CONSUMERS AS TAX AUDITORS

Joana Naritomi, London School of Economics*

April, 2015

Abstract

Access to third-party information trails is widely believed to be critical to the development of modern tax systems, but there is limited direct evidence of the effects of changes in information trails. This paper investigates the enforcement effect of an increased availability of third-party information, and sheds light on how governments can harness this information despite collusion opportunities. I exploit unique administrative data on firms and consumers from an anti-tax evasion program in Sao Paulo, Brazil (*Nota Fiscal Paulista*) that created monetary rewards for consumers to ensure that firms report final sales transactions, and establishes an online verification system that aids consumers in whistle-blowing firms. Using variation in intensity of exposure to the policy, I estimate that firms' reported revenue increased by at least 22% over four years. The compliance effect is stronger for firms that face a high volume of consumers, consistent with positive shifts in audit probability from whistle-blower threats. I also investigate the effect of whistle-blowers directly: firms report 14% more receipts and 6% more revenue after receiving the first consumer complaint. To study the role of the value of rewards in improving enforcement, I show evidence consistent with the possibility that lottery incentives amplify consumer responses due to behavioral biases, which would make it more costly for firms to try to match government incentives in a collusive deal. I conclude by discussing policy implications.

*J.Naritomi@lse.ac.uk. I am very grateful to my advisors Alberto Alesina, Raj Chetty, Michael Kremer and Andrei Shleifer. I thank Tim Besley, Ciro Biderman, Raluca Dragusanu, Itzik Fadlon, François Gérard, Nathan Hendren, Nate Hilger, Martin Kanz, Asim Khwaja, Henrik Kleven, Anders Jensen, Olivia S. Mitchell, Sendhil Mullainathan, Arash Nekoei, Dina Pomeranz, Francisco Queiro, Tristan Reed, Edson Severnini, Heather Schofield, Joel Slemrod, Ian Tomb, Matthew Weinzierl and Danny Yagan for invaluable comments and advice. I also thank seminar and conference participants. I am extremely grateful to Adrea Calabi, André Luis Grotti Clemente, Igor Baremboin, and Euripedes De Oliveira for the opportunity to work with the Department of Finance of Sao Paulo (SEFAZ/SP). I am very thankful to Paulo Yamada and to the SEFAZ/SP staff for outstanding collaboration. I thank Sabrina Naritomi and Tiffany Blackman for legal advice, and SOX Consult for assistance with the Brazilian tax law. Research support from the LEAP and Lemann Foundation is gratefully acknowledged. This work does not necessarily reflect the views of SEFAZ/SP.

I Introduction

Tax revenue as a share of GDP is substantially higher in modern advanced economies than in the early 20th century or in present day developing countries (Besley and Persson, 2014; Peacock and Wiseman, 1961). A key source of the variation in tax revenue is the enforcement capacity of governments.[1] In particular, a growing literature emphasizes that information on taxable transactions shared with third-parties can be leveraged by governments to ensure more accurate self-reporting,[2] and that the increased availability of third-party information trails as countries develop could help explain the dynamics of government revenue among advanced economies during the last century (Gordon and Li, 2009; Kleven, Kreiner and Saez, 2015).

Despite the empirical literature on the deterrence effect of third-party reporting, there is little direct evidence on whether changes in availability of information trails can improve compliance, and on the mechanisms through which third-party reporting deters evasion as it hinges on avoiding collusion opportunities among the informed parties.[3] This paper exploits quasi-experimental variation and unique administrative data on firms and consumers from an anti-tax evasion program in Sao Paulo, Brazil – *Nota Fiscal Paulista* (NFP) - that created monetary rewards for consumers to ensure that firms report final sales transactions. The program provides tax rebates and monthly lottery prizes for consumers who ask for receipts, and establishes a direct communication channel between the tax authority and consumers through an online account system, where consumers can verify receipts reported by establishments, and can act as whistle-blowers by filing complaints.

The program was designed to address the 'last mile' problem of the self-enforcing mechanism of the Value Added Tax (VAT). Along the supply chain, the tax credit and debit system of the VAT generates third-party reporting in transactions across firms.[4] At the final consumer stage, however, these self-enforcing incentives break down since consumers typically derive no direct

[1]Musgrave (1969) emphasized the historical relevance of tax administration for tax collection. In the policy debate, tax administration and the enforcement capacity of developing country governments are central issues (Slemrod and Gillitzer, 2014; Bird and Gendron, 2007; IMF, 2011).

[2]Audit experiments typically detect near zero evasion in income subject to third-party reporting (Kleven, 2014; Pomeranz, 2014; Slemrod et al 2014). For instance, wage earners would face a much higher risk of audit relative to the self-employed if they under-report income, as firms typically also report wages paid to the government (Slemrod, 2007). More generally, information trails shared with third-parties such as employees, suppliers, banks or customers could have a deterrence effect even if they are not systematically reported to the government. Evidence from Denmark and the U.S. suggests that even when income is not subject to systematic third-party reporting, compliance is well below full evasion, which could be explained in part by the existence of derivative information shared with third-parties (Kleven, 2014).

[3]This is a well-known issue in the mechanism design literature (e.g., Tirole, 1986): once more than one person is informed about evasion then there are many mechanisms that can be used to elicit that information (Besley and Persson, 2013). A key assumption in these cases is that there is no scope for collusion among the informed parties. For instance, Yaniv (1993) argues that employers and employees can find mutually beneficial opportunities to reduce their tax liabilities, which would result in limited enforcement effect on self-reports of individual income subject to cross-reporting by firms.

[4]Most countries in the world adopted the VAT instead of sales tax, perhaps because of its enforcement advantage (Keen and Lockwood, 2010). Kopczuk and Slemrod (2006) argue that retail sales tax and the VAT are theoretically equivalent, but the VAT has built-in enforcement incentives along the supply chain. Pomeranz (2014) provides empirical evidence for the self-enforcing properties of the VAT.

monetary benefit from asking for receipts.[5] The NFP policy introduced incentives similar to the VAT for final sales: it aims to affect both the likelihood that a transaction is reported at all, and the accuracy of reporting, since rewards to consumers are an increasing function of the value of receipts.[6]

I begin the analysis describing a conceptual framework to discuss how incentives to consumers can affect firm behavior. The NFP policy is effectively increasing the availability of third-party information trails through rewards to consumers, but collusion between consumers and firms could hinder the self-enforcing effect of third-party information. However, in order to collude with consumers and continue evading, firms would need to transfer part of evasion rents to consumers through discounts. Moreover, firms would reveal evasion information to many third parties by conditioning the discount on not accurately reporting the transaction to the government. As in Kleven, Kreiner and Saez (2015), the difficulty in sustaining collusion with a large number of informed economic agents who can act as whistle-blowers might be important to deter evasion. Therefore, the effect of consumer monitoring should be stronger the higher the threat of whistle-blowing, and the more firms need to transfer to consumers to match the value of the rewards offered by the government.

In order to empirically investigate the extent to which rewards to consumers can affect firm compliance, I construct unique administrative data on establishment-level monthly tax returns from over a million establishments, monthly individual-level data on receipt collection and overall participation in the NFP program from over 40 million consumers based on administrative records from the tax authority of the state of Sao Paulo.[7] I divide my analysis into three parts. First, I study the effect of consumer monitoring on establishments' compliance by exploiting variation in the intensity of exposure to the policy. I compare reported revenue changes in establishments that sell mostly to final consumers (retail) versus establishments that sell mostly to other firms (wholesale). I estimate that reported revenue in retail increased on average by 22% over four years as a result of NFP. This estimate is likely to be a lower bound for the effect of the program, given that wholesale establishments may also have been affected by the change in consumers' decisions to ask for receipts. [8]

In the second part, I shed light on mechanisms by examining the implications from the conceptual framework for firms subject to higher whistle-blower threats, and by discussing the role of rewards offered by the government on consumer participation in the enforcement policy. I find evidence consistent with the argument that collusion might be difficult to sustain if consumers can

[5]Slemrod (2007) refers to the enforcement problem at the final consumer stage as the *Achilles heel* of administering a retail sales tax: if firms collude to underreport transactions, the self-enforcing mechanism can unravel, and may hinder tax collection across the entire chain.

[6]Both the tax rebate and the number of lottery tickets with which consumers are rewarded are a function of the total amount they spend in a given month. For more details, see section III.B.

[7]A number of measures were taken to de-identify the data in order to protect confidential tax records. See Section II.C.

[8]Wholesale firms can sell to final consumers directly, in which case the rewards program applies. Additionally, improving compliance among retail firms can affect compliance by wholesalers through the self-enforcing mechanism of the VAT.

blow the whistle. Establishments in sectors that are characterized by a large number of transactions that would be more exposed to potential whistle-blowers are relatively more affected by the consumer rewards program. Furthermore, I link consumer participation to firm compliance by exploiting the timing of consumers' whistle-blowing and find that firms report 14% more receipts and 6% more revenue after receiving the first complaint.

Next, I turn to the effects of rewards on consumer participation. As suggested by the conceptual framework, the more consumers value the rewards, the more costly it will be for firms to try to match the government's incentives. I exploit variation from lottery prize rewards from NFP to analyze changes in the number of receipts for which individuals ask and the number of different businesses in which they ask for receipts. I find that consumers condition their decisions to ask for receipts on past lottery wins. Even when prizes are as small as 5 dollars, winners ask for receipts more often and in a larger set of establishments for at least three months after the lottery result relative to non-winners with the same odds of getting a prize. Moreover, consumers sharply increase their overall participation in the program around the time they become eligible for lotteries even though the expected value of lottery rewards is substantially lower than the tax rebates for which they were already eligible.[9] The results are consistent with the possibility that lotteries amplify consumer engagement due to behavioral biases.

In the final part of the paper I discuss policy implications. The evidence shows that rewards to consumers can increase compliance by establishments. However, an increase in tax enforcement may affect exit and employment decision of firms, and the government is foregoing a fraction of both marginal and infra-marginal revenue through tax rebates and lotteries. I study the impact of the consumer monitoring policy on exit rates by comparing establishments in retail and wholesale sectors in Sao Paulo. For the employment analysis, I use a matched employer-employee dataset on the universe of formal firms in Brazil, and I compare employment in retail establishments in Sao Paulo to employment in retail establishments in other Brazilian states. I find no effects of the program on these margins during the period of analysis. In a cost-benefit analysis I discuss other potential costs and benefits from the consumer rewards policy.

This paper contributes to a growing literature that argues that third-party information is key for compliance (Gordon and Li, 2009; Kleven et al., 2011; Kumler et al., 2013; Pomeranz, 2014; Slemrod et. al., 2014). In particular, it provides evidence from changes in the availability of third-party information trails and the results are consistent with the whistle-blower mechanism suggested by the theoretical model in Kleven, Kreiner and Saez (2015).

Additionally, the paper contributes to the literature on the challenges of taxation in developing countries. In particular, a growing strand of the literature (e.g., Kumler et al., 2013; Pomeranz, 2014) sets aside non-compliance due to firm non registration at the tax authority – the formal-informal margin – and instead examines non-compliance among formal firms.[10] More generally,

[9] Any SSN holder in Brazil is eligible for tax rebates, but one needs to create an online account at the tax authority's website to opt in for lottery prizes. See section III.A.

[10] De Paula and Scheinkman (2010) argue that better enforcement in the intensive margin of VAT systems can endogenously generate incentives to formalize by creating supply chains of formal firms. See Bruhm and McKenzie (2013) for

the paper is related to a vast literature on tax evasion and enforcement (e.g., Andreoni et al., 1998; Slemrod and Yitzhaki, 2002).

Finally, the paper contributes to the policy debate on sales tax enforcement. A number of countries reward consumers to ask for receipts to address the last-mile problem of the VAT.[11] This paper provides, to my knowledge, the first direct evidence of consumer behavioral responses to rewards from asking for receipts. The results also reinforce existing findings on individual responses to lotteries that are used as levers in other contexts, such as lottery-linked savings (Tufano, 2008; Kearney et al., 2011). Moreover, the evidence from the NFP lotteries adds to the literature on the behavioral effects of lottery wins such as the lucky store effect (Guryan and Kearney, 2008). More generally, the paper sheds light on how participatory policies can be used as a monitoring tool.

The remainder of the paper is organized as follows. Section II outlines a simple conceptual framework to guide the empirical analysis. Section III describes the institutional background of the *Nota Fiscal Paulista* program, the relevant datasets, sample definitions and summary statistics. Section IV investigates the enforcement effect of the introduction of third-party information through consumer rewards on firms's reported revenue, and sheds light on mechanisms suggested by the conceptual framework regarding whistle-blower threats and the value of monetary rewards. Section V examines implications on exit and the employment decision of firms, and discusses a cost-benefit analysis. Section VI concludes.

II Conceptual framework

I begin by describing a simple conceptual framework that examines the degree to which consumer monitoring can affect the evasion decision by firms. I follow a Becker (1968) crime model developed by Allingham and Sandmo (1972) to analyze tax evasion. In particular, the framework uses a variant of this model discussed by Kleven et al. (2011), in which the probability of audit is increasing in the amount evaded.[12] First, I present a baseline case with government monitoring only that combines government audits and penalties for enforcement. Then, I introduce consumer monitoring as an additional enforcement tool that gives monetary incentives for consumers to ensure firms report final sales transactions, and allows consumers to act as whistle-blowers. [13]

a review of the literature on the formalization of firms.

[11] For instance, Argentina, Bolivia, Brazil, China, Chile, Colombia, Indonesia, Italy, Portugal, Puerto Rico, South Korea and Slovakia, among other countries, have introduced policies to address the enforcement problem downstream through monetary incentives – through tax refunds, lotteries, or fines – for consumers to collect receipts (Bird, 1992; Cowell, 2004; Fabbri, 2013; Marchese, 2009). Wan (2010) argues that a program that turns receipts into lottery tickets in China was effective in raising tax revenue, but the evidence for policies such as the NFP on tax collection in Brazilian states is mixed (Barroso and Cortez, 2009; Mattos et. al, 2013).

[12] This assumption is in line with anecdotal evidence that tax authorities may audit more often the higher the potential evasion it uncovers.

[13] I take a positive approach to understand the effects of different monitoring tools on firms evasion decision. For a normative approach, see Arbex and Mattos (2015) that investigate how the Ramsey equation is modified once consumers are rewarded to ask for receipts. They find that welfare is higher when consumer auditing is added to the usual direct government auditing.

II.A A tax evasion model

Consider a risk-neutral firm that pays a tax $\tau \in [0,1]$ proportional to their reported revenue $Y \geq 0$. Suppose firms sell a single product, and that each firm has N consumers who each make one purchase that generates revenue $\bar{y} \geq 0$. Firms have a true pre-tax revenue $\bar{Y} = N\bar{y}$, and choose to report revenue Y to maximize profits π. As is standard in the tax evasion literature, I assume that when the government audits a firm, it can perfectly observe the firm's true revenue $\bar{Y}$, and it applies a fine $\theta \geq 0$ in proportion to the evaded tax $\tau(\bar{Y} - Y)$.

II.A.1 Government monitoring only

Let $p \in [0,1]$ be the probability a firm is audited by the government. I assume that this probability is increasing in the amount evaded $p \equiv p(E)$, $p'(E) > 0$, $E = \bar{Y} - Y$. Thus, it is endogenous to the reported revenue decision. For simplicity, assume that in the absence of monetary incentives, consumers do not ask for receipts and have no impact on the evasion decision of firms. Thus, firms report revenue Y to maximize:

$$\pi = (\bar{Y} - \tau Y)(1 - p(\bar{Y} - Y)) + [\bar{Y}(1-\tau) - (\bar{Y} - Y)\theta\tau]p(\bar{Y} - Y) \qquad (1)$$

An interior optimal solution for Y satisfies the first order condition $d\pi/dY = 0$:[14]

$$[p(E) + p'(E).E](1+\theta) = 1 \qquad (2)$$

where $E = \bar{Y} - Y$. The right hand side of equation (2) is the marginal benefit of evading an extra dollar, and the left-hand side is the marginal cost of evading that extra dollar.[15] Firms choose the optimal Y^* that satisfies equation (2).

II.A.2 Adding consumer monitoring

Consider now the case where consumers can be used to monitor firms in addition to the government monitoring. Consumers are rewarded with $\alpha \in [0,1]$ of the tax τ firms pay on the transaction reported to the government. Consumers can ensure they receive this reward with the paper trail provided by receipts,[16], and they can act as whistle-blowers by informing the government about firms' non-compliance. The aim of adding these features is to discuss the role of monetary rewards in increasing information trails about firms' evasion, and the relevance of whistle-blower threats as a device to harness this information.

[14] The second-order condition is $-2p'(E) - p''(E)E < 0$. It is sufficient that $p(E)$ is convex.

[15] As discussed in Kleven et al (2011), the firm that evades an extra dollar incurs in a higher probability of detection of all infra-marginal dollars evaded.

[16] I assume that if a firm issues the receipt, it will report the transaction to the government. This is reasonable empirically as it is explained in detail in section III.A: the paper trail alone can deter evasion, and consumers can cross-check online each purchase the firm is supposed to report to ensure they get the monetary rewards associated with the receipt.

I describe a case in which firms may try to collude with consumers to avoid issuing receipts with the true value of the transaction. As the government is rewarding consumers with a fraction of what firms pay in taxes, firms and consumers could potentially agree to a mutually-beneficial deal and not issue receipts.[17] In particular, firms can propose a discount such that consumers accept that firms will report y instead of the true amount $\bar{y}$. In order to convince consumers to trade the government benefit for the discount, the discount should be at least $\alpha\tau(\bar{y}-y)$ to match the government's benefit.

It is important to note that not only must the firm share part of their evasion rents with the consumer, the firm reveals to a third party that it evades taxes by conditioning the discount on not reporting the true amount of the transaction.[18] Consumers, therefore, become informed third-parties. As consumers can act as whistle-blowers, the firm might face an increased audit probability if consumers cannot commit not to whistle-blow.[19]

Kleven, Kreiner and Saez (2015) argue that a key deterrent of collusion is the sheer number of internal or external parties to which a firm that evades taxes exposes itself. I will consider the case where there is a probability of a random shock between the parties that can trigger a consumer to blow the whistle. A random shock could be generated by some conflict between the consumer and the shopkeeper, or a moral concern of the consumer. Therefore, the larger the number of consumers N, the higher the additional risk of audit introduced by consumes acting as whistle-blowers.

Assume $\varepsilon > 0$ is the probability that such a random shock occurs; let ε be i.i.d. across consumers. Assume an audit will be triggered with certainty if one consumer blows the whistle on the firm, and that all the N consumers may blow the whistle. The probability that a firm is audited due to whistle-blowing is $1-(1-\varepsilon)^N$. Therefore, the relevant probability of audit under consumer monitoring p_c will be given by $p_c = 1-(1-p(E))(1-\varepsilon)^N$.

Now, therefore, firms choose Y to maximize:[20]

$$\pi = (\bar{Y}-\tau Y)(1-p_c) + [\bar{Y}(1-\tau) - (\bar{Y}-Y)\theta\tau]p_c - (\bar{Y}-Y)\alpha\tau \tag{3}$$

As mentioned above, under the new policy, firms have to transfer part of the evasion rents to consumers through discounts. An interior optimal solution Y^{**} satisfies the first order condition $d\pi/dY = 0$:

$$(1+\theta)[p_c - p_c'(\bar{Y}-Y)] = 1-\alpha \tag{4}$$

[17] Tirole (1986) discusses the collusion problem in auditing contracts in which a group of informed parties - the auditor and the agent - can manipulate the information reported to the principal. This context is also similar to the case of corruption with theft in Shleifer and Vishny (1993), and administrative corruption in Flatters and MacLeod (1995).

[18] In the empirical context, it is particularly salient that a collusive deal will allow the firm not to report the transaction since the government is giving a reward for consumers to ask for receipts in a campaign against tax evasion.

[19] As will be described in detail in section III.A, in the empirical setting consumers can file complaints about specific firms to the government through a website.

