

# CONSUMERS AS TAX AUDITORS

Joana Naritomi, Harvard University\*

November, 2013

JOB MARKET PAPER

## Abstract

This paper uses quasi-experimental evidence to shed light on whether consumers can help governments monitor firms and collect taxes. I construct administrative datasets on monthly tax returns for over 1 million establishments, 40 million people, and 2.7 billion receipts from an anti-tax evasion program in the state of Sao Paulo, Brazil – *Nota Fiscal Paulista* (NFP) – that gives monetary rewards to final consumers to ask for receipts. I exploit variation in treatment intensity across sectors to evaluate the effects of incentivizing consumers to monitor firms. I estimate that the NFP program increased the revenue reported in retail sectors by at least 23% over four years. I find no effects on exit rates or formal employment decisions as a result of the rise in enforcement. The estimated effect of the program is stronger for sectors with a high volume of transactions and small receipt values, consistent with a model in which there are fixed costs to negotiate collusive deals to avoid issuing receipts. Furthermore, the effect of the program has an inverted-U shape with respect to establishment size. This result is consistent with a model of higher baseline compliance among larger firms, and in which shifts in audit probability from consumer monitoring increase in firm size. Additionally, I exploit lottery rewards to show that consumers condition their participation on past lottery wins, and respond to increases in the expected rewards of the program. The evidence is consistent with the possibility that lotteries amplify consumer responses due to behavioral biases. I conclude by discussing the cost-effectiveness of the rewards program and implications for tax policy.

---

\*naritomi@fas.harvard.edu. An updated version can be found at <http://scholar.harvard.edu/jnaritomi/research>. I am very grateful to my advisors Raj Chetty, Alberto Alesina, Michael Kremer and Andrei Schleifer. I thank Ciro Biderman, Itzik Fadlon, François Gérard, Nathan Hendren, Nate Hilger, Martin Kanz, Asim Kwaja, Sendhil Mullainathan, Arash Nekoei, Dina Pomeranz, Francisco Queiro, Tristan Reed, Edson Severini, Ian Tomb, Matthew Weinzierl and Danny Yagan for invaluable comments and advice. I am extremely grateful to Adrea Calabi, André Luis Grotti Clemente, and Igor Baramboin, 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 enforcement is a central public finance issue in developing countries. Barriers to collecting tax revenue can severely constrain funding for basic public services and investments (Burgess and Stern, 1993; Besley and Persson, 2013), and pervasive tax evasion can cause significant distortions in the economy (Gordon and Li, 2009; La Porta and Shleifer, 2008). A government’s ability to collect revenue depends on two key components: information on taxable activities, and the enforcement capacity of the government.<sup>1</sup> For these reasons, the challenge of enforcing tax collection for transactions between retail firms and consumers is particularly acute: consumers have no incentives to collect receipts, and there is typically a large number of small tax-paying establishments to monitor.

In this context, enforcement strategies that incentivize both the creation of information on taxable transactions and compliance with tax laws in a self-enforcing manner can have an important role in overcoming governmental constraints. The Value Added Tax (VAT) was adopted instead of sales tax by most countries in the world due in part to its built-in enforcement incentives along the supply chain (Keen and Lockwood, 2010).<sup>2</sup> At the final consumer stage, however, these self-enforcing incentives break down since consumers typically derive no direct monetary benefit from asking for receipts.<sup>3</sup>

In order to address this enforcement problem, some countries have experimented with policies designed to change incentives at final sales through monetary rewards to consumers who collect receipts.<sup>4</sup> Nonetheless, the effectiveness of such a self-enforcing approach may be limited by the possibility of collusion between the buyer and the seller. Since the benefit offered by the government is typically lower than the tax collected, firms could offer a collusive deal to consumers and avoid issuing the receipt.<sup>5</sup> This paper uses evidence from an anti-tax evasion program in the state of Sao Paulo, Brazil – *Nota Fiscal Paulista* (NFP) – to study how consumers can help the government monitor firms and elicit relevant information on final sales.

The NFP program provides tax rebates and monthly lottery prizes for consumers who ask for

---

<sup>1</sup>Musgrave (1969) emphasized the relevance of tax administration for tax collection through history. In the current policy debate, tax administration and enforcement capacity of developing country governments are central issues (Slemrod, 2005; Bird and Gerdron, 2007; IMF, 2011), and a growing literature shows that information is crucial for tax enforcement (Gordon and Li, 2009; Kleven et al., 2011; Kumler et al., 2013; Pomeranz, 2013; Carrillo et al., 2013).

<sup>2</sup>Kopczuk and Slemrod (2006) argue that retail sales tax and the VAT are theoretically equivalent, but the VAT generates a paper trail from transactions across firms that can be cross-checked, and it has self-enforcing properties: a tax credit and debit system along the supply chain. Pomeranz (2013) shows that the self-enforcing mechanism of the VAT is empirically relevant.

<sup>3</sup>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.

<sup>4</sup>Brazil, China, Chile, Portugal, Puerto Rico, Bolivia, South Korea, Argentina, Italy, Slovakia, among other countries, have created monetary incentives – through tax refunds, lotteries, or fines – for consumers to collect receipts (Bird, 1992; Cowell, 2004; Marchese, 2009).

<sup>5</sup>As in the case of corruption with theft in Shleifer and Vishny (1993), for instance, a shop can offer a discount on a purchase conditional on not issuing a receipt. Similarly, Yaniv (1993) argues that employers and employees can find mutually beneficial opportunities to reduce their tax liabilities.

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 policy is designed to affect both the likelihood that a transaction is reported at all, and the accuracy of reporting, since rewards are an increasing function of the value of receipts.

In order to evaluate the extent to which consumer monitoring can overcome potential collusion opportunities, and improve compliance by firms, I construct establishment-level data on 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, and transaction-level data on over 2.7 billion receipts based on administrative records from the tax authority of the state of Sao Paulo.

I divide my analysis into three parts. First, I provide evidence that the program is salient, that it increased establishments' compliance, and that the higher effective tax rates had no impact on firm exit or employment. Second, I investigate potential mechanisms through which collusion may break down by examining heterogeneous increases in establishment compliance, and by analyzing consumer responses to lottery incentives. Third, I discuss the cost-effectiveness of the rewards program.

To study the effect of consumer monitoring on establishments' compliance I exploit 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 provide direct evidence in support of the identifying assumption of common pre-trends between the two groups for the difference-in-differences estimation: I show that retail and wholesale sectors behaved similarly before the program was introduced, and that the reported revenue in retail sectors increased relatively more than in wholesale sectors after the implementation of the NFP policy.

I estimate that reported revenue in retail increased on average by 23% 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.<sup>6</sup> I provide further supporting evidence that the change observed in reported revenue from the Sao Paulo tax data is a compliance effect by examining survey data on establishment revenue, which may more accurately capture actual revenue. The retail-wholesale revenue ratio in the survey does not change around the time NFP is implemented.

Since the rise in tax enforcement increases the effective tax rate, I study the impact of the consumer monitoring policy on exit and employment decisions of establishments. I compare exit rates of 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

---

<sup>6</sup>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.

Brazilian states. I find no effects of the program on those margins during the period of analysis.

To investigate potential mechanisms, I begin with a simple conceptual framework in which consumer monitoring may affect firms' evasion decisions through costs of collusion and an increased audit probability from consumers acting as whistle-blowers. I then show empirically that establishments in sectors that are characterized by a large number of transactions and that sell low-price items are relatively more affected by the NFP program, which suggests that there may be relevant fixed costs to negotiate a deal to avoid paying taxes. Furthermore, the effect of the program has an inverted-U shape with respect to establishment size, consistent with the model's predictions of higher baseline compliance among larger firms, and with shifts in audit probability from consumer monitoring that increase in firm size.

Next, I turn to the key mechanism of the policy: consumer participation. I investigate changes in the number of receipts for which individuals ask by exploiting variation from lottery prize rewards from NFP. I show 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 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.<sup>7</sup> The results are consistent with the possibility that the lotteries amplify consumer responses due to behavioral biases.

The evidence shows that rewards to consumers can increase compliance by establishments. However, the government is foregoing a fraction of both marginal and infra-marginal revenue through tax rebates and lotteries. From the estimates of the impact of NFP, I calculate an increase in tax revenue of approximately U.S. \$400 million net of consumers' rewards over four years. In a cost-benefit analysis I calculate participation costs and discuss other potential benefits from the consumer-rewards policy. I find that the program can be fairly cost-effective, and I consider potential ways to reduce costs.

This paper contributes to a growing empirical literature which shows that information - through third-party reporting or paper trail - is key for compliance, and that self-enforcing incentives can be effective for tax enforcement (Gordon and Li, 2009; Kleven et al., 2011; Kumler et al., 2013; Pomeranz, 2013; Carrillo et al., 2013).<sup>8</sup> More generally, the paper is related to a vast literature on tax evasion and enforcement (e.g., Andreoni et al., 1998; Slemrod and Yitzhaki, 2002).

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., 2012; Pomeranz, 2013) 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.<sup>9</sup> Even though the

---

<sup>7</sup> 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.

<sup>8</sup> In particular, the paper contributes to understanding how incentives for final consumers can help enforce tax collection. 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).

<sup>9</sup> See Bruhm and McKenzie (2013) for a review of the literature on formalization of firms.

informal sector is often large in developing countries, the degree to which the government can enforce taxation in already-registered firms is a key lever to make registration meaningful for tax revenue purposes. Further, better enforcement in a VAT system can endogenously generate incentives to formalize by creating supply chains of formal firms (De Paula and Scheinkman, 2010).

Finally, this paper is, to my knowledge, the first direct evidence of consumer behavioral responses to rewards from collecting 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 behavioral effects of lottery wins such as the lucky store effect (Guryan and Kearney, 2008).

The remainder of the paper is organized as follows. Section II describes the institutional background of the *Nota Fiscal Paulista* program, presents evidence that the program is salient, describes the relevant datasets, sample definitions and summary statistics. Section III investigates the effects of the program on reported revenue by sector, and examines changes in exit and employment decisions by establishments. Section IV outlines a simple conceptual framework to guide the empirical analysis of potential mechanisms, describes heterogeneous results by establishment characteristics and consumer responses to monetary rewards. Section V introduces a cost-benefit analysis and discusses implications for tax policy. Section VI concludes.

## II Institutional Background and Data

This section provides institutional background for 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.<sup>10</sup>

### II.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 revenue of the state of Sao Paulo depends mostly on its own tax revenue, as opposed to federal transfers.<sup>11</sup> States in Brazil have two main tax instruments: a tax on goods and certain services (ICMS) and a property tax on motor vehicles (IPVA).<sup>12</sup> The ICMS

---

<sup>10</sup>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).

<sup>11</sup>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 a taxes on manufactured products.

<sup>12</sup>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")

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 decided by the states. The average ICMS rate is 18% over the valued added in Sao Paulo.<sup>13</sup> Due to compliance costs - 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.<sup>14</sup> The ICMS average rate in the SIMPLES is 3.5% over gross revenue. Across all establishments, the average tax paid over reported revenue was 4% in 2007.

In 2007 the state of Sao Paulo collected U.S. \$34.5 billion with the ICMS which is 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).<sup>15</sup> Nonetheless, there are many reasons to believe that tax compliance is not perfect in Brazil. According to Schneider et al. (2010), 40% of the country's economy is informal. Unregistered firms are invisible to the tax authority, and no taxes are levied directly from them. Formal firms have to report their activity to the tax authority on a monthly basis, and pay the ICMS associated with 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.<sup>16</sup>

In the case of establishments that sell to final consumers, the tax evasion problem is more severe. Anecdotally, tax authorities tradeoff the cost of audit and the potential extra revenue from an audit. This tradeoff implies that small retail establishments may face a low probability of audit when they evade taxes. The percentage of sales that are underreported or not reported at all reaches 30-40% among smaller firms in Brazil (World Bank Investment Climate, 2003). When the NFP program was implemented, the Secretary of Finance 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).

## ***II.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.<sup>17</sup> The idea of the NFP program is to use consumers as tax auditors by introducing a set of monetary incentives and a system of cross-checks. The policy was implemented together

---

and fees for public services.

<sup>13</sup>The value added is the total value of sales net of inputs.

<sup>14</sup>Some sector restrictions apply. For more details see De Paula and Scheinkman (2011) or Monteiro and Assunção (2012).

<sup>15</sup>The average tax revenue as a share of GDP in developing countries is 17.6% (Gordon and Li, 2009).

<sup>16</sup>The question in WBES (2003) is: "Recognizing that 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%)."

<sup>17</sup>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").

with 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.<sup>18</sup> 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 ask for receipts.

The incentives provided by the NFP program aim to directly affect two forms of underreporting. First, establishments may not report a transaction at all. Second, establishments may falsely claim a lower transaction value.<sup>19</sup> Since rewards are increasing in value of the purchase, consumers have incentives to ask for receipts and to make sure that the value of is reported correctly by the establishment. Consumers can create an online account at the tax authority’s website, which allows them to collect rewards and to cross-check the receipts issued with their SSN. The online system substantially reduces the cost for whistle-blowers to file complaints, and introduces a threat of cross-checks by consumers that may prevent establishments from giving receipts to consumers but not reporting them to the government.

*Eligibility.* Any person that holds a Brazilian Social Security Number (SSN) equivalent is eligible to participate in the program.<sup>20</sup> No pre-registration is needed for consumers to be eligible for tax rebates. In order to be rewarded with lottery tickets for monthly cash prizes, consumers have to create an online account.

*The reward system.* At the moment of purchase, the consumer may ask for the receipt, and give the cashier her SSN.<sup>21</sup> Establishments must send all receipts - with and 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.<sup>22</sup> If the consumer has an online account and opted in for lotteries, the system also automatically generates lottery tickets. 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.<sup>23</sup>

---

<sup>18</sup>The new system is called TD-REDF (“Transmissor de Dados para o Registro Eletrônico de Documento Fiscal”). For transactions between firms, another policy of electronic reporting of receipts was implemented: the NF-e (“Nota Fiscal Eletronica”). For more details on electronic reporting of receipts see De Mello et al (2010).

<sup>19</sup>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, 2009).

<sup>20</sup>Throughout the paper I will refer to the CPF (“Cadastro de Pessoa Física”) as SSN. I will focus on CPF holders only. They are the overwhelming majority of participants in the program. Some NFP participants have a CNPJ (“Cadastro Nacional de Pessoa Jurídica”), which is a firm SSN. Charitable institutions and condominiums also have CNPJ and receive the exact same benefits as final consumers. Also, firms in the SIMPLES tax regime that have yearly gross revenue of less than US\$120,000 are eligible for tax rebates, but they do not participate in the lotteries, and their tax rebate is capped by the amount paid in ICMS in the period. The disbursement is infrequent in this case: from 2007 to 2013 there have only been 3 disbursements of tax rebates to SIMPLES firms.

<sup>21</sup>Note that SSNs in Brazil are not highly-sensitive (as they are in the U. S.). For example, SSNs are written on checks under the signature line, and consumers are frequently asked for their SSN in business transactions.

<sup>22</sup>Throughout the paper I will refer to the receipts with SSN as *NFP receipts*.

<sup>23</sup>The snapshot of the online account and the receipt in figures 2a and 2b are the author’s own online account.

