Monitoring and its Interaction with Punishment in Tax Enforcement: Evidence from a Regression Discontinuity Design[1]

(Preliminary, Please do not Quote)

Abstract:

We provide credible evidence on two important issues in tax reform: (i) effects of better monitoring and (ii) possible complementarity between monitoring and punishment. In spite of the importance of monitoring in tax enforcement; there is no estimate of the causal effect of relying on Large Taxpayer Units (LTUs) to intensively monitor large firms. Although complementarity between monitoring and punishment is a feature of theoretical models, there is no evidence on its validity. We exploit a regression discontinuity design (RD) in Ecuador to estimate the tax revenue effects of higher monitoring of LTU firms. The RD is based on a discontinuity in LTU eligibility in 2005 and 2007, when a firm was eligible if its 'index' was greater than zero. The index incorporates firm characteristics, such as prior revenue and taxes. To test the complementarity hypothesis, we combine RD with a difference-in-difference framework based on a reform instituting prison for tax evasion in 2007. We compare RD estimates of the tax payments between LTU and non-LTU firms for pre- and post-reform periods. Results show that higher monitoring induces firms to more than double their mean taxes. While mean and median estimates show no complementarity, strong complementarity is observed above the 70th quantile.

Keywords: Tax evasion, corporate tax compliance, tax reform, developing country, monitoring, complementarity, punishment, Ecuador

JEL Codes: H26, H32, O12

[1] We thank conference participants at the 2010 LACEA Annual Conference for helpful comments and discussions. We are especially grateful to the Ecuadorian Tax Authority (Servicio de Rentas Internas, SRI) staff for their collaboration and for providing access to the data used in this study. All confidential data has been handled according to the Ecuadorian data confidentiality laws.

1. Introduction

A key implication of the standard economic model of law enforcement a la Becker (1968) is that the optimal form of deterrence is maximal fines and minimal levels of costly monitoring. The subsequent theoretical literature, however, shows that better monitoring may be important for enforcement, especially when the threat of fines is limited due to information and liquidity constraints (see Polinsky and Shavell 2000). How important is better monitoring for tax policy in developing countries? One can argue that monitoring will be more important in developing countries, because the information base for tax policy is limited by a large informal sector (Emran and Stiglitz 2005; Munk 2008; Stiglitz 2010), and credit (wealth) constraint is much more severe due to underdeveloped credit markets and low levels of income (Bardhan et al 2006; Banerjee and Duflo 2010). One of the widely adopted innovations in tax administration in last few decades that takes monitoring seriously is the Large Taxpayer Units (LTU), specialized tax units devoted to the monitoring of the tax liabilities of large firms. By 2004, LTUs had been established in more than 50 developing and transition countries (McCarten 2006).[2] Although increasingly popular with the practitioners and international tax experts, surprisingly, there is, to the best of our knowledge, no rigorous analysis of the effectiveness of LTUs in developing countries.[3] The LTUs provide an excellent opportunity to study the effects of higher monitoring on firms' compliance and the governments' tax revenue.

The standard models of tax evasion in the tradition of Allingham and Sandmo (1972) assume that there is complementarity between monitoring and punishment in affecting tax-payer behavior (see Polinsky and Shavell (2000), Cowell (1990), Mookherjee and Png (1994), among others). In fact, the complementarity assumption is shared by almost all models in the law and economics literature. If there is significant complementarity between monitoring and punishment, the impact of higher monitoring on tax compliance would be larger in a high punishment environment relative to a low punishment environment.[4] This implies, in particular, that the effectiveness of LTUs will vary significantly depending on the

[2] The empirical literature on tax compliance is limited due to difficulties in getting actual tax-return data and identification challenges arising from unobserved heterogeneity. One exception is Pomeranz (2011), which measures the response of taxpayers in Chile to "news" from the tax authority about an increase in their probability of being audited (Slemrod et al (2001) conducted a similar study for Minnesota).

[3] The IMF has provided technical assistance for setting-up LTUs in over 40 countries, and LTUs have been recommended by the IMF, WB and USAID (World Bank 1996; Baer et al 2002; Gallagher 2005).

[4] In the economics literature, X and Y are complementary in the function f (x,y) if f(.) is supermodular in X and Y, for continuous functions, which implies a positive cross partial derivative (see, Milgrom et al 1991; Topkis 1998). Intuitively, complementarity in our context implies that the revenue impact of a simultaneous increase in monitoring and punishment is larger than the sum of the two individual effects when implemented separately (i.e. the interaction term is positive).

punishment regime in place. It is surprising that complementarity between monitoring and punishment has been a standard assumption in the theoretical tax literature for more than 40 years since Allingham and Sandmo (1972), but, to the best of our knowledge, there is no empirical evidence on the validity and strength of such potential complementarity in tax enforcement. The existence of strong complementarity would imply that package reforms, that simultaneously raise the level of punishment and monitoring, may be more effective than piecemeal tax administration reform. We provide evidence on the complementarity assumption by analyzing the effectiveness of LTUs in low and high punishment regimes.

This paper thus provides empirical evidence on two important issues in tax reform: (i) importance of better monitoring of large firms and (ii) possible complementarity between monitoring and punishment.[5] We use a regression discontinuity design (RD) in Ecuador to estimate the causal effect of higher monitoring of LTU firms on their tax payments. The RD design exploits a discontinuity in the LTU selection process. We formulate a test for the complementarity hypothesis by combining the RD with a difference-in-difference framework based on a policy reform that instituted prison as a punishment for tax evasion in 2007. Our panel dataset includes actual annual tax-return and financial-statement information for large firms between 2003 and 2008. Tax return data is supplemented with the actual criteria used by the tax authority to choose which firms are subject to higher monitoring through their inclusion in the LTU.

In 2005 and 2007, LTU firms in Ecuador were selected using an index that was constructed as a function of firms' characteristics. A firm was eligible for the LTU only if its 'index' was greater than a cutoff value. The index assigned to firms by the tax authority incorporates firm characteristics, such as prior sales, size and tax payments. LTU firms are much more likely to be audited (i.e. higher monitoring) than non-LTU firms (see Table A1). In 2007, Ecuador enacted a major tax reform introducing harsher punishment for tax evasion. In particular, tax evasion practices were coded as a felony (rather than a misdemeanor) and punished with imprisonment of anyone involved in these (mal) practices. We compare the estimated effects of monitoring on tax revenue (using the regression discontinuity design) between the pre- and post-reform periods (2005 vs. 2007). This design effectively combines the RD design with a difference-in-difference approach to estimate the interaction effect of increasing monitoring and punishment simultaneously, thus providing a test for the existence and an estimate of the magnitude of complementarity (when it exists). Our empirical strategy

[5] In an earlier paper, we studied the effect of increased punishment on firms' compliance (Aparicio et al 2011).

relies on instrumental variables to account for the bias arising when the tax authority does not perfectly follow its LTU eligibility criteria (i.e., fuzzy RD design).

Another important aspect of our study is that we do not assume a constant policy effect across the firm distribution.[6] There are plausible theoretical reasons to expect that the direct effect of higher monitoring of LTU firms and the strength of the interaction effect may vary across different firms. The first and simplest reason for heterogeneity in the effects of higher monitoring is that the tax authority may devote more resources to the largest firms among those selected into LTU status. The large firms may also respond more to a given monitoring increase because they likely evaded (and avoided) more prior to LTU selection. The large firms may have more flexibility to misclassify in order to lower tax burden, and may be able to stretch the line between evasion and avoidance with the help of accountants and lawyers (Alm et al 2006). The heterogeneity in the monitoring intensity among the LTU firms also implies that the interaction effect for the lower tail of the tax distribution may be small. Further, the fact that smaller firms may be less risk averse and discount more the cost of higher punishment, suggests that the interaction effect may be much smaller for the firms in the lower tail. Another important reason for differential response to the higher punishment in the form of imprisonment is that some firms may be already situated at the corner solution after being selected into LTU (because of intensive monitoring even without the threat of prison).[7] To estimate the heterogeneous effects along the tax distribution, we apply the IV-Quantile framework of Chernozhukov and Hansen (2005; 2006) as adapted by Guiteras (2008) to the case of RD design.

The evidence presented later in the paper shows the following: (i) With respect to the effects of higher monitoring of LTU firms, selection into the LTU induces significant improvements in firms' tax compliance. The effects of higher monitoring are heterogeneous, with much larger effects being observed at higher quantiles of the tax distribution. (ii) The evidence on the interaction between monitoring and punishment is mixed. Estimates for the mean and median of the tax distribution show no significant complementarity. Instead, the evidence from quantile regressions shows that there is strong complementarity at the upper tail of the tax distribution. The evidence has important implications for tax policy reform in developing countries.

---

[6] The recent literature has emphasized the potential heterogeneity in "treatment effects". See for example, the discussion by Heckman, and Ravallion (2008).The standard approach to analyze potentially heterogeneous effects of a policy reform is to use quantile regressions (See, for example, Bitler et. al. (2006; 2008)).

[7] At the corner solution firms no longer evade taxes and further improvements in enforcement, such as higher punishment, have no additional effects. Smaller firms are likely to reach the corner solution first because they are likely to evade less initially, as they have less resources (lawyers and accountants) for 'creative accounting' and avoidance.

The remaining of the paper is organized as follows: Section 2 describes taxation in Ecuador, and provides additional details about the LTU eligibility criteria. Section 3 describes the data. Section 4 presents the empirical strategy. Section 5 presents the results. Section 6 presents robustness checks. Finally, section 7 concludes.

2. Taxation in Ecuador

This section describes the most relevant aspects of corporate taxation in Ecuador. All firms in the dataset pay taxes as corporations (thus, we use the terms "firm" and "corporation" interchangeably). All corporations in Ecuador are taxed 25% of their profits.[8] Reinvested profits are taxed at a rate of 15%.[9] Firms are required to distribute an additional 15% of pre-tax profits among their employees. The tax base is defined as the sum of ordinary and extraordinary revenues subject to tax, less production costs and other discounts and deductions. Profits are taxed equally regardless of whether they are retained or distributed. Each corporation assesses its own profit tax, but the tax authority may revise those assessments on subsequent inspections within specified time limits. The tax year coincides with the calendar year ("tax years" refer to the year for which tax-returns apply, rather than the year in which returns are filed). Tax returns must be filed between February 1st and May 10th following the end of the tax year. The precise deadline depends on the firms' tax registration number (Deloitte 2010).

Higher Monitoring: The Selection into LTU

Ecuador's LTU (locally known as "*Contribuyente Especial*" or "Special Taxpayer") is of great economic importance for the country. The number of LTU firms nationwide has increased from about 700 firms at program's inception (in 1995) to about 6,000 in 2009 (see Table A2). In 2009, LTU taxpayers represented more than 80% of corporate tax revenue collected by the tax administration (SRI 2008). The tax authority enhances the level of monitoring of LTU firms by assigning them more tax-administration resources and more specialized tax-staff; performing more frequent audits and cross-matching of returns with third party information; and imposing extra reporting requirements via e-filing.

Ecuador's tax administration is decentralized into eight regional offices. The Northern Office serves the Northern Region of the country, where Ecuador's capital city (Quito) is located, and it accounts for about 40 percent of the corporations in the country and

[8] A flat tax ensures that no discontinuities in reported taxes arise from changes in marginal tax rates.
[9] We control for tax discounts from capital investment by using the ratio of fixed to total assets as an additional control.

roughly the same share of LTU firms (see Table A2). In this study, we focus on firms within the Northern Office ("Regional Norte").[10] LTU firms in the Northern Office are monitored much more intensively than non-LTU firms. For instance, although LTU firms account for roughly 10 percent of firms, they account for over 80 percent of audits (see Table A1).

Between 2005 and 2007 the Northern Office added more than 700 firms into the LTU pool. As shown in Table A2, 372 firms were added in December 2005, and 292 firms were added in 2007. During these periods, the Northern Office determined LTU eligibility based on an index that it assigned to taxpayers. Firms with an index above a pre-specified cutoff were selected into the LTU pool, while firms with an index below the cutoff remained in the general tax-scheme.[11] The tax authority calculated the index as a weighted average of four variables:

$$Index_{i,t} = \phi_1 Sales_{i,t} + \phi_2 Costs_{i,t} + \phi_3 Suppliers\ and\ Customers_{i,t} + \phi_4 Taxes_{i,t}$$

Where ϕ_j is a weight for variable *j*. The decision variables are reported sales, reported costs, number of suppliers and customers, and actual tax collection. As is customary in the literature, we have normalized the index calculated by the tax authority so that the cutoff for LTU selection occurs when $Index_{i,t}$ equals zero. The normalized index takes values between *-0.85* and *0.2*.[12] Note that since firm performance changes over time and the index was calculated based on the most recently filed tax return, firms received different index values in 2005 and 2007. Although the tax authority followed its index rule quite closely (the rule was followed for over 95 percent of firms), some firms were not included in the LTU in spite of meeting eligibility criteria, and vice versa (see
Figure D1, and the bottom of Table A4).

Higher Punishment: The 2007 Tax Reform

In December 2007, the Ecuadorian constitutional assembly passed a major tax reform bill, known as "*Reform for Tax Equity*". At its core, the reform toughened sanctions for non-compliance with tax law and granted new enforcement tools to the tax authority (EIU 2009).

[10] Other offices also select LTU firms and impose higher monitoring levels. While LTU selection rules are based on the same type of index, the weights and the threshold varies from region to region. We did not have access to the specific formulas and thresholds used to select LTU firms in other regions. For this reason, we focus our study on firms within the Northern Office.

[11] During the period 2005-07, the Northern Office set a fixed cutoff for LTU eligibility (the cutoff did not adjust to reflect tax administration resources or the desired level of tax enforcement). We exclude 2006 from the analysis as few firms were eligible to the LTU (see Table A3). In 2008 the formulas and thresholds for calculating the index changed, and we do not have access to the new values

[12] Prior to the normalization, weights were chosen to ensure that the index ranged between zero and one. The contribution of each variable to the index was capped (weights became zero after a decision variable reached a given value). LTU selection generally occurred at an index value of 0.85 units.

