Consumers as Tax Auditors

Joana Naritomi

February 14, 2024

Introduction

- Access to third-party information and paper trails believed to be critical for administration of modern tax systems.
- Despite collusion opportunities, how can government still effectively utilize this paper trail generated by economic activities.
- We discussed in this class that third-party reporting leads to substantial decrease in evasion (Gordon & Li 2009; Kleven et al. 2016).
- Therefore extensive margin effect of paper trail is known, but what are the mechanisms and how it avoids collusion among informal transcting parties.

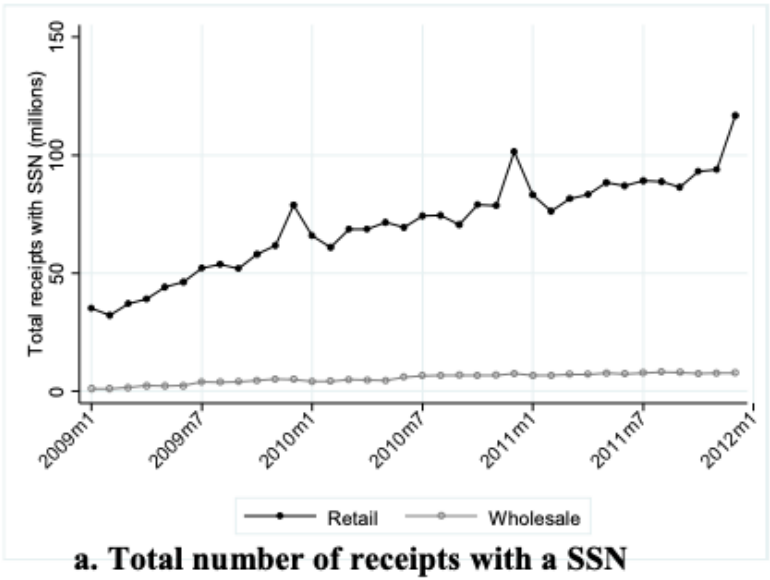
This paper

- This paper exploits quasi-experimental variation and unique administrative data on firms and consumers from an anti-tax evasion program in Sao Paulo, Brazil – Nota Fiscal Paulista (NFP) – that created monetary rewards for consumers to ensure that firms report final sales transactions.
- The program provides tax rebates and monthly lottery prizes for consumers who ask for receipts, and establishes a direct communication channel between the tax authority and consumers through an online account system, where consumers can verify receipts reported by firms and can act as whistle-blowers by filing complaints.
- She constructs unique administrative data on firm-level monthly tax returns, monthly individual-level data on requested receipts and overall participation in the NFP program, based on administrative records from the tax authority of the state of Sao Paulo.

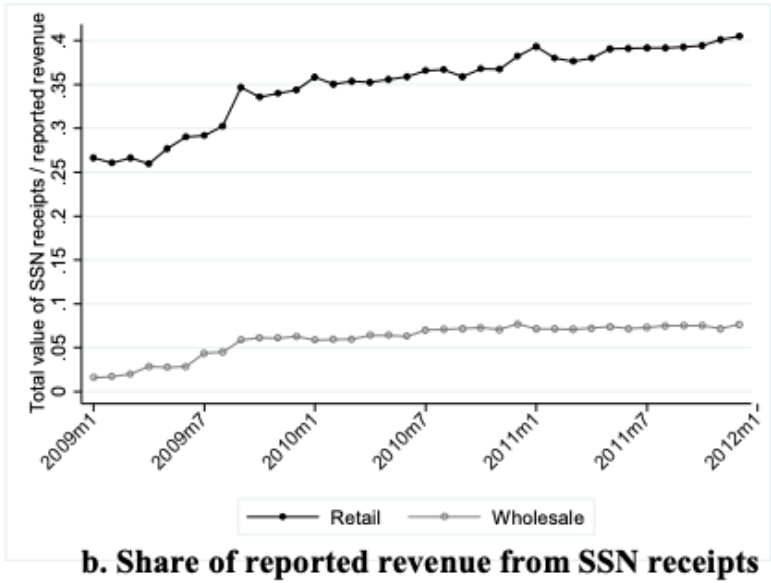
Empirical Approach

- Compare retail firms vs wholesalers.
- Reported revenue by retail firms increased by 21% within 4 years of reform. This is likely a lower bound.
- Study mechanisms through winners of lottery and volume of firm sales or mass of consumers of a firm.