[20] I assume that if the firm is audited the government will consider as tax evasion the amount not reported based on the posted price $\bar{y}$, not the discounted price. Therefore, $\bar{Y}$ will be the true revenue of the firm, instead of the revenue net of transfers to consumers.

The marginal benefit of evading an extra dollar is reduced by α. Therefore, the costs of collusion enter as an extra penalty for each dollar evaded. If consumers cannot commit not to whistle-blow, the new audit probability will be increasing in the number of consumers $N : \frac{\partial p_c}{\partial N} > 0$.

II.B Comparative statics for empirical analysis

The reported revenue Y^* that satisfies the optimal compliance decision with government monitoring only (equation 2) is lower than the reported revenue Y^{**} that satisfies the optimality condition once consumer monitoring is included (equation 4) in both cases considered. The increase in reported revenue $\Delta Y = Y^{**} - Y^*$ is driven by the reduction in the expected benefit of evasion. This reduction is due to the costs of collusion and to the fact that the audit probability might be higher if consumers can act as whistle-blowers. In this subsection, I describe two relevant mechanisms behind the effectiveness of a consumer monitoring program, which I investigate further in the empirical analysis.

Volume of consumers The enforcement change introduced by consumer monitoring is sharper the larger the increase in the audit probability. Under whistle-blower threat, therefore, the increase in reported revenue is higher for firms that have a large number of consumers N. This comparative statics follows from the increase in audit probability induced by a higher risk of a whistle-blower. The increase in reported revenue should increase with the 'foot traffic' of the firm, or volume of consumers for a given firm size or true revenue $\bar{Y}$. Also, for a given 'foot traffic' the relative change of reported revenue should be smaller for larger firms. This distinction between firm size and volume of consumers is relevant to shed light on one of the specific mechanisms of third-party information: the enforcement effect may result from exposure to whistle-blower threats.

Value of rewards In the framework described above, higher rewards to consumers should have stronger effects on compliance. In a collusive deal, firms try to match the rewards provided by the government through a discount. Therefore, the reward to consumers α directly reduces firms' benefits from evasion. If consumers have behavioral biases in assessing the odds of winning prizes, the value of rewards α may be even higher than the monetary value of the program's reward and there could be heterogeneity in the population in how much consumers value the program, making it particularly costly for a firm to replicate α through a discount.

III Institutional Background and Data

This section provides institutional background on the *Nota Fiscal Paulista* (NFP) policy, and the details of the program that are important for the empirical analysis. First, I briefly introduce the relevant features of the Brazilian tax system and the NFP policy. Then, I describe the datasets I use and sample definitions.[21]

III.A Institutional Background

The State of Sao Paulo is the largest state in Brazil: it accounts for 34% of the country's GDP, and has a population of 42 million people. The metropolitan area of Sao Paulo is the second most populous in the Americas. The state of Sao Paulo depends mostly on its own tax revenue, as opposed to federal transfers.[22] States in Brazil have two main tax instruments: a tax on goods and certain services (ICMS) and a property tax on motor vehicles (IPVA).[23]

The ICMS is a value added tax (VAT), and it is the most important source of revenue in Sao Paulo. Because the ICMS is a state-level tax in Brazil, its legislation and enforcement policies are determined by the states. The average ICMS rate is 18% over the valued added in Sao Paulo.[24] As is common in VAT across the world (Keen and Mintz, 2004), there is a threshold below which firms pay taxes over gross revenue instead of the value added.[25] Across all establishments, the average VAT paid as a share of reported gross revenue was 4% in 2007.

In 2007 the state of Sao Paulo collected U.S. $34.5 billion with the ICMS, equivalent to 7.6% of the state's GDP. Overall, tax revenue in Brazil is very high for developing country standards. Considering all taxes, tax revenue amounts to 34% of the country's GDP (IMF, 2011).[26] Nonetheless, there are many reasons to believe that tax compliance is not perfect in Brazil. According to La Porta and Shleifer (2014), estimates of size of the country's informal economy range from 19% to 34% of GDP. Unregistered firms are invisible to the tax authority, and no taxes are levied directly on them. Formal firms have to report their activity to the tax authority on a monthly basis, and pay the ICMS in relation to their reported activity. Despite the tax authority's monitoring, compliance by formal firms is also limited. In the World Business Environment Survey (2003), on average Brazilian formal firms claim that 20-30% of sales are not reported to the tax authority by a typical firm in their area of activity.[27] When the NFP program was implemented, the Secretary of Finance

[21] Throughout the paper I will convert Brazilian Reais to dollars using US$1=R$2 exchange rate, which is the average exchange rate during the period of analysis (2004 - 2011).

[22] When the NFP policy was implemented in 2007, Sao Paulo's own tax revenue was 75% of its total revenue according to the balance sheets of the Brazilian Treasury Department. Moreover, Luque et al. (2011) argue that Sao Paulo state generated more than 40% of the Federal tax revenue, while receiving less than 35% of Federal transfers in 2005. Federal taxes include, for instance, individual and corporate income taxes, payroll taxes and taxes on manufactured products.

[23] The IPVA ("Imposto sobre Propriedade de Veículos Automotores") and ICMS ("Imposto sobre Circulação de Mercadorias e Serviços") typically account for 95% of the total tax collected by states. The other two sources of tax revenue are a tax on bequests and donations called ITCMD ("Imposto sobre Transmissão Causa Mortis e Doações") and fees for public services.

[24] The value added is the total value of sales net of inputs.

[25] Firms that have yearly gross revenue of less than U.S. $1.2 million can choose to be in a simplified tax regime called SIMPLES in which firms pay taxes based on gross revenue. The ICMS average rate in the SIMPLES is 3.5% of gross revenue. For more details about SIMPLES see De Paula and Scheinkman (2010) or Monteiro and Assunção (2012).

[26] The average tax revenue as a share of GDP in developing countries is 17.6% (Gordon and Li, 2009).

[27] The question in WBES (2003) is: "Recognizing the difficulties many enterprises face in fully complying with taxes and regulations, what percentage of the total sales would you estimate a typical firm in your area of activity keeps off the books: 1 (none); 2(1-10%); 3 (11-20%); 4 (21-30%); 5 (31 - 40%); 6 (41 -50%); 7 (over 50%)." In the case of establishments that sell to final consumers, the tax evasion problem is likely to be more severe since firms are smaller than in upstream sectors. The percentage of sales that are underreported or not reported at all reaches 30-40% among smaller firms in

of Sao Paulo at the time argued that the retail sector in the state evaded taxes on approximately 60% of its sales (Jornal Estado de São Paulo, 2007).

III.B The Nota Fiscal Paulista program

The *Nota Fiscal Paulista* (NFP) program was created by the government of the state of Sao Paulo in October 2007 in order to reduce tax evasion of the state's VAT, and to foster a culture of tax compliance.[28] The idea behind the NFP program is to use consumers as tax auditors by introducing a set of monetary incentives and a system of cross-checks. Moreover, as is explained in more detail in this section, the incentives provided by the NFP program are increasing in the value of the purchase such that consumers have incentives to ask for receipts, and to make sure that the value of the purchase is reported correctly by the establishment. Therefore, the NFP program directly affects two forms of underreporting: (i) establishments may not report a transaction at all, or (ii) establishments may falsely claim a lower transaction value.[29]

NFP leverages the new availability of information technology in the developing world,[30] and the fact that individual identification numbers in Brazil - Social Security Number (SSN) equivalents - are not considered sensitive information.[31] The policy took advantage of a new system of data transmission of receipts: instead of keeping receipts in their books, establishments have to send the government all receipts they issue electronically.[32] In order to make sure retail establishments report their transactions truthfully, the government introduced the possibility of identifying the SSN of the buyer on each receipt, and created a system of tax rebates and monthly lotteries so that final consumers have incentives to request receipts.

Since the process of reporting receipts to the tax authority is done by establishments, and the consumer's SSN is attached to it, consumers do not need to send their receipts to the tax authority to get the rewards, which markedly reduces consumer participation costs. Consumers can create an online account at the tax authority's website, which allows them to collect rewards and cross-check the receipts issued with their SSNs. The online system also allows consumers to file complaints, which introduces a threat that consumers may act as whistle-blowers.

Eligibility. Any person that holds a Brazilian SSN equivalent is eligible to participate in the program.[33] No pre-registration is needed for consumers to be eligible for tax rebates. In order

Brazil.

[28]The NFP policy was framed as an incentive to improve tax morale. The official slogan of the policy was "Incentive Program for Fiscal Citizenship" ("Programa de Incentivo à Cidadania Fiscal").

[29]A common way to evade taxes in Brazil is to underreport the value of a sale. This type of evasion is informally known as "meia-nota" or "half-receipt" (Amaral et al, 2009).

[30]Bird and Zolt (2008) highlight that information technologies may play an important role in influencing tax administration and tax design in developing countries.

[31]For example, the Brazilian SSN equivalent (CPF) is written on checks under the signature line, and consumers are frequently asked for their SSN in business transactions.

[32]The system is called TD-REDF ("Transmissor de Dados para o Registro Eletrônico de Documento Fiscal" or "Data Transmitter for Electronic Registration of Fiscal Document"). For transactions between firms, another policy of electronic reporting of receipts was implemented: the NF-e ("Nota Fiscal Eletronica" or "electronic receipt"). For more details on electronic reporting of receipts see De Mello et al (2010).

[33]Throughout the paper I will refer to the CPF ("Cadastro de Pessoa Física") as SSN. I will focus on CPF holders only.

to be rewarded with lottery tickets for monthly cash prizes, consumers must create their online accounts.

The reward system. At the moment of purchase, the consumer may ask for the receipt, and give the cashier her SSN. Establishments must send all receipts - with or without SSNs - to the tax authority on a monthly basis. As the tax authority receives the receipts, it creates an account for each SSN where it stores all receipt information and the tax rebates due from each receipt.[34] If the consumer has an online account and has opted in for lotteries, the system also automatically generates lottery tickets for every total of U.S. $50 spent. During the registration, a consumer may also opt to receive an email every time a receipt is issued with her SSN. The online account displays how much consumers are rewarded for each transaction, and has tabs where a consumer can click to manage rewards and file complaints. Figure 1a shows an online account example, and Figure 1b displays a receipt with a consumer's SSN.[35]

Tax rebates. For a given receipt, consumers receive a tax rebate of 30% of the VAT collected by the final sale establishment in a month, shared among all consumers of that establishment who provided their SSN that month in proportion to their expenditure in that establishment and month. The calculation of the benefit is a function of an entire month's worth of SSN receipts and resultant tax revenue.[36]

Lotteries. NFP has held monthly lotteries since December 2008. For every U.S. $50 a consumer spends in NFP receipts per month, she receives one lottery ticket.[37] If the consumer opts in for these lotteries while enrolling online, lottery tickets are automatically generated based on the consumer's total expenditures in NFP receipts.[38] Lotteries are held around the 15th of each month, and each month 1.5 million prizes are distributed on average. Most prizes range from 5 to 25 dollars, and there are usually 3 large prizes from 15,000 to 500,000 dollars.

Collecting rewards. Rewards can be: (i) direct deposited into the consumer's bank account, (ii) used to pay other state taxes, (iii) transferred to another person with an online account or to a charity. Consumers must have an online account to manage the rewards, and there is a U.S. $12 minimum requirement for any type of transfer. Tax rebates are disbursed biannually. In April, tax rebates from July to December of the previous year are made available to consumers; in October the tax authority disburses tax rebates from purchases between January and June of the same year.

They are the overwhelming majority of participants in the program. Some NFP participants have a CNPJ ("Cadastro Nacional de Pessoa Jurídica"), which is a SSN for firms. Charitable institutions and condominiums also have CNPJ and receive the exact same benefits as final consumers.

[34]Throughout the paper I will refer to the receipts with SSN as *NFP receipts.*

[35]The snapshot of the online account and the receipt in figures 1a and 1b are the author's own online account.

[36]If the firm has N consumers in a month, the benefit consumer i receives from an NFP receipt depends directly on the total ICMS collected from establishment e in month m ($ICMS_{em}^{total}$), the total value of NFP purchases associated with consumer i and establishment e in month m (V_{iem}) and inversely on the total value of NFP purchases in establishment e in month m ($\sum_{j=1}^{N} V_{jem}$). Also, there is a cap on how much an individual consumer can receive: 7.5% of the total expenditure, which is 30% of the highest VAT rate (of 25%). Thus, $TaxRebate_{ime} = \min\{0.3 \cdot [ICMS_{em}^{total} \times \frac{V_{iem}}{\sum_{j=1}^{N} V_{jem}}], 0.075 \cdot V_{ime}\}$.

[37]The lottery draw in month m uses lottery tickets generated in month $m-4$. This 4-month gap is necessary in order to make sure that all disputes over missing or incorrect receipts are resolved before the lottery.

[38]Therefore, 50 receipts of 1 dollar value, and 1 receipt of 50 dollar value are equivalent, and generate 1 lottery ticket.

Lottery prizes can be collected soon after the results are released. Consumers have up to five years to claim the benefits.

Complaints. Consumers may file complaints regarding a purchase made at a specific establishment up to the 15th of the month following the purchase. The consumer must identify the establishment and select a reason for the complaint from a 5-option menu: (i) the establishment did not issue a receipt; (ii) the establishment refused to write the consumer's SSN on the receipt; (iii) the establishment issued the receipt but did not register it electronically; (iv) there is a discrepancy between the information on the receipt issued to the consumer and the receipt registered electronically at the tax authority; and (v) other reasons.[39] Consumers receive a part of the fines paid by the firm as rewards instead of the usual monetary reward when they file a complaint.

Fines. Establishments that do not issue the NFP receipt correctly are subject to penalties and potentially more comprehensive audits by the tax authority. Under tax law, establishments can pay up to 100% of the evaded tax, and there are additional penalties for for misreporting documents and receipts.[40] If a firm issues a receipt with an individual SSN and misreports the transaction, the process of punishing firms is straightforward if the consumer has a SSN receipt as proof of purchase. Note that under-reporting in this case would hurt consumers as they will not receive the reward associated with the purchase.[41] In this case, there are fines applied by the consumer's protection bureau PROCON (Fundação de Proteção e Defesa do Consumidor). Under consumer law, establishments must pay up to U.S. $740 per receipt issued with a SSN but not reported to the government.[42]

Implementation. NFP was implemented in the retail sector between October 2007 and December 2008. The tax rebate system and electronic submission of receipts was phased-in by groups of sectors between October 2007 and May 2008. The online system to file complaints was available starting in October 2008; the first lottery occurred in December 2008. In April 2009, the tax authority disbursed tax rebates for the first time from all purchases since October 2007, and every 6-months thereafter the government disbursed tax rebates according to the schedule described above.[43]

[39] At that point the consumer does not need to provide evidence to support her complaint, and she can describe details of her case in a text box. The establishment is notified that a complaint was filed via email or letter, and it has 10 days to respond to the complaint. If the consumer is not satisfied with the response, she can file an official complaint. Before this point, the tax authority is not involved in the case. If the consumer decides to file an official complaint, she has to submit supporting evidence by scanning or taking a picture of the receipt or any other proof of purchase. From that point onward, the tax authority and the Consumer Protection Bureau will review the case and apply fines accordingly.

[40] For the legislation on tax penalties *Part IV "violations concerning fiscal documents and tax forms"* of Decree 45490/00.

[41] Dyck, Morse and Zingales (2010) find that, in the context of U.S. corporate fraud, access to information and monetary rewards play an important role in encouraging whistle-blowing.

[42] For details of the legislation regarding the consumer protection law can be found at *Decreto Estadual 53.085, de 11 de Junho de 2008*.

[43] Even though the NFP program is targeted at final consumer sales, consumers who purchased directly from wholesalers and manufacturers could enjoy the same reward benefits as in retail purchases.

III.B.1 Statistics of NFP

Figure 2 shows the time series of the total number of receipts reported to the tax authority from the beginning of the program until the end of 2011. The three vertical lines indicate the beginning of phase-in, the end of phase-in, and the first lottery in December 2008. The purpose of the figure is to show the mechanical increase in the total number of receipts reported electronically by establishments to the tax authority as the program was being implemented. After May 2008, the total number of receipts submitted to the tax authority follows the seasonality of consumption.

Figure 3a describes the evolution of consumer participation in the program. The dashed line is the total number of consumers that ask for NFP receipts by month, and the solid line is the total number of consumers with online accounts. In any given month there are more people asking for SSN receipts than there are people who have online accounts. This gap highlights the fact that the cost to start participating in the program is relatively small: no pre-registration is needed since one just needs to have a social security number; but enrolling online might be more costly. The graph also shows that online registration took time to pick up: the online enrollment rate was 14% of the economically active population in 2008 and 60% in 2011.

Figure 3b shows the evolution of average expenditures captured in SSN receipts per month by consumers who enrolled online by the end of 2011. The figure is based on the *consumer sample* I describe in section III.C. It shows that the average monthly expenditure is increasing across time, suggesting that consumers may be increasing the share of their consumption expenditure for which they ask for SSN receipts.

During the period of analysis from October 2007 to December 2011, 13 million people enrolled online at the tax authority's website, and over 40 million people asked for SSN receipts more than once.[44] Over U.S. $1.1 billion has been distributed in tax rebates and lottery prizes. 740,000 establishments have submitted over 3.5 billion receipts with consumers' SSNs to the tax authority. During the period of analysis there was a total of 1,151,518 complaints sent to the tax authority by 135,102 different consumers regarding 134,054 different establishments to the tax authority during the period of analysis.

III.C Data and Sample Definition

In this section, I briefly describe each data source, and the summary statistics of the data. First, I present the establishment-level data and the main outcomes I examine in section IV.A and section IV.B. Second, I explain the datasets at the consumer level, and the key variables I use in section IV.C.2. In both cases, I focus on features of the data most relevant for my empirical analysis. Additional details on variable definitions and sample choices can be found in the Appendix A.

[44] The total number of different SSN numbers in the data is 50 million, but I exclude consumers that only asked for SSN receipts once from January 2009 to December 2011. Since any SSN holder in Brazil is eligible for the rewards, people in neighboring states and tourists may also participate (the total population of Sao Paulo is 42 million). Over 500,000 consumers with online accounts are from municipalities outside the state of Sao Paulo.

III.C.1 Establishment Data

I use administrative data on establishment-level tax returns and registry information from the Department of Finance of the state of Sao Paulo, Brazil from January 2004 to December 2011.[45]

Reported revenue. The NFP program aims to ensure that firms accurately report final sales. Accordingly, the gross revenue reported by an establishment is the key variable directly affected by NFP. Therefore this variable is the primary outcome in my empirical analysis of establishment compliance. All establishments must report their gross revenue to the tax authority on a monthly basis. For more details on the specific forms used to construct this variable, see the Appendix A.1.

In the empirical analysis, I focus on changes in reported revenue. I use log reported revenue in all regressions. In order to reduce the influence of outliers, I winsorize reported revenue by its 99^{th} percentile value.[46]

The total amount of taxes paid by an establishment in a month is an important variable, but it is not used directly as an outcome. The amount of taxes remitted or due by establishments has important measurement problems. For instance, it is affected by changes in tax collection rules or by changes in the tax forms. These types of changes generate mechanical increases and decreases in the time series of tax payments, even though the total tax liability of a establishment may not change. Importantly, in some cases, there is not a one-to-one relationship between an establishment's tax liability and how much it remits to the tax authority, due to tax withholding policies.[47]

Establishment characteristics. From the registry of firms of Sao Paulo, the main variable I use is the establishment sector of activity. Sectors are defined according to a 7-digit code of the Brazilian National Classification of Economic Activity (CNAE version 2.1). The retail sectors are all the sectors that start with 47 plus motor vehicle retail under sectors that start with 45. Wholesale is defined by all sectors that start with 46, plus motor vehicle wholesale under sectors that start with 45. The sector definition is very detailed. For instance: 472 is Retail food, beverages, tobacco; 4722-9 is Retail meat and fish; and 4722-9/01 is Retail meat (butchery). Throughout the paper, *sector* refers to the 7-digit definition, unless otherwise noted.