The reform introduced prison from 1 to 6 years as a punishment for tax evasion. In addition, legal actions can no longer be extinguished through the payment of tax liabilities as occurred prior to 2007. While in the past only a company's legal representative or accountant were responsible for tax-crimes, after the reform this responsibility extends to anyone involved in a tax evasion scheme, and in particular the CFO. The new legislation also increased the fines charged to firms that fail to pay their tax obligations.[13] Note that since firms file their 2007 tax returns at the beginning of 2008, the 2007 Reform affected how firms filled their 2007 returns.[14]

3. The Data

The data were obtained from SRI administrative records. We have access to panel data from every line-item of firms' actual tax-returns between 2003 and 2008. Regardless of profitability, all corporations are required to file annual profit-tax returns, which contain information on the firm's balance sheet, income statement, and reported taxation.[15] Data on tax returns is complemented with the index value used by the tax authority to classify taxpayers into LTU and non-LTU tax schemes. As previously explained, we focus on firms in the Northern Region.

To exploit the regression discontinuity design, our analysis relies on firms for which an index value is available. An index could be obtained for over 90 percent of the population of interest. The population of interest consists of firms in the highest decile of revenue of the Northern Region, excluding firms already in the LTU pool. Small firms do not belong to the population of interest because it is unlikely that they will be eligible for the LTU. Firms already in the LTU set do not belong to the population of interest because the index eligibility rule does not apply to them; once a firm is selected to the LTU it rarely returns to the general tax-scheme regardless of its performance. Further, even if an index value were available for these firms, they would contribute little to the econometric analysis, as they may be far from the cutoff (either too small or too large). Notice, however, that since our data only includes newly selected LTU firms, our estimates are interpreted as short term impacts of LTU selection (for the first two years of program participation).

It is helpful to classify our dataset into two different sub-samples. The "low-punishment" sub-sample includes firms for which an index was calculated in 2005, two years

[13] The interest rate on arrears increased from 1.1 to 1.5 times the reference rate, and a 20% surcharge on non-declared income discovered by the SRI was introduced. We include tax arrears as a covariate in our regressions.
[14] More details about the reform can be found in Aparicio et al. 2011.
[15] Firms file profit taxes using Tax-Form 101 available at www.sri.gov.ec. Nominal variables are deflated using the CPI from the Central Bank of Ecuador.

before the reform, presumably when punishment for tax crimes was low. For this sub-sample, we expect that higher monitoring affected firms' tax reporting for fiscal years 2005 and 2006.[16] Earlier years may still be used in the empirical analysis to increase efficiency. The "high-punishment" sub-sample includes firms for which an index was calculated in 2007, when punishment increased due to the 2007 Reform.[17] For this sub-sample, we expect that higher monitoring affected firms' tax reporting for fiscal years 2007 and 2008. For comparability, only two years after LTU selection are considered for each sub-sample (as the effects of monitoring may vary over time). Each sub-sample is further divided into firms that were selected into the LTU and those that were not. To test the complementarity hypothesis we use a pooled sample covering the low punishment and the high punishment sub-samples.

The dependant variable of interest is a firm's tax compliance as measure by the level of profit taxes reported in firms' tax returns (measured in real thousands of dollars).[18] This measure also allows us to estimate the revenue implications of selecting a firm into LTU. The determinants of reported taxes are selected according to the literature on tax compliance, the literature on effective tax rates (ETRs),[19] and tax law in Ecuador. These variables include revenue and total assets as measures of firm size and performance. According to prior empirical evidence, the effect of these variables on tax compliance of a firm is, however, not unambiguous (Rice 1992; Hanlon et al 2007). Other covariates include the ratio of fixed to total assets, leverage, and costs. These variables should enter with negative signs in the tax payments equation because of the deductibility of capital investments, interest payments, and costs (Gupta and Newberry 1997). Lastly, we include tax arrears as an additional covariate.

Table A4: presents descriptive statistics for each sub-sample of data. First, we present different quantiles of the tax distribution. Then, we present mean values of the tax determinants discussed above. We notice that, in both sub-samples, firms selected to the LTU are larger and pay higher taxes than the firms that were not selected. Table A5 presents descriptive statistics exclusively for firms located near the cutoff (we define proximity to cutoff as a distance of 0.05 index-units). Comparing the tax payments of LTU and non-LTU

[16] LTU selection might have affected firms' tax reporting for the 2005 fiscal year, because firms were notified of their LTU status prior to filing their 2005 taxes (which are filed in 2006). For the year 2006 it is plausible that firms responded not only by changing their tax reporting, but also their actual behavior. As shown in Figure D6, the effect of higher monitoring on reported taxes appears to be larger in 2006 than in 2005, suggesting that changes in firm behavior may play an important role. Additional details about firms' tax reporting vs. changes in actual behavior are provided in the empirical section.

[17] Given the small number of firms that became LTUs in 2006 (63), we exclude them from the analysis and focus on the 2005 and 2007 fiscal periods.

[18] Reported taxes may be used as a measure of tax compliance because, (ceteris paribus) firms would only increase reported taxes as a response to higher punishment if they were previously under-reporting.

[19] ETRs, defined as total tax expense divided by pre-tax earnings, are a proxy for aggressive tax positions. The literature on ETRs is relevant because covariates of ETRs may also explain other measures of tax non-compliance (Hanlon et al 2007).

firms near the cutoff provides preliminary estimates of the effects of higher monitoring. Thus, higher monitoring appears to have increased tax payments by about $16,000 on average. Firms near the cutoff do not differ systematically with respect to other observable characteristics.

Comparing the impact of higher monitoring for the low-punishment and high-punishment sub-samples provides estimates of the interaction effect. However, these estimates may be unreliable if the firms in the low-punishment sub-sample differ systematically from the firms in the high-punishment sub-sample. Panel B of Table A5 tests whether firms in each of these sub-samples are comparable. Results show that firms just below the threshold in the low-punishment sub-sample have similar observable characteristics to firms just below the threshold in the high-punishment sub-sample.[20] Firms just above the threshold also have similar characteristics in both sub-samples.

4. Empirical Strategy

As explained in the introduction, the quasi-experimental design exploited in this paper combines three different quasi-experiments: two RD designs and an exogenous tax reform. The exogenous tax reform increased the level of punishment for all firms starting in 2007; thus, the reform effectively splits tax years according to the prevailing level of punishment. While the first RD design occurred during a low punishment environment (2005), the second one occurred during a high punishment environment (2007). Both RD designs arise from the discontinuity in the LTU eligibility criteria discussed earlier. The basic set-up can be summarized in the following matrix:

		Punishment	
		1st RD: Low Punishment Environment	**2nd RD: High Punishment Environment**
Monitoring	**High**	Selected into the LTU based on an Index calculated in 2005 (pre 2007 Reform)	Selected into the LTU based on an Index calculated in 2007 (post 2007 Reform)
	Low	Not selected into the LTU based on an Index calculated in 2005 (pre 2007 Reform)	Not selected into the LTU based on an Index calculated in 2007 (post 2007 Reform)

4.1. Identification

Let $Tax_{i,t}$ denote the taxes reported by firm i at time t. $Monit_{i,t}$ is a binary variable equal to unity when monitoring is high due to LTU selection, and zero otherwise. $Punish_t$ is a

[20] The only exception is total assets, as firms in the high-punishment sub-sample appear to be somewhat larger. Thus, we control for a polynomial in assets in all regressions. There are also small differences in arrears, but this can be explained by changes in the penalties and interest rates applied to late payments.

binary variable equal to unity when punishment is high due to the enactment of the 2007 reform, and zero otherwise. Firms' eligibility to higher monitoring via LTU selection is established according to the deterministic rule $1(Index_{i,t} > 0)$. $Index_{i,t}=0^{+}$ and $Index_{i,t}=0^{-}$ denote firms marginally above and marginally below the eligibility cutoff, respectively

Monitoring

In the simplest case, the mean causal effect of higher monitoring on tax compliance (γ^{M0}) can be retrieved by comparing mean reported taxes for firms marginally above and marginally below the cutoff:

$$\gamma^{M0} = E\{Tax_{i,t} \mid Index_{i,t} = 0^{+}, Punish_t = 0\} - E\{Tax_{i,t} \mid Index_{i,t} = 0^{-}, Punish_t = 0\}$$

However, if the eligibility rule is not perfectly followed by the tax authority, the discontinuity in the probability of receiving higher monitoring (φ^{M0}) no longer jumps from zero to unity at the cutoff, and it needs to be taken into account. In this case, the causal effect of higher monitoring is given by the ratio $\gamma^{M0} / \varphi^{M0}$.[21]

The main identifying assumption is that while most determinants of tax compliance vary smoothly at the eligibility cutoff, the level of monitoring varies discontinuously at that point because of the selection rule used by the tax authority. Thus, the existence of discontinuities in tax compliance at the cutoff can only be attributed to the higher monitoring that LTU selection entails. Another identifying assumption is that firms should be unable to manipulate their index to determine their own treatment status. This assumption is valid for our application because (i) the index and cutoff rules were confidential, (ii) the computation of the index was complex, (iii) firms cannot perfectly predict treatment because of the randomness in the number of new LTU firms selected yearly (see Table A2), and (iv) there is no evidence of firms bunching below the cutoff to avoid LTU eligibility.[22] Evidence supporting the validity of the requirements for identification is presented in the robustness checks section.

When the identifying assumptions hold, it has been shown that the probability of LTU selection near the cutoff should behave as if randomized (Black et al 2007). For our application, Table A5 shows that firms' observable characteristics do not differ systematically in the neighborhood of the cutoff, with the only exception of total assets. This suggests that the treatment status of the firms around the cutoff can be treated as effectively

[21] Where $\phi^{M0} = E\{Monit_{i,t} \mid Index_{i,t} = 0^{+}, Punish_t = 0\} - E\{Monit_{i,t} \mid Index_{i,t} = 0^{-}, Punish_t = 0\}$.

[22] In contrast, there is strong evidence of bunching for publically known tax thresholds in Ecuador (Carrillo et al 2011).

randomized.[23] To allay any concerns about the differences in assets, we provide a number of checks.

Note that the identifications of quantiles effects have a similar intuition. If the tax authority perfectly follows its LTU eligibility rule, the τ^{th} quantile effect of higher monitoring may be recovered by comparing the τ^{th} quantile of reported taxes for firms just below and just above the cutoff (see Figure D3 for a graphical explanation). An additional identifying assumption is that firms should not change rank after receiving treatment.[24]

Interaction between Monitoring and Punishment

The interaction between monitoring and punishment may be recovered as a difference-in-difference (DD) parameter (see Figure D4 for a graphical explanation). We compare the discontinuity in tax payments for firms around the cutoff during a low punishment environment (γ^{M0}) relative to the discontinuity for firms around the cutoff during a high punishment environment (γ^{M1}):[25]

$$\begin{aligned}\gamma^{M*P} &= \gamma^{M1} - \gamma^{M0} \\ &= \left[E\{Tax_{i,t} \mid Index_{i,t} = 0^{+}, Punish_t = 1\} - E\{Tax_{i,t} \mid Index_{i,t} = 0^{-}, Punish_t = 1\}\right] - \\ &\quad \left[E\{Tax_{i,t} \mid Index_{i,t} = 0^{+}, Punish_t = 0\} - E\{Tax_{i,t} \mid Index_{i,t} = 0^{-}, Punish_t = 0\}\right]\end{aligned}$$

As in any DD analysis, the main identifying assumption is that firms in the pre-reform period should be a good counterfactual for the behavior of firms in the post-reform period in the absence of the reform (parallel trends assumption).

This assumption is likely to hold in our application because (i) the cutoff for LTU eligibility did not change over time, and (ii) throughout the span of our dataset most of the relevant tax policy parameters have remained unchanged (other than the 2007 Reform). Further, Panel B of Table A5 shows that firms' observable characteristics around the cutoff do not differ systematically between the low punishment and high punishment sub-samples (with the exception of total assets).[26] Another identifying assumption is that higher punishment should not affect firms above and below the cutoff differently, for reasons other than the interaction between monitoring and punishment (i.e. potential complementarity). This assumption would be violated if the LTU firms are less responsive to changes in punishment because they have more accountants and lawyers. However, the evidence shows

[23] Note that means tests of covariates are stricter than the weak smoothness conditions needed for identification.
[24] The approach is somewhat more complex when the tax authority does not perfectly follow its eligibility rule.
[25] As before, it is also possible to recover the parameters of interest when the LTU eligibility is not perfectly followed by the tax authority, and to recover quantile estimates.
[26] Differences in arrears can be explained by higher interest rates on arrears introduced in 2007.

that this is not the case; as shown in Figure D2 there is no discontinuity at the cutoff in the number of high skilled employees.

4.1. Regression Analysis

We now turn to a more formal regression analysis. Relative to direct comparisons of means (or quantiles) around the cutoff, a regression framework has the advantage of exploiting more data to gain efficiency, and relaxing the assumption that the index does not affect tax compliance in the arbitrary interval near the cutoff. However, the intuition for the identification of parameters of interest remains the same as discussed above (Angrist and Lavy 1999; Hahn et al 2001). The analysis presented below builds upon a fully parametric approach, which we found to provide a good fit for the data at hand.

We introduce additional notation. $Trend_t$ is a linear trend that accounts for any unobserved determinants of $Tax_{i,t}$ changing systematically over time.[27] The continuous relationship between $Tax_{i,t.}$ and the forcing variable $Index_{i,t}$ is captured by a p^{th} order parametric polynomial function $f(Index_{i,t})$, whose parameters may be restricted to be the same or allowed to differ to the right and left of the cutoff. $Above_{i,t}$ is a binary indicator equal to unity if a firm qualifies to the LTU according to the index rule and zero else. Note that if the tax authority does not follow the eligibility rule perfectly, $Monit_{i,t}$ may be unity when $Above_{i,t}$ is zero, and vice versa. In the robustness checks section, we also account for additional covariates described in the data section, $x_{i,t}$.