First Stage



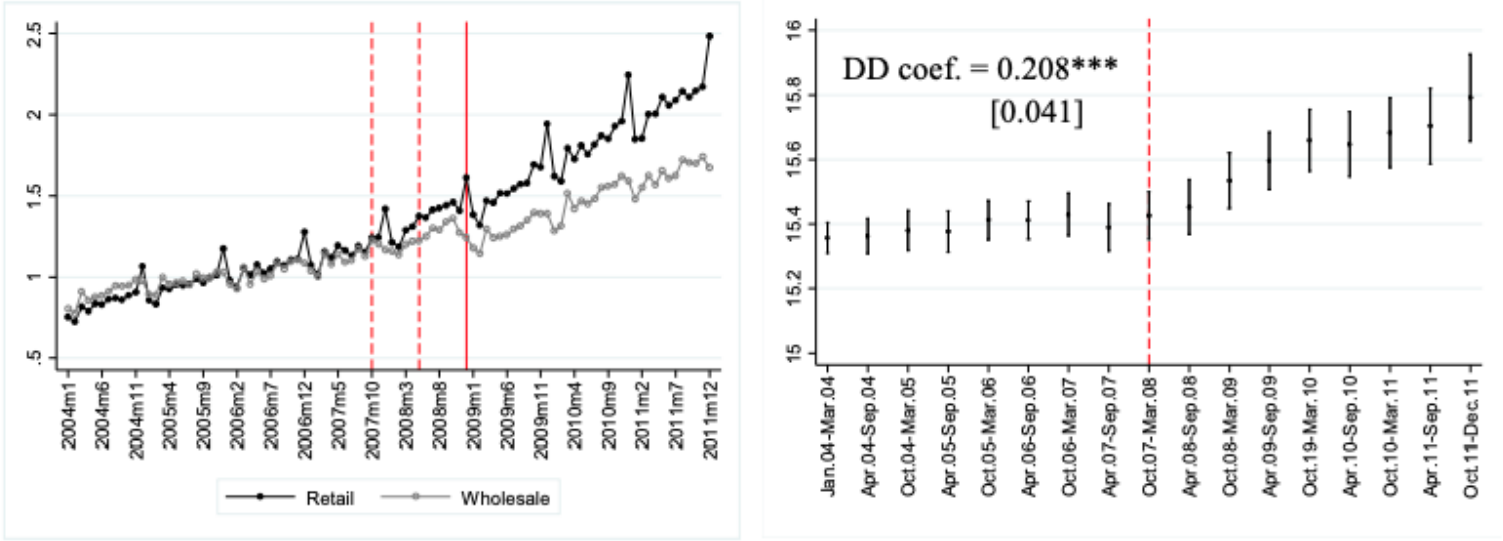
a. Total number of receipts with a SSN



b. Share of reported revenue from SSN receipts

FIGURE 1: NFP RECEIPTS AND REVENUE SHARES – RETAIL VS. WHOLESALE

Effect of Reform



a. Raw data: reported revenue changes **b. Difference coefficients for 6-month time bins**

FIGURE 2: EFFECT OF THE POLICY ON REPORTED REVENUE – RETAIL VS. WHOLESALE

Note: Figure 2a shows reported revenue changes for retail and wholesale sectors. Each line is the revenue reported by all firms aggregated by retail or wholesale scaled by the average monthly reported revenue before Oct. 07 for each sector group. The figure plots the raw data. The 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 2b plots regression coefficients from estimating specification (5) in Section 3 using a sample of 212 sectors between Jan 2004 and Dec 2011. The sector sample has 20,352 observations. The difference in differences (DD) coefficient displayed in the figure is estimated using the specification (6) in Section 3 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. See Table A1 Panel B and Figure A2 in Online Appendix A for robustness checks. Significance levels *** 1%, ** 5%.

Who responds to the reform?