*Tax rebates.* For a given receipt, consumers receive a tax rebate of 30% of the tax collected by the relevant establishment in a month shared among all consumers of that establishment that asked for SSN receipts in that month in proportion to their expenditure in that establishment and month. The calculation of the benefit is complex, and neither the consumer nor the establishment knows precisely the size of the tax rebate from an individual transaction. It is a function of an entire months' worth of SSN receipts and resultant tax revenue.<sup>24</sup>

*Lotteries.* NFP has monthly lotteries since December 2008. For every U.S. \$50 a consumer spends in NFP receipts in a month, she receives one lottery ticket.<sup>25</sup> 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.<sup>26</sup> 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 12 dollar 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. Lottery prizes can be collected soon after the results are released. Consumers have up to 5 years to claim any benefits.

*Complaints.* Consumers can file complaints regarding a purchase made at a specific establishment up to the 15th in the month following the purchase. The consumer needs to 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; (v) other reasons.<sup>27</sup> Consumers do not receive monetary re-

<sup>24</sup>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 a consumer can receive: 7.5% of the total expenditure, which is 30% of the highest VAT rate (of 25%). Thus,  $TaxRebate_{ime} = \min\{0.3 \cdot [ICMS_{em}^{total} \times \frac{V_{iem}}{\sum_{j=1}^N V_{jem}}], 0.075 \cdot V_{ime}\}$ .

<sup>25</sup>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.

<sup>26</sup>Therefore, 50 receipts of 1 dollar value, and 1 receipt of 50 dollar value are equivalent, and generate 1 lottery ticket.

<sup>27</sup>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



wards above and beyond the tax rebate and lottery tickets due from the purchase 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. Additionally, there are fines applied by the consumer's protection bureau PROCON (Fundação de Proteção e Defesa do Consumidor). From consumer's law, establishments must pay up to U.S. \$740 per receipt that was issued with a SSN but not remitted or incorrectly reported to the government. Under tax law, establishments can pay up to 100% of the evaded tax.<sup>28</sup>

*Implementation.* The NFP program 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.<sup>29</sup>

### II.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.08, 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 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 II.C. It shows that the average monthly expenditure is increasing across time,

---

and apply fines accordingly.

<sup>28</sup>Details of the legislation regarding the consumer protection law can be found at *Decreto Estadual 53.085, de 11 de Junho de 2008*. Regarding the tax penalties see *Part IV "violations concerning fiscal documents and tax forms" of Decreto 45490/00*.

<sup>29</sup>Even though the NFP program is targeted to final consumer sales, wholesalers and manufacturers became part of the NFP program in 2009: April and September respectively. Therefore, final consumers that purchased directly from wholesalers and manufacturers could enjoy the same benefits as in retail purchases.

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.<sup>30</sup> Over U.S. \$1.1 billion had been distributed in tax rebates and lottery prizes. 740,000 establishments have submitted over 3.5 billion receipts with consumers' SSN to the tax authority. There was a total of 1,151,518 complaints sent by 135,102 different consumers regarding 134,054 different establishments to the tax authority during the period of analysis.

## ***II.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 III.B and section IV. Second, I explain the datasets at the consumer level, and the key variables I use in section IV.E. In both cases, I focus on the features of the data that are relevant for my empirical analysis. Additional details on variable definitions and sample choices can be found in the Appendix.

### **II.C.1 Establishment Data**

I obtained access to 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.<sup>31</sup>

*Reported revenue.* The NFP program creates incentives for consumers to collect receipts such that the paper-trail deters establishments from underreporting sales. Therefore, the gross reported revenue by a 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 Appendix.

In the empirical analysis I focus on changes in reported revenue. I use log reported revenue in the sector-level analysis in section III.B. In the establishment-level regressions in section IV I scale reported revenue by the average monthly reported revenue the year before the NFP implementation, i.e., between October 2006 and September 2007. In order to reduce the influence of outliers I winsorize this scaled variable by its 95<sup>th</sup> percentile value. This scaling choice delivers coefficients that can be interpreted as proportional changes with the advantage of allowing for zeros in the

---

<sup>30</sup>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.

<sup>31</sup>Due to confidentiality reasons, the data was de-identified, and no establishment data was provided from sectors that have fewer than 5 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 in the period of analysis.

outcome variable, and many establishments report zero revenue in some month of the year. The results are similar when using log reported revenue at the establishment level.<sup>32</sup>

*Tax payments.* 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 paid 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 drops in the time series of tax payments even though the total tax liability of an establishment may not change. Importantly, in some cases there is not a one-to-one relationship between the taxes an establishment actually pays, and how much it remits to the tax authority due to tax withholding policies.<sup>33</sup>

*Establishment characteristics.* From the registry of firms of Sao Paulo the main variable I use is establishments' 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 of food, beverages, tobacco; 4722-9 is Retail of meat and fish; and 4722-9/01 is Retail of 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 to this study includes all establishments that have reported at least one employee between 2004 and 2011. The employment data has a 5-digit sector definition for all establishments. See Appendix for more details.

## II.C.2 Establishment sample and Summary Statistics

*Establishment sample.* From a total of 1,428,531 unique establishments registered in Sao Paulo in the period of analysis, I restrict attention to the 1,035,933 classified as retail or wholesale as described in the previous section. I impose two further restrictions to obtain my primary analysis sample. First, I exclude observations from establishments that were created after October 2007, when the NFP program begins. Second, I only consider establishments that were already operating by January 2004 and that were still active the quarter before NFP implementation.<sup>34</sup> Therefore, I focus on the effect of the program among active establishments that were created at least 3 years before the program started. These restrictions are relevant to keep a relatively stable sample during

---

<sup>32</sup>The coefficients from log regressions are positive and statistically significant, but are typically smaller since a regression featuring log reported revenue systematically excludes observations for smaller establishments that often report zero revenue.

<sup>33</sup>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.

<sup>34</sup>I exclude all establishments that report zero revenue across all months between January 2004 and December 2011, and I also exclude establishments that exited the quarter before the treatment.

the period of analysis. In case of exit, the monthly reported revenue was replaced by zero.<sup>35</sup> The final panel has approximately 20 million observations for 224,528 unique establishments between January 2004 and December 2011. There are 204,072 retail establishments and 20,456 wholesale establishments.

*Sector sample.* I aggregate the reported revenue from the final *establishment sample* described above by 7-digit sectors of activity between January 2004 and December 2011. There are 210 sectors, 90 in retail and 220 in wholesale. The sector sample has 25,200 observations.

*Employment sample.* Because the data from the Department of Finance of Sao Paulo is de-identified, I cannot match the employer-employee dataset (RAIS/CAGED) with the tax data. In order to analyze the impact of the program on employment, I construct a sample of establishments that follows as close as possible the *establishment sample* from Sao Paulo.<sup>36</sup> 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. The summary statistics for the *establishment sample* include the gross reported revenue by establishment for the key groups I use in the empirical analysis. It also shows that, on average, the revenue from SSN receipts accounts for a substantial share of the average revenue in retail, and that over 40% of receipts reported by the establishments in the sample have a SSN attached to it.<sup>37</sup> The summary statistics for the *sector sample* display the average reported revenue across 7-digit retail and wholesale sectors I analyzed in section III.B. 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 and in the other 26 states in Brazil.

### II.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.<sup>38</sup> Here I describe the datasets I use in section VII. Importantly, the source of the consumer-level data is the NFP program. Therefore, there is no "pre-NFP" data on receipts, or any other individual characteristic.<sup>39</sup>

*Receipts data.* This data captures purchases for which final consumers asked for SSN receipts between January 2009 and December 2011. For these receipts the data includes: 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 enroll

---

<sup>35</sup>I investigate the effects of the NFP program on establishment exit separately in section III.C.

<sup>36</sup>I exclude observations from establishments that were born after October 2007, and I only include establishments that were already operating by January 2004.

<sup>37</sup>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.

<sup>38</sup>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.

<sup>39</sup>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.

online. This feature of the data is important in the empirical analysis of section IV.E 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 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 99<sup>th</sup> percentile value.

*Online account data.* This dataset collects information on month and year of enrollment, the timing of monetary reward collection, and the amount collected by 13 million individuals who created an online account at the tax authority’s website from October 2007 to December 2011.<sup>40</sup> Additionally, the dataset contains participation in monthly lotteries: the total number of tickets each consumer holds and the associated prizes she receives.

The main variables I derive 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) *lottery prizes*: the number of lottery prizes and the value of lottery prizes per month.

#### II.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* 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 that participated in one of the 12 monthly lotteries between June 2010 and May 2011. I restrict attention to consumers that hold fewer than 40 lottery tickets in each lottery, which is relevant to assure common support between lottery winners and non-winners in the event-study analysis I describe in section IV.E.<sup>41</sup> 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 I consider 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

---

<sup>40</sup>All data on approximately 90 consumers who won one of the top 3 lottery prizes of over U.S. \$500 dollars were excluded from the datasets available to this study for confidentiality reasons. See Appendix for more details.

<sup>41</sup>See the Appendix for a more detail description of the *lottery sample*.

for least U.S. \$50 dollars 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*.

### III The effect of consumer monitoring on establishment compliance

This section evaluates whether NFP increases tax compliance by establishments. First, I provide evidence that the program is salient by exploiting variation in the disbursement schedules of the monetary rewards. Second, I measure the effect of NFP on the revenue reported by establishments using a difference-in-differences (DD) design.

#### III.A Are consumers paying attention to the program?

In this section I show that consumers are tuned to the schedule of lottery prizes and tax rebates 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 can meaningfully capture salient social patterns that other survey methods cannot capture as easily (Stephens-Davidowitz, 2013). 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 4a pools Google search data from the first to the last day of each month between 2008-2011, and scales each data point by the first day of the month.<sup>42</sup> 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 biannually disburses tax rebates. Figure 4b 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.<sup>43</sup>

---

<sup>42</sup>I exclude the months of April and October - during which the government does the tax rebate disbursements - to make sure the search pattern is related to the lotteries. Including these two months does not change the pattern in the graph.

<sup>43</sup>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.

### III.B Compliance effects: retail vs. wholesale

In order to estimate the effect of NFP on compliance by establishments, I compare two groups of sectors that were 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 that sell 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 growth in the sector groups. Thus, I can shed light on whether the key identification assumption in a DD holds: that trends in potential reported revenue changes are parallel for retail and wholesale sectors. Figure 5a displays changes in total raw reported revenue by group of sectors from January 2005 to December 2011.<sup>44</sup> In this figure each data point is scaled by the average monthly reported revenue in 2004 for the group.<sup>45</sup>

In figure 5a, 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 II.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.

To make sure 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).<sup>46</sup> Two steps are taken to ensure that the survey elicits accurate information on establishments' activities. First, micro-level data is 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 falsify 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;<sup>47</sup> (ii) an actual increase

---

<sup>44</sup>See *sector sample* description in section II.C.

<sup>45</sup>The purpose of scaling the data by this average in a period before the series reported in the graph - instead of the pre-treatment average - is to avoid a mechanical visual divergence of the two lines.

<sup>46</sup>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.

<sup>47</sup>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 5a is indeed a compliance effect.

in retail revenue in Sao Paulo, relative to wholesale.

Figure 5b compares changes in the revenue ratio of retail to wholesale,  $r \equiv \frac{\text{retail revenue}}{\text{wholesale revenue}}$ , from the Sao Paulo administrative data to changes in the same ratio from the census survey.<sup>48</sup> 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.

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  $k$ .<sup>49</sup>

$$R_{st} = \eta_s + \gamma_t + \sum_{k=-8}^8 \beta^k (Treat_s \cdot Period_t^k) + u_{st} \quad (1)$$

where  $R_{st}$  is the log of reported revenue in sector  $s$  and time  $t$ ;  $\eta_s$  are 7-digit sector fixed effects and  $\gamma_t$  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 6a shows the raw log difference between revenue reported in retail vs. wholesale sectors across 6-month windows from Figure 5a without controlling for time and sector fixed effects. The vertical lines mark the beginning and the end the program's implementation. Figure 6b plots the coefficients and the 95% confidence intervals from estimating equation (1) without a constant and controlling for time and sector effects. As in the graph without controls, 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 increase. This effect, averaged across all post-implementation periods, can be estimated from a standard DD specification:

$$R_{st} = \eta_s + \gamma_t + \beta Treat_s \cdot Post_t + u_{st} \quad (2)$$

where  $Post_t = 1$  if  $t \geq \text{October 2007}$  and  $u_{st}$  is clustered by sector. Estimates of equation (2) suggest that the NFP program induced a positive and significant 23% increase in reported revenue

<sup>48</sup>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*.

<sup>49</sup>For instance,  $Period_t^0 = 1$  if  $t \in [\text{Oct.07}, \text{Mar.08}]$ ,  $Period_t^{-1} = 1$  if  $t \in [\text{Apr.07}, \text{Sep.07}]$ , and  $Period_t^1 = 1$  if  $t \in [\text{Apr.08}, \text{Sep.08}]$ .



by establishments across the 4-year period following implementation. As I discuss in more detail in the cost benefit analysis (section V), this effect implies that the NFP program increased tax revenue by U.S. \$400 million net of rewards to consumers. 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.<sup>50</sup>

### III.C *Establishment exit and employment*

Changes in tax enforcement affect the effective tax rate: establishments must pay higher taxes under the new policy. This increase in tax payments may affect establishments that were on the margin of exiting the market or firing an employee. 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 comprises all establishments in retail and wholesale, i.e., without the restrictions from *establishment sample*. Figure 7a shows yearly exit rates by retail and wholesale sectors, where the exit rate is define 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  $s$ .<sup>51</sup> The figure also shows the DD coefficient from estimating a specification similar to equation (2) in a 7-digit sector yearly panel, and in which 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.

*Employment.* To investigate employment effects I use the *employment sample* described in section II.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 7b displays log employment in retail in Sao Paulo and in 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 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.<sup>52</sup>

The evidence above indicates that the NFP policy did not affect establishments’ formal employment and exit decisions, which suggests a small cost from behavioral responses of establishments to the policy in these dimensions. Even though the increase in tax enforcement may change effective tax rates, the change implied by the NFP policy may not be large enough to affect employment

---

<sup>50</sup>I conduct a number of robustness checks. The results are robust to different sample choices, and to various changes in top coding to deal with the influence of outliers. They are also robust to running a two-period DD at the establishment level in a collapsed version of the data by *pre* and *post* periods to account for serial correlation, and control for establishment fixed effects.

<sup>51</sup>I exclude 2011 because my sample period ends at the end of 2011, so I do not have exit data for the month of December in 2011. There is typically a larger number of exits in December of each year relative to other months, so the total of exits in 2011 without December exits misses a relevant share of exits that year.

<sup>52</sup>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.

or exit. The average tax paid over reported revenue was 4% before NFP, so a 23% increase in the effective tax rate is possibly not a large change in the net of taxes revenue.<sup>53</sup> Importantly, 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. However, it is possible that changes in employment and exit may occur after the period of analysis.

## IV Mechanisms: Heterogeneity and Consumer responses

The previous section shows that the program is effective despite opportunities for collusion: the monetary benefit is lower than the tax collected, so firms and consumers could potentially agree to a mutually-beneficial deal and not issue receipts.<sup>54</sup> In order to investigate the mechanisms through which the consumer monitoring program affected establishments' compliance I now turn to the micro data on establishments, receipts and consumers.