To exploit the panel nature of the dataset and to account for fixed effects the time-varying variables have been demeaned. Thus, the dependant variable used in the regression analysis is $\tilde{T}ax_{i,t} = Tax_{i,t} - \sum Tax_{i,t} / t = Tax_{i,t} - \overline{Tax_i}$. The symbol "~" is used to indicate that variables have been demeaned in order to control for individual fixed-effects. In the RD design, an important rationale for using fixed effects is to reduce sampling variance.[28]

Table A5 and Figure D9 suggest that assets may not be continuous at the cutoff. The relationship between assets and the index may exhibit jumps because the sample at hand focuses on the country's largest firms, which are likely to be quite heterogeneous. Further, while the index incorporates variables such as prior revenue and prior costs, it does not

[27] Given that the 2007 Reform affected all private firms, its effect can only be estimated as a deviation from trend. Thus, it is essential that the underlying trend in corporate profit taxes is adequately accounted for. The trend may be driven by macroeconomic factors or overall improvements in tax administration. It has been shown that a linear trend is a good fit for the underlying trend in corporate profit taxes in Ecuador (see Aparicio et al 2011).

[28] Another alternative to reduce sampling variance is including a lag of the dependant variable as a covariate (Lee and Lemieux (2009)), as done in the robustness checks section.

include any measure of assets. Thus, the index alone may not be enough to properly account for asset differences between groups of firms. Since one may argue that any observed discontinuity at the cutoff in reported taxes may be driven by differences in asset values at that point; we control for assets in all of our regressions. We model the potentially non-linear relationship between $Tax_{i,t}$ and $Assets_{i,t}$, as a symmetric third order polynomial $f(Assets_{i,t})$.[29]

Monitoring

The econometric model used to estimate the impact of higher monitoring on tax compliance during the low-punishment environment is presented below:

$$\tilde{T}ax_{i,t} = \delta^{M} Monit_{i,t} + f(Index_{i,t}) + \alpha_0 + \alpha_1 trend_t + x'_{i,\bar{t}}\,\beta + f(Assets_{i,\bar{t}}) + u_{i,t} \quad (1)$$

The main parameter of interest (δ^M) captures any discontinuity in reported taxes at the cutoff arising from higher monitoring, and is interpreted as a weighted treatment effect.

Since the tax authority does not follow the LTU eligibility criteria perfectly, the level of monitoring is potentially endogenous.[30] The tax authority chooses (possibly based on criteria correlated with outcomes) which firms are not selected to the LTU in spite of meeting the eligibility criteria, and vice versa. Thus, OLS estimation of Equation (1) results in biased estimates. Unbiased estimates may be obtained, however, by taking advantage of the link between RD and instrumental variables (IV). $Monit_{i,t}$ in Equation (1) is instrumented by $Above_{i,t}$ using a first-stage regression of the form:

$$\begin{aligned} Monit_{i,t} = {} & \phi\, Above_{i,t} + f(Index_{i,t}) \\ & + \alpha_0 + \alpha_1 trend_t + x'_{i,t}\,\beta + f(Assets_{i,t}) + \varepsilon_{i,t} \end{aligned} \quad (2)$$

We correct the estimated variance-covariance matrix for heteroskedasticity of arbitrary form, which is likely to arise under the linear probability model above. The main parameter of interest (δ^M) is now interpreted as a local average treatment effect (LATE).

We also estimate the impact of higher monitoring on tax compliance at different quantiles of the tax distribution. We think of Equation (1) as a simplification of the structural quantile response function into a piecewise linear model (where each parameter depends on the specific quantile being estimated). The model is first estimated using Koenker and Basset's (1978) quantile regression (QR). Then, we account for any possible bias arising from the endogeneity of $Monit_{i,t}$ by exploiting the link between IV and RD. We implement a

[29] See also the discussion in the Robustness Checks Section.

[30] This gives rise to a Fuzzy Regression Discontinuity design (see Angrist and Lavy 1999; Hahn et al 2001; Lee and Lemieux 2010; Cook and Wong 2008; Imbens and Lemieux 2008).

recent econometric approach developed by Guiteras (2008) that extends the IV-Quantile framework of Chernozhukov and Hansen (2005; 2006) to regression discontinuity (IQR).[31]

At this stage, it is important to consider the timing when firms received notifications of their selection to the LTU (see Table A3). During the low punishment environment, most firms received LTU notifications in December 2005, allowing them to adjust how they filed their 2005 tax-returns but not their actual behavior for that tax year (behavior could be adjusted for 2006). Instead, during the high punishment environment, most firms received LTU notifications in March, September or October[32] of 2007, giving them more time to adjust their behavior. Fortunately, we can use the variation in the months in which LTU notifications were received in 2007 to test whether small differences in the timing of notifications matters. Relying on the high punishment sub-sample, we interact $Monit_{i,t}$ in Equation (1) with a dummy equal to unity if LTU notifications were received in March, as follows:

$$\begin{aligned}\widetilde{Tax}_{i,t} = {} & \delta^{M} Monit_{i,t} + \delta^{MD} Monit_{i,t} * March_{i} \\ & + f(Index_{i,t}) + \alpha_0 + \alpha_1 trend_t + x_{i,t}\beta_T + f(Assets_{i,t}) + u_{i,t}\end{aligned} \quad (3)$$

If δ^{MD}=0 small differences in the timing of notifications are irrelevant, and comparing RD estimates obtained from the low and high punishment sub-samples should be reliable. As before, we also account for endogeneity by instrumenting $Monit_{i,t}$ and $Monit_{i,t}$*$March_{i,t}$ with $Above_{i,t}$ and $Above_{i,t}$*$March_{i,t}$.

Interaction between Monitoring and Punishment

The econometric model used to estimate the potential complementarity between monitoring and punishment is presented below:

$$\begin{aligned}\widetilde{Tax}_{i,t} = {} & \delta^{M} Monit_{i,t} + \delta^{P} Punish_t + \delta^{P*M} Monit_{i,t} * Punish_t \\ & + \alpha_0 + \alpha_1 trend_t + f(Index_{i,t}) + x_{i,t}\beta + f(Assets_{i,t}) + u_{i,t}\end{aligned} \quad (4)$$

The parameter δ^{M} measures the causal effect of higher monitoring on tax compliance; δ^{P} measures the impact of higher punishment as a deviation-from-trend; δ^{P*M} captures the interaction between monitoring and punishment. δ^{P*M} is a difference-in-difference parameter, effectively comparing the discontinuous jumps in $\widetilde{Tax}_{i,t}$ arising from two RD

[31] The econometric literature on quantile regression discontinuity is of recent origin (Frolich and Melly 2008).

[32] Note that the March, 2007 notification were received during the period for filing tax returns for 2006. We find no evidence that 2006 returns were affected for firms receiving the LTU notification in March. Given that the actual deadline for filing varies by firm, it is possible that the deadline had already passed or there was not enough time for firms to change their tax returns. Index values for firms selected in October are not available; consequently, these firms were excluded from the analysis.

designs. Any potential complementary between monitoring and punishment, beyond the additive individual effects, would imply that $\delta^{P*M}>0$.[33] As before, we also account for endogeneity by instrumenting $Monit_{i,t}$ and $Monit_{i,t}*Punish_t$ with $Above_{i,t}$, and $Above_{i,t}*Punish_t$. We also estimate the interaction effect at different quantiles of the tax distribution using QR and IQR.

Estimators that do not use instruments (OLS and QR) may be biased downwards for two reasons. First, bias may arise because the tax authority does not follow the LTU eligibility rule perfectly. However, this source of bias is likely to be small as the tax authority follows the eligibility rule quite closely. Second, there may be attenuation bias due to measurement error. Our measure of the level of monitoring is quite precise because we have detailed records indicating when notifications of LTU status were received by firms. In contrast, our measure of the level of punishment may be imprecise (firms may interpret the 2007 law differently). Thus, OLS and QR estimates of the interaction between monitoring and punishment are likely to suffer from attenuation bias.

5. Results

Monitoring: The Effects of Selection into LTU

We start by presenting preliminary evidence regarding the impact of higher monitoring on tax compliance. Panel A of Figure D5 shows the average of reported taxes for different intervals of the index, relying exclusively on data from the low-punishment sub-sample. Thanks to the availability of panel data, the same firms are shown in different years. We notice a discontinuity in reported taxes at the cutoff after LTU selection (2005 and 2006), but not in prior years (2003 and 2004). Similar evidence is obtained for the high punishment sub-sample as well (see Panel B of Figure D5). Panel C of Figure D5 shows that there is a large discontinuity in median reported taxes at the cutoff, but there is no apparent discontinuity for the 25th quantile, suggesting that the impact of higher monitoring is heterogeneous.

We now discuss the effects of higher monitoring of LTU firms from the estimation of Equation (1) using regression analysis. Results are shown in Table B1 and Table B3. For brevity, we report coefficient estimates exclusively for the effects of higher monitoring on reported taxes (δ^{M0}). In all models, time-varying variables are demeaned to exploit the panel nature of the data, and a cubic polynomial of assets is included in order to control for

[33] Note that this specification forces the coefficients on $f(Index_{i,t})$ to be the same across sub-samples. Similar results are obtained from more flexible models allowing for different coefficients.

heterogeneity in this variable. We model $f(Index_{i,t})$ using eight different polynomial functions. The order of the polynomials ranges from linear to fourth order. For the first four models, the parameters of the polynomial function are forced to be identical on both sides of the cutoff, while for the remaining four models they are allowed to vary.[34] At this stage we do not include any additional controls, $x_{i,t.}$

Table B1 presents results obtained from OLS and 2SLS estimation. Estimates of the causal effect of higher monitoring on tax compliance are statistically significant different from zero at the 1% level across the board. With respect to the magnitude of the effects, all models suggest that firms more than doubled their tax payments after being selected into the LTU, an increase of roughly USD 10,000 (see Figure D6). Table B3 presents results obtained from QR and IQR estimation. The impact of higher monitoring on reported taxes is, in general, positive and statistically significant at the median; but it is much larger at the 90th quantile, with estimates over USD 20,000 (see Figure D6).[35]

We also present results from a falsification test. This test thus compares the tax payments between firms that will later become LTU and those that will not, but prior to the LTU selection. Results show that, prior to the LTU selection (i.e. before the firms received the treatment), there was no statistically significant discontinuity it reported taxes at the cutoff. This is strong evidence that the response observed after selection into LTU is not driven by any unobserved heterogeneity.

As shown in Table B2, similar evidence is obtained for the high punishment sub-sample as well. Further, results presented in Table B2 show that there is no evidence that firms receiving LTU notifications in different months respond differently to higher monitoring. The coefficients for the interaction term $Monitoring_{i,t}*March_i$ in Equation (3) are not statistically significant.

Interaction between Monitoring and Punishment

In this section we test for complementarity between monitoring and punishment by pooling together firms proposed for LTU selection in 2005 (low-punishment sub-sample) and firms proposed for LTU selection in 2007 (high punishment sub-sample). Results from the

[34] Thus p index parameters are estimated for the first four models of Table 1, while $2p$ index parameters are estimated for the last four models of Table 1 for a specification with a p^{th} order polynomial.

[35] The preferred estimates are based on 2nd and 3rd order polynomials. They are sufficiently flexible to avoid model misspecification, and they are, with asymmetric polynomial functions to the right and left of the cutoff, provide better fit for the data These asymmetric polynomial functions are also more flexible and less likely to be miss-specified. The coefficient of the 4th-order polynomial term is not statistically significant (and figures suggest that the fourth order term is redundant). Tests of symmetric polynomials versus asymmetric polynomials reject the null hypothesis of a symmetric polynomial function about the threshold at the 10% level.

OLS and 2SLS estimation of Equation (4) are presented in Table B4. For brevity, we only report the estimates for the main parameters of interest: the impact of higher monitoring on tax compliance (δ^{M0}), the impact of higher punishment (δ^{P}), and potential complementarity between the two (δ^{M*P}). Interestingly, the evidence indicates that there is no statistically significant complementarity between monitoring and punishment at the mean.[36] However, as discussed in the introduction, there are plausible theoretical reasons to expect that the impact of complementarity is heterogeneous along the tax distribution. Table B5 presents estimates of Equation (4) for different quantiles (median, 70th quantile and 90th quantile) using QR and IQR. We find that the interaction effect is not statistically significant at the median, but it is positive and significant at the 70th and 90th quantiles.

6. Robustness Checks

We perform two types of robustness checks. First, we test for possible miss-specification of the benchmark models discussed in the previous section. Second, we test that the requirements for the validity of the RD methodology are met.

Miss-specification

We start by replicating some of our benchmark results using smaller windows of data around the cutoff.[37] Using data in smaller windows makes it less likely that the polynomial function $f(Index_{i,t})$ is miss-specified, and it contrasts our parametric polynomial estimation with a kind of non-parametric local linear regression approach where the bandwidth is varied (Lee and Lemieux 2010). The benchmark window for OLS and 2SLS estimates presented in the previous section ranged from -.85 to .2. Alternative windows used in this section include -.4 to .2 and -.2 to .2.[38] Table C1 shows that our coefficient estimates of interest remain robust to changes in the window size.

We replicate some of our benchmark results including additional determinants of reported taxes as covariates. Model 1 exploits the panel structure of the data by including a lagged value of the dependant variable as another baseline covariate in period *t* to reduce

[36] By construction, the monitoring effect is very similar to the one obtained when using only the low-punishment sub-sample.

[37] We focus on 2nd and 3rd-order polynomials, which provide a somewhat better fit, as they are jointly statistically significant.

[38] Given that the majority of the data is below the threshold, rather than using symmetric windows for the index variable, only observations at the left of the threshold are excluded.

sampling error (Lee and Lemieux 2010).[39] Model 2 includes revenue, and costs. Model 3 adds financial ratios (leverage and fixed-to-total-assets) and tax arrears to the previous model. Table C2 shows that in general results remain robust to including additional covariates (coefficients of interest remain statistically significant). However, the estimated impacts are often smaller in magnitude.

Validity of the RD Methodology

For brevity, most tests for the validity of the requirements of the RD design are presented graphically (the corresponding tables are available from the authors upon request). First, we test for a potential discontinuity in the distribution of firms around the cutoff. As previously discussed, we are particularly worried about bunching of taxpayers below the cutoff to avoid LTU treatment. A histogram of the index and McCrary's test of continuity are presented in Figure D8. This figure suggests that the firm distribution does not exhibit any unusual discontinuities.

We also test for discontinuities in covariates. Figure D9 presents parametric graphs of the relationship between covariates and the index. In all cases, this relationship is modeled as a symmetric second order polynomial. Results show that there is no statistically significant discontinuity at the cutoff for any of the covariates. However, we notice that, although not statistically significant, the magnitude of the discontinuity in assets at the cutoff is quite large. Nevertheless, the magnitude of the discontinuity in assets is greatly reduced once we control for additional variables such as a lagged value of assets. This suggests that the relationship between assets and the index is actually smooth at the cutoff once the heterogeneity in assets is properly accounted for, and consequently our results are reliable.