TABLE 2: REPORTED REVENUE EFFECT – RETAIL VS. WHOLESALE

	log Reported Revenue				
	[1]	[2]	[3]	[4]	[5]
DD (Post Oct 07 * Retail)	0.254***				
	[0.0722]				
DD * Large firms		0.253***			
		[0.0732]			
DD * Small firms		0.350***			
		[0.0511]			
DD * High volume of different consumers			0.246***		
			[0.0705]		
DD * Low volume of different consumers			0.0329		
			[0.0919]		
DD * High volume of transactions				0.253***	
				[0.0335]	
DD * Low volume of transactions				0.0181	
				[0.0391]	
DD * High value of transactions					0.0969
					[0.0689]
DD * Low value of transactions					0.285***
					[0.0754]
3rd-order polynomial of firm size * DD			X	X	X
Time FE	X	X	X	X	X
Firm FE	X	X	X	X	X
Observations	1,035,268	1,035,268	1,035,268	1,035,268	1,035,268
Adjusted R-squared	0.907	0.907	0.908	0.909	0.908

Do they deduct more?

TABLE 3: TAX LIABILITY AND REPORTED EXPENSES – RETAIL VS. WHOLESALE

Panel A: Tax sample

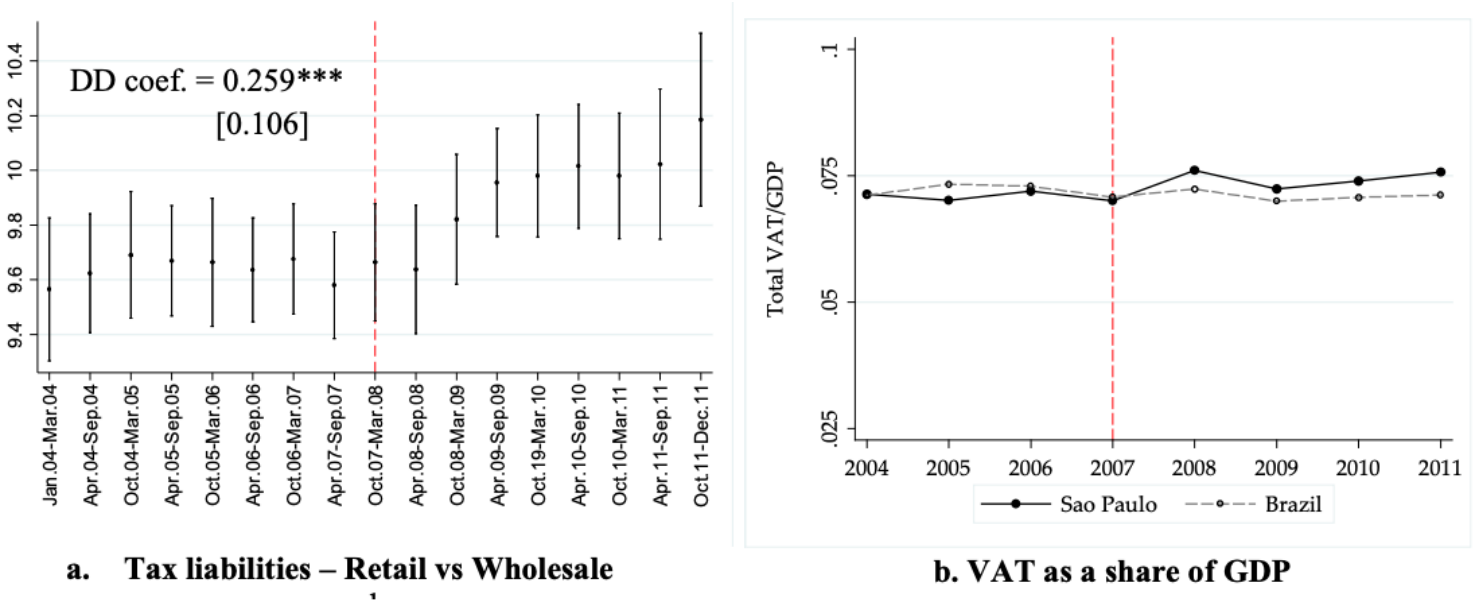
	Log of Reported Revenue [1]	Log of Tax Liability [2]	Positive tax liability [3]
DD (Post Oct 07 * Retail)	0.311**	0.316**	0.0434
	[0.151]	[0.137]	[0.0350]
Firm FE	X	X	X
Time FE	X	X	X
Observations	167,110	133,950	167,110
Adjusted R-squared	0.85	0.876	0.801