I begin by describing a simple conceptual framework that emphasizes two dimensions of the evasion decision by firms that can be affected by consumer monitoring: collusion costs and shifts in audit probability from consumers acting as whistle-blowers. Following the predictions of the model, I examine heterogeneity within the NFP effect across establishments by volume of transactions, receipt value and establishment size. Moreover, I investigate how behavioral considerations can affect how much consumers value rewards, which directly affects the scope for collusion.

### IV.A *Conceptual framework*

I follow the Becker (1968) crime model developed by Allingham and Sandmo (1972) to analyze tax evasion decisions by firms. In particular, the framework uses a variant of this model discussed by Yitzhaki (1978), in which the probability of audit is increasing in the amount evaded. I first consider the baseline case with government monitoring only, and then I introduce rewards to consumers for monitoring firms.

In the case with consumer monitoring, I consider a setting in which firms need to collude with consumers in order to evade taxes for a given transaction. In this case, firms may choose to set a wedge between the prices with versus without a receipt such that consumers forego the receipt and the government's rewards. However, negotiating a collusive deal with consumers at each transaction may be costly. In addition, by proposing a collusive deal to consumers, the firm reveals to a third party that it evades taxes and, in the current context, consumers can act as whistle-blowers.<sup>55</sup>

---

<sup>53</sup>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 relevant to understand the incidence of the policy.

<sup>54</sup>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 obtained to the principal.

<sup>55</sup>Dyck et al., (2009) find that in the context of U.S. corporate fraud, access to information and monetary rewards play a role in encouraging whistle-blowing. In order to factor in the increase in audit probability induced by whistle-blowing threat, I discuss the idea of random shock from the third-party reporting model proposed by Kleven et

## IV.B A tax evasion model

Consider a set of risk-neutral firms that each pay 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. As is standard in the tax evasion literature, I assume that once 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  $\tau(\bar{Y} - Y)$ . Let  $p \in [0, 1]$  be the probability a firm is audited.

The government has two enforcement tools: government audits and consumer monitoring. I consider the case where the probability of government audit 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. In addition, the government can reward consumers that ask for receipts with a share  $\alpha$  of a purchase, where  $\alpha \in (0, \tau)$ .

### IV.B.1 Government monitoring only

For simplicity, assume that in the absence of monetary incentives consumers do not ask for receipts. Thus, firms report revenue  $Y$  to maximize:

$$\pi = (\bar{Y} - \tau Y)(1 - p(\bar{Y} - Y)) + [\bar{Y}(1 - \tau) - (\bar{Y} - Y)\theta\tau]p(\bar{Y} - Y) \quad (3)$$

An interior optimal solution for  $Y$  satisfies the first order condition  $d\pi/dY = 0$ .<sup>56</sup>

$$[p(E) + p'(E) \cdot E](1 + \theta) = 1 \quad (4)$$

Where  $E = \bar{Y} - Y$ . The right hand side of equation (4) is the marginal benefit of evading an extra dollar, and the left-hand side is the marginal cost of evading that extra dollar.<sup>57</sup> Firms choose the optimal  $Y^*$  that satisfies equation (4).

### IV.B.2 Adding consumer monitoring

Consider now the case where the government introduces a reward  $\alpha \in (0, \tau)$  proportional to a purchase  $y$  if a consumer asks for a receipt.<sup>58</sup> Therefore, under the new policy firms that evade taxes need to collude with consumers when they ask for receipts: firms may set a wedge between the price with receipt  $y^r = \bar{y}$ , and a price without receipt  $y^w = \delta\bar{y}$ ,  $\delta \leq 1$ .<sup>59</sup>

---

al. (2009).

<sup>56</sup>The second-order condition is  $-2p'(E) - p''(E)E < 0$ . It is sufficient that  $p(E)$  is convex.

<sup>57</sup>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.

<sup>58</sup>I will assume that firms face a probability of being caught of close to one if they issue a receipt and do not report the transaction to the tax authority. The paper trail alone makes evasion more difficult and, in the NFP policy case, consumers can cross-check the submission of receipts online and file complaints.

<sup>59</sup>I assume that each consumer makes a single purchase based on the posted price  $\bar{y}$ , so the decision is unchanged by the reward policy. This assumption is reasonable if the reward or discount are not salient for the purchase decision.

The price difference must be large enough to convince consumers to trade the government benefit for the deal proposed by the firm. When firms make an offer  $(y^r, y^w) = (\bar{y}, \delta\bar{y})$  to a consumer, they pay a fixed cost  $\phi > 0$ . The fixed cost can be thought of as a concealment cost paid to collude at each transaction.<sup>60</sup> Thus, the total negotiation cost is increasing in the number of collusive deals a firm makes given by  $(\frac{\bar{Y}-Y}{\bar{y}})$ .<sup>61</sup>

*Consumers.* Consumers choose the discount  $\delta$  instead of the government's reward  $\alpha$  if the benefit from colluding,  $(1 - \delta)\bar{y}$ , is higher than  $\alpha$ . I assume that it suffices for firms to match the government's monetary incentives by setting  $\delta = (1 - \alpha)$ .<sup>62</sup>

*Firms.* First, suppose consumers can commit not to whistle-blow, so collusion does not affect the audit probability faced by the firm. Now, firms choose  $Y$  to maximize<sup>63</sup>:

$$\pi = (\bar{Y} - \tau Y)(1 - p(\bar{Y} - Y)) + [\bar{Y}(1 - \tau) - (\bar{Y} - Y)\theta\tau]p(\bar{Y} - Y) - (\bar{Y} - Y)\alpha - (\frac{\bar{Y} - Y}{\bar{y}})\phi \quad (5)$$

Under the new policy, firms have to transfer part of the evasion rents to consumers and have to pay a cost to negotiate a collusive deal in all transactions that the firm does not report. An interior optimal solution  $Y^{**}$  satisfies the first order condition  $d\pi/dY = 0$ :

$$(1 + \theta)[p(E) - p'(E)(E)] = 1 - \frac{1}{\tau}(\alpha + \frac{\phi}{\bar{y}}) \quad (6)$$

The marginal benefit of evading an extra dollar is reduced by  $\frac{1}{\tau}(\alpha + \frac{\phi}{\bar{y}})$ . This amount is equal to the transfer firms need to make to consumers in the collusive deal plus the cost per dollar negotiated with consumers in a given transaction. Therefore, the costs of collusion enter as an extra penalty for each dollar evaded.

Second, consider the case in which consumers cannot commit not to whistle-blow. In this case, collusion may affect the audit probability of a firm since the deal reveals to a third party that the firm is evading taxes. Kleven et al. (2009) argue that a key deterrent of collusion is the sheer

---

Chetty et al. (2009) find that consumers under-react to taxes that are not salient.

<sup>60</sup>The negotiation cost is arguably convex: a super market with a high volume of transactions will likely have a higher cost of negotiation per transaction than a flower shop that has a lower volume of transactions. For expositional reasons, I will consider the case where  $\phi$  is linear in the number of collusive deals the firm needs to make to evade taxes.

<sup>61</sup>The number of collusive deals the firm makes is given by the amount evaded  $E = \bar{Y} - Y$  divided by each individual transaction  $\bar{y}$ . Given the fixed cost  $\phi$ , it will be less costly for firms to evade by colluding with a subset of consumers, than by colluding with all  $N$  consumers to underreport the value of each transaction.

<sup>62</sup>An alternative model would be one in which each consumer could threaten the firm with whistle-blowing to extract part of the total evasion rent of the firm. Kofmann and Lauree (1996) argue that by introducing multiple supervisors that visit the agent in sequence but are imperfectly informed about their position can deter collusion. In the case of consumer monitoring, the sheer number of supervisors make coordination harder. However, there might be noise in the evasion information consumers receive once the firm offers a deal, and consumers may not be willing to extort the firm.

<sup>63</sup>I assume that if the firm is audited the government will consider as tax evasion the amount not reported based on the posted price  $\bar{y}$ . Therefore,  $\bar{Y}$  will be the true revenue of the firm, instead of the revenue net of transfers to consumers.

number of internal or external parties to which a firm that evades taxes exposes itself given the probability of a random shock between the parties. In the final sales case, a random shock could be generated by some conflict between the consumer and the shopkeeper, or a moral concern of the buyer. Therefore, the larger the number of consumers  $N$ , the higher the additional risk of audit introduced by consumers acting as whistle-blowers.<sup>64</sup>

#### *IV.C Comparative statics for empirical analysis*

The reported revenue  $Y^*$  that satisfies the optimal compliance decision with government monitoring only (equation 4) is lower than the reported revenue  $Y^{**}$  that satisfies the optimality condition once I add consumer monitoring (equation 6). The increase in reported revenue  $\Delta Y = Y^{**} - Y^*$  is driven by the reduction in the expected benefit of evasion, which is due to the costs of collusion, and possibly to the fact that the audit probability might be higher if consumers can act as whistle-blowers. In this subsection, I describe three mechanisms that are likely to determine the effectiveness of a consumer monitoring program, and that I further investigate in the empirical analysis.

**Negotiation Costs** The increase in reported revenue  $\Delta Y$  is larger for firms that have a higher negotiation cost  $\phi$ . Moreover, because the cost  $\phi$  is fixed for every transaction, firms that sell low value items  $\bar{y}$  should be more affected than firms that sell high-value items. These comparative statics follow from differentiating equation (6) with respect to  $\phi$  and  $\bar{y}$ , and from the convexity of  $p(E)$ .<sup>65</sup> In the empirical analysis below I exploit variation from establishment characteristics related to negotiation cost and transaction value to test the relevance of the negotiation cost channel.

**Firm size** In the case of government monitoring only, the optimal tax evasion amount  $E^*$  that satisfies equation (4) is equal across firms, so if a firm has 1 dollar more of true revenue  $\bar{Y}$ , the optimal reported revenue  $Y^*$  also increases by one dollar. Therefore, in relative terms, larger firms under-report a smaller share of their revenue if compared to small firms. In this context, the introduction of consumer monitoring implies a larger proportional change  $\frac{\Delta Y}{Y}$  for small firms with low  $\bar{Y}$ . If, in addition, consumers can act as whistle-blowers, there is an increase in the audit probability  $p$ , implying a further increase in reported revenue.<sup>66</sup>

<sup>64</sup> As in Kleven et al. (2009), assume  $\varepsilon > 0$  is the probability that such a random shock occurs; let  $\varepsilon$  be i.i.d. across consumers. Assume an audit will be triggered with certainty if one consumer blows the whistle on the firm, and assume that if a firm evades taxes all its  $N$  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  $p_c = 1 - (1 - p(E))(1 - \varepsilon)^N$ . The new audit probability will be increasing in the number of consumers  $N$ :  $\frac{\partial p_c}{\partial N} > 0$ .

<sup>65</sup> From the convexity of  $p(E)$ , the sign of the comparative static of the total amount evaded  $E^*$  from equation 6 with respect to  $\phi$  will be determined by  $\frac{1}{\tau \bar{y}} > 0$ , and the sign of the comparative static  $\frac{dE}{d\bar{y}}$  will be determined by  $-\frac{\phi}{\tau \bar{y}^2} < 0$ .

<sup>66</sup> Another way to think about the fact that large firms evade less (or relatively less) taxes is to consider a model where the audit probability is increasing in firm size  $\bar{Y}$  instead of being increasing in the amount evaded. Due to third-party reporting, for instance, larger firms are more visible to the tax authority to begin with. In this case, there will be a threshold above which no firms will evade taxes, and below which all firms would evade taxes. Such model would deliver similar predictions.

However, both the shift in audit probability and the burden of negotiation costs are likely to be increasing in  $\bar{Y}$  since the number of customers or potential whistle-blowers is higher for larger firms. Hence, on one hand, larger firms were complying relatively more in the baseline. On the other hand, the enforcement change introduced by consumer monitoring is sharper for larger firms.<sup>67</sup> Therefore, it is theoretically ambiguous whether the effect of the program is increasing or decreasing in firm size. In the empirical analysis I test how the effect of the program varies by establishment size.

**Value of rewards** In the framework described above, firms match the rewards provided by the government through a discount. Therefore, the reward to consumers  $\alpha$  directly reduces firms' benefits from evasion, so higher rewards to consumers should have stronger effects on compliance. In addition, as mentioned above, the more costly it is for a firm to negotiate a deal to match  $\alpha$  (high  $\phi$ ), the more effective the program is. In section IV.E I exploit the lottery component of the rewards to shed light on potential behavioral biases consumers may exhibit. In this case, the value of rewards  $\alpha$  may be even higher than the monetary value of the program's reward, making it particularly costly for a firm to replicate  $\alpha$  through a discount  $\delta$ .

#### IV.D *Heterogeneous responses of establishments*

In this section, I examine heterogeneity in the responses of establishments to the NFP policy in order to shed light on the relevance of negotiation costs and the whistle-blower threat detailed in the previous section.

**Receipt value** If there are fixed negotiation costs, the values of transactions matter for the effectiveness of the NFP policy as discussed in the conceptual framework. In particular, the lower the value of a purchase  $\bar{y}$ , the more establishments should change their compliance levels. To analyze the heterogeneity of the NFP effect on reported revenue by purchase value, I first calculate the median receipt value for each establishment from the *receipts data* described in section II.C, and then I rank retail sectors by the median establishment value.<sup>68</sup>

Table 2 then displays the results from estimating the following equation

$$R_{its} = \eta_i + \gamma_t + \beta_1 DD_{ts} \cdot I_s\{< p50\} + \beta_2 DD_{ts} \cdot I_s\{\geq p50\} + f(x_i)^* DD_{ts} + \varepsilon_{its} \quad (7)$$

where  $R_{its}$  is reported revenue in establishment  $i$  in month-year  $t$  and and sector  $s$ , scaled by the average monthly reported revenue during the year before the NFP program implementation,

<sup>67</sup>Optimal revenue evasion amount  $E^* = \bar{Y} - Y^*$  from equation (4) is decreasing in the audit probability  $p$ , as has been extensively documented in the literature. If the change in audit probability is increasing in  $\bar{Y}$ , the change in compliance is smaller for small firms.

<sup>68</sup>Even though I have receipt data at the establishment level, I use a sector-level definition of receipt value instead to mitigate endogeneity concerns from the fact that I use data from after the program was implemented (between 2009 and 2011). Examples of sectors classified as low receipt value: coffee shops and drugstores. Examples of sectors classified as high receipt value: car dealers and home furnishings.

i.e., between October 2006 and September 2007.<sup>69</sup> Establishment fixed effect are denoted by  $\eta_i$ ,  $\gamma_t$  is a month-year fixed effect,  $I_s\{< p50\}$  is a dummy for retail sectors below the median of the receipt value distribution across sectors (Low receipt value) and  $I_s\{\geq p50\}$  is a dummy for retail sectors above the median (High receipt values). The term  $f(x_i)$  is a 5th-order polynomial of establishment size as measured by the average reported revenue during the year 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. I flexibly control for establishment size effect through an interaction of  $DD_{ts}$  with  $f(x_i)$  because the receipt value heterogeneity could pick up establishment size effects. The error  $\varepsilon_{its}$  is clustered by sector.