We test for discontinuities at placebo cutoff values, other than the actual cutoff used by the tax authority to determine LTU eligibility, i.e. $I(Index_{i,t}=0)$. For this purpose we run a large number of regressions, as follows:

$$\tilde{T}ax_{i,t} = \delta\, Test_{i,t} + f(Index_{i,t}) + \alpha_0 + \alpha_1 trend_t + \beta\, Tax_{i,t-1} + f(Assets_{i,t}) + u_{i,t} \quad (5)$$

Where the variable $Test_{i,t}$ is a dummy equal to unity for firms with an index value above a given placebo cutoff. The placebo cutoff value changes for each regression. We cover the entire domain of possible values of the index, in turn. Note that Equation (5) is very similar to Equation (1), except that the placebo dummy variable ($Test_{i,t}$) replaces the actual dummy for firms selected to the LTU ($Monit_{i,t}$). We also test for discontinuities in the probability of LTU

[39] A two-period lag is used because firms may take over one year to fully respond to the impact of higher monitoring.

selection at different placebo cutoffs. Figure D10 shows that large discontinuities occur only when using the actual cutoff used by the tax authority for LTU eligibility.

Although we do not have the necessary data to test that the treatment dose is sufficiently large (i.e. that monitoring increases at the cutoff due to LTU selection), we have evidence that audit probability is much higher for LTU firms (see Table A1).

7. Conclusions

This paper provided credible evidence on two important issues in tax reform: (i) importance of better monitoring of large firms and (ii) possible complementarity between monitoring and punishment. (i) With respect to the effects of higher monitoring of LTU firms, we find that selection into the LTU induces significant improvements in firms' tax compliance. The effects of higher monitoring are heterogeneous, with much larger effects being observed at higher quantiles of the tax distribution. (ii) The evidence on the interaction between monitoring and punishment is mixed. Estimates for the mean and median of the tax distribution show no significant complementarity. Instead, the evidence from quantile regressions shows that there is strong complementarity at the upper tail of the tax distribution.

There are plausible theoretical explanations for the results obtained in this paper. With respect to the effectiveness of better monitoring, large firms may improve compliance the most because they evaded more prior to improvements in tax enforcement and/or because audit probability may be particularly concentrated on the largest LTU firms. With respect to the interaction between monitoring and punishment, we may not observe any complementarity for the lower tail of the tax distribution because the prison effect is small for the smallest LTU firms. Alternatively, the increase in monitoring from LTU treatment may be enough to take smaller LTU firms to the corner solution. At the corner solution firms no longer evade and further improvements in enforcement, such as higher punishment, have no additional effects.

The evidence presented has important implications for tax policy reform in developing countries. First, we conclude that LTUs may be a cost-effective approach for improving tax compliance. While monitoring is costly, the large impacts obtained in this paper, suggest that the benefits of higher monitoring of LTU firms are likely to offset the costs of the LTU program.[40] However, since the benefits of LTU selection are smaller for smaller firms, the marginal benefits of continuing to expand the LTU will eventually become negative. Thus, it is critical to develop a sound LTU eligibility criterion that considers both

[40] Unfortunately, however, we do not have an exact measure of the actual costs.

the benefits and costs of such a program. Note, however, that the results obtained in this paper apply only to large firms, and they may not generalize to other populations such as informal firms.[41]

Second, we conclude that piece-wise tax administration reform may be ineffective. Given that the success of one piece of reform depends on the value of other tax parameters, tax administration reform should take into account multiple aspects of the tax environment.

[41] In the case of small informal firms, the benefits of higher monitoring may be small and it is possible that the gains may not be enough to offset any costs of higher monitoring. However, further research is needed in this area.

A. Descriptive Tables

Table A1: Audit Processes, Northern Office (2005-08)

	Share of Corporations	*Share of Audits*
LTU	11%	83%
Non-LTU	89%	17%
Total	100%	100%

Table A2: New LTU Firms (1995-09)

	Largest Office (Northern)	2nd Largest Office (Litoral)	Other (6 Offices)	Total (8 Offices)	Northern/ Total
1995	395	303	25	723	55%
1996	121	297	305	723	17%
1997	77	3	0	80	96%
1998	117	69	20	206	57%
1999	237	253	196	686	35%
2000	267	8	156	431	62%
2001	28	28	74	130	22%
2002	3	53	25	81	4%
2003	32	3	29	64	50%
2004	137	64	70	271	51%
2005	372	13	69	454	82%
2006	63	55	43	161	39%
2007	292	239	139	670	44%
2008	321	393	143	857	37%
2009	131	204	39	374	35%
Statistics					
cumulative (all years)	2,593	1,985	1,333	5,911	44%
mean (all years)	173	132	89	394	

Notes: The table shows the number of new taxpayers selected to the LTU each year. The cumulative figure refers to the sum of new LTU taxpayers selected over 1995-09. The paper focuses on fiscal years 2005 and 2007.

Table A3: New LTU Firms, Northern Office (2005-07)

	Jan	Feb	March	April	May	June	July	Aug	Sept	Oct	Nov	Dec	Total
2005	0	0	0	3	0	0	0	0	7	0	0	362	372
2006	4	0	0	0	2	1	1	0	54	0	0	1	63
2007	0	0	134	0	0	0	0	0	42	116	0	0	292

Notes: The light-gray shading shows months when returns for the previous tax year are due. The dark-gray shading shows months in which a large number of firms received notifications of LTU selection.

Table A4: Descriptive Statistics

	Low Punishment (2005/06)			High Punishment (2007/08)		
	LTU	non-LTU	Diff.	LTU	non-LTU	Diff.
	A	**B**	**A-B**	**C**	**D**	**C-D**
Real Taxes $000						
Mean	31	7	24.2***	40	14	25.9***
Median	13	4	9.7***	14	5	8.9***
60th	19	5	14.0***	29	8	21.4***
70th	28	7	20.8***	41	12	29.3***
80th	43	10	33.0***	55	19	36.1***
90th	84	16	67.7***	103	33	70.6***
99th	254	59	195.6***	499	130	369.1***
Covariates (Mean)						
Revenue $000	4,918	987	3,931.5***	3,613	1,353	2,259.5***
Local Purchases	0.7	0.7	0.0	0.5	0.6	-0.1***
Imports	0.3	0.4	-0.0**	0.2	0.3	-0.1***
Leverage	0.7	0.8	-0.1***	0.7	0.8	-0.1
Fixed to total assets	0.3	0.2	0.1***	0.3	0.2	0.1***
Arrears $000	30	9	21.3***	17	6	11.5***
Total Assets $000	4,498	1,107	3,391.7***	3,562	1,578	1,983.5***
Firm Obs.	335	1,670		120	2,024	
Follows rule	305	1,651		120	2,021	
Not follows rule	30	19		0	3	
Obs. (2 years *x* firms)	670	3,340		240	4,048	

Notes: The table presents descriptive statistics for the data at hand, focusing on firms that were considered for LTU selection in fiscal years 2005 and 2007. The panel to the left focuses on firms that were considered for LTU selection by the tax authority in 2005, two years before the enactment of harsher punishment for tax crimes. The panel to the right focuses on firms that were considered for LTU selection in 2007, after harsher punishment was enacted. For comparability, each panel in the table relyes on two years of data after firms were considered for LTU selection (i.e. after an index was assigned). The bottom of the table provides information about the sample size.

.

Table A5: Descriptive Statistics near the Cutoff (Window -0.05 to 0.05)

Panel A

	Low Punishment (2005/06)			High Punishment (2007/08)		
	LTU	non-LTU	Diff.	LTU	Non-LTU	Diff.
	A	**B**	**A-B**	**C**	**D**	**C-D**
Real Taxes $000						
Mean	23	7	16.6**	37	20	17.0**
Median	11	2	9.0***	18	9	8.3
60th	15	5	10.3***	28	18	10
70th	22	7	15.0**	42	26	15.6*
80th	30	11	19.5**	53	36	17.4**
90th	50	23	27.2	103	54	49.4**
99th	253	33	219.9***	185	111	74.1
Covariates (Mean)						
Revenue ($000s)	2,709	2,242	466.8	3,409	3,031	378
Local Purchases	0.7	0.8	-0.1	0.6	0.6	0.0
Imports	0.3	0.3	0.0	0.3	0.3	0.0
Leverage	9.8	7.0	8.8	0.7	0.9	-0.2*
Fixed to total assets	0.3	0.2	0.1*	0.2	0.2	0.0
Arrears ($000s)	17.6	17.7	-0.2	10.8	8.9	1.9
Total Assets ($000s)	1,798	694	1,104***	2,720	1,529	1,191**
Obs. (years *x* firms)	150	32		107	102	

Panel B

	Low Punishment (2005/06)		High Punishment (2007/08)		Differences	
	LTU	non-LTU	LTU	non-LTU	Diff.	Diff.
	A	**B**	**C**	**D**	**C-A**	**D-B**
Covariates (Mean)						
Revenue ($000s)	2,709	2,242	3,409	3,031	700	789
Local Purchases	0.7	0.8	0.6	0.6	0.0	-0.1
Imports	0.3	0.3	0.3	0.3	0.0	0.0
Leverage	9.8	7.0	0.7	0.9	-9.1	-0.1
Fixed to total assets	0.3	0.2	0.2	0.2	0.0	0.0
Arrears ($000s)	17.6	17.7	10.8	8.9	-6.8**	-8.9*
Total Assets ($000s)	1,798	694	2,720	1,529	922**	835**
Obs. (years *x* firms)	150	32	107	102		

Notes: Firms for which the tax authority does not follow the index eligibility rule are excluded.

B. Main Result Tables

Table B1: Monitoring – Low Punishment Sub-sample

Panel A. Benchmark Results

	Dependant Variable: Tax $000 (demeaned)							
	Index: Symmetric Polynomials				Index: Asymmetric Polynomials			
	Linear	2nd	3rd	4th	Linear	2nd	3rd	4th
OLS								
Monitoring	**7.54*** **	**8.88*** **	**9.01*** **	**9.10*** **	**8.28*** **	**9.06*** **	**9.35*** **	**9.48*** **
	(1.13)	(1.24)	(1.26)	(1.26)	(1.19)	(1.26)	(1.29)	(1.31)
2SLS								
Monitoring	**7.75*** **	**9.33*** **	**9.51*** **	**9.61*** **	**8.65*** **	**9.60*** **	**9.96*** **	**10.14*** **
	(1.21)	(1.35)	(1.38)	(1.38)	(1.29)	(1.38)	(1.44)	(1.46)
Obs.	8,364	8,364	8,364	8,364	8,364	8,364	8,364	8,364

Panel B. Falsification Tests

	Dependant Variable: Tax $000 (demeaned)							
	Index: Symmetric Polynomials				Index: Asymmetric Polynomials			
	Linear	2nd	3rd	4th	Linear	2nd	3rd	4th
OLS								
Monitoring	**-0.11**	**-0.09**	**-0.12**	**-0.15**	**-0.12**	**-0.11**	**-0.17**	**-0.17**
	(0.58)	(0.67)	(0.76)	(0.72)	(0.66)	(0.70)	(0.71)	(0.73)
2SLS								
Monitoring	**-0.08**	**0**	**-0.04**	**-0.08**	**-0.07**	**0**	**-0.11**	**-0.05**
	(0.65)	(0.99)	(1.37)	(1.30)	(0.91)	(1.41)	(1.74)	(2.44)
Obs.	4,354	4,354	4,354	4,354	4,354	4,354	4,354	4,354

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively. The table presents coefficient estimates based on Equation (1). Fixed-effects are accounted for by demeaning time-varying variables. A third order polynomial of assets is included as an additional control. In the first four columns, the polynomial functions are forced to have identical parameters to the left and the right of the threshold, but they are allowed to differ in the last four. Each of these eight models is estimated using OLS and 2SLS. For the 2SLS regression, the instrument is a dummy equal to 1 if $Index_{i,t}>0$. First-stage regressions are linear probability models. Heteroskedasticity robust standard errors are shown in parenthesis.

Only data from the low punishment sub-sample is used. The window is set between -.85 to .2. The low-punishment sample has slightly over 2,000 firms per year. Panel A includes data for years 2003-06, thus there are over 8,000 observations (2,000 firms per year * 4 years = 8,000 observations). Panel B only includes data prior to firms being selected to the LTU, 2003-04, thus there are roughly 4,000 observations (2,000 firms per year * 2 years = 4,000 observations).