Panel B: Expenses, output and value added - firms that were always VAT

	Log of Reported Revenue [1]	Log of Reported Inputs [2]	Log of Reported Value Added [3]	Positive Value Added [4]
DD (Post Oct 07 * Retail)	0.363***	0.302***	0.387***	0.0192
	[0.0824]	[0.0833]	[0.105]	[0.0153]
Firm FE	X	X	X	X
Time FE	X	X	X	X
Observations	88,422	88,422	70,845	88,422
Adjusted R-squared	0.87	0.85	0.90	0.71

Note: Table 3 displays the main coefficients form regressions described in Section 5 using the firm-level data. The variable

Taxes go up

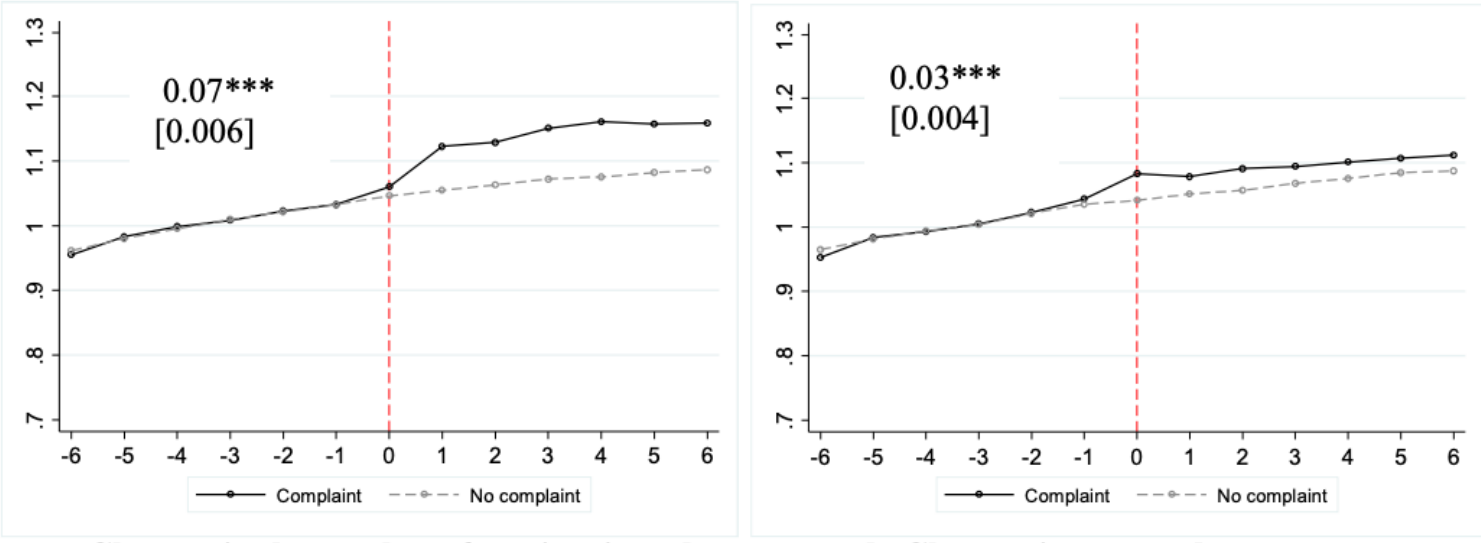


a. Tax liabilities – Retail vs Wholesale

b. VAT as a share of GDP

FIGURE 6: EFFECT OF THE POLICY ON TAX REVENUE – RETAIL VS. WHOLESALE

Effect of Whistle Blowers



a. Changes in the number of receipts issued **b. Changes in reported revenue**

FIGURE 3: WHISTLE-BLOWER EFFECT ON FIRM COMPLIANCE

Note: Figures 3a and 3b plot the changes in the total raw number of receipts firms issue and changes in reported revenue to the government before and after a firm receives the first complaint. Both graphs display changes across event-time where each data point is scaled by the outcome's average before the first complaint (event-time zero). The 'Complaints' group is composed by firms that received their first complaint at event-time zero. The 'No complaint' group is composed by firms that did not receive their first complaint at event-time zero and firms that did not receive a complaint until Dec. 2011. The outcome is averaged across groups and event times using weights based on quartiles of the propensity score to get the first complaint in a given calendar time. The propensity score is estimated using 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, reported receipts, SSN receipts and number of consumers (see Online Appendix B for more details). The estimated DD coefficient displayed