Employer-employee data. From the Brazilian Department of Labor, I use annual reported employment for all formal establishments in Brazil (RAIS/CAGED). The data available for this study include all establishments that have reported at least one employee between 2004 and 2011. The employment data have a 5-digit sector definition for all establishments. See the Appendix for more details.

[45]Due to confidentiality reasons, the data were de-identified, and no establishment data were provided from sectors that have fewer than five establishments, or from sectors in which one establishment is responsible for over 90% of the sector's tax revenue of that sector. In the groups of sectors I analyze - retail and wholesale - only 126 establishments were excluded from a total of 1,035,933 establishments registered in Sao Paulo over the period of analysis.

[46]I replace all values above 99th percentile of the reported revenue distribution by the 99th percentile value. The results are robust to winsorizing the top 5% or the top 0.1%

[47]Reported revenue is based on an accounting definition - monthly gross revenue generated by all of an establishment's sales, before deductions for expenses -, so it is not subject to mechanical changes from filling out different tax forms or changes in tax rules.

III.C.2 Establishment sample and Summary Statistics

Establishment sample. From the total of 1,283,777 unique establishments registered in Sao Paulo in the period of analysis,[48] I restrict attention to the 632,751 establishments classified as retail or wholesale - as described in the previous section . The final panel has approximately 20 million observations between January 2004 and December 2011.

For robustness checks I consider two other samples. One that only considers establishments already operating by January 2004 and that were still active the quarter before NFP implementation. This sample focuses on the effect of the program on active establishments that were created at least three years before the program started to keep a relatively stable number of establishments during the period of analysis. The second alternative sample restricts the previous sample further by only considering establishments that existed throughout the period of analysis, i.e., were active up to December 2011.

Sector sample. I aggregate the reported revenue of the *establishment sample* by 7-digit sectors of activity between January 2004 and December 2011. There are 210 sectors: 90 in retail and 120 in wholesale. The sector sample has 24,990 observations. As a robustness check for the aggregate results, I also consider the two alternative samples described above.

Employment sample. Because the data from the Department of Finance of Sao Paulo is de-identified, it cannot be matched with the employer-employee dataset (RAIS/CAGED). In order to analyze the impact of the program on employment, I construct a sample of establishments that follows as closely as possible the *establishment sample* from Sao Paulo.[49] I aggregate the employment sample by 5-digit sector definition to analyze the effect of the program on employment.

Table 1a describes the *establishment*, *sector* and *employment* samples. Statistics for the *establishment sample* include the monthly gross reported revenue by establishment for the key groups I use in the empirical analysis. On average, the revenue from SSN receipts accounts for a substantial share of the average revenue in retail, and over 40% of receipts reported by the establishments in the sample have SSNs attached to them.[50] Statistics for the *sector sample* display the average reported revenue across 7-digit retail and wholesale sectors I analyzed in section IV.A. The *employment sample* aggregates the annual employer-employee data by 5-digit sector, and the table shows the average number of formal employees per establishment for retail sectors registered in Sao Paulo, as well as in the other 26 states in Brazil.

[48]I exclude all establishments that report zero revenue across all months between January 2004 and December 2011

[49]I exclude observations from establishments active after October 2007, and I only include establishments that were already operating by January 2004.

[50]

As I explain in the next section, the data from the NFP program available to this project starts in January 2009. Therefore, establishment-level datasets generated from the program are not available before that.

III.C.3 Consumer Data

Consumer-level datasets are based on de-identified administrative data from NFP receipts and from online account activity at the tax authority's website.[51] Here I describe the datasets I use in section IV. Importantly, the consumer-level data are provided by the NFP program. Therefore, there is no "pre-NFP" data on receipts, or any other individual characteristic.[52]

Receipts data. This data file captures purchases for which final consumers asked for SSN receipts between January 2009 and December 2011. For these receipts the data include: month and year it was issued, the total amount spent, and an establishment identifier. The receipt dataset has information for all consumers that have made purchases with their SSN, even before they enrolled online. This feature of the data is important in the empirical analysis of section IV.C.2, in which I estimate the effect of online enrollment.

The main variables I derive from the receipts dataset are: (i) *number of receipts*: the total number of SSN-identified receipts for which a consumer asks per month; (ii) *number of establishments*: the number of different establishments for which a consumer asks for SSN-identified receipts per month; (iii) *total expenditures with a SSN*: the total amount of money spent associated with the SSN-identified receipts, aggregated by consumer, per month; (iv) *average receipt value*: the average value among all purchases represented by a consumer's SSN-identified receipts in a given month. In order to reduce the influence of outliers I winsorize the *number of receipts* and *total expenditure in SSN receipts* by their 99^{th} percentile value.

Online account data. This dataset contains information on month and year of enrollment, the timing of monetary reward collection, and the amount received by 13 million individuals who created an online account at the tax authority's website from October 2007 to December 2011.[53] Additionally, the dataset contains participation in monthly lotteries: the total number of tickets each consumer held and the associated prizes she received.

The main variables derived from the online account dataset are: (i) *total amount claimed*: the total value of rewards claimed by consumers through bank account deposits; (ii) *number of lottery tickets*: the total number of lottery tickets a consumer holds per month; (iii) and *lottery prizes*: the number of lottery prizes and the value of lottery prizes per month.

III.C.4 Consumer sample and Summary Statistics

Consumer sample. I take a 10% random sample of consumers who enrolled online by the end of 2011 - around 1.3 million people - and I construct a balanced monthly panel from the *receipts data*

[51]For confidentiality reasons, no information that may identify individuals was available to this study. A "fake" unique identifier was created for each individual SSN, and no information on names or addresses was provided. Also, for a given receipt, the total amount spent is rounded to the nearest integer, and the final data contains no information on prices or products that were purchased.

[52]The state tax authority has no information on individual income tax records or any other federal tax data. Apart from motor vehicle property information, state tax authorities do not usually collect data on individuals.

[53]All data on approximately 90 consumers who won one of the top 3 lottery prizes of over U.S. $500 were excluded from the datasets available to this study for confidentiality reasons. See Appendix for more details.

of consumer's participation in the program containing 46,505,268 observations between January 2009 and December 2011. A balanced panel is crucial to correctly count the number of receipts, since in the original data I only observe individuals when they ask for receipts. Table 1b displays descriptive statistics of the *consumer sample*.

Lottery sample. This sample covers consumers who participated in one of the twelve monthly lotteries between June 2010 and May 2011. I restrict attention to consumers holding fewer than 40 lottery tickets per lottery, which is relevant to assure common support between lottery winners and non-winners in the event-study analysis I describe in section IV.C.2.[54] I merge the data on lottery ticket holdings and lottery prizes from this sample with the *receipts data*. The combined dataset of lotteries and receipts covers the time period between January 2010 and November 2011, i.e., 6 months before and after the first and last lottery considered in this analysis.

The second panel of Table 1b displays the descriptive statistics of the lottery sample. Since consumers need to be enrolled online in order to be eligible for lotteries, and since they must ask at for least U.S. $50 in SSN receipts to get one lottery ticket, consumers in the lottery sample have much higher participation rates than the consumers in the *consumer sample*.

IV The effect of third-party information trails on establishment compliance

To investigate the degree to which the availability of third-party information trails introduced by consumer rewards for requesting receipts can improve firm compliance, I begin by exploiting the impact of the introduction of the NFP program on revenue reported by establishments using a difference-in-differences (DD) research design.

IV.A Compliance effects: retail vs. wholesale

The identification strategy exploits variation in treatment intensity from the policy change. I compare sectors affected differently by the consumer monitoring program: retail and wholesale. NFP targets final consumer sales, so establishments that sell mostly to final consumers are more affected than establishments selling mostly to other establishments. To exploit this difference I compare "treated" retail sectors to "control" wholesale sectors. I use a DD design to estimate changes in reported revenue by establishments in each group before and after the implementation of the program.

One advantage of the data is that I have a long time series of pre-NFP observations of reported revenue changes in the sector groups. Thus, I can shed light on whether a key identification assumption in a DD holds: that trends in potential reported revenue changes are parallel for retail and wholesale sectors. Figure 4a displays changes in total raw reported revenue by group of

[54]See the Appendix for a more detailed description of the *lottery sample*.

sectors from January 2005 to December 2011. In this figure each data point is scaled by the average monthly reported revenue before the introduction of the NFP in October 2007 for the group.

In figure 4a, retail and wholesale reported revenue changes closely trace each other until program implementation. The vertical lines highlight the key moments in the implementation of the program discussed in section III.B. After implementation, change in reported revenue gradually increases in retail sectors, relative to wholesale sectors. This gradual change is consistent with the fact that the program was not implemented at once, and consumer participation increased steadily over time. Since the figure displays raw data, there is quite a bit of variation across months of the year due to the seasonality of consumption. In particular, in retail sectors, reported revenue spikes each December, consistent with increased holiday-related consumption.

In order to measure the effect of the program across time, I run a flexible DD specification that includes 17 time dummies for 6-month windows from 2004 - 2011, using October 2007 (the starting point of the program's implementation) as a reference point. Each 6-month window, denoted by k, is associated with a dummy variable $Period_t^k$, which equals one if time period t falls within window k: [55]

$$lnR_{st} = \eta_s + \gamma_t + \sum_{k=-8}^{8} \beta^k (Treat_s \cdot Period_t^k) + u_{st} \tag{5}$$

where lnR_{st} is the log of reported revenue in sector s and time t; η_s are 7-digit sector fixed effects and γ_t are dummies for each month of each year. $Treat_s = 1$ if sector s is a retail sector, and u_{st} is clustered by sector. This specification allows me to show the treatment effect across time, while controlling for finely-defined time and sector effects.

Figure 4b plots the coefficients and the 95% confidence intervals from estimating equation (5) without a constant. The difference between the two groups is relatively constant before NFP. By the time the program is fully implemented - after the second dashed line - the difference in log reported revenue between the two groups begins to grow. This effect, averaged across all post-implementation periods, can be estimated from a standard DD specification:

$$lnR_{st} = \eta_s + \gamma_t + \beta Treat_s \cdot Post_t + u_{st} \tag{6}$$

where $Post_t = 1$ if $t \geq October$ 2007 and u_{st} is clustered by sector. Estimates of equation (6) suggest that the NFP program induced a positive and significant 22% increase in reported revenue by establishments across the 4-year period following implementation. Because I am exploiting differences in the treatment intensity across establishments, the estimated effect is a lower bound of the program's impact. The control group was also potentially affected by the policy: either directly from sales to final consumers or indirectly from the self-enforcing properties of the VAT.[56]

[55] For instance, $Period_t^0 = 1$ if $t \in [Oct.07, Mar.08]$, $Period_t^{-1} = 1$ if $t \in [Apr.07, Sep.07]$, and $Period_t^1 = 1$ if $t \in [Apr.08, Sep.08]$.

[56] I conduct a number of robustness checks. The results are robust to the alternative samples described in section III.C.2 for establishment samples, and to changes in top coding at top 5% or top 0.1% instead of top 1% to deal with the influence of outliers.

To make sure that the retail-wholesale comparison is indeed capturing an increase in compliance, rather than an increase in actual revenue, I use trade sector annual survey from the Brazilian Census Bureau (IBGE).[57] Two steps are taken to ensure that the survey elicits accurate information on establishments' activities. First, micro-level data are kept confidential. Second, Brazilian law ensures that no information reported in this survey can be used as evidence in a legal proceeding against an establishment. By comparing two independent sources of information on establishments' reported revenue - administrative data from Sao Paulo and the census data - I can address two alternative explanations for the differential change in reported revenue between retail and wholesale: *(i)* a nationwide differential change in the revenue ratio between the two groups;[58] *(ii)* an actual increase in retail revenue in Sao Paulo, relative to wholesale.

Figure 5 compares changes in the revenue ratio of retail to wholesale, $r \equiv \frac{\textit{retail revenue}}{\textit{wholesale revenue}}$, from the Sao Paulo administrative data to changes in the same ratio from the census survey.[59] Each data point is scaled by the ratio r in 2004. Until the introduction of NFP in 2007, the three ratios follow similar time trends. After 2007, the ratio derived from reported revenue in Sao Paulo tax data increase, whereas the ratios derived from survey data - in Sao Paulo state and nationwide - remain relatively unchanged. This is inconsistent with the notion that changes in reported revenue in Sao Paulo tax data are due to an increase in actual revenue in retail relative to wholesale sectors, or by a nationwide change.

IV.B Mechanisms: whistle-blower threats and consumer rewards

In order to investigate the mechanisms through which the government can credibly harness the information consumers have on firms' evasion to improve enforcement, I turn to the micro data on establishments, receipts, and consumers following the predictions from the conceptual framework in section II. First, I study the role of whistle-blower threats by examining heterogeneous effects of the program, and by analyzing the behavior of firms after consumers blow the whistle. Second, I investigate the role of consumer rewards and discuss how behavioral biases may amplify individual responses to rewards and make it more costly for firms to match the government incentives in a collusive deal.

[57]PAC ("Pesquisa anual do comércio") is an annual national survey conducted by IBGE based on a registry of all formal establishments in Brazil ("Cadastro Central de Empresas"). The survey is a census of establishments with more than 20 employees and a stratified sample of establishments with fewer than 20 employees. The retail-wholesale revenue ratio was calculated from aggregate tables of the survey. The micro-data is confidential.

[58]The time period post-NFP overlaps with the great recession in the U.S. that could have potentially affected retail and wholesale sectors differently. Therefore, the fact that the revenue ratio from the nationwide survey data is constant is an important indication that the difference in reported revenue between retail and wholesale after NFP implementation from Figure 4a is indeed a compliance effect.

[59]The national ratio is based on the total gross revenue from sales ("Receita Bruta de Venda"). Retail revenue includes the retail and motor-vehicle trade. Because the national data adds up all revenue within both groups of sectors, in this graph I consider all reported revenue in each group of sector from administrative data from Sao Paulo, instead of the reported revenue from the *establishment sample*.

IV.B.1 Whistle-blowers threats

I examine the effect of heterogeneity in the responses of establishments to the NFP policy in order to shed light on the role of whistle-blower threats discussed in section II. I use the *volume of transactions* to capture the increased threat of audit under consumer monitoring: the larger the number of consumers the more likely it may be that one of those consumers will blow the whistle when the firm evades taxes. Therefore, the effect of the program should be increasing in the volume of transactions.

To define the *volume of transactions* I count the number of receipts per establishment from the *receipts data*, and I rank retail sectors by the average number of transaction per establishment.[60]

$$lnR_{its} = \eta_i + \gamma_t + \sum_{m=1}^{k} \alpha_m (d_{ms}.DD_{ts}) + f(x_i)^* DD_{ts} + \varepsilon_{its} \tag{7}$$

where lnR_{its} is the log of reported revenue where in establishment i in period t and sector s,. Establishment fixed effects are denoted by η_i, γ_t is a month-year fixed effect. The term $f(x_i)$ is a 3rd-order polynomial of establishment size as measured by the average reported revenue three years before the program, and DD_{ts} variable is defined by the interaction between a dummy for retail sectors and a dummy that equals 1 for time periods after October 2007. The error ε_{its} is clustered by sector.[61] The term $d_{ms} = 1$ if sector s is in quintile k of the transaction volume distribution across sectors instead of the receipt value distribution. I flexibly control for establishment size effect through an interaction of DD_{ts} with $f(x_i)$ to separate the size effect from the effect of volume of consumers as discussed in the conceptual framework.

The establishment-level regression is run in a two-period DD, for which the data is collapsed by *pre* and *post*. The *pre* period is between January 2004 and September 2007, and the *post* period is between October 2007 and December 2011. This precaution helps to ignore serial correlation when computing standard errors (Bertrand, Duflo, and Mullainathan, 2004). The regressions are dollar-weighted - i.e., each observation is weighted by its pre-NFP value - such that each observation contributes to all regression estimates according to its economic scale to best approximate the sector aggregate-level analysis.[62]

Figure 6 plots the coefficients and 95% confidence interval from the estimating equation (7). The figure suggests that there is a monotonic increasing relationship between the transaction volume quintiles and the effect of the program: the effect of the program is stronger in sectors with a high volume of transactions. This pattern is consistent with the prediction described in section II.B.

Table 2 shows regression results of specifications similar to equation (7). Column (1) shows

---

[60] Examples of sectors classified as having low volume of transactions: art supplies and second hand shops. Examples of sectors classified as having high volume of transactions: supermarkets and gas stations.

[61] The results are robust to clustering by establishement, by firm or by time.

[62] The heterogeneity results do not depend on weighting. I find qualitatively similar results with attenuated magnitudes relative to the aggregate results when not dollar weighting because of the relatively high weight placed firms that collectively generate very little revenue.

the the simple DD coefficient for the full sample, and column (2) restricts attention to establishments that were already active in January 2004. The results are very similar in both samples, and are also similar to the aggregate result discussed in the previous section. In Column (3) I look at heterogeneity in the DD coefficient if a establishment is in a sector above or below median volume of transactions controlling for differential effects according to a 3rd order polynomial of pre-treatment firm size. The results simply reflect the patterns already observed in Figure 6. Column (4) shows heterogeneity in the DD coefficient according to weather a establishment is below or above the median pre-program size distribution, controlling for differential effects by volume of transactions. The results suggest that the effect of the program is relatively stronger for smaller firms, which is consistent with a higher baseline compliance among larger firms.

The results speak to a broader theory that information trails can have an enforcement effect through whistle-blower threats (Kleven, Kreiner and Saez, 2015). Kleven (2014) argues that derivative information makes full evasion infeasible even when there is no systematic third-party reporting. In the case of final consumer sales, firms are self-reporting income since the systematic VAT third-party reporting in business-to-business transactions along the supply chain is absent at the final consumer stage. This program is changing the availability of information trails, and the threat imposed by potential whistle-blowers might help to explain how this program can work despite collusion opportunities between the buyer and the seller.

IV.B.2 Whistle-blower event study

The evidence above indicates that whistle-blower threats could be an important device to improve compliance. In order to further examine how whistle-blowers affect firm behavior, I exploit a direct link between the participation of consumers in the enforcement effort and firm behavior. I use a dataset with over 1 million complaints to analyze how firms respond after a consumer blows the whistle. Once a firm receives a consumer complaint, it may increase the perceived probability of being audited due to whistle-blower threats.

Every month a firm may receive a complaint from a consumer through the NFP website. Typically a firm is notified by a complaint one month after the purchase. In order to study the effect of consumers' complaints I examine the impact of the first complaint. Different firms received their first complaints at different points in time, and I can exploit the timing of the first complaint to assess the response of firms. The likelihood of receiving a complaint in a given point in time, however, may be driven by the volume of sales leading up to the first complaint. It is possible that firms that have a large volume of sales in a given month may be followed by a lower reported revenue in the next period due to mean reversion or other seasonal characteristics. Therefore, exploiting the timing of the complaint alone might not be ideal, as subsequent changes in reporting patterns after the first complaint might reflect real changes in economic activity of the firm.

In order to circumvent mean-reversion and other seasonal effects, I build a counterfactual for each complaint event. Following Hilger (2014) I create an "event-control" group composed of firms that did not receive their first complaints by a given date. I use a subset of the establishment sample defined in section III.D: I consider only retail firms, and within retail I only retain the firms that did not exit before 2009. The data is monthly between June 2009 and May 2011. Throughout this period, 134,054 or 25% of establishments received at least one complaint. I use a re-weighting method based on quartiles of the propensity score of getting a complaint in a given period to control for firm characteristics and past outcomes.[63]

Let $g \in \{T, C\}$ index each firm as "complaint" T or a "no-complaint" C in a given month. Let t_O index the month in which an outcome is observed, and t_E index the month in which a consumer blows the whistle on the firm for the first time (the "event-month"). Define $k \equiv t_O - t_E$ as the number of "periods" or months after/before the first complaint. I performed this re-weighting exercise separately for each month between June 2009 and May 2011, and I collapsed the data by event-month $k, k \in [-6, 6]$, using the propensity score weights.

