(0.071)
500k	2.5	14,286	**22,600**	**76,200**	0.017	**0.282**	**0.036**	**0.388**
			(5,291)	(11,493)	(0.013)	(0.036)	(0.016)	(0.048)
600k	2.5	17,647	**27,400**	**67,200**	**0.016**	**0.147**	**0.036**	**0.208**
			(1,930)	(5,964)	(0.004)	(0.013)	(0.004)	(0.016)
Panel B: Full Population								
300k	2.5	8,108	**30,000**	**40,100**	**0.129**	**0.233**	**0.181**	**0.318**
			(5,699)	(8,663)	(0.037)	(0.075)	(0.048)	(0.100)
400k	2.5	11,111	**31,600**	**46,700**	**0.075**	**0.172**	**0.111**	**0.238**
			(5,809)	(8,410)	(0.020)	(0.038)	(0.026)	(0.050)
500k	2.5	14,286	**31,400**	**40,100**	**0.042**	**0.074**	**0.069**	**0.111**
			(4,506)	(6,704)	(0.010)	(0.018)	(0.013)	(0.024)
600k	2.5	17,647	**42,500**	**44,700**	**0.053**	**0.060**	**0.085**	**0.094**
			(3,120)	(8,064)	(0.006)	(0.018)	(0.008)	(0.023)

Notes: the table presents estimates of the earnings response and the structural elasticity of taxable income with respect to the marginal net-of-tax rate at different notch points for self-employed individuals and firms. Panel A shows results for non-rounders while Panel B shows result for the full population of tax filers. The bunching/hole method refers to the method that scales observed bunching B b*y* the inverse of the hole in the dominated range 1/(1-a*) in order to estimate responses that are not attenuated by optimization frictions. The convergence method estimates responses based on the point of convergence between the observed density and the counterfactual density. Those two methods provide lower and upper bounds on the true structural elasticity. Standard errors are shown in parentheses and estimates in bold are significant at the standard 5% level.

TABLE 2
Structural Earnings Elasticities for Wage Earners

Notch Point	ATR Jump (%-points)	Dominated Range	Bunching/Hole Method		
			Earnings Response	Structural Elasticity	Reduced-form Elasticity
(1)	(2)	(3)	(4)	(5)	(6)
500k	1.0	5,236	**11,300** (1,875)	**0.016** (0.003)	**0.024** (0.004)
600k	1.5	9,574	**26,900** (2,997)	**0.044** (0.003)	**0.063** (0.005)
700k	1.5	11,351	**12,400** (1,987)	**0.002** (0.000)	**0.010** (0.001)
850k	1.5	14,011	**20,700** (2,348)	**0.008** (0.000)	**0.018** (0.001)

Notes: the table presents estimates of the earnings response and the structural elasticity of taxable income with respect to the marginal net-of-tax rate at different notch points for wage earners (non-rounder sample). The bunching/hole method in columns refers to the method that scales observed bunching B b*y* the inverse of the hole in the dominated range 1/(1-a*) in order to estimate responses that are not attenuated by optimization frictions. Standard errors are shown in parentheses and estimates in bold are significant at the standard 5% level.

TABLE 3
Determinants of Strictly Dominated Choice

	Self-Employed and Firms		Wage Earners	
	(2007)	**(2008)**	**(2007)**	**(2008)**
	(1)	(2)	(3)	(4)
Constant	0.0361***	0.0301***	0.0346***	0.0130***
	(0.0022)	(0.0021)	(0.0022)	(0.0018)
Registered filer 2006	-0.0064***	-0.0020***	-0.0063***	-0.0023***
	(0.0004)	(0.0004)	(0.0004)	(0.0003)
Registered filer 2007		-0.0050***		-0.0089***
		(0.0005)		(0.0003)
Dominated 2006	0.1301***	0.0392***	0.0080**	0.0091***
	(0.0022)	(0.0023)	(0.0038)	(0.0030)
Dominated 2007		0.1127***		0.0015
		(0.0024)		(0.0022)
Income quantile 25-50	-0.0283***	-0.0271***	-0.0260***	0.0005
	(0.0006)	(0.0005)	(0.0006)	(0.0004)
Income quantile 50-75	-0.0192***	-0.0185***	-0.0271***	0.0016***
	(0.0006)	(0.0006)	(0.0006)	(0.0004)
Income quantile 75-90	-0.0233***	-0.0219***	-0.0313***	0.0000
	(0.0007)	(0.0006)	(0.0006)	(0.0005)
Income quantile 90-95	-0.0225***	-0.0190***	-0.0323***	-0.0016**
	(0.0010)	(0.0009)	(0.0009)	(0.0007)
Income quantile 95-100	-0.0231***	-0.0214***	-0.0342***	-0.0034***
	(0.0010)	(0.0009)	(0.0009)	(0.0007)
Firm	-0.0026***	-0.0035***		
	(0.0008)	(0.0007)		
Capital & property income	-0.0020	-0.0061***	-0.0015	-0.0015
	(0.0015)	(0.0016)	(0.0014)	(0.0012)
Rural	-0.0043***	-0.0030***	-0.0078***	-0.0035***
	(0.0004)	(0.0004)	(0.0005)	(0.0004)
Male	-0.0027	0.0022	0.0060***	0.0024
	(0.0022)	(0.0021)	(0.0022)	(0.0017)
R-squared	0.0276	0.0229	0.0128	0.0031
Mean of dependent variable	0.0102	0.0090	0.0139	0.0080
# Obs.	241,525	259,299	360,677	376,747

Notes: the table reports coefficients of OLS regressions of a dummy for locating in a 500 Rupee range above a notch on dummy covariates. The dummy covariates are (i) dummies for being a registered tax filer in prior years, (ii) dummies for strictly dominated choice in prior years, (iii) income percentile dummies, (iv) firm dummy, (v) capital & property income dummy, (vi) rural region dummy, and (vii) male gender dummy. Standard errors are shown in parentheses. Bottom rows show R-squared, mean of the dependent variable (unconditional probability of strictly dominated choice), and the number of observations in the regressions.