Column (1) of Table 2 shows the average DD effect from estimating equation (7) without splitting the retail sector into two groups. The coefficient can be interpreted as an 18% increase in reported revenue by establishment. Column (2) presents the results from estimating equation (7). The increase in reported revenue positive and significant for both groups of sectors, but is stronger for sectors characterized by low receipt values. The two coefficients are statistically different at the 1% level.

The comparative static described in the previous section with respect to the value of a purchase  $\bar{y}$  implies that the effect of the policy should monotonically decrease with the value of the purchase. In order to further investigate whether the heterogeneity effect by receipt value is monotonic, I bin sectors in quintiles of the receipt value distribution across sectors and estimate the equation:

$$R_{its} = \eta_i + \gamma_t + \sum_{m=1}^k \alpha_m (d_{ms} \cdot DD_{ts}) + f(x_i) * DD_{ts} + \varepsilon_{its} \quad (8)$$

where  $R_{its}$ ,  $\eta_i$ ,  $\gamma_t$ ,  $DD_{ts}$  and  $\varepsilon_{its}$  are defined as in equation (7). The term  $d_{ms} = 1$  if sector  $s$  is in quintile  $m$  in the receipt value distribution across sectors. Figure 8a plots the coefficients and 95% confidence interval from the estimating equation (8). The figure shows that the change in reported revenue is, on average, decreasing in receipt value, and that the effect seems to come from small value transactions, consistent with the prediction described in section IV.C.

**Volume of transactions** The volume of transactions an establishment typically generates is related to both the negotiation cost  $\phi$  and the increased threat of audit  $p$  under consumer monitoring. If an establishment needs to negotiate a collusive deal with each consumer, the sheer volume of transactions may increase the cost per transaction.<sup>70</sup> Moreover, the volume of transactions is related to the volume of consumers, and 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. Even though the variation in the data cannot distinguish the two effects, both mechanisms suggest that the effect of

<sup>69</sup> As I discuss in section II.C this definition allows for zero reported revenue in a given month, and the coefficient can be interpreted as a proportional change.

<sup>70</sup> For instance,  $\phi$  may be a convex function of the number of customers an establishment has: negotiating with consumers in a supermarket may be costly since other consumers may be forced to wait; this is a cost likely lower in a flower shop.

the program should be increasing in the volume of transactions.

To study how the effect of the program varies with 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.<sup>71</sup> Column (3) of Table 2 presents the results from estimating a version of equation (7), in which  $I_s\{< p50\}$  is defined as a dummy for retail sectors below the median of the transaction volume distribution across sectors (low volume of receipts) and  $I_s\{\geq p50\}$  is a dummy for retail sectors above the median (high volume of receipts). The increase in reported revenue is positive and significant for both groups of sectors, but is stronger for sectors characterized by high volume of receipts. As in column (2) of Table 2, the estimated coefficients in column (3) are reported after flexibly controlling for establishment size effects. The two coefficients are statistically different at the 1% level.

The comparative statics from the conceptual framework with respect to negotiation cost  $\phi$  and the increase threat of audit  $p$  imply that the effect of the NFP policy should monotonically increase with the volume of receipts. Figure 8b plots the coefficients and 95% confidence intervals from estimating a version of equation (8) in which  $\beta_k$  varies by each quintile  $k$  of the transaction volume distribution across sectors instead of the receipt value distribution. 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 transaction volume of transactions as predicted by the model.

**Establishment size** Now I turn to the heterogeneity of the reported revenue effect by establishment size as measured by the average reported revenue the year before the program implementation. Column (4) of Table 2 displays the results from splitting establishments into two groups according to the distribution of establishment size, and estimate an equation below

$$R_{its} = \eta_i + \gamma_t + \beta_1 DD_{ts} \cdot I_i\{< p50\} + \beta_2 DD_{ts} \cdot I_i\{\geq p50\} + \varepsilon_{its} \quad (9)$$

where  $R_{its}$ ,  $\eta_i$ ,  $\gamma_t$ ,  $DD_{ts}$  and  $\varepsilon_{its}$  are defined as in equation (7). The indicator variable  $I_i\{< p50\}$  is a dummy for establishments below the median size distribution, and  $I_i\{\geq p50\}$  is a dummy establishments above the median size distribution, and the other variable are defined in the same way as in equation (7). The increases in reported revenue captured by estimates of  $\beta_1$  and  $\beta_2$  from equation (9) are positive and significant for both groups of establishments. Note that the difference between the two coefficients is not statistically significant at the 5% level.

As discussed in section IV.C, the heterogeneity of the effect of the program by establishment size may not be monotonic. In order to verify how the effect varies across establishment size, I bin establishments into 10 deciles according to the establishment size distribution. Then I interact each size dummy with the DD indicator:

---

<sup>71</sup>Examples of sectors classified as low volume of transactions: art supplies and second hand shops. Examples of sectors classified as high volume of transactions: supermarkets and gas stations.



$$R_{its} = \eta_i + \gamma_t + \sum_{m=1}^{10} \alpha_m (d_{mi} \cdot DD_{ts}) + \varepsilon_{its} \quad (10)$$

where  $R_{its}$ ,  $\eta_i$ ,  $\gamma_t$ ,  $DD_{ts}$  and  $\varepsilon_{its}$  are defined as above, and  $d_{mi} = 1$  if establishment  $i$  is in decile  $m$  in the establishment size distribution.

Figure 9 plots the estimates coefficients  $\hat{\alpha}_m$  from estimating equation (10) using the establishment sample described in section II.C. The graph suggests that the effect of the program has an inverted-U shape with respect to establishment size. These results are consistent with the argument outlined in the conceptual framework: even though the effect of the program may be decreasing in firm size because larger firms were complying relatively more before the program implementation, the enforcement change introduced by consumer monitoring is likely to be weaker for small firms.

#### **IV.E Consumer responses to rewards**

As detailed in section IV.C the more consumers value the rewards  $\alpha$ , the more effective NFP will be in preventing tax evasion. In addition, the negotiation cost to cut a deal with consumers to avoid issuing the receipt may be directly affected by the composition of program rewards.

The lottery component of the rewards may leverage consumers' taste for gambling or individual behavioral biases. Friedman and Savage (1947) noted that individuals may even be willing to pay for lotteries with negative expected gain. Filiz-Ozbay et al.(2013) find evidence that prize-linked savings offered by commercial banks and governments across the world may be 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 in comparisons of 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 issuing receipts.

In this section, 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 behavioral responses of consumers due to lottery rewards, and then I show suggestive evidence of the effect of lottery eligibility on consumer participation.

##### **IV.E.1 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 increase their estimate of the probability a ticket bought from the store that sold a winning ticket in the past will be a lottery winner (the "lucky store effect").<sup>72</sup>

Following Hilger (2013) 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 have 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 get 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 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.<sup>73</sup> 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.<sup>74</sup>

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 10 displays the average number of receipts for which lottery winners and non-winners ask by 6 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  $k \in [-6, 6]$ :

$$y_{gj} = \gamma_j + \pi_g + \beta \cdot I\{j \geq 0, g = T\} + u_{gj}, \quad (11)$$

where  $y_{gj}$  is the average number of SSN receipts group  $g$  asks in "event-month"  $j$ . Figure 10

---

<sup>72</sup>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.

<sup>73</sup>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.

<sup>74</sup>For detailed description of DFL-reweighting, see Yagan (2013). For details of this specific application, see Appendix.

shows that there is a significant difference in behavior between lottery winners and non-winners for all prize levels it displays. 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 tuned to the lottery results.

The evidence is consistent with a behavioral explanation, given that there is a significant .6% difference between the number of receipts lottery winners and non-winners ask for even for a 5-dollar prize. Since the odds of winning are independent of past wins, the change in behavior observed in Figure 10 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. However, the effect 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 lotteries wins as a signal that the program works as advertised. Figure 11a shows the effect of a 5-dollar 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 ask for this subsample. Figure 11b 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 11c 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.

#### IV.E.2 The effect of lottery eligibility

In order to shed light on the relative effectiveness of tax rebates and lottery prizes I 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 a consumer creates an 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*, *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.

$$y_{i,t_O,t_E} = \sum_{j=-6}^6 \beta_j \cdot I\{j = k\} + \gamma_i + \pi_{t_O} + u_{i,t_O,t_E} \quad (12)$$

where  $y_{i,t_O,t_E}$  is one of the four outcomes of interest,  $I\{j = k\}$  is a dummy variable that is equal to 1 when  $j = t_O - t_E$ ,  $\gamma_i$  is individual fixed effects,  $\pi_{t_O}$  is 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  $\beta_j$  by exploiting variation in the timing of enrollment. Figure 12 displays the estimated  $\hat{\beta}_j$  from Equation (12) 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 12a shows that consumers would ask on average for 4 receipts with SSN before they enroll online, while they are only eligible for tax rebates. After enrolling at period zero, the average number of receipts they ask for per month doubles, and the change seems to be a permanent level shift in participation. The number of different establishments in which a consumer asks for receipts also increases considerably in Figure 12b. Figure 12c shows that there is an upward trend in the total amount spent in SSN receipts before an individual enrolls 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 12d; so consumers are asking for receipts more often and more broadly.

The effects described in figures 12a to 12d 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 condition their participation in the program on past lottery wins discussed in the previous section, are consistent with lotteries being an effective device to create incentives for consumers. Since lotteries are relatively less costly than tax rebates, a composition of rewards that puts more resources in lotteries relatively to rebates may potentially be more cost-effective. However, more research is needed in order to pin down the relative effectiveness of 1 dollar in tax rebate versus 1 dollar in lottery prizes.

## V Implications for tax policy

The empirical analysis performed in sections III and IV suggests that the policy of creating incentives for consumers to collect receipts is an effective way to improve compliance by establishments. However, by carrying out this policy, the government of Sao Paulo is forgoing part of its tax revenue: both incremental revenue from the program and infra-marginal 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 ways to make such policies more cost-effective.

The total tax paid by the retail sector between 2004 and 2007 in Sao Paulo was U.S. \$9 billion. Thus, the 23% increase in compliance reported in section III.B and Figure 6 would imply an increase

in tax collection of approximately U.S. \$2 billion over four years due to NFP. In the same period, U.S. \$1.6 billion dollars of rewards were distributed to consumers with online accounts.<sup>75</sup> Therefore, the program generated a U.S. \$400 million net gain in tax revenue in the period of analysis. This increase in revenue needs to be balanced against the costs of obtaining this revenue, including not only the cost of disbursement but also the costs borne by the program participants.<sup>76</sup>

In order to shed light on participation costs, I take two approaches: an opportunity cost calculation and a revealed preference calculation. For the opportunity cost I approximate the cost of time for a consumer to spell out the SSN based on the average hourly wage in Sao Paulo.<sup>77</sup> Since the time to spell out the SSN is a cost for both the consumer and the cashier, I add the average hourly wage in Sao Paulo and the average hourly wage in retail, which equals U.S. \$6.8.<sup>78</sup> Considering that it takes 11 seconds to spell out 11 digits of the Brazilian SSN, and that between 2007 and 2011 there were 3.2 billion SSN-identified receipts, the cost adds up to U.S. \$ 66 million. Assuming it takes 30 minutes to set up an online account, and considering that 13 million people in Sao Paulo have enrolled in the period of analysis, there is an additional U.S. \$28 million in costs. From this approach, the total participation cost would be around U.S. \$ 94 million.

In the revealed preference approach, I exploit individual receipt data. Arguably, consumers deciding whether or not to ask for a receipt weight the monetary benefits of the program against the costs from asking for receipts. Therefore, by observing the individual receipt value I can approximate the threshold below which consumers may not consider it beneficial to ask for receipts. I calculate the minimum of the receipt value distribution for each individual and the average reward associated with the lowest receipt value, and then I multiply this threshold by the number of transactions by individual in the period of analysis.<sup>79</sup> By trimming total individual cost distribution

---

<sup>75</sup>Out of the U.S. \$1.6 billion, U.S. \$1.1 billion was actually claimed by final consumers over four years after the program started: either through bank account deposits or through payments of other state taxes. As of December 2011, U.S. \$500 million have not been claimed by consumers enrolled online, but this amount will likely be collected eventually since these consumers have already created an online account.

<sup>76</sup>Berhan and Jenkin (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 physically keep the receipts because 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 and Turkey. Moreover, in these cases, there is an incentive to break down the same purchase into multiple receipts or create fake receipts. NFP attempted to reduce participation costs to consumers and to avoid incentives to create fake receipts. It leverages the fact that establishments need to report receipts electronically to the tax authority regardless of whether it has a consumer's SSN attached to, and it depends on the fact that Brazilians are willing to put their SSN on receipts. 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. Consumers may want to hold on to their receipt in order to file complaints, but it is not strictly necessary to do so to derive the benefits from the program.

<sup>77</sup>It is possible, however, that participating in the program has a leisure value to consumers, since they may extract utility from playing lotteries, or from participating in the program more-generally.

<sup>78</sup>I measure hourly wages from the administrative employer-employee dataset (RAIS), which contains data from only formal employment. Formal sector wages are typically higher than informal sector wages.

<sup>79</sup>A formal approach to this estimate would apply a model that allows for extreme value censoring, since the threshold is arguably below the minimum receipt I observe in the data. While this argument would suggest that the calculation proposed here is an over-estimate, it is also possible that the threshold rule followed by consumers may be stochastic, in which case estimating costs based on the minimum receipt value may under-estimate the overall cost.

at the 90<sup>th</sup> percentile,<sup>80</sup> the total cost adds up to U.S. \$ 174 million.<sup>81</sup>

In order to analyze the cost-benefit of NFP, I calculate the marginal cost of public funds (MCPF) based on a cost-benefit ratio proposed by Hendren (2013) for non-budget neutral policies. The ratio is the marginal social welfare impact of the policy per unit of government revenue expended. In the NFP case,  $MCPF = \frac{\text{net costs to society}}{\text{tax revenue gain}}$ . The tax revenue gain is U.S. \$400 million. Abstracting from other potential benefits to society, the numerator of the MCPF is the extra U.S. \$400 million the government is taxing society plus the cost of participating in the program. Using cost calculations of U.S. \$ 94 million or U.S. \$ 174 million described above, the MCPF from this policy would be between 1.23 and 1.43 dollars; i.e., the program would cost 1.23 to 1.43 dollars in welfare for every dollar it collects.<sup>82</sup> These numbers fall within the general range found in the literature for the marginal value public funds of tax policies in the U.S: Hendren (2013) finds a MVPF of \$2 for high earners and 0.88 for EITC beneficiaries in the context of income tax in the U. S..<sup>83</sup>

Nonetheless, there may be other social benefits to be considered in the cost-benefit analysis. 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 the consumer propensity to ask for receipts even if the government eventually discontinues the rewards. The program also allows consumers to provide the SSN of charities instead of their own, which may have "warm glow" effects.<sup>84</sup>

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.<sup>85</sup> 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.<sup>86</sup> Another potential benefit is the presence of

---

<sup>80</sup>The total compliance cost would be over a 3 billion dollars since it is heavily driven by outliers. Individuals above the 90th percentile in the cost distribution correspond to more than 96% of the total cost and less than 30% of the total number of transactions.