Table B2: Monitoring – High Punishment Sub-sample

Panel A. Benchmark Results

	Dependant Variable: Tax $000 (demeaned)							
	Symmetric Polynomials				Asymmetric Polynomials			
	Linear	2nd	3rd	4th	Linear	2nd	3rd	4th
OLS								
Monitoring	**14.68****	**15.89****	**16.63****	**16.67****	**16.31****	**17.11****	**17.24****	**17.43****
	(6.92)	(6.89)	(6.85)	(6.83)	(6.80)	(6.93)	(6.94)	(6.93)
Monitoring *x* March	**-7.9**	**-7.98**	**-8.06**	**-8.08**	**-8.22**	**-8.33**	**-8.14**	**-8.16**
	(7.32)	(7.32)	(7.32)	(7.31)	(7.31)	(7.31)	(7.24)	(7.24)
2SLS								
Monitoring	**13.69***	**15.02****	**15.91****	**15.97****	**15.59****	**16.41****	**16.51****	**16.76****
	(7.14)	(7.12)	(7.07)	(7.04)	(7.00)	(7.14)	(7.14)	(7.13)
Monitoring *x* March	**-6.93**	**-7.17**	**-7.4**	**-7.43**	**-7.55**	**-7.68**	**-7.48**	**-7.55**
	(7.51)	(7.52)	(7.51)	(7.49)	(7.48)	(7.49)	(7.40)	(7.41)
Obs.	13,274	13,274	13,274	13,274	13,274	13,274	13,274	13,274

Panel B. Falsification Tests

	Dependant Variable: Tax $000 (demeaned)							
	Symmetric Polynomials				Asymmetric Polynomials			
	Linear	2nd	3rd	4th	Linear	2nd	3rd	4th
OLS								
Monitoring	**-0.06**	**-0.05**	**-0.03**	**-0.05**	**-0.07**	**-0.05**	**-0.03**	**-0.05**
	(1.16)	(1.18)	(1.39)	(1.48)	(1.44)	(1.72)	(1.87)	(1.93)
Monitoring *x* March	**-0.01**	**-0.01**	**-0.01**	**0.01**	**0.01**	**-0.01**	**-0.01**	**-0.01**
	(1.43)	(1.42)	(1.42)	(1.41)	(1.41)	(1.40)	(1.42)	(1.45)
2SLS								
Monitoring	**-0.03**	**-0.02**	**0.01**	**0.01**	**-0.03**	**0.01**	**0.04**	**0.03**
	(1.24)	(1.29)	(1.64)	(1.79)	(1.72)	(2.21)	(2.47)	(2.81)
Monitoring *x* March	**-0.03**	**-0.03**	**-0.04**	**-0.03**	**-0.03**	**-0.04**	**-0.04**	**-0.04**
	(1.49)	(1.48)	(1.47)	(1.47)	(1.48)	(1.50)	(1.50)	(1.60)
Obs.	8,800	8,800	8,800	8,800	8,800	8,800	8,800	8,800

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively. The table presents coefficient estimates based on Equation (1). Fixed-effects are accounted for by demeaning time-varying variables. A third order polynomial of assets is included as an additional control. In the first four columns, the polynomial functions are forced to have identical parameters to the left and the right of the threshold, but they are allowed to differ in the last four. Each of these eight models is estimated using OLS and 2SLS. For the 2SLS regression, the instrument is a dummy equal to 1 if $Index_{i,t}>0$. First-stage regressions are linear probability models. Heteroskedasticity robust standard errors are shown in parenthesis.

Only data from the high punishment sub-sample is used. The window is set between -.85 to .2. The high-punishment sample has about 2,200 firms per year. Panel A includes data for years 2003-08, thus there are over 13,200 observations (2,200 firms per year * 6 years = 13,200 observations). Panel B only includes data prior to firms being selected to the LTU, 2003-06, thus there are roughly 8,800 observations (2,200 firms per year * 4 years = 8,800 observations).

Table B3: Heterogeneous Monitoring – Low Punishment Sub-sample

	Dependant Variable: Tax $000							
	Symmetric Polynomials				Asymmetric Polynomials			
	Linear	2nd	3rd	4th	Linear	2nd	3rd	4th
QR, Median								
Monitoring	**0.79*** **	**1.16*** **	**1.11*** **	**1.15*** **	**0.98*** **	**1.14*** **	**1.18*** **	**1.22*** **
	(0.15)	(0.13)	(0.14)	(0.13)	(0.15)	(0.12)	(0.14)	(0.12)
IQR, Median								
Monitoring	**0.85*** **	**0.32*** **	**na**	**na**	**na**	**na**	**na**	**na**
	(0.03)	(0.03)	na	na	na	na	na	na
QR, 70th								
Monitoring	**3.23*** **	**3.23*** **	**3.23*** **	**3.23*** **	**3.23*** **	**3.23*** **	**3.23*** **	**3.23*** **
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
IQR, 70th								
Monitoring	**3.52*** **	**3.52*** **	**1.05*** **	**3.35*** **	**na**	**na**	**na**	**na**
	(0.03)	(0.03)	(0.03)	(0.03)	na	na	na	na
QR, 90th								
Monitoring	**8.89*** **	**9.16*** **	**8.88*** **	**8.98*** **	**9.13*** **	**9.10*** **	**8.52*** **	**7.52*** **
	(0.82)	(0.83)	(0.94)	(0.92)	(0.85)	(0.91)	(0.88)	(0.83)
IQR, 90th								
Monitoring	**9.66*** **	**9.04*** **	**9.43*** **	**9.67*** **	**na**	**na**	**na**	**na**
	(0.02)	(0.02)	(0.02)	(0.03)	na	na	na	na
Obs.	3,421	3,421	3,421	3,421	3,421	3,421	3,421	3,421

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively. The table presents coefficient estimates based on Equation (1). Fixed-effects are accounted for by demeaning time-varying variables. A third order polynomial of assets is included as an additional control. In the first four columns, the polynomial functions are forced to have identical parameters to the left and the right of the threshold, but they are allowed to differ in the last four. Each of these eight models is estimated using QR and IQR for different quantiles of the tax distribution. For the IQR regressions, the instrument is a dummy equal to 1 if $Index_{i,t}>0$. NA indicates that the IQR method does not converge.

Only data from the low punishment sub-sample is used. The window is set between -.4 to .2 in order to avoid convergence problems. The sample size is smaller than in Table B1 because a smaller window is used.

Table B4: Interaction – Pooled Sample

	Dependant Variable: Tax $000 (demeaned)							
	Symmetric Polynomials				Asymmetric Polynomials			
	Linear	2nd	3rd	4th	Linear	2nd	3rd	4th
OLS								
Monitoring *x* Prison	**2.5**	**2.44**	**2.31**	**2.22**	**2.14**	**2.37**	**2.55**	**2.6**
	(2.63)	(2.63)	(2.63)	(2.63)	(2.63)	(2.64)	(2.66)	(2.66)
Monitoring	**6.86*** **	**8.03*** **	**8.87*** **	**9.06*** **	**8.63*** **	**9.23*** **	**9.35*** **	**9.59*** **
	(1.13)	(1.28)	(1.42)	(1.44)	(1.40)	(1.47)	(1.48)	(1.50)
Prison	**1.21*** **	**1.20*** **	**1.19*** **	**1.19*** **	**1.19*** **	**1.19*** **	**1.19*** **	**1.18*** **
	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
2SLS								
Monitoring *x* Prison	**2.39**	**2.13**	**1.86**	**1.74**	**1.76**	**1.83**	**1.91**	**1.92**
	(2.66)	(2.65)	(2.64)	(2.65)	(2.65)	(2.65)	(2.66)	(2.66)
Monitoring	**6.52*** **	**7.90*** **	**8.93*** **	**9.17*** **	**8.62*** **	**9.41*** **	**9.67*** **	**9.97*** **
	(1.26)	(1.46)	(1.68)	(1.71)	(1.65)	(1.74)	(1.77)	(1.81)
Prison	**1.44*** **	**1.44*** **	**1.45*** **	**1.44*** **	**1.44*** **	**1.44*** **	**1.44*** **	**1.44*** **
	(0.39)	(0.39)	(0.39)	(0.39)	(0.39)	(0.39)	(0.39)	(0.39)
Obs.	21,638	21,638	21,638	21,638	21,638	21,638	21,638	21,638

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively. The table presents coefficient estimates based on Equation (4). Fixed-effects are accounted for by demeaning time-varying variables. A third order polynomial of assets is included as an additional control. The window is set between -.85 to .2. In the first four columns, the polynomial functions are forced to have identical parameters to the left and the right of the threshold, but they are allowed to differ in the last four. Each of these eight models is estimated using OLS and 2SLS. For the 2SLS regressions, the instruments are $Above_{i,t}$ and $Punish_t*Above_{i,t}$. First-stage regressions are linear probability models. Heteroskedasticity robust standard errors are shown in parenthesis.

Data from the pooled sample is used. Thus there are about 21,600 observations (8,600 in the low-punishment sub-sample + 13,000 in the high-punishment sub-sample = 21,600 observations)

Table B5: Heterogeneous Interaction – Pooled Sample

	Dependant Variable: Tax $000s (demeaned)							
	Symmetric Polynomials				Asymmetric Polynomials			
	Linear	2nd	3rd	4th	Linear	2nd	3rd	4th
QR, Median								
Monitoring *x* Prison	**-0.39****	**-0.40*****	**-0.45*****	**-0.45*****	**-0.42*****	**-0.50*****	**-1.03*****	**-0.21****
	(0.15)	(0.10)	(0.10)	(0.10)	(0.14)	(0.10)	(0.11)	(0.11)
QR, 70th								
Monitoring *x* Prison	**0.38*****	**0.37*****	**0.39*****	**0.41*****	**0.40*****	**0.44*****	**0.42*****	**0.43*****
	(0.11)	(0.09)	(0.10)	(0.10)	(0.12)	(0.09)	(0.12)	(0.10)
QR, 90th								
Monitoring *x* Prison	**3.49*****	**3.63*****	**3.68*****	**4.41*****	**4.28*****	**4.35*****	**2.29****	**3.75*****
	(1.01)	(0.94)	(0.94)	(0.97)	(1.06)	(0.93)	(1.00)	(0.94)
Obs.	21,457	21,457	21,457	21,457	21,457	21,457	21,457	21,457

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively. The table presents coefficient estimates based on Equation (4). Fixed-effects are accounted for by demeaning time-varying variables. A third order polynomial of assets is included as an additional control. In the first four columns, the polynomial functions are forced to have identical parameters to the left and the right of the threshold, but they are allowed to differ in the last four. Each of these eight models is estimated using QR and IQR, for different quantiles of the tax distribution. For the IQR regressions, the instruments are $Above_{i,t}$ and $Punish_t*Above_{i,t}$. NA indicates that the IQR method does not converge.

Data from the pooled sample is used. The window is set between -.8 to .2. The sample size is the same as in previous tables.

C. Robustness Checks

Table C1: Smaller Windows

Panel A. Window -.4 to .2 – Low Punishment Sub-sample

	Dependant Variable: Tax $000 (demeaned)							
	Symmetric Polynomials				Asymmetric Polynomials			
	Linear	2nd	3rd	4th	Linear	2nd	3rd	4th
OLS								
Monitoring	**8.49*** **	**8.74*** **	**8.97*** **	**9.14*** **	**8.53*** **	**9.13*** **	**9.29*** **	**9.33*** **
	(1.19)	(1.24)	(1.25)	(1.26)	(1.20)	(1.26)	(1.29)	(1.30)
2SLS								
Monitoring	**8.99*** **	**9.33*** **	**9.62*** **	**9.87*** **	**9.05*** **	**9.86*** **	**10.13*** **	**10.22*** **
	(1.34)	(1.41)	(1.42)	(1.45)	(1.35)	(1.44)	(1.49)	(1.51)
Obs.	3,421	3,421	3,421	3,421	3,421	3,421	3,421	3,421

Panel B. Window -.2 to .2 – Low Punishment Sub-sample

	Dependant Variable: Tax $000 (demeaned)							
	Symmetric Polynomials				Asymmetric Polynomials			
	Linear	2nd	3rd	4th	Linear	2nd	3rd	4th
OLS								
Monitoring	**7.70*** **	**7.85*** **	**8.03*** **	**8.12*** **	**7.92*** **	**8.21*** **	**8.23*** **	**8.23*** **
	(1.32)	(1.33)	(1.32)	(1.35)	(1.35)	(1.39)	(1.40)	(1.40)
2SLS								
Monitoring	**8.36*** **	**8.56*** **	**8.88*** **	**9.01*** **	**8.65*** **	**9.21*** **	**9.31*** **	**9.32*** **
	(1.58)	(1.59)	(1.59)	(1.63)	(1.62)	(1.70)	(1.73)	(1.74)
Obs.	1,980	1,980	1,980	1,980	1,980	1,980	1,980	1,980

Panel C. Window -.2 to .2 – Pooled Sample

	Dependant Variable: Tax $000 (demeaned)							
	Symmetric Polynomials				Asymmetric Polynomials			
	Linear	2nd	3rd	4th	Linear	2nd	3rd	4th
QR, 80th Monitoring *x* Prison	**2.553*** **	**2.207*** **	**2.542*** **	**2.582*** **	**2.343*** **	**2.644*** **	**2.420*** **	**na**
	(0.51)	(0.51)	(0.50)	(0.43)	(0.45)	(0.51)	(0.53)	na
Obs.	10,785	10,785	10,785	10,785	10,785	10,785	10,785	na

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively. Panels A and B present estimates from Equation (1). Panel C presents estimates from Equation (4). NA indicated that the IQR estimation does not converge. The sample size is smaller than in earlier tables because the window size is smaller.

Table C2: Additional Controls

Panel A. Model 1

	Dependant Variable: Tax $000 (not demeaned)							
	Window -.85 to .2				Window -.2 to .2			
	Symmetric		Asymmetric		Symmetric		Asymmetric	
	2nd	3rd	2nd	3rd	2nd	3rd	2nd	3rd
2SLS								
Monitoring	**8.8*** **	**6.2***	**6.9***	**16.3*** **	**9.9** **	**4.5**	**13.5**	**27.6**
	(2.59)	(3.75)	(4.02)	(5.92)	(4.34)	(8.21)	(11.44)	(19.84)
Obs.	4,010	4,010	4,010	4,010	954	954	954	954
QR, 80th								
Monitoring *x*	**3.310*** **	**2.319** **	**3.658*** **	**4.703*** **	**6.21*** **	**5.943*** **	**6.18*** **	**5.909*** **
Prison	(1.13)	(1.07)	(1.18)	(1.05)	(1.27)	(1.48)	(1.44)	(1.36)
Obs.	14,759	14,759	14,759	14,759	6,440	6,440	6,440	6,440

Panel B. Model 2

	Dependant Variable: Tax $000 (demeaned)							
	Window -.85 to .2				Window -.2 to .2			
	Symmetric		Asymmetric		Symmetric		Asymmetric	
	2nd	3rd	2nd	3rd	2nd	3rd	2nd	3rd
2SLS								
Monitoring	**11.2*** **	**10.1*** **	**10.6*** **	**11.6*** **	**10.8*** **	**9.9*** **	**10.8*** **	**10.8*** **
	(2.34)	(2.34)	(2.35)	(2.44)	(2.91)	(2.97)	(3.12)	(3.14)
Obs.	8,364	8,364	8,364	8,364	1,980	1,980	1,980	1,980
QR, 80th								
Monitoring *x*	**1.025*** **	**1.015*** **	**1.009*** **	**1.252*** **	**0.984** **	**1.05*** **	**0.988*** **	**0.843** **
Prison	(0.38)	(0.37)	(0.37)	(0.33)	(0.42)	(0.41)	(0.38)	(0.35)
Obs.	24,907	24,907	24,907	24,907	14,128	14,128	14,128	14,128

