ONLINE APPENDIX A: ADDITIONAL TABLES AND FIGURES

TABLE A1: ROBUSTNESS TOP CODING AND STANDARD ERRORS – MAIN ESTIMATES

Panel A: Log of Reported revenue - Firm-level regressions

	p99	p99.9	p95
DD (Post Oct 07 * Retail dummy)	0.254	0.291	0.2
s.e. clustered by firm	[0.0722]	[0.107]	[0.0521]
s.e. clustered by sector	[0.0360]	[0.0690]	[0.0251]
firm FE	x	x	x
Time FE	x	x	x
obs	1,035,268	1,035,268	1,035,268
Adjusted R-squared	0.907	0.907	0.88

Panel B: Log of Reported revenue - Sector-level regressions

	p99	p99.9	p95
DD (Post Oct 07 * retail)	0.208	0.186	0.249
s.e. clustered by sector	[0.0411]	[0.0488]	[0.0340]
sector FE	x	x	x
Time FE	x	x	x
obs	20,352	20,352	20,352
Adjusted R-squared	0.982	0.976	0.987

Note: *Panel A* displays the robustness of the coefficient reported in Table 2 column [1] from regressions using the firm-level data described in Section 3. The variable DD is defined as the interaction between a dummy for retail sectors (*Retail dummy*) and a dummy that equals 1 for time periods after Oct 2007 (*Post Oct 07*). The dependent variable is log of reported revenue by firm, and the data is collapsed into two periods: before and after Oct. 2007. Time and firm fixed effects are included in all regressions. The regressions are dollar- weighted (each observation is weighted by the pre-policy reported revenue) such that each observation contributes to all regression estimates according to its economic scale to best approximate the aggregate effect. The columns indicate the threshold used to winsorize the dependent variable in order to mitigate the influence of outliers. The results for p99 that are reported in the paper are presented first, then p99.9 and p.95 are also shown. Each column reports standard errors clustered by sector and clustered by firms. *Panel B* displays the robustness of the DD coefficient reported in Figure 2b estimated using the sector-level monthly data. The dependent variable is log of total reported revenue by sector. Time and sector fixed effects are included in all regressions. The columns indicate the threshold used to winsorize the dependent variable in order to mitigate the influence of outliers. The results for p99 that are reported in the paper are presented first, then p99.9 and p.95 are also shown. Standard errors are clustered by sector. Significance levels *** 1%, ** 5%.

TABLE A2: ROBUSTNESS TOP CODING AND STANDARD ERRORS – HETEROGENEITY ANALYSIS

	Log of Reported Revenue											
	p99	p99.9	p95	p99	p99.9	p95	p99	p99.9	p95	p99	p99.9	p95
DD * Large firms	0.253	0.292	0.191									
s.e. clustered by sector	[0.0732]	[0.107]	[0.0524]									
s.e. clustered by firm	[0.0363]	[0.0693]	[0.0254]									
DD * Small firms	0.350	0.350	0.470									
s.e. clustered by sector	[0.0511]	[0.0827]	[0.0453]									
s.e. clustered by firm	[0.0287]	[0.0615]	[0.0225]									
DD * High volume of different consumers				0.246	0.275	0.202						
s.e. clustered by sector				[0.0705]	[0.108]	[0.0508]						
s.e. clustered by firm				[0.0335]	[0.0727]	[0.0242]						
DD * Low volume of different consumers				0.0329	0.0436	0.0712						
s.e. clustered by sector				[0.0919]	[0.115]	[0.0649]						
s.e. clustered by firm				[0.0380]	[0.0701]	[0.0251]						
DD * High volume of transactions							0.253	0.289	0.208			
s.e. clustered by sector							[0.0707]	[0.106]	[0.0516]			
s.e. clustered by firm							[0.0335]	[0.0717]	[0.0241]			
DD * Low volume of transactions							0.0181	0.00226	0.0584			
s.e. clustered by sector							[0.0865]	[0.130]	[0.0614]			
s.e. clustered by firm							[0.0391]	[0.0881]	[0.0268]			
DD * High value of transactions										0.0969	0.153	0.0938
s.e. clustered by sector										[0.0689]	[0.121]	[0.0531]
s.e. clustered by firm										[0.0424]	[0.0868]	[0.0274]
DD * Low value of transactions										0.285	0.319	0.224
s.e. clustered by sector										[0.0754]	[0.123]	[0.0558]
s.e. clustered by firm										[0.0325]	[0.0655]	[0.0244]
3rd-order polynomial of firm size * DD				x	x	x	x	x	x	x	x	x
Time FE	x	x	x	x	x	x	x	x	x	x	x	x
Firm FE	x	x	x	x	x	x	x	x	x	x	x	x
Observations	1,035,356	1,035,356	1,035,356	1,035,356	1,035,356	1,035,356	1,035,356	1,035,356	1,035,356	1,035,356	1,035,356	1,035,356
Adjusted R-squared	0.907	0.906	0.88	0.908	0.907	0.881	0.908	0.907	0.881	0.908	0.907	0.881

Note: The table displays the robustness of the coefficients columns [2] to [5] in Table 2. The variable DD is defined as the interaction between a dummy for retail sectors (*Retail dummy*) and a dummy that equals 1 for time periods after Oct 2007 (*Post Oct 07*). The dependent variable is log of reported revenue by firm, and the data is collapsed into two periods: before and after Oct. 2007. Time and firm fixed effects are included in all regressions. The regressions are dollar- weighted (each observation is weighted by the pre-policy reported revenue) such that each observation contributes to all regression estimates according to its economic scale to best approximate the aggregate effect. The columns indicate the threshold used to winsorize the dependent variable in order to mitigate the influence of outliers. The results for p99 that are reported in the paper are presented first, then p99.9 and p.95 are also shown. Each column reports standard errors clustered by sector and clustered by firms. Definition of large vs small firms; high vs low volume of different consumers; high vs low volume of transactions; and high vs low value of transactions are the same as in Table 2 and Section 4 of the paper. See notes of Table 2 for more details.

Table A3: Robustness Top coding and Standard Errors – Tax liability and Reported expenses

Panel A: Tax sample - Firm-level regressions

	Log of Reported Revenue			Log of Tax Liability			Positive tax liability
	p99	p99.9	p95	p99	p99.9	p95	
DD (Post Oct 07 * Retail)	0.311	0.295	0.266	0.316	0.284	0.420	0.0434
s.e. clustered by sector	[0.151]	[0.139]	[0.0741]	[0.135]	[0.117]	[0.116]	[0.0350]
s.e. clustered by firm	[0.125]	[0.127]	[0.0518]	[0.137]	[0.149]	[0.0638]	[0.00597]
Observations	167,110	167,110	167,110	133,950	133,950	133,950	167,110
Adjusted R-squared	0.851	0.868	0.839	0.88	0.88	0.86	0.801