Figure 7a displays changes in the total number of transactions *complaint* and *no-complaint* firms report to the tax authority and $k \in [-6, 6]$, and figure 7b shows changes in reported revenue relative to 6 months before the first consumer blows the whistle. The x-axis shows the distance in months to the first complaint or "event-month." The graph displays the estimated DD coefficient from estimating the following equation on the collapsed data for $k \in [-6, 6]$:

$$lnY_{gj} = \gamma_j + \pi_g + \beta \cdot I\{j \geq 0, g = T\} + u_{gj}, \tag{8}$$

where lnY_{gj} is either the log of the number of receipts group g reports to the government in "event-month" j or the log of reported revenue in "event-month" j. The graph displays the estimated DD coefficient from estimating equation 8 on the collapsed data for $k \in [-6, 6]$. I find a significant 14% increase in the number of receipts establishments issue and a 6% increase in reported revenue over 13 months. The impact of the first complaint is capturing the overall impact of complaints, as some firms received additional complaints after time zero. It can be interpreted as an increase in the perceived probability of a government audit as the complaint may be used to flag non-compliant firms.

Together, the impact of consumers blowing the whistle and the heterogeneous effect of the consumer-monitoring based on whistle-blower threats is consistent with the argument that whistle-blowers can be an important part of the explanation for why third-party reporting is so effective to ensure compliance. In the context of NFP, it can be a tool for the government to tap into the wealth of information that consumers elicit when asking for receipts from hard-to-tax firms that self report final sales.

[63]The propensity score of a firm receiving its first complaint at a given time is estimated using time specific trends for each sector, age of the firm, number of establishments by firm, dummy for establishment-headquarter, dummies for legal nature of the firm, sector and time fixed effects, dummy for location in the metropolitan region of Sao Paulo, and the three lags of third-order polynomials of reported revenue, reported receipts, SSN receipts and number of consumers. For a detailed description of the propensity score and reweighting see the Appendix B1.

IV.C Consumer responses to rewards

As discussed in the conceptual framework, the more consumers value rewards, the more firms need to compensate consumers in a collusive deal, which will increase the effectiveness of consumer monitoring. First, I provide evidence that the program is salient by exploiting variation in the disbursement schedules of the monetary rewards. Then, I exploit variation from the monthly lotteries to investigate whether potential behavioral biases with respect to lotteries may amplify the response consumers have from rewards.

IV.C.1 Are consumers paying attention to the rewards?

In this section I show that consumers are paying attention to the schedule of lottery prizes and tax rebate disbursements. First, I verify that the release of monthly lottery results is salient to consumers by examining changes in the volume of Google searches about NFP. Google data aggregates information from millions of searches, and they can meaningfully capture salient social patterns that other survey methods cannot capture as easily (Stephens-Davidowitz, 2014).[64] Around the 15th of each month, the tax authority performs the lottery draws and releases information on lottery winners. A consumer can only check her lottery results by logging in to her online account at the tax authority's website. The actual address is not straightforward to remember (http://www.nfp.fazenda.sp.gov.br); as a result, consumers looking for this address may search for the program's name or initials.

Figure 8a pools Google search data from the first to the last day of each month between 2008-2011, and it scales each data point by the first day of the month.[65] From the figure, it is clear that there is an increase in search volume around the time the tax authority releases the results of the lotteries: it is 16% higher than on the first day of the month. The gray line displays data from searches with the word "futebol" (soccer in Portuguese) which provides a metric of how the general volume of Google searches varies within a month.

Second, I examine whether consumers are paying attention to the tax rebate schedule of the program. As described in the previous section, the tax authority disburses tax rebates biannually. Figure 8b shows that the timing of disbursement is salient: the total amount of rewards requested for bank account deposits spike as soon as tax rebates become available every April and October.[66]

IV.C.2 The lottery effect

As detailed in section II.B, the more consumers value the rewards α, the more effective NFP will be in preventing tax evasion.

[64] Hoopes et. al. (2014) use Google and Wikipedia searches about U.S. income tax to show that the propensity to search varies systematically with tax salience.

[65] I exclude the months of April and October - during which the government disburses the tax rebates - to make sure that the search pattern is related to the lotteries. Including these two months does not change the pattern in the graph.

[66] Consumers can use rewards in other ways - e.g., they can be transferred to a third party, used to pay other taxes or saved for a later deposit - so the total amount in the graph will not necessarily add up to the total amount available to consumers at that point in time.

The lottery component of the rewards may leverage consumers' taste for gambling or individual behavioral biases. Friedman and Savage (1948) noted that individuals may even be willing to pay for lotteries paying a negative expected value. Filiz-Ozbay et al. (2013) find evidence that prize-linked savings offered by commercial banks and governments around the world may be more effective at increasing savings than regular interest payments with the same expected value.

In addition, the NFP monthly lotteries typically have three very large prizes - the top prize can be as large as U.S. $500,000 - and millions of small prizes, which is a payoff structure commonly seen in gambling games and prize-linked savings accounts (Guillen and Tschoegl, 2002). The skewness of the prize values may be a tool to create salience. Bordalo et al. (2013) argue that when comparing alternative risky lotteries, individuals pay attention to the payoffs that are most different relative to their objective probabilities. If consumers exhibit behavioral biases with respect to the NFP lotteries, it would be more difficult for firms to try and replicate the government's rewards to avoid truthfully reporting sales.

Next, I use two different sources of variation in lottery rewards to investigate how lottery prizes affect consumers' participation in the program above and beyond the tax rebate effect, and I discuss potential behavioral considerations that emerge from the empirical analysis. I first focus on the random variation in lottery wins to document consumer behavioral responses to lottery rewards, and then I show suggestive evidence of the effect of lottery eligibility on consumer participation.

The effect of lottery wins I examine how consumers react when they win a lottery prize. Consumers may use past wins as a signal of their likelihood of getting a lottery prize, which would be consistent with misperception of randomness and the use of heuristics in making choices under uncertainty. Guryan and Kearney (2008) find that consumers increased their estimate of the probability a ticket bought from the store that sold a winning ticket in the past would be a lottery winner (the "lucky store effect").[67]

To analyze the effect of lotteries I create a natural "event-control" group composed of people that held the same number of lottery tickets in a given lottery but did not win prizes. I use the lottery sample defined in section III.D: all consumers that participated in monthly lotteries between June 2010 and May 2011. Each of the 12 monthly lotteries had 1.5 million lottery prizes and over 50 million lottery tickets. There are typically 1,407,394 prizes of U.S. $5, 76,303 prizes of U.S. $10, 15,000 prizes of U.S. $25, 1,000 prizes of U.S. $125, and 300 prizes of U.S. $500. Because it is common for individuals to hold more than one lottery ticket in a given month, there are many cases of consumers that received a total of U.S. $15, U.S. $20, U.S. $30, U.S. $35, by winning more than one prize.

Let $g \in \{T, C\}$ index each consumer as "winners" T or a "non-winners" C in a given month. I use a re-weighting method based on DiNardo, Fortin, and Lemieux (1996) to flexibly control for

[67]They argue that consumers may rationalize the observed streaks by inferring heterogeneity in the data generating process. In the context of financial investments, Kaustia and Knüpfer (2008) find evidence of reinforcement learning in investors' behavior: personally experienced outcomes are overweighted in future choices.

the number of lottery tickets individuals hold. I create bins for each possible number of lottery ticket holdings up to 40 tickets, which is the set of lottery tickets for which there is common support between the two groups.[68] I then re-weight the non-winners group such that each bin carries the same relative weight as the analogous bin in the winner group distribution across lottery ticket holdings. This method ensures that I use the random component of the lottery by matching the two groups based on the odds of winning prizes.[69]

Let t_O index the month in which an outcome is observed, and t_E index the month in which the consumer wins the lottery (the "event-month"). Define $k \equiv t_O - t_E$ as the number of "periods" or months after/before the lottery win. I performed this re-weighting exercise separately for each of the 12 lotteries, and each prize level between U.S. \$10 and U.S. \$35, and also for U.S. \$125 and U.S. \$500 prizes. I then collapsed the data for each lottery by group g and period $k, k \in [-6, 6]$, using the DFL weights, and I took the average number of SSN receipts across the 12 lotteries.

Figure 9 displays the average number of receipts for which lottery winners and non-winners ask by six different prizes levels and $k \in [-3, 3]$. The x-axis shows the distance in months to the lottery or "event-month." Each graph displays the estimated DD coefficient from estimating the following equation on the collapsed data for $k \in [-6, 6]$:

$$y_{gj} = \gamma_j + \pi_g + \beta \cdot I\{j \geq 0, g = T\} + u_{gj}, \tag{9}$$

where y_{gj} is the average number of SSN receipts group g asks in "event-month" j. Figure 9 shows that there is a significant difference in behavior between lottery winners and non-winners for all prize levels displayed. As the size of the lottery win grows, the estimated effect is larger. This pattern indicates that the change in behavior is indeed due to the lottery win, and that consumers are attuned to the lottery results.

The evidence is consistent with a behavioral explanation, given that there is a significant 0.6% difference between the number of receipts lottery winners and non-winners ask for even for a U.S. \$5 prize. Since the odds of winning are independent of past wins, the change in behavior observed in Figure 9 is consistent with consumers using the past lottery win as a signal of luck, and therefore perceiving a higher expected return from participating in the program. Since the effect is increasing in the lottery size, it is possible that the size of the prize matters for strength of the signal. The effect, however, is confounded with the fact that larger prizes are more relevant cash shocks that can increase the level of overall consumption.

An alternative explanation is that consumers use lottery wins as a signal that the program works as advertised. Figure 10a shows the effect of a U.S. \$5 win for a sample of individuals that won the lottery once before, in which the effect of confirming that the program works should not be as relevant. I find a .7% statistically significant difference in the number of receipts consumers

[68] Figure A shows two examples of the distribution of lottery ticket holdings among winners and non-winners. It is clear that the winner group typically holds more lottery tickets. Since the number of lottery tickets is determined by consumers' participation 4 months before the draw, it is important to carefully control for the odds of winning.

[69] For a detailed description of DFL-reweighting, see Yagan (2013). For details of this specific application, see Appendix.

ask for this subsample. Figure 10b shows that the 5-dollar win affects not only the number of receipts a consumer asks for, but also the number of different businesses in which a consumer asks for receipts. In order to further investigate the effect of winning a lottery, Figure 10c plots the effect of 5 dollar prize differences for different levels of prizes. There is a concave relationship between the change in the number of receipts consumers ask for and the prize level, i.e., the effect of winning a lottery prize at all is relatively larger than the effect of incremental prizes of similar size on consumer participation.

The effect of lottery eligibility In order to shed light on the relative effectiveness of tax rebates and lottery prizes, I next exploit the timing in which consumers enroll online and become eligible for lotteries. Consumers do not need to enroll online to start accumulating tax rebates, so the rebate incentive does not vary around the time the consumer creates her online account. Even though the timing of enrollment is not random, the most relevant incentive change when a consumers enrolls online is that she can opt in for lottery prizes.

I follow Jacobson, LaLonde, and Sullivan (1993) in estimating the dynamic effects of online enrollment on consumer participation in the program using an event-study design. I use the *consumer sample* described in section II, and I focus on four outcomes defined in section II: *number of receipts*, *number of establishments, total expenditures with a SSN, and average receipt value*. Let t_O index the month in which an outcome is observed, and t_E index the month in which the consumer enrolled or "event-month." Define $k \equiv t_O - t_E$ as "period" or months after/before event, then:

$$y_{i,t_O,t_E} = \sum_{j=-6}^{6} \beta_j \cdot I\left\{j = k\right\} + \gamma_i + \pi_{t_O} + u_{i,t_O,t_E} \tag{10}$$

where y_{i,t_O,t_E} is one of the four outcomes of interest, $I\left\{j = k\right\}$ is a dummy variable that is equal to 1 when $j = t_O - t_E$, γ_i refers to individual fixed effects, π_{t_O} models time fixed effects, and u_{i,t_O,t_E} is the error term that is clustered by municipality. This specification has individual and calendar time fixed effects, so I identify β_j by exploiting variation in the timing of enrollment. Figure 11 displays the estimated $\hat{\beta}_j$ from Equation (10) and 95% confidence intervals, where $k = -1$ is the omitted category. In each graph, I add the sample mean before enrollment to facilitate interpretation.

Figure 11a shows that consumers asked for four receipts on average with SSNs before they enrolled online, when they were only eligible for tax rebates. After enrolling at period zero, the average number of receipts that they asked for per month doubled, and the change seems to reflect a permanent level shift in participation. The number of different establishments in which a consumer asked for receipts also increases considerably in Figure 11b. Figure 11c shows that there is an upward trend in the total amount spent in SSN receipts before an individual enrolled online, but there is a sharp jump in the total SSN expenditure at the moment of registration. The average receipt value drops from U.S. \$120 to U.S. \$25 in Figure 11d; so consumers were asking for receipts more often and more widely.

The effects described in figures 11a to 11d cannot be exclusively attributed to the lottery eligibility since the decision to enroll online could be explained by a shock that affects both the decision to enroll and the change in behavior documented in the figures. Nonetheless, the sharp differences in participation between before and after online enrollment, and the fact that consumers conditioned their participation in the program on past lottery wins, are consistent with lotteries being an effective incentive device for consumers. Since lotteries are relatively less costly than tax rebates, a composition of rewards that puts more resources into lotteries relative to rebates may potentially be more cost-effective. More research is needed in order to pin down the relative effectiveness of $1 in tax rebate versus $1 in lottery prizes.

V Implications for tax policy

The empirical analysis performed in section IV shows that incentives for consumers to ensure that firms accurately report transactions can be an effective way to improve firm compliance in final sales transactions. In this section, I explore two dimensions of policy implications beyond the compliance effect. First, I examine the impact of this policy on exit and employment decisions of firms, as the increase in enforcement can imply an increase in the effective tax rate.

Second, I discuss the cost-benefit of the NFP. By carrying out this policy, the government of Sao Paulo is forgoing part of its tax revenue: both incremental revenue from the program, and inframarginal revenue. Moreover, there are likely non-trivial compliance costs and benefits associated with NFP. In this section I perform a cost-benefit analysis of the policy, and I discuss potential ways to make such policies more cost-effective.

V.A Establishment exit and employment

The observed enforcement affect could increase in firms' tax liability, i.e., it can imply an increase in the effective tax rate. This change may affect establishments that were on the margin of exiting the market or firing employees. I analyze each effect in turn.

Exit. I define an establishment's month of exit as the last month I observe that establishment in the data, and I consider all exits between 2005 and 2010. The data for the analysis of exits comprise all establishments in retail and wholesale, i.e., without the restrictions from *establishment sample*. Figure 12a shows yearly exit rates by retail and wholesale sectors, where the exit rate is defined as the total number of exits in year t and sector s, divided by the total number of establishments in year $t-1$ in sector s.[70] The figure also shows the DD coefficient from estimating a specification similar to equation (6) in a 7-digit sector yearly panel, where the exit rate is the dependent variable. The coefficient is not statistically distinguishable from zero, which indicates that on average the policy did not affect establishments' decisions to exit during the period of analysis.[71]

[70] I exclude 2011 because my sample period ends at the end of 2011. Therefore, I allow at least consecutive 12 months without observing a firm in the data to be sure that the firm exited.

[71] This result is robust to an alternative exit measure, where I consider as exit the last date I observe a firm report non-zero revenue. Many firms that do exit may still submit forms with zero activity to avoid or postpone the paperwork required to close a firm in Brazil.

Employment. To investigate employment effects I use the *employment sample* described in section III.C. As opposed to the tax data from Sao Paulo, this sample covers the entire country. As a result, I can use retail sectors in other states as a counterfactual for retail sectors in Sao Paulo. Figure 12b displays log employment in retail in Sao Paulo and in other states. The figure also shows the DD coefficient from estimating a specification similar to equation (6) in a 5-digit sector yearly panel, using log employment as the dependent variable and adding state fixed effects along with time and sector fixed effects. The coefficient is very close to zero, suggesting that the policy, on average, had no effect on establishments' formal employment decisions.[72]

The evidence above indicates that the increase in tax enforcement did not affect employment or exit decision of firms. However, it is possible that changes in employment and exit may occur after the period of analysis. The fact that I find no effect on employment is consistent with the increase in reported revenue being a reporting effect, rather than an actual increase in sales, in which case I could potentially observe an increase in employment or a drop in exit. As this is a reporting effect, it could impact the effective tax rate firms face and negatively affect the ability of firms to keep employees or survive in the market. The null effect indicates that the implied increase in the effective tax rate may not be large enough to affect the firm along these margins, and may just reduce evasion rents. The average tax paid over reported revenue was 4% before NFP, so a 22% increase in the effective tax rate might not be a large change in the net of taxes revenue.[73] Another explanation is that firms can potentially adjust other margins that I do not observe in the data such as, for instance, firing informally-hired workers.

V.B Cost-benefit analysis

The government of Sao Paulo is forgoing part of the tax revenue collected at at the final consumer stage: both incremental revenue from the program, and infra-marginal revenue. Therefore, it is not clear that the program is able to increase revenue net of transfers. The government is rewarding consumers with 33% of the tax collected in final sales transactions: 30% in tax rebate and 3% in lottery prizes. Thus the 22% increase in compliance reported in section IV would imply that the potential increase in tax collection would not be enough for the government to break even. However, the NFP is arguably relying on the fact that some consumers may never collect rewards. As of 2011, 50% of the rewards were not collected. In particular, there are 27 million consumers that asked for SSN receipts but did not enroll online in the first four years of the program, which is the only way one can claim rewards. If one only takes into account rewards claimed by consumers, the program would be breaking even. It is important to consider, in addition, that the 22% increase is a lower bound for the effect of the program, so the total potential increase in tax collection could be larger and the program could be generating additional revenue net of rewards.

The rewards, however, should arguably be considered as transfers and not costs. There might

[72]Since the *employment sample* only covers formal employment, this exercise does not rule out the possibility that informal employment changed as a result of NFP.

[73]Moreover, establishments may be able to pass-through this tax increase to consumers. Data on prices and quantities - which have not been available for this project - would be needed to understand the incidence of the policy.

be actual costs borne by the program participants. A number of countries have experimented with policies designed to change incentives at final sales through monetary rewards to consumers who ask for receipts as a solution to the 'last mile' problem of the self-enforcing mechanism of the VAT. Berhan and Jenkins (2005) argue that policies that introduce incentives for consumers to ask for receipts usually have high participation costs. The administrative burden is normally borne by consumers that need to submit receipts or use receipts to get rewards[74] NFP's implementation is innovative in reducing participation costs to consumers. It leverages the fact that establishments need to report receipts electronically to the tax authority regardless of whether the receipts have a consumers SSN attached, and it depends on the fact that Brazilians are used to frequently disclosing their SSN. Consumers need to pay a time cost to spell out their SSN and to enroll online, but the lottery tickets and tax rebates are automatically generated by the tax authority once the receipt is reported by the establishment.[75] It is important to consider that, in addition to consumers' participation costs, there are also administrative costs for the government and compliance costs borne by firms. It is also possible that there are other margins of evasion that increased as a result of this policy. [76]

Similarly, there may also be other social benefits to be considered in a cost-benefit analysis beyond the direct effect of the policy. Consumer-reward programs are often framed as a way to encourage a culture that values tax compliance (or "tax morale"). If this channel is relevant, the program could potentially generate a shift in consumer's propensity to ask for receipts even if the government eventually discontinues the rewards. The program also allows consumers to provide the SSNs of charities instead of their own, which may increase utility from altruistic motives.[77]

Moreover, by improving enforcement, consumer monitoring may also help tilt the playing field in retail away from firms that evade taxes toward the most-efficient firms. There may also be redistributive benefits if lower income individuals value lotteries relatively more and consume a larger share of their income; or if the government puts a higher weight on consumers than on tax evaders.[78] In the case of the NFP program, a likely important benefit is the new information elicited by consumers from establishments through their complaints and participation in the program, which can be used by the tax authority to optimize audits.[79] Another potential benefit could

[74]Often the receipts themselves are lottery tickets - as in the Chinese lottery program - or consumers need to submit receipts to the tax authority to get tax refunds - as in Bolivia.