<sup>81</sup>I do not have receipt data from before 2009. In order to calculate the total cost between 2008 and 2011, I calculate the average yearly cost between 2009 and 2011 and apply this average to 2008. Since the number of individual transactions increased over the years, this extrapolation may be an over - estimate of the total cost.

<sup>82</sup>One aspect of costs could be tax-paying workers losing their job or firm exit, but I find no effects of the program in these margins.

<sup>83</sup>NFP is arguably relying on the fact that some consumers may never collect rewards. 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. Taking the tax rebates of non-enrolled consumers into account, the program would be generating about US\$5 million dollars above and beyond all of its rewards liabilities. Hence, considering all the rewards generated by the policy - irrespective of whether it will be collected or not - the net benefit becomes US\$5 million dollars and the MCPF would be between 19 and 35 dollars.

<sup>84</sup>Since the option of indicating the SSN of a non-profit organization was introduced in May 2009 until December 2011, 6% of all SSN-identified receipts had a non-profit's SSN.

<sup>85</sup>According to Andreoni et al (1998) it is an unresolved issue whether governments put different weights in cheaters vs. honest tax payers.

<sup>86</sup>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

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.<sup>87</sup> Importantly, millions of people are already participating and have paid the fixed cost of setting up online accounts. Given the take-up NFP achieved already, the program could potentially change some of the incentives to make the program 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 IV.E, it is possible that lotteries provide a stronger incentive than tax rebates per dollar spent by the government.

## VI Conclusion

Many developing countries rely on firm-level taxation to finance public goods but face important enforcement problems. The VAT is a popular tax instrument in these contexts due to its self-enforcement features, but the built-in incentives break down at the final consumer stage since consumers usually have no direct monetary benefit from asking for receipts. I show that a program designed to change incentives at final sales from Sao Paulo, Brazil increased revenue reported in retail sectors by at least 23%: approximately U.S. \$400 million net of consumers' rewards over four years. The implied increase in effective tax rates did not affect firm exit or employment.

In order to study mechanisms I describe a simple conceptual framework where a consumer reward program can be effective despite opportunities for collusion due to increases in audit risk and costs of negotiation. Based on the predictions of the model, I examine heterogeneity across establishments and consumer responses to rewards. The estimated effect is stronger for sectors with a high volume of transactions and small receipt values, consistent with fixed costs to negotiate collusive deals to avoid issuing receipts. Moreover, the effect of the program has an inverted-U shape with respect to establishment size, consistent with a higher baseline level of compliance among larger firms, and shifts in audit probability from consumer monitoring that increase in firm size.

Furthermore, I show that consumers are finely tuned to the incentives of the program, and I exploit the random component of lottery rewards to show that consumers condition their participation on past lottery wins. Even small prizes generate a significant and steady increase in the number of receipts consumers ask, and in the number of different business in which they ask for receipts. The results are consistent with the possibility that lotteries amplify consumer responses

---

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.

<sup>87</sup>The bar code cards are not widely used yet, partially because bar coding is still not standardized across shops.

due to behavioral biases, indicating that lotteries might be an effective incentive per dollar spent by the government. I find that on net the program is fairly cost-effective, and perhaps could be made less costly by lowering participation costs and by relying more on lotteries.

The findings of the paper have broader implications for self-enforcing policies and consumer-rewards programs beyond the Sao Paulo experience. Evidence presented in the paper is consistent with the idea that collusion is a potential barrier to the effectiveness of a self-enforcing reward system when there is scope for a mutually beneficial deal: the NFP effect is stronger where collusion is less likely to pay off. Moreover, the paper provides supporting evidence that consumers respond to lottery incentives. Most governments only use lotteries in such policies, and this paper is, to my knowledge, the first direct evidence of consumer responses to rewards from collecting receipts.

Finally, the results raise several questions for future research. The paper provides a positive analysis of using consumers as tax auditors. A normative approach would be a natural next step in understanding the optimal design of final sales enforcement policies, given that the government can audit firms directly or use consumer monitoring, and there may be relevant trade-offs regarding collusion opportunities in both enforcement tools. A more comprehensive evaluation of a consumer-rewards program may also need to consider redistributive impacts, and social interactions effects in consumer participation. From a policy perspective, it would be important to build more evidence on the relative cost-effectiveness of different rewards options - tax rebates, lottery in-kind prizes or cash lottery prizes - that are common in policies that incentivize 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, Leticia 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. **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.
5. **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.*
6. **Becker, Gary S.**1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, 76 (2): 169-217.
7. **Berhan, Bahro A. and Glenn P. Jenkins.** 2005. "The High Costs of Controlling GST and VAT Evasion." *Canadian Tax Journal*, 53, 720–36.
8. **Bird, Richard M.**1992. "Tax reform in Latin America: a review of some recent experiences." *Latin American Research Review*, 27 (1): 7-36.
9. **Bird, Richard and Pierre-Pascal Gendron.** 2007. *The VAT in Developing and Transitional Countries.* Cambridge University Press.
10. **Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer.** 2013. "Salience and Asset Prices." *American Economic Review: Papers & Proceedings*, 103(3): 623–628.
11. **Bruhm, Miriam and David McKenzie.** 2013. "Entry Regulation and Formalization of Microenterprises in Developing Countries." *The World Bank Policy Research Working Paper* 6507.
12. **Burgess, Robin and Nicholas Stern.** 1993. "Taxation and development." *Journal of economic literature*, 31(2): 762-830.
13. **Carrillo, Paul, Dina Pomeranz and Monica Singhal.** 2013. "Tax Me if You Can: Evidence on Firm Misreporting Behavior and Evasion Substitution." *Working Paper.*
14. **Chetty, Raj, Adam Looney, and Kory Kroft.** 2009. "Salience and Taxation: Theory and Evidence." *American Economic Review*, 99(4): 1145–1177.
15. **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.
16. **International Monetary Fund.** 2011. "Revenue Mobilization in Developing Countries." *Fiscal Affairs Department.*

17. **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.
18. **De Paula, Aureo and José A. Scheinkman.** 2010. "Value Added Taxes, Chain Effects and Informality." *American Economic Journal: Macroeconomics*, 2(4):195-221.
19. **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.
20. **Dyck, Alexander, Adair Morse and Luigi Zingales.** 2010. "Who Blows the Whistle on Corporate Fraud?" *Journal of Finance*, 65 (6): 2213-2254.
21. **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*.
22. **Friedman, Milton, and Leonard J. Savage.**1948. "The utility analysis of choices involving risk." *The Journal of Political Economy* 56 (4): 279-304.
23. **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.
24. **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.
25. **Guryan, Jonathan and Melissa S. Kearney.** 2008. "Gambling at Lucky Stores: Evidence from State Lottery Sales." *American Economic Review*, 98 (1): 458-73.
26. **Hendren, Nathaniel.** 2013. "Policy Elasticity." *NBER Working Paper 19177*.
27. **Hilger, Nathaniel G.** 2013. "How does family income affect college enrollment? Evidence from timing of parental layoffs", *Working Paper, Brown University*.
28. **Jacobson, Louis S., Robert J. LaLonde and Daniel G. Sullivan.** 1993. "Earnings Losses of Displaced Workers," *The American Economic Review*, 83 (4): 685-709.
29. **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>
30. **Kaustia, Markku and Samuli Knüpfer.** 2008. "Do Investors Overweight Personal Experience? Evidence from IPO Subscriptions." *The Journal of Finance*, 63 (6): 2679-2702.
31. **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.

32. **Keen, Michael and Jack Mintz.** 2004. "The optimal threshold for a value-added tax." *Journal of Public Economics*, 88 : 559– 576.
33. **Keen, M., and B. Lockwood.** 2010 "The Value Added Tax: Its Causes and Consequences," *Journal of Development Economics*, 92(2): 138 151.
34. **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.
35. **Kleven, Henrik J., Martin B. Knudsen, Claus T. Kreiner, and Emmanuel Saez.** 2009. "Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries." *NBER Working Paper 15218*.
36. **Kofman, Fred and Jacques Lawarree.** 1996. "A prisoner's dilemma model of collusion deterrence." *Journal of Public Economics*, 59 : 117– 136.
37. **Kopczuk, Wojciech, and Joel Slemrod.** 2006. "Putting firms into optimal tax theory." *The American Economic Review*, 96 (2): 130-134.
38. **Kumler, Todd, Eric Verhoogen and Judith Frias.** 2013. "Enlisting Workers in Monitoring Firms: Payroll Tax Compliance in Mexico." *NBER Working Paper 19385*.
39. **La Porta, Rafael and Andrei Shleifer.** 2008. "The Unofficial Economy and Economic Development." *Brookings Papers on Economic Activity*, 275-352.
40. **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*.
41. **Marchese, Carla.** 2009. "Rewarding the Consumer for Curbing the Evasion of Commodity Taxes?" *FinanzArchiv: Public Finance Analysis*, 65 (4): 383-402.
42. **Mattos, Enlison, 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.
43. **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.
44. **Musgrave, Richard A..** 1969. *Fiscal systems*. New Haven: Yale university press.
45. **Pomeranz, Dina.** 2013. "No taxation without information". *Harvard Business School Working Paper 13-057*.
46. **Shleifer, Andrei and Vishny, Robert W.** 1993. "Corruption." *The Quarterly Journal of Economics*, 108 (3): 599-617.
47. **Schneider, Friedrich, Andreas Buehn, and Claudio Montenegro.** 2010. "Shadow Economies All Over the World: New Estimates for 162 Countries from 1999 to 2007." *World Bank Policy Research Working Paper Series*, 5356.

48. **Slemrod, Joel.** 2007. "Cheating ourselves: the economics of tax evasion." *The Journal of Economic Perspectives*, 21 (1): 25-48.
49. **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.
50. **Stephens-Davidowitz, Seth.** 2013. "The cost of racial animus on black presidential candidate: Using google search data to find what surveys miss." *Working Paper*.
51. **Tirole, Jean.** 1986. "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations." *Journal of Law, Economics, & Organization*, 2 (2): 181-214.
52. **Tufano, Peter.** 2008. "Saving whilst Gambling: An Empirical Analysis of UK Premium Bonds." *American Economic Review: Papers & Proceedings*, 98 (2): 321-326.
53. **Yagan, Danny.** 2013. "Capital Tax Reform and the Real Economy: the Effects of the 2003 Dividends Tax Cut." *Working paper*.
54. **Yaniv, Gideon.** 1993. "Collaborated Employee-Employer Tax Evasion." *Public Finance/Finances Publiques*, 47: 312-321.
55. **Yitzhaki, Shlomo.** 1987. "On the excess burden of tax evasion." *Public Finance Review*, 15 (2): 123-137.
56. **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

<b>a: Establishments</b>				
	Number of Obs.	Mean	Std. Dev.	Time Period
<b>Establishment sample</b>				
<i>Retail establishments</i>				
Reported revenue	19,038,509	25,640	83,549	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
Revenue from SSN receipts	8,083,008	16,308	2,042,087	Jan.2009-Dec.2011
Share of receipts with SSN	8,083,008	0.44	0.41	Jan.2009-Dec.2011
<i>Wholesale establishments</i>				
Reported revenue	1,942,310	116,033	196,069	Jan.2004-Dec.2011
<b>Sector sample</b>				
Reported revenue				
<i>Retail sectors</i>	10,800	51,000,104	132,623,090	Jan.2004-Dec.2011
<i>Wholesale sectors</i>	14,400	18,185,621	26,373,370	Jan.2004-Dec.2011
<b>Employment sample</b>				
Number of employees				
<i>Retail - Sao Paulo</i>	351	7.3	8.2	2004 - 2011
<i>Retail - other Brazilian States</i>	9,117	6.7	6.4	2004 - 2011
<b>b. Consumers</b>				
	Number of Obs.	Mean	Std. Dev.	Time Period
<b>Consumers sample</b>				
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
<b>Lottery sample</b>				
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 SSN 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 the establishment sample 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. *Total amount claimed*: the total value of rewards claimed by consumers through bank account deposits; *number 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: Mechanisms and Heterogeneous effects of NFP

	Reported revenue, scaled by average reported revenue during the pre-implementation year (Oct 2006-Sep 2007)			
	(1)	(2)	(3)	(4)
DD (Post Oct 07 * retail)	0.183*** [0.0398]			
DD * Low receipt value		0.280*** [0.0406]		
DD * High receipt value		0.091** [0.0390]		
DD * Low volume of receipts			0.113*** [0.0405]	
DD * High volume of receipts			0.286*** [0.0412]	
DD * Small establishments				0.161*** [0.0450]
DD * Large establishments				0.200*** [0.0405]
5th-order polynomial of establishment size * DD		x	x	
Time FE	x	x	x	x
Establishment FE	x	x	x	x
R-squared	0.32	0.33	0.33	0.32
Number of observations	20,980,819	20,980,819	20,980,819	20,980,819

Notes: standard errors are clustered at the 7-digit sector classification level (210 clusters). Significance levels \*\*\* 1%, \*\* 5%. The sample is composed by monthly observations between Jan 2004 and Dec 2011 of establishments classified in retail or wholesale as described in Section II. 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. The dependent variable is the reported revenue of an establishment, scaled by its average reported revenue the year before program implementation. The dependent variable is winsorized at the 95<sup>th</sup> percentile. Column (1) shows the average DD estimate controlling for time and establishment fixed effects. Column (2) splits retail sectors into two groups: sectors below the median of the receipt value distribution across sectors (Low receipt value) and sectors above the median (High receipt values). Receipt value is defined by ranking 7-digit retail sectors by the median receipt value issued by an establishment between 2009 and 2011. 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 5<sup>th</sup> 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). Size is defined by the total reported revenue by establishments during the year 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: 000000000-00  
 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".