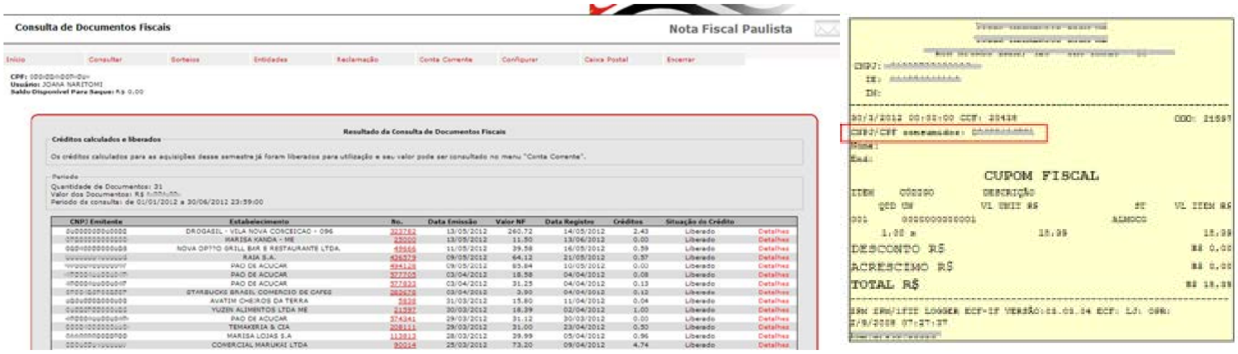
Panel B: Tax sample - Sector level regressions

	Log of Reported Revenue			Log of Tax Liability		
	p99	p99.9	p95	p99	p99.9	p95
DD (Post Oct 07 * Retail)	0.280	0.253	0.282	0.259	0.234	0.319
s.e. clustered by sector	[0.0665]	[0.0755]	[0.0502]	[0.106]	[0.106]	[0.104]
Observations	5,088	5,088	5,088	5,088	5,088	5,088
Adjusted R-squared	0.983	0.976	0.99	0.956	0.954	0.967

Panel C: Expenses, output and value added - VAT firms

	Log of Reported Revenue			Log of Reported Inputs			Log of Reported Value Added			Positive Value Added
	p99	p99.9	p95	p99	p99.9	p95	p99	p99.9	p95	
DD (Post Oct 07 * Retail)	0.363	0.351	0.314	0.302	0.345	0.268	0.387	0.258	0.354	0.0192
s.e. clustered by sector	[0.108]	[0.139]	[0.0931]	[0.107]	[0.151]	[0.0970]	[0.147]	[0.129]	[0.102]	[0.0176]
s.e. clustered by firm	[0.0824]	[0.110]	[0.0540]	[0.0833]	[0.115]	[0.0611]	[0.105]	[0.156]	[0.0592]	[0.0153]
Observations	88,422	88,422	88,422	88,422	88,422	88,422	70,845	70,845	70,845	88,422
Adjusted R-squared	0.87	0.89	0.87	0.85	0.88	0.83	0.90	0.91	0.90	0.71

Note: Table A3 displays the main coefficients form regressions described in Section 5. The variable DD is defined as the interaction between a dummy for retail sectors (*Retail*) and a dummy that equals 1 for time periods after Oct 2007 (*Post Oct 07*). *Panel A* reports the results for a sample of firms that are in sectors where there is little tax withholding and, therefore, the firm-level reported tax liabilities in the data best approximates their own tax liabilities (see Section 2 and Appendix B for more details). The results in *Panel A* show the robustness of the results from Table 3 *Panel A*. The results in *Panel B* show the robustness of the results displayed in Figure 6a using sector-level monthly data. The results in *Panel C* reports the robustness of the results from Table 3 *Panel B* for a sample of firms that are always registered in the VAT. In all panels, the columns indicate the threshold used to winsorize the dependent variable in order to mitigate the influence of outliers. The results for p99 that are reported in the paper are presented first, then p99.9 and p95 are also shown. Standard errors clustered by sector and clustered by firms are reported for Panels A and C, and standard errors clustered by sector are reported in Panel B. For more details, see the notes of Table 3 and the notes of Figure 6.



Consulta de Documentos Fiscais

Nota Fiscal Paulista

CPF: [redacted]
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$ [redacted]
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	
[redacted]	DROGASIL - VILA NOVA CONCEICAO - 096	323782	13/05/2012	260.72	14/05/2012	2.43	Liberado	Detalhes
[redacted]	MARISA KANDA - ME	25000	13/05/2012	11.50	13/06/2012	0.00	Liberado	Detalhes
[redacted]	NOVA OPTTO GRILL BAR E RESTAURANTE LTDA.	49666	11/05/2012	39.58	16/05/2012	0.59	Liberado	Detalhes
[redacted]	RAIA S.A.	426579	09/05/2012	64.12	21/05/2012	0.57	Liberado	Detalhes
[redacted]	PAO DE ACUCAR	494128	09/05/2012	85.84	10/05/2012	0.00	Liberado	Detalhes
[redacted]	PAO DE ACUCAR	577705	03/04/2012	18.58	04/04/2012	0.08	Liberado	Detalhes
[redacted]	PAO DE ACUCAR	577833	03/04/2012	31.25	04/04/2012	0.13	Liberado	Detalhes
[redacted]	STARBUCKS BRASIL COMERCIO DE CAFES	282678	03/04/2012	3.90	04/04/2012	0.12	Liberado	Detalhes
[redacted]	AVATIM CHEIROS DA TERRA	5839	31/03/2012	15.80	11/04/2012	0.04	Liberado	Detalhes
[redacted]	YUZEN ALIMENTOS LTDA ME	21597	30/03/2012	18.39	02/04/2012	1.00	Liberado	Detalhes
[redacted]	PAO DE ACUCAR	374341	29/03/2012	31.12	30/03/2012	0.00	Liberado	Detalhes
[redacted]	TEMAKERIA & CIA	208111	29/03/2012	31.00	23/04/2012	0.50	Liberado	Detalhes
[redacted]	MARISA LOJAS S.A	113813	28/03/2012	39.99	05/04/2012	0.96	Liberado	Detalhes
[redacted]	COMERCIAL MARUKAI LTDA	90014	25/03/2012	73.20	09/04/2012	4.74	Liberado	Detalhes

[redacted]

CNPJ: [redacted]
IE: [redacted]
IM:

30/3/2012 00:00:00 CCF: 20438 COO: 21597
CNPJ/CPF consumidor: [redacted]
Nome:
End.:

CUPOM FISCAL

ITEM CÓDIGO DESCRIÇÃO
QTD UN VL UNIT R$ ST VL ITEM R$
001 0000000000001 ALMOCO
1,00 x 18,39 18,39
DESCONTO R$ R$ 0,00
ACRESCIMO R$ R$ 0,00
TOTAL R$ R$ 18,39

[illegible] LOGGER ECF-IF VERSÃO:[illegible] ECF: [illegible]
[illegible] 07:27:37
[redacted]

a. Online Account **b. Receipt**

FIGURE A1: ONLINE ACCOUNT AND *NOTA FISCAL PAULISTA* RECEIPT

Note: Figures a and b are snapshots of an online account example at https://www.nfp.fazenda.sp.gov.br. The snapshot in Figure a is from the author's online account. Tabs on the top of the figure can be translated (from left to right) as: *Home*, *Queries*, *Lotteries*, *Charities*, *Complaints*, *Current Account*, *Settings*, *Inbox*, *Sign out*. 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. The interface of the account is similar to a credit card statement with a list of all receipts, the issuing date, total value of each receipt, tax rebate, and a link to the details of each receipt. Figure b shows the receipt if one clicks on the last column in Figure a for *details* of one of the purchases listed. The receipt has a field to fill in the consumers' SSN – as highlighted in Figure b.