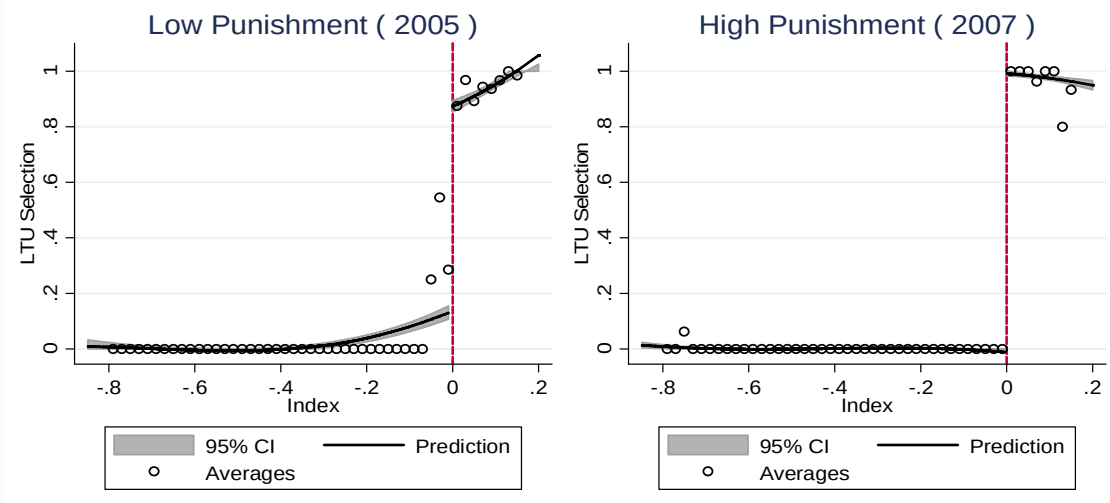
Panel C. Model 3

	Dependant Variable: Tax $000 (demeaned)							
	Window -.85 to .2				Window -.2 to .2			
	Symmetric		Asymmetric		Symmetric		Asymmetric	
	2nd	3rd	2nd	3rd	2nd	3rd	2nd	3rd
2SLS								
Monitoring	**9.8*** **	**8.8*** **	**9.7*** **	**11.0*** **	**9.5*** **	**9.4*** **	**10.0*** **	**10.0*** **
	(2.06)	(2.06)	(2.07)	(2.15)	(2.54)	(2.56)	(2.68)	(2.72)
Obs.	8,364	8,364	8,364	8,364	1,980	1,980	1,980	1,980
QR, 80th								
Monitoring *x*	**1.387*** **	**0.748** **	**0.729** **	**-5.318*** **	**0.734** **	**0.738** **	**0.95*** **	**0.856*** **
Prison	(0.35)	(0.35)	(0.35)	(0.34)	(0.35)	(0.35)	(0.36)	(0.30)
Obs.	24,907	24,907	24,907	24,907	14,128	14,128	14,128	14,128

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively. Model 1 includes a (two-period) lagged dependent variable as an additional covariate. Model 2 includes assets revenue, and costs. Model 3 includes the same covariates as Model 2 plus financial ratios (leverage and fixed-to-total-assets) and tax arrears.

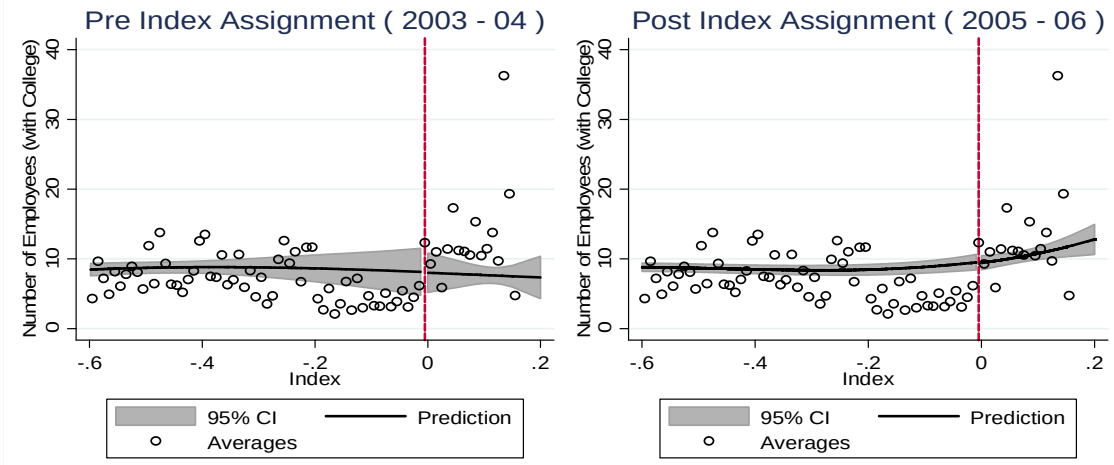
D. Figures

Figure D1: Probability of LTU selection



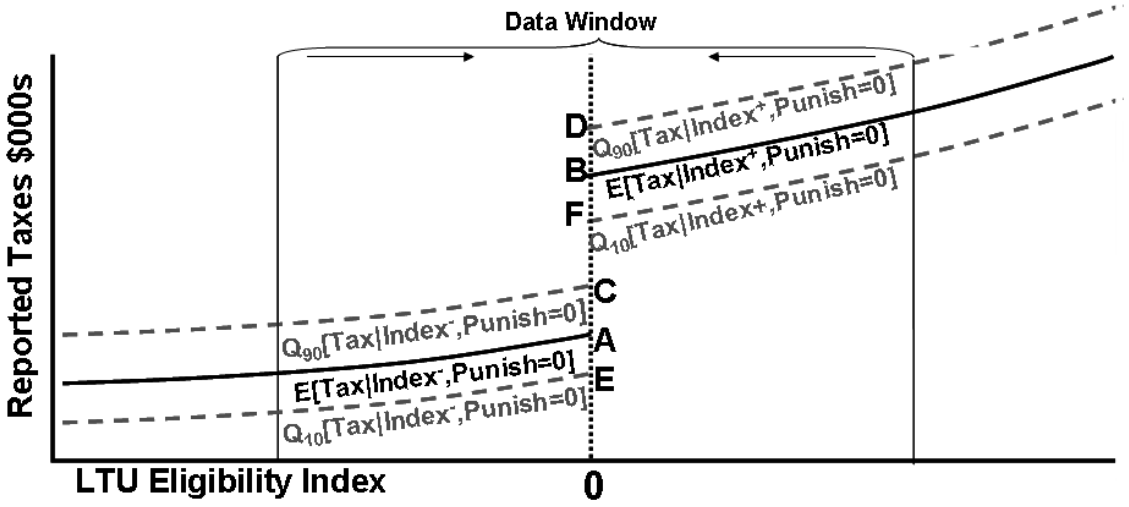
Notes: The figure focuses on firms considered for LTU selection in fiscal year 2005 (two years before a tax reform made punishment for tax crimes harsher); and 2007 (after the enactment of the reform). The dependant variable is a dummy equal to unity for firms receiving higher monitoring due to LTU selection, and zero else. The dots are simple averages of the dependant variable at given intervals of the index (non-parametric). The solid line is the prediction of a linear probability model, where the relationship between the dependant variable and the index is modeled as a symmetric 3rd order polynomial (parametric). Specifically, we esimtae simplified versions of Eq. (2), relying on a single time period of data (i.e. 2005 and 2007, respectively), as follows: $Monit_i = \phi\, Above_i + \alpha_0 + f(Index_i) + \varepsilon_i$.

Figure D2: Number of Employees with College Degree – Low Punishment Sample



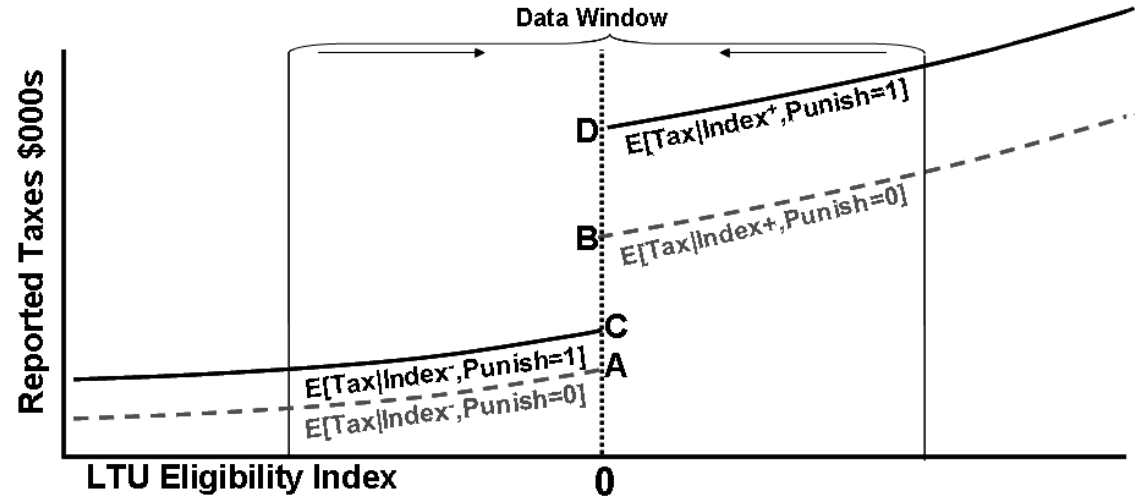
Notes: The figure focuses exclusively on firms considered for LTU selection in fiscal year 2005 (two years before a tax reform instituted prison term for tax crimes); however, these firms are observed over multiple years (i.e. pre- and post-index assignment). After the index was assigned, firms with an index value above the cutoff were selected to the LTU. The dependant variable (y_i) is the number of firm workers with a college degree. The dots are simple averages of the dependant variable at given intervals of the index (non-parametric). The solid line is the prediction of an OLS regression, where the relationship between the dependant variable and the index is modeled as a symmetric 3rd order polynomial (parametric). We estimate the following model: $y_i = \phi\, Above_i + \alpha_0 + \alpha_1 y_{t-1} + f(Index_i) + \varepsilon_i$. A lagged dependant variable is added to reduce heterogenity. Results show that the is no discontinuity at the cutoff in the number of workers with college degree (i.e., $Above_i$ is not statistically significant) Outliers are excluded from the analysis..

Figure D3: Identification Strategy – Monitoring



Notes: In the simplest case, the mean causal effect of higher monitoring on tax compliance can be retrieved by comparing the mean taxes for firms just below and just above the LTU eligibility cutoff ($\gamma^{M0} = B - A$). The figure shows tax heterogeneity as variation along the vertical axis. The monitoring effect at the 90th quantile can be retrieved by comparing the 90th quantile of taxes for firms just below and firms just above the cutoff ($\gamma_{90}^{M0} = D - C$). In practice a window of data close enough to the threshold needs to be selected.

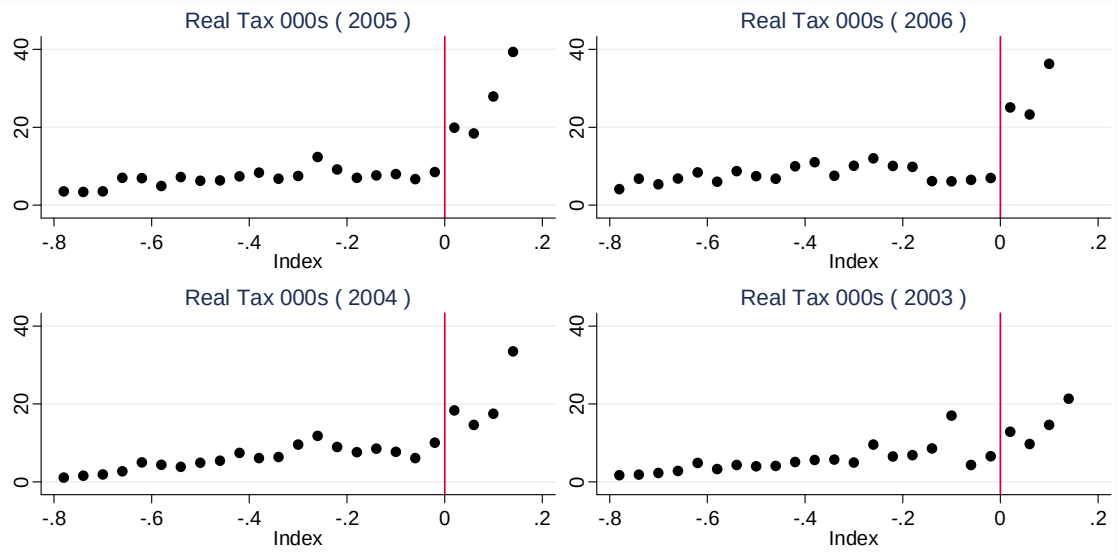
Figure D4: Identification Strategy – Interaction



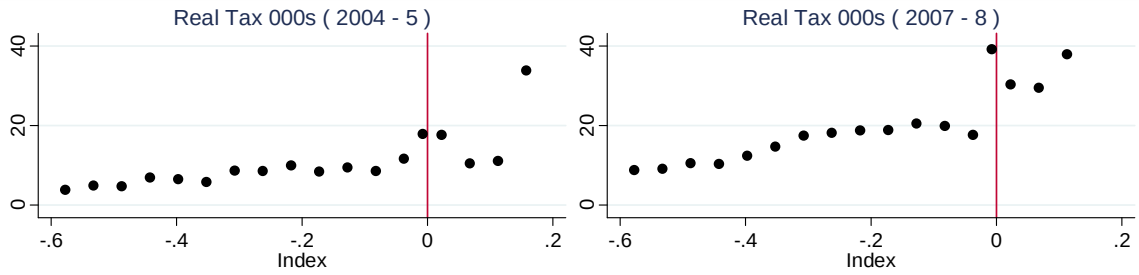
Notes: The mean interaction effect between monitoring and punishment can be retrieved as a difference-in-difference parameter, by comparing the discontinuity in taxes during the low and high punishment environments ($\gamma^{M*P} = [D - C] - [B - A]$).