[75]Consumers may want to hold on to their receipt in order to help file complaints, but it is not strictly necessary to do so to derive the benefits from the program.

[76]For instance, firms could be firing informal labor or they could be evading other taxes that are not affected by this new monitoring policy. Carrillo, Pomeranz and Singhal (2014) find that an increase in enforcement was followed by taxpayers making offsetting adjustments on less verifiable margins in Ecuador.

[77]Since the option of indicating the SSN of a non-profit organization was applied in May 2009 until December 2011, 6% of all SSN-identified receipts had a non-profit's SSN.

[78]According to Andreoni et al (1998) it is an unresolved issue whether governments put different weights on cheaters vs. honest tax payers.

[79]For instance, the patterns of receipt reporting by establishments can be used to flag potential tax evaders. Consumers often forget or decide not to ask for receipts, so establishments are not expected to report 100% of receipts with SSN. Nonetheless, frequently establishments report 100% of receipts with SSN across all months of the year, which suggests that establishments might be selectively reporting transactions to the tax authority that have a higher probability of being caught, while underreporting the transactions for which consumers did not ask for receipts.

be the fiscal externalities to the federal tax authority that levies other taxes on firms that could benefit from the increase in compliance and information generated by the state-level program.

There are arguably ways to make the policy more cost-effective. For instance, Sao Paulo has created a SSN barcode card to mitigate participation costs: consumers may scan the card at the moment of the purchase instead of verbally reporting the SSN for every transaction. Importantly, millions of people are already participating and have paid the fixed cost of setting up online accounts. Given the take-up NFP has already achieved, the program could potentially change some of the incentives to become more cost-effective in the future. Perhaps the rewards could rely more on lotteries, relative to tax rebates. Lotteries cost less than 15% of the total amount of rewards to consumers, and due to consumers' potential behavioral biases discussed in section above lotteries may provide a stronger incentive than tax rebates per dollar spent by the government. It would be important to build more evidence on the relative cost-effectiveness of different reward options - tax rebates, lottery in-kind prizes or cash lottery prizes - that are common in policies that incentivize consumers against tax evasion.

VI Conclusion

Access to substantial third-party information trails is widely believed to be critical for modern tax enforcement. This paper has investigated how the availability of third-party information can improve firms' compliance. I exploit administrative data and quasi-experimental variation from a policy that rewards consumers for ensuring that firms accurately report final sales transactions to the government in Sao Paulo, Brazil.

I find that the program increased revenue reported in retail sectors by at least 22% over four years. I examine heterogeneity across establishments and consumer responses to rewards to shed light on the mechanism through which third party reporting can improve compliance despite collusion opportunities. I find that the estimated effect is stronger for sectors with a high volume of transactions, consistent with shifts in audit probability that increase in firm size due to whistle-blower threats. I also provide direct evidence on the enforcement effect triggered by consumers blowing the whistle: firms report 14% more receipts and 6% more revenue after receiving their first complaint.

Furthermore, I show that consumers are finely tuned to the incentives of the program, and I exploit the random component of lottery rewards to investigate the effect of lotteries on consumer engagement with the policy. I find that that consumers condition their participation on past lottery wins. Even small prizes generate a significant and steady increase in the number of receipts consumers request, and in the number of different businesses in which they ask for receipts. The results are consistent with the possibility that lotteries amplify consumer responses due to behavioral biases, which would make it more costly for firms to try to match government incentives in order to collude with consumers.

The findings of the paper are consistent with the argument that whistle-blower threats and collusion costs could help to explain how self-enforcing incentives can be effective to harness third-party information despite collusion opportunities. From a policy perspective, this study sheds light on how citizen engagement can be used as a monitoring tool in a participatory program. In the context of VAT systems, the results indicate that incentives to consumers can potentially help address the lack of cross-reporting at the final consumer stage, which is a well-known shortcoming of one of the most important and prevalent tax instruments in the world. In particular, the paper provides supporting evidence that consumers respond to lottery incentives to ask for receipts, which is the most common policy reward used by governments to mobilize consumers against tax evasion.

References

1. **Allingham, M. and A. Sandmo**. 1972. "Income tax evasion: a theoretical analysis", *Journal of Public Economics*, Vol. 1, 1972, 323-338.

2. **Amaral, Gilberto L., João E. Olenike, Letícia M. F. Amaral and Fernando Steinbruch.** 2009. "Estudo sobre a Sonegação Fiscal das Empresas Brasileiras. " *Instituto Brasileiro de Planejamento Tributario.*

3. **Andreoni, J., B. Erard and J. Feinstein**.1998. "Tax Compliance", *Journal of Economic Literature, 36*: 818-60.

4. **Arbex, Marcelo and Enlinson Mattos**. 2015. "Optimal sales tax rebates and tax enforcement consumers?," *Oxford Economic Papers, 67* (2): 479-493.

5. **Besley, Timothy and Torsten Persson.** 2013."Taxation and Development." in Alan J. Auerbach, Raj Chetty, Martin Feldstein and Emmanuel Saez, eds., *Handbook of Public Economics, 5*: 51-110.

6. **Besley, Timothy and Torsten Persson.** 2014."Why Do Developing Countries Tax So Little?", *Journal of Economic Perspectives, 28*(4): 99120

7. **Barroso, Rafael and João Pedro Cortez**. 2009. "Devolvendo impostos para arrecadar mais: uma avaliação dos efeitos da Nota Fiscal Paulista na arrecadação do ICMS no Estado de São Paulo." *Working Paper*.

8. **Becker, Gary S.**.1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy, 76* (2): 169-217.

9. **Berhan, Bahro A. and Glenn P. Jenkins**. 2005."The High Costs of Controlling GST and VAT Evasion." *Canadian Tax Journal, 53*, 720–36.

10. **Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan**. 2004. "How Much Should We Trust Differences-In-Differences Estimates?" *The Quarterly Journal of Economics, 119*, (1): 249-275.

11. **Bird, Richard M.** 1992. "Tax reform in Latin America: a review of some recent experiences." *Latin American Research Review, 27* (1): 7-36.

12. **Bird, Richard and Pierre-Pascal Gendron.** 2007. *The VAT in Developing and Transitional Countries*. Cambridge University Press.

13. **Bird, Richard and Eric Zolt.** 2008. "Technology and Taxation in Developing Countries: From Hand to Mouse." *National Tax Journal*, Vol. LXI, No. 4, Part 2

14. **Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer**. 2013. "Salience and Asset Prices."*American Economic Review: Papers & Proceedings, 103*(3): 623–628.

15. **Bruhm, Miriam and David McKenzie.** 2013. "Entry Regulation and Formalization of Microenterprises in Developing Countries." *The World Bank Policy Research Working Paper* 6507.

16. **Carrillo, Paul, Dina Pomeranz and Monica Singhal.** 2014. "Tax Me if You Can: Evidence on Firm Misreporting Behavior and Evasion Substitution." *Working Paper*.

17. **Cowell, F. A**. 2004. "Carrots and Sticks in Enforcement." In: Aaron, Henry J., and Joel Slemrod, eds. *The crisis in tax administration*. Brookings Institution Press.

18. **De Mello, Newton O., Marcelo L. A. Fernandez, Vidal Augusto Z. C. Melo, Eduardo M. Dias, Caio F. Fontana.** 2010. "New Technologies for Nota Fiscal Paulista (São Paulo Tax Invoice): Automation of the Tax Documents issue Process in the Retail of the State of São Paulo - Brazil." *World Scientific and Engineering Academy and Society Transactions on Systems and Control*, 12 (5): 910 -917.

19. **De Paula, Aureo and José A. Scheinkman.** 2010."Value Added Taxes, Chain Effects and Informality." *American Economic Journal: Macroeconomics, 2*(4):195-221.

20. **DiNardo, John, Nicole Fortin, and Thomas Lemieux**. 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica, 64* (5):1001-1044.

21. **Dyck, Alexander, Adair Morse and Luigi Zingales.** 2010. "Who Blows the Whistle on Corporate Fraud?" *Journal of Finance, 65* (6): 2213-2254.

22. **Fabbri, Marco.** 2013. "Shaping Tax Norms Through Lotteries." Working Paper.

23. **Filiz-Ozbay, Emel , Jonathan Guryan, Kyle Hyndman, Melissa Kearney and Erkut Y. Ozbayk.** 2013. "Do Lottery Payments Induce Savings Behavior? Evidence from the Lab." *Working Paper*.

24. **Flatters, Frank and MacLeod** 1995. "Administrative Corruption and Taxation" *International Tax and Public Finance, 2*: 397-417.

25. **Friedman, Milton, and Leonard J. Savage.** 1948. "The utility analysis of choices involving risk." *The Journal of Political Economy 56* (4): 279-304.

26. **Gordon, Roger, and Wei Li.** 2009. "Tax structures in developing countries: Many puzzles and a possible explanation." *Journal of Public Economics, 93*(7): 855-866.

27. **Guillén, Mauro F. and Adrian E. Tschoegl.** 2002. "Banking and Gambling: Banks and Lottery-Linked Deposit Accounts." *Journal of Financial Services Research, 21* (3) : 219 - 231.

28. **Guryan, Jonathan and Melissa S. Kearney.** 2008. "Gambling at Lucky Stores: Evidence from State Lottery Sales." *American Economic Review, 98* (1): 458-73.

29. **Hilger, Nathaniel G.** 2014."How does family income affect college enrollment? Evidence from timing of parental layoffs", *Working Paper, Brown University*.

30. **Hoopes, Jeffrey, Daniel Reck and Joel Slemrod** (2014) "Taxpayer Search for Information: Implications for Rational Attention" *American Economic Journal: Economic Policy*, forthcoming

31. **International Monetary Fund**. 2011. "Revenue Mobilization in Developing Countries." *Fiscal Affairs Department*.

32. **Jacobson, Louis S., Robert J. LaLonde and Daniel G. Sullivan.** 1993. "Earnings Losses of Displaced Workers," *The American Economic Review, 83* (4): 685-709.

33. **Jornal Estado de São Paulo.** September 8th 2007. "Sonegação de impostos equivale a 30% do PIB." http://www.estadao.com.br/noticias/economia,sonegacao-de-impostos-equivale-a-30-do-pib,48525,0.htm

34. **Kaustia, Markku and Samuli Knüpfer.** 2008. "Do Investors Overweight Personal Experience? Evidence from IPO Subscriptions." *The Journal of Finance, 63* (6): 2679-2702.

35. **Kearney, Melissa S., Peter Tufano, Jonathan Guryan and Erik Hurst.** 2011."Making Savers Winners: An Overview of Prize-Linked Savings Products." In Olivia S. Mitchell and Annamaria Lusardi, eds., *Financial Literacy: Implications for Retirement Security and the Financial Marketplace*. Oxford University Press.

36. **Keen, Michael and Jack Mintz.** 2004."The optimal threshold for a value-added tax." *Journal of Public Economics, 88* : 559– 576.

37. **Keen, M., and B. Lockwood.** 2010 "The Value Added Tax: Its Causes and Consequences," *Journal of Development Economics, 92*(2): 138 151.

38. **Kleven, Henrik J.** 2014. "How Can Scandinavians Tax So Much?" *Journal of Economic Perspectives, 28*(4): 77-98.

39. **Kleven, Henrik J., Martin B. Knudsen, Claus T. Kreiner, Søren Pedersen, and Emmanuel Saez.** 2011. "Unwilling or Unable to Cheat? Evidence from a Randomized Tax Audit Experiment in Denmark."*Econometrica, 79* (3): 651-692.

40. **Kleven, Henrik J., Claus T. Kreiner, and Emmanuel Saez.** 2015. "Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries."*Working Paper.*

41. **Kopczuk, Wojciech, and Joel Slemrod.** 2006. "Putting firms into optimal tax theory." *The American Economic Review, 96* (2): 130-134.

42. **Kumler, Todd, Eric Verhoogen and Judith Frias**. 2013. "Enlisting Workers in Monitoring Firms: Payroll Tax Compliance in Mexico." *NBER Working Paper 19385.*

43. **La Porta, Rafael and Andrei Shleifer.** 2008. "The Unofficial Economy and Economic Development." *Brookings Papers on Economic Activity*, 275-352.

44. **La Porta, Rafael, and Andrei Shleifer.** 2014. "Informality and Development. " *Journal of Economic Perspectives* 28, no. 3: 109-126.

45. **Luque, Carlos A., Eduardo A. Haddad, Gilberto T. Lima, Sergio N. Sakurai and Silvio M. Costa.** 2011. "Impact assessment of interregional government transfers in Brazil: an input-output approach." *MPRA Paper 31292*.

46. **Marchese, Carla.** 2009. "Rewarding the Consumer for Curbing the Evasion of Commodity Taxes?" *FinanzArchiv: Public Finance Analysis, 65* (4): 383-402.

47. **Mattos, Enlinson, Fabiana Rocha and Patricia Toporcov.** 2013. "Programas de Incentivos Fiscais São Eficazes? Evidência a Partir da Avaliação do Impacto do Programa Nota Fiscal Paulista Sobre a Arrecadação de ICMS." *Revista Brasileira de Econometria, 67* (1): 97 - 120.

48. **Monteiro, Joana and Juliano Assunção**. 2012. "Coming out of the shadows? Estimating the impact of bureaucracy simplification and tax cut on formality in Brazilian microenterprises" *Journal of Development Economics, 99* (1):105–115.

49. **Musgrave, Richard A.** 1969. *Fiscal systems*. New Haven: Yale university press.

50. **Pomeranz, Dina.** 2014. "No taxation without information". *Harvard Business School Working Paper 13-057*.

51. **Shleifer, Andrei and Robert W. Vishny**. 1993. "Corruption." *The Quarterly Journal of Economics, 108* (3): 599-617.

52. **Slemrod, Joel.** 2007. "Cheating ourselves: the economics of tax evasion." *The Journal of Economic Perspectives, 21*(1): 25-48.

53. **Slemrod, Joel., Brett Collins, Jeffrey Hoopes, Daniel Reck, and Michael Sebastiani**. 2014. "Does Credit-Card Information Reporting Improve Small-Business Tax Compliance?" *Working paper*.

54. **Slemrod, Joel and S. Yitzhaki.** 2002. "Tax Avoidance, Evasion and Administration", in Auerbach, Alan J., and Martin S. Feldstein, eds. *Handbook of Public Economics, 4*: 1423-1470.

55. **Stephens-Davidowitz, Seth.** 2014. "The cost of racial animus on black presidential candidate: Using google search data to find what surveys miss." *Journal of Public Economics*, Volume 118, Pages 26–40

56. **Tirole, Jean.** 1986. "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations." *Journal of Law, Economics, & Organization, 2* (2): 181-214.

57. **Tufano, Peter.** 2008. "Saving whilst Gambling: An Empirical Analysis of UK Premium Bonds." *American Economic Review: Papers & Proceedings, 98* (2): 321–326.

58. **Yagan, Danny.** 2013. "Capital Tax Reform and the Real Economy: the Effects of the 2003 Dividends Tax Cut." *Working paper*.

59. **Yaniv, Gideon.** 1993. "Collaborated Employee-Employer Tax Evasion." *Public Finance/Finances Publiques, 47*: 312-321.

60. **Wan, Jummin.** 2010. "The Incentive to Declare Taxes and Tax Revenue: The Lottery Receipt Experiment in China." *Review of Development Economics, 14* (3), 611-624.

Table 1: Descriptive Statistics

a: Establishments

	Number of Obs.	Mean	Std. Dev.	Time Period
Establishment sample				
Retail establishments				
Reported revenue	66,382,628	13,916.99	83,549.03	Jan.2004-Dec.2011
Number of receipts	8,083,008	648.47	5261.08	Jan.2009-Dec.2011
Number of consumers - SSN receipts	8,083,008	84.65	1736.94	Jan.2009-Dec.2011
Share of receipts with SSN	8,083,008	0.44	0.41	Jan.2009-Dec.2011
Wholesale establishments				
Reported revenue	6,888,041	67,169.69	196,069	Jan.2004-Dec.2011
Employment sample				
Number of employees				
Retail - Sao Paulo	351	7.3	8.2	2004 - 2011
Retail - other Brazilian States	9,117	6.7	6.4	2004 - 2011

b. Consumers

	Number of Obs.	Mean	Std. Dev.	Time Period
Consumers sample				
Number of receipts	46,505,268	4.5	7.6	Jan.2009 - Dec.2011
Number of businesses	46,505,268	3.0	4.2	Jan.2009 - Dec.2011
Tax Rebate	46,505,268	2.9	7.7	Jan.2009 - Dec.2011
Total expenditure in SSN receipts	46,505,268	356.4	1,346.4	Jan.2009 - Dec.2011
Lottery sample				
Number of receipts	37,237,148	10.4	9.4	Jun. 2010 - May.2011
Number of businesses	37,237,148	6.1	5.0	Jun. 2010 - May.2011
Tax Rebate	37,237,148	5.6	10.1	Jun. 2010 - May.2011
Total expenditure in SSN receipts	37,237,148	653.6	1,688.1	Jun. 2010 - May.2011
Number of lottery tickets	37,237,148	4.0	26.3	Jun. 2010 - May.2011
Lottery prize value	37,237,148	2.2	10.9	Jun. 2010 - May.2011

Notes: Tables present the number of observations, means, standard deviations, and time periods of the key variables for each sample. All values are in US dollars (US$1=R$2). Table 1a describes the establishment, sector and employment samples. Reported revenue is the gross reported revenue by establishment. *Number of receipts* is the total number of receipts an establishment reports to the tax authority. *Number of consumers - SSN receipts* is the total number of different Social Security Numbers (SSNs) to which an establishment issues a receipt. *Revenue from SSN receipts* is a sum of the total value of SSN receipts an establishment issues. *Share of receipts with SSN* is the total count of reported receipts with SSN over the total number of reported receipts by establishment. The sector sample aggregates revenue reported by all establishments by 7-digit sectors as described in section II.C.2 in the text. The employment sample aggregates the annual employer-employee data by 5-digit sectors. The table displays the average number of formal employees by establishment in retail sectors registered in Sao Paulo and in the other 26 states in Brazil. Table 1b describes the *consumer sample* and the *lottery sample* (see section II.C.4). *Number of receipts: the total number of SSN receipts for which a consumer asks per month*; *number of establishments*: the number of different establishments for which a consumer asks for SSN receipts per month; *total expenditures with a SSN*: the total amount of money spent associated with the SSN receipts; *average receipt value*: the average value among all purchases represented by consumer's SSN receipts in a given month. N*umber of lottery tickets*: the total number of lottery tickets a consumer holds per month; *lottery prizes*: the number of lottery prizes and the value of lottery prizes per month.