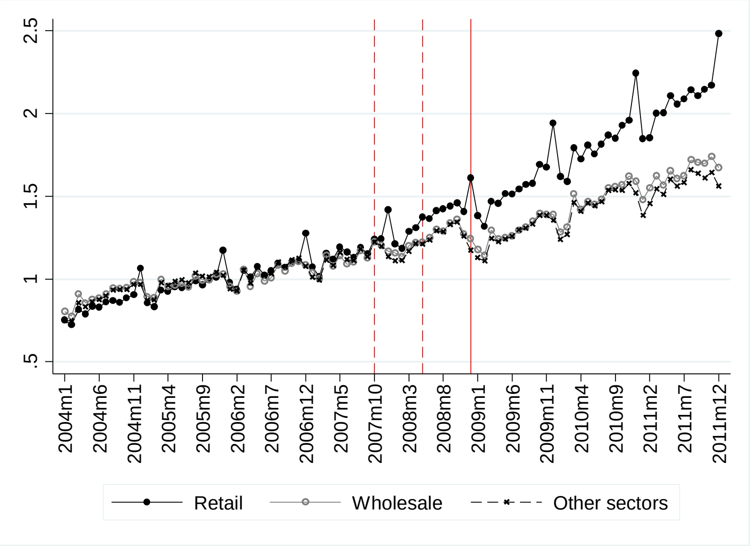


FIGURE A2: REPORTED REVENUE EFFECT – RETAIL VS. WHOLESALE VS. OTHER SECTORS

Note: Figure A2 is similar to Figure 2a: it shows reported revenue changes for retail and wholesale sectors, but it also adds all the remaining sectors as a third category. Each line is defined by the reported revenue by all firms aggregated by retail or wholesale or other sectors scaled by the average monthly reported revenue before Oct. 07 for each sector group. The figure plots the raw data, so there are spikes around December of each year follows the seasonal variation in consumption. The vertical lines highlight the key dates for the implementation of the NFP program: phase-in of sectors begins in Oct.07 and ends in May.08, and the first lottery based on the purchases with SSN receipts was introduced in Dec.2008.

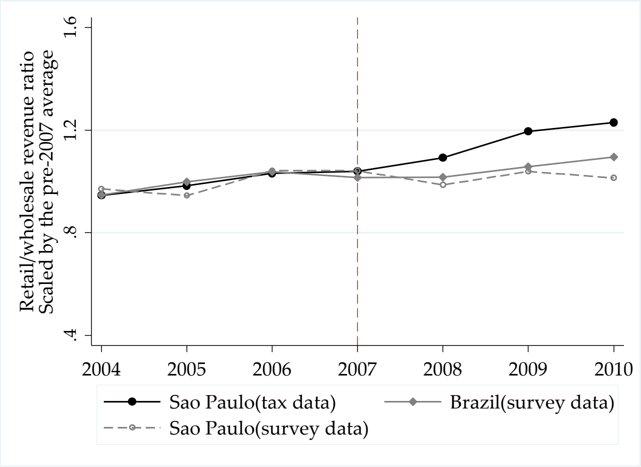


FIGURE A.3: COMPARING SAO PAULO WITH BRAZIL - CHANGES IN REVENUE RATIO

Note: The figure 3 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 (excluding Sao Paulo), 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 pre-2007 retail-wholesale revenue ratio. The national ratio is based on the total gross revenue from sales, and retail revenue considers retail and motor-vehicles trade. Figure 3a compares changes in the revenue ratio of retail to wholesale, $r = \frac{retail\ revenue}{wholesale\ revenue}$, from the Sao Paulo administrative data to changes in the same ratio from the census survey. 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. The retail-wholesale revenue ratio was calculated from aggregate tables of the survey.

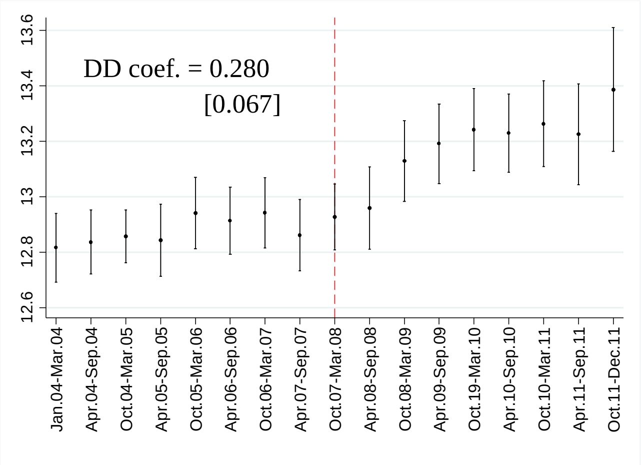


FIGURE A4: REPORTED REVENUE EFFECT – TAX SAMPLE

Note: Figure A4 is similar to Figure 2b: plots regression coefficients from estimating specification (5) using a sample of sectors for between Jan 2004 and Dec 2011. The difference in differences (DD) coefficient displayed in the figure is estimated using the specification (6) 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. The difference with respect to Figure 2b is that it restricts attention to a sample of firms that are in sectors where there is little tax withholding and, therefore, the firm-level reported liabilities in the data best approximates their own tax liabilities (see Online Appendix B for more details).

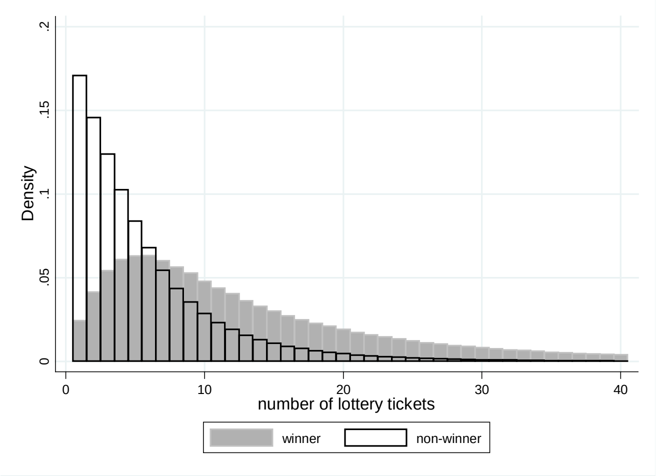
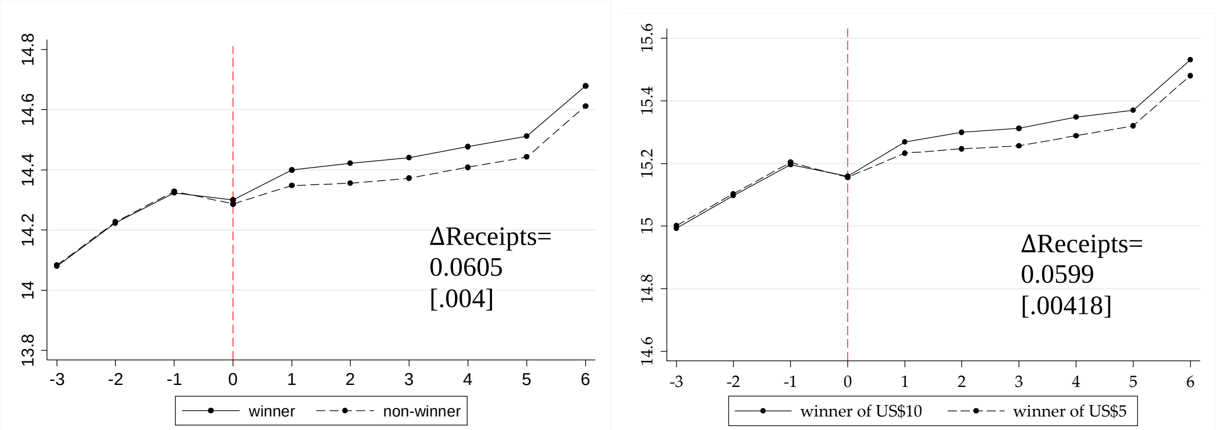