Período  
 Quantidade de Documentos: 31  
 Valor dos Documentos: R\$ 1.000,00  
 Período 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	
000000000000	ALMOCO	222782	13/05/2012	260,72	14/05/2012	2,43	Liberado	Detalhes
000000000000	ALMOCO	23000	13/05/2012	11,50	13/06/2012	0,00	Liberado	Detalhes
000000000000	NOVA OPTIC UNIVERSIDADE NARITOMI LTDA	49666	11/05/2012	39,58	16/05/2012	0,59	Liberado	Detalhes
000000000000	ALMOCO	436579	09/05/2012	64,12	21/05/2012	0,57	Liberado	Detalhes
000000000000	ALMOCO	494128	09/05/2012	85,84	10/05/2012	0,00	Liberado	Detalhes
000000000000	ALMOCO	377208	09/04/2012	18,58	04/04/2012	0,08	Liberado	Detalhes
000000000000	ALMOCO	377823	03/04/2012	31,25	04/04/2012	0,13	Liberado	Detalhes
000000000000	STARBUCKS UNIVERSIDADE NARITOMI LTDA	283678	03/04/2012	3,90	04/04/2012	0,12	Liberado	Detalhes
000000000000	AVATAR UNIVERSIDADE NARITOMI	3838	31/03/2012	15,80	11/04/2012	0,04	Liberado	Detalhes
000000000000	ALMOCO	21597	30/03/2012	18,39	02/04/2012	1,00	Liberado	Detalhes
000000000000	ALMOCO	374341	29/03/2012	31,12	30/03/2012	0,00	Liberado	Detalhes
000000000000	ALMOCO	208111	29/03/2012	31,00	23/04/2012	0,50	Liberado	Detalhes
000000000000	ALMOCO	113813	28/03/2012	39,99	05/04/2012	0,96	Liberado	Detalhes
000000000000	ALMOCO	80014	25/03/2012	73,20	09/04/2012	4,74	Liberado	Detalhes

## b.: Nota Fiscal Paulista Receipt

**Secretaria da Fazenda** Nota Fiscal Paulista

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

CPF: 000000000-00  
 Usuário: JOANA NARITOMI  
 Saldo Disponível Para Saque: R\$ 0,00

**Cupom Fiscal**

CPF: 000000000-00  
 IE: 0000000000  
 IM: 0000000000

30/3/2012 00:00:00 CCF: 20428 COD: 21597

**CNPJ/CPF consumidor: 000000000**

Nome:  
 End:

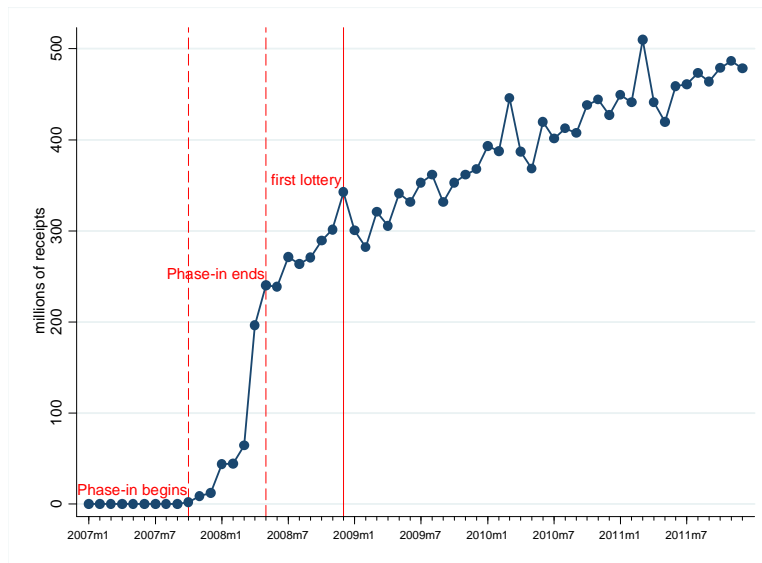
**CUPOM FISCAL**

ITEM	CÓDIGO	DESCRIÇÃO	ST	VL ITEM R\$
001	00000000000001	ALMOCO		18,39
	1,00 x			18,39
		DESCONTO R\$		R\$ 0,00
		ACRESCIMO R\$		R\$ 0,00
		<b>TOTAL R\$</b>		<b>R\$ 18,39</b>

SPM 2RM/ITIP LOGGER ECF-IF VERSÃO:03.03.04 ECF: L3: OPR:  
 3/3/2008 07:27:37

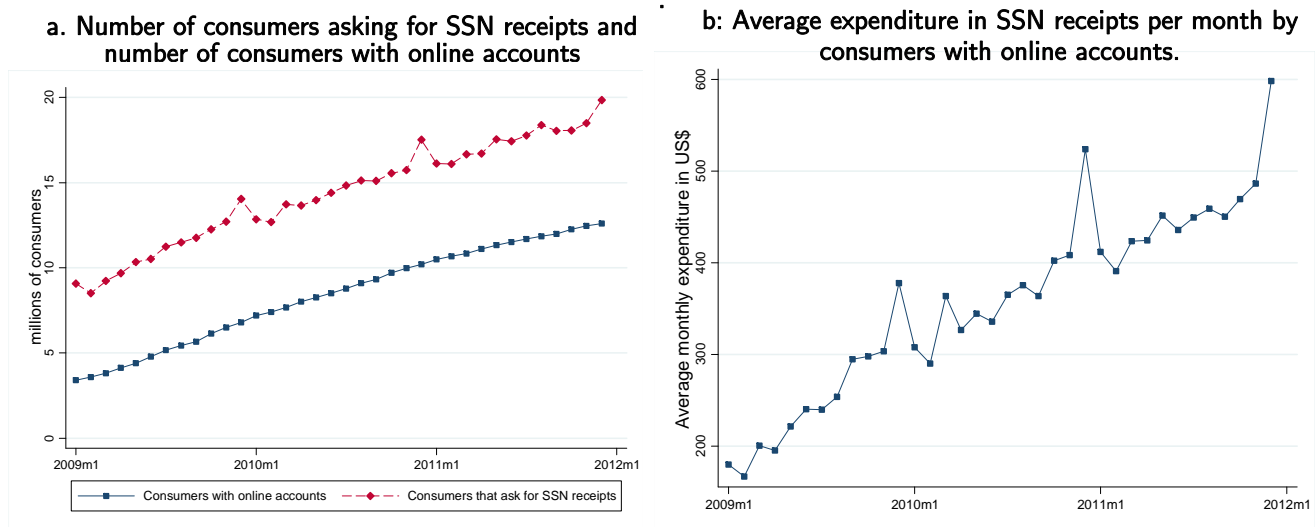
Notes: Figures 1a and 1b are snapshots of an online account example at <https://www.nfp.fazenda.sp.gov.br/login>. The snapshot is from the author's online account. Tabs on the top of both figures can be translated as: Home ("Início"), Check receipts ("Consultar"), Lotteries ("Sorteios"), Charities ("Entidades"), Complaints ("Reclamação"), 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: Number of receipts reported electronically to the tax authority – Timeline of NFP



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) 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: Nota Fiscal Paulista – consumer participation

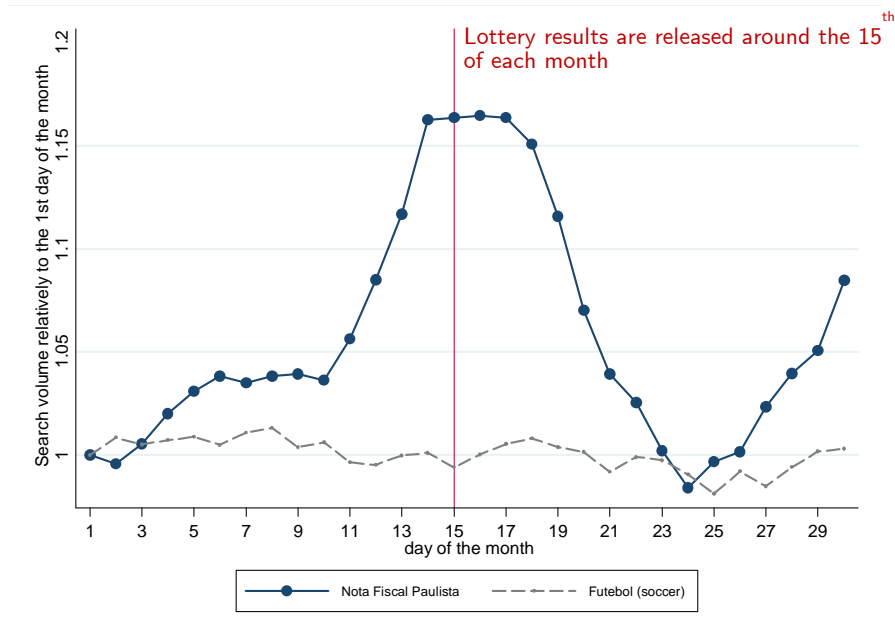


Notes: The dashed line in Figure 3a displays total number of consumers asking for SSN receipts each month, and the solid line is the total number of consumers that had set up online account at the tax authority’s website between Jan. 2009 and Dec. 2011. Any person holding a Brazilian SSN is eligible 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.

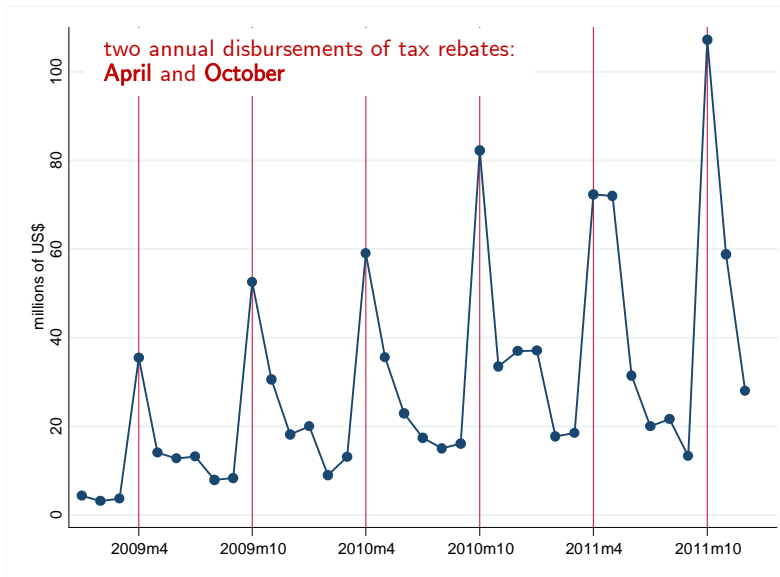


Figure 4: Are consumers paying attention to the schedule of NFP benefits?

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



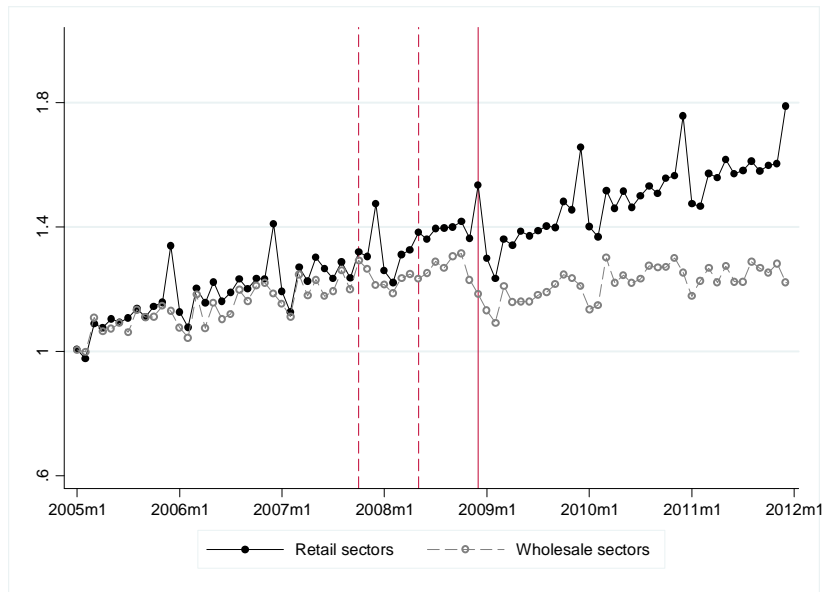
b: Timing of tax rebate disbursements - Rewards requested for deposit in consumers' bank accounts



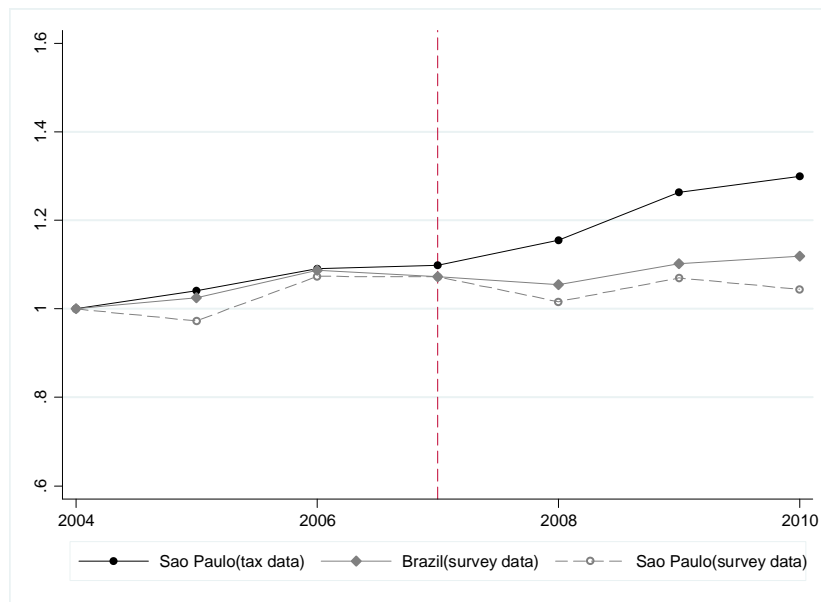
Notes: Figure 4a displays the search volume from Google Trends website for Google searches with terms related to "nfp" or "nota fiscal paulista" or "nota paulista" pooled by day of the month from IPs addresses in the state of Sao Paulo between Oct. 2007 and Dec. 2011. It also displays searches for "futebol" (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 15<sup>th</sup> of each month marked by the solid vertical line. In Figure 4b 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 5: Compliance effect – Retail v. Wholesale

a. Raw data: reported revenue changes



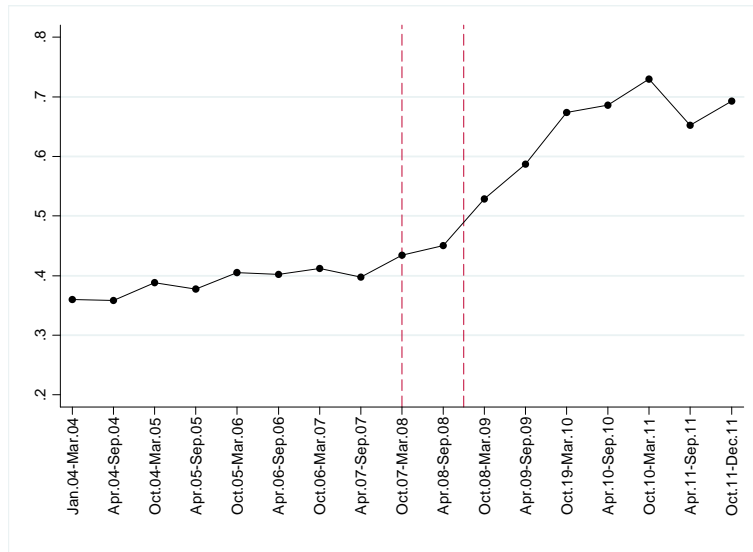
b. Changes in (Retail revenue/wholesale revenue) ratio



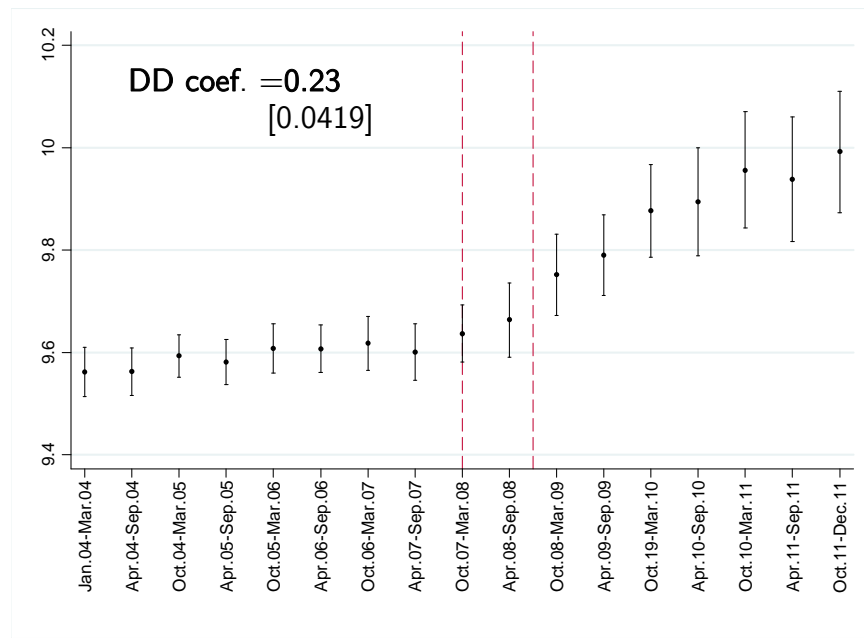
Notes: Figure 5a shows reported revenue changes. Each line is defined by the reported revenue from the establishment sample - described in section II.C.2 - aggregated by Retail or Wholesale scaled by the average monthly reported revenue in 2004 for each group. The figure plots the raw data, so there are spikes around December of each year follows the seasonal variation in consumption due to Christmas shopping. 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 5b 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. Because the national data adds up all revenue from both groups of sectors, in Figure 5b 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.