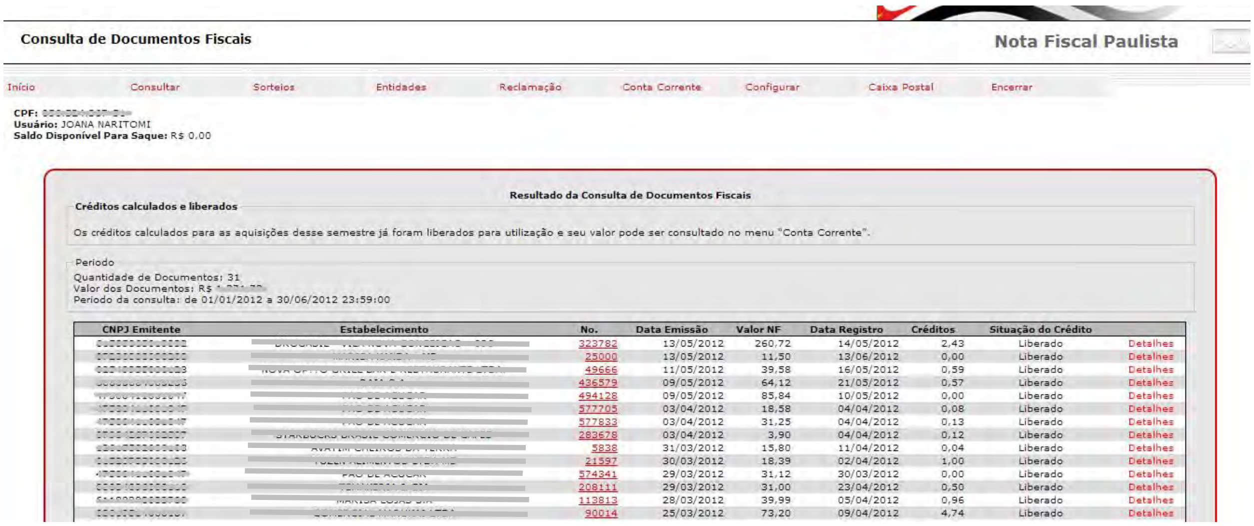
Table 2: Compliance Effect – Retail vs. Wholesale

	log Reported revenue			
	(1)	(2)	(3)	(4)
DD (Post Oct 07 * retail)	0.253***	0.244***		
	[0.0629]	[0.0664]		
DD * High volume of consumers			0.484***	
			[0.0614]	
DD * Low volume of consumers			0.182**	
			[0.0701]	
DD * Small establishments				0.526***
				[0.0909]
DD * Large establishments				0.0722
				[0.0700]
3rd-order polynomial of establishment size * DD			X	
Dummy for high volume of transactions*DD				X
Time FE	X		X	X
Establishment FE	X		X	X
R-squared	0.78	0.79	0.78	0.78
Number of observations	1,080,676	681,151	1,080,674	1,080,676

Notes: Table 2 displays the main coefficients form regressions described in section IV.B.1. Standard errors are clustered at the 7-digit sector classification level (210 clusters). Significance levels *** 1%, ** 5%. The variable DD is defined as the interaction between a dummy for retail sectors and a dummy that equals 1 for time periods after Oct 2007. This table displays the regression coefficient of a DD regression using the establishment-level data. The dependent variable is log of reported revenue by establishment, and the data is collapsed into two periods: before and after Oct. 2007. The regressions are dollar- weighted (each observation is weighted by its pre-NFP value) such that each observation contributes to all regression estimates according to its economic scale to best approximate the aggregate effect. Column (1) shows the average DD estimate controlling for time and establishment fixed effects. Column (2) runs the same regression in a sample that restricts attention to firms that have been active since Jan. 2004. Column (3) splits retail sectors into two groups: sectors below the median volume of transactions distribution across sectors (Low volume of receipts) and sectors above the median of volume of transactions (High volume of receipts). Volume of receipts is defined by ranking 7-digit retail sectors by the average number of transaction by establishment between 2009 and 2011. In order to control for establishment size effects the regressions in columns (2) and (3) include a 3rd order polynomial interacted with the DD variable. Column (4) splits the retail establishments in two groups: establishments below the median establishment size distribution (Small establishments) and establishments above the median (Large establishments), and flexibly control for differential effects from volume of consumers by interacting the DD variable with five dummies for quintiles of volume of consumers. Size is defined by the total reported revenue by establishments during a four-year period before program implementation.

Figure 1: Online Account Example

a. Checking all receipts issued with one's SSN



Consulta de Documentos Fiscais

Nota Fiscal Paulista

Início | Consultar | Sorteios | Entidades | Reclamação | Conta Corrente | Configurar | Caixa Postal | Encerrar

CPF: [illegible]
Usuário: JOANA NARITOMI
Saldo Disponível Para Saque: R$ 0,00

Resultado da Consulta de Documentos Fiscais

Créditos calculados e liberados

Os créditos calculados para as aquisições desse semestre já foram liberados para utilização e seu valor pode ser consultado no menu "Conta Corrente".

Periodo
Quantidade de Documentos: 31
Valor dos Documentos: R$ [illegible]
Periodo da consulta: de 01/01/2012 a 30/06/2012 23:59:00

CNPJ Emitente	Estabelecimento	No.	Data Emissão	Valor NF	Data Registro	Créditos	Situação do Crédito	
[illegible]	[illegible]	323782	13/05/2012	260,72	14/05/2012	2,43	Liberado	Detalhes
[illegible]	[illegible]	25000	13/05/2012	11,50	13/06/2012	0,00	Liberado	Detalhes
[illegible]	[illegible]	49666	11/05/2012	39,58	16/05/2012	0,59	Liberado	Detalhes
[illegible]	[illegible]	436579	09/05/2012	64,12	21/05/2012	0,57	Liberado	Detalhes
[illegible]	[illegible]	494128	09/05/2012	85,84	10/05/2012	0,00	Liberado	Detalhes
[illegible]	[illegible]	577705	03/04/2012	18,58	04/04/2012	0,08	Liberado	Detalhes
[illegible]	[illegible]	577833	03/04/2012	31,25	04/04/2012	0,13	Liberado	Detalhes
[illegible]	[illegible]	283678	03/04/2012	3,90	04/04/2012	0,12	Liberado	Detalhes
[illegible]	[illegible]	5838	31/03/2012	15,80	11/04/2012	0,04	Liberado	Detalhes
[illegible]	[illegible]	21597	30/03/2012	18,39	02/04/2012	1,00	Liberado	Detalhes
[illegible]	[illegible]	574341	29/03/2012	31,12	30/03/2012	0,00	Liberado	Detalhes
[illegible]	[illegible]	208111	29/03/2012	31,00	23/04/2012	0,50	Liberado	Detalhes
[illegible]	[illegible]	113813	28/03/2012	39,99	05/04/2012	0,96	Liberado	Detalhes
[illegible]	[illegible]	90014	25/03/2012	73,20	09/04/2012	4,74	Liberado	Detalhes

b.: *Nota Fiscal Paulista* Receipt

Secretaria da Fazenda

Cupom Fiscal

Nota Fiscal Paulista

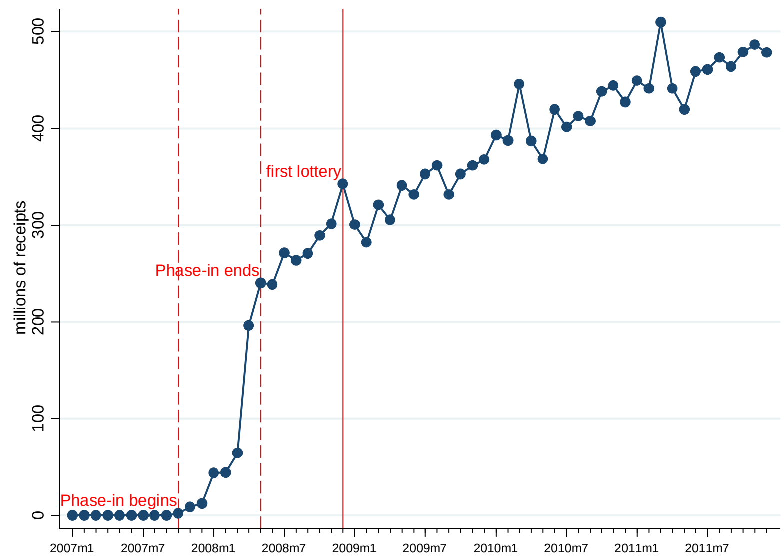
Início | Consultar | Sorteios | Entidades | Reclamação | Conta Corrente | Configurar | Caixa Postal | Encerrar

CPF: [illegible]
Usuário: JOANA NARITOMI
Saldo Disponível Para Saque: R$ 0,00


```
CNPJ:
  IE:
  IM:
30/3/2012 00:00:00 CCF: 20438                  COO: 21597
CNPJ/CPF consumidor:
Nome:
End:
                 CUPOM FISCAL
ITEM    CÓDIGO          DESCRIÇÃO
    QTD UN              VL UNIT R$          ST    VL ITEM R$
001     0000000000001                   ALMOCO
    1,00 x                  18,39                     18,39
DESCONTO R$                                        R$ 0,00
ACRESCIMO R$                                       R$ 0,00
TOTAL R$                                           R$ 18,39
ZPM ZPM/1FIT LOGGER ECF-IF VERSÃO:03.03.04 ECF: LJ: OPR:
2/9/2005 07:27:37
```

Notes: Figures 1a and 1b are snapshots of an online account example at https://www.nfp.fazenda.sp.gov.br/login. The snapshot is from the author's online account. Tabs on the top of both figures can be translated as: Home ("Início"), Check receipts ("Consultar"), Lotteries ("Sorteios"), Charities ("Entidades"), Complaints ("Reclamacao"), Current Account ("Conta Corrente"), Settings ("Configurar"), Inbox ("Caixa Postal"), Sign out ("Encerrar"). Figure 1a shows an example of an online account at the tax authority website under the tab "Check receipts": a list of all receipts, the issuing date, total value of each receipt, tax rebate, and a link to the details of each receipt. The tabs allow consumers to file complaints, verify whether they got a prize in a lottery, request deposits in a bank account, transfers to other enrolled consumers or transfers to charity. Figure 1b shows the receipt if one clicks on the last column in Figure 1a for details of one of the purchases listed ("detalhes"). The receipt is standardized across establishments where there is a field to fill in the consumers' SSN – as highlighted in the picture.

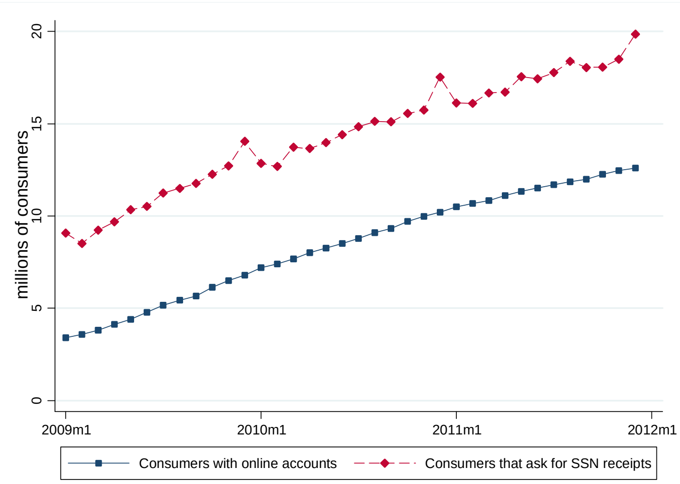
Figure 2: Timeline of Program Implementation



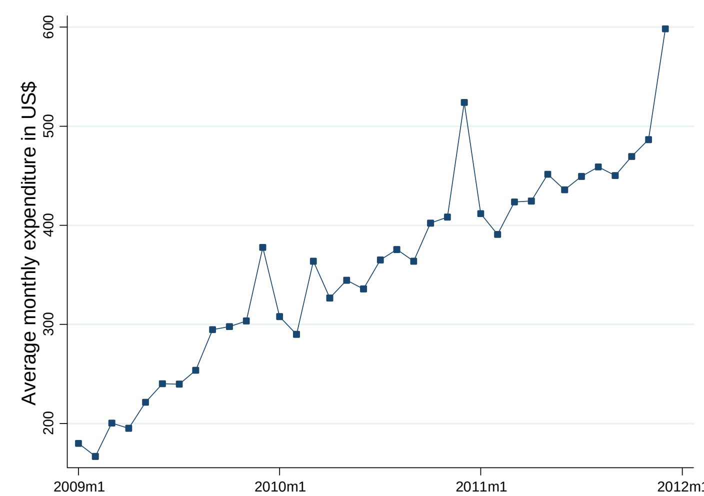
Notes: The figure shows the total number of receipts (millions of receipts) – with and without a SSN – electronically reported to the tax authority by month by establishments in Sao Paulo. The vertical lines highlight the key dates for the implementation of the NFP program. Between Oct.07 (Phase-in begins line) and May.08 (Phase-in ends line) 8 groups of sectors were phased-in in the policy of submitting receipts electronically to the tax authority. The possibility of inserting a SSN in the receipt for tax rebate purposes was introduced along with the electronic submission of receipts. The first lottery based on the purchases with SSN receipts was introduced in Dec.2008.

Figure 3: Consumer Participation

a. Number of consumers asking for SSN receipts and number of consumers with online accounts



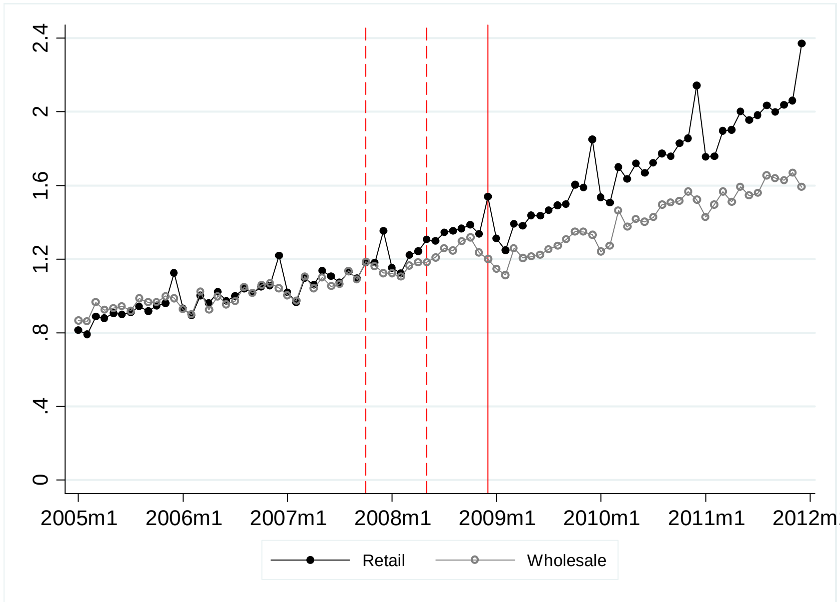
b: Average expenditure in SSN receipts per month by consumers with online accounts.



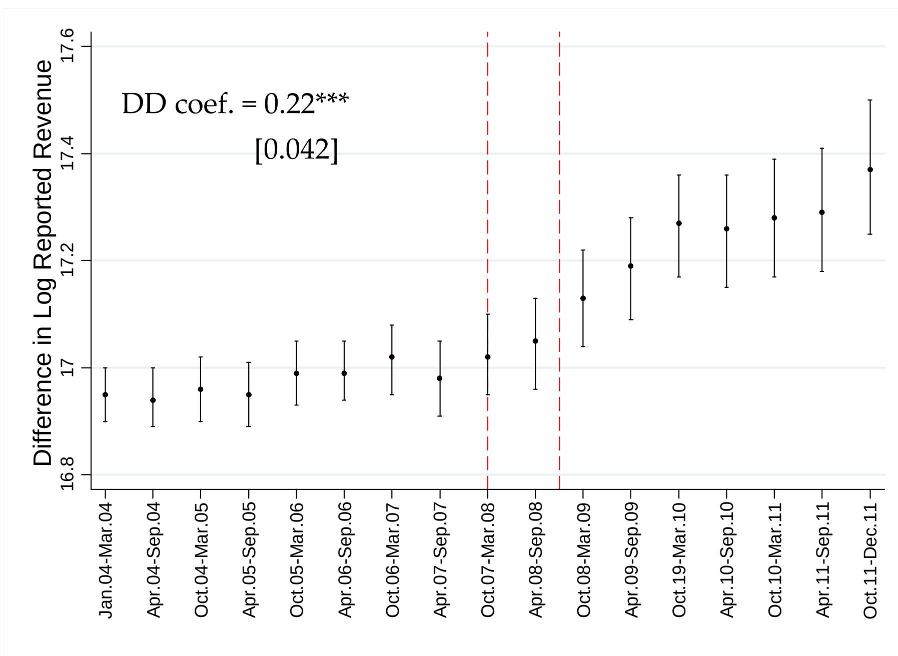
Notes: The dashed line in Figure 3a displays total number of consumers asking for SSN receipts each month, and the solid line is the total number of consumers that had set up online account at the tax authority's website between Jan. 2009 and Dec. 2011. Any person holding a Brazilian SSN is eligible to ask for receipts. In order to collect rewards and opt in for lotteries consumers need to enroll online. Figure 3b shows the average monthly expenditure by consumers with online accounts between Jan. 2009 and Dec. 2011. Monthly expenditure is the sum of the total value of SSN receipts by individual. The spikes of expenditure around December of each year follows the seasonal variation in consumption due to Christmas shopping.

Figure 4: Compliance Effect – Retail vs. Wholesale

a. Raw data: reported revenue changes

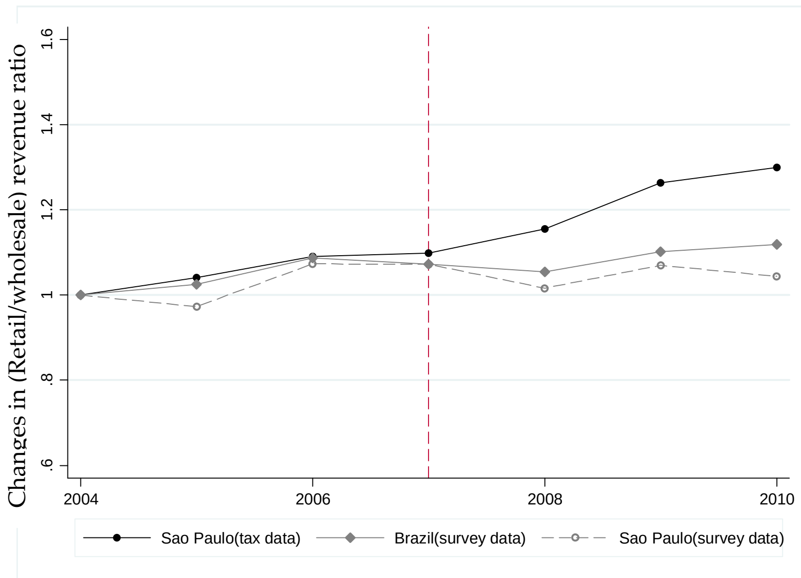


b. Difference coefficients for 6-month time bins - Log reported revenue by sector



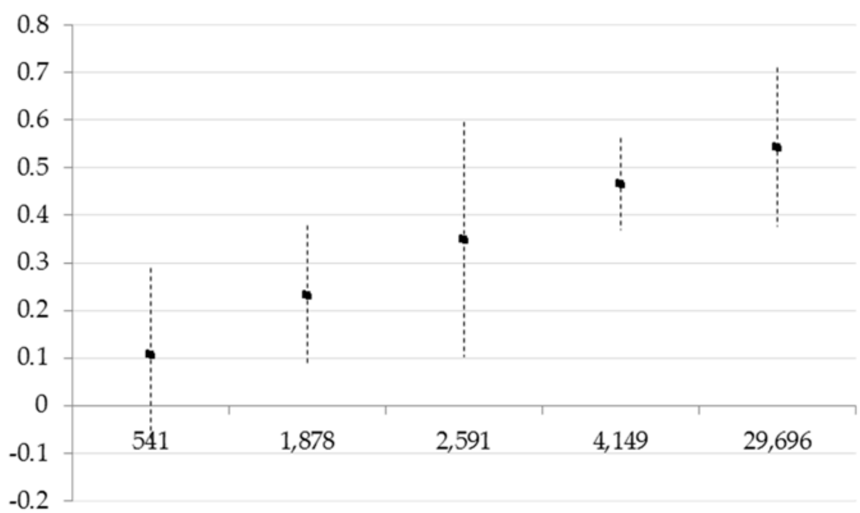
Notes: Figure 4a shows reported revenue changes for retail and wholesale sectors. Each line is defined by the reported revenue by all establishments aggregated by retail or wholesale scaled by the average monthly reported revenue before Oct. 07 for each sector group in constant prices. The figure plots the raw data, so there are spikes around December of each year follows the seasonal variation in consumption. The vertical lines highlight the key dates for the implementation of the NFP program: phase-in of sectors begins in Oct.07 and ends in May.08, and the first lottery based on the purchases with SSN receipts was introduced in Dec.2008. Figure 4b plots regression coefficients from estimating specification (1) using a sample of 210 sectors between Jan 2004 and Dec 2011. The sector sample has 24,990 observations. The difference in differences (DD) coefficient displayed in the figure is estimated using the specification (2) where the DD variable is defined by the interaction between a dummy for retail sectors and a dummy that equals 1 for time periods after Oct 2007. Standard errors are clustered by sector.