FIGURE A.5: LOTTERY TICKETS DISTRIBUTION – DECEMBER 2009

Notes: Figures A.5. shows a histogram for the number of lottery tickets winners and non-winners hold in December 2009 as an example for the re-weighting of observations for the analysis in Section 4 and described in Online Appendix B. 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. 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. The graphs were constructed from the consumer sample described in Section 2 and Online Appendix B.



a: The effect of a 5-dollar win (R$10) on consumers that won a lottery in the past

b. The effect of a 5-dollar difference in prize (winning R$10 vs R$20)

FIGURE A6: THE EFFECT OF A SMALL LOTTERY WIN – WINNERS ONLY

Note: The graphs show total raw number of receipts consumers ask by month aggregating all lotteries from June 2009 to June 2009. The x-axis is the number of months since the consumer participated in a lottery. Figure A6a shows the effect of a U.S. $5 lottery win (R$10) for consumers that have won a lottery once before, in which case they have already verified that the program works as advertised. Figure A6b compares winners of two different prize amounts that differ by U.S. $5 (R$20 vs R$10). In this graph, both sets of consumers won a prize, so the difference is driven by the size of the prize and not a discouragement effect from not winning. Before taking the averages in each case, I create bins for each possible number of lottery ticket holdings from 1-40 tickets in each monthly lottery for 12 lotteries between June 2009 and June 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 (see Online Appendix B for more details). The DD coefficient displayed in each graph is based on estimating specification (10) using the micro-data and the lottery ticket weights. Standard errors are clustered by lottery draw.

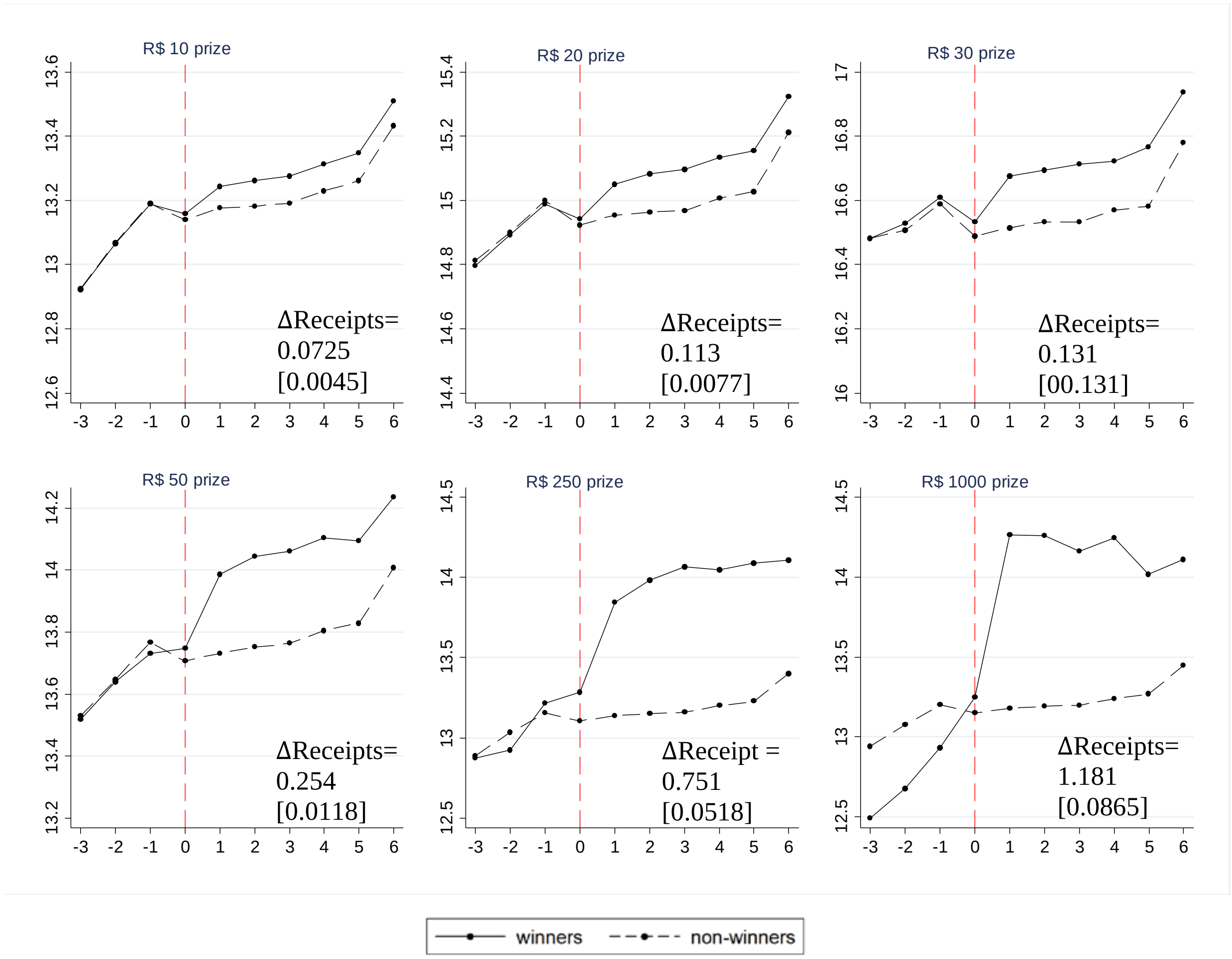


FIGURE A7: THE EFFECT OF LOTTERY WINS ON THE NUMBER OF RECEIPTS

Note: The graphs show the total raw number of receipts consumers ask by month aggregating all lotteries from June 2009 to June 2011. The x-axis is the number of months since the consumer participated in the lottery. The figures show the results for different cash prizes. In each of the lotteries there are 1,407,394 prizes of R$10, 76,303 prizes of R$20, 15,000 prizes of R$50, 1,000 prizes of R$250, and 300 prizes of R$1000 (US$1=R$2). 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 R$20 or R$30 by winning a combination of a R$10 and/or R$20 prizes. Before taking the averages in each case, I create bins for each possible number of lottery ticket holdings from 1-40 tickets in each monthly lottery for 24 lotteries between June 2009 and June 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 (for more details see Online Appendix B). The DD coefficient displayed in each graph is based on estimating specification (10) using the micro-data and the lottery ticket weights. Standard errors are clustered by lottery draw.

ONLINE APPENDIX B: DATA AND RE-WEIGHTING

This appendix provides additional information on the datasets and variables discussed in Section 2, and on the re-weighting exercise from Sections 4: whistle-blower effect and lottery effect.

B.1. Firm Data

Firm panel. Tha data was de-identfied, and a scrambled identifier was created for each establishment and firm. For the firm-level analysis in the paper, I aggregate all the data from establishments by firm. This is possible because a dataset based on the registry of firms allowed the link between the scrambled ids of establishments and firms.

The first data source are tax forms from establishments in the tax regime RPA ("Regime Periódico de Apuração") that requires establishments to file monthly a detailed tax return of all sales called GIA/ICMS ("Guia de Informação e Apura ção do ICMS") to assess the total VAT due by the establishment in a given month. These tax returns include how much of sales and input transactions fall into different categories of the tax code (e.g., different rates and withholding). These are called CFOP ("Código Fiscal de Operações e Prestações). Due to confidentiality concerns, the data available to this project does not include all the transaction details by each tax code. A few aggregations of such codes were added to the data as I explain below.