Figure 6: Reported revenue difference between Retail v. Wholesale

a. Raw difference between retail and wholesale in log reported revenue across 6-month time bins



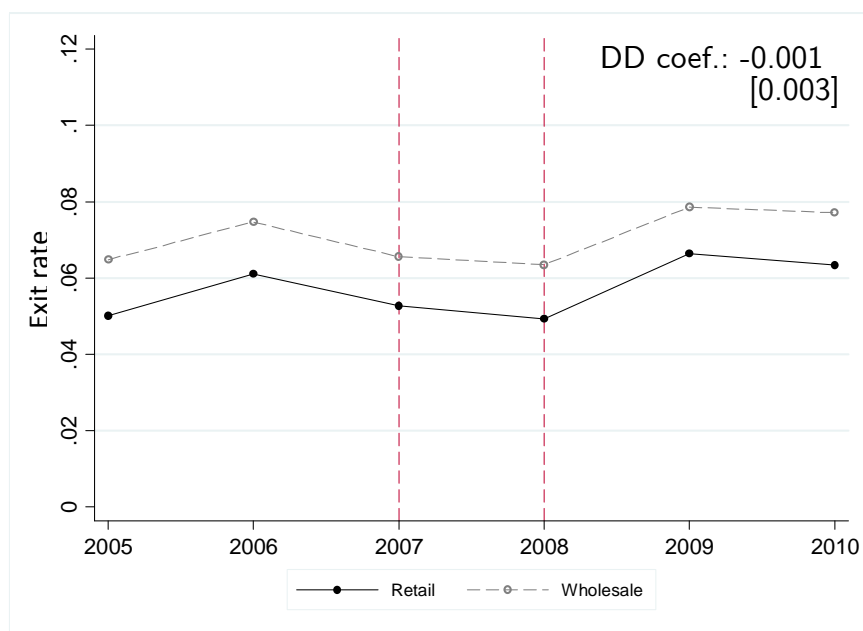
b. DD coefficients for 6-month time bins - Log reported revenue by sector



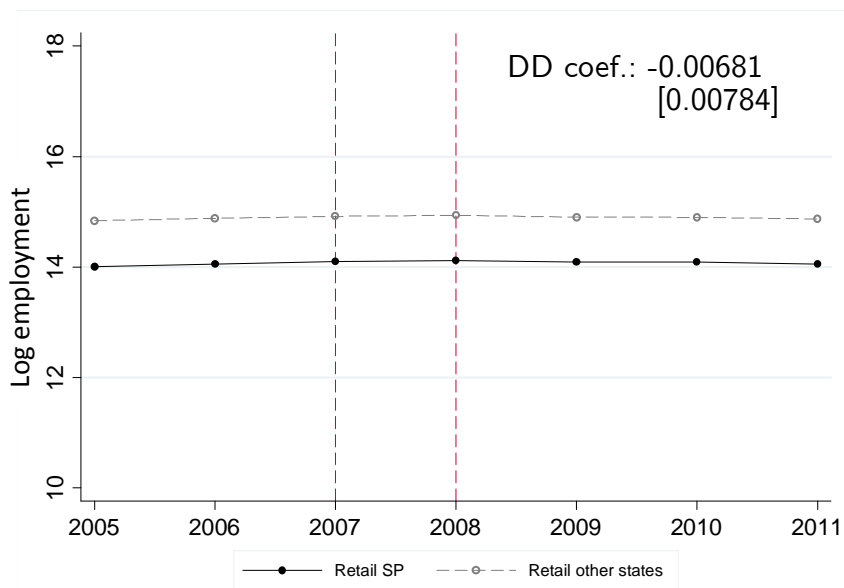
Notes: Figure 6a is the raw difference in log reported revenue between retail and wholesale sectors that is described in Figure 5a pooled across 6-month windows without controls. Figure 6b plots regression coefficients from estimating specification (1) using a sample of 210 sectors between Jan 2004 and Dec 2011. The sector sample has 20,160 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 7: The effect of an increase in the effective tax rate on exit and employment

a. Exit rate: retail vs. wholesale



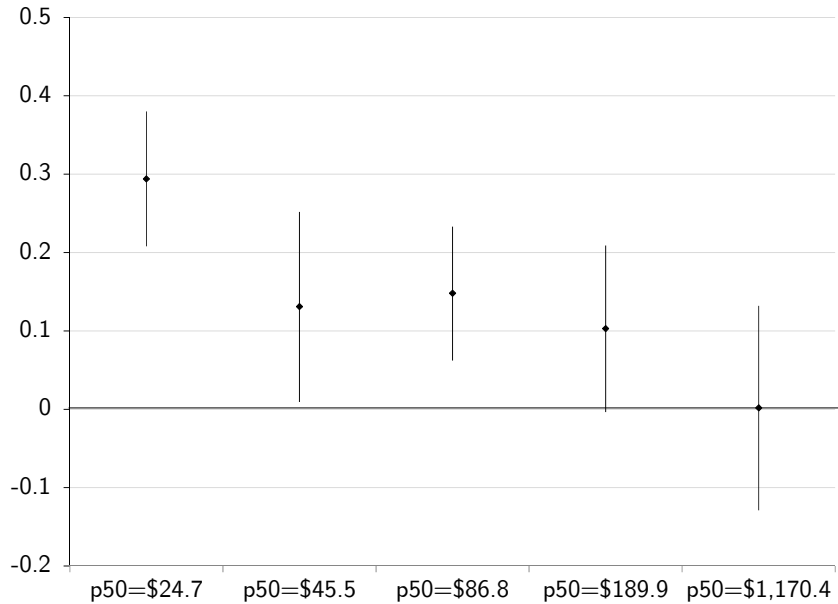
b. Employment changes: retail in Sao Paulo vs. retail in other Brazilian states



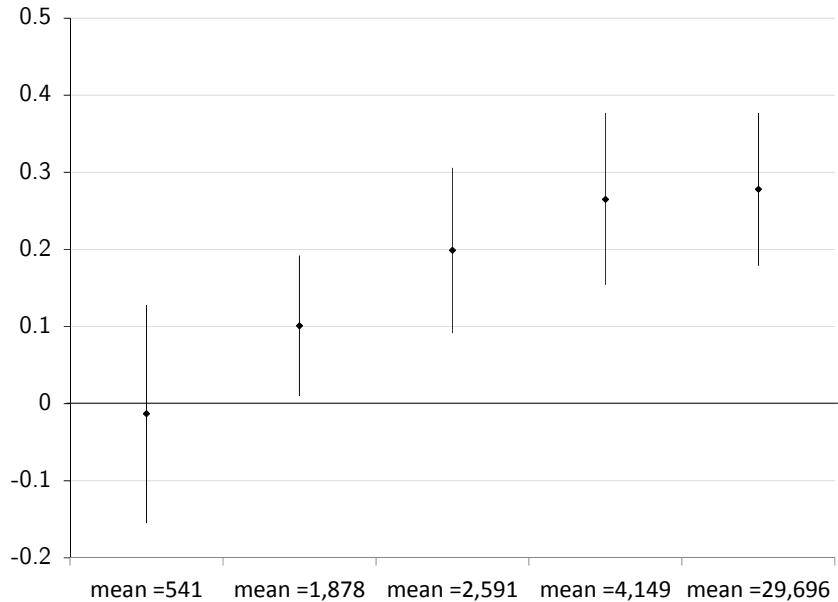
Notes: Figure 7a displays exit rates in retail and wholesale. Exit rate is defined as the total number of exits in year  $t$  and sector  $s$ , divided by the total number of establishments in year  $t-1$  in sector  $s$ . Exit is the year I observe an establishment in the data, and I consider all exits between 2005 and 2010. The data for the exit analysis comprises all establishments in retail and wholesale. Figure 7a 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 7b 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 7b 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 8: Mechanisms and Heterogeneous effects of NFP

a. Percentage changes in reported revenue by receipt value quintile

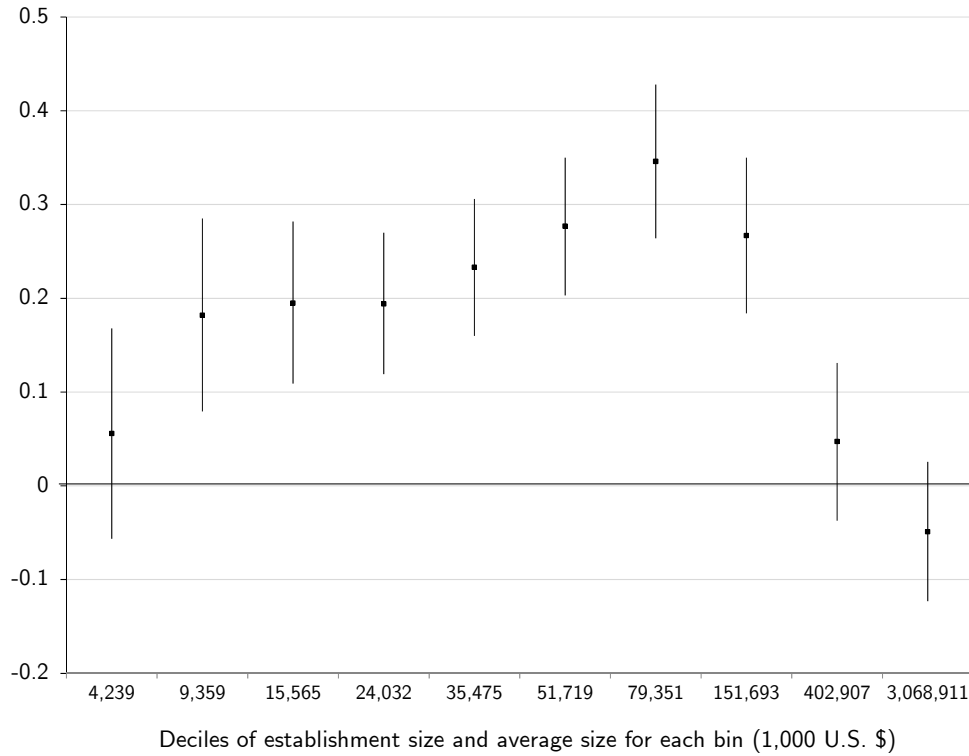


b. Percentage changes in reported revenue by transaction volume quintile



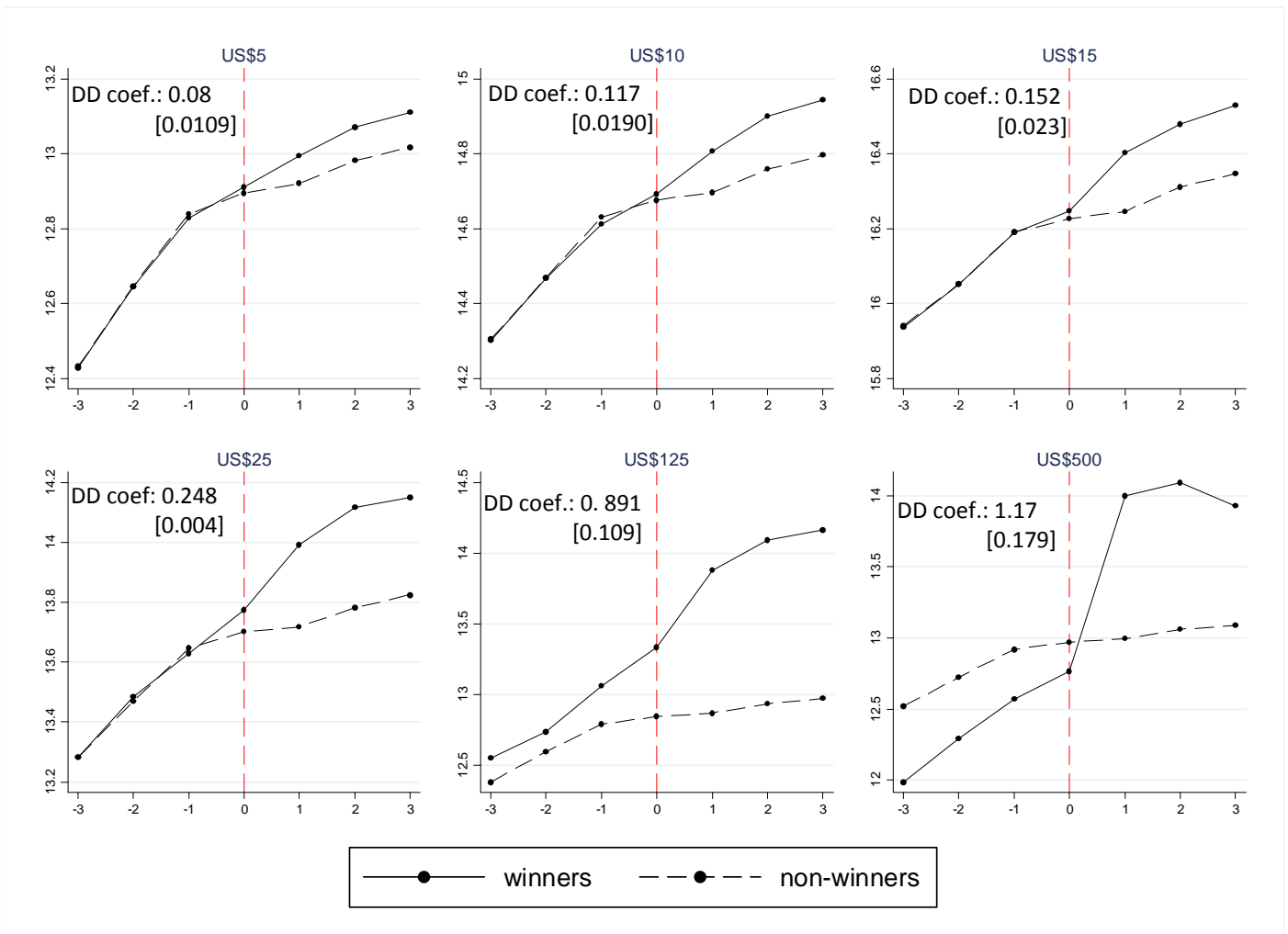
Notes: Figures 8a and 8b plot the coefficients and a 85% confidence interval from estimating equation (8) in section IV.D. Figure 8a plots the effects of the program by sector quintiles of the receipt value distribution. Receipt value is defined by ranking retail sectors by receipt value as described in Section IV.d. The x-axis displays the values of the median sector receipt value in each bin. Figure 8b 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 the reported revenue by an establishment scaled by average reported revenue the year before the program. The dependent variable is Winsorized at the 85<sup>th</sup> percentile. In order to control for establishment size effects the regression includes a 5<sup>th</sup> order polynomial interacted with the DD variable.