Figure 5: Reported Revenue Difference between Retail v. Wholesale



Notes: The figure shows changes in the retail-wholesale reported revenue ratio from the Sao Paulo tax data (black line), and changes in the retail-wholesale actual revenue ratio from a national-wide survey on the trade sector (gray lines). The dashed vertical line marks the beginning of the NFP program in 2007. The solid gray line displays the national-wide ratio, and the dashed gray line shows the retail-wholesale actual revenue ratio for the state of Sao Paulo from the survey data. Each line is scaled by the 2004 retail-wholesale revenue ratio. The national ratio is based on the total gross revenue from sales, and retail revenue considers retail and motor-vehicles trade.

Figure 6: Heterogeneous Compliance Effects – whistle-blower threats

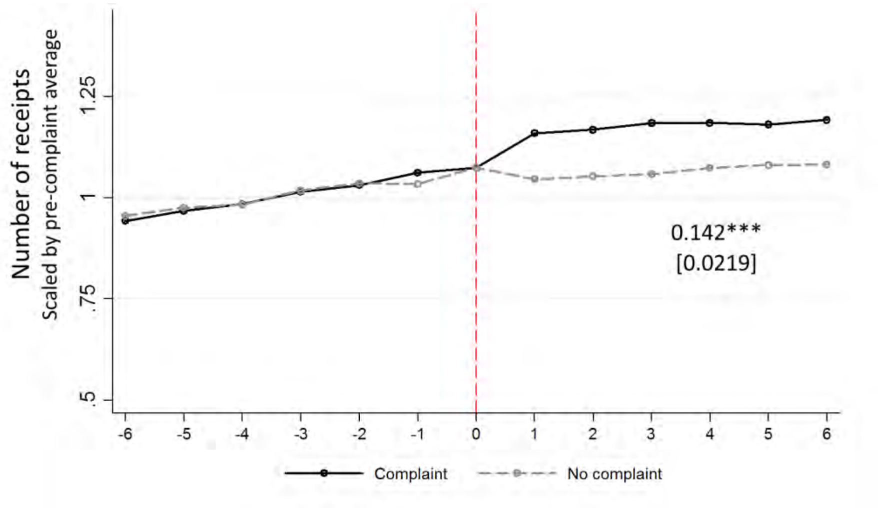


Quintiles of transaction volume (mean # of receipts in each bin)

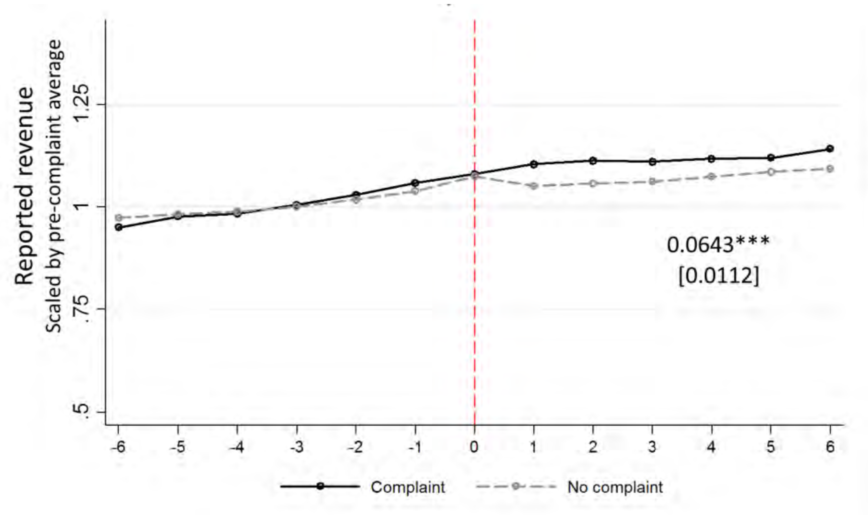
Notes: Figures 6 plots the coefficients and a 95% confidence interval from estimating equation (7) in section IV.B.1 The figure plots the effects of the program by quintiles of the volume of receipts distribution. Volume of receipts is defined by ranking retail sectors by the volume of transaction as described in Section IV.d. The x-axis displays the average number of receipts across sectors in each bin. In both graphs standard errors are clustered by 7-digit sector classification (210 clusters). The sample is composed by monthly observations between Jan.2004 and Dec.2011 of establishments classified as either retail or wholesale, as described in section II. The dependent variable is log reported. In order to control for establishment size effects the regression includes a 3rd order polynomial interacted with the DD variable.

Figure 7: Whistle-blower Effect on Firm Compliance

a. Changes in the number of receipts issued



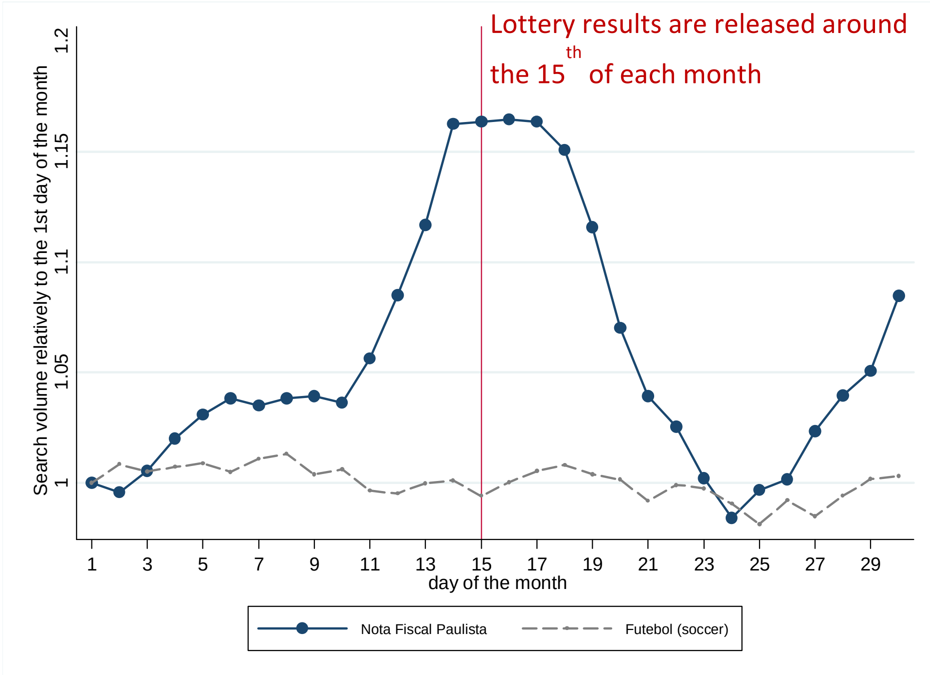
b. Changes in reported revenue



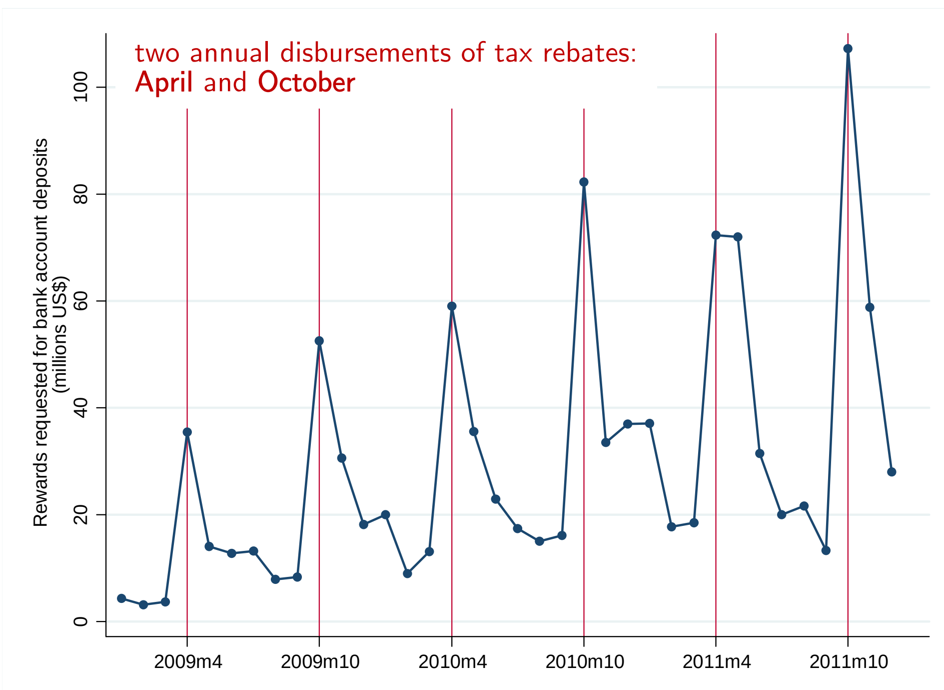
Notes: Figures 7a and 7b plot the changes in the total number of receipts a firm reports and the changes in revenue reported to the government after the firm receives the first complaint. Consumers can file complaints about specific firms every month. Both graphs display changes across event-time where each data point is scaled by the outcome's average before the first complaint (event-time zero). The 'Complaints' group is composed by firms that received their first complaint at event-time zero. The 'No complaint' group is composed by firms that did not receive their first complaint at time zero, and firms that did not receive a complaint until Dec. 2011. The outcome is averaged across groups and event times using weights based on quartiles of the propensity score to get the first complaint in a given time period. The propensity score is estimated using time specific trends for each sector, age of the firm, number of establishments by firm, dummy for establishment-headquarter, dummies for legal nature of the firm, sector and time fixed effects, dummy for location in the metropolitan region of Sao Paulo, and the three lags of third-order polynomials of reported revenue, reported receipts, SSN receipts and number of consumers. For more details see Appendix. The estimated DD coefficient displayed in each graph is based on estimating specification (8) using the weighted averaged data by group 6 months before and after the first complaint.

Figure 8: Are Consumers Paying Attention to the Rewards Schedule?

a: Timing of lottery results - Google searches for *Nota Fiscal Paulista*.

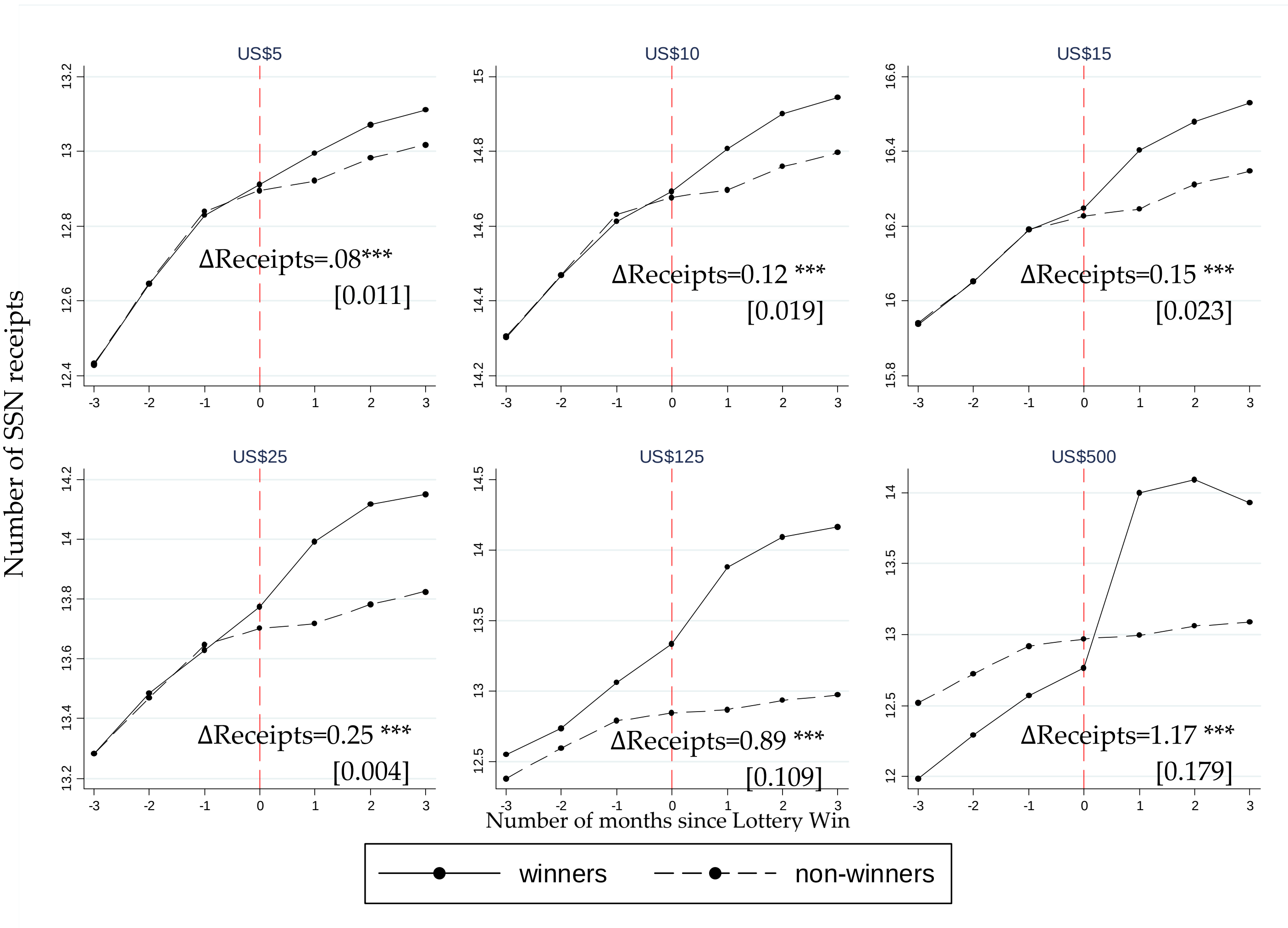


b: Timing of tax rebate disbursements - Rewards requested for deposit in consumers' bank accounts



Notes: Figure 8a displays the search volume from Google Trends website for Google searches with terms related to "nfp" or "nota fiscal paulista" or "nota paulista" pooled by day of the month from IPs addresses in the state of Sao Paulo between Oct. 2007 and Dec. 2011. It also displays searches for "futebol" (soccer in Portuguese) pooled by day of the month from IPs addresses in the state of Sao Paulo for the same time period. The lottery results are released around the 15th of each month marked by the solid vertical line. In Figure 8b each data point is the total amount in millions of US$ requested for direct deposit in consumer's bank accounts. As described in section II.B, the tax authority does a biannual disbursement of the tax rebates: every April and October. The disbursement dates are marked by vertical lines in the x-axis.

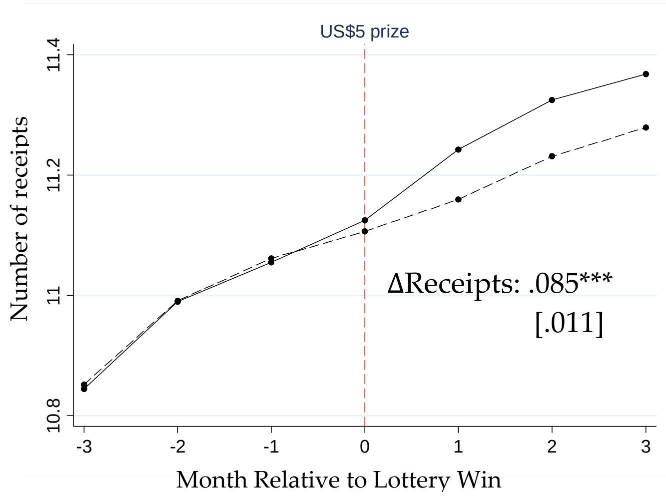
Figure 9: The Effect of Different Sizes of Lottery Wins on the Number of Receipts



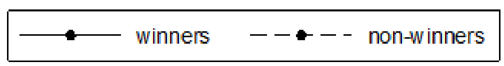
Notes: The figure shows the average number of receipts a consumer asks for per month if she is in the lottery winners group vs. the non-winners group. The x-axis is the number of months since the individual won or did not win a lottery. The y-axis in each graph is the average number of SSN receipts consumers ask for in a month. Before taking the averages I create bins for each possible number of lottery ticket holdings from 1-40 tickets in each monthly lottery for 12 lotteries between June 2010 and May 2011. Then I re-weight the non-winners group such that each bin carries the same relative weight as the winner group distribution across lottery ticket holdings. The graphs were constructed from the lottery sample described in section III.D and Appendix B. In each of the lotteries there are 1,407,394 prizes of U.S. $5, 76,303 prizes of U.S. $10, 15,000 prizes of U.S. $25, 1,000 prizes of U.S. $125, and 300 prizes of U.S. $500. Because it is common for individuals to hold more than one lottery ticket in a month, there are many cases of consumers that get a total of US$15 by winning a combination of a U.S. $5 and a U.S. $10 prizes. The estimated DD coefficient displayed in each graph is based on estimating specification (11) using the weighted averaged data by group 6 months before and after the lottery.