The second source of data is composed by tax forms from establishments in a simplified tax regime called SIMPLES. As is common in VAT systems across the world, there is a threshold below which firms do not pay taxes over the value added. In the period of analysis, 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. [1]

Sector. The sector definition used in the analysis is a 7-digit CNAE (Classificação Nacional de Atividades Econômicas) based on a snapshot of the registry of establishments in Sao Paulo in 2011. It was not feasible to rebuild the registry as of Oct.2007. However, the policy did not generate any incentive to change registration from a retail CNAE to a wholesale CNAE, so any change in CNAE since the policy implementation is likely to be orthogonal to the main variation used to identify the impact of the policy.

[1] 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 system changes did not generate mechanical changes in reporting.

Because firms are the unit of analysis in the micro-data, and sector is defined by establishment, I assigned CNAEs to firms based on the establishment CNAE. 94.08% of firms have a single establishment and among firms that are multi-establishment, most of them had multiple establishments in the same CNAE, so this process was straightforward for a vast majority of cases. 1.35% of firms have more than one establishment registered in more than one CNAE. For these firms, I assigned to the firm the CNAE of the establishment that is registered as the firm's headquarter. For 0.07% of firms, there was no headquarter indicated in the registry of Sao Paulo. In these cases, I assigned the CNAE of the establishment with most revenue during the period of analysis.

Tax liabilities measurement. As explained in Section 2 of the paper, there are measurement problems in the data for firms' tax liabilities. The variable I observe in the data is the amount of tax a firm is due to remit to the tax authority, not their tax liability. There could be substantial differences between these two quantities. An important driver of this difference is tax withholding within the VAT chain: part of the tax that is due by an firm is withheld and remitted by a upstream or downstream trade partner. If I had access to all the transaction values itemized by tax code in firms' tax returns, it could be possible, in principle, to recover the tax liability of each firm. However, this detailed data from tax returns was not available to this study for confidentiality concerns as I mentioned above.

This measurement problem introduces mechanical drops and increases in tax liabilities by firms that are difficult to control for as withholding rules are based on products and I do not observe products. Yet, there are some sectors that are less affected by withholding than others. Therefore, I restrict attention to all firms in a set of sectors with little withholding for which the reported tax due best approximates tax liabilities.

To give a concrete example: products like gasoline and some pharmaceuticals have upstream withholding policies, i.e., the producers or distributors withholds the tax for the entire supply chain. For firms that buy and sell such products downstream, the tax remitted to the government will depend on changes in withholding policies and the composition of goods that they sell or buy. Therefore, the idea is to remove from the data drugstores or gas stations as a relevant share of the trades in those sectors are affected by withholding.

In addition, backing out the tax liability from reported revenue and reported inputs is not straightforward because, depending on the tax code that applies to a given transaction, the tax liability it generates can be different. The VAT system is a credit-invoice method. With a single rate and no exemptions, it could be possible to back out the liability by simply looking at reported revenue minus reported inputs. However, this is not the case in Sao Paulo as there are multiple rates and exemptions.

To identify sectors less affected by withholding I proceed as follows: for firms in the VAT, which submit more detailed tax returns, I obtained an aggregation of the total values of input and output transactions that are in tax codes related to withholding. I aggregate these firms by sector, and calculate how much of total inputs and sales transactions are affected by withholding during the period of analysis. Then I restrict attention to sectors for which neither the input or output transactions affected by withholding represent more than 1% of the total input or output reported by VAT firms in those sectors.

I focus on sectors instead of firms because I do not observe this withholding information for all firms, only for firms that file VAT forms. Firms in the same finely defined 7-digit sector will likely be similarly effected by withholding as they will sell a similar set of products. In addition, there are concerns about some firms making systematic mistakes in reporting withholding values in exemptions tax codes when the sale of a product does not generate a refund, in which case both tax codes would have the same null implication for the tax due.

One concern with the procedure above is that the set of sectors I look at for the tax liabilities analysis is not necessarily representative. I address this concern by allowing a direct comparison between this sample and the overall sample. In the empirical analysis (see Section 5 and Table 3) I present the results for reported revenue for the sample with little withholding alongside the results for tax liabilities to allow for a direct comparison between the this subsample of sectors and the main sample. In addition, Figure A4 in the Online Appendix A shows the main DD graph in Figure 2b for reported revenue (Difference coefficients for 6-month time bins between Jan. 2004 to Dec. 2011)

Reported inputs data. Only firms in the VAT regime (RPA regime mentioned above) file information about inputs. Therefore, in order to analyze the effect of the policy on reported inputs, I have to restrict attention to firms that are in the VAT and have not switched tax regimes during the period of analysis. The reported inputs data was extracted as a separate file. The scrambled identifiers used in this extraction were different than the ones used in the original data. However, the procedure do de-identify and clean the data was the same. The files included the GIA/ICMS forms, and the history of tax regimes of firm ids that are in these files to help identify firms that never switched out of the VAT regime between Jan 2004 and December 2011. By imposing this restriction the sample size shrinks to 44,211 firms. In the analysis I present the results for reported revenue for this sample alongside the results for reported expenses to allow for a direct comparison between the this sample and the main sample.

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. The data was de-identified, and a scrambled 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 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. For retail firms, there is also a data with the total count of receipts issued - with or without SSN - between January 2009 and December 2011.

B.2. Re-weighting, Complaints and Lotteries

First, I introduce the re-weighting methods used in Section 4 of the paper. Then I provide further details and discussion of the sampling and empirical design of the impact of the first complaint for firms and the impact of lottery wins for consumers in Section 4.

B.2.1 Re-weighting

In both event studies I use in the paper - complaints and lotteries - I use a re-weighting method based on DiNardo *et al.* (1996) (DFL) to flexibly control for the odds of the event. The Appendix B of Yagan (2015) presents a thorough description of an application of DFL re-weighting. In this section, I explain how I use the weights in the applications of the paper, but it largely based on the description in Yagan (2015).

First, define the groups g that are being compared (e.g., consumers that are lottery winners vs. losers, or firms that received complaints vs not). Then divide all observations into bins b according to the relevant traits for the realization of the event (e.g. lottery ticket holding or probability of getting a complaint). Then use the number of observations in every group-bin to create weights so that the within-group distribution of weights across bins equals the original cross-bin distribution of weights in some base group g (e.g. lottery winners in time t_0 or firms that received a complaint in time t_0). Intuitively, DFL holds fixed the distribution of observable traits across groups by inflating or deflating the control group to match the distribution of the treatment group.

To explain the details, consider first the case of lottery wins in Section 4, there are two groups - winners and non-winners - and 10 event-time periods between [-3, +6] around each event date t_0. I DFL-reweight across 20 groups g (= 2 groups and 10 periods). I define the base group g to be the the winner group at event date t_0. I want to compare two consumers with the same odds of winning, i.e., that hold the same number of lottery tickets. I therefore use each lottery ticket holding between 1 and 40 (to ensure common support) to bin observations into one of 40 bins b. Let b denote the bin and let g denote the group that observation j falls in. The final weight w on observation j equals:

$$w_{jbg} = \left(\frac{\sum_{j' \in b \cap j' \in g} 1}{\sum_{j' \in b \cap j' \in g} 1} \right) \left(\frac{\sum_{j' \in g} 1}{\sum_{j' \in b} 1} \right) \tag{1}$$

where j' denotes firm-year observations generally.