Robustness

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)	0.254***	0.291***	0.200***
s.e. clustered by sector	[0.0722]	[0.107]	[0.0521]
s.e. clustered by firm	[0.0360]	[0.0690]	[0.0251]
Firm FE	x	x	x
Time FE	x	x	x
Observations	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
Observations	20,352	20,352	20,352
Adjusted R-squared	0.982	0.976	0.987

Program Participation and Role of Incentives

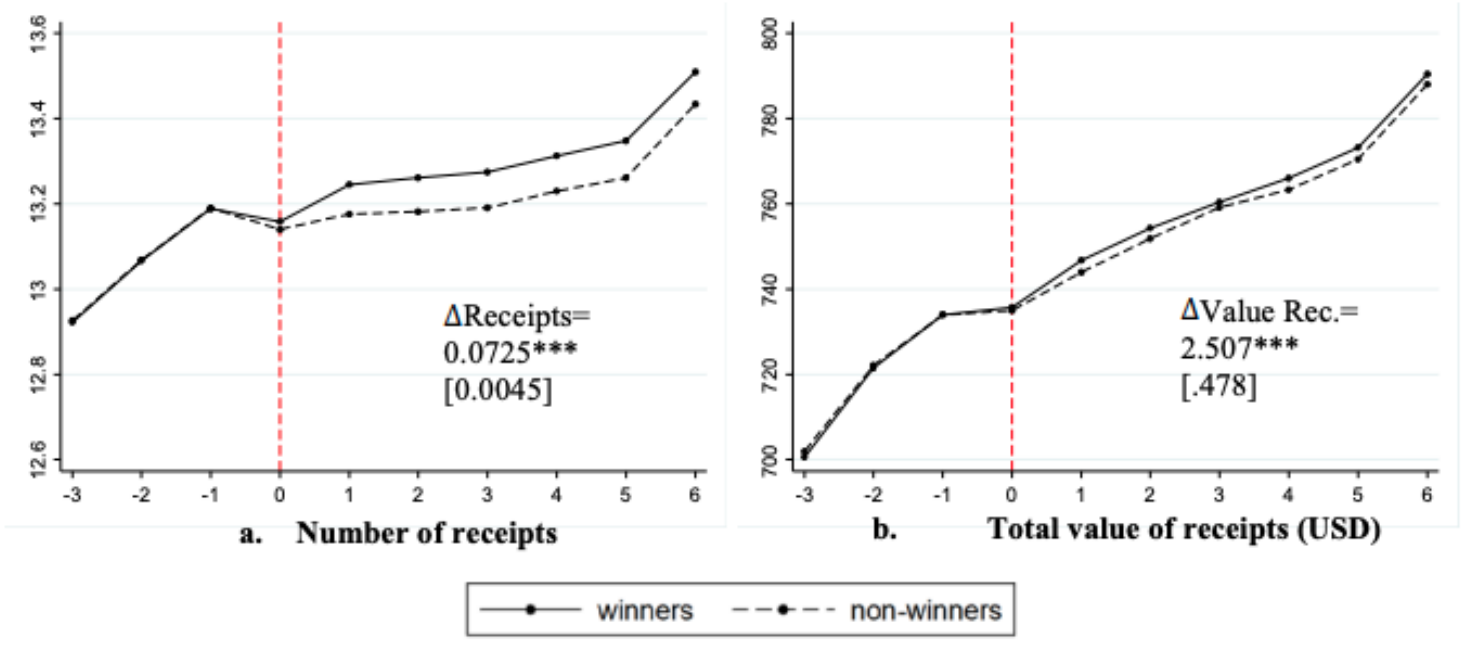


FIGURE 5: THE EFFECT OF A 5-DOLLAR LOTTERY WIN ON CONSUMER PARTICIPATION

Note: The graphs show the raw data by month aggregating all lotteries from June 2009 to June 2011. The x-axis is the number of months since the individual participated in a lottery. The winner group got a cash prize of US $5 (R$10) and the non-winner group did not get any prize. Figure 5a plots the total number of receipts consumers ask in each group before and after the lottery draw at event-time zero. Figure 5b shows the total value of receipts (in USD) for each group before and after the lottery draw at event-time zero. 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) in Section 2 using the micro-data and the lottery ticket weights. Standard errors are clustered by lottery draw. Significance levels *** 1%, ** 5%.

Conclusions

- What would be marginal value of public funds?
- Does program pay for itself? tax revenue increased by 9.3% net of rewards.
- However, we do not know the cost of running this program!