Figure 9: Heterogeneity by establishment size



Notes: The figure plots coefficients and the 95% confidence intervals from estimating Equation (10). The x-axis shows average size in 1,000 US\$ in each decile bin from the establishment size distribution. Size is defined as the total revenue 12 months before program implementation in Oct 2007. Standard errors are clustered at the 7-digit sector classification level (210 clusters). The sample is composed of monthly observations between Jan 2004 and Dec 2011 of establishments classified as retail or wholesale as described in Section II. The dependent variable is the reported revenue of an establishment, scaled by average reported revenue the year before the program. The dependent variable is Winsorized at the 95<sup>th</sup> percentile.

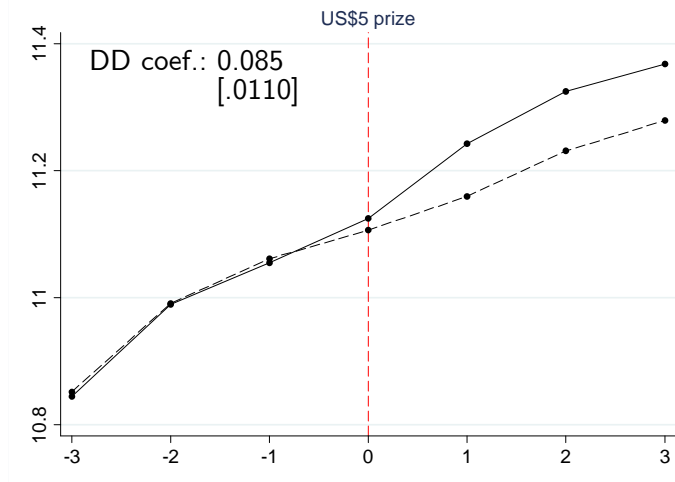
Figure 10: The effect of different sizes of lottery wins on the number of SSN receipts a consumer asks for



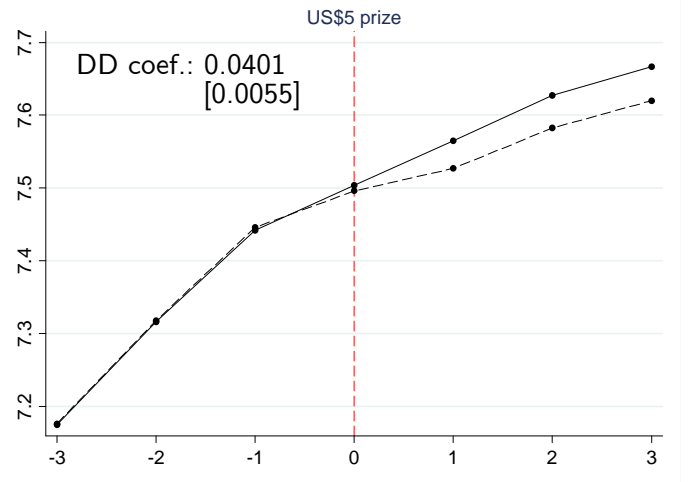
Notes: The x-axis is the number of months since the individual won or did not win a lottery. The y-axis in each graph is the average number of SSN receipts consumers ask for in a month. Before taking the averages I create bins for each possible number of lottery ticket holdings from 1-40 tickets in each 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 11: The effect of lottery wins on consumer participation

a. The effect of winning a 5-dollar prize on the number of receipts among consumers that won a prize before

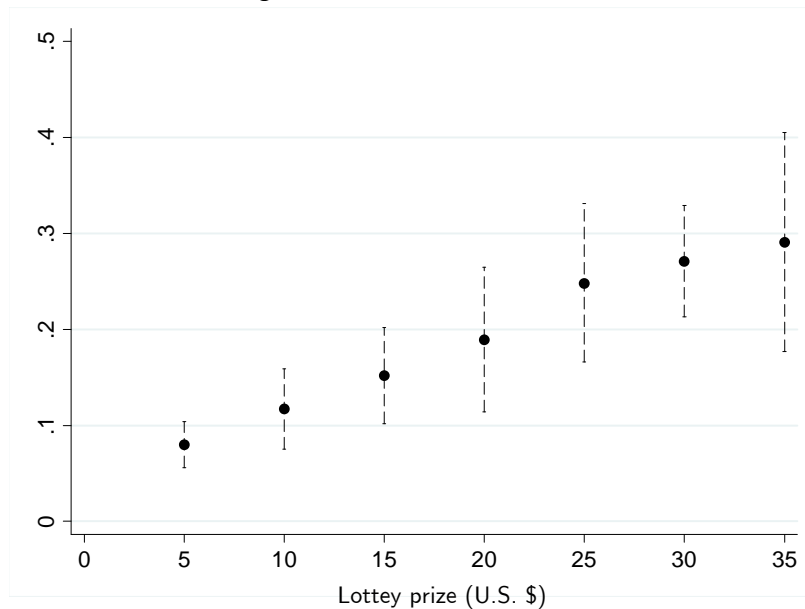


b. The effect of winning a 5-dollar prize on the number of different businesses in which consumers ask for receipts



—●— winners    - - -●- - - non-winners

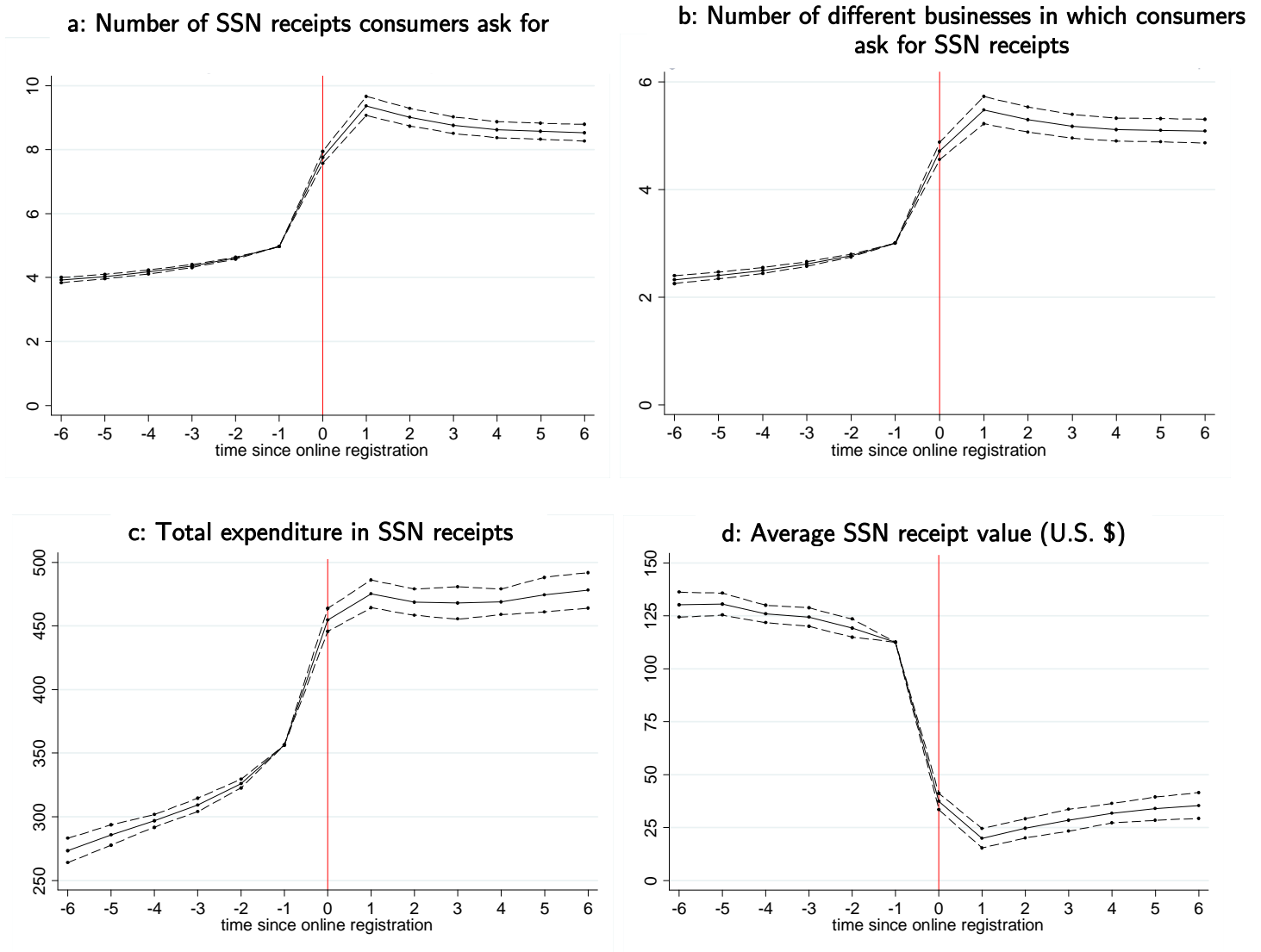
c. The estimated effect of winning different sizes of lotteries on the number of SSN receipts



Notes: The x-axis in Figure 11a and Figure 11b is the number of months since the individual won or did not win a lottery. Figure 11a 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 11b 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 11. 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 11c 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 12: Effect of enrolling online on consumer participation



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 99<sup>th</sup> percentile value.

## Appendix: Data on Establishments and Consumers

Section II.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 Section IV.E.

### 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.<sup>88</sup>

*Employment sample.* The employer-employee data RAIS/CAGED covers all formal establishments that have at least one employee. Because the Sao Paulo tax data is de-identified, I cannot match the Sao Paulo tax data with the RAIS/CAGED data. As I discussed in section II.C, I apply in this data similar sample restrictions as in the *establishment sample*: the establishment must have reported at least one employee at any point before 2004. Moreover, all the variables defined for 7-digit sector in the Sao Paulo tax data were mapped into 5-digits sector definitions in the RAIS/CAGED data.<sup>89</sup>

### 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 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

---

<sup>88</sup>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.

<sup>89</sup>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 7b, I include motor vehicles trade in retail. The results do not change if I exclude these sectors.

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 II.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.

### A.3. 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.E I merge this data with the *receipts data* described in section II.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*.

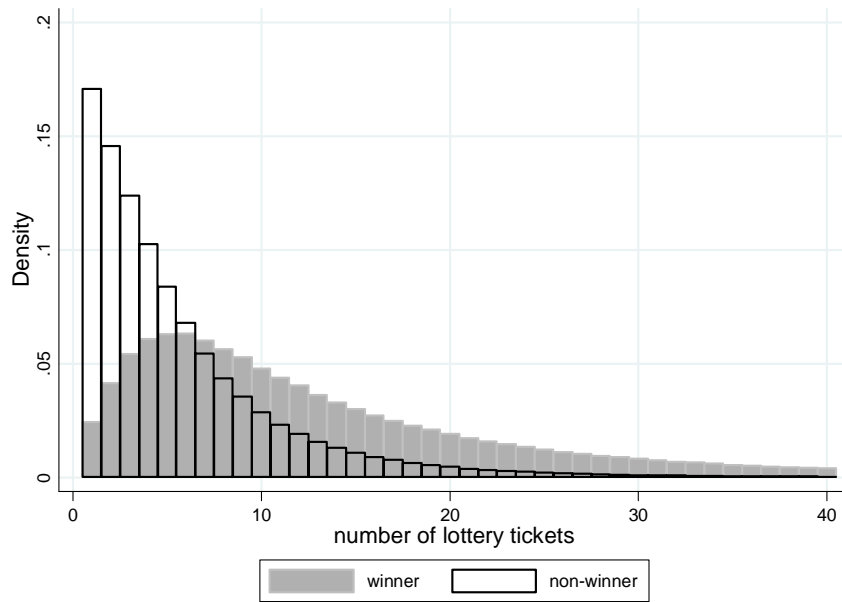
*Re-weighting.* Since the number of lottery tickets is determined by the total value of a consumer's purchase 4 months before the lottery draw, the more a consumer participates in the program by asking for receipts, the higher are the odds she will get a prize in a given lottery. Therefore, it is important 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.E, I use a re-weighting method based on DiNardo, Fortin, and Lemieux (1996) to flexibly control for the number of lottery tickets individuals hold. This method ensures that I use the random component of the lottery by matching the two groups based on the odds of winning prize, i.e., on the number of lottery tickets a consumer holds.

I perform this re-weighting exercise separately for each lottery win I study in Section IV.E. 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). The pool of non-winners is the same across the datasets I create for each prize level, but they are re-weighted differently depending on the prize I am considering because the winners group of different prize levels may have slightly different distributions of lottery ticket holdings.

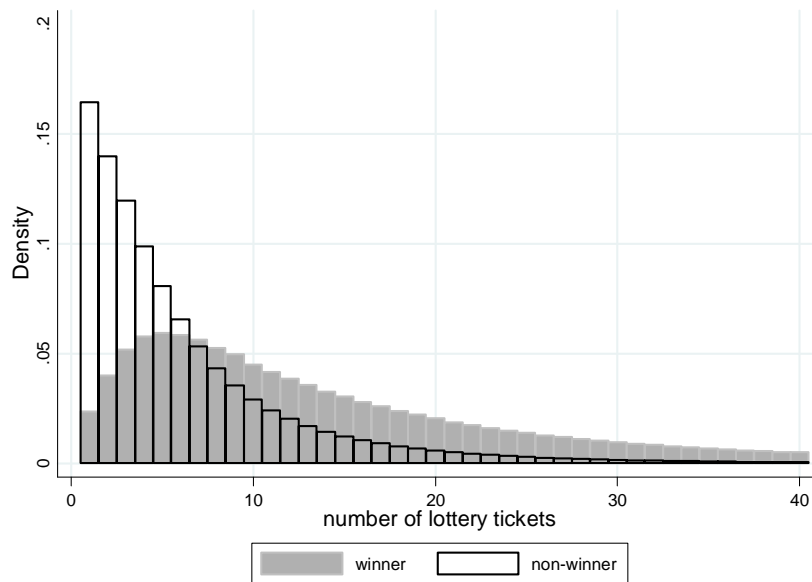
Figure A shows two examples of the distribution of lottery ticket holdings among winners and non-winners in monthly lotteries. The two examples look very similar, and it is clear that the winner group typically holds more lottery tickets. I create bins for each possible number of lottery ticket holdings 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. 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.

Figure A: Histogram of lottery tickets – two examples

i. Distribution of lottery tickets between lottery winners and non-winners in the December 2009 Lottery



ii. Distribution of lottery tickets between lottery winners and non-winners in the 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 generated 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.