Using this formula, every observation j that is in the base group g will have a final weight equal to 1. Every observation that is not in the base group will received a weight smaller or greater than 1 depending on whether it is over represented or underrepresented when compared to the base group g. The first factor in parentheses ensures that within every group g, the ratio of the sum of observations for a given number of lottery tickets b to the sum of observations in any other other number of lottery ticket holding b' is identical to the corresponding ratio in for the lottery winners in time t_0 (the base group g). The second factor ensures that the sum of each groups' final weight equals the sum of that group's original weight.

For the case of complaints, I first calculate a propensity score for getting a complaint at time t_0. Then I create the bins for firms b based on quartiles of the propensity score distribution and follow the same procedure described above to calculate the weights using formula (1). In the next two subsections I explain in more detail the construction of the data and a discussion of the role of the DFL weights in each exercise.

B.2.2 Complaints

Complaints data. 25% of firms received at least one complaint in the period of analysis. I begin defining the complaints sample by looking at firms in the retail sector that issued at least one receipt before June 2009. For each firm i I identify the time of the first complaint any of their establishments received. Then, for each complaint date e between July 2009 and June 2011, I build a panel data with 6 month window around the complaint date where firms that received a complaint at that event date are in the treatment group and firms that did not receive their first complaint by that event date are in the control group. The same firm i can be in T or C depending on the event as the control group draws from firms that did not yet receive a complaint by event e. Then, I restrict attention to firms that have positive revenue and receipt data to make sure I have a balanced panel. The combined dataset that appends all events covers the time period between Jan. 2009 and Dec. 2011, i.e., at 6 months before and after the earliest and latest first complaint respectively.

Once I create the propensity score, I construct a dataset for each month within this period, where I keep all firm that received their first complaint that month and all firms that did not receive their first complaint that month and I re-weight the no-complaints group to match the complaints group within each quartile of the propensity score distribution. I collapse each cohort of complaints by each group and “event-month" using the weights to show the raw data in Figure 3, and use the micro-data to run specification (9) using the weights.

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

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

B.2.2 Lottery wins

Lottery data. The lottery sample covers consumers that hold fewer than 40 lottery tickets in a given month for 12 lotteries between July 2009 and June 2011. In order to perform the empirical exercises on the effects of lottery wins on consumer participation in Section 4.2 I merge this data with the *receipts data* described above The combined dataset of lotteries and receipts covers the time period between January 2009 and December 2011, i.e., 6 months before and after the first and last lottery. I balance the panel of consumer participation and replace missing values by zero for the two key variables I use *number of receipts* and*total value of receipts*

Consumer sample. This sample only includes consumers that participated in at least one lottery between July 2009 and June 2011. For each monthly lottery, I restrict attention to lottery winners and 10% random sample of consumers who did not win the lottery in each draw. Table 2 creates summary statistics for a balanced panel of individual SSNs with a count of the number of receipts they ask, the total value of receipts, lottery tickets and lottery prizes. The data includes zeros in months where no receipts are asked or the individual does not participate in lotteries.

For the lottery analysis, I take the winners and the sample of non-winners of each lottery, and I append all lotteries. I create a balanced sample for three months before and six months after the lottery draw, and the lottery draw month. The same consumer can be a winner a non-winner depending on the lottery. The estimation use individual-lottery FE in order to only use variation within individual-lottery, and also include calendar time FE and event-time FE.

Re-weighting. Since the number of lottery tickets is determined by the total value of a consumer' purchase 4 months before the lottery draw, the more a consumer participates in the program by asking for receipts, the higher are the odds she will get a prize in a given lottery. Therefore, it is important to carefully control for the odds of winning a prize in order to study the effect of lottery wins on consumer participation. As I describe in Section 4.2, I use the DFL re-weighting method to flexibly control for the number of lottery tickets individuals hold to ensure I use the random component of the lottery by matching the two groups based on the odds of winning prize.

Figure A.5 shows an example of the distribution of lottery ticket holdings among winners and non-winners in monthly lotteries. 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 for every lottery holding. 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.

I perform this re-weighting exercise separately for each lottery win I study in Section 4.2. I construct a dataset for each prize level as described in *lottery data* above, where I keep all consumers that won a given prize (winners) and all consumers that do not win any prize (non-winners). When I compare the effect of different prize values, the pool of non-winners is the same in each lottery across the datasets I create for each prize level but they are re-weighted differently depending on the prize I am considering since the winners group of different prize levels may have slightly different distributions of lottery ticket holdings.

Once I calculate these weights for each monthly lottery, I append the panels of consumers and 10 event-time periods between [-3, +6] around each event date t_0. Then, I collapse the data by event-time using the weights for the graphs in Figure 5, Figure A5 and Figure A6. 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.

Online Appendix C: Conceptual framework - Extensions

In this section, the goal is to extend the discussion of the conceptual framework to further discuss relevant policy dimensions: (i) imperfect take up and (ii) collusion costs.

C.1. Imperfect take up

Consider the same conceptual framework as in Section 1, where the government provides a targeted incentive α for consumers to ask for receipts. As before, suppose consumers derive benefit $\kappa(\alpha)\tau y$ from the governments' incentives to ask for receipts, but allow consumers to be heterogeneous in their participation cost ϕ_i. Let $\phi_i = \underline{\phi}$ for a share $\lambda \geq 0$ of the N consumers, and $\phi_i = \bar{\phi}$ for a share $(1 - \lambda)$, where $\underline{\phi} < \bar{\phi}$.

If $\kappa(\alpha)\tau y \geq \bar{\phi}$, all consumers will take up the policy, i.e., will ask for receipts, and the problem will be the same as the compliance decision with consumer monitoring in Section 1. In order to highlight the relevance of take-up, let $\kappa(\alpha)$ be such that $\underline{\phi} < \kappa(\alpha)\tau y < \bar{\phi}$. Therefore, a share $\lambda > 0$ of consumers will respond to the incentives and ask for receipts. Imperfect take up will weaken the enforcement effect of the policy. I will consider two alternative responses of firms: selective reporting and collusion.

Selective Reporting. Suppose firms can strategically choose how each of the N transactions will be reported to the government. In this case, the degree to which the policy can affect firms at all may depend on the baseline compliance. Let λ^* be such that $\lambda^* \bar{Y} = Y^*$. If $\lambda \leqslant \lambda^*$, the share of transactions for which consumers ask for receipt will be smaller than what would be reported in the absence of the policy $\lambda \bar{Y} \leqslant Y^*$. In this case, the receipts consumers begin to ask would be entirely infra-marginal if firms can selectively under-report transactions for which no receipt was asked, and report transactions for which consumers did ask for receipts as there is a higher detection probability if firms misreport these transactions.[2]

Thus, if consumer take up is sufficiently low, compliance may not change irrespective of the size of rewards. Moreover, the availability of whistle-blower channels absent of rewards would make no difference as consumers would not observe evidence of evasion. This stylized case emphasizes how important take up can be if firms do not under-report the entirety of their sales and can do selective reporting.[3] In addition, the rewards offered by the government would be infra-marginal as no additional sale would be reported as a result of the policy. Thus, the baseline level of compliance matters for the cost of the policy: in the limit, if the baseline compliance close to perfect, rewards could be entirely infra-marginal even with large take-up.[4]

[2] This is similar to the argument used in Kleven *et al.* (2011) to describe reporting incentives for incomes that differ in third-party reporting coverage: because the tax rate and penalty are the same across different revenue items, this strategic selective reporting would be the optimal sequence for the taxpayer. The detection probability would have a S shape: it sharply increases once firms start underreporting income subject to third-party reporting.

[3] If selective reporting imposes a concealment cost, the policy could have an effect on compliance by increasing the cost of evasion.