Figure D5: Monitoring Effect – Preliminary Evidence

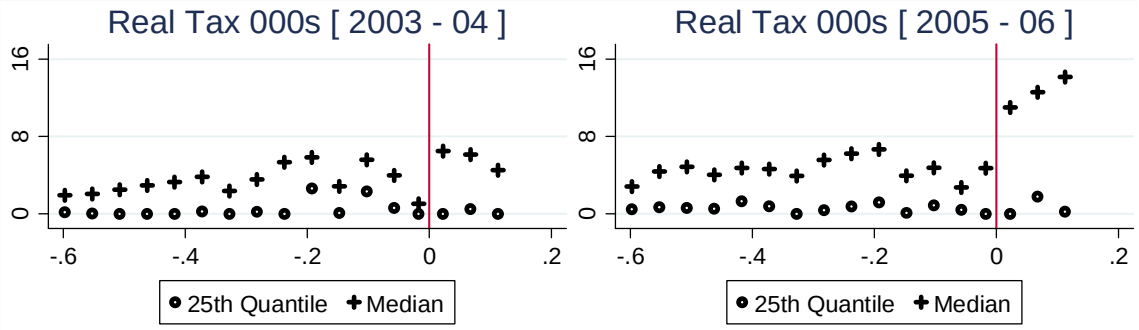
Panel A: Low Punishment Sub-sample (Mean)



Panel B: High Punishment Sub-sample (Mean)



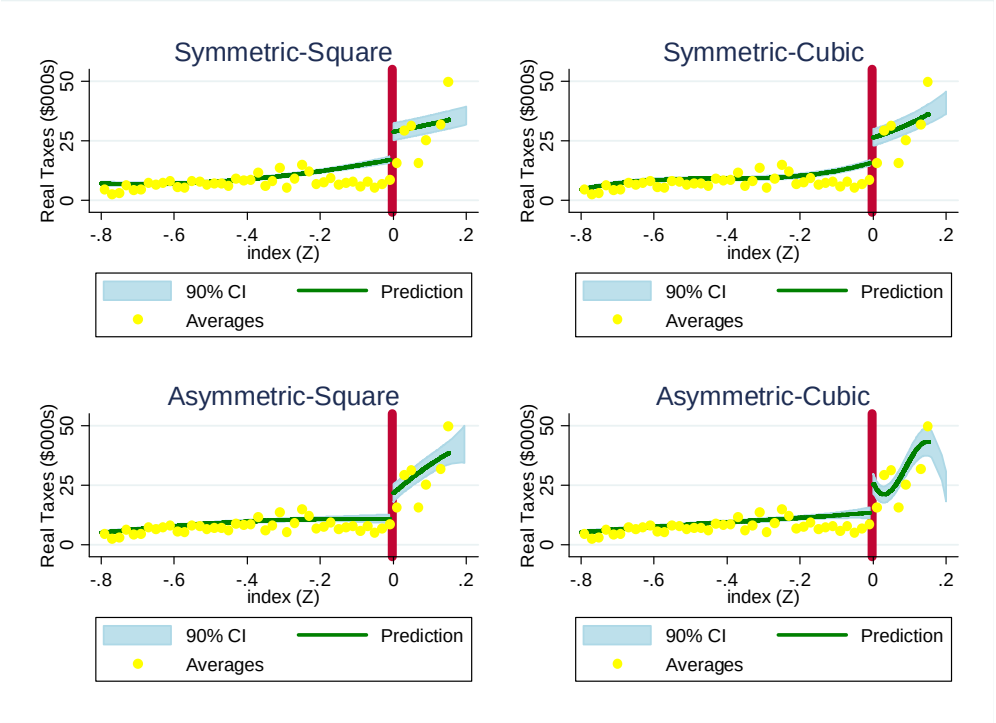
Panel C. Low Punishment Sub-sample (Quantiles)



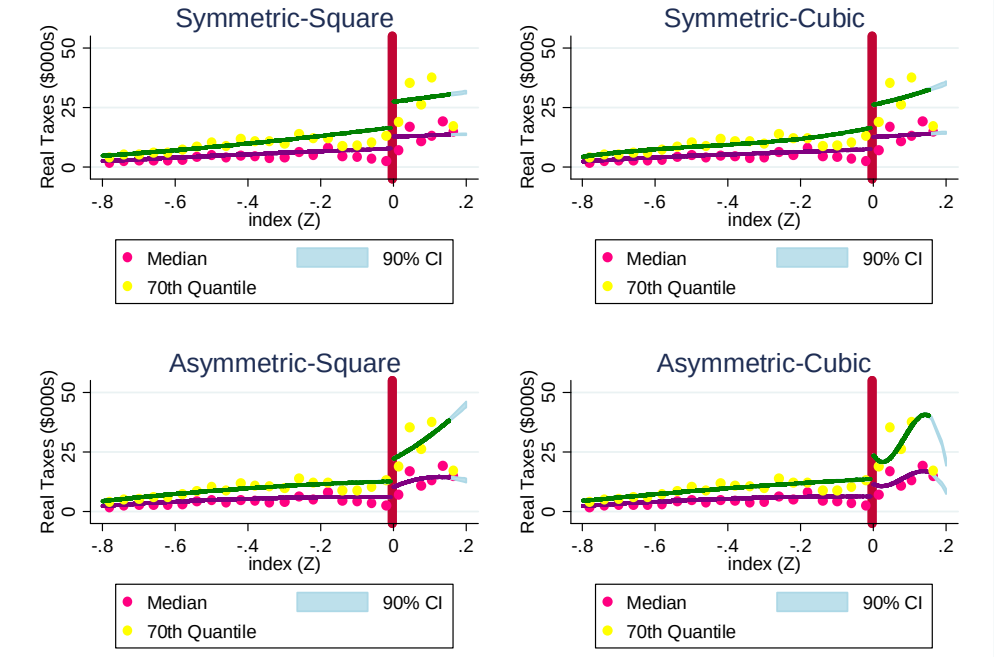
Notes: Only firms assigned according to the index eligibility rule are shown. Note that due to the panel nature of the data the same firma are shown each year. Filled circles represent mean values, other symbols represent different quantiles. For Panel C, note that there is little or no effect of higher monitoring at the 25th quantile, but the effect is much larger at the median, suggesting that the effect of higher monitoring may be heterogeneous.

Figure D6: Monitoring Results – Low Punishment Sub-sample

Panel A. Mean Results



Panel B. Quantile Results



Notes: The figure shows estimates of Equation (1) using QR. No additional controls $X_{i,t}$ are included. The lines shown in the figure are the prediction for the years 2005-06 only.

Figure D7: Complementarity

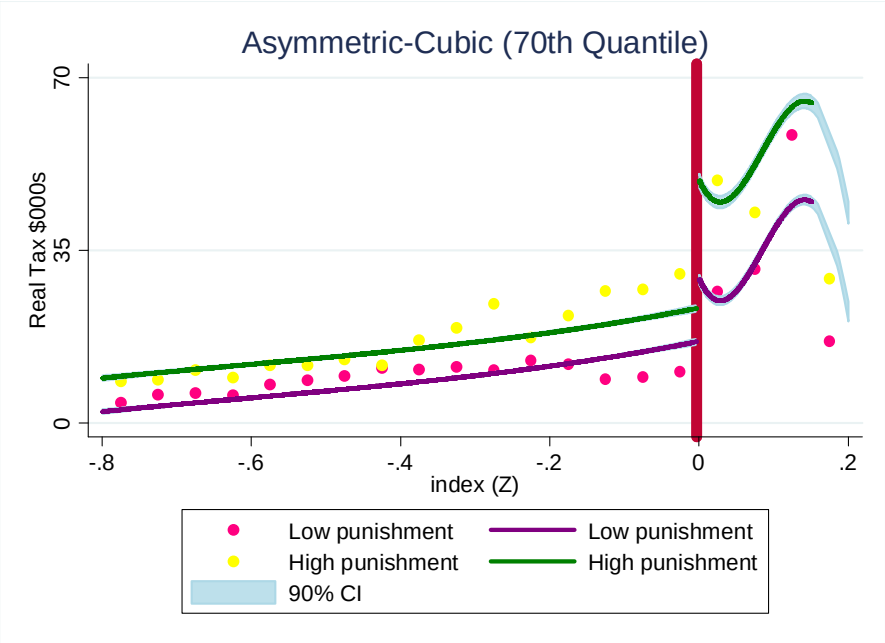
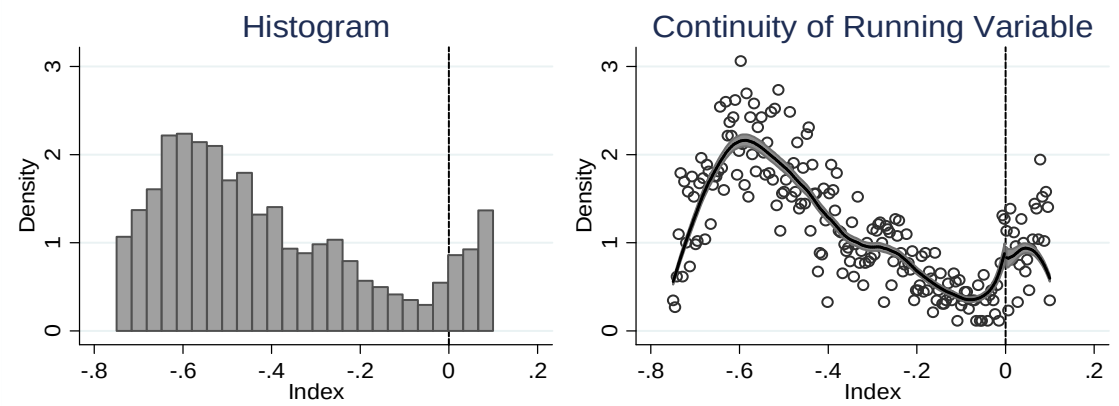


Figure D8: Continuity of the Density of the Index (McCrary's Test)

Panel A. Low Punishment Sub-sample



Panel B. High Punishment Sub-sample

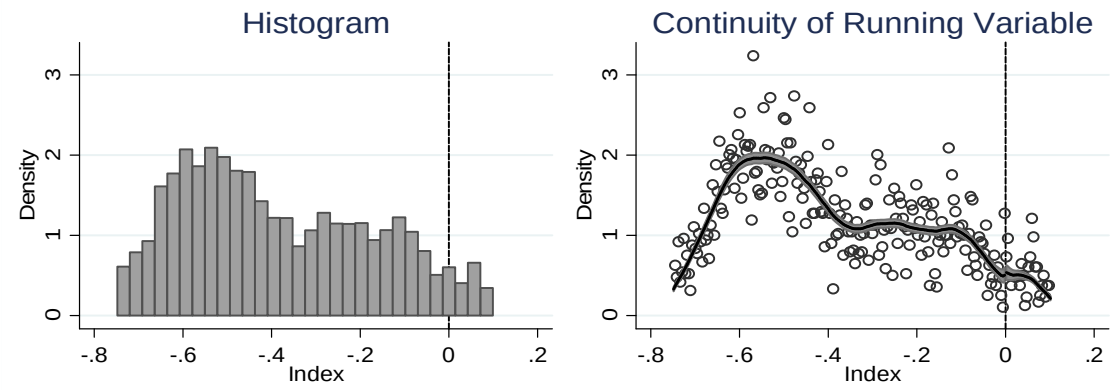
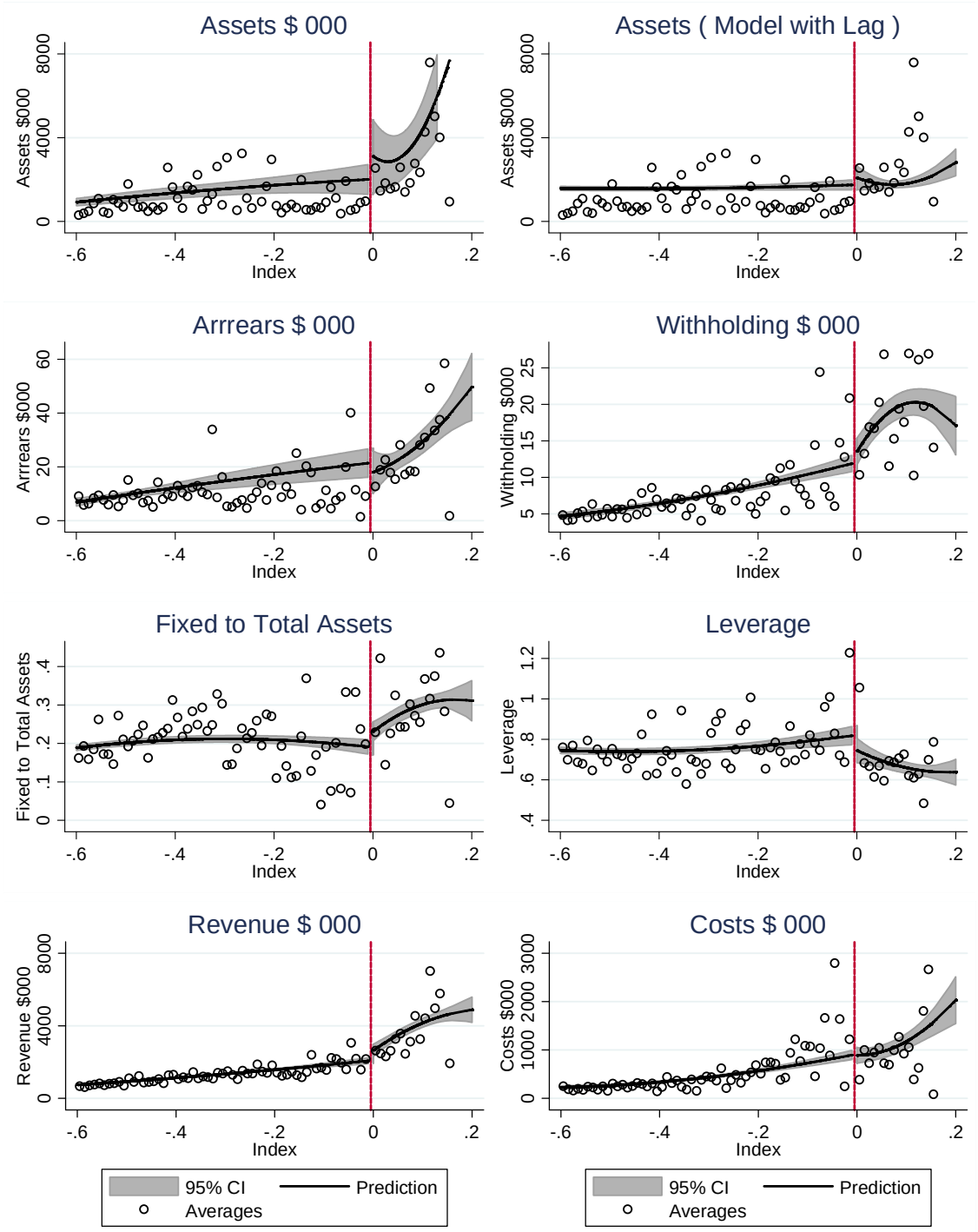
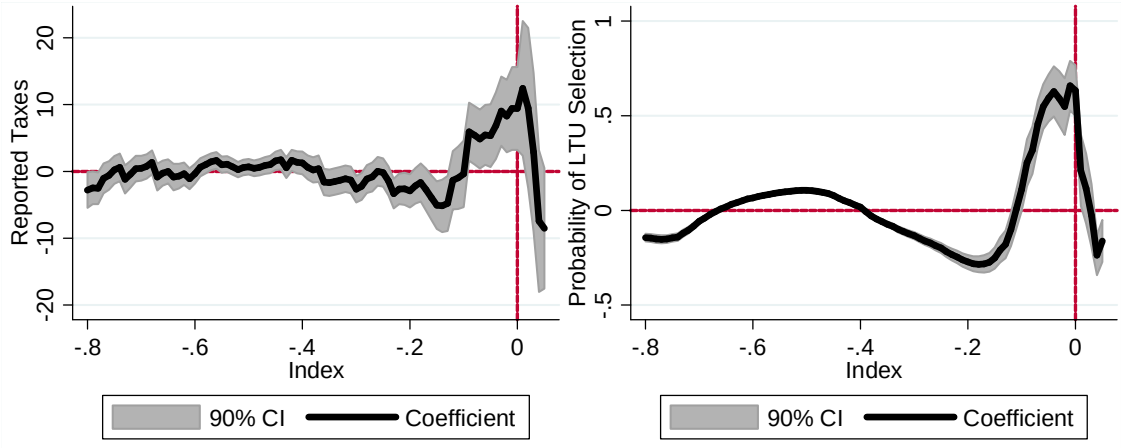