Figure 10: The Effect of Lottery Wins on Consumer Participation

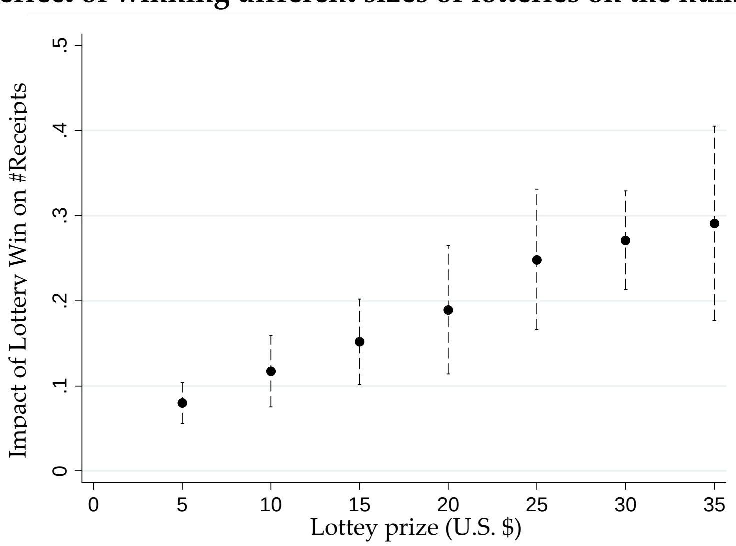
a. Impact of winning US$5 on the number of receipts among consumers that won a prize before



b. Impact of winning US$5 on the number of different establishments in which consumers ask for receipts

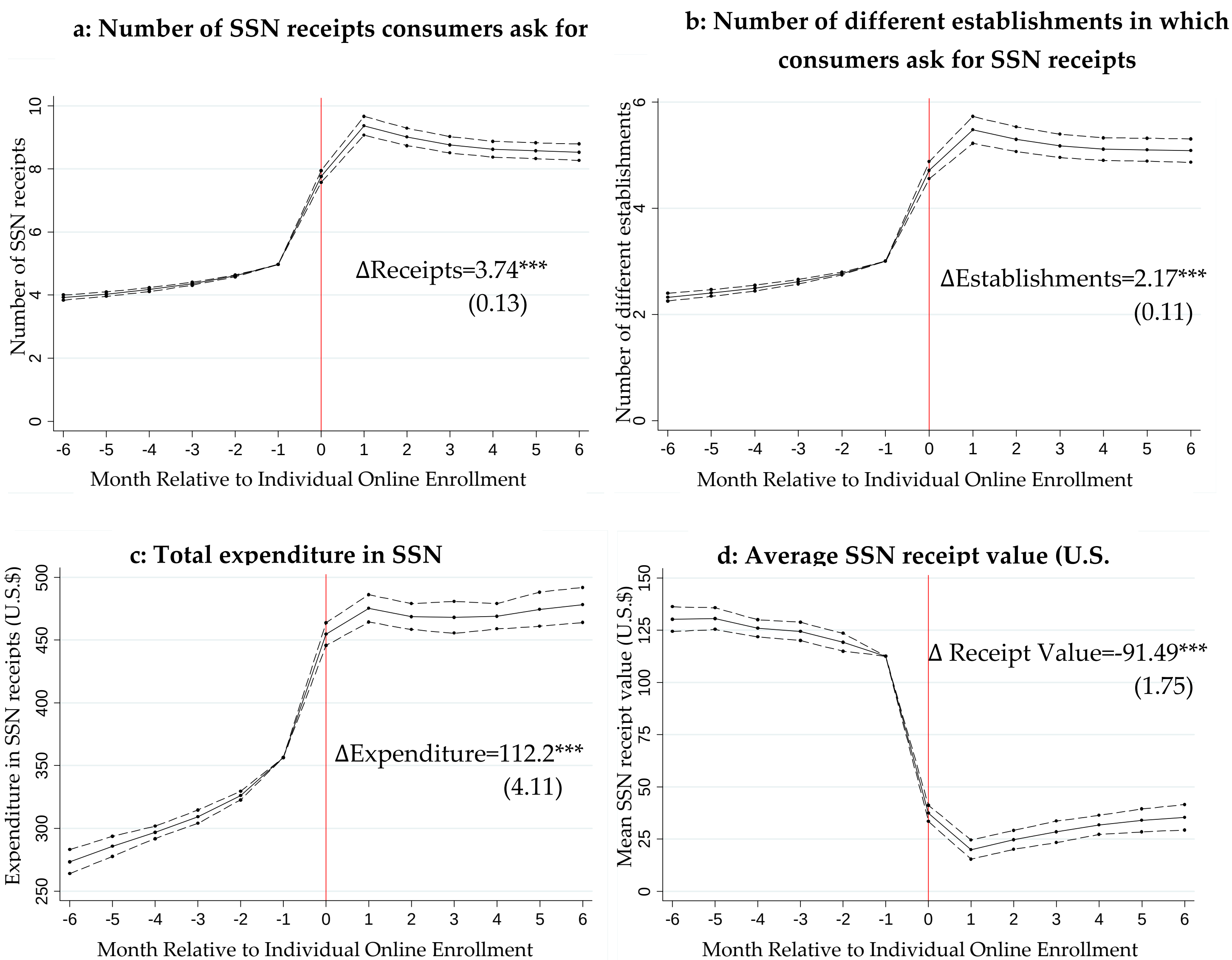


c. The estimated effect of winning different sizes of lotteries on the number of SSN receipts



Notes: Figure 10a displays the effect of winning a U.S. $5 dollar prize on the average number of different businesses consumers ask for SSN receipts in a month. The DD coefficient displayed in the graph is estimated using specification (11) using the weighted averaged data by group and by 6 months before and after the lottery. Figure 10b shows the effect of a U.S. $5 lottery win for consumers that have won a lottery once before. Both graphs are constructed in the same way as the graphs in Figure 9. The x-axis in Figure 10a and Figure 10b is the number of months since the individual won or did not win a lottery. Because it is common for individuals to hold more than one lottery ticket in a month, there are many cases of consumers that get a total of U.S. $15, U.S. $20, U.S. $30, U.S. $35, by winning more than one prize. Figure 10c plots the DD coefficients and a 95% confidence intervals from estimating specification (11) using the weighted averaged data by group 6 months before and after the lottery for each prize level.

Figure 11: Lottery Eligibility and Consumer Participation – The Effect of Online Enrollment

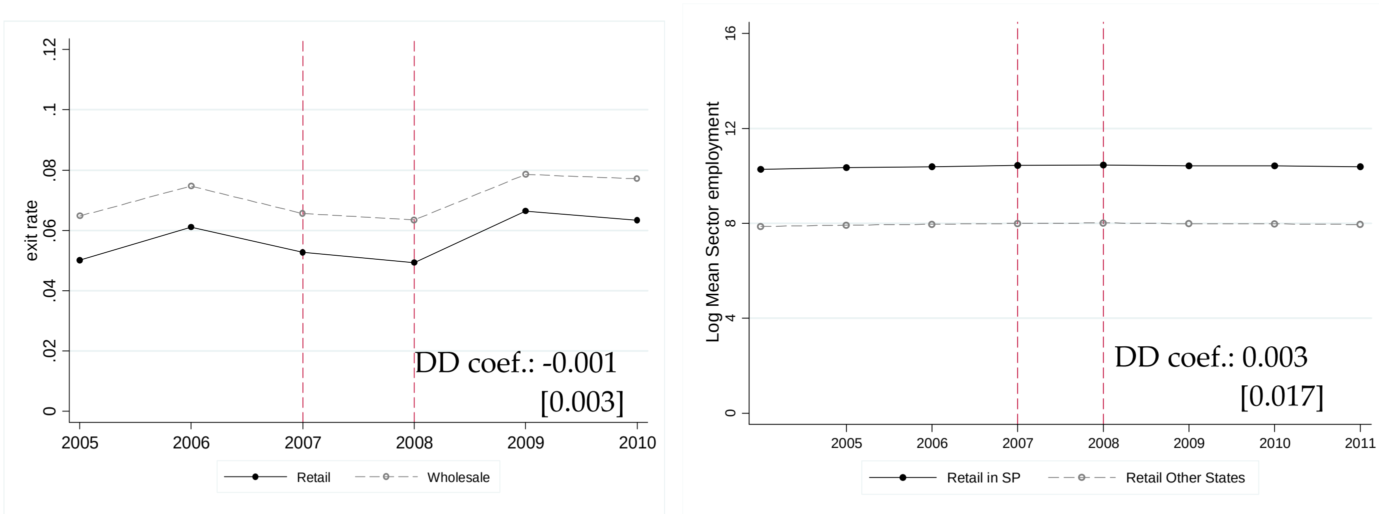


Notes: The graphs plot coefficients and 95% confidence intervals from estimating equation (12) in a panel of a 10% random sample of consumers that were registered online by the end of 2011 - around 1.3 million people -, and 46,505,268 observations between Jan 2009 and Dec 2011. *Number of receipts*: the total number of SSN-identified receipts for which a consumer asks per month; *number of establishments*: the number of different establishments for which a consumer asks for SSN-identified receipts per month; *total expenditures with a SSN*: the total amount of money spent associated with the SSN-identified receipts, aggregated by consumer, per month; *average receipt value*: the average value among all purchases represented by consumer's SSN-identified receipts in a given month. In order to reduce the influence of outliers I winsorize the number of receipts and total expenditure with a SSN by their 99th percentile value.

Figure 12: The Impact on Exit and Formal Employment

a. Exit rate: retail vs. wholesale in Sao Paulo

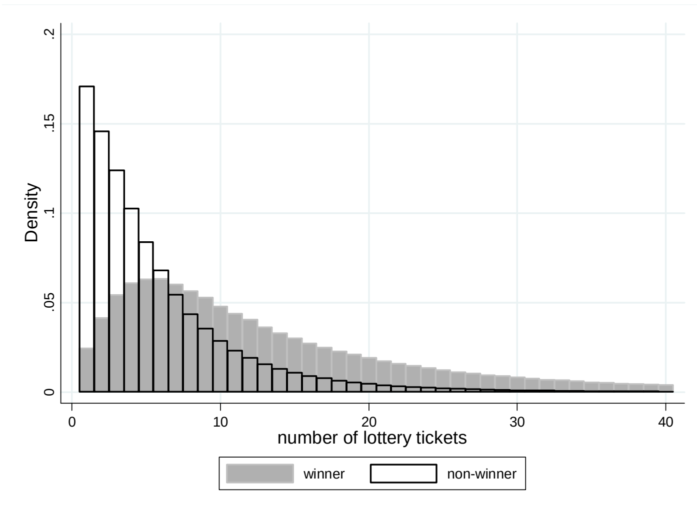
b. Formal employment changes: retail in Sao Paulo vs. retail in other Brazilian states



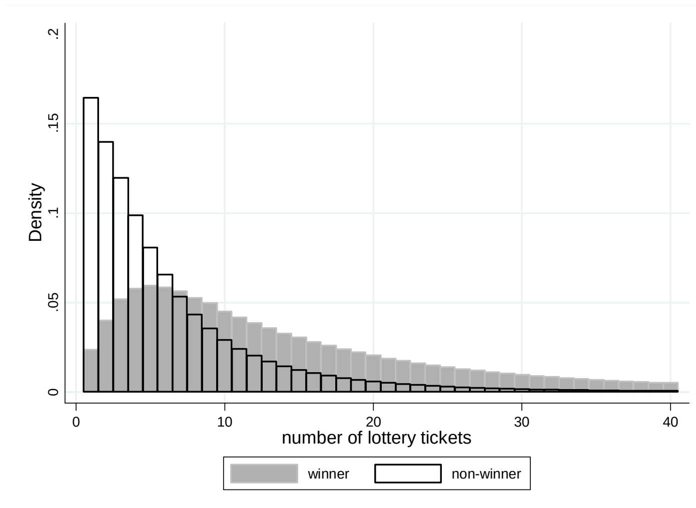
Notes: Figure 12a displays exit rates in retail and wholesale. Exit rate is defined as the total number of exits in year *t* and sector *s*, divided by the total number of establishments in year *t -1* in sector *s*. Exit is the year I observe an establishment in the data, and I consider all exits between 2005 and 2010. The data for the exit analysis comprises all establishments in retail and wholesale. Figure 12a also displays the DD coefficient from estimating a specification similar to equation (2) in a 7-digit sector yearly panel with 1,260 obs., and using exit rates as the dependent variable. Standard errors are clustered at the 7-digit sector level. Figure 12b uses a different data source: a nationwide annual administrative formal employer-employee dataset that allows a within retail comparison. The employment sample aggregates the employer-employee data by 5-digit sectors. Figure 12b displays changes in the log employment in retail in Sao Paulo and retail in other states. The figure also shows the DD coefficient from estimating a specification similar to equation (2) in a 5-digit sector yearly panel and using log employment as the dependent variable, and adding state fixed effects. The data has 9,392 observations and covers all years between 2004 and 2011 in 27 states. Standard errors are clustered at the 5-digit sector level.

Figure A: Histogram of Lottery Tickets – two examples
Distribution of lottery tickets between lottery winners and non-winners

i. December 2009 Lottery



ii. January 2011 Lottery



Notes: Figures A.i and A.ii shows examples of histogram for the number of lottery tickets winners and non-winners hold in two different lotteries: December 2009 and January 2011. The figures only consider lottery ticket holdings under 40. A lottery ticket is genereted for every 50 dollars a consumer spends in SSN receipts; so 50 receipts of 1 dollar or 1 receipt of 50 dollars are equivalent, and generate 1 lottery ticket. The two figures display a very similar pattern: there is common support between the two groups for lottery ticket holdings below 40, and the winner group holds more lottery tickets than the non-winner group. In order to make the two groups comparable in the empirical analysis in section IV.E.1, I re-weight the non-winners group such that each lottery holding carries the same weight as the winners group.

Appendix A: Data on Establishments and Consumers

Section III.C describes establishment-level and consumer-level variables and samples used in this paper. This appendix provides additional information on the datasets and variables, as well as further details on the re-weighting exercise from Sections IV.B.2 and IV.C.2.

A.1. Establishments

Establishment data. In order to construct the establishment sample I combine two different sources of data. Due to confidentiality reasons, each dataset used to construct the *establishment data* was de-identified, and a "fake" identifier was created for each establishment. The first data source are tax forms from establishments in the tax regime RPA ("Regime Periódico de Apuração") that requires establishments to report their gross revenue, tax credits and tax debits monthly through a form called GIA/ICMS ("Guia de Informação e Apuração do ICMS") to assess the total VAT due by the establishment in a given month. The second source of data is composed by tax forms from establishments in a simplified tax regime called SIMPLES. As is common in the VAT across the world (Keen and Mintz, 2004) - there is a threshold below which firms do not pay taxes over the value added. In the case of Brazil, firms that have yearly gross revenue of less than U.S. $1.2 million can choose to be in a simplified tax regime called SIMPLES in which firms pay taxes over gross revenue.

For the SIMPLES establishments I combined monthly data for establishments in Sao Paulo from three different sources: (i) tax returns from the state's *SIMPLES Paulista* in all months between 2004 and until June 2007; (ii) tax returns for the DASN-SP ("Declaração do Simples Nacional-SP") from July 2007 until the end of 2008; (iii) tax returns from DASN ("Declaração anual do Simples Nacional") between 2009 and 2011. The changes in data sources are due to the fact that there was a separate SIMPLES regime for federal and state taxes before June 2007. After that, states and federal government centralized in a single system all SIMPLES tax information, and there was a transition period in which states and the federal government kept separate records.[80]

Employment data. The employer-employee data RAIS/CAGED covers all formal establishments that have at least one employee. All formal firms must report to the Department of Labor their employment information in a yearly basis. It comprises individual information of employees such as wages, hours, years of education, date of hiring, date of firing, and type of contract. As I discussed in section III.C, I use a version of this data that aggregates the total number of employees by 5-digits sector definition.[81]

A.2. Consumers

Receipt data. The receipt data is constructed from a dataset that has transactions with SSN-identified receipts between January 2009 and December 2011. The transaction level data is a linked establishment-consumer data and has over 2.7 billion observations. The data was de-identified, and a "fake" identifier was created for each establishment and consumer. The datasets between

[80] The datasets listed in (i) - (iii) have some months of overlap, which allowed me to cross-check the information available in each of them, and verify that these changes did not generate mechanical changes in revenue reporting.

[81] The sector of activity listed in RAIS/CAGED has 5-digits until 2008 (CNAE 2.0), and it has 7-digits from 2009 onward (CNAE 2.1). In order to construct a time series from 2004 to 2011 by sector, I used the CNAE 2.0. definition of sector with 5-digits. The change from 7-digits to 5-digits do not affect the retail and wholesale classification for most of the sectors. Only in the case of motor vehicle trade it is not possible to distinguish retail and wholesale with a 5-digit definition. In the results reported in Figure 12b, I include motor vehicles trade in retail. The results do not change if I exclude these sectors.

October 2007 and December 2008 were not available to this study. The available data restricts attention to final consumers SSN ("CPF" holders), i.e., I do not have information on receipts issued with the SSN of other establishments or charities. Also, the data on approximately 90 consumers who won one of the top 3 lottery prizes of over U.S. $500 dollars in each monthly lottery between January 2009 and December 2011 were excluded from the dataset for confidentiality reasons.

Consumer sample. I take a 10% random sample of consumers who enrolled online by the end of 2011 - around 1.3 million people- and I construct a balanced monthly panel of consumer's participation in the program from the *receipts data* containing 46,505,268 observations between January 2009 and December 2011. In the case of three variables I describe in Section III.C - *number of receipts*, *number of establishments, total expenditures with a SSN* - for each individual I replace with zero the missing data in time periods in which no receipt with her SSN is reported by an establishment. The variable *average receipt value* is conditional on at least one receipt being issued with the consumer's SSN.

Appendix B: Complaints and Lotteries re-weighting Consumers

B.1. Complaints and re-weighting

Complaints data. For each firm I identify the time of the first complaint they ever received. In order to perform the empirical exercises on the effects of complaints on firms reported revenue in Section IV.B.2 I merge the *Establishment sample* data with the *receipts data* described in section III.C. The combined dataset of reported revenue and reported receipts covers the time period between Jan. 2009 and Dec. 2011. For the event-studies, I consider all first complaints between June 2009 and May 2011, i.e., at least 6 months before and after the earliest and latest first complaint respectively. The complaints sample covers firms in the retail sector that issued at least one receipt before June 2009. 25% of firms received at least one complaint in the period of analysis.

Re-weighting.I use a propensity re-weighting method to flexibly control for the probability of getting a complaint such that I use a quasi-random component of the timing of the first complaint by matching groups that have similar odds of getting a complaint. I estimate a propensity-score of a firm receiving the first complaint for every month-year between June 2009 and May 2011 based on pre-event characteristics.[82]. Then I use quartiles of the propensity score to re-weight establishments that did not received their first complaint in the given month-year to compare with establishments that received their first complaint in that month-year.

I perform this re-weighting exercise separately for each period between June 2010 and May 2011. For each case, I restrict attention to the sectors that had at least one firm that received a complaint in a given date and I draw a 10 percent random sample of firms that did not receive the first complaint on that date to build the no-complaint sample. This sample includes both establishments that did not receive their first complaint in a given date and establishments that did not receive any complaint by Dec. 2011. The propensity score is estimated using a logit model on time specific trends for each sector, age of the firm, number of establishments by firm, dummy for establishment-headquarter, dummies for legal nature of the firm, sector and time fixed effects, dummy for location in the metropolitan region of Sao Paulo, and the three lags of third-order polynomials of reported revenue, total number of receipts issued, total number of SSN receipts issued and total number of consumers.

Once I create the propensity score, I construct a dataset for each month within this period, where I keep all firm that received their first complaint that month and all firms that did not

[82]I follow a similar method to Hilger (2014) that uses propensity-score re-weighting in the context of parental job loss effects on children's long-term outcomes.

receive their first complaint that month and I re-weight the no-complaints group to match the complaints group within each quartile of the propensity score distribution. I collapse each cohort of complaints by each group and "event-month" using the weights.

B.2. Lottery wins and re-weighting

Lottery data. The lottery sample covers consumers that hold fewer than 40 lottery tickets in a given month for 12 lotteries between June 2010 and May 2011. In order to perform the empirical exercises on the effects of lottery wins on consumer participation in Section IV.C.2 I merge this data with the *receipts data* described in section III.C. The combined dataset of lotteries and receipts covers the time period between January 2010 and November 2011, i.e., 6 months before and after the first and last lottery. As in the *consumer sample*, I balance the panel of consumer participation and replace missing values by zero for the two key variables I use *number of receipts* and *number of establishments.*

Re-weighting. Since the number of lottery tickets is determined by the total value of a consumer' purchase 4 months before the lottery draw, the more a consumer participates in the program by asking for receipts, the higher are the odds she will get a prize in a given lottery. Therefore, it is important to carefully control for the odds of winning a prize in order to study the effect of lottery wins on consumer participation. As I describe in Section IV.C.2, I use a re-weighting method based on DiNardo, Fortin, and Lemieux (1996) to flexibly control for the number of lottery tickets individuals hold to ensure I use the random component of the lottery by matching the two groups based on the odds of winning prize.

Figure A shows two examples of the distribution of lottery ticket holdings among winners and non-winners in monthly lotteries. The two examples look very similar, and it is clear that the winner group typically holds more lottery tickets. I create bins for each possible number of lottery ticket holdings up to 40 tickets, which is the set of lottery tickets for which there is common support between the two groups. In the case of prizes that are only possible by winning a combination prizes - e.g., a U.S. $15 total prize is always a result of winning a US$5 prize and a US$10 prize -, I restrict attention to lottery ticket holdings between 2 and 40 tickets. I then re-weight the non-winners group such that each bin carries the same relative weight as the analogous bin in the winner group distribution across lottery ticket holdings.[83] Once I create the DFL weights I collapse the *lottery wins data* by each group and "event-month" using the weights for each prize level I display in Figures 11 and 12.

I perform this re-weighting exercise separately for each lottery win I study in Section IV.C.2. I construct a dataset for each prize level as described in *lottery wins data* above, where I keep all consumers that won a given prize (winners) and all consumers that do not win any prize (non-winners). When I compare the effect of different prize values, the pool of non-winners is the same in each lottery across the datasets I create for each prize level but they are re-weighted differently depending on the prize I am considering since the winners group of different prize levels may have slightly different distributions of lottery ticket holdings.

[83] See Appendix B of Yagan (2013) for a thorough description of DFL re-weighting.