Essays on Taxation and Transfers in Middle-Income Countries

by

Pierre Bachas

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Emmanuel Saez, Chair
Professor Edward Miguel
Professor Paul Gertler
Professor Andres Rodriguez-clar

Spring 2016
Essays on Taxation and Transfers in Middle-Income Countries

Copyright 2016
by
Pierre Bachas
Abstract

Essays on Taxation and Transfers in Middle-Income Countries
by
Pierre Bachas
Doctor of Philosophy in Economics
University of California, Berkeley
Professor Emmanuel Saez, Chair

At a time of growing inequality and under-investment in public infrastructure, my research has focused on understanding governments’ constraints in raising tax revenue and providing redistribution. These challenges are particularly important for low and middle-income countries: despite improvements in their institutional capacity in the last decades, their ratios of tax revenue to GDP remain much lower than OECD countries’ (Besley and Persson 2013a), and their tax and transfer systems are often distributionally neutral, instead of progressive.

In the first chapter, I ask whether developing countries with limited information and tax capacity can use the corporate income tax to raise additional revenue, and design it optimally given these constraints. I explore this question for Costa Rica, using the universe of corporate tax returns and a novel methodology which exploits the country’s unique tax design: firms with marginally different revenue face discontinuously higher average tax rates. This notch feature allows me to first estimate the elasticity of profits with respect to the tax rate, and second to separate it into its components, namely the revenue and cost elasticities. I find that firms facing a higher tax rate slightly decrease reported revenue, but considerably increase reported costs, leading to a large drop in reported profits. Using additional data sources, firms’ behavioral responses appear to occur through tax evasion, with no evidence of production responses. Taken together, this implies that Costa Rican firms evade taxes on a massive 70% of their profits when faced with a 30% tax rate. In this context, lowering corporate tax rates could increase tax revenue, since we estimate the revenue maximizing rate to be below 25%. Alternative tax rules, that limit the deductibility of costs could be preferable since they would reduce evasion opportunities on this crucial margin. The results highlight the limitations of standard business taxation as an instrument to raise revenue in developing countries.

The first chapter points to limitations for revenue collection, when given the current enforcement environment. In the Second Chapter, I study how third-party information might spread to the government, in order to improve tax enforcement. Firm level tax compliance depends on the stock of information accessible to the government. Theoretical work
(Gordon and Li 2009b, Kleven, Kreiner, and Saez 2016) highlights two specific information trails: access to formal finance and the number of employees. I test whether firms with more employees and with access to formal finance are more likely to be audited and less prone to evading taxes. I use firm-level data on 108,000 firms across 79 countries in the World Bank Enterprise Surveys and construct instruments for finance and worker-size at the industry level, using an out of sample extrapolation strategy related to Rajan and Zingales (1998). The instruments isolate variation in industry technological demand for labour and formal finance by taking the US industry distributions as undistorted benchmarks. I find that firms with more employees are more likely to be audited and to comply, but find no evidence that firms using the financial sector are under higher scrutiny.

Finally, in the third chapter, I turn to the redistributive role of governments, and study how a new technology to deliver cash transfers can be used to impact transfer beneficiaries’ trust in financial institutions and their savings behavior. It is well documented that trust is an essential element of economic transactions, however trust in financial institutions is especially low among the poor, which may explain in part why the poor do not save formally. Debit cards provide not only easier access to savings (at any bank’s ATM as opposed to the nearest bank branch), but also a mechanism to monitor bank account balances and thereby build trust in financial institutions. I study a natural experiment in which debit cards were rolled out to beneficiaries of a Mexican conditional cash transfer program, who were already receiving their transfers in savings accounts through a government bank. Using administrative data on transactions and balances in over 300,000 bank accounts over four years, I find that after receiving a debit card, the transfer recipients do not increase their savings for the first 6 months, but after this initial period, they begin saving and their marginal propensity to save increases over time. During this initial period, however, they use the card to check their balances frequently; the number of times they check their balances decreases over time as their reported trust in the bank increases. Using household survey panel data, I find the observed effect represents an increase in overall savings, rather than shifting savings; I also find that consumption of temptation goods (alcohol, tobacco, and sugar) falls, providing evidence that saving informally is difficult and the use of financial institutions to save helps solve self-control problems.
To my godmother, my biggest fan, my parents for their guidance, and Natalie for all her support and attention.
# Contents

## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>iv</td>
<td></td>
</tr>
</tbody>
</table>

## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>vi</td>
<td></td>
</tr>
</tbody>
</table>

## 1 Not(ch) Your Average Tax System: Corporate Taxation Under Weak Enforcement

1.1 Introduction 1
1.2 Tax System and Theoretical Framework 5
1.3 Behavioral Responses and Tax Elasticities 9
1.4 Model-Based Estimation of Revenue and Cost Elasticities 18
1.5 Mechanisms: Evasion Responses 21
1.6 Mechanisms: Dynamic, Real and Avoidance Responses 24
1.7 Conclusion 28

## 2 An Empirical Test of Information Trails: Do Firm Finance and Size Matter for Tax Collection?

2.1 Introduction 52
2.2 Theory 55
2.3 Data and Identification Strategy 59
2.4 Regression Results 62
2.5 Conclusion 66
2.6 Appendix A: Additional Results 74
2.7 Appendix B: Data Sources 79

## 3 Banking on Trust: How Debit Cards Enable the Poor to Save

3.1 Introduction 82
3.2 Institutional Context 85
3.3 Data 87
3.4 Effect of Debit Cards on Stock of Savings 88
3.5 Effect of Debit Cards on Marginal Propensity to Save 90
3.6 Mechanisms ................................................................. 93
3.7 Increase in Overall Savings vs. Substitution ......................... 95
3.8 Does Money Burn a Hole in Your Pocket? .......................... 98
3.9 Robustness .................................................................... 103
3.10 Conclusion .................................................................... 107

Bibliography ................................................................ 129
# List of Figures

1.1 Corporate Tax Revenue and Income ................................................. 31
1.2 Corporate Tax Rate and Income .................................................... 31
1.3 Costa Rica’s Corporate Tax Schedule ............................................. 32
1.4 Bunching Theory ............................................................................ 32
1.5 Firm Density and Average Profit Margin ........................................ 33
1.6 Revenue Bunching Estimation ....................................................... 34
1.7 Cost Discontinuity .......................................................................... 35
1.8 Linear Relation of Average Costs by Revenue ................................ 36
1.9 Effective Tax Rate on Revenue ...................................................... 37
1.10 Quartiles of Profit Margin by Revenue .......................................... 37
1.11 Industry Results ........................................................................... 38
1.12 Excess Mass and Profit Discontinuity by Industry ......................... 39
1.13 Structural Assumption - Stable Profit Margin Distributions ............ 39
1.14 Numerical Estimation of Bunching Behavior ................................ 40
1.15 Dynamic Firm Behavior by Tax Bracket ........................................ 41
1.16 Profit Margin Change for Growing Firms ....................................... 41
1.17 Correction Letters by Revenue (2012) .......................................... 42
1.18 Share of Subsidiaries by Revenue .................................................. 42
1.19 Employment and Wage Bill by Revenue ........................................ 43
1.20 Cost Categories Breakdown ........................................................... 44
1.21 Worldwide Corporate Tax Systems by Income Quintiles ................. 44

2.1 Tax inspection and Sales Tax Compliance Across Firm Size ............... 70
2.2 Tax Inspection and Sales Tax Compliance Across External Financial Use 70
2.3 Distributions of Firm Size Conditional on External Financial Use .......... 71
2.4 Distributions of External Financial Use Conditional on Firm Size .......... 71
2.5 Firm Size Rank-Rank Relation: U.S. Industries on WB Countries ....... 72
2.6 Impact of Firm Size and External Finance on Tax Inspection ............. 72
2.7 Impact of Firm Size and External Finance on Sales Tax Compliance ....... 73

3.1 Low Trust in Banks Around the World .......................................... 108
3.2 Low Trust in Banks by Education Level in Mexico .......................... 108
3.3 Cross-Country Comparison of Trust in Banks and Saving in Financial Institutions 109
3.4 Timing of Roll-out and Data ................................................. 110
3.5 Geographic Coverage and Expansion of Debit Cards .................... 111
3.6 Evolution of Average Balances ............................................. 111
3.7 Difference between Treatment and Control in Average Balances .......... 112
3.8 Withdrawal to Deposit Ratio per Bimester .................................. 112
3.9 Distribution of Withdrawals .................................................. 113
3.10 Difference between Treatment and Control in Net Balances ............. 114
3.11 Difference between Treatment and Control in Marginal Propensity to Save Out of Transfer .................................................. 115
3.12 Trust and Knowledge Over Time with the ATM Card .................... 116
3.13 Balance Checks (Administrative Data) ...................................... 117
3.14 Parallel Pre-Treatment Trends in Household Survey Data ............... 117
3.15 Effect of the debit card on consumption, income, total savings, purchase of durables, and assets ............................................. 118
3.16 Effect of the debit card on different categories of consumption ........ 119
3.17 Placebo Tests ....................................................................... 120
3.18 Difference between Treatment and Control in Marginal Propensity to Save Out of Transfer, Separated by Time with Account ..................... 127
## List of Tables

1.1 Summary of Estimates and Assumptions ........................................ 45  
1.2 Robustness of Bunching Estimates .............................................. 46  
1.3 Cost on Revenue Distance to Threshold Relation ............................ 47  
1.4 Alternative Models for Cost Discontinuity by Revenue ..................... 48  
1.5 Elasticity Estimates - Point of Convergence .................................... 48  
1.6 Model-Based Numerical Estimation: Iteration Steps .......................... 49  
1.7 Industry Level Results 1st Threshold ........................................... 49  
1.8 Dynamic Firm Behavior .................................................................. 50  
1.9 Revenue Shifting at the End of the Fiscal Year ................................. 50  
1.10 Threat of Audit Impact at the Industry Level ................................... 51  
2.1 Tax Inspection Results ................................................................... 68  
2.2 Sales Tax Compliance Results ....................................................... 69  
2.3 Tax Inspection Results with Controls ............................................. 75  
2.4 Sales Tax Compliance Results with Controls ................................... 75  
2.5 Tax Inspection Results when Tax Inspection is Number of Days .......... 76  
2.6 Sales Tax Compliance when Tax Inspection is Number of Days .......... 76  
2.7 Tax Inspection Results Using Alternative Sample for US External Finance . 77  
2.8 Sales Tax Compliance Results Using Alternative Sample for US External Finance . 77  
2.9 Tax Inspection Results with Alternative Non-Rank Measures of Firm Size and External Finance .................................................. 78  
2.10 Sales Tax Compliance Results with Alternative Non-Rank Measures of Firm Size and External Finance ........................................ 78  
2.11 Summary of Financing Sources ..................................................... 80  
3.1 Comparison of Baseline Means ....................................................... 121  
3.2 Trust and Knowledge Over Time with the ATM Card ........................ 122  
3.3 Other Barriers to Saving Informally ............................................... 123  
3.4 Supply-Side Response ................................................................. 124  
3.5 Crime as a Barrier to Saving Informally ......................................... 125  
3.6 Change in Savings and Assets After Receiving Card ......................... 126  
3.7 Computation of Mechanical Effect ................................................ 128
Acknowledgments

I thank my advisors Emmanuel Saez and Edward Miguel for their guidance and many insightful suggestions and for providing such a fruitful and open environment for research. I also have benefited greatly from comments of faculty at UC Berkeley, including Alan Auerbach, Fred Finan, Paul Gertler, Andres Rodriguez Clare, Danny Yagan and Gabriel Zucman.

I thank my co-authors Sean Higgins, Anders Jensen, Enrique Seira and Mauricio Soto for their hard work and motivation. I have benefited from many work sessions and discussions with my classmates, Juan-Pablo Atal, Youssef Benzarti, Dorian Carloni, Natalie Cox, Simon Galle, David Silver and Moises Yi. I am also indebted to Miguel Almunia, Jorge Richard Munoz Nunez, Gerard Padro-i-Miquel, Dina Pomeranz, Alvaro Ramos-Chavez and Oscar Fonseca Villalobos who greatly helped me at various stages of the PhD process.

I have received financial support from the Burch Center for Tax Policy, the Center for Effective Global Action, the Center for Equitable Growth and UC MEXUS.

Finally, I am very grateful to my family for their teaching, advice and support through all these years. I thank my parents, Costas and Pascale, my brother, Alexis, for frequent visits, my godmother, Deppy, and my cousins, Angelina, Georgina and Xavier.
Chapter 1

Not(ch) Your Average Tax System: Corporate Taxation Under Weak Enforcement

with Mauricio Soto

1.1 Introduction

Boosting tax revenue is an important challenge for lower-income countries, which only collect 20% of their GDP in taxes, compared to 35% on average for OECD countries (Besley and Persson 2013a). Low levels of tax revenue limit countries’ ability to redistribute income and invest in public goods\(^1\) and have been highlighted as major impediments to inclusive growth (United Nations 2014).

Recent research shows that tax revenue can be increased by designing tax systems and incentives which consider the constraints faced by tax administrations in low-income countries (Best et al. 2014, Pomeranz 2015a, Khan, Khwaja, and Olken 2015, Naritomi 2015). In particular, tax administrations have difficulty monitoring income and transactions, due to the structure of their economy and lower fiscal capacity\(^2\), which leads to informality and tax evasion. Given these constraints, we ask whether developing countries can rely on the

\(^1\)Taxation can also relax aid-dependence: in the last decade revenue increases have dwarfed foreign aid flows, and even in Sub-Saharan Africa, governments collect $10 in own-revenue for every dollar in foreign aid (World Bank 2013).

\(^2\)Information and capacity constraints are certainly not the only reasons for differences in tax revenue across income levels: an exciting new literature studies preferences for redistribution and tax morale as other candidate explanations (Luttmer and Singhal 2011, Luttmer and Singhal 2014, Kleven 2014). Moreover, Olken and Singhal 2011 show that households in low-income countries pay substantial informal taxes, which could get substituted for formal taxes through the development path.
corporate income tax to raise revenue, and how should they optimally design it? Two important features in the design of the corporate income tax are the tax rate and definition of the tax base. When tax evasion is a primary concern, firms’ behavioral response to higher tax rates can be large, which limits the range of optimal tax rates. In addition, the standard corporate tax base, which allows for all costs to be deducted, might not be desirable since it provides evasion opportunities both on the revenue and on the cost margin.

In this paper we make four contributions. First, we show that even for a middle-income country, here Costa Rica, the elasticity of declared profit with respect to the tax rate for small and medium firms is very large and the top of the Laffer curve appears to correspond to a tax rate below 25%.

Second, we separate the profit elasticity into changes in declared revenue versus declared cost, and show that the cost elasticity is substantially larger than the revenue elasticity. This highlights a novel mechanism: even though manipulating business revenue is relatively difficult, observing business costs is so hard for the tax administration, that the standard profit tax collects little income. Firms’ ease in understating profit by misreporting cost rationalizes the use of broad tax bases and policies determined by revenue, instead of profits. Such policies, while rare in rich countries, are in fact often observed in lower-income countries\(^3\).

Third, we provide evidence that tax evasion is a driver of our results, while real effects and avoidance responses seem limited. Taken together, the results imply that Costa Rican firms evade up to 70% of taxes on their profits when faced with a 30% tax rate.

Fourth, we develop a new methodology, which combines bunching at tax notches and a discontinuity, to separate the profit response into revenue and cost responses.

Estimating the parameters needed to evaluate the design of tax policy in developing countries is challenging. However, new estimation techniques in public economics (Saez 2010, Chetty et al. 2011, Kleven and Waseem 2013), combined with improved access to large and high quality administrative datasets, are increasing researchers’ capacity to address these questions. Even five years ago, many low and middle income countries had neither online tax fillings with automatic quality checks, nor integrated data systems covering the universe of taxpayers. As part of this small but growing literature, we use rich administrative data and the design of the corporate tax in Costa Rica to study small and medium firm’ behavioral responses to taxation.

Our unique setup allows us to estimate the elasticity of corporate profits with respect to the net of tax rate and to separate the profits response into revenue and cost responses. While most corporate tax systems tax profits at a flat rate, Costa Rica imposes increasing average tax rates on profits as a function of firms’ revenue. The average tax rate jumps from 10 to 20% at the first revenue threshold and from 20 to 30% at the second threshold. The change in average tax rate generates two distinct behavioral responses. First, some firms reduce their revenue below the threshold, in order to lower the tax rate they face on their

\(^3\)Examples of such policies are presumptive taxes, which tax revenue instead of profits, enforcement and registration thresholds determined by revenue (e.g. Large Taxpayers Units), and corporate tax systems with different rates as a function of revenue.
whole profit base. This generates excess mass in the firm distribution just below the threshold and missing mass above it, which we use to measure the elasticity of revenue, following the notch estimation technique in Kleven and Waseem 2013. Second, firms remaining above the revenue threshold respond to the higher tax rate by sharply dropping their reported profits. This is evidenced by a large discontinuity in average profits on either side of the threshold, when plotted against revenue. The response of profits to higher average tax rates is a mix of revenue and cost responses. Using the revenue elasticity, estimated with the bunching methodology, we hold revenue responses constant, such that the remaining profits discontinuity only identifies changes in reported costs. Finally, by combining the revenue and cost responses we estimate the elasticity of profits with respect to the net of tax rate. The resulting elasticities are very large: 5 at the first threshold and 3 at the second threshold. These are an order of magnitude higher than those of small firms in OECD countries, estimated at around 0.5 (Devereux, Liu, and Loretz 2014, Patel, Seegert, and Smith 2015) and severely constrain the range of optimal tax rates: given the current policy and enforcement environment, rates above 25% are on the wrong side of the Laer curve for Costa Rica. 

The reduced form estimation provides a robust measure of the profit elasticity: an overestimate of the revenue responses would mechanically underestimate cost responses, leaving profit responses practically unchanged. In contrast, the relative shares of revenue and cost responses are not as robust to this estimation. Under heterogeneity in revenue elasticities the bunching method recovers the response of the highest revenue elasticity firm, hence providing an upper bound to the true revenue elasticity. This mechanically implies that the estimated cost elasticity is a lower bound of the average cost elasticity. If we make no assumption about the counterfactual distribution of profit margin by revenue, then our estimates are the tightest possible bounds on the revenue and cost elasticities. However, if we assume that the counterfactual distribution of profit margin by revenue is constant around the threshold, then these elasticities are rejected by the data, as they predict substantially more bunching than observed. With this new counterfactual, we develop a model-based numerical method to estimate the revenue and cost responses, as a joint function of the revenue distance to the threshold and the costs of the firm. This new estimation recovers the average elasticity of revenue, while dealing with two-dimensional selection into bunching. Our preferred estimates show that cost responses account for 71% of the drop in reported profits, while revenue responses only for 29%. The relative ease to manipulate cost, compared to revenue, rationalizes the use of tax bases with few deductions and policies based on revenue instead of profits, often observed in lower-income countries.

Behavioral responses due to evasion responses, real production responses or avoidance have the same impact on revenue collection, but call for different policy action. To study the mechanisms of firms' responses we draw on rich administrative datasets. In addition to the corporate tax returns, we use information on audits, the central bank's firm registry, social security data on wages and employment, and monthly sales tax receipts. We find evidence that tax evasion is a key driver of responses for bunching firms: these firms are significantly more likely to display inconsistencies with third-party reported information and adjust revenue upwards following audit threats at the industry level. Furthermore, all statistical tests
for real effects and avoidance responses are rejected: the social security data shows no discontinuity in the number of employees and wage bill at the threshold, monthly sales receipts do not display evidence of revenue shifting across fiscal years and the registry of economic groups only shows very modest evidence of firms dividing themselves into smaller firms. In a literature that often remains agnostic on the mechanisms of behavioral response, our paper takes innovative steps to support tax evasion as the key mechanism. Further, if we assume that profit responses are only due to evasion, then firms facing a 30% tax rate evade taxes on as much as 70% of their profits.

Our paper contributes to the literature on tax design and tax enforcement in low and middle income countries. We are the first to estimate a corporate tax elasticity in this context and find that it is substantially higher than in rich countries: Devereux, Liu, and Loretz 2014 for the UK and Patel, Seegert, and Smith 2015 for the US both estimate corporate elasticities of 0.5. These estimates are particularly relevant for comparison since they also concern small and medium firms. To make sense of the magnitude of our results, it is important to consider the weak enforcement environment - an expanding empirical literature (Pomeranz 2015a, Naritomi 2015) shows that difficulties in monitoring transactions and missing third-party information lead to large evasion rates in developing countries. Even in Denmark, Kleven et al. 2011a estimate tax evasion rates as high as 40% on income not subject to third-party reporting. In Costa Rica, firms facing a 30% tax rate have an implied evasion rate on profits of 70%, which is comparable to the 60% evasion rate estimated with micro-data for large Pakistani firms by Best et al. 2014 and the 65% average evasion for Costa Rican firms, estimated with aggregate data by the IMF 2012.

Our paper also builds on a recent literature on the two-dimensional aspect of the corporate tax declaration and provides the first separation of the elasticity of profit into cost and revenue responses. The relative ease to manipulate costs, compared to revenue, complements the findings of Carrillo, Pomeranz, and Singhal 2014 and Slemrod et al. 2015: both studies show that following tighter enforcement on revenue by the tax administration, firms’ reported revenue increases to limit the risk of an audit, but reported costs also increase, such that overall profits and tax liability are practically unchanged. Regarding tax policy, the result supports theoretical work on the desirability of “production inefficient” tax instruments under evasion (Emran and Stiglitz 2005, Gordon and Li 2009a) and the empirical findings of Best et al. 2014: when evasion opportunities are large, limiting deductions or switching to a turnover tax with no deductions could be optimal.

Finally, from a methodological standpoint, we contribute to the literature using discon-


\[4\text{A large literature summarized in Auerbach, Devereux, and Simpson 2010 studies firms' responses to the tax code but few studies estimate corporate tax elasticities. An exception is Gruber and Rauh 2007 who estimate an elasticity of 0.2 for large US corporations, using panel data and an instrument for the effective tax rate change. With a similar methodology, Dwenger and Steiner 2012 estimates a corporate tax elasticity of 0.5 in Germany.}\]

\[5\text{Kleven et al. 2011a and Slemrod, Blumenthal, and Christian 2001 use randomized audits to estimate tax evasion - the latter study also finds that tax evasion is concentrated among self-employed individuals in Minnesota.}\]
timinuities in tax design to identify structural parameters. Saez 2010 and Chetty et al. 2011 developed the framework to recover taxable income elasticities from kink points, which was extended to notches by Kleven and Waseem 2013. In our setting, tax notches are determined by revenue, but the tax rate applies to profit: we show how to use this variation to estimate the profit elasticity, and for the first time in the literature, to separate revenue and cost elasticities. Revenue-dependent policies, such as registration and enforcement thresholds⁶, are common in low-income countries, and our methodology could be applied to these settings.

The remainder of the paper is organized as follows. Section 1.2 introduces the tax system and provides a theoretical framework. Section 1.3 presents the data, methods and results. Section 1.4 adds structure to refine the previous results. Section 1.5 shows evidence of evasion as a key mechanism, while Section 1.6 rejects some specific real and avoidance mechanisms. Section 1.7 discusses policy implications and concludes.

### 1.2 Tax System and Theoretical Framework

#### Corporate Tax System in Costa Rica

Figure 1.3 presents the Costa Rican corporate tax schedule. A corporation pays an average tax rate of 10%, 20% or 30% on its profit as a function of its revenue - firms with revenue below the first threshold⁷ face a 10% average tax rate, firms with revenue in between the two thresholds face a 20% rate and firms with revenue above the second threshold face a 30% tax rate. A unique feature of this tax design is that the determinant of the tax rate, revenue, is different from the tax base, profits. Importantly the revenue thresholds only determine tax liability and are not used to determine any other policy. Loss carry-forwards are limited to the manufacturing sector and a three year period, while loss carry-backs are never allowed. The current tax design was implemented in 1988 and has been unchanged since then. Prior to 1988, corporations were taxed at increasing marginal tax rates on profits, with multiple brackets ranging from 5 to 50%: the original tax reform of 1987 proposed a flat 30% tax rate on profits, but was strongly contested by small and medium enterprises who would, on average, face an increase to their tax liability (Naranjo and Zúñiga 1990). The political pressure to apply preferential rates to these firms led to the addition of two tax rates, determined by firms’ revenue.

#### Theoretical Framework: Baseline

We develop a simple theoretical framework of firm behavior, to consider simultaneous revenue and cost responses to tax changes, and allow for heterogeneity in revenue and in cost for a

⁶For example, Almunia and Lopez-Rodriguez 2015 study the impact of an enforcement threshold, the Large Taxpayer Unit, on Spanish firms’ reporting behavior.

⁷In 2014, the revenue thresholds are 49,969,000 and 10,0513,000 Colones, corresponding to 150,000 and 300,000 USD in Purchasing Power Parity. The thresholds are indexed on inflation and therefore grow 5% yearly, on average.
given revenue. A representative firm decides how much to produce and can simultaneously evade taxes by under-reporting revenue and over-reporting cost. When evading taxes, the firm incurs resource costs and risks detection. Under this simple framework, we discuss the impact of the Costa Rican corporate tax system on firm behavior. We then derive the empirical predictions of the model and extend it to discuss heterogeneity in revenue and cost elasticities.

Consider a firm that produces good \( y \), subject to a convex cost function \( c(y) \). The costs incurred by the firm are fully tax-deductible and therefore a flat tax rate on profit is non-distortionary\(^8\). The firm can under-report revenue, such that revenue evasion is \( (\tilde{y}_i - y_i) \), where \( \tilde{y}_i \) is declared revenue, and over-report cost, such that cost evasion is \( (\tilde{c}_i - c_i) \), where \( \tilde{c}_i \) is declared cost. In doing so it incurs resource costs\(^9\) and risks detection: this generates a convex cost of evasion \( R(y_i - \tilde{y}_i, \tilde{c}_i - c_i) \). The convexity captures the idea that detection is increasingly likely for large amounts evaded. Finally, the firm faces the tax rate \( \tau \) that applies to declared profit, \( \pi = \tilde{y}_i - \tilde{c}_i \). The firm’s expected profits are therefore:

\[
E\pi_i = y_i - c(y_i) - \tau (\tilde{y}_i - \tilde{c}_i) - R(y_i - \tilde{y}_i, \tilde{c}_i - c_i) \tag{1.1}
\]

To generate heterogeneity and simplify the exposition we make two more assumptions. First, we assume that the cost function takes the following form \( c(y_i; \phi_i, \alpha_i) = \alpha_i + \frac{k(y_i)}{\phi_i} \), where \( \alpha_i \) is the fixed cost, equivalent to a demand shifter, and \( \phi_i \) is a productivity parameter, which scales variable costs \( k(y_i) \). Second, we assume that the cost of evasion function is separable in revenue and cost evasion such that \( R(y_i - \tilde{y}_i, \tilde{c}_i - c_i) = h(y_i - \tilde{y}_i) + g(\tilde{c}_i - c(y_i)) \). Under these conditions the firm’s expected profits are:

\[
E\pi_i = y_i - c(y_i; \phi_i, \alpha_i) - \tau (\tilde{y}_i - \tilde{c}_i) - h(y_i - \tilde{y}_i) - g(\tilde{c}_i - c(y_i)) \tag{1.2}
\]

The firm maximizes expected profits, by choosing the triple of revenue to produce, revenue to declare and costs to declare \( (y_i, \tilde{y}_i, \tilde{c}_i) \). An interior optimum satisfies the following first order conditions:

\[
1 = \frac{k'(y_i)}{\phi_i} \tag{1.3}
\]

\[
h'(y_i - \tilde{y}_i) = \tau \tag{1.4}
\]

\[
g'(\tilde{c}_i - c_i) = \tau \tag{1.5}
\]

Equation (1.3) determines the revenue produced \( y \). Since, in our model, taxation is non-distortionary, the production decision is independent of the tax rate. Equations (1.4) and

\(^8\)We do not pretend that corporate taxation is generally non-distortionary but make this assumption for the tractability of the model. The corporate tax is non-distortionary in a cost of capital model (Jorgenson and Hall 1967) with immediate expensing: if all costs, including returns to capital, are immediately deductible, then the corporate income tax is a tax on pure profits and does not impact production decisions.

\(^9\)Resource costs from evasion include, forgoing business opportunities with formal firms, keeping multiple sets of accounting records and limiting interactions with the financial sector. See Chetty 2009a for a discussion.
(1.5) state that the marginal return to revenue and cost evasion, \( \tau \), equals the marginal cost, which is a function of the amount evaded. Firm revenue is a function of its productivity draw \( \phi_i \) but independent of the fixed cost draw \( \alpha_i \), such that \( \frac{d\pi_i^e}{d\phi_i} > 0 \) and \( \frac{d\pi_i^e}{d\alpha_i} = 0 \). Firm costs are given by \( c^i(y^i; \phi_i, \alpha_i) = \alpha_i + k(y^i) \) and depend both on the productivity draw \( \phi_i \) and the fixed cost \( \alpha_i \), such that \( \frac{d\pi_i^e}{d\phi_i} > 0 \) and \( \frac{d\pi_i^e}{d\alpha_i} > 0 \). Finally, we define profit margin as profit over revenue, \( \pi_{\text{margin}} = \frac{y^r - c^f}{y^r} \), and is determined jointly by \( \phi_i \) and \( \alpha_i \).

Under a continuous and differentiable joint distribution of productivity and fixed cost parameters \( f_0(\phi, \alpha) \) the distribution of revenue and cost is smooth. We assume that the cost of evasion functions \( h(y_i - \bar{y}_i) \) and \( g(c_i - c) \) are continuous and differentiable and therefore the distributions of reported revenue, reported costs and reported profit margins are also smooth.

**Theoretical framework: Impact of the tax system**

A noteworthy aspect of Costa Rica’s corporate schedule is that the average tax rate applied on profits increases from \( \tau \) to \( \tau + d\tau \) when firms declare revenue above the threshold \( y^T \). Tax liability is a function of declared revenue \( \bar{y} \) and declared costs \( \bar{c} \):

\[
T(\bar{y}_i - \bar{c}; \bar{y}_i) = \tau(\bar{y}_i - \bar{c}) \quad \text{if} \quad \bar{y}_i \leq y^T \\
T(\bar{y}_i - \bar{c}; \bar{y}_i) = (\tau + d\tau)(\bar{y}_i - \bar{c}) \quad \text{if} \quad \bar{y}_i > y^T \\
T(\bar{y}_i - \bar{c}; \bar{y}_i) = 0 \quad \text{if} \quad \bar{y}_i - \bar{c} \leq 0
\]

We consider that the above tax system is imposed as a tax reform over a previously flat corporate tax at rate \( \tau \). Since only the productivity parameter \( \phi_i \) determines firm revenue and all firms face the same cost of evasion, there exists a productivity threshold \( \bar{\phi} \) such that a firm with productivity \( \phi_i = \bar{\phi} \) reports revenue exactly equal to the threshold \( \bar{y}_i = y^T \), and all firms with \( \phi_i \leq \bar{\phi} \) declare revenue below the threshold \( \bar{y} \leq y^T \). These firms are not affected by the tax change. For firms with \( \phi_i > \bar{\phi} \) there are two possible responses: (1) reduce revenue, declared or real, by an amount such that the new revenue equals the threshold (the “bunchers”) or (2) stay above the threshold and face a higher tax rate. These firms then change their reporting revenue and cost such that the marginal cost of evasion equals the new tax rate.

Firms choose one of two responses depending on their productivity and fixed cost draw: for every productivity draw \( \phi_i \) in an interval \([\underline{\phi}, \bar{\phi}_{\alpha}]\) there exists a fixed cost \( \alpha_i \) such that all firms within the interval \([\phi_i, \bar{\phi}_\alpha]\) bunch at the threshold. \( \bar{\phi}_\alpha \) is determined by the indifference condition between expected profits at the threshold and expected profits at the interior solution, \( E_\pi^{\text{Threshold}}(y, \bar{y}^T, \bar{c}; \bar{\phi}_\alpha, \alpha) = E_\pi^{\text{Interior}}(y, \bar{y}, \bar{c}; \bar{\phi}_\alpha, \alpha) \). Firms with \( \phi_i > \bar{\phi}_\alpha \) remain above the threshold and adjust their reporting behavior.

To illustrate the effect of costs on the bunching decision we consider a firm with productivity \( \phi_i > \bar{\phi} \) and fixed costs \( \alpha_i \) mapping into true revenue and cost \((y_0, c_0)\) and reported revenue and cost \((\bar{y}_0, \bar{c}_0)\) such that \( \bar{y}_0 > y^T \) before the tax change. To reach the threshold the
firm can reduce declared income with a combination of real and reporting behavior. Real income reduction is \(dy\) and reported income is \(d\tilde{y}\). The total change in revenue is \(\Delta y = dy + d\tilde{y}\) such that \(\Delta y\) is the revenue distance to the threshold, \(\Delta y = \tilde{y}_0 - y^T\). We compare the firm’s utility when it reports revenue at the threshold versus when it report its pre-tax change revenue - and approximate the expected gains from bunching as:

\[
E \text{ Gains} \approx d\tau (y^T - \tilde{c}_0) + \Delta y (\tau + d\tau) - d\tilde{y}.h'(y_0 - \tilde{y}_0 + d\tilde{y}) - dy.[1 - c'(y_0 - dy)] \quad (1.7)
\]

Where we have used the envelope condition and ignored intensive margin changes past the threshold. The first term of equation (1.7) is a noteworthy feature of the Costa Rican setting: it shows that the gains from lowering revenue to reach the threshold are proportional to the change in the tax rate \(d\tau\) and to the firm’s declared tax base at the threshold, \(y^T - \tilde{c}_0\). Therefore variation in cost, due to fixed cost heterogeneity, generate different incentives to bunch for firms of equal productivity.

The other terms of equation (1.7) state that the firm directly gains by not paying taxes on undeclared and non-produced revenue \(\Delta y\), but incurs larger resource costs, due to the additional revenue under-reporting (evasion responses) and losess profit due to its lower production level (real responses). Note that if all responses are due to revenue under-reporting, equation (1.7) simplifies to:

\[
E \text{ Gains} \approx d\tau (y^T - \tilde{c}_0) + d\tilde{y}(\tau + d\tau) - h'(y_0 - \tilde{y}_0 + d\tilde{y})
\]

**Prediction 1: Bunching at the revenue thresholds**

From the distribution of productivity and fixed cost parameters \(f(\phi, \alpha)\) we obtain a direct mapping into the distribution of declared revenue and declared costs \(\psi_0(\tilde{y}_0, \tilde{c}_0)\) such that the total number of firms bunching at the revenue threshold is:

\[
B = \int_{\tilde{c}_0}^{c_0} \int_{\tilde{y}_0 = y^T}^{y^T + \Delta y(\tilde{c}_0)} \psi_0(\tilde{y}, \tilde{c})d\tilde{y}.d\tilde{c} \quad (1.8)
\]

With knowledge of the joint distribution of revenue and cost we can estimate the elasticity of revenue \(\epsilon_y\) that generates a given amount of bunching. Absent the counterfactual cost distribution we can still estimate the revenue response of the marginal buncher, defined as the firm with the maximal revenue change. For firms with the same revenue distance to the threshold the marginal buncher is the firm with the lowest declared costs: given a support of costs \([c_0; \bar{c}_0]\) the marginal buncher’s revenue response is \(\Delta y^{mb} = \Delta y(\tilde{c}_0 = c_0)\). With knowledge of the lower support of the distribution of \(c_0\), we can identify the revenue response of the marginal buncher as the maximum revenue response, which in the model corresponds to the response of the firm with the lowest costs. Under homogeneous revenue elasticities the marginal buncher’s revenue elasticity and the average revenue elasticity are the same.

**Prediction 2: Missing mass above the thresholds but no strictly dominated region**
A corollary to the first prediction states that some revenue intervals past the threshold display missing density, which corresponds to the excess density at the threshold. At each revenue level past the threshold, the missing density is a function of the distance to the threshold and the cost distribution at that revenue level. In the standard notch setting (Kleven and Waseem 2013) there is a deterministic dominated revenue interval just above the threshold: firms that report revenue in that interval are making an irrational decision under any preferences, since lowering production would increase their after tax profits. Whereas, for the Costa Rican notches, only a subset of firms with sufficiently low costs are dominated. For example, a firm with zero profits has no incentives to lower its revenue since its tax liability is already null. Being dominated is not only a feature of the revenue distance to the threshold but also of costs and hence firm specific. As a consequence, even in a frictionless world, there will be firm density in revenue intervals just past the threshold.

**Prediction 3: Increased Revenue and Cost Evasion Past the Thresholds**
Infra-marginal firms do not bunch at the revenue threshold but face an increase in the marginal return to evasion, which jumps from $\tau$ to $\tau + d\tau$. They respond to the tax hike by increasing revenue and cost evasion such that the marginal resource costs of each evasion type equals the new tax rate. As a consequence firms above the threshold declare less revenue and more costs than under the counterfactual. As a consequence observed profits and profit margins by revenue, jump downwards discontinuously at the threshold.

**Prediction 4: Excess Profit at the Thresholds (Under evasion responses)**
On the one hand, firms selecting into bunching have higher profit than the average firm (Selection effect). On the other hand, by lowering declared revenue to reach the threshold, bunchers lower their profits (Evasion effect). Theoretically, the average declared profit margin of firms at the threshold could be higher or lower than that of firms below the threshold, depending on the variance of the distribution of costs, in the revenue bins above the threshold. Under homogeneous costs (no variance) then the Evasion effect dominates and the average observed profit margin of bunchers is lower than that of firms below the threshold. With sufficient heterogeneity in the cost distribution the Selection effect dominates and bunching firms display excess profit at the threshold. The domination of the selection effect is better understood from equation 1.7: while gains from bunching are linear in the firm’s costs, the cost of bunching are convex in the firm’s revenue distance to the threshold, due to the convexity of the resource cost of evasion. Therefore, for a sufficiently large revenue distance to the threshold, the revenue change is small compared to the cost difference between selected bunchers and the average firm.

### 1.3 Behavioral Responses and Tax Elasticities

The goal of this section is to estimate the firms’ elasticity of profit with respect to the net of tax rate and separate the profit response between changes in revenue versus changes in costs. To this end, we develop an estimation procedure that deals with the interlinked
revenue and cost responses, without assuming a functional form for firms’ utility. Elasticities are defined with respect to the net of tax rate and when discussing size and bounds we refer to their absolute value. Under mild assumptions, we can estimate the profit elasticity and obtain an upper bound on the revenue elasticity and a lower bound on the cost elasticity. In section (1.4) we impose additional structure and obtain tighter estimates for each of the revenue and cost elasticities. The profit elasticity, which combines revenue and cost responses, is stable across the different estimation strategies\(^{10}\). We summarize the different methodologies, elasticity estimates and assumptions in Table 1.1.

Our identification relies on two assumptions: first, absent the tax rate increase past the threshold, the distribution of firms by revenue would be smooth and continuous and therefore can be approximated by a flexible polynomial\(^{11}\). Second, average cost by revenue would not jump discontinuously precisely at this revenue-size. Under these assumptions we develop a three step methodology: in a first step, we use bunching at the revenue thresholds to identify revenue responses to higher tax rates. In a second step, we use the discontinuity in average cost by revenue, on both sides of the thresholds, to estimate the cost response. The novelty of the approach is to adjust the cost discontinuity to take intensive margin revenue responses into account, using the revenue elasticity estimated in the first step. We then recover the average increase in reported cost at the threshold, holding revenue responses constant. In a third and final step, we combine the revenue and cost responses to compute the profit response at the threshold.

**Setting and Data**

Costa Rica is a middle-income country, with GDP per capita slightly under 15,000 US dollars at purchasing power parity, and is considered to have stable and well-functioning institutions for its income level. Its government collects 21\% of its GDP in revenue, of which 14\% is tax revenue and 7\% social security contributions. Following several failed attempts at tax reforms, Costa Rica’s revenue collection has stagnated in the last decade: increasing it is considered a key priority by the main political parties, in order to reduce the large deficit.

We base our study on administrative data from the Ministry of Finance (Ministerio de Hacienda) and have access to the universe of corporate tax returns over the 2008-2014 period. All registered corporations are required to submit yearly tax declaration D101 (“Declaracion Jurada del Impuesto Sobre la Renta”) and report their profits, revenue and costs. The tax declaration can be filled electronically since 2008, and a large majority of firms have opted for this format.

The data consists of 617,929 firm-year observations and 222,352 unique firms. As a whole, the

---

\(^{10}\)The estimation strategy induces a mechanical negative correlation between the estimates of the revenue and cost elasticities. Since the profit elasticity combines the two, it is very robust to their respective separation.

\(^{11}\)The ability to approximate the counterfactual density with a flexible polynomial is the standard assumption in the bunching literature and we show that in our data the parametric choices (polynomial order, bunching interval limits, etc.) have little effect on estimated parameters.
Corporate income tax raises 18% of tax revenue, around 2.5% of GDP. The firms we study are small enterprises with yearly revenue below 150 million Costa Rican Colones ($450,000 in PPP). They represent 85% of the 80,000 firms filling taxes in a given year and declare 25% of total profits, which generates 15% of corporate tax revenue.

Figure 1.5 presents the key features of the data by revenue bins of half million CRC, pooling all years together. Panel A shows the number of firms by revenue. We observe a clear excess mass below each revenue threshold and missing mass just above, as predicted by theory. Panel B shows the average profit margin by revenue, where profit margin is defined as profit over revenue. Profit margin by revenue resembles a downward step function - constant within a given tax bracket and jumping down at the thresholds. Average profit margin within the first tax bracket is 16%, 7-8% in the second bracket and 4-5% in the third bracket. We also observe that firms reporting revenue at the thresholds display profit margins in excess of 22% and 9%, respectively at the first and second thresholds. As discussed in theory (Prediction 4), this could arise from the selection into bunching of firms with low costs.

The estimation strategy combines the distributions of Figure 1.5 to estimate profit elasticities and separate them between revenue and cost responses. Intuitively, the excess mass at the revenue thresholds provides evidence of revenue responses while the jump in profit margin combines revenue and cost responses. Therefore, in a first step we apply the bunching methodology to the firm density to estimate revenue elasticities. In a second step, we use the discontinuity in profit margin on either sides of the thresholds to estimate the cost elasticity, holding constant revenue responses. In a third step, we combine the revenue and cost responses to obtain profit elasticities.

**Revenue elasticity estimation**

**Bunching methodology**

To estimate the revenue elasticity, we use the distribution of firms by revenue and the point of convergence method described in Kleven and Waseem 2013. We slice the data in half million CRC bins. To obtain the counterfactual density, we fit a flexible polynomial of degree five:

\[
F_j = \sum_{k=0}^{5} \beta_k (y_j)^k + \sum_{i=y_l}^{y_u} \delta_i 1(y_j = i) + \nu_j
\]

where \(F_j\) is the number of firms in revenue bin \(j\), \(y_j\) is the revenue midpoint of interval \(j\), \([y_l, y_u]\) is the excluded region and \(\delta_i\)'s are dummy shifters for the excluded region. We use

---

12 This share concerns tax revenue only and excludes social security contributions
13 The notch estimation builds upon the kink method of Saez 2010 and Chetty et al. 2011
14 The order of the polynomial is chosen to maximize Akaike's criteria. Table 1.2 shows the impact on the results of using different orders of polynomial.
Chapter 1: Corporate Taxation Under Weak Enforcement

the estimated $\beta_k$'s to obtain the counterfactual firm distribution by revenue absent the tax change:

$$\hat{F}_j = \sum_{k=0}^{5} \hat{\beta}_k (y_j)^k \quad (1.10)$$

The estimation procedure requires that the excess mass below the threshold ($E$) equals the missing mass past the threshold ($M$), defined as:

$$\hat{E} = \sum_{j=y_l}^{y^*} (F_j - \hat{F}_j) \quad \text{and} \quad \hat{M} = \sum_{j=y^*_l}^{y_u} (\hat{F}_j - F_j) \quad (1.11)$$

Where $y^*$ is the revenue threshold and the bounds of the excluded region $[y_l, y_u]$ are obtained as follows: the lower limit $y_l$ is chosen by the researchers as the revenue bin where excess density starts appearing\(^{15}\). The upper limit, $y_u = y^* + dy$, is estimated using the identity that the excess mass ($E$) has to equal the missing mass ($M$). Starting from $y_u$ just above the threshold, we estimate equation (1.9) and compute $\hat{E}$ and $\hat{M}$. For a low value of $y_u$, the excess density is much larger than the missing density ($\hat{E} > \hat{M}$). We iteratively increase $y_u$ until the excess mass equals the missing mass ($\hat{E} = \hat{M}$). The estimated upper bound, $y_u$, is the revenue of the marginal firm responding to the tax change. Under heterogeneity in revenue elasticities, this is the response of the highest elasticity individual and therefore provides an upper bound on revenue responses.

By forcing the excess mass to equal the missing mass, the point of convergence method generates two potential concerns\(^{16}\). First, it assumes that there are no extensive margin responses. Extensive responses could occur if firms decided to become informal when faced with higher tax rates. This would generate additional missing mass past the threshold and imply that $E < M$. In our setting extensive margin responses should play a limited role, as Costa Rica is one of Latin America’s country with the least informality (ILO 2012), and it is unlikely that firms with growing revenue and already registered decide to reverse back to informality, after increasing their revenue past the threshold. In terms of results, if extensive margin responses exist, then the true revenue elasticity is smaller than the estimated one. Second, the standard bunching method ignores intensive margin revenue responses past the threshold. Intensive responses imply that, above the threshold, the counterfactual firm distribution should be higher than the observed distribution. We take into account this second order effect by shifting the counterfactual distribution above the threshold with the factor implied from the estimated revenue elasticity. The intensive margin adjustment occurs simultaneously with the point of convergence method and the iterative process of determining the upper bound on revenue $y_u$. In our setting where elasticities are substantial, this adjustment does have a modest impact on the results, reducing slightly the estimated revenue response.

In the case of a notch, and in particular of a notch with two-dimensional incentives,

\(^{15}\)We show in table 1.2 that changing $y_l$ does not impact the results.

\(^{16}\)These limitations are also noted in Kleven and Waseem 2013.
Chapter 1: Corporate Taxation Under Weak Enforcement

revenue distance to threshold and costs, obtaining the change in the marginal tax rate is less straightforward than with a kink. Given the tax liability \( T(y - c; y) \), we define the implicit marginal tax rate \( \tau^* \), for an increase in revenue \( dy \), as the change in tax liability over the change in revenue:

\[
\tau^* = \frac{T(y^* + dy) - T(y^*)}{dy} = \frac{(\tau_0 + d\tau)(y^* + dy - c) - \tau_0(y^* - c)}{dy}
\]

\[
\tau^* = (\tau_0 + d\tau) + \frac{d\tau(y^* - c)}{dy}
\]

(1.12)

Where \( \tau_0 \), \( d\tau \) and \( y^* \) are known parameters and \( dy \) is estimated with the bunching point of convergence method. However, the cost of the marginal buncher \( c \) is unknown. From the theory section we know that the marginal buncher is the firm with the lowest cost, within its revenue bin. Therefore, the marginal buncher should have costs in the 1st percentile of the cost distribution for its revenue bin. To ensure that we obtain an upper bound on the revenue elasticity we assume the cost of the marginal buncher are at the 10th percentile. With this assumption, we obtain \( c \) and can compute the implicit marginal tax rate \( \tau^* \), given the estimate of the revenue response \( dy \). The revenue elasticity is then defined as:

\[
\epsilon_{y,1-t} = \frac{\text{%change revenue}}{\text{%change (net of tax rate)}} = \frac{dy}{y^*} \frac{(1 - t_0)}{(t^* - t_0)}
\]

(1.13)

Bunching results

Figure 1.6 shows the distribution of firms by revenue and the counterfactual density, estimated from the polynomial fit around each threshold\(^{17}\). The estimated parameters are displayed in the top right corner of each panel. For the first threshold (Panel A), the excess mass is 2.3 times the counterfactual, meaning that there is 3.3 times the density that should be expected. In the absence of the notch, the marginal buncher would have an income of 58.3 million CRC, 16% higher than the threshold. For the second threshold (Panel B), the excess mass is 1.1 times the counterfactual and the marginal buncher has revenue of 107.7 M CRC, 7.6% higher than the threshold.

Given the estimated revenue responses, we compute with equation (1.12) the implicit marginal tax rate faced by the marginal buncher: at the first threshold the resulting revenue elasticity with respect to the net of tax rate is 0.33. This implies that firms respond to a 10% reduction in the net of tax rate by reducing reported revenue by 3.3%. At the second threshold, the elasticity of revenue is 0.08. Table 1.5 reports the parameters and the resulting revenue elasticities at each threshold. Standard errors are estimated from 1,000 bootstrap iterations from residuals resampling of the polynomial fit\(^{18}\).

\(^{17}\)Due to the intensive margin adjustment above the threshold, the counterfactual does not exactly fit the observed density for revenue bins above \( y_u \).

\(^{18}\)Since the data contains the universe of corporate tax returns, the source of uncertainty arises from the functional form of the polynomial. When running the bootstrap iterations we therefore draw from the sample
Despite being graphically compelling, the large behavioral responses to the revenue notches produce moderate revenue elasticities. Three points are worth mentioning. First, on a small profit base a modest change in revenue could generate a large profit elasticity, holding costs constant. Second, notches differ from kinks in that they generate sizable changes in implicit marginal tax rates and therefore large behavioral responses are consistent with moderate elasticities. Third, firms can also reduce their tax liability by increasing reported costs and therefore lowering revenue is only one of two possible margins of response to an increase in the tax rate. We investigate the latter point in the next section.

Cost discontinuity

Figure 1.5, Panel B, presented the step-like pattern of average profit margins by revenue. Profit margins by revenue are visually attractive since unit free and, in our data, very stable within tax brackets. However, to quantify the jump in costs at the threshold, caused by the tax rate increase, we turn to the relation between reported costs and revenue. Figure 1.7 plots average reported costs by revenue at the first threshold. Importantly, some firms have selected into the revenue range around the threshold, as a function of their costs. From the bunching analysis, we know that selection occurs precisely in the revenue bins corresponding to the excess and missing mass intervals, \([y_l, y_u]\). Therefore, we exclude these intervals from the cost discontinuity analysis, with dummy variables for the excess and missing mass areas.

We measure the discontinuity in cost at the threshold with the following specification:

\[
\text{cost}_j = \alpha + \delta \mathbb{1}(\tilde{y}_j > 0) + \beta_1 \tilde{y}_j + \beta_2 \tilde{y}_j \mathbb{1}(\tilde{y}_j > 0) + \sum_{j=\text{in}} \gamma_j \mathbb{1}(\tilde{y}_j = j) + \epsilon_j
\]

where \(\text{cost}_j\) is firms’ average cost in bin \(j\), \(\tilde{y}_j = y_j - y^*\) is the revenue distance to the threshold and \(\gamma_j\) are dummy shifters for firms with revenue in the excluded excess and missing mass intervals. \(\beta_1\) provides the slope of average cost on revenue below the threshold and \(\beta_1 + \beta_2\) the slope past the threshold. The parameter of interest is \(\delta\), the discontinuity in costs at the threshold. The specification directly provides the percentage change in cost at the threshold as \(dc = \frac{\delta}{\alpha}\).

Our objective is to measure the discontinuity in costs, holding revenue responses constant. However, the cost discontinuity estimated from equation (1.14) could entirely be due to intensive margin responses of revenue. To understand this, note that the “running” variable is revenue, which is also distorted by the change in the tax rate: absent the tax change, firms in the upper tax bracket would have declared larger revenue. Since we have estimated the revenue elasticity in Section (1.3), we can adjust for intensive margin revenue responses. To illustrate the revenue adjustment, we consider firms belonging to revenue bin \(j\), with revenue

of residuals of equation 1.9 and obtain new firm densities, with which we repeat the point of convergence method.
Chapter 1: Corporate Taxation Under Weak Enforcement

midpoint $y_j$. Absent the tax change their counterfactual revenue would be:

\[
\begin{align*}
\hat{y}_{j}^{adj} &= y_j & \text{ if } y_j \leq y^* \\
y_j^{adj} &= y_j + \epsilon_{y,1-t} y_j \cdot \frac{dt}{1-t} = y_j + 0.33 \cdot y_j \cdot \frac{0.1}{0.9} = y_j \cdot 1.037 & \text{ if } y_j > y^* 
\end{align*}
\]  

(1.15)

We clarify three aspects of the revenue adjustment. First, the adjustment only applies to firms with revenue above the threshold. Second, for firms with revenue sufficiently above the threshold the increase in the average tax rate is equivalent to an increase in the marginal tax rate. Since we exclude firms from the missing mass interval, this holds for the vast majority of firms. Third, the revenue adjustment assumes that the revenue elasticity is the average revenue elasticity. Since under heterogeneity in revenue responses the revenue elasticity is an upper bound of the average elasticity, the revenue adjustment is an upper bound of the true adjustment. We return to this point when interpreting the cost elasticity.

We apply the revenue adjustment, that is we replace $\tilde{y} = y_j - y^*$ with $\hat{y}_{j}^{adj} = y_{j}^{adj} - y^*$, and then estimate equation (1.14). $\delta$ now measures the increase in reported costs due to the tax change, but holding revenue responses constant. The discontinuity in costs, with and without the revenue adjustment\(^{19}\), are reported in Table 1.3. Figure 1.7 presents graphically the results for the first threshold: Panel A plots average cost by revenue and shows the revenue adjustment, which shifts costs horizontally for firms past the threshold. We then fit separate lines to the right and to the left of the threshold, excluding the interval affected by bunching responses between $z_l$ and $z_u$. The linear extrapolation to the threshold on the left provides a counterfactual average cost for firms at the threshold under a 10% tax rate and absent the notch. The linear extrapolation to the right, provides the average cost for a 20% tax rate, assuming no revenue responses. The resulting discontinuity is the change in reported costs that arises at the threshold due to the change in the tax rate. In Panel B, we zoom in on the discontinuity in the predicted average cost at the first threshold. We estimate a large jump in average cost of 2.5 million on a cost base of 42 million. Given the net of tax rate increase of 11%, the elasticity of cost is:

\[
\epsilon_{c,1-t} = \frac{dc}{c^*} \frac{(1-t_0)}{dt} = \frac{-2.55}{41.97} \cdot \frac{0.9}{0.1} = -0.55
\]

This implies that when the net of tax rate is reduced by 10%, firms respond by increasing their reported costs by 5.5%. At the second threshold costs jump by 1.2 Million on a 92 Million base. Together with the net of tax rate increase of 12.5% this implies a cost elasticity of -0.11.

The estimation is equivalent to a donut RD, with a local linear fit. Linearity is an important assumption to which we provide empirical support. First, figure 1.8 shows the linear and

\(^{19}\)The revenue adjustment method shifts horizontally average costs, such that under a sufficiently large revenue elasticity, the entire cost discontinuity could arise due to intensive margin revenue responses. For this to be the case, the elasticity of revenue would have to be 0.83 at the first threshold and 0.21 at the second, slightly under three times what we estimated.
quadratic fits of average costs by revenue, above and below each threshold. The quadratic fit is indistinguishable from the linear fit\(^{20}\). Second, table 1.4 displays the coefficients from Equation (1.14) using a quadratic fit instead of a linear fit: the cost discontinuity is even larger than under the linear model, and therefore if assuming linearity introduces bias, we would be underestimating the discontinuity in cost and the cost and profit elasticities. In addition, table 1.4 shows that the results are robust to variation in the revenue interval used to estimate the model and to the assumption that the revenue elasticity falls with revenue\(^{21}\).

### Profit elasticity

By combining the revenue and cost responses, we can now estimate the elasticity of profit with respect to the net of tax rate. The elasticity of profit is a central parameter to set optimal tax rates and a sufficient statistic for revenue collection under a flat tax rate. It is defined as:

\[
\epsilon_{\pi, 1-t} = \frac{\text{% change profit}}{\text{% change (net of tax rate)}} = \frac{\Delta \pi}{\pi} \ast \frac{1 - \tau}{\Delta \tau} = \frac{(\Delta y - \Delta c)}{\pi} \ast \frac{1 - \tau}{\Delta \tau}
\]

We already estimated the change in cost at the threshold \(\Delta c\) and compute the change in revenue \(\Delta y\) using the revenue elasticity: \(\Delta y = y \ast \epsilon_{y, 1-t} \ast \frac{\Delta y}{\pi}\).

Table 1.5 summarizes the elasticity estimates and changes in revenue, cost and profit, at each threshold. At the first threshold, we estimate a profit elasticity with respect to the net of tax rate of 4.9, and at the second threshold an elasticity of 2.9. These are very large elasticities and imply that the revenue maximizing rate is 17% for micro firms and 25% for small firms\(^{22}\). Tax rates above these are on the wrong side of the Laffer curve and Pareto dominated, since government revenue would fall, as the base diminishes faster than the rate increases. These large elasticities are a function of the current policy environment and of evasion and avoidance opportunities, which we investigate in further detail in Sections 1.5 and 1.6. However, we highlight in Figure 1.9 that the estimated elasticities correspond to an interesting reporting behavior. The figure shows average tax payment as a share of revenue by revenue: despite the 10% tax rate increase at each threshold, tax liability as a share of revenue is continuous and stable, at roughly 1.5% of revenue. It appears that faced with a tax hike, firms adjust their reported profits such that tax payments represent a near constant share of their revenue.

---

\(^{20}\)When comparing the adjusted R-squared from the linear, quadratic and cubic regressions we find that the linear model has the highest adjusted R-squared.

\(^{21}\)The revenue adjustment uses the estimated elasticity at the threshold and applies it homogeneously to all firms with revenue above the threshold. Since the revenue elasticity is larger at the first than second threshold, an alternative is to assume a linearly decreasing elasticity as a function of the firm’s revenue, with a slope proportional to the drop in revenue elasticities between the first and second threshold.

\(^{22}\)Under a flat corporate tax, the government revenue maximizing rate is \(\tau^* = \frac{1}{1 + \epsilon_{\pi, 1-t}}\).
Another novel result in Table 1.5 is the comparison between the cost and revenue elasticities. Slightly over 60% of the discontinuity in profit is due to an increase in costs and 40% from an increase in revenue\(^{23}\). The difference is statistically significant at the first threshold, and holds qualitatively at the second: reported costs react stronger to a change in the tax rate than reported revenue. This holds despite estimating a lower bound on the cost elasticity and an upper bound on the revenue elasticity. In Section 1.4 we add a counterfactual assumption on the distribution of costs by revenue to obtain average revenue and cost elasticities.

**Robustness and heterogeneity**

We discuss three important dimensions of robustness and heterogeneity: elasticity estimates, distributional results and industry variation.

The large drop in average profit margin on either side of the threshold is the key identifying variation for the total profit response, while the two step estimation decomposes this variation into revenue and cost responses. In doing so, a near perfect negative correlation mechanically arises, due to the revenue adjustment term applied to the cost discontinuity: a larger revenue elasticity implies a larger revenue adjustment, which reduces the cost elasticity. Accordingly, the profit elasticity is robust to results from the bunching estimation and hinges upon the assumption that average reported costs by revenue would have been smooth around the threshold, absent the tax change. Following the same logic, the cost to revenue elasticity ratio estimate is less robust and a lower bound of the true ratio: our preferred estimate of this ratio is this of Section (1.4).

Another robustness is whether the decrease in profit margins is driven by a few profitable firms or by an entire shift in the distribution. Figure 1.10 shows the quartiles of profit margin: the median profit margin starts at 6% below the threshold and drops to 3% above it. We observe similar proportional falls at the 25th and 75th percentiles. It appears that profit margin discontinuities arise from an entire downward shift of the distribution of profit margin and not only from a change in profit of a few high profitability firms.

Finally, some of the results could be driven by industry variation. Our estimation methodology relies on large sample sizes such that the firm distribution\(^{24}\) and average cost are smooth. As a consequence we can not estimate revenue and cost elasticities at the industry level with precision. Instead, to summarize the revenue response at the industry level, we turn to the excess mass of bunchers\(^{25}\) at the threshold, \(\hat{E}_s\) for industry \(s\). The industry excess ratio \(\hat{E}_s\) is robust to the polynomial fit and provides a proxy for the size of revenue

\(^{23}\)In addition, costs are measured on a smaller base than revenue, therefore the cost elasticity is larger in absolute terms than the revenue elasticity at both thresholds \((\epsilon_{c,1-t} > \epsilon_{y,1-t})\).

\(^{24}\)The main limitation to apply the two-step estimation method for each industry arises from the point of convergence method, which iteratively fits the polynomial to equate the excess mass with the missing mass. With small sample sizes the iterative process can fail to converge.

\(^{25}\)We fit a new polynomial for each industry but do not re-estimate, \(y_u\) the upper bound of the excluded interval.
responses. We also estimate the profit margin discontinuity for each industry using equation (1.14). Table 1.7 shows the industry level responses\textsuperscript{26} and Figure 1.12 plots for the first threshold the excess mass on the percentage change in profit margins by industry. On the one hand, we observe a large variation in excess mass ranging from just above one for retailers and wholesalers to six for consultancies. On the other hand, the proportional change in profit margin is very homogeneous: most sectors declare margins 40 to 50\% lower above the threshold. Despite large variation in revenue responses, total profit responses are fairly homogeneous. All industries display excess mass at the threshold, however sectors with high evasion potential such as construction, real estate and legal and economic consultancies exhibit stronger bunching than retailers and manufacturers. Most importantly, average profit margin drops discontinuously for all industries and the proportional downward fall in margins appears homogeneous. The only sector which does not display any discontinuity is NGOs and public administration\textsuperscript{27}.

1.4 Model-Based Estimation of Revenue and Cost Elasticities

The elasticity of profits with respect to the tax rate is robust to the separation into revenue and cost responses and if we make no assumption about the counterfactual distribution of cost by revenue, the estimation of Section 1.3 provides the tightest possible bounds on the revenue and cost elasticities. This estimation faces however several limitations: first, under heterogeneity in revenue elasticity, it provides an upper bound to the true revenue elasticity and a lower bound on the cost elasticity. Second, it does not take into account selection into bunching as a function of costs. Third, it does not consider the feedback effect of cost responses on revenue responses.

To address these limitations, we assume a counterfactual distribution of profit margins: absent the tax change, the distribution of profit margins by revenue remains constant within the revenue intervals around the threshold. We then combine the counterfactual firm density by revenue with the new counterfactual profit margin distribution, to determine jointly the revenue and cost elasticities at which the number of counterfactual firms which should be bunching corresponds to the observed bunching mass. The additional structure allows us to model responses to the notch as a joint function of the firms’ revenue distance to the threshold and costs, as suggested by the model. It also considers the impact of cost responses on the bunching decision: we assume that firms’ cost responses, if not bunching, would equal the average cost response, estimated with the discontinuity approach. With an additional counterfactual assumption, the numerical model-based method allows us to estimate aver-

\textsuperscript{26}Graphically shown in Figure 1.11
\textsuperscript{27}Public administration includes all local levels of government, which are required to withhold taxes on all transactions
age revenue and cost elasticities and hence obtain a more precise decomposition of profit responses into cost and revenue responses.

**Main Estimation**

We impose the restriction that absent the tax change, the entire distribution of profit margin by revenue would stay constant in an interval around the threshold. Under this restriction, we use the profit margin distribution of firms with revenue below the threshold, and away from the bunching zone, as the counterfactual for firms with revenue just below or just above the threshold. Specifically, this assumes that the profit margin distribution of firms with 45 million CRC in revenue, would apply to firms with revenue 10 to 20% larger, absent the tax change. We test if this assumption holds away from the first threshold. Figure 1.13, Panel A, plots several distributions of profit margin for revenue intervals below the threshold. The distributions of profit margin appear extremely stable across different revenue intervals 10 to 20% lower than the threshold, and we can never reject the Kolmgorov-Smirnov tests that profit margins are generated from identical distributions. Figure 1.13, Panel B, also shows that within revenue intervals 20 to 30% higher than the threshold, the new distributions of profit margins are stable. Firms in these revenue intervals are not impacted by selection into bunching: therefore the stability of profit margin distributions across these intervals gives supports the assumption of locally constant counterfactual profit margin distributions. Thereafter, we use the distributions in Panel A as the counterfactual profit margin distribution for firms in the bunching impacted zone, under a flat 10% corporate tax rate.

By combining the constant counterfactual profit margin distribution within each revenue bin, with the distribution of firms by revenue from the polynomial fit, we obtain a joint counterfactual distribution of revenue and costs. This allows us to model selection into bunching as a common function of the firms’ costs and revenue distance to the threshold. For each revenue bin past the threshold, and given an elasticity of revenue, we can compute the cost threshold at which the firm is indifferent between bunching and remaining above the threshold. To compute the cost threshold, we first return to the expression of the implicit marginal tax rate, \( t_i \), and model cost responses, \( dc \):

\[
\tau_i^* = \frac{T(y_i^* + dy_i) - T(y_i^*)}{dy_i} = \frac{(\tau_0 + d\tau)(y_i^* + dy_i - c_i - dc_i) - \tau_0(y_i^* - c_i)}{dy_i}
\]

\[
\tau_i^* = (\tau_0 + d\tau) + \frac{d\tau(y_i^* - c_i) - (\tau_0 + d\tau).dc_i}{dy_i}
\]

The above equation states that the implicit marginal tax rate \( t_i^* \), faced by firm \( i \), is a function of the firm’s revenue distance to the threshold \( dy_i \), costs \( c_i \), and the change in reported costs conditional on facing the higher tax rate \( dc_i \). To understand this later term, consider a firm that can easily evade costs: facing the higher average tax rate hardly increases its implicit marginal tax rate, since it can freely adjust its tax liability by reporting higher costs (large \( dc \)). On the contrary, a firm facing large resource costs of evasion on costs, can
not adjust it tax liability by over-reporting costs, and therefore faces a large increase in its implicit marginal tax rate (low $dc$). With the explicit marginal tax rate we can now express the revenue elasticity as:

$$\epsilon_{y,1-\tau} = \frac{dy}{y} \frac{1 - \tau_0}{d\tau} = \frac{dy}{y} \frac{1 - \tau_0}{\tau^* - \tau_0}$$

$$\epsilon_{y,1-\tau} = \frac{(dy)^2}{y} \frac{(1 - \tau_0)}{d\tau(dy) + d\tau(y^* - c) - (\tau_0 + d\tau).dc} \quad (1.18)$$

With knowledge of $dc$, then for a given elasticity of revenue, $\epsilon_{y,1-\tau}$, and distance to the threshold, $dy$, we can measure the cost threshold, $\tilde{c}$, such that all firms with cost lower than $\tilde{c}$ bunch. $dc$ is the change in reported costs of bunchers, had they remained above the threshold and faced the higher tax rate. In practice $dc$ is unknown: our preferred estimation assumes that the cost response of bunchers would have equaled the average cost response ($dc = d\bar{c}$). This is equivalent to say that bunchers, would have had the same cost response as infra-marginal firms, had they not selected into bunching²⁸. The average cost response, $d\bar{c}$, is estimated from the cost discontinuity, adjusted for revenue responses, following the methodology of section 1.3.

We rewrite equation 1.18 such that the cost threshold is a function of all parameters:

$$\tilde{c}_j = y^* + dy_j - \frac{(dy_j)^2.(1 - \tau_0)}{d\tau.\epsilon_{y,1-\tau}.(y^* + dy_j) - (\tau_0 + d\tau).d\bar{c}} \quad (1.19)$$

Equation 1.19 states that firm $i$ in revenue bin $y_j$, with distance $dy_j$ to the threshold, will bunch under revenue elasticity $\epsilon_{y,1-\tau}$, if its costs are below the cost threshold, $c_{ij} < \tilde{c}_j$. With the counterfactual revenue distribution, we know the number of firms that would have declared revenue in bin $y_j$, absent the tax change. With the counterfactual profit margin distribution, we know the distribution of costs within each revenue bin. Therefore we can numerically estimate the number of bunching firms for a given elasticity of revenue.

The estimation is an iterative procedure: the initial values of the revenue elasticity, $\epsilon_{y,1-\tau}^{Step1}$, and of the cost response, $d\bar{c}^{Step1}$, are taken from our estimation in Section 1.3. This combination of revenue and cost responses predict substantially more bunching than observed. We estimate the new revenue elasticity $\epsilon_{y,1-\tau}^{Step2}$, such that the number of numerically estimated bunchers equates the excess mass at the threshold. With the resulting revenue elasticity, $\epsilon_{y,1-\tau}^{Step2}$, we measure the average cost response from the cost discontinuity equation (Equation 1.14), adjusted for the newly estimated revenue elasticity, and obtain a new average cost response, $d\bar{c}^{Step2}$. $d\bar{c}^{Step2}$ is then used as the average cost response to measure a new revenue

²⁸In our model, under heterogeneous cost elasticities, this assumption provides an upper bound on the elasticity of revenue: bunchers select into bunching because their revenue elasticity is large and because their cost elasticity is low. Therefore, assuming the average cost response over-estimates bunchers potential response, and under-estimates bunching for a given elasticity. However, with a resource cost of evasion function non-separable into revenue and cost responses the estimation is not necessarily an upper bound: if revenue and cost evasion are easily substitutable, then firms with large cost evasion can substitute for revenue evasion to reach the threshold.
elasticity $\epsilon_{y,1-t}^{Step3}$. We iterate this process\footnote{At each iteration we also re-estimate the counterfactual density distribution of firms, since we applied a correction for intensive margin responses. This only has second order effects} until we converge to the fixed point $(\hat{\epsilon}_{y,1-t}, \hat{\epsilon}_{c,1-t})$, where the revenue and cost elasticities are consistent with each other.

We report the iteration steps in table 1.6 and represents graphically the last iteration in Figure 1.14. In Panel A, The number of bunchers are represented by the area between the elasticity curve and the counterfactual density. We show the elasticity curves for three values of the revenue elasticities. On the one hand, even under a very small revenue elasticity, some firms with revenue just above the threshold bunch, since their implicit marginal tax rate is above one. On the other hand, even with a very large revenue elasticity some firms do not bunch since they have very large costs. As the revenue elasticity increase firms further away from the threshold bunch. Panel B shows the number of bunching firms by revenue bins, for the equilibrium revenue and cost elasticities.

At the first threshold, the resulting revenue elasticity is 0.22, while the cost elasticity is -0.65 and the profit response is practically unchanged at 4.9. This implies that 71% of the total profit responses are due to an increase in reported costs and only 29% to a decrease in reported revenue. The share of cost to revenue responses is substantially larger than the one estimated from the point of convergence bunching method. This result was expected since under heterogeneity in revenue elasticity, the point of convergence method estimated the revenue elasticity of the highest revenue elasticity firm. Instead the model-based estimation measures the average revenue elasticity, in a frictionless environment, which mechanically gives us the average cost elasticity via the discontinuity method. This comes at the price of two additional assumptions: first the profit margin counterfactual distribution by revenue is stable around the threshold and second the counterfactual cost responses of bunchers, had they not bunched, would equal the average cost response.

1.5 Mechanisms: Evasion Responses

Behavioral responses generated by evasion, real production effects or tax avoidance responses have similar impacts on the government’s tax revenue collection, but entail different policy responses. In the literature, it has often been difficult to separate the different mechanisms: evasion responses are by nature secretive and measuring real responses requires precise production data linked to tax records. Using additional dimensions of the data and new data sources we perform a series of empirical test of the mechanisms. Even though these tests are not exhaustive, we believe that taken together they provide a convincing picture that evasion and misreporting is a key driver while real effects and avoidance responses appear limited.

In this section we study two direct channels to uncover evasion behavior. First we look at the percentage of firms being flagged in the internal data cross-checks and show that firms reporting revenue at the threshold are significantly more likely to be selected for discrepancies. Second we look at variation in audit intensity by sectors across years. We find evidence...
that bunching firms under higher scrutiny increase their declared revenue and move past the threshold but also increase cost by a similar amount. These empirical tests both support the idea that reporting responses are strong drivers of revenue bunching.

**Audits and tax corrections in Costa Rica**

Costa Rica undertakes five hundred in depth taxpayer audits every year, of which three hundred are targeted to firms. As a consequence, only 0.4% of all firms are audited in a given year. Taxpayers are selected following a risk based analysis which incorporates information from third-parties, deviation from industry averages and the taxpayers’ history. The small and medium firms we study are rarely audited: just above 50 audits a year corresponding to a 0.2% probability of being audited. The percentage of audited firms increases with firm revenue to reach over 3% for large firms.

The low capacity to conduct in depth audits is partly mitigated by the extensive automatic warning system: auditors send notification letters to firms raising “red flags” in the internal data intelligence process. Specifically, anytime a computer operated system observes discrepancies between self-declared revenue and revenue based on third-parties reports, it generates a correction letter. A large part of third-party information is collected through the D151 informative tax form, which requires all individuals and firms to declare purchases and sales to the same entity when the value within the tax year is above two million Colones ($6,000 in PPP) and any commission, professional fee or rental agreement above fifty thousand Colones ($150). Other third-party information such as sales tax retentions, credit card payments and insurance policies also enter the database.

The letters sent by tax auditors ask for a correction or justification of the tax declaration in order to match the amount assessed by the tax administration. Importantly, this process does not treat bunching firms differently. For legal reasons this information can not be linked to the individual tax records, however we obtained the number of correction letters sent by revenue bins for the year 2012. Figure 1.17 shows the proportion of firms receiving a letter by revenue bins of two Million CRC. The green line shows the linear fit excluding the revenue intervals around the thresholds. Around a third of small Costa Rican firms receive correction letters, which highlights that tax declarations are often incomplete and that the environment is prone to evasion. Bunchers face large and significant increase in the probability of receiving correction letters compared to their expected probability: they are 8.3% more likely to receive correction letters at the first threshold and 11.5% at the second. This result highlights that bunching and revenue changes are most likely the result of tax evasion and that bunchers get noticed for inconsistencies at a higher rate but are willing to incur the expected costs\(^{30}\).

\(^{30}\)Measuring precisely the expected costs from tax evasion is a difficult task and from our discussions it is unclear how many firms actually make significant adjustments to their tax payment following a correction letter. Firms can certainly revise their tax payment at minimal cost and do not get systematically prosecuted. However failure to comply increases the risk of an in depth audit which is considered to be very costly for the firm.
Two other results are worth noticing. First, firms declaring revenue just above the threshold (potentially dominated firms) are less likely to get flagged and the joint F-test shows that the difference is significant at both thresholds. This might indicate that firms that do not adjust their revenue to the threshold do so partly because of honesty. Second the proportion of correction letters by revenue is fairly constant on either side of the threshold contrarily to profit margins. However we saw that the increase in declared costs explains a large share of the discontinuity. A possible explanation is that third-party information on costs is not sufficient to establish evasion as it only provides a lower bound on the true cost. On the contrary, third-party information on revenue provides an upper bound on true revenue, hence a clear signal of tax evasion. Carrillo, Pomeranz, and Singhal 2014 also observe this asymmetry in the case of Ecuador.

**Sectors of special audit attention**

A second test of tax evasion uses the variation in audit probability generated at the industry level by the program of “Special audit attention”. In 2012 the tax agency determined during the first semester of the calendar year a list of industries assigned to special audit attention, which was posted on the ministry of finance website. In practical terms it implied that the selected industries are assigned a dedicated group of auditors and that their risk of an audit increased. Industries are not randomly selected but determined by the underlying evasion risk and the industry’s growth rate compared to its tax revenue growth. The twelve sectors selected in 2012 were real estate, private education, hotels and tour agencies, transport of merchandise, sale of vehicles, sports, production of pineapple, yucca, flowers and plants, casinos and betting, performances and recycling. The difference in difference analysis of firms within the audit sectors versus other sectors shows significant growth in reported profits following the assignment of the sector to “special audit attention”. However it is difficult to establish causality due to the endogenous selection mechanism.

Instead we use a triple difference strategy to study firms’ evasion behavior at the threshold: we compare the change in revenue, costs and profits reported by bunchers in the sectors of special audit attention with bunchers in other sectors and non-bunchers in the same sectors. Our hypothesis is that bunchers are evading more revenue compared to smaller firms and dominated firms. Therefore, when faced with a higher audit probability, bunching firms should lower revenue evasion and increase reported revenue by more than non-bunchers. We estimate the following equation:

$$y_{ist} = \alpha_i + \beta \cdot Bunch_i \cdot Audit_j \cdot Post_t + \gamma \cdot Bunch_i \cdot Post_t + \delta \cdot Audit_s \cdot Post_t + Post_t + \epsilon_{ijt}$$ (1.20)

Where depending on the specification $y_{ijt}$ is revenue, costs or profits of firm $i$ in sector $j$ at time $t$. Bunch is a firm level dummy equal to one if the firm declared revenue in the two Million revenue interval below the threshold in 2011 and zero otherwise. Audit is a sector

---

31The Sectors selected in 2013 were quasi-identical and therefore we do not perform a sector event study for each year of the variation.
level dummy that equals one if the firm belongs to the sectors of special audit attention and zero otherwise. *Post* is a time dummy equal to one in 2012 and 2013 and zero in 2011.

Table 1.10 presents two sets of results: one using firms with revenue in 2011 below the bunching interval (declared revenue 4 to 8 Million CRC below the threshold) as a control group in columns (1)-(3) and one using dominated firms (revenue 0 to 3 M Colones above the threshold) as a control group in columns (4)-(6). Columns (1) and (4) presents the main result: bunchers in special audit sectors increase revenue by 10% or more of their initial size compared to non-bunchers. However they simultaneously increase their reported costs by a large amount in columns (2) and (5), leading to a statistical significant decrease in declared profits of above 1 M Colones column (3) and (6). Placebo treatment effects for previous years are not significant on any of the three margin.

These results support the idea that bunching firms are evading revenue, since bunchers belonging to audit sectors increase their declared revenue substantially more than firms in the control groups following an increase in their audit probability. Interestingly, reported cost also increase by a large amount, which indicates that away from the bunching segment firms lower their tax liability by increasing cost. Firms with revenue just above the threshold have strong incentives for revenue evasion in order to decrease both their tax base and tax rate, while for firms with revenue further past the threshold it is equivalent to under-report revenue or over-report costs which only decreases the tax base.

### 1.6 Mechanisms: Dynamic, Real and Avoidance Responses

In this section we investigate four other dimensions of firm behavior: dynamic responses, firm division, labor input usage and time-shifting of monthly revenue.

First, we show that over time firms slightly increase their profit margin, conditionally on staying within the same tax bracket. However when changing tax bracket, firms display large jumps in profit margin mirroring the cross-sectional results. Second, we investigate whether larger firms divide themselves and create subsidiaries with revenue below the threshold on which to offload their profits. We show that this is unlikely to be an important mechanism. Few firms repeatedly bunch which would be a prediction of this model. More importantly a dataset of economic groups collected by the central bank shows very limited excess profit of subsidiaries when compared to non-subsidiaries and no excess mass of subsidiaries at the threshold. Third, we explore responses to employment and wage bill using data from social security, which is considered to be well-reported information. Neither employment nor wage bill show any discontinuity at either threshold. Finally, looking at monthly revenue from sales tax receipts we do not observe reduced economic activity in the last month of the fiscal yea nor any time-shifting to the first month of the next fiscal year.
Dynamic Responses

Do the discontinuous profit margins in the cross-section also apply within the same firms across years? To answer this question we use the panel dimension of the data and look at the difference in reported profit margin as a function of the firm’s tax bracket in a given year. Figure 1.15 shows the average profit margin difference between years $t+1$ and $t$, conditional on firms tax bracket in those years. Firms remaining within the same bracket in consecutive years do not change on average their profit margins. Whereas, growing firms jumping to a higher tax bracket declare lower profit margins and symmetrically, shrinking firms falling to a lower bracket declare higher profit margins.

Nonetheless, profit margins could decrease with revenue for structural reasons, in particular if firms are facing decreasing returns to scale.$^{32}$ To investigate this claim, we focus on a subsample of firms with similar revenue growth: Figure (1.16) shows that firms that grew but stayed within their own tax brackets slightly increased their reported profit margin, while firms that jumped past the threshold significantly decreased their reported profit margin. A systematic way to test the within-firm drop in profit margin when crossing the threshold is to regress firm revenue on profit margins while controlling for the tax bracket:

$$
margin_{it} = \alpha_i + \gamma_t + \beta y_{it} + \delta \mathbb{1}(\tau_{it} = \tau + d\tau) + \sum_{j=\hat{y}_i} \psi_j \mathbb{1}(\hat{y}_j = j) + \epsilon_{it}
$$

Where $\alpha_i$ and $\gamma_t$ are respectively firm and year fixed effects, $y_{it}$ is the revenue of firm $i$ at time $t$, $\tau_{it}$ is the average tax rate faced by firm $i$ at time $t$ and the dummy variables are shifters for the revenue intervals impacted by bunching. Table (1.8) presents the results from the above regression. The coefficients on revenue $\beta$ shows that conditional on staying within the same tax bracket, growing firms increase their declared profit margins rejecting the possibility that profit margin should be falling for a growing firms. The dummy coefficients for changing tax bracket measures the discontinuity that occurs at the threshold and are significant and negative at each threshold. A firm growing past the first threshold decreases its profit margin by 3.06% and a firms growing past the second threshold decreases its profit margin by 0.86%. For these results to be consistent with real responses under distortionary taxation firms would need to have increasing returns to scale such that lowering production would also lower profit margin.

Firm Division

Given the design of the corporate tax system, it seems attractive for a large “mother” firm to create a subsidiary firm on which to “offload” its profits: the small subsidiary then declares high profits, taxed at a 10% rate, while the larger firm declares low profits, taxed at a 30% rate. If there are large profits to shift, then the subsidiary should locate its revenue just

$^{32}$In the model this corresponds to a positive correlation between changes in productivity $\phi$ and changes in fixed costs $\alpha$. 
below the threshold. Therefore, sufficient firm division and profit shifting could explain the revenue bunching and the discontinuities in profit margins. To test for firm division we use the registry of economic groups, a unique dataset, compiled by the Central Bank exclusively for statistical purposes. It links corporate groups and their subsidiaries by combining the registry of corporate ownership, the census and direct visits and calls to firms offices.\footnote{This data project, named REVEC, was motivated by the updating of the input-output matrix for Costa Rica in 2012 and to obtain an accurate view of corporate ownership structure.}

Firms operating under common ownership are defined as forming part of an economic group. On the one hand, shared ownership structure of firms can exist for structural reasons and the existence of subsidiaries does not provide in itself evidence of tax avoidance. On the other hand, if avoidance motivations are important, the following hypothesis should hold: First, firms in the 10% tax bracket should be more likely to be subsidiaries compared to firms in the 20% tax bracket since they represent better tax instruments, while subsidiaries in the 30% tax bracket do not serve any tax goal. Second, there should be an excess number of subsidiaries with revenue in the bunching interval and few subsidiaries with revenue just above the threshold: if subsidiaries are tax-related vehicles, then changing their declared revenue should produce minimal resource costs, while generating large tax gains. Third, the profitability of subsidiaries should be large since these are profit-shifting vehicles while the profitability of mother firms should be low.

Figure 1.18 shows the share of subsidiaries by revenue bin and fits a linear relation on both sides of the threshold, excluding revenue bins in the bunching and dominated intervals. On average 4 to 5% of firms are subsidiaries of larger firms, however the relation appears rather continuous on either sides of the threshold: The estimated drops in the number of subsidiaries at the first threshold of 0.39% and 0.42% at the second threshold are not significant. Qualitatively this still represents a 10% decrease in the probability of being a subsidiary past the threshold, which indicates that the strategy might exist but is marginal in explaining our results. Bunchers do not exhibit a significant particularly large response. Finally, when comparing subsidiaries to non-subsidiary firms we only find very modest evidence of excess profitability.

To summarize, we only find modest evidence of firms dividing themselves and gaming the tax system. The results can not explain the excess bunching nor the large difference in profitability on each side of the thresholds. The absence of firm division could be a combination of several factors. The monetary and non-monetary costs of setting up a corporation and keeping it active are not trivial: In addition to cumbersome administrative work, Costa Rica has a registration fee and a yearly stamp duty payment. related to the above, the relative ease of evading taxes by inflating costs might make this avoidance strategy suboptimal. Finally, the data might not reveal the full extent of firm division. Economic groups might have hard to detect dilution ownership strategies, and the Central Bank dataset is a (large) subset of the universe of tax filling corporations (around 80% of firms filling a tax declaration).
Employment and wage bill

The breakdown of costs into the five categories reported on the tax return brings limited insights. In figure 1.20 we show the cost discontinuity by revenue, for three cost categories. The two main categories, “Administrative and Operational Costs” and “Material and Production Costs” explain respectively 60% and 40% of the cost discontinuity. The other three categories, interest deductions, depreciation and other costs do not display a discontinuity and together represent less than 15% of total costs.

In the above cost categorization, wages are reported as part of administrative and operational costs and can not be separately identified. In order to study employment and wage bill we turn to data from social security records, a separate database obtained from the central bank. There are reasons to believe that labor inputs in social security records are better reported than deductible costs from the corporate tax returns. First, employees have incentives for accurate reporting of their wages as social security benefits depend upon it and are generous in Costa Rica. Second estimated evasion on payroll tax and personal income tax of wage earners is much lower than evasion on other margins: the International Labor Organization estimates that among formal firms, only 9% of employees are informal. Finally, the sum of personal income taxes and payroll taxes is larger than the corporate income tax within all tax brackets and, if anything, a firm has incentives to under-report labor and not over-report it.

Figure 1.19 plots the average number of employees and average wage bill by revenue around each threshold and shows the linear fit of employment on revenue on each side of the threshold, excluding revenue bins in the bunching and dominated intervals. We run the following regression:

\[
E_j = \alpha + \delta \mathbb{1}(y_j > 0) + \sum_{j=yl}^{yu} \gamma_j \mathbb{1}(y_j = j) \beta_1 y_j + \beta_2 y_j \mathbb{1}(y_j > 0) + \epsilon_j
\]

(1.21)

Where \(E_j\) is average employment or average wage bill in revenue bin \(j\), \(y_j\) is the midpoint of revenue bin \(j\) and \([y_l, y_u]\) is the excluded revenue interval around the threshold. The estimated discontinuity and the bunchers dummy are reported on Figure 1.19: neither average employment nor the wage bill discontinuity are significant. The only possible indication of differential labor inputs usage is the significantly lower wage bill for bunchers at the second threshold\(^{34}\). Overall the absence of a dip in labor inputs for bunchers and of a downward discontinuity past the threshold suggests that firms are not limiting their labor input use and distorting their production. It also rules out a strategy where firm owner-managers increase their own wages when faced with a higher tax rate.

\(^{34}\)We are not controlling for average profitability which could minimize the significant lower tax bill estimated at the second threshold.
Production and timing responses

Another potential strategy for bunching is to either limit production or shift revenue across fiscal years. For example, firms could limit their operations in September, the last month of the fiscal year and/or date September revenue in October, such that their revenue in the fiscal year remains below the threshold. We use the subsample of firms liable for sales taxes to obtain a monthly revenue for the years 2008 to 2013. We run the following specification:

\[ y_{imt} = \beta_1 \mathbb{1}(m = \text{Sept}) * Bunch_{it} + \beta_2 \mathbb{1}(m = \text{Oct}) * Bunch_{it} + \delta * Bunch_{it} + \alpha_m + \gamma_t + \epsilon_{imt} \]  \hspace{1cm} (1.22)

where \( \beta_1 \) measures the differential monthly revenue of bunchers in September and \( \beta_2 \) in October. Under limitation of production at the end of the fiscal year, we should observe \( \beta_1 \) to be negative and significant and under time-shifting we should observe \( \beta_2 \) to be positive and significant. Table 1.9 reports the results for the first threshold for all firms in columns (1)-(4) and for firms who’s corporate income tax revenue equals the sum of the monthly sales tax in column (5)-(8) and who therefore have no revenue not subjected to sales tax. The regressions reject evidence of either responses: bunchers do not report revenue differently in September and October compared to other months and other non-bunching firms. The only significant coefficient shows a positive revenue increase of bunchers revenue in September - possibly a sign of reverse time-shifting: as firm realize their revenue fall in the lower tax bracket they decide to bring revenue forward due to uncertainty about next year’s revenue. A caveat to keep in mind when interpreting this non-result is that sales tax liable firms belong to sectors with lower bunching intensity such as retails, restaurants and car dealers. Nonetheless those results support the limiting role of real production effects and time-shifting responses in bunching behavior.

1.7 Conclusion

In this paper we highlighted the role of firms’ behavioral responses to higher tax rates as a channel contributing to low tax revenue collection in developing countries. We focused on the corporate income tax and showed that even in a middle-income country, the capacity to tax small and medium enterprises is limited due to large evasion responses to higher tax rates. We also quantified a new mechanism: even though business revenue is somewhat difficult to manipulate, observing business costs is so hard that the standard profit tax, which allows for all costs to be deducted, collects little tax revenue. To obtain this result we used the unique design of Costa Rica’s corporate income tax, administrative data on the universe of formal firms and a novel estimation strategy, which combined bunching at revenue thresholds with a discontinuity in profits on either sides of the thresholds.

Two key dimensions should be considered for the external validity of the results: the

---

35The sales tax in Costa-Rica works like an incomplete VAT: it exempts some industries (e.g. liberal professions) and only allows deductions of physically observable inputs, such as materials, and is therefore different from a full value-added tax.
size of firms in our study and Costa Rica’s institutional environment. First, our elasticity estimates concern small and medium firms and might not apply to large firms. Based on only two observations, one at each threshold, the elasticity of profits with respect to the net of tax rate appears to decrease with firm revenue. However large firms could have access to more elaborate evasion and avoidance schemes and it is therefore hard to conclude that the elasticity of profit is necessarily falling with firm revenue. Second, tax elasticities are always dependent on the institutional and policy environment. On the one hand, Costa Rica’s institutions are solid given its per-capita income: for example, the country ranks above its income level on Transparency International’s corruption perception index. In weaker institutional environments, profit elasticities could be even larger. On the other hand, Costa Rica’s tax structure is complex and fragmented\textsuperscript{36}, which could contribute to the large estimated profit elasticity and the ease of over-reporting costs.

With these caveats about interpretation in mind, how should low and middle-income countries design their corporate income tax? We discuss four types of corporate tax design, and their potential implications for tax collection and efficiency. Across countries, the most common corporate schedule taxes profits at a flat rate and permits the deduction of most production costs. As a direct consequence of the large estimated elasticities, flat rates above 17-25% are on the wrong side of the Laffer curve and should be excluded\textsuperscript{37}. Some countries have increasing marginal tax rates on profits, which reduce bunching incentives but generate a loss in tax revenue, as infra-marginal firms with large profits would see their tax bill decrease on the initial portion of their earnings. These concerns are particularly important if the profit elasticity falls with firm revenue, such that large firms with a lower profit elasticity, obtain a reduction of their tax liability without substantially increasing reported profits. In light of this plausible assumption, Costa Rica’s tax system permits to tag\textsuperscript{38} firms based on revenue, which we show is relatively hard to manipulate, and to assign increasing tax rates, potentially satisfying an inverse elasticity rule. In addition, the low initial tax rate for small firms might contribute to firm registration and formalization: once firms are registered it could be difficult to return to informality and at this stage increasing enforcement could yield revenue gains. However, the current system does not satisfy horizontal equity (i.e. quasi-identical firms face vastly different tax treatment) and only imperfectly deals with the large cost over-reporting. To this end the introduction of presumptive tax schemes could be beneficial. Under such schemes, tax liability is the maximum amount of a low rate applied on revenue and a higher rate on profits. Best et al. 2014 show that in Pakistan, the revenue gains from a presumptive scheme could outweigh the production distortion it generates. The

\textsuperscript{36}In addition to the corporate and personal income taxes, Costa Rica has a self-employed regime and a micro-sellers regime, which applies to firms with revenue substantially below the firms we study. In addition it does not have a fully fledged VAT system (Pomeranz 2015a), even though its sales tax mimics certain aspects of a VAT.

\textsuperscript{37}Gorodnichenko, Martinez-Vazquez, and Sabirianova Peter 2009 and Kopczuk 2012 both find that flat tax reforms in Eastern Europe, which decreased substantially the rate and simplified the tax code, led to large increases in reported income.

\textsuperscript{38}See Ito and Sallee 2014 for a model of attribute based regulation and enforcement.
large ratio of cost to revenue elasticity we estimate, further supports their empirical results on the desirability of using revenue as the tax base. More generally, a tax design which limits deductions and/or requires tangible evidence of incurred costs, such as electronic receipts, could generate a substantial increase in tax revenue.

To conclude, we show in Figure 1.21 the types of corporate tax systems used worldwide, grouping countries in per-capita income quintiles. Tax system with flat rates and increasing marginal rates are the most common and ubiquitous in the top-quintile of the income distribution. However, within low and middle-income countries, we observe large variation in corporate income tax policies: presumptive schemes are common, especially in the bottom quintile, while systems with revenue dependent rates, like Costa Rica’s, are sometimes used in middle-income countries. The observed heterogeneity in corporate tax systems highlights that developing countries do use non-standard tax instruments, potentially as a response to the behavioral responses constraining revenue collection which we documented in this study.
Figure 1.1: Corporate Tax Revenue and Income

![Corporate Tax Revenue and Income](image)

Data: tax revenue from the ICTD. Countries with pop>1M & data after 2005

Figure 1.1 shows the relation between corporate tax revenue and log of per capita GDP at the country level. Tax revenue data from ICTD and GDP data from the world bank indicators.

Figure 1.2: Corporate Tax Rate and Income

![Corporate Tax Rate and Income](image)

Data: Corporate tax rates from Abramovsky, Bachas & Jensen(2015). Countries with pop>1M & data after 2005

Figure 1.2 shows the relation between statutory corporate tax rates and log of per capita GDP at the country level. Statutory tax rates data collected by the authors for the year 2014 and GDP data from the world bank indicators.
Figure 1.3 shows the design of the corporate income tax in Costa Rica. Firms pay increasing average tax rates on their profits as a function of their revenue. When revenue exceeds the first threshold, the average tax rate jumps from 10% to 20% and from 20% to 30% past the second threshold.

Figure 1.4 displays the theoretical density distribution. Under a flat 10% tax rate the counterfactual firm density follows a smooth distribution. The notch induces some firms, with counterfactual revenue above the threshold, to reduce their revenue and bunch just below the threshold. The bunching decision is a joint function of the firm’s revenue distance to the threshold and costs, such that at each revenue bin past the threshold, only firms with sufficiently low costs bunch.
Figure 1.5: Firm Density and Average Profit Margin

Panel A: Firm Density

Panel B: Profit Margin

Figure 1.5 presents the key patterns of the corporate tax returns, pulling together the years 2008 to 2014. Panel A shows the density of firms by revenue. Panel B displays the average profit margin by revenue, where profit margin is defined as profits over revenue. The size of the revenue bins is 575,000 CRC.
Figure 1.6: Revenue Bunching Estimation

Figure 1.6 displays the density of firms by revenue and fits the counterfactual distribution for the first and second thresholds. In the boxes on the top right, B is the excess mass as a share of the counterfactual, and $y_u$ the revenue of the marginal buncher, obtained with the point of convergence method. The counterfactual is obtained from the regression of a polynomial of degree 5 (maximizes Akaike criteria), on all data points outside the $[y_l, y_u]$ interval. The lower bound $y_l$ is chosen by the researchers as the revenue bin which starts exhibiting excess. The upper bound $y_u$ is estimated from an iterative process: starting from $y_u$ close to the threshold, we obtain the counterfactual and estimate the excess mass (B) below the threshold and missing mass (M) above the threshold. For low $y_u$, the excess mass is larger than the missing mass, $B > M$. We increase $y_u$ until the two masses are equal, $B = M$. 

Panel A: First threshold

Panel B: Second threshold
Figure 1.7: Cost Discontinuity

Panel A: First Threshold Cost by Revenue

Panel B: Cost discontinuity (zoom)

Figure 1.7 displays the average declared cost for each revenue bin around the first threshold. To estimate the cost discontinuity at the threshold, absent revenue responses, we adjust for intensive margin revenue responses: firms declaring revenue above the threshold reduced their declared revenue, due to the tax rate increase. To take intensive responses into account, we horizontally shift firms’ costs proportionally to the elasticity of revenue, estimated from bunching. For example, given an elasticity of revenue of 0.25 and a firm with revenue of 60M: $\text{revenue}_{\text{counter}} = 60 + \epsilon_{y, 1-t} y \frac{dy}{1-t} = 60 + 0.25 \times 60 \times \frac{1}{3} \approx 61.6$. We linearly fit costs by revenue, below and above the threshold. On either side, we exclude revenue bins impacted by bunching behavior. We then extrapolate the linear fits to the threshold. The resulting cost discontinuity represents the average increase in declared costs, for a firm at the threshold, due to an increase in the tax rate from 10 to 20%.
Figure 1.8 shows the linearity of the relation between average costs and revenue. For each of the four revenue intervals considered (below and above the first and second threshold), we plot the linear and quadratic fits of the data. Quadratic fits are practically indistinguishable from Linear fits providing support to our linear extrapolation assumption.
Figure 1.9: Effective Tax Rate on Revenue

Figure 1.9 plots average tax payment as a share of revenue for revenue bins of half million CRC. The large profit elasticity implies that despite a 10% tax rate increase past each threshold, taxes effectively paid are constant at 1 to 1.5% of the firm’s revenue for all tax brackets.

Figure 1.10: Quartiles of Profit Margin by Revenue

Figure 1.10 shows the distribution of profit margins by revenue for each quartile within a revenue bin. The discontinuous step pattern observed for average profit margins is observed for all percentiles and is generated by an entire downwards shift of the distribution and not only by the reduced reported profitability of a few high profitability firms.

Note: The y-axis scale is not constant across figures
Figure 1.11: Industry Results

Within industry density (Blue) & profit margin (Green)

Figure 1.11 presents the firm density and profit margin by revenue, separating the economy in fifteen industries. In blue, the within industry firm density by revenue and in green the average profit margin by revenue. We observe for all sectors a large discontinuous drop in average profit margins, while bunching behavior is more heterogeneous.
Chapter 1: Corporate Taxation Under Weak Enforcement

Figure 1.12: Excess Mass and Profit Discontinuity by Industry

Figure 1.12 shows the relation between the profit margin discontinuity and excess mass by industry, at the first threshold.

Figure 1.13: Structural Assumption - Stable Profit Margin Distributions

Figure 1.13 shows that the distribution of profit margin is stable across revenue intervals 10 to 20% below the threshold (Panel A), and revenue intervals 20 to 30% above the threshold (Panel B). The legend indicates the revenue intervals to the threshold considered. We use an Epanechnikov kernel with bandwidth of 0.04 across all distributions. Within each panel, we never reject the Kolmgorov-Smirnov tests, that profit margins are sampled from populations with identical distributions across all pairs of revenue intervals.
Figure 1.14: Numerical Estimation of Bunching Behavior

Panel A: Revenue Elasticity Scenarios

Panel B: Estimated Revenue Elasticity

Figure 1.14 displays the results from the numerical model-based estimation, when assuming a joint counterfactual distribution of revenue and costs and that bunchers cost response correspond to the average cost response. For a given revenue elasticity, $e_y$, and cost elasticity, $e_c$, the area between the counterfactual density and the curves represents the number of bunching firms. In panel A, we display the profile of these curves for several values of the revenue elasticity. Panel B, displays the result for the last iteration: the number of estimated bunchers equals the observed bunchers and the revenue and cost elasticity are in equilibrium.
Figure 1.15: Dynamic Firm Behavior by Tax Bracket

Figure 1.15 shows the average change in firms’ profit margins between year $t$ and $t+1$ as a function of their tax brackets in year $t$ and $t+1$. On the one hand, firms remaining within their initial tax bracket hardly change reported profits. On the other hand, firms jumping to higher tax brackets drop their profit margins and symmetrically firms falling to lower tax bracket increase their profit margins.

Figure 1.16: Profit Margin Change for Growing Firms

Figure 1.16 plots the average change in firms’ profit margins between year $t$ and $t+1$ for the subset of firms who’s revenue grew by 3 to 5 Million (Left Panel) and by 7-9 Million (Right Panel). This figure visually shows the difference in differences across the group of firms that jumped past the threshold versus the firms that stayed within the same tax bracket, while controlling for revenue growth.
Figure 1.17: Correction Letters by Revenue (2012)

Figure 1.17 displays the percentage of firms receiving correction letters due to an inconsistency in their tax declaration or a discrepancy with third party information. Bunching firms are significantly more likely to receive a letter showing an inconsistency with their third-party reports. Each revenue bin represents a 2 Million CRC interval. The fitted line excludes the revenue intervals impacted by the bunching selection.

Figure 1.18: Share of Subsidiaries by Revenue

Figure 1.18 shows the percentage of firms that are subsidiaries of a larger firm, by revenue. The lack of discontinuity in the share of subsidiaries around the thresholds empirically shows that firm division to evade taxes is not a first order strategy.
Figure 1.19: Employment and Wage Bill by Revenue

Figure 1.19 shows the average number of employees and wage bill by revenue around the first and second thresholds. The data is obtained from a merge of social security records with corporate income tax returns. We display the coefficient and standard errors from the discontinuity regression at the threshold and the dummy coefficient for firms in the bunching interval.
Figure 1.20: Cost Categories Breakdown

Figure 1.20 shows the cost discontinuity by revenue, broken down into the three main cost categories reported on the tax returns (“Formulario D101”). Each cost category is displayed as a percentage of revenue. The five categories on the corporate tax returns are: administrative and operational costs, material and production costs, depreciation, interest deductions and other costs and we group the later three categories together.

Figure 1.21: Worldwide Corporate Tax Systems by Income Quintiles

Figure 1.21 shows the worldwide distribution of corporate income tax types, dividing country into three income groups. The sample contains 120 countries, with data on corporate incomes tax types collected from international tax guides by Abramovsky, Bachas, and Jensen 2015 and per-capita income data from the Penn Tables. Presumptive tax systems are very common for low-income countries, while CIT with revenue dependent rate, such as Costa Rica’s, are sometimes used in middle-income countries. In rich countries, the corporate income is almost always taxed at a flat or increasing marginal rate.
Table 1.1: Summary of Estimates and Assumptions

<table>
<thead>
<tr>
<th>Method</th>
<th>Estimation strategy</th>
<th>Elasticity</th>
<th>Bound</th>
<th>Assumptions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reduced form</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1 → $\epsilon_y$</td>
<td>Bunching: empirical point of convergence $\epsilon_{y,1-\tau} = 0.33$</td>
<td>Upper bound: $\epsilon_y$ of marginal buncher</td>
<td>Counterfactual firm density by revenue is regular</td>
<td></td>
</tr>
<tr>
<td>↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2 → $\epsilon_c(\epsilon_y)$</td>
<td>Cost discontinuity $\epsilon_{c,1-\tau} = -0.55$</td>
<td>Lower bound: adjustment for $\hat{c}_y \rightarrow \hat{c}$ negative relation with $\hat{c}_y$</td>
<td>Linear cost extrapolation is correct</td>
<td></td>
</tr>
<tr>
<td>↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 3 → $\epsilon_\pi(\epsilon_y, \epsilon_c)$</td>
<td>$d\pi = dy - dc$ $\epsilon_{\pi,1-\tau} = 4.93$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Added structure**
Profit margin distribution of firms with revenue 10% below the threshold would apply to firms with revenue just above the threshold absent the tax change

**With $dc = \bar{dc}$**

<table>
<thead>
<tr>
<th>Method</th>
<th>Estimation strategy</th>
<th>Elasticity</th>
<th>Bound</th>
<th>Assumptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1 → $\epsilon_y(\epsilon_c)$</td>
<td>Bunching: numerical elasticity $\epsilon_y$ $\epsilon_{y,1-\tau} = 0.22$</td>
<td>Bunchers would have $dc = \bar{dc}$ if not bunching</td>
<td></td>
<td></td>
</tr>
<tr>
<td>↓ Iterate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2 → $\epsilon_c(\epsilon_y)$</td>
<td>Cost discontinuity $\epsilon_{c,1-\tau} = -0.65$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 3 → $\epsilon_\pi(\epsilon_y, \epsilon_c)$</td>
<td>$d\pi = dy - dc$ $\epsilon_{\pi,1-\tau} = 4.90$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**With $dc = 0$**

<table>
<thead>
<tr>
<th>Method</th>
<th>Estimation strategy</th>
<th>Elasticity</th>
<th>Bound</th>
<th>Assumptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1 → $\epsilon_y$</td>
<td>Bunching: numerical elasticity $\epsilon_y$ $\epsilon_{y,1-\tau} = 0.095$</td>
<td>Lower bound: Assumption $dc=0$</td>
<td>Bunchers would have $dc = 0$ if not bunching</td>
<td></td>
</tr>
<tr>
<td>↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2 → $\epsilon_c$</td>
<td>Cost discontinuity $\epsilon_{c,1-\tau} = -0.79$</td>
<td>Upper bound: $\hat{c}$ negative relation with $\hat{c}_y$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 3 → $\epsilon_\pi(\epsilon_y, \epsilon_c)$</td>
<td>$d\pi = dy - dc$ $\epsilon_{\pi,1-\tau} = 4.90$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1.2: Robustness of Bunching Estimates

<table>
<thead>
<tr>
<th>Panel A: Varying the order of the Polynomial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order of Polynomial</td>
</tr>
<tr>
<td>First Threshold</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>$y_u$</td>
</tr>
<tr>
<td>$\epsilon_{y,1-\tau}$</td>
</tr>
<tr>
<td>Second Threshold</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>$y_u$</td>
</tr>
<tr>
<td>$\epsilon_{y,1-\tau}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Varying the excluded zone, $y_l$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of excluded bins</td>
</tr>
<tr>
<td>First Threshold</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>$y_u$</td>
</tr>
<tr>
<td>$\epsilon_{y,1-\tau}$</td>
</tr>
<tr>
<td>Second Threshold</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>$y_u$</td>
</tr>
<tr>
<td>$\epsilon_{y,1-\tau}$</td>
</tr>
</tbody>
</table>

Table 1.2 shows under different scenarios the estimates of the excess mass $B$, the revenue of the marginal buncher $y_u$ and the resulting revenue elasticity $\epsilon_{y,1-\tau}$. Panel A varies the order of the polynomial and Panel B the number of excluded bins on the lower side, which corresponds to $y_u$. 
Table 1.3: Cost on Revenue Distance to Threshold Relation

<table>
<thead>
<tr>
<th></th>
<th>1st Threshold</th>
<th>2nd Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cost (revenue adjust.)</td>
<td>Cost (revenue adjust.)</td>
</tr>
<tr>
<td>Jump in cost $\delta$</td>
<td>4.203**</td>
<td>2.548**</td>
</tr>
<tr>
<td></td>
<td>(0.212)</td>
<td>(0.226)</td>
</tr>
<tr>
<td>Slope below T. $\beta_1$</td>
<td>0.834**</td>
<td>0.834**</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>Slope change above T. $\beta_2$</td>
<td>0.103**</td>
<td>0.069**</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Intercept $\alpha$</td>
<td>41.971</td>
<td>41.971</td>
</tr>
<tr>
<td>Observations</td>
<td>80</td>
<td>80</td>
</tr>
</tbody>
</table>

% Jump in Cost $\frac{\delta}{\alpha}$: +10.01% +6.07% +2.37% +1.36%

Table (1.3) shows the results from the regression of average costs by revenue on revenue distance to the threshold, estimated from equation (1.14). For each threshold we report the discontinuity in cost $\delta$ with and without the revenue adjustment. The revenue adjustment holds revenue responses above the threshold constant, using the revenue elasticity estimated with bunching, such that the discontinuity in the cost to revenue relation, after adjustment, only identifies cost responses: the results from column (2) & (4) are our main estimate of the cost discontinuity. An observation is a revenue bin of 0.575 Million Colones. Standard errors are shown in parentheses and stars indicate statistical significance level. *=5% level, **=1% level.
Table 1.4: Alternative Models for Cost Discontinuity by Revenue

<table>
<thead>
<tr>
<th>Model Specification</th>
<th>1st Threshold</th>
<th>2nd Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Narrow Window</td>
<td>Wide Window</td>
</tr>
<tr>
<td>Jump in cost $\delta$</td>
<td>2.688**</td>
<td>2.465**</td>
</tr>
<tr>
<td></td>
<td>(.249)</td>
<td>(.204)</td>
</tr>
<tr>
<td>Slope below T.</td>
<td>.823**</td>
<td>.841**</td>
</tr>
<tr>
<td></td>
<td>(.012)</td>
<td>(.007)</td>
</tr>
<tr>
<td>$\Delta$ Slope above T.</td>
<td>.079**</td>
<td>.063**</td>
</tr>
<tr>
<td></td>
<td>(.017)</td>
<td>(.011)</td>
</tr>
<tr>
<td>Quadratic below T.</td>
<td>- .009**</td>
<td>.003</td>
</tr>
<tr>
<td>$\Delta$ Quadratic above T.</td>
<td>.009**</td>
<td></td>
</tr>
<tr>
<td>Intercept, $\alpha$</td>
<td>41.86</td>
<td>42.046</td>
</tr>
<tr>
<td>Observations</td>
<td>70</td>
<td>90</td>
</tr>
</tbody>
</table>

Table 1.4 shows the regressions of average costs by revenue on revenue for different model specifications. The parameter of interest is the jump in declared costs at the threshold, $\delta$, from Equation (1.14). Compared to the main specification of Table (1.3), Rows (1)-(2) & (5)-(6) vary the revenue interval over which the line is fitted. Rows (3) & (7) assume that the revenue elasticity is falling with revenue, at the speed estimated between the first and second threshold. Rows (4) & (8) assume a quadratic fit instead of a linear fit. An observation is a revenue bin of 0.575 Million Colones. Standard errors are shown in parentheses and stars indicate statistical significance level. *=5% level, **=1% level.

Table 1.5: Elasticity Estimates - Point of Convergence

<table>
<thead>
<tr>
<th>$y^*$</th>
<th>Parameters</th>
<th>Elasticity</th>
<th>Threshold jump</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$dy^*$</td>
<td>Revenue</td>
<td>Cost</td>
</tr>
<tr>
<td></td>
<td>$1 - \tau_0$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\tau^*$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\Delta y$</td>
<td>$\Delta c$</td>
</tr>
<tr>
<td>50</td>
<td>8.3**</td>
<td>0.9</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>(1.3)</td>
<td>(0.07)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>100.5</td>
<td>7.2**</td>
<td>0.8</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>(1.8)</td>
<td>(0.04)</td>
<td>(0.05)</td>
</tr>
</tbody>
</table>

Table (1.5) shows the elasticity estimates obtained with the point of convergence method. Cost elasticities are estimated with the threshold discontinuity, holding constant revenue responses. The profit elasticity combines revenue and cost responses. Standard errors are obtained through 1,000 bootstrap iterations. $y^*$ is the revenue threshold in Million CRC and $dy^*$ is the revenue response of the marginal buncher estimated with bunching. $1 - \tau_0$ is the average tax rate below each threshold and $\tau^*$ is the implicit marginal tax rate faced by the marginal buncher. Standard errors are shown in parentheses and stars indicate statistical significance level. *=5% level, **=1% level.
Table 1.6: Model-Based Numerical Estimation: Iteration Steps

<table>
<thead>
<tr>
<th>Iteration Step</th>
<th>Revenue Elasticity $\epsilon_{g,1-t}$</th>
<th>Cost Jump $dc$</th>
<th>Cost Elasticity $\epsilon_{c,1-t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.33</td>
<td>2.55</td>
<td>-0.55</td>
</tr>
<tr>
<td>2</td>
<td>0.17</td>
<td>3.25</td>
<td>-0.70</td>
</tr>
<tr>
<td>3</td>
<td>0.26</td>
<td>2.83</td>
<td>-0.61</td>
</tr>
<tr>
<td>4</td>
<td>0.19</td>
<td>3.17</td>
<td>-0.68</td>
</tr>
<tr>
<td>5</td>
<td>0.24</td>
<td>2.93</td>
<td>-0.63</td>
</tr>
<tr>
<td>6</td>
<td>0.21</td>
<td>3.07</td>
<td>-0.66</td>
</tr>
<tr>
<td>7</td>
<td>0.23</td>
<td>2.98</td>
<td>-0.64</td>
</tr>
<tr>
<td>Final</td>
<td><strong>0.22</strong></td>
<td><strong>3.02</strong></td>
<td><strong>-0.65</strong></td>
</tr>
</tbody>
</table>

Table 1.6 shows the iteration steps of the model-based numerical bunching estimation. With a counterfactual firm density by revenue and a counterfactual profit margin distribution, the method numerically estimates the number of bunching firms as a joint function of their revenue distance to the threshold and costs. Step 1 uses as initial values the revenue and cost elasticity from section 1.3. Under those parameters the revenue elasticity of Step 2 is sufficient such that the number of bunchers equal that of the excess mass. With this new revenue elasticity we re-estimate the cost elasticity using the discontinuity method of 1.3. We iterate this procedure until we find the fixed point at which the revenue and cost elasticities are consistent with the observed bunching behavior.

Table 1.7: Industry Level Results 1st Threshold

<table>
<thead>
<tr>
<th>Sector</th>
<th>Profit Margin (%)</th>
<th>Bunching</th>
<th># Firms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Drop Base % Drop</td>
<td>Excess Mass Total % Below T1</td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>-4.1 8.4 48.8</td>
<td>1.95 33,095 59.5</td>
<td></td>
</tr>
<tr>
<td>Manufacture</td>
<td>-3.2 6.8 47.1</td>
<td>2.34 34,799 45.4</td>
<td></td>
</tr>
<tr>
<td>Construction</td>
<td>-5.2 9.5 54.7</td>
<td>3.17 26,410 51</td>
<td></td>
</tr>
<tr>
<td>Wholesale &amp; Motor Vehicle</td>
<td>-3.5 7 -50</td>
<td>1.11 63,544 45.1</td>
<td></td>
</tr>
<tr>
<td>Retail</td>
<td>-4.9 8.5 57.6</td>
<td>1.17 100,552 47.9</td>
<td></td>
</tr>
<tr>
<td>Hotel &amp; Restaurants</td>
<td>-3.5 7 -50</td>
<td>1.41 21,483 49</td>
<td></td>
</tr>
<tr>
<td>Transport</td>
<td>-4.1 9.9 41.4</td>
<td>2 36,294 54.7</td>
<td></td>
</tr>
<tr>
<td>Financial Activities</td>
<td>-10.3 21.8 47.2</td>
<td>3.93 26,366 71.9</td>
<td></td>
</tr>
<tr>
<td>Real Estate</td>
<td>-13 36.4 35.7</td>
<td>4.05 91,525 85.1</td>
<td></td>
</tr>
<tr>
<td>Legal &amp; Econ. Consultants</td>
<td>-9.5 17 55.9</td>
<td>6.27 64,617 73.3</td>
<td></td>
</tr>
<tr>
<td>Other Services</td>
<td>-9.3 14.6 63.7</td>
<td>4.37 37,091 69.3</td>
<td></td>
</tr>
<tr>
<td>Education &amp; Culture</td>
<td>-1.3 5.8 22.4</td>
<td>2.64 14,228 56.8</td>
<td></td>
</tr>
<tr>
<td>Health</td>
<td>-8.4 17.1 49.1</td>
<td>3.23 19,611 65.2</td>
<td></td>
</tr>
<tr>
<td>NGO &amp; Public Admin.</td>
<td>8 28.1 2.8</td>
<td>1.9 10,068 68.8</td>
<td></td>
</tr>
<tr>
<td>Undetermined</td>
<td>-9.6 19.9 48.2</td>
<td>3.66 36,044 80.8</td>
<td></td>
</tr>
</tbody>
</table>
### Table 1.8: Dynamic Firm Behavior

<table>
<thead>
<tr>
<th></th>
<th>Profit Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st Threshold</td>
</tr>
<tr>
<td>Revenue (Million CRC)</td>
<td>0.0115**</td>
</tr>
<tr>
<td>Higher Tax Bracket</td>
<td>-3.06**</td>
</tr>
<tr>
<td>Buncher (Narrow)</td>
<td>1.56**</td>
</tr>
<tr>
<td>Bunching (Broad)</td>
<td>0.84**</td>
</tr>
<tr>
<td>Above threshold (Narrow)</td>
<td>-0.33</td>
</tr>
<tr>
<td>Above threshold (Broad)</td>
<td>-0.12</td>
</tr>
<tr>
<td>Constant (Avg across years)</td>
<td>14.63</td>
</tr>
</tbody>
</table>

All firms with revenue in a 70 Million CRC window centered around the thresholds are included in the sample. Profit margin is defined as profit over revenue. The Bunching and above threshold are dummies for declaring revenue in the intervals around the threshold. Bunching narrow is defined as the having revenue in the half Million interval below the threshold. Bunching wide as having revenue between 4 and 0.5 Million below the threshold. Above threshold narrow is defined as having revenue between 0 to 3 Million above the threshold and wide as having revenue 3 to 9M above threshold. Standard errors are shown in parentheses and stars indicate statistical significance level. *=5% level, **=1% level.

### Table 1.9: Revenue Shifting at the End of the Fiscal Year

<table>
<thead>
<tr>
<th></th>
<th>Dependent Variable: Monthly Revenue (Million CRC)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All firms</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Buncher*Sept</td>
<td>0.10</td>
</tr>
<tr>
<td>Buncher*Oct</td>
<td>-0.02</td>
</tr>
<tr>
<td>Firm FE</td>
<td>NO</td>
</tr>
<tr>
<td>Observations</td>
<td>596,705</td>
</tr>
</tbody>
</table>

Table 1.9 tests for revenue shifting at the end of the fiscal year using the revenue declared for the monthly sales tax payment. Robust standard errors are shown in parentheses and stars indicate statistical significance level. *=5% level, **=1% level.
Table 1.10: Threat of Audit Impact at the Industry Level

<table>
<thead>
<tr>
<th>Outcome (Million CRC)</th>
<th>Control 1: firms too small to bunch</th>
<th>Control 2: Dominated firms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Bunch<em>Audit</em>Post</td>
<td>4.87</td>
<td>6.43</td>
</tr>
<tr>
<td></td>
<td>(2.39)**</td>
<td>(2.24)***</td>
</tr>
<tr>
<td>Bunch*Post</td>
<td>-0.22</td>
<td>-0.48</td>
</tr>
<tr>
<td></td>
<td>(2.13)</td>
<td>(2.00)</td>
</tr>
<tr>
<td>Audit*Post</td>
<td>5.55</td>
<td>5.88</td>
</tr>
<tr>
<td></td>
<td>(3.10)*</td>
<td>(3.01)*</td>
</tr>
<tr>
<td>Firm FE</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>7,203</td>
<td>7,203</td>
</tr>
</tbody>
</table>

Table 1.10 shows the results of the triple difference regression estimated from equation 1.20. The coefficient of interest is the triple interaction $\text{Bunch} \times \text{Audit} \times \text{Post}$ which shows the change in reported revenue, costs and profits of bunchers following an increase in their audit risk at the industry level. Standard errors are clustered at the industry level and shown in parentheses. Stars indicate statistical significance level. *=10% level, **=5% level, ***=1% level.
Chapter 2

An Empirical Test of Information Trails: Do Firm Finance and Size Matter for Tax Collection?

with Anders Jensen

2.1 Introduction

Governments enforcing taxes are constrained by the stock and flows of information on transactions in the economy. Third-party reporting, verifiable paper trails, and whistle-blowing mechanisms have been shown empirically to create important information flows that facilitate tax enforcement (Kopczuk and Slemrod 2006, Kleven et al. 2011b, Kumler, Verhoogen, and Frias 2012, Pomeranz 2015b, Carillo, Pomeranz, and Singhal 2015).

In the theoretical literature, two settings in particular have been emphasized as powerful generators of tax-relevant information: the firm’s interaction with the financial system and the firm’s worker size. The financial trail is based on the insight that when accessing formal banking firms generate a paper trail of information, often including financial statements and lists of customers Gordon and Li (2009b). If this financial information trail is shared by the bank with the tax authority, it may constitute the basis for stronger tax enforcement on financially connected firms.

The firm-size information trail is related to the idea that firms with more employees are more complex and thus require more accurate book-keeping (Kleven, Kreiner, and Saez 2016). In the hands of a tax auditor, such book-keeping constitutes the paper trail necessary to enforce tax liability. Larger firms can still evade by keeping parallel books (one for their own use, one for the tax authority) but evasion will be more costly and collusion to under-report activities between employer and employees more likely to break down. This could be due
either to random shocks that reveal the true internal records or to whistle-blowing. Despite the importance of the worker-size and the financial information trails in the theoretical literature no studies has tested their empirical relevance. We provide the first empirical test to study whether larger-sized firms and more financially externally reliant firms face higher tax inspection and comply more with sales taxes. To motivate the empirical specification, we formulate a reduced-form model in which firm behavior generates information trails that tax authorities can use to enforce tax collection.

We use the World Bank Enterprise Surveys (WBES) 2002-2010, a firm level survey with information on firm-size, external financial reliance, tax inspection and reported sales tax compliance for 108,000 firms in 79 countries. In Fig.1, pooling all countries together, we group firms in twenty equal-sized groups (vingtiles) of firm-size distribution, and plot mean tax inspection and sales tax compliance across the twenty vingtiles. Enforcement and compliance clearly increase monotonically by firm-size. In Fig.2 we construct the same graph, with twenty quantiles of external financial reliance. The upwards-sloping relationship is much less clear, both with regards tax inspection and sales tax compliance. These figures summarize a general trend in all our regressions results: we systematically uncover a strong positive effect of the firm-size information trail on tax inspection and sales tax compliance and an insignificant effect of the external finance trail.

In the regression setting, we isolate exogenous demand-driven variation in external funding and firm-size at the industry-level, by adopting an identification strategy related to Rajan and Zingales (1998), (RZ hereafter). The identification strategy consists in predicting the size-ranks and the external finance-ranks of an industry within a WBES country from the size and external finance rankings of the same set of industries in the U.S. This out-of-sample extrapolation technique relies on the assumption that the labor and finance markets of the U.S. are relatively undistorted and thus that the ranking of industries in external reliance and number of workers in the U.S. reflects “benchmark” technological differences in demand for labour and funding inputs. The first stage shows the extent to which US industry-ranks in firm-size and funding choice explain the WBES countries’ corresponding industry-ranks. As RZ did not have data on external funding for firms in developing countries, we believe our study is the first to test the hypotheses of industry-specific country-invariant demand technologies for finance and firm-size in a large dataset of both developed and developing countries. We find strong evidence in favor of the industry-specific technology for workers, but weak evidence in the case of financial external reliance. In a double IV specification, which simultaneously instruments external reliance rank and firm-size rank by the U.S. rankings in external reliance and the US firm-size, we uncover strong effects for the worker-size trail on both tax inspection and sales tax compliance, but no discernible impact of external reliance. Finally, we study the impact of the information trails on tax evasion through the tax inspection channel. The exclusion restriction is less strong in this setting, as information trails could impact tax compliance through other channels than solely tax inspection. With that caveat in mind, we find that exogenous changes in firm-size and external reliance predict large positive changes to tax compliance through increased tax enforcement. These results speak to the importance of information trails and third party information in understanding.
firm level tax compliance behavior. The results also relate to the revenue versus production efficiency trade-off the tax authority faces when choosing enforcement instruments. On the one hand, the tax administration may consciously choose to leave aside relevant information trails generated from firm-size and external funding, since any enforcement based on firm size and funding will distort the firm input choices and generate socially costly production inefficiencies. On the other hand, using the information trails would help to maximize revenue for a given amount of enforcement effort: tax collection yield will be higher for the information-intensive firms since their paper trails allow the discovery of more evaded taxes. Empirically we find a significant firm-size effect on the probability of tax inspection, suggesting tax administrations do actively levy the size-information trail via targeted enforcement on larger firms. We also find that tax compliance increases significantly when due to size-based enforcement, which suggests that levying the firm-size information trail generates large tax revenue returns.

The non-results concerning the financial trail are interesting in their own right. The insignificance in all regressions could point to the lack of willingness to use financial information by tax administrations or to the lack of access to such information. Despite an important push for the availability of bank information to tax authorities (OECD 2000, 2006) bank secrecy laws have often prevented significant improvements. In many countries, even in the case of an audit, accessing bank information can be long and tedious and typically requires the signature of a judge.

This paper contributes to the literature focusing on the importance of administrations’ information constrains in understanding tax enforcement and tax structure policy. Brockmeyer et al. (2015) study the minimum tax scheme in Pakistan where firms are taxed either on profits or turnover depending on whichever liability is larger. The authors show that this production inefficient policy can be rationalized by revenue-efficiency considerations, since it is typically harder to evade on the broader turnover tax base than on the profit tax base. Wingender (2008) tests of the relation between finance and evasion at the industry level, using an empirical strategy related to ours, uncovers a much larger impact of the finance information trail than we do. Other information trails have also been discussed. For example, Pomeranz (2015b) studies the paper trail generated by VAT payments among firms in Chile. She shows that an audit threat to a firm producing at the end of the supply chain affects tax payments for the whole chain of suppliers. In addition she shows that firms already covered by a paper trail are not affected by the audit threat. She concludes that the VAT chain acts as a substitute for the paper trail in case of an audit. Artavanis, Morse, and Tsoutsoura (2012) show that a large Greek bank had informally embedded into its credit models total income instead of reported income. However, the authors argue that the lack of tax enforcement does not arise because of a lack of information but because of a lack of political will. In this case the information trail is left on the side, not for production-efficiency reasons, but due to political factors. Finally, Besley and Persson (2013b) develop a series of models which study the government’s choice of investment in tax capacity, recognizing the practical limits imposed on administration policy, including information constraints and political instability. Section 2 outlines an industry-representative firm model of external reliance, firm-size, tax
enforcement and tax compliance, and derives the two sets of empirical predictions. Section 3 discusses the data and the identification strategy. Section 4 presents the results for tax inspection and sales tax evasion and then performs robustness checks. Section 5 concludes.

2.2 Theory

Industry-representative firm specification

In order to study the relationship between information trails and tax under-reporting, we model the reduced-form interactions between access to external funds and tax enforcement on the one hand, and between firm-size and tax enforcement on the other hand.

Number of Workers

The number of workers channel, formalized by Kleven, Kreiner, and Saez (2016), highlights random shocks and whistle blowing as two mechanisms which could reveal the true business records to the tax authority. The model generates a critical threshold firm-size, \( N \), below which full evasion can be the optimal strategy. Further, the optimal amount of under-reporting on the intensive margin will be decreasing in \( N \) above the threshold: in the whistle-blowing case, the reward to the whistle blower will be a share of total uncovered revenue, which is increasing in the size of the firm; in the random shocks case, the probability that any employee reveals the true, internal records, is also increasing in the total number of workers employed.

External Finance Reliance

The second information channel is generated by a firm’s interaction with the financial markets. The important distinction is between the use of internal and external sources of funding, where the latter is characterized to generate an information trail on firm’s real activities. We model the choice of internal and external cash flow based on Kaplan and Zingales (1997), a workhorse model in the finance literature. In a one-period setting, a firm chooses a level of investment \( W \) to maximize (net-of-tax) profits. Any investment \( W \) can be financed with internal and with external funds, denoted respectively \( I \) and \( E \):

\[
W = I + E. 
\]

The cost of internal funds is its opportunity cost in capital markets, which we normalize to 1. There exists a wedge, \( r \), between the firm’s internal and external cost of funds: \( r \geq 1 \). This wedge could arise due to information or agency problems, which are passed on as costs to the firm. Investment financed internally is limited by the upper bound of available internal funds, \( \bar{I} \). The firm should therefore use external finance only once it has exhausted its internal funds and reached the \( \bar{I} \) limit.

Formal Setting

Our model incorporates the internal and external funding choice and number of workers choice into a standard representative-firm setting of under-reporting Chetty (2009b). It models the firms tax-gain benefit of evasion \( e \) weighted against the evasion cost \( c(e) \). Later
in the text we discuss some possible functional forms for \( c(e) \). Under convexity of \( c(e) \) there will be an interior solution for evasion \( e \).

Based on our above discussion, and a novel feature to this type of reduced-form evasion-cost models, we model the probability and cost of getting detected \( c(e; E; N) \) as increasing in the level of evasion \( e \); in the amount of external funds \( E \); and in the number of workers \( N \) weakly above the threshold \( \bar{N} \). The industry-representative firm faces the optimization problem:

\[
\max_{\{I,E,N,e\}} \ (1 - t) \ (f(W,N) - e) + e - c(e, E, 1 \ (N \geq \bar{N}) \cdot N) - (rE + I) - wN \tag{2.1}
\]

subject to \( W = I + E \ I \leq \bar{I} \ e \leq f(W,N) \) We define two empirical variables directly observable in the data and which will allow us to derive the empirical predictions:

- \( \alpha \) is external finance reliance, the share of a firm’s finance that is not met by internal funds: \( \alpha = 1 - \frac{I}{W} = \frac{E}{E + I} \). This is equivalent to the Rajan and Zingales empirical measure of external financial reliance.
- \( \gamma \) is the share of production reported for tax purposes: \( \gamma = \frac{f(W,N) - e}{f(W,N)} \).

Given the above structure, we can characterize the optimal choice of internal and external funds, number of workers, and evasion. The representative firm makes optimal choices based on three first-order conditions, which are characterized by

\[
\text{FOC wrt } e: t = c_e(e, E, N) \tag{2.2} \\
\text{FOC wrt } E: (1 - t)f_W(W,N) = r + c_E(e, E, N) \tag{2.3} \\
\text{FOC wrt } N: (1 - t)f_N(W,N) = w + c_N(e, E, N) \tag{2.4}
\]

**Modeling the probability of tax inspection**

The reduced-form evasion-cost \( c(e, E, N) \) contains the first central prediction of our study, namely that firms with more workers and with more external finance face higher tax enforcement. There are a number of distinct ways to model these enforcement information trails. One possibility is to specify a form for \( c(\cdot) \) as the expected monetary loss under a tax audit, such as

\[
c(e, E, N) = p(E, N) \cdot (\vartheta e) \tag{2.5}
\]

where \( p(E, N) \) is the probability a firm will receive a tax inspection - assumed to be independent of evasion - and \( \vartheta \) is the amount of evaded taxes that have to be paid in fines. Under this formulation, a firm which is perfectly compliant \( e = 0 \) faces no compliance-cost, independently of how large and externally reliant it is. On the other hand, it may be that larger and more externally dependent firms face higher administrative compliance costs simply from the time spent dealing with tax inspectors, such that \( c(0, E, N) > 0 \). This could be modeled by introducing an firm administrative cost of compliance \( \varphi \), such that:

\[
c(e, E, N; \varphi) = p(E, N) \cdot (\varphi + \vartheta e) \tag{2.6}
\]
Finally, the amount evaded can enter directly the probability of tax inspection, $p(e, E, N)$, so that evading is more costly both because detection is more likely and because fines are larger.

One policy that a large number of administrations have followed to manage enforcement of taxes on subsets of the tax-payer population is the establishment of special dedicated units, usually referred to as Large Taxpayer Units (LTU hereafter). A firm’s inclusion in the LTU implies an increase in the auditing probability and intensity, corresponding to a discrete upwards jump of $p(\cdot)$ in our model. As anecdotal policy-evidence to support modeling $p = p(N)$, firm size $N$ seems to constitute one of the main criteria to assign a firm to the LTU. Pooled survey responses from administrative revenue authorities in 67 countries\(^1\) indicate that a threshold criteria in terms of number of employees is amongst the 6 most common criteria for allocating a firm to the country LTU.\(^2\) In both Denmark and the United Kingdom a firm is assigned to the LTU once it employs in excess of 250 people; in Ghana, the threshold lies at 500, while in Sweden it is at 800; in Lithuania, allocation to the LTU is based on jointly meeting the criteria of 10+ employees and large sales revenue. Other countries seem in practice to also target audits based on size. Goyette (2012) shows using firm-level data in Uganda that sales tax audits are effectively based on number of employees rather than the official cut-off rule which is in terms of sales. In the Finnish setting, Harju, Kosonen, and Ropponen (2014) discuss how the desk audit inspection probability for VAT returns is increasing in size for labor intensive firms. Similarly information on finance is used to select audit cases. The World Bank Risk-based tax audits report discusses the case of Turkey which automatically cross checks VAT receipts with credit card information (Risk-Based Tax Audits: Approaches and Country Experiences 2011).

**Empirical predictions of the model**

There are two sets of empirical testable implications that the model delivers. The first concerns the relationship between tax inspection probability and the firm information trails, $E$ and $N$.

**Empirical prediction (1)**

The industry-representative firm tax inspection probability is increasing in firm-size and the amount of external funds

$$\frac{\partial p(E, N)}{\partial N} \geq 0 \text{ and } \frac{\partial p(E, N)}{\partial E} \geq 0$$

(2.7)

The positive derivative of the tax inspection probability with respect to the two information trails is motivated by the discussions in Section 2.1. In the case where $N$ and $E$ are chosen as

---


\(^2\)Other criteria cited include annual gross turnover; potential revenue contribution of the firm; size of assets; and significant international business activity.
interior solutions (i.e. \( N > \bar{N} \) and \( E > 0 \)), the first-order conditions (2) – (3) – (4) allow us to directly derive the second set of empirical predictions of our model. Supposing some form of underlying heterogeneity cause some firms to optimally choose higher levels of funding and size, this will impact the firm’s compliance decision towards larger reliance on external financing, \( \alpha \), and towards an increase in the optimal size of the firm, \( N \); these changes lead to higher tax-compliance, partly through increased tax inspection intensity. Formally

**Empirical prediction (2)**

An exogenous increase in firm-demand for labor input causes sales tax compliance to increase in general

\[
\frac{\partial \gamma}{\partial N} \geq 0
\]

and in particular through the channel of tax inspection

\[
\frac{\partial \gamma}{\partial p(E,N)} \frac{\partial p(E,N)}{\partial N} \geq 0
\]

Similarly, the total impact of an exogenous increase in firm reliance on external finance is for sales tax compliance to increase

\[
\frac{\partial \gamma}{\partial \alpha} \geq 0
\]

and a specific channel is through increased tax inspection

\[
\frac{\partial \gamma}{\partial p(E,N)} \frac{\partial p(E,N)}{\partial E} \geq 0
\]

In our empirical setting we will try to isolate exogenous changes in demand for labor and external finance input which should drive sales tax compliance through the information trails and the change to the cost of evasion: these are the direct tests (8) and (10). But we also provide the empirical tests for the impact of information trails on sales tax compliance through changes in tax inspection probability, given by (9) and (11). In both empirical predictions, we simply assume that there exists heterogeneity at the industry-level causing variation in input demands for size and external funding. One way to model this is to suppose that there exists an industry-specific productivity-parameter \( \theta_i \), and that industries differ in the following sense

\[
\theta_i > \theta_j \Rightarrow f_W(W,N;\theta_i) > f_W(W,N;\theta_j) \text{ and } f_N(W,N;\theta_i) > f_N(W,N;\theta_j)
\]

In the empirical exercise, we will treat differences in demand for labor and external reliance between two U.S. industries \( i \) and \( j \) as reflective of underlying technological differences.
2.3 Data and Identification Strategy

Data description

Throughout the analysis the unit of observation is a country-industry, where industries are classified using the ISIC 3 digit methodology, produced by the U.N. Statistics Division. The World Bank Enterprise Survey is the primary data source: it is a firm level survey collected by the World bank between 2002 and 2010 in more than a hundred countries. To our knowledge this is the only standardized firm data set spanning countries across the development spectrum. Since the survey was clearly administered by a third party not related to the government or the statistical office it could ask unique questions on tax behavior and corruption.

We rely on two data sources to construct our instrumental variables: the Compustat database of publicly listed firms in the U.S. and the U.S. Bureau of Labour Statistics Quarterly Census of Employment and Wages. The final sample is 2597 country-ISIC3 observations from 77 countries.\(^3\)

Tax inspection and tax compliance

To operationalize a measure of tax enforcement, we use the firm’s reported answer to the question: “Total days spent with officials from: tax inspectorate.” In particular we constructed a tax inspection dummy equals to one if the firm has been visited for at least one day by tax auditors. This proxies for the extensive margin of tax inspection. The industry-average tax inspection probability is .71, with a standard deviation of .35. This is the measure of tax enforcement which we used to produce Figs.1-2 in the introduction. In , we also report the results from the ‘intensive margin’ measure of total number of days. These results using the ‘intensive margin’ number of days visited by tax inspectors are discussed in Section 4.3.\(^4\)

In a second stage of the analysis we use the WB survey question on sales tax evasion: “Recognizing the difficulties many enterprises face in fully complying with taxes and regulations, what percentage of total sales would you estimate the typical establishment in your area of activity reports for tax purposes?” There are obvious difficulties in assuming that the firm’s answer to this question reflects its own industry tax evasion and even harder that it is the firm’s own evasion level, but since our theory suggests that information trails ultimately impact tax compliance, we also use this measure of sales tax compliance (in Section 3.4).

The average share of sales reported for tax purpose is 84.21% with a standard deviation of 20.79%.

\(^3\)For details on construction of the dataset, please see Appendix B.

\(^4\)Regression results from using the continuous number of days measures are in Appendix A, Tables 5-6.
External reliance

To construct a measure of external reliance, we follow the methodology presented in Rajan and Zingales (1998). The measure of external dependence proxies for the amount of investment that is not be financed through internal sources such as retained earnings. For the U.S. values of this measure, we use Compustat over the years 2000-2012. In the World Bank sample, firms’ balance sheets give information on the shares of financing from different sources of earnings. The measure of external reliance is defined as the ratio of external sources’ share over the sum of external and internal shares:

\[ \text{share ext}_i = \frac{\text{external}_i}{\text{external}_i + \text{internal}_i} \]

Internal finance is the sum of retained earnings, friends and family and informal finance. External finance is defined as banking finance which includes loans, overdrafts and credit card finance. In Appendix B we provide some summary statistics on the different sources of financing. Retained earnings is the dominant method of financing. A compelling feature of our data is the possibility to construct the same measure of external reliance in both the World Bank and the Compustat sample.

Firm size

We choose as number of workers the Enterprise Survey firm’s reported average number of permanent workers. Workers which are permanent employees are more likely to have access to the internal books of the firm and thus are more able than temporary employees to 'blow the whistle' and report tax under-reporting. The sample-wide distribution of firms has a long right tail: while the median is at 17 workers, the mean of 105 is largely influenced by a few very large firms. Given the shape of the marginal density, we always use the natural log of firm size in the regressions. U.S. data is taken from the 2002 Census of Employment and Wages.

Figure 3 uses the pooled Enterprise Survey firm-level observations to plot the density distributions of firm-size in terciles of the external financial reliance distribution: zero, low, and high. The size-distribution appears increasingly right skewed as external funding reliance decreases which could suggest non-zero cross-derivatives of the production function and of the evasion-cost function with respect to inputs \( N \) and \( E \). Though less clear, this 'interaction' between the inputs is also reflected in the density distributions of external reliance across terciles of the firm-size distribution, plotted in Figure 4. The possibility of interactions in the production and evasion functions highlights the importance of estimating jointly the two first stages to predict exogeneous variation in firm-size and external reliance, as detailed in the following section.

\(^5\) Detailed description of this ratio is in Appendix B

\(^6\) We believe that these sources all generate an information trail in the case of external funding, and are non-visible in the case of internal funds. The main findings are robust to only including retained earnings as internal funding source and local bank access as external funding source.
Chapter 2: An Empirical Test of Information Trails

Identification strategy

A key identification issue is that a firm’s choice of workers and finance in our sample of surveyed countries could be distorted by the information trails. Our identification strategy relies on out-of-sample extrapolation techniques at the industry-level: if firms across countries, but within the same industry, share a common underlying technological demand for inputs, then we can use the demand for workers and finance of firms in a relatively undistorted market, like the US, as instruments for firms’ demand for workers and external funding in our sample. This idea was developed by Rajan and Zingales (1998) who studied whether sectors reliant on external finance grew faster in countries with more developed financial institutions.

The between-industry regressions thus rely on two assumptions for identification. First, the US labor and financial markets must represent a benchmark, where firm-choices of workers and external funding are undistorted, in particular with respect to the information-trail enforcement channels. Rajan and Zingales (1998) provide a series of arguments to support this assumption, which ultimately remains hard to test.

The second assumption is that there exist an industry-specific technological demand for external finance and workers. In a regression with country fixed effects, the identifying assumption holds if the following type of statement is true: “If Drugs and Pharmaceuticals require more external funding than Tobacco in the US, then this rank-condition also holds in Bangladesh.” Figure 3 provides evidence to support this assumption. For four selected countries from the World Bank sample across all levels of per capita income (Burkina Faso, Argentina, Slovenia, Germany), the graph plots the firm-size rank of a given industry in the U.S. sample against the industry’s rank in the surveyed country sample. In each country, the upward-slope is clear with a linear trend, although there are some outliers. The identifying assumption is equivalent to a significant upward-slope being empirically verified in the set of surveyed countries.

We estimate jointly the two first-stages

\[(\text{Size Rank})_{i,c} = \gamma (\text{Size Rank})_{i-US,c} + \beta (\text{External Fin. Rank})_{i-US,c} + \delta_c + \varepsilon_{ic} \] (2.13)

\[(\text{External Fin. Rank})_{i,c} = \omega (\text{External Fin. Rank})_{i-US,c} + \pi (\text{Size Rank})_{i-US,c} + \tau_c + \varepsilon_{ic} \] (2.14)

where e.g. \((\text{Size Rank})_{i,c}\) is the firm-size rank of ISIC3 industry \(i\) in country \(c\), \((\text{Size Rank})_{i-US,c}\) is the US-rank of the same industry, and \(\delta_c\) is a country fixed effect. Significant \(\gamma > 0\) and \(\omega > 0\) coefficients provide support for the assumption of country-invariant but industry-dependent demand functions for labor and finance inputs, with coefficients equal to 1 implying perfect rank-ordering of country \(c\)’s industries relative to U.S. industries. To the best of our knowledge, this study is the first to provide direct evidence on the assumption underlying the RZ study and the ensuing literature in finance and development building on their results.

In a second stage, we study the impact of demand for labor and external finance on the probability of tax enforcement. That is, we estimate

\[1 (\text{TaxInspection})_{ic} = \beta \cdot (\text{Firm Size Rank})_{ic} + \alpha \cdot (\text{External Reliance Rank})_{ic} + \theta_c + \varepsilon_{ic} \] (2.15)
where the rank of firm size and external reliance may be replaced by their predicted values from the first stage.
Using the rank of firm size as opposed to actual firm size values to predict tax enforcement is motivated by the idea that in the absence of administrative and information constraints within a country, tax authorities will effectively target enforcement and auditing towards industry-firms in descending order of these industries’ reliance on external finance and labor inputs. Thus actual firm-size and external finance only matter to the extent that they determine an industry’s relative location in terms of the country-wide distributions of labor input and external dependency. A rank-rank regression also provides a more direct test of the identification assumption than a regression using actual values.\(^7\)

\section*{2.4 Regression Results}

In 4.1 we present the impact of information trails on the probability of tax inspection. In 4.2 we go one step further and study the effect of an increase in tax enforcement, predicted by the information trails, on sales tax compliance. In the final section 4.3 we test the robustness of the results.

\subsection*{Main results}

Table 2 presents the results of information trails on tax inspection. Column 1 presents the linear regression of tax inspection probability on the surveyed country’s industry-rank in terms of workers and finance.\(^8\) The results show that while the rank of firm-size positively predicts the probability of tax inspection, there is no effect for the external reliance channel. The joint explanatory power of the two trails in determining tax enforcement is strong (F-test equal to 16.07).

The non-impact of external reliance may arise if the true relationship between external funding and tax inspection is highly non-linear, while our OLS assumes a linear relationship. To get at this, Figure 6 illustrates graphically the regression result of Col.1 of Table 1 using binned scatter plots. To construct the left (right) plot, we first take the residuals of the the tax-inspection variable and the firm-size rank (external reliance rank) variable with respect to external reliance rank (firm-size rank) and country fixed effects. We then divide the residual firm-size rank (external reliance rank) into twenty equal-sized groups (vingtiles) and plot the means of the tax-inspection residual within each bin against the mean value of residual firm-size rank (external reliance rank). The solid line shows the best linear fit estimated on the

\(^7\)In Appendix A Tables 9-10, we show that using actual industry-average values of size and external reliance yield very similar results.

\(^8\)All standard errors are clustered at the country level, to account for country-wide correlation. We use as analytical weights the ratio of country-industry number of firm observations to the total number of firm observations in the raw Enterprise Survey sample; country-industry averages which reflect a larger underlying number of firms, and thus are likely to be more important and more accurate, receive higher weight in the regression.
underlying micro-data using OLS and corresponds to the regression coefficient for firm-size (external reliance) in Col.1 Table 1.\(^9\) The scatter-plot provides no evidence of a non-linear underlying relationship, and the flat linear OLS provides a good fit. The linear OLS also seems to fit well the size-tax inspection scatter plot. This linear relationship substantiates our choice of modeling the inspection dummy outcome as a linear probability model over size and external finance inputs.

In column 2, we present the reduced form results based on the extrapolation technique where the US-industry ranks of firm size and external reliance are used to directly predict cross-industry variation in tax enforcement in the WBES country. The coefficient on firm size is small but highly significant at the 1\% level; a one standard deviation increase in the US firm-size industry-ranking leads to a 2.38 percentage points increase in the probability of tax inspection for that industry in the surveyed country. However, the coefficient on external reliance again fails to explain tax inspection.

In columns 3 and 4, we study the identifying assumption that industry-differences in US-ranks of firm size and external reliance predict differences in the industry-ranks of the World Bank countries. Column 3 shows that a one-rank increase in the US-ladder of a given industry predicts a 0.24 rank increase in the industry-ladder of the WB surveyed country, significant at 1\%. The U.S. industry of firm-size rank also seems to positively predict external reliance in the World Bank country industry-rank, again significant at 1\%. The US-industry rank in external dependency explains positively, albeit with a small coefficient, firm-size, and even negatively external reliance; this last result however, is no longer as significant when controls are added. The joint F-test on the U.S. values in predicting cross-industry ranking in firm size and external reliance remain large, respectively at 43.34 and 10.13, implying a strong first stage.

We estimate the joint instrumental variables model (13) – (14) – (15) in column 5. The firm-size information trail is again significant at 1\% while the external reliance channel is not significant, despite a first stage F-test in excess of 10. This estimate suggests that a one standard deviation in the surveyed country size-ranking caused by exogenous industry-differences in size and external finance leads to an 8.34 percentage points industry average tax auditing probability, or a 12\% increase at a sample mean tax inspection probability of 71\%.

In Appendix A (Tables 3-4), we report the results for Tables 1-2 when controlling for age, legal status (a dummy for sole proprietorship), exporter status, and reported difficulty in accessing finance.\(^10\) There is a small loss in significance, but the results are largely unchanged.

---

\(^9\)The same methodology is used in Chetty et al. (2014)

\(^10\)We do not include the log of sales for several reasons: first, it is likely to be endogenous to our setting; second, log workers and external finance are arguably amongst the primitives which determine sales volume; third, total sales is not reported for a large set of firms which do disclose worker size and external reliance.
Chapter 2: An Empirical Test of Information Trails

Extension to sales tax compliance

As an extension, we study the impact of the information trails on sales tax compliance, both directly and through tax inspection. The industry-average measure of sales tax compliance comes with the caveats discussed in Section 3; on the other hand, it should be noted that studies of corruption and informality (Svensson 2003, Almeida and Carneiro 2012) have used the same survey formulation of the question and interpreted the values as the true measure of own firm behavior.

The results from this extension are presented in Table 3. In column 1, the simple OLS specification shows that within country industry-ranked firm-size significantly predicts sales tax compliance, while external reliance does not. Again, the non-significance on the external finance trail may arise if the true relationship is highly non-linear. Constructed similarly to Figure 6, in Figure 7 we present the graphical results of the regression in Column 1. The external reliance - tax compliance relationship seems to be well approximated by a linear OLS, alleviating concerns of mis-specification in the econometric model.

In the reduced-form specification of column 2, US-industry rank differences in firm size are found to have a significant reduced-form impact on sales tax compliance, while there is no pattern across external reliance ranks. In the double IV estimation, a one standard deviation increase in the surveyed country’s size-industry rank, caused by exogenous industry-differences in demand for size and finance, leads to a 4.22 percentage points increase in sales tax compliance, significant at 5%.

In columns 4-6, we study whether exogenous differences in firm-size and external funding reliance explain sales tax compliance through tax inspection. Column 4 simply documents a positive correlation between higher tax inspection and sales tax compliance, which is potentially affected by endogeneity. In column 5, we therefore use directly the US-industry ranks in firm-size and external reliance; that is, we estimate the IV model in which the second stage is:

\[(\text{Sales Tax Compliance})_{ic} = \theta \cdot \text{(Tax Inspection)}_{ic} + \mu_c + \varepsilon_{ic} \quad (2.16)\]

and the first stage is:

\[\text{(Tax Inspection)}_{ic} = \beta \cdot \text{(Firm Size Rank)}_{i-US,c} + \alpha \cdot \text{(External Reliance Rank)}_{i-US,c} + \psi_c + \varepsilon_{ic} \quad (2.17)\]

The estimated impact $\theta^{IV}$ is very large, and significant at 5%: increased tax inspection caused by exogenous industry-differences in labor and finance inputs causes sales tax compliance to increase by 43.5 percentage points. While the extrapolation technique almost by construction guarantees orthogonal instruments, it is harder to entirely validate the exclusion restriction in this IV-setting. There are other plausible mechanisms than tax inspection, through which firm-size and external reliance would cause sales tax compliance to increase.

Finally in column 6 we estimate the full three stage least squares model. We use the values

---

11In the firm-level survey data, there are 3330 ISIC3-matched observations which do not report an answer to the sales tax compliance question; this leads to a loss of 88 ISIC3-country observations. Average tax inspection in the two groups are not significantly different at the 5% level.
of external reliance and firm-size in the surveyed country that are predicted by variation in U.S. industry rankings (cols.3-4 of Table 2) to estimate changes in tax inspection, which in turn explains variation in sales tax compliance. Formally, while the third-stage is now (16), the second stage becomes:

\[ 1(\text{Tax Inspection})_{ic} = \omega \cdot (\text{Firm Size Rank})_{ic} + \phi \cdot (\text{External Reliance Rank})_{ic} + \eta_{ic} + \varepsilon_{ic} \]  

(2.18)

and we add two first stages (13) – (14). The estimated \( \theta^{3SLS} \) equals 43.50 and is significant at 1%.\(^{12}\) Note that \( \theta^{IV} \) and \( \theta^{3SLS} \) are very close, suggesting U.S. variation in industry-size and external finance explain tax inspection and sales tax compliance mainly through impacting the surveyed country’s variation in industry size and external reliance.

**Robustness checks**

**Number of days spent with tax inspector as tax enforcement measure**

As a continuous measure of tax enforcement, this intensive margin variable of the firm’s reported number of days spent with tax inspectors provides a robustness check to the tax inspection and compliance results. The mean number of days spent with tax inspectors is at 2.23, with a standard deviation of 2.27. The results carry over using this intensive margin variable.\(^{13}\) Full results are reported in Appendix Tables 5-6. The double-IV estimates show that from exogenous changes to technological demand for labour and finance input, increased firm-size leads to an increase of 1.18 tax inspection days, significant at 5% (Table A5, Col.5). When instrumented by the U.S. industry-ranks in firm-size and external reliance, one more day spent with tax inspectors causes sales tax compliance to increase by 2.28 percentage points, just below significance at 1%. This ‘intensive margin’ effect of increased tax inspection is almost 20 times smaller than the ‘extensive margin’ effect estimated using the binary tax inspection dummy (Table 2, Col.5), suggesting that broadening the set of industries inspected and audited may contribute significantly more to revenue collection than intensifying enforcement on a smaller set of firms.

\(^{12}\)We have not been able to find other papers which estimate a 3SLS with two first stages. Our estimate of \( \theta^{3SLS} \) is essentially an IV in the second and third stages ((18) & (16)) and a 2SLS in the first and second stage; we plug the predicted values from (13) and (14) into (20), then perform an IV using (16) and (18). Parameter estimates are consistent in each stage. If instead we estimated the whole system through a series of 2SLS, the coefficient \( \theta^{2SLS} \) reported in column (7) is equal to 43.51 and significant at 1%.

\(^{13}\)In this double instrumental variables model, the second stage is now:

\[ 1(\# \text{ Days spent with tax inspectors})_{ic} = \beta \cdot (\text{Firm Size Rank})_{ic} + \alpha \cdot (\text{External Reliance Rank})_{ic} + \vartheta_{ic} + \varepsilon_{ic} \]  

(2.19)

and the first stages remain (13) – (14)
Alternative U.S. sample to compute external dependency

It may be the case that the US industries of the 1990’s are better proxies for the current position of developing countries’ industries in their finance life-cycles. If so, then using older samples of industry-variation of U.S. firm-financing pattern may reveal a more significant impact of external reliance on tax inspection. We therefore re-compute our measure of ISIC-3 external reliance, but using instead the 1990-1999 sample of Compustat US listed firms. Results are reported in the Appendix, Tables 7-8. Both the results on the significant impact of the firm-size information trail both on tax enforcement and tax inspection and on the non-impact of external reliance hold.14

2.5 Conclusion

Due to the rising internal demand for public goods and transfers, increasing tax revenue has become a priority for developing countries. Taxing firms is particularly important, not only because firms directly collect an important share of revenue through sales and profit taxes, but also due to their role in monitoring employees and transactions in the economy. The theoretical literature (Gordon and Li 2009b, Kleven, Kreiner, and Saez 2016) has focused on firms and suggested two important information channels: the use of finance and the number of workers. As a firm grows in size and uses external sources of finance it generates trails about its operations and becomes more visible to the tax authorities. We construct a simple model of tax evasion and tax enforcement which provides the reduced-form predictions that tax enforcement and tax compliance should increase for these firms which are larger in employee-size and more externally reliant.

We test the importance of the finance and worker channels by using 80 country-surveys from the World Bank Enterprise Survey, covering 108,500 firms, with data on size, source of funding, tax inspection and compliance. Identification relies on within-country between-industry variation, which uses the U.S. distribution of firm size and external finance as an instrument for undistorted and exogenous variation in technological demand for workers and finance (a technique similar to Rajan and Zingales 1998,). We find that a larger firm-size affects significantly the probability of tax inspection both in the OLS and IV specifications: a one standard deviation exogenous increase in the industry size-ranking leads to a 8.34 percentage points higher industry tax auditing probability. External finance does not seem to have an impact on tax inspection. In a second part we study the impact of information trails on sales tax compliance, both directly and through its effect on tax inspection: in both cases we find a large effect for the size information trail, but no effect for external financial

---

14This alternative U.S-industry ranking of external finance predicts positively but insignificantly the survey-country rank. The double instrumental variables model predicting tax inspection, equations (13) – (14) – (15), still estimates a significant coefficient for firm-size but an insignificant coefficient for external reliance. The instrumental variables model predicting sales tax compliance, equations (16) – (17), retains the same signs but loses some significance. The estimated $\theta^{IV}$ comes out at 39.31, just below significance at 1%; and the $\theta^{3SLS}$ is at 38.74, significant at 1%.
dependence. Instrumented by exogenous industry-differences in labor and external finance, an increase in tax inspection leads sales tax compliance to increase by 43.51 percentage points.

The non-result regarding external funding reliance leaves open the question on why don’t tax administrations use financial information. One possibility is that the induced misallocation from levying the finance information trail is prohibitively large. Another is that political economy factors prevent tax administrations from systematically accessing this information. A final possibility is that tax authorities of developing countries have not yet made the investments in tax capacity to efficiently manage financial information.
### Table 2.1: Tax Inspection Results

<table>
<thead>
<tr>
<th></th>
<th>Reduced Forms</th>
<th>First Stages</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tax Inspection</td>
<td>Rank WB Worker</td>
<td>Rank WB ExtRel</td>
</tr>
<tr>
<td></td>
<td>OLS (1)</td>
<td>OLS (2)</td>
<td>OLS (3)</td>
</tr>
<tr>
<td>Rank WB Worker</td>
<td>.003</td>
<td>.006</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.000) ***</td>
<td>(.002) ***</td>
<td></td>
</tr>
<tr>
<td>Rank WB External Reliance</td>
<td>.000</td>
<td>.003</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.000)</td>
<td>(.006)</td>
<td></td>
</tr>
<tr>
<td>Rank US Worker</td>
<td>.002</td>
<td>.248</td>
<td>.085</td>
</tr>
<tr>
<td></td>
<td>(.000)***</td>
<td>(.033)***</td>
<td>(.021) ***</td>
</tr>
<tr>
<td>Rank US External Reliance</td>
<td>.000</td>
<td>.075</td>
<td>-.039</td>
</tr>
<tr>
<td></td>
<td>(.000)</td>
<td>(.035)**</td>
<td>(.019)**</td>
</tr>
<tr>
<td>F test joint significance</td>
<td>16.07</td>
<td>8.02</td>
<td>43.43</td>
</tr>
<tr>
<td></td>
<td>(.000)</td>
<td>(.000)</td>
<td>(.000)</td>
</tr>
<tr>
<td>Observations</td>
<td>2597</td>
<td>2597</td>
<td>2597</td>
</tr>
<tr>
<td>Clusters</td>
<td>78</td>
<td>78</td>
<td>78</td>
</tr>
</tbody>
</table>

Standard errors clustered at the country level. All regressions include country fixed effects

* p < 0.10, ** p < 0.05, *** p < 0.01

Column (5) is a double IV which uses Cols. (3)-(4) as first stages. See Section 4.2 for more details.
## Chapter 2: An Empirical Test of Information Trails

### Table 2.2: Sales Tax Compliance Results

<table>
<thead>
<tr>
<th></th>
<th>OLS (1)</th>
<th>OLS (2)</th>
<th>IV (3)</th>
<th>OLS (4)</th>
<th>IV (5)</th>
<th>3SLS (6)</th>
<th>2SLS (7)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tax Inspection</strong></td>
<td>5.65</td>
<td>43.51</td>
<td>43.51</td>
<td>43.51</td>
<td>43.52</td>
<td></td>
<td>43.52</td>
</tr>
<tr>
<td></td>
<td>(2.57)**</td>
<td>(17.37)**</td>
<td>(17.37)**</td>
<td>(14.39)***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rank WB Worker</strong></td>
<td>.202</td>
<td>.317</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.034)***</td>
<td>(.145)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rank WB Ext Rel</strong></td>
<td>-.031</td>
<td>-.032</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.027)</td>
<td>(.300)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rank US Worker</strong></td>
<td>.076</td>
<td>.025</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.026)***</td>
<td>(.028)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rank US Ext Rel</strong></td>
<td></td>
<td></td>
<td></td>
<td>.025</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(.028)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>F test joint significance</strong></td>
<td>17.28</td>
<td>4.64</td>
<td>4.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.000)</td>
<td>(.012)</td>
<td>(.012)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>1st Stage</strong></th>
<th>(3)-(4)Table1</th>
<th>(2) Table2</th>
<th>(3)-(4)-(5) Table1</th>
<th>(3)-(4)-(5)Table1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observations</strong></td>
<td>2509</td>
<td>2509</td>
<td>2509</td>
<td>2509</td>
</tr>
<tr>
<td><strong>Clusters</strong></td>
<td>77</td>
<td>77</td>
<td>77</td>
<td>77</td>
</tr>
</tbody>
</table>

Standard errors clustered at the country level. All regressions include country fixed effects.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Column (6) is a 3SLS where Cols.(3)-(4) Table 1 are first stages, and Col.5 Table 1 is the second stage. Column (7) estimates the 3SLS as a set of 2SLS. See Section 4.2 for more details.
Figure 2.1: Tax inspection and Sales Tax Compliance Across Firm Size

Figure 2.1 plots averages of sales tax compliance (X-series) and tax inspection (circle-series) across 20 quantiles of firm size. The quantiles are constructed based on the pooled sample of country-ISIC3 industry observations in the World Bank Enterprise database (2,597 observations). Firm size is measured by the reported number of permanent employees in the firm. Tax inspection is the dummy indicator for whether the firm had a tax inspection visit over the last year. Sales tax compliance is the share of sales that a firm estimates its competitors report for tax purposes. See section 3.1 for further details.

Figure 2.2: Tax Inspection and Sales Tax Compliance Across External Financial Use

Figure 2.2 plots averages of sales tax compliance (X-series) and tax inspection (circle-series) across 20 quantiles of share of external financial reliance. The quantiles are constructed based on the pooled sample of country-ISIC3 industry observations in the World Bank Enterprise database (2,597 observations). External reliance is measured as the percent share of financing which is sourced in retained earnings, family and friends and informal finance. Firms with zero external reliance represent 34% of the total number of firms. Tax inspection is the dummy indicator for whether the firm had a tax inspection visit over the last year. Sales tax compliance is the share of sales that a firm estimates its competitors report for tax purposes. See section 3.1 for details.
Figure 2.3: Distributions of Firm Size Conditional on External Financial Use

Figure 2.3 This graph superposes external financial reliance kernel density distributions where sample-restriction is based on the number of permanent employees at the firm-level. Respectively the blue, red dashed, and green lines depict the external reliance density distribution in the first, second and third tercile of the firm-size distribution. External reliance is measured as the percent share of financing which is sourced in retained earnings, family and friends and informal finance. Total firm-level sample-size: 108,538.

Figure 2.4: Distributions of External Financial Use Conditional on Firm Size

Figure 2.4 This graph superposes firm-size kernel density distributions where sample-restriction is based on external financial reliance at the firm-level. Respectively the blue, red dashed, and green lines depict the firm-size density distribution for firms in the first, second, and third terciles of the external reliance distribution. The first group has zero percent external reliance, the second external reliance strictly positive but below 34%, and the third group has external reliance greater than 34%. External reliance is measured as the percent share of financing which is sourced in retained earnings, family and friends and informal finance. Total firm-level sample-size: 108,538.
Figure 2.5 plots the rank-rank relation for firm-size in 4 selected WB surveyed countries: Burkina Faso, Argentina; Slovenia; and Germany. Each observation represents the ranking-position of a given ISIC-3 industry, in the WB country industry-rank and in the U.S. industry-rank. Blue lines are linear fits. See Section 4.2 for details.

Figure 2.6: Impact of Firm Size and External Finance on Tax Inspection

In Figure 2.6 the right and left panels are binned scatter plots of tax inspection versus respectively firm-size rank and external financial reliance rank. These plots correspond to regression Col.1 of Table 1 and use the same sample and variable definitions. To construct the left (right) plot, we first residualize the tax-inspection variable and the firm-size rank (external reliance rank) variable with respect to external reliance rank (firm-size rank) and country fixed effects. We then divide the residual firm-size rank (external reliance rank) into twenty equal-sized groups (vingtiles) and plot the means of the tax-inspection residual within each bin against the mean value of residual firm-size rank (external reliance rank). The solid line shows the best linear fit estimated on the underlying microdata using OLS and corresponds to the regression coefficient for firm-size (external reliance) in Col.1 Table 1.
In Figure 2.7 the right and left panels are binned scatter plots of sales tax compliance versus respectively firm-size rank and external financial reliance rank. These plots correspond to regression Col.1 of Table 2 and use the same sample and variable definitions. To construct the left (right) plot, we first residualize the sales tax compliance variable and the firm-size rank (external reliance rank) variable with respect to external reliance rank (firm-size rank) and country fixed effects. We then divide the residual firm-size rank (external reliance rank) into twenty equal-sized groups (vingtiles) and plot the means of the sales tax compliance residual within each bin against the mean value of residual firm-size rank (external reliance rank). The solid line shows the best linear fit estimated on the underlying microdata using OLS, and corresponds to the regression coefficient for firm-size (external reliance) in Col.1 Table 2.
2.6 Appendix A: Additional Results

- Table 3: Tax inspection results with a set of country-industry specific controls: age, legal status, exporter status, self-reported financial constraints.

- Table 4: Sales tax compliance results with a set of country-industry specific controls: age, legal status, exporter status, self-reported financial constraints.

- Table 5: Tax inspection results using self-reported number of days spent with tax inspector as outcome variable (instead of binary tax inspection variable used in main text).

- Table 6: Sales tax compliance results using self-reported number of days spent as tax inspection variable.

- Table 7: Tax inspection results using Compustat 1990-99 sample to compute US industry-averages of financial external reliance.

- Table 8: Sales tax compliance results using Compustat 1990-99 sample to compute US industry-averages of financial external reliance.

- Table 9: Tax inspection results using actual firm size and share of external reliance.

- Table 10: Sales tax compliance results using actual firm size and share of reliance.
## Chapter 2: An Empirical Test of Information Trails

### Table 2.3: Tax Inspection Results with Controls

<table>
<thead>
<tr>
<th></th>
<th>Reduced Forms</th>
<th>First Stages</th>
<th>Double IV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS (1)</td>
<td>OLS (2)</td>
<td>OLS (3)</td>
</tr>
<tr>
<td>Tax Inspection Rank WB Worker</td>
<td>.003</td>
<td>.007</td>
<td>(.003)***</td>
</tr>
<tr>
<td></td>
<td>(.000)***</td>
<td>(.004)*</td>
<td>(.000)</td>
</tr>
<tr>
<td>Rank WB External Reliance</td>
<td>.000</td>
<td>.001</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>(.000)</td>
<td>(.000)</td>
<td>(.000)</td>
</tr>
<tr>
<td>Rank US Worker</td>
<td>.001</td>
<td>.166</td>
<td>.089</td>
</tr>
<tr>
<td></td>
<td>(.000)***</td>
<td>(.028)***</td>
<td>(.022)***</td>
</tr>
<tr>
<td>Rank US External Reliance</td>
<td>.000</td>
<td>.049</td>
<td>-.037</td>
</tr>
<tr>
<td></td>
<td>(.000)</td>
<td>(.028)*</td>
<td>(.020)*</td>
</tr>
<tr>
<td>F test joint significance</td>
<td>10.34</td>
<td>4.35</td>
<td>26.54</td>
</tr>
<tr>
<td></td>
<td>(.000)</td>
<td>(.016)</td>
<td>(.000)</td>
</tr>
</tbody>
</table>

Controls included: Age, legal status, exporter status, reported financial constraints

### Table 2.4: Sales Tax Compliance Results with Controls

<table>
<thead>
<tr>
<th></th>
<th>OLS (1)</th>
<th>OLS (2)</th>
<th>IV (3)</th>
<th>OLS (4)</th>
<th>IV (5)</th>
<th>3SLS (6)</th>
<th>2*2SLS (7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tax Inspection</td>
<td>4.88</td>
<td>42.41</td>
<td>42.41</td>
<td>42.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.53)*</td>
<td>(24.67)*</td>
<td>(24.67)*</td>
<td>(19.84)***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rank WB Worker</td>
<td>.161</td>
<td>.291</td>
<td>.019</td>
<td>.056</td>
<td>.010</td>
<td>.056</td>
<td>(.025)**</td>
</tr>
<tr>
<td></td>
<td>(.042)***</td>
<td>(.221)</td>
<td>(.028)</td>
<td>(.025)</td>
<td>(.026)</td>
<td>(.025)**</td>
<td>(.026)</td>
</tr>
<tr>
<td>Rank WB Ext Rel</td>
<td>.019</td>
<td>.072</td>
<td>.028</td>
<td>.056</td>
<td>.010</td>
<td>.028</td>
<td>(.026)</td>
</tr>
<tr>
<td></td>
<td>(.028)</td>
<td>(.369)</td>
<td>(.025)</td>
<td>(.026)</td>
<td>(.026)</td>
<td>(.026)</td>
<td>(.026)</td>
</tr>
<tr>
<td>F test joint significance</td>
<td>7.52</td>
<td>2.42</td>
<td>2.61</td>
<td>2.61</td>
<td>2.61</td>
<td>2.61</td>
<td>2.61</td>
</tr>
<tr>
<td></td>
<td>(.001)</td>
<td>(.095)</td>
<td>(.080)</td>
<td>(.080)</td>
<td>(.080)</td>
<td>(.080)</td>
<td>(.080)</td>
</tr>
</tbody>
</table>

Controls included: Age, legal status, exporter status, reported financial constraints

Standard errors clustered at the country level. All regressions include country fixed effects

* p < 0.10, ** p < 0.05, *** p < 0.01
### Table 2.5: Tax Inspection Results when Tax Inspection is Number of Days

<table>
<thead>
<tr>
<th></th>
<th>Reduced Forms</th>
<th>First Stages</th>
<th>Double IV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Rank WB Worker</td>
<td>.047</td>
<td>.089</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.009)***</td>
<td>(.037)***</td>
<td></td>
</tr>
<tr>
<td>Rank WB External Reliance</td>
<td>.010</td>
<td>.135</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.009)</td>
<td>(.127)</td>
<td></td>
</tr>
<tr>
<td>Rank US Worker</td>
<td>.033</td>
<td>.248</td>
<td>.085</td>
</tr>
<tr>
<td></td>
<td>(.007)***</td>
<td>(.033)***</td>
<td>(.021)***</td>
</tr>
<tr>
<td>Rank US External Reliance</td>
<td>.001</td>
<td>.075</td>
<td>-.039</td>
</tr>
<tr>
<td></td>
<td>(.005)</td>
<td>(.035)***</td>
<td>(.019)***</td>
</tr>
<tr>
<td>F test joint significance</td>
<td>15.05</td>
<td>10.02</td>
<td>43.34</td>
</tr>
<tr>
<td></td>
<td>(.000)</td>
<td>(.016)</td>
<td>(.000)</td>
</tr>
<tr>
<td>Observations</td>
<td>2597</td>
<td>2597</td>
<td>2597</td>
</tr>
<tr>
<td>Clusters</td>
<td>78</td>
<td>78</td>
<td>78</td>
</tr>
</tbody>
</table>

Standard errors clustered at the country level. All regressions include country fixed effects.

* p < 0.10, ** p < 0.05, *** p < 0.01

### Table 2.6: Sales Tax Compliance when Tax Inspection is Number of Days

<table>
<thead>
<tr>
<th></th>
<th>Sales Tax Compliance (Percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS (1)</td>
</tr>
<tr>
<td>Tax Inspection Nr. Days</td>
<td>.093</td>
</tr>
<tr>
<td></td>
<td>(.090)</td>
</tr>
<tr>
<td>Rank WB Worker</td>
<td>.161</td>
</tr>
<tr>
<td></td>
<td>(.042)***</td>
</tr>
<tr>
<td>Rank WB Ext Rel</td>
<td>-.019</td>
</tr>
<tr>
<td></td>
<td>(.028)</td>
</tr>
<tr>
<td>Rank US Worker</td>
<td>.056</td>
</tr>
<tr>
<td></td>
<td>(.025)***</td>
</tr>
<tr>
<td>Rank US Ext Rel</td>
<td>.010</td>
</tr>
<tr>
<td></td>
<td>(.026)</td>
</tr>
<tr>
<td>F test joint significance</td>
<td>7.52</td>
</tr>
<tr>
<td></td>
<td>(.001)</td>
</tr>
</tbody>
</table>

1st Stage

<table>
<thead>
<tr>
<th>1st Stage</th>
<th>(3)-(4)Table5</th>
<th>(2)Table6</th>
<th>(3)-(4)Table5</th>
<th>(3)-(4)-(5)Table5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>2509</td>
<td>2509</td>
<td>2509</td>
<td>2509</td>
</tr>
<tr>
<td>Clusters</td>
<td>77</td>
<td>77</td>
<td>77</td>
<td>77</td>
</tr>
</tbody>
</table>

Standard errors clustered at the country level. All regressions include country fixed effects.

* p < 0.10, ** p < 0.05, *** p < 0.01
Table 2.7: Tax Inspection Results Using Alternative Sample for US External Finance

<table>
<thead>
<tr>
<th></th>
<th>Reduced Forms</th>
<th>First Stages</th>
<th>Double IV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tax Inspection</td>
<td>Rank WB Worker</td>
<td>Rank WB Ext Rel</td>
</tr>
<tr>
<td></td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Rank WB Worker</td>
<td>.003</td>
<td>.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.000)***</td>
<td>(.002)*</td>
<td></td>
</tr>
<tr>
<td>Rank WB External Reliance</td>
<td>.000</td>
<td>.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.000)</td>
<td>(.010)</td>
<td></td>
</tr>
<tr>
<td>Rank US Worker</td>
<td></td>
<td>.001</td>
<td>.274</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.000)***</td>
<td>(.031)***</td>
</tr>
<tr>
<td>Rank US Ext Rel 1990-99</td>
<td>.000</td>
<td>.132</td>
<td>-.010</td>
</tr>
<tr>
<td></td>
<td>(.000)</td>
<td>(.031)***</td>
<td>(.031)</td>
</tr>
<tr>
<td>F test joint significance</td>
<td>16.07</td>
<td>11.60</td>
<td>44.98</td>
</tr>
<tr>
<td></td>
<td>(.000)</td>
<td>(.000)</td>
<td>(.000)</td>
</tr>
<tr>
<td>Observations</td>
<td>2597</td>
<td>2597</td>
<td>2597</td>
</tr>
<tr>
<td>Clusters</td>
<td>78</td>
<td>78</td>
<td>78</td>
</tr>
</tbody>
</table>

Standard errors clustered at the country level. All regressions include country fixed effects

*p < 0.10, ** p < 0.05, *** p < 0.01

Table 2.8: Sales Tax Compliance Results Using Alternative Sample for US External Finance

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>OLS</th>
<th>IV</th>
<th>OLS</th>
<th>IV</th>
<th>3SLS</th>
<th>2*2SLS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
</tr>
<tr>
<td>Tax Inspection</td>
<td>5.65</td>
<td>39.31</td>
<td>39.31</td>
<td>39.31</td>
<td>38.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.57)**</td>
<td>(14.99)**</td>
<td>(14.99)**</td>
<td>(12.73)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rank WB Worker</td>
<td>.202</td>
<td>.119</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.034)**</td>
<td>(.177)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rank WB Ext Rel</td>
<td>-.031</td>
<td>.603</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.027)</td>
<td>(.694)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rank US Worker</td>
<td>.081</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.026)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rank US Ext Rel 1990-99</td>
<td>.011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.022)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F test joint significance</td>
<td>17.28</td>
<td>4.71</td>
<td>5.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.000)</td>
<td>(.011)</td>
<td>(.006)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1st Stage

Observations         | 2509          | 2509         | 2509      | 2509      | 2509      | 2509      |
| Clusters             | 77            | 77           | 77        | 77        | 77        | 77        |

Standard errors clustered at the country level. All regressions include country fixed effects

*p < 0.10, ** p < 0.05, *** p < 0.01
Table 2.9: Tax Inspection Results with Alternative Non-Rank Measures of Firm Size and External Finance

<table>
<thead>
<tr>
<th></th>
<th>Reduced Forms</th>
<th>First Stages</th>
<th>Double IV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tax Inspection</td>
<td>Lg Firm Size WB</td>
<td>Share Ext Rel WB</td>
</tr>
<tr>
<td></td>
<td>OLS (1)</td>
<td>OLS (2)</td>
<td>OLS (3)</td>
</tr>
<tr>
<td>Lg Firm Size WB</td>
<td>.040</td>
<td>.172</td>
<td>(.006)***</td>
</tr>
<tr>
<td>Share Ext Rel WB</td>
<td>.018</td>
<td>-.781</td>
<td>(.017)</td>
</tr>
<tr>
<td>Lg Firm Size US</td>
<td>.025</td>
<td>.303</td>
<td>(.007)***</td>
</tr>
<tr>
<td>Share Ext Rel US</td>
<td>.006</td>
<td>-.03</td>
<td>(.003)</td>
</tr>
<tr>
<td>F test joint significance</td>
<td>19.23</td>
<td>8.82</td>
<td>37.90</td>
</tr>
<tr>
<td>Observations</td>
<td>2597</td>
<td>2597</td>
<td>2597</td>
</tr>
<tr>
<td>Clusters</td>
<td>78</td>
<td>78</td>
<td>78</td>
</tr>
</tbody>
</table>

Standard errors clustered at the country level. All regressions include country fixed effects

* p < 0.10, ** p < 0.05, *** p < 0.01

Table 2.10: Sales Tax Compliance Results with Alternative Non-Rank Measures of Firm Size and External Finance

<table>
<thead>
<tr>
<th></th>
<th>OLS (1)</th>
<th>OLS (2)</th>
<th>IV (3)</th>
<th>OLS (4)</th>
<th>IV (5)</th>
<th>3SLS (6)</th>
<th>2*2SLS (7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tax Inspection</td>
<td>5.65</td>
<td>49.37</td>
<td>49.37</td>
<td>49.37</td>
<td>50.85</td>
<td>16.35***</td>
<td>16.35***</td>
</tr>
<tr>
<td>Lg Firm Size WB</td>
<td>2.415</td>
<td>(.371)***</td>
<td>.527</td>
<td>(.423)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share Ext Rel WB</td>
<td>-.367</td>
<td>-.9343</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lg Firm Size US</td>
<td>.753</td>
<td>(.483)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share Ext Rel US</td>
<td>.677</td>
<td>(.207)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F test joint significance</td>
<td>21.97</td>
<td>6.82</td>
<td>1.04</td>
<td>(.000)</td>
<td>(.001)</td>
<td>(.359)</td>
<td></td>
</tr>
<tr>
<td>1st Stage</td>
<td>(3)-(4)Table9</td>
<td>(2)Table10</td>
<td>(3)-(4)Table9</td>
<td>(3)-(4)-(5)Table9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>2509</td>
<td>2509</td>
<td>2509</td>
<td>2509</td>
<td>2509</td>
<td>2509</td>
<td></td>
</tr>
<tr>
<td>Clusters</td>
<td>77</td>
<td>77</td>
<td>77</td>
<td>77</td>
<td>77</td>
<td>77</td>
<td></td>
</tr>
</tbody>
</table>

Standard errors clustered at the country level. All regressions include country fixed effects

* p < 0.10, ** p < 0.05, *** p < 0.01
2.7 Appendix B: Data Sources

Throughout the analysis the unit of observation is a country-industry, where industries are classified using the ISIC 3 digit methodology, produced by the U.N. Statistics Division. The ISIC3 methodology classifies distinct and refined groups; per example, 'Manufacture of dairy products' is classified as ISIC152, while 'Manufacture of grain mill products, starches and starch products, and prepared animal feed' is classified as ISIC153.

The World Bank Enterprise Survey is the primary data source: this is a firm level survey collected by the World bank between 2002 and 2010 in more than a hundred countries. To our knowledge this is the only standardized firm data set spanning countries across the development spectrum. Since the survey was clearly administered by a third party not related to the government or the statistical office it could ask unique questions on tax behavior and corruption.

We also rely on two other data sources to construct our instrumental variables: the Compustat database of firms listed in the stock exchange in the US and the U.S. Bureau of Labour Statistics Quarterly Census of Employment and Wages. We mapped the U.S. industry coding, NAICS, into the ISIC classification using the appropriate U.N. tables and merged U.S. industry-averages with the industry averages from the Enterprise Survey based on common ISIC3 codes. The final sample is 2597 country-ISIC3 observations from 77 countries.

Tax inspection and tax compliance

We use the firm’s reported answer to the question: “Total days spent with officials from: tax inspectorate.” In particular we construct a tax inspection dummy equal to one if the firm has been visited for at least one day by tax auditors. The industry-average tax inspection probability is .71, with a standard deviation of .35. This is the main measure of tax enforcement that we use throughout the main text.

We also use the question on sales tax evasion: “Recognizing the difficulties many enterprises face in fully complying with taxes and regulations, what percentage of total sales would you estimate the typical establishment in your area of activity reports for tax purposes?” There are obvious difficulties in assuming that the firm’s answer to this question reflects its own industry tax evasion and even harder that it is the firm’s own evasion level.

External reliance

To construct a measure of external reliance, we strictly follow the methodology presented in Rajan and Zingales 1998 (RZ hereafter). The measure of external dependence proxies for the amount of investment that cannot be financed through internal sources, i.e. the retained earnings generated by the business. For the U.S. values of this measure, we use Compustat over the years 2000-2012, and compute external reliance as the ratio of capital expenditures minus cash flow from operations divided by capital expenditures. For each firm, we sum the numerator and denominator over time before dividing, with the intention of smoothing
temporal fluctuations. As in RZ we use the industry median, to reduce the impact of larger firms. For the decades after 1990, the reporting in Compustat changes from the original RZ sample. We follow the instructions in RZ and the conventions in the finance literature to construct the measure of external reliance.

In the World Bank sample, firm’s balance sheets give information on the shares of financing from different sources of earnings. The measure of external reliance is defined as the ratio of external sources’ share over the sum of external and internal shares: \( \text{share}_{ext_i} = \frac{\text{external}_i}{\text{external}_i + \text{internal}_i} \)

Internal finance is the sum of retained earnings, friends and family and informal finance. External finance is defined as banking finance which includes loans, overdrafts and credit card finance. An important advantage of the Enterprise Surveys is the possibility to construct the same measure of external reliance in both the World Bank and the Compustat sample. In the Enterprise Survey, the industry-average external reliance is .38, with a standard deviation of .47. Table 1 provides some summary statistics for sources of firm-financing, distinguishing between finance-sources for working capital and for new investment. In both cases, retained earnings is the dominant method of financing, with respectively 63.9 and 63.8 percent of working capital and new investment financed through the internal cash flow of the firm. Bank finance comes second for both capital and investment, at respectively 15.4 and 19.6 percent.

For some sources of financing it is ambiguous whether they generate any tax-relevant information trail (such as trade credit and leasing arrangements). In all results presented, they are excluded when constructing our measure of external reliance.

<table>
<thead>
<tr>
<th>Table 2.11: Summary of Financing Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Working Capital</td>
</tr>
<tr>
<td>mean</td>
</tr>
<tr>
<td>Retained earnings</td>
</tr>
<tr>
<td>Bank finance</td>
</tr>
<tr>
<td>Trade credit</td>
</tr>
<tr>
<td>Friends</td>
</tr>
<tr>
<td>Non bank finance</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>(N)</td>
</tr>
</tbody>
</table>

**Firm size**

We choose as number of workers the Enterprise Survey firm’s reported average number of permanent workers. Workers which are permanent employees are more likely to have access to the internal books of the firm and thus are more able than temporary employees to “blow
the whistle” and report tax under-reporting. The sample-wide distribution of firms has heavy negative skew, with a long right tail: while the median is at 17 workers, the mean of 105 is largely influenced by a few very large firms. Given the shape of the marginal density, we always use the natural log of firm size in the regressions. We retrieve data in the U.S. from the 2002 version of the Census of Employment and Wages to U.S. ISIC3-specific average firm-size.
Chapter 3

Banking on Trust: How Debit Cards Enable the Poor to Save

with Paul Gertler, Sean Higgins & Enrique Seira

3.1 Introduction

Trust is an essential element of economic transactions and an important driver of economic development (Banfield 1958; Knack and Keefer 1997; Porta et al. 1997; Narayan and Pritchett 1999; Algan and Cahuc 2010). Trust is the “subjective probability with which an agent assesses that another . . . will perform a particular action (Gambetta 1988, p. 217). It is particularly important in financial transactions where people pay money in exchange for promises, and essential where the legal institutions that enforce contracts are weak (McMil- lan and Woodruff 1999; Karlan et al. 2009). Given the nature of financial decisions, it is not surprising that trust has been shown to be key to stock market participation (Guisco, Sapienza, and Zingales 2008), use of checks instead of cash (Guisco, Sapienza, and Zingales 2004), and decisions to not withdraw deposits from financial institutions in times of financial crisis (Iyer and Puri 2012; Sapienza and Zingales 2012).

Trust in financial institutions is low, as evidenced by the fact that majorities in 40 percent of countries included in the World Values Survey report lack of confidence in banks (Figure 3.1). Trust is especially low among the poor. In Mexico, for example, 71% of those with less than primary school report low trust in banks, compared to 55% of those who completed primary school and 46% of those who completed university (Figure 3.2). Along with fees and minimum balance requirements, trust is frequently listed as a primary reason for not saving in formal bank accounts (e.g., Dupas et al. 2016). At the country level, low trust in financial institutions is strongly correlated with the proportion of the population without bank accounts (Figure 3.3). Despite its importance, trust as a potential barrier to
Chapter 3: Banking on Trust

the poor saving in financial institutions has not been extensively studied (Karlan, Ratan, and Zinman 2014).¹

Lack of trust in financial institutions may not be unfounded. Cohn, Fehr, and Marchal (2014) provide evidence that the banking industry fosters a culture of dishonesty relative to other industries. In Mexico in particular, bankers have been found to loot money by directing a large portion of bank lending to “related parties,” i.e. shareholders of the bank and their firms (La Porta, Silanes, and Zamarripa 2003). Mexican newspapers report many instances of outright bank fraud where depositors have lost their savings. For example, an extensively covered scandal involved Ficrea whose majority shareholder reportedly stole USD 200 million from savers (CNBV 2014).² It is also telling that articles with financial advice in Mexican newspapers have titles like “How to Save for Your Graduation and Avoid Frauds” and “Retirement Savings Accounts, with Minimal Risk of Fraud.” When contract enforcement is poor and fraud is rampant, trust becomes even more important (Guiso, Sapienza, and Zingales 2004; Karlan et al. 2009) and people are understandably even more reluctant to use untrustworthy financial institutions (Bohnet, Herrmann, and Zeckhauser 2010).

While trust is important, it is not an innate characteristic but rather can be influenced by experience and information (Hirschman 1984; Williamson 1993; Attanasio, Pellerano, and Reyes 2009). Debit cards (and mobile money) provide a low cost technology to monitor account balances and thereby build trust that a bank will neither explicitly steal deposits nor charge unexpectedly large hidden fees. Previous studies on debit cards and mobile money have focused on the effect of the lower transaction costs facilitated by these technologies to make purchases (Zinman 2009), access savings and remittances (Suri, Jack, and Stoker 2012; Schaner forthcoming), and transfer money (Jack, Ray, and Suri 2013; Jack and Suri 2014), but not their capacity to monitor and build trust in financial institutions. We hypothesize that new debit card clients first use the cards to check balances and thereby establish trust, after which they take advantage of the cards lower transaction costs to use the services of formal financial institutions. In this sense, we argue that building trust in a financial institution is a necessary condition for the use of formal financial services; i.e., financial inclusion requires trust. Indeed, a lack of trust could explain why a number of randomized field experiments have found that even when take-up of accessible and affordable formal savings products is high, use is low in that most opened accounts have few transactions after the first 6 to 12 months (Ashraf, Karlan, and Yin 2006; Dupas and Robinson 2013a; Karlan and Zinman 2014; Schaner 2015).

We examine this hypothesis in the context of a natural experiment in which debit cards were rolled out geographically over time to beneficiaries of the Mexican conditional cash transfer program Oportunidades. The beneficiaries had been receiving their transfers into

¹Increased trust is proposed—but not explored further—as one channel through which no-fee savings accounts led to saving in Prina (2015).

²This type of fraud is not uncommon: we scraped the online news archives of all electronic newspapers and news websites we could find in Mexico (129 total) using several keywords, then filtered the results by hand to keep only relevant stories. We found 1338 news stories associated with savings fraud in 2014 and 2015 alone.
Chapter 3: Banking on Trust

savings accounts for five years on average before debit cards were attached to their accounts, but typically did not use the accounts to save as they immediately withdraw most if not all of the transfer.\(^3\) The phased geographic rollout provides plausibility exogenous variation in assignment of debit cards to beneficiaries in a difference in difference context. For the analysis, we use high frequency administrative data on bank transactions for over 340,000 beneficiary accounts in 370 bank branches over 4 years as well as several household surveys of a sample of the same beneficiaries.

Using the high frequency administrative data, we find that beneficiaries initially used debit cards to check account balances without any increase in savings, but over time the frequency of account balance checks fell and savings rates rose. We estimate that after one year, the share of total income saved each payment period increased by 5 percentage points and that after nearly two years those with cards saved 8 percentage points more per period.

The delayed initiation of savings suggests some kind of learning. We explore three kinds of learning that may be occurring: (i) learning to trust the bank, (ii) learning to use the debit cards and ATMs, and (iii) learning that the program will not drop beneficiaries who accumulate savings. Using household survey data, we find support for the learning to trust hypothesis but not for the other two types of learning. Specifically, we find that 27 percent of beneficiaries who have had the debit card for less than 6 months report that they do not trust the bank, compared to just 17 percent of those who have had the card for more than 6 months. We find very few beneficiaries who report not knowing how to use the technology or fear the program will drop them if they accumulate savings, and no change over time comparing those that have had the debit card less than and more than 6 months. We also find that those who have had the card more than 6 months report checking their balances significantly less frequently than those who have had the card less than 6 months, consistent with our finding from administrative data that when beneficiaries first get the debit card, they check their balances often, but the frequency of checking falls over time.

We then test whether the increase in the bank account balances is an increase in total savings or a substitution from other forms of saving, both formal and informal. Using panel household survey data, we find that after one year the treatment group increases total savings by about 5 percent of income relative to the control group, which is close in magnitude to the effect we see in the administrative account data. We find no differential change in income or assets in the treatment group compared to the control. These results suggest that the increase in saving is not driven by higher income but by (voluntarily) lowering current consumption and that that the increase in bank savings does not crowd out other forms of saving (consistent with Ashraf, Karlan, and Yin 2015; Dupas and Robinson 2013a; Kast, Meier, and Pomeranz 2012).

Finally, a portion of the increase in savings is achieved through a decrease in the consumption of alcohol, tobacco, and sugar—the most frequently mentioned temptation goods

\(^3\)This is consistent with findings from other countries such as Brazil, Colombia, and South Africa, in which cash transfers are paid through bank accounts, but recipients generally withdraw the entire transfer amount each pay period and do not save in the account (Bold, Porteous, and Rotman 2012).
in Banerjee and Mullainathan (2010). Indeed, this is the only consumption category with a statistically significant decrease after receiving the card. Although the poor do save via cash at home (Collins et al. 2009), saving informally is harder as the money is hot and susceptible to temptation spending, either by the beneficiary herself or by her husband if she lacks control over his access to her savings (Ashraf 2009). Indeed, we also find that among beneficiaries living with a spouse or partner, those with lower baseline bargaining power relative to their spouse have a higher increase in savings after receiving the debit card. Our results suggest that saving in formal financial institutions may help solve some of the intra-household bargaining and self-control problems associated with trying to save informally.

These results are important for public policy as building savings in formal financial institutions has positive welfare effects for the poor and nearly half of the world’s adults do not use financial institutions (Demirg-Kunt et al. 2015). The poor have used savings products to decrease income volatility (Chamon, Liu, and Prasad 2013), accumulate money for microenterprise investments (Dupas and Robinson 2013a), invest in preventative health products and pay for unexpected health emergencies (Dupas and Robinson 2013b), and invest in childrens education (Prina 2015). Various randomized experiments have found that providing affordable and accessible savings accounts to the poor increases their future agricultural/business output and household consumption (Brune et al. 2016; Dupas and Robinson 2013a), decreases debt (Kast, Meier, and Pomeranz 2012; Atkinson et al. 2013), and improves their ability to cope with shocks (Prina 2015). For these reasons, Mullainathan and Shafir (2009) conclude that access to formal savings services may provide an important pathway out of poverty.

Given our results, government cash transfer programs could be a promising channel to increase financial inclusion and enable the poor to save, not only because of the sheer number of the poor that are served by cash transfers, but also because many governments are already embarking on digitizing their cash transfer payments through banks and mobile money. Furthermore, the technologies of debit cards and ATMs or point of sale (POS) terminals—which can be used to check balances and access savings—are simple, prevalent, and potentially scalable to millions of government cash transfer recipients worldwide.

### 3.2 Institutional Context

We examine the the rollout of debit cards to urban beneficiaries of Mexico’s conditional cash transfer program Oportunidades whose benefits were already being deposited directly into savings accounts without debit cards. Oportunidades is one of the largest and most well-known conditional cash transfer programs worldwide with a history of rigorous impact evaluation (e.g., Gertler 2004; Parker and Teruel 2005). The program provides bimonthly cash transfers to poor families in Mexico, seeking to alleviate poverty in the short term and break the intergenerational poverty cycle in the long term by requiring families to invest in the human capital of children by sending their children to school and having health check-ups.
It began in rural Mexico in 1997 under the name Progresa, and later expanded to urban areas starting in 2002. Today, nearly one-fourth of Mexican households receive benefits from Oportunidades (Levy and Schady 2013).

Oportunidades opened savings accounts in banks for a portion of beneficiaries in urban localities and began depositing the transfers directly into those accounts. The original motives for paying through bank accounts were to (i) decrease corruption as automatic payments through banks lowers both the ability of corrupt local officials to skim off benefits and of local politicians to associate themselves with the program through face-to-face contact with recipients when they receive their transfers, (ii) decrease long wait times for recipients who previously had to show up to a “payment table” on a particular day to receive their benefits, (iii) decrease robberies and assaults of program officers and recipients transporting cash on known days, and (iv) increase the financial inclusion of poor households. By the end of 2004, over one million families received their benefits directly deposited into savings accounts in Bansefi, a government bank created to increase savings and financial inclusion of underserved populations (Figure 3.4).

The Bansefi savings accounts have no minimum balance requirement or monthly fees and pay essentially no interest. Before the introduction of debit cards, beneficiaries could only access their money at Bansefi bank branches. Because there are only about 500 Bansefi branches nationwide, many beneficiaries live far from their nearest branch, meaning that accessing their accounts involved large transaction costs for many beneficiaries. Overall, the savings accounts were barely used prior to the introduction of debit cards. In 2008, the year before the rollout of debit cards, the average number of deposits per bimester was 1.05 including the deposit from Oportunidades, the average number of withdrawals was 1.02, and 98.9 percent of the transfer was taken during the first withdrawal following payment.

In 2009, the government announced that they would issue Visa debit cards to beneficiaries that were receiving their benefits directly deposited into Bansefi savings accounts. The cards enabled account holders to withdraw cash from, make deposits into, and check balances of their account at any banks ATM as well as make electronic payments at any store accepting Visa. The cards included two free ATM withdrawals every bimester at any banks ATM, after which ATM withdrawal fees averaged 13 pesos (about $1 using 2009 exchange rates) but varied by bank.

Opportunities used direct deposit into savings accounts for its beneficiaries in 275 out of Mexico’s 550 urban localities. Of these, debit cards were rolled out to approximately 100,000 beneficiaries in 143 localities in 2009 (wave 1) and to an additional 75,000 beneficiaries in 88 localities in late 2010 (wave 2). Another 170,000 beneficiaries in the remaining localities

---

4 Originally Oportunidades partnered with two banks: Bansefi, a government bank, and Bancomer, a commercial bank. However, working with a commercial bank proved to be difficult, and Oportunidades phased out the Bancomer accounts and transferred them to Bansefi by mid-2006.

5 Nominal Interest rates were between 0.09 and 0.16 percent per year compared to an inflation of around 5 percent per year during our sample period.

6 The program is paid in two-month intervals, which we refer to throughout the paper as bimesters. (The Spanish word *bimestre* is more common than its English cognate, and is used by Bansefi and Oportunidades.)
were scheduled to receive cards between November 2011 and February 2012 (control group) after the end date of our data period. The map in Figure 3.5 shows that the treatment and control waves had substantial geographical breadth and that some treatment and control localities were physically close.

The sequence with which localities switched was determined as a function of the proportion of households in the locality that were eligible for the program but were not yet receiving benefits. This is because the introduction of debit cards to existing recipients was coupled with an effort to incorporate more beneficiaries. Table 3.1 compares the means of locality-level variables and account-level variables from the control, wave 1, and wave 2 localities using data from the population census from 2005, poverty estimates from Oportunidades from 2005, Bansefi branch locations from 2008, and the administrative account data on average balances and transactions from Bansefi in 2008. Column 6 shows the p-value of an F-test of equality of means. Because the rollout was not random, it is not surprising that there are some differences across treatment and control localities: treatment localities are slightly larger and beneficiaries in these localities receive higher transfer amounts. The percent of the transfer withdrawn also differs (it is lower in wave 1 than the control and insignificantly different but with a higher point estimate in wave 2), but is high in all cases (ranging from 97.5 percent to 99.6 percent of the transfer), indicating very low savings in the account prior to receiving the card. In Sections 3.4 and 3.9, we will test and show that trends of saving, income and consumption were parallel across waves.

3.3 Data

We use a rich combination of administrative and survey data sources. To examine the effect of rollout of the debit cards on savings we use administrative data from Bansefi at the account level for 342,709 accounts at 380 Bansefi branches for a four-year period, from November 2007 to October 2011. These data include the bimonthly transfer amount the timing and amount of transactions made in the account, bimonthly average savings balances, the date the savings account was opened, and the month the card was awarded to the account holder. The average account had been opened 5.3 years before getting the card.

To test whether the delayed savings effect and increasing propensity to save over time can be explained by learning to use the technology, learning the program rules, or building trust in the bank, we use the Survey of Urban Households Sociodemographic Characteristics (ENCASDU), conducted by Oportunidades at the end of 2010. We also use the Payment Method Survey, a household survey conducted by Oportunidades in 2012 aimed at eliciting satisfaction with and use of the debit cards.

To explore whether the increased savings in the Bansefi accounts is an increase in overall savings or a substitution from other forms of saving, we use the Survey of Urban Household Characteristics (ENCERLUB), a panel survey with three pre-treatment waves in 2002, 2003, and 2004, and one post-treatment wave conducted from November 2009 to early 2010. This survey has comprehensive modules on consumption, income, and assets. We merge these data
with administrative data from Oportunidades on the transfer histories for this sample—which we use to add transfer income into total income and to identify which households are Oportunidades recipients, given the common misreporting of transfer receipt in surveys (Meyer, Mok, and Sullivan 2015)—and on the dates that debit cards were distributed in each locality.

Because the final pre-treatment wave of ENCELURB in 2004 is five years prior to wave 1 of the debit card rollout, we supplement our parallel trends test in ENCELURB with data for the intervening period (2004-2008) from the National Household Income and Expenditure Survey (ENIGH), a repeated cross-section; we merge the publicly available ENIGH with restricted-access locality identifiers provided by the National Institute of Statistics and Geography (INEGI) to determine which surveyed households were in treatment and control localities, and restrict the analysis to the poorest 20 percent of surveyed households to proxy for Oportunidades recipients.

Figure 3.4a shows the timing of the administrative Bansefi account balance and transaction data, while Figure 3.4b shows the timing of the household survey data (merged with additional administrative data) we use, both relative to the rollout of debit cards.

### 3.4 Effect of Debit Cards on Stock of Savings

Figure 3.6 presents average balances over time; even the raw data are very telling. Panel (a) compares the first wave of debit card recipients to the control group, with a dashed vertical line indicating the time when wave 1 localities received debit cards, while Panel (b) compares the second wave to the control, with a dashed vertical line indicating the time when wave 2 localities received debit cards. Strikingly, average balances increase sharply for the first wave after receiving the card, but the effect is not immediate: it begins about four bimesters after receiving the card and the larger increase happens after a year with the card. By October of 2011, wave 1 has average balances of around 2000 pesos, over three times that of the control group. Average balances also increase over time with the card in wave 2, although we have information for less bimesters after wave 2’s later switch to debit cards.

Although our data on average balances is by bimester, some payments get shifted to the end of the prior bimester, so we group adjacent bimesters into four-month periods for the remainder of the analysis. Because we have four years of data, this leaves us with 12 four-month periods. To compare the stock of savings in the treatment and control groups while controlling for individual observables and unobservables, as well as any common time shocks, we use a period-by-period difference-in-differences (DID) strategy and estimate:

\[
\text{Balance}_{it} = \lambda_i + \delta_t + \sum_{k=1}^{12} \phi_k T_{j(i)} \times \mathbb{I}(t = k) + \varepsilon_{it} \tag{3.1}
\]

where \(\text{Balance}_{it}\) is the average balance in account \(i\) over period \(t\) (specifically, end of day balances were averaged over the number of days in the bimester by Bansefi, and we average
the average balances over the two adjacent bimesters that make up the four-month period). Following other papers measuring savings (chetty14; e.g., Kast, Meier, and Pomeranz 2012; Mel, McIntosh, and Woodruff 2013; Karlan and Zinman 2014; Akbas et al. 2015; Dupas et al. 2016), we winsorize average balances to avoid results driven by outliers; our main results winsorize at the 95th percentile, and the results are robust to other cut-offs.7 The λi are account-level fixed effects which control for observable and unobservable time-invariant characteristics of the beneficiaries, δt are time-period dummies that control for general macro trends such as bimester-specific shocks that affect both treatment and control groups, Tj(i) = 1 if locality j in which account holder i lives is a treatment (i.e., wave 1 or 2) locality, and I(t = k) are time period dummies. Thus, the Tj(i) × I(t = k) terms pick up the difference in balances between treatment and control localities in each period. We estimate cluster-robust standard errors, εist, clustering by Bansefi branch. Since one time period dummy must be omitted from (3.1), we follow the standard procedure of omitting the four-month period immediately preceding the change to cards. We estimate (3.1) separately for wave 1 and wave 2.

The coefficients of interests are the φk’s, which measure the average difference in balances between the control and treatment group in bimester k. The raw data clearly suggest that pre-treatment trends of savings were parallel across control and treatment groups before getting the card; we test this statistically by testing φ1 = ··· = φℓ−1 = 0 where ℓ is the period of switch. (In wave 1, ℓ is the period November 2009-February 2010, and in wave 2 it is the period November 2010-February 2011.) Figure 3.7 plots the φk’s and shows that pretreatment coefficients are, in most periods, not individually different from zero, and we cannot reject that pre-trends are equal between treatment and control: the p-value for the F-test of φ1 = ··· = φℓ−1 = 0 is 0.823 for wave 1 and 0.110 for wave 2.

The cards also led to an increase in use of the accounts: even though clients’ average number of deposits remained constant at one, corresponding to the Oportunidades transfer, their number of withdrawals increased. Figure 3.8 plots the number of withdrawals per Oportunidades deposit per bimester: prior to receiving the debit card, both the treatment and control groups received a single deposit and made a single withdrawal, and hence the number of withdrawals per deposit was stable at very close to 1. After receiving the card, clients made on average 1.4 withdrawals per deposit, per bimester.8 Figure 3.9 shows the distribution of the number of withdrawals, before and after receiving the card. Prior to receiving the card, over 90% of clients just made a single withdrawal. After receiving the card, 65% of beneficiaries continue to make just one withdrawal, but 27% make 2 withdrawals, 6% make 3 withdrawals, and 2% make 4 or more withdrawals. This immediate increase in use of the account after a decrease in the transaction costs of accessing money agrees with the prediction of the Baumol (1952) and Tobin (1956) model of money demand in the face of transaction costs, and with empirical evidence that ATMs and debit cards lead to reduced

---

7To avoid truncating a true treatment effect or time trend, we winsorize within each wave and time period.

8The reason we measure withdrawals per deposit, rather than simply withdrawals, is that some beneficiaries receive two payments in some bimesters.
transaction costs and an increased number of withdrawals (Attanasio, Guiso, and Jappelli 2002