Figure D9: Continuity of Covariates – Low Punishment Sub-sample



Notes: All predictions are based on parametric regressions where the relationship between the dependant variable and the index is modeled as a 2nd order asymmetric polynomial. Outliers excluded.

Figure D10: Discontinuities Other than the Cutoff – Low Punishment Sub-sample



Notes: We test for discontinuities at placebo cutoff values, other than the actual cutoff used by the tax authority to determine LTU eligibility, i.e. $I(Index_{i,t}=0)$. Each point on the black lines (shown in the figure) plots the coefficient from a test of discontinuity at a given placebo cutoff, against the placebo cutoff value used for that test. For this purpose we run a large number of regressions, as shown in Equation (5): $Tax_{i,t} = \delta\, Test_{i,t} + f(Index_{i,t}) + \alpha_0 + \alpha_1 trend_t + \beta\, Tax_{i,t-2} + f(Assets_{i,t}) + u_{i,t}$. Where the variable $Test_{i,t}$ is a dummy equal to unity for firms with an index value above a given placebo cutoff. The placebo cutoff value changes for each regression. We cover the entire domain of possible values of the index in turn. Note that Equation (5) is very similar to Equation (1), except that the placebo dummy variable ($Test_{i,t}$) replaces the actual dummy for firms selected to the LTU ($Monit_{i,t}$). We also test for discontinuities in the probability of LTU selection at different placebo cutoffs.

References

Allingham, M. and A. Sandmo. 1972. "Income Tax Evasion: A Theoretical Analysis." *Journal of Public Economics*. 1: 323-338.

Alm, James, Jorge Martinez-Vazquez, and Mark Rider. 2006. *The Challenges of Tax Reform in a Global Economy*. New York, NY: Springer. http://surveyor.gelman.gwu.edu/?hreciid=%7clibrary%2fm%2fWRLC%7c106534645.

Angrist, Joshua D. and Victor Lavy. 1999. "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics* 114 (2): 533-575. doi:http://www.mitpressjournals.org/loi/qjec. http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=0495006&site=ehost-live; http://www.mitpressjournals.org/loi/qjec.

Aparicio, Gabriela, Paul E. Carrillo, and M. Shahe Emran. 2011. "Taxes, Prisons, and CFOs: The Effects of Increased Punishment on Corporate Tax Compliance in Ecuador." *IIEP-WP-2011-2*. doi:10.2139/ssrn.1808945. http://www.gwu.edu/~iiep/assets/docs/papers/Carrillo_IIEPWP2011-02.pdf.

Baer, Katherine, Olivier Benon, and Juan Toro. 2002. *Improving Large Taxpayers' Compliance : A Review of Country Experience*. Vol. 215. Washington, DC: International Monetary Fund. http://surveyor.gelman.gwu.edu/?hreciid=%7clibrary%2fm%2fWRLC%7c103996944; http://leagle.wcl.american.edu/search/i?=1589061187.

Banerjee, Abhijit V. and Esther Duflo. 2010. "Giving Credit Where it is due." *Journal of Economic Perspectives* 24 (3): 61-80. doi:http://www.aeaweb.org/jep/. http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=1119933&site=ehost-live; http://dx.doi.org/10.1257/jep.24.3.61; http://www.aeaweb.org/jep/.

Bardhan, Pranab, Samuel Bowles, and Herbert Gintis. "Wealth Inequality, Wealth Constraints and Economic Performance."Center for International and Development Economics Research, Institute for Business and Economic Research, UC Berkeley, Center for International and Development Economics Research, Working Paper Series: 1020, . http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=0905044&site=ehost-live; http://repositories.cdlib.org/cgi/viewcontent.cgi?article=1020&context=iber/cider.

Becker, G. S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76: 169-217. http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=1224966&site=ehost-live.

Bitler, Marianne P., Jonah B. Gelbach, and Hilary W. Hoynes. 2008. "Distributional Impacts of the Self-Sufficiency Project." *Journal of Public Economics* 92 (3-4): 748-765. doi:http://www.elsevier.com/locate/inca/505578/. http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=0968867&site=ehost-live; http://dx.doi.org/10.1016/j.jpubeco.2007.07.001; http://www.elsevier.com/locate/inca/505578/.

———. 2006. "What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments." *American Economic Review* 96 (4): 988-988. http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=0877335&site=ehost-live; http://dx.doi.org/10.1257/aer.96.4.988.

Black, Daniel, Jose Galdo, and Jeffrey Smith. 2007. "Evaluating the Bias of the Regression Discontinuity Design using Experimental Data." *Working Paper*.

Carrillo, P. E., M. S. Emran, and A. Rivadeneira. 2011. "Do Cheaters Bunch Together? Profit Taxes, Withholding Rates and Tax Evasion." *Working Paper*.

Chernozhukov, Victor and Christian Hansen. 2006. "Instrumental Quantile Regression Inference for Structural and Treatment Effect Models." *Journal of Econometrics* 132 (2): 491-525. doi:http://www.sciencedirect.com/science/journal/03044076. http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=0879311&site=ehost-live; http://dx.doi.org/10.1016/j.jeconom.2005.02.009; http://www.sciencedirect.com/science/journal/03044076.

———. 2005. "An IV Model of Quantile Treatment Effects." *Econometrica* 73 (1): 245-261. doi:http://www.econometricsociety.org. http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=0769817&site=ehost-live; http://www.econometricsociety.org.

Cook, Thomas D. and Vivian C. Wong. 2008. "Empirical Tests of the Validity of the Regression Discontinuity Design." *Annales d'Economie Et De Statistique* (91-92): 127-150. doi:http://annales.ensae.fr/. http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=1134782&site=ehost-live; http://annales.ensae.fr/.

Cowell, Frank A. 1990. *Cheating the Government: The Economics of Evasion*. Cambridge, Mass. and London: MIT Press. http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=0237873&site=ehost-live.

Deloitte. 2010. *International Tax and Business Guide: Ecuador*. Quito: Ecuador. http://www.deloitte.com/assets/Dcom-Ecuador/Local%20Assets/Documents/Doing%20Business%20in%20Ecuador%202008.pdf.

Economist Intelligence Unit (EIU). 2009. *Ecuador Risk Report*. Quito, Ecuador: The Economist.

Emran, M. S. and Joseph E. Stiglitz. 2005. "On Selective Indirect Tax Reform in Developing Countries." *Journal of Public Economics* 89 (4): 599-623. doi:http://www.elsevier.com/locate/inca/505578/. http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=0774125&site=ehost-live; http://dx.doi.org/10.1016/j.jpubeco.2004.04.007; http://www.elsevier.com/locate/inca/505578/.

Frolich, M. and B. Melly. 2008. "Quantile Treatment Effects in the Regression Discontinuity Design." *IZA Discussion Paper no.3638*.

Gallagher, M. 2005. "Benchmarking Tax Systems." *Public Administration and Development* 25 (2): 125-144.

Guiteras, R. 2008. "Estimating Quantile Treatment Effects in a Regression Disconitnuity Design." *Unpublished Manuscript*.

Gupta, Sanjay and Kaye Newberry. 1997. "Determinants of the Variability in Corporate Effective Tax Rates: Evidence from Longitudinal Data." *Journal of Accounting & Public Policy* 16 (1): 1-34. http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=9712031996&site=ehost-live.

Hahn, Jinyong, Petra Todd, and der Klaauw Van. 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica* 69 (1): 201-209.

doi:http://www.econometricsociety.org. http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=0556573&site=ehost-live; http://www.econometricsociety.org.

Hanlon, Michelle, Lillian Mills, and Joel Slemrod. 2007. "An Empirical Examination of Corporate Tax Noncompliance." In *Taxing Corporate Income in the 21st Century*, edited by Alan J. Auerbach, James R. Hines Jr. and Joel Slemrod, 171-210: Cambridge and New York: Cambridge University Press.

Imbens, Guido W. and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142 (2): 615-635. doi:http://www.sciencedirect.com/science/journal/03044076. http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=0963180&site=ehost-live; http://dx.doi.org/10.1016/j.jeconom.2007.05.001; http://www.sciencedirect.com/science/journal/03044076.

Koenker, Roger W. and Gilbert Bassett Jr. 1978. "Regression Quantiles." *Econometrica* 46 (1): 33-50. doi:http://www.econometricsociety.org. http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=0103223&site=ehost-live; http://www.econometricsociety.org.

Lee, David S. and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48 (2): 281-355. doi:http://www.aeaweb.org/jel/index.php. http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=1108108&site=ehost-live; http://dx.doi.org/10.1257/jel.48.2.281; http://www.aeaweb.org/jel/index.php.

McCarten, W. 2006. "Focusing on the Few: The Role of Large Taxpayer Units in the Revenue Strategies of Developing Countries." In *The Challenges of Tax Reform in a Global Economy*, edited by James Alm, Jorge Martinez-Vazquez and Mark Rider, 24-25. New York, NY: Springer.

Milgrom, Paul R., Yingyi Qian, and John Roberts. 1991. "Complementarities, Momentum, and the Evolution of Modern Manufacturing." *American Economic Review* 81 (2): 84-88. doi:http://www.aeaweb.org/aer/. http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=0244308&site=ehost-live; http://www.aeaweb.org/aer/.

Mookherjee, Dilip and I. P. L. Png. 1994. "Marginal Deterrence in Enforcement of Law." *Journal of Political Economy* 102 (5): 1039-1066. doi:http://www.journals.uchicago.edu/JPE/. http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=0352093&site=ehost-live; http://www.journals.uchicago.edu/JPE/.

Munk, Knud J. "On the use of Border Taxes in Developing Countries."Universite catholique de Louvain, Departement des Sciences Economiques, Universite catholique de Louvain, Departement des Sciences Economiques Working Paper: 2008005, . http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=0959116&site=ehost-live; http://www.ires.ucl.ac.be/DP/IRES_DP/2008-5.pdf.

Polinsky, A. M. and Steven Shavell. 2000. "The Economic Theory of Public Enforcement of Law." *Journal of Economic Literature* 38 (1): 45-76. doi:http://www.aeaweb.org/jel/index.php. http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=0526947&site=ehost-live; http://www.aeaweb.org/jel/index.php.

Pomeranz, Dina. 2011. "No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax." *Unpublished Manuscript*.

Ravallion, Martin. 2008. "Evaluating Anti-Poverty Programs, Ch. 59." In *Handbook of Development Economics*, edited by T. Paul Schultz and John A. Strauss. Vol. 4, 3787-3846. http://EconPapers.repec.org/RePEc:eee:devchp:5-59.

Rice, Eric M. 1992. "The Corporate Tax Gap: Evidence on Tax Compliance by Small Corporations." In *Why People Pay Taxes: Tax Compliance and Enforcement*, edited by Joel Slemrod, 125-661: Ann Arbor:; University of Michigan Press.

Servicio de Rentas Internas (SRI). 2008. *Metodología De Selección Masiva De Contribuyentes Especiales*. Quito, Ecuador: Internal Document.

Slemrod, Joel, Marsha Blumenthal, and Charles Christian. 2001. "Taxpayer Response to an Increased Probability of Audit: Evidence from a Controlled Experiment in Minnesota." *Journal of Public Economics* 79 (3): 455-483. doi:http://www.elsevier.com/locate/inca/505578/. http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=0559868&site=ehost-live; http://www.elsevier.com/locate/inca/505578/.

Stiglitz, J. 2010. "Development Oriented Tax Policy." In *Taxation in Developing Countries: Six Case Studies and Policy Implications*, edited by R. Gordon.

Topkis, Donald M. 1998. *Supermodularity and Complementarity* Frontiers of Economic Research series; Princeton:; Princeton University Press. http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=0475186&site=ehost-live.

World Bank. 1996. *Ecuador Poverty Report*. A World Bank Country Study, 0253-2123. Washington, D.C.: World Bank. http://surveyor.gelman.gwu.edu/?hreciid=%7clibrary%2fm%2fWRLC%7c103124548; http://sfx.wrlc.org/gw/OpenURL?sid=sfx:e_collection&isbn=9780821336656&pid=serviceType=getFullTxt.

-