[4] Similarly, if there is a share of consumers that require receipts even absent of the a reward policy (if they have a negative ϕ_i because they value compliance directly), their effort to ensure firms issue receipts would only affect firms if the number of consumers with negative costs is high enough.

If $\lambda > \lambda^*$, the consumer policy will affect the reporting decision of the firm. If firms continue reporting any transaction for which consumers ask for receipts, the reported revenue would be given by $Y = \lambda\bar{Y}$. In this case, the higher the take up, the larger the amount of sales reported by firms, so firms would mechanically increase the amount of reported sales depending on the number of consumers that take up the policy.

Since take-up depends on $\kappa(\alpha)\tau y$ being higher than the participation cost, goods with high τ and high price y would be relatively more effected by the policy as it is more likely that the reward will be higher than the participation cost for a larger share of the population. It also highlights that if participation cost are very high (e.g., $\kappa(\alpha)\tau y < \underline{\phi}$), the policy may have no effect even if α is generous. Note that under selective reporting, the whistle-blower channel would have no effect as firms would not reveal to consumers evasion information. In terms of government revenue, the policy would be net positive if $\frac{\lambda-\lambda^*}{\lambda^*} > \frac{\alpha}{1-\alpha}$.

Collusion policy. Suppose firms could offer the collusion deal to consumers that request receipts: a discount of $\kappa(\alpha)\tau y$. As before, assume that consumers take any offer that matches their valuation $\kappa(\alpha)$ of the policy. Because only consumers that ask for receipts learn about the evasion decision of firms, the risk of whistle-blowers will depend on take up as well. The detection probability would be affected less than in the case of full take up: $d_c^\lambda = 1-(1-d)(1-\varepsilon)^{\lambda N} < d_c$ as there are fewer potential whistle-blowers, so $p_c^\lambda = a(E)[1-(1-d)(1-\varepsilon)^{\lambda N}]$. Firms will have to transfer $\lambda(\bar{Y}-Y)\kappa(\alpha)\tau$ to consumers. Firms choose Y to maximize:[5]

$$(\bar{Y}-\tau Y)(1-p_c^\lambda) + [(1-\tau)\bar{Y} - \theta\tau(\bar{Y}-Y)]p_c^\lambda - \kappa(\alpha)\tau\lambda(\bar{Y}-Y) \tag{2}$$

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

$$[a + a'(E).E]d_c^\lambda(1+\theta) = 1 - \lambda\kappa(\alpha) \tag{3}$$

For the comparative statics, c can be re-written as $c = \frac{1-\lambda\kappa(\alpha)}{d_c^\lambda(1+\theta)}$. As before, changes in c translate into comparative statics for evasion E. The main difference under imperfect take-up is that a lower λ decreases the change in detection ability and the transfers to consumers, so $Y_\lambda^{**} < Y^{**}$. Compliance, thus, increases with the share of consumers λ that take up the policy.

Similarly to the case of selective reporting, take-up depends on $\kappa(\alpha)\tau y$ being higher than the participation cost, so goods with high τ and high price y could be relatively more effected by the policy. But in the collusion case, any share $\lambda > 0$ would change the evasion decision of the firm as firms are underreporting all transactions by the same amount, instead of selectively reporting some transactions in full and evading the remaining ones. In terms of government revenue, the net effect will be positive if $\frac{Y_\lambda^{**}-Y^*}{Y^*} > \frac{\alpha}{(1-\alpha)}$.

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

The data available to this study does not allow to distinguish between these two strategies of misreporting: selective underreporting and collusion. However, these two cases are helpful to emphasize the relevance of take up and participation costs in affecting the effectiveness of the policy.

C.2. Collusion costs

Another mechanism that is consistent with the evidence is that there is a fixed cost in setting a collusive deal with consumers that acts as a concealment cost, so the larger the number of consumers a firm interacts with, the more this policy may increase compliance as it dis-proportionally affects firms that need to collude with a large number of consumers.

In the paper, I assume that when firms make an offer to a consumer, they pay a fixed cost $\rho > 0$. In order to focus on the effect of collusion costs, assume perfect take up and homogeneous consumers. The fixed cost can be thought of as a concealment cost paid to collude at each transaction. If the firm follows a collusion policy with all its consumers, the total cost of collusion ρN would affect the extensive margin decision between evasion and full compliance, but not the intensive margin of compliance as discussed in the paper.

Given the fixed cost ρ, if firms can optimize the number of collusive deals and offer discounts to some consumer but not others, they could restrict collusion to a smaller set of consumers. The total number of collusive deals a firm makes is given by the amount evaded $E = \bar{Y} - Y$ divided by the value of an individual transaction $\bar{y}$. Thus, the total negotiation cost is increasing in the number of collusive deals a firm makes given by $(\frac{\bar{Y}-Y}{\bar{y}})$. Because firms will only reveal evasion information to a subset of consumers, $\hat{d}_c = 1 - (1-d)(1-\varepsilon)^{\frac{\bar{Y}-Y}{\bar{y}}} < d_c$ as there are fewer potential whistle-blowers. So $\hat{p}_c = a(E)[1 - (1-d)(1-\varepsilon)^{\frac{\bar{Y}-Y}{\bar{y}}}]$

$$\pi = (\bar{Y} - \tau Y)(1 - \hat{p}_c) + [\bar{Y}(1-\tau) - (\bar{Y} - Y)\theta\tau]\hat{p}_c - (\bar{Y} - Y)\kappa(\alpha)\tau - (\frac{\bar{Y} - Y}{\bar{y}})\rho \tag{4}$$

$$[\hat{p}_c - \hat{p}_c'(E).E](1+\theta) = 1 - \kappa(\alpha) - \frac{\rho}{\tau\bar{y}} \tag{5}$$

The marginal benefit of evading an extra dollar is reduced further by $\frac{\rho}{\tau\bar{y}}$. The total change in the marginal benefit of evasion amount is equal to the transfer firms need to make to consumers in the collusive deal plus the cost per dollar negotiated with consumers. Therefore, the costs of collusion enter as an extra penalty for each dollar evaded.

Here, it becomes more explicit that the value of the transaction $\bar{y}$ will matter for this policy as firms with the same level of evasion may be affected differently as firms with a low ticket items would have to collude with a higher number of consumers for a given total revenue $\bar{Y}$. Since the number of potential whistle-blowers depend on how many collusive deals are done, $\hat{p}_c$ is also higher for lower $\bar{y}$, further pushing firms towards compliance.

Therefore, the size of the transaction could matter through collusion costs and the higher risk of whistle-blowers for a given firm size. It is worth noticing that predictions about the size of the transaction should also depend on take up. If take up is imperfect as the cost of asking for receipts is too high for some consumers, fewer consumers would be willing to ask for receipts for small ticket items. In the data, the effect is stronger for low ticket items, which is consistent with the transaction cost for consumers not being very high and the collusion costs (through fixed costs and/or additional whistle-blower threat from volume of transactions).

Online Appendix D: Effects on Formal employment and number of firms

The observed enforcement affect generated an increase in the effective tax rate. This change may affect firms that were on the margin of exiting the market, entering the market or firing employees. In this appendix, I analyze the implications of the policy on formal employment and the number of firms.

D.1. Formal employment

Employer-employee data. From the Brazilian Department of Labor, I use annual reported employment for all formal establishments in Brazil (RAIS/CAGED). The data covers all formal establishments that have at least one formal employee. All formal firms must report to the Department of Labor their employment information in a yearly basis. I use a version of this data that aggregates the total number of employees by 7-digits sector definition. Because the data from the Department of Finance of Sao Paulo is de-identified, it cannot be matched with the employer-employee dataset (RAIS/CAGED). In order to analyze the impact of the program on employment, I construct a panel of CNAE-year for each Brazilian state.

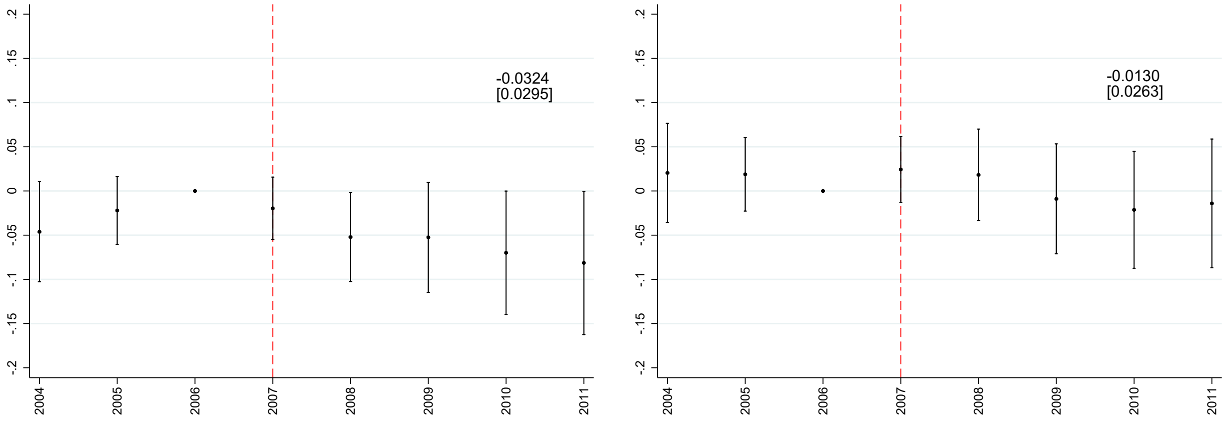
Employment effect. I employ two strategies. First, I apply the same difference in differences (DD) design in equation (6) of the paper comparing retail and wholesale sectors. Because the employment data is measured yearly I define *after* as $t \geq 2007$. Figure D1 shows the DD coefficient from estimating a specification similar to (6) in the sector yearly panel with standard errors clustered by sector. It also shows the point estimate and 95% confidence intervals for the coefficients of a more flexible DD specification in (5) using yearly dummies interacted with the retail dummy, and using the 2007 interaction as the omitted category.

Then, because this sample covers the entire country, I can use retail-wholesale difference in Sao Paulo and compare with the retail-wholesale difference in other states in a triple difference in differences (DDD).

$$\begin{aligned} lnEmpl_{stm} =& \pi_m + \eta_s + \lambda_t + \psi SP_m \cdot Post_t + \alpha Treat_s \cdot Post_t + \gamma SP_m \cdot Treat_s + \\ & \beta Treat_s \cdot Post_t \cdot SP_m + u_{stm} \end{aligned}$$

Where $lnEmpl_{stm}$ is the natural log of total formal employment in sector s, year t and state m. π_m is state fixed effects, η_s is sector fixed effects, λ_t is year fixed effects. Figure D1.B. shows the coefficient β estimated using the above specification, and also displays the point estimate and 95% confidence intervals for a more flexible specification where the interactions time is done with year dummies separately for each year, using 2007 as the omitted category.

Even though the point estimate is negative in both cases, the coefficients are not statistically different than zero, suggesting that the policy, on average, had no effect on formal employment. The fact that I find no or slightly negative 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.



(a) (DD) - SP

(b) Retail vs. Wholesale in SP vs. Other States(DDD)

FIGURE D.1: LOG OF NUMBER OF FORMAL EMPLOYEES

Note: The figure plots the log of number of firms in retail vs. wholesale. The count of number of firms considers all firms that that reported positive revenue each year. The figure displays the DD coefficient from estimating a specification similar to equation (6) in a 7-digit sector yearly panel with 1,680 obs., and using the log of total number of firms as the outcome. Standard errors are clustered at the 7-digit sector level.

D.2 Number of firms

The increase in enforcement could affect both entry and exit decisions. To assess whether these margins were affected by the policy, I use the log of total number of firms observed in each year and sector as a simple measure of potential changes in entry and exit decisions.

Number of firms To calculate the number of firms, I define as the year of exit the last year a firm had any sales during the period of analysis. Firms may submit forms with zero activity due to the slow process of closing a firm in Brazil, so defining exit as when they stop reporting positive economic activity would better capture the timing of exit. The results are robust to alternative ways of counting firms, where I count any firm that files a tax form at least once each year even if it has zero sales or inputs.

Figure D.2 plots the log of number of firms in retail and wholesale sectors. The figure also displays the DD coefficient from estimating a specification similar to equation (6) in the paper but using the log of number of firms as the outcome. The coefficient is not statistically distinguishable from zero, which indicates that, on average, the policy did not affect the number of firms in retail sectors differently than in wholesale sectors.

The evidence above indicates that the increase in tax enforcement did not affect employment or the number of firms of affected sectors. The null effect indicates that the implied increase in the effective tax rate may not be large enough to affect firms along these margins, and may just

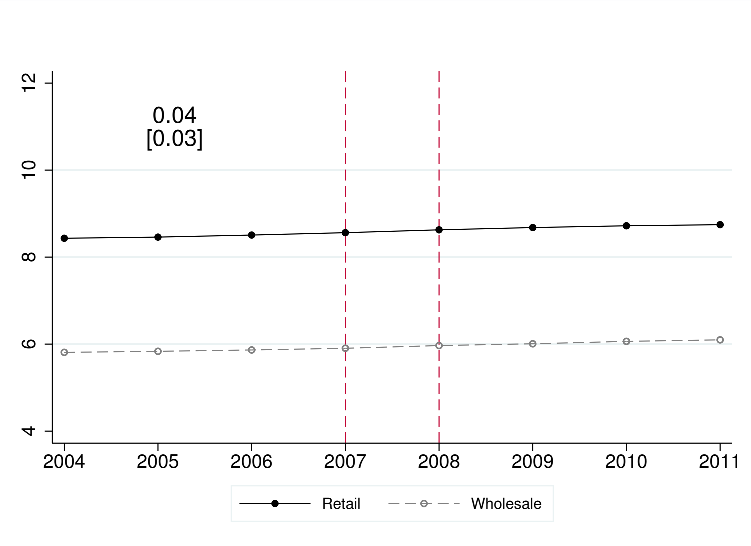


FIGURE D.2: LOG OF NUMBER OF FIRMS

Note: The figure plots the log of number of firms in retail vs. wholesale. The count of number of firms considers all firms that that reported positive revenue each year. The figure displays the DD coefficient from estimating a specification similar to equation (6) in a 7-digit sector yearly panel with 1,680 obs., and using the log of total number of firms as the outcome. Standard errors are clustered at the 7-digit sector level.

reduce evasion rents.[6] However, it is possible that changes may occur after the period of analysis. Another possibility is that firms can potentially adjust other margins that I do not observe in the data such as, for instance, informally-hired workers.

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

References

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

Kleven, Henrik Jacobsen, Knudsen, Martin B, Kreiner, Claus Thustrup, Pedersen, Søren, & Saez, Emmanuel. 2011. Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica*, **79**(3), 651–692.

Yagan, Danny. 2015. Capital tax reform and the real economy: The effects of the 2003 dividend tax cut. *American Economic Review*, **105**(12), 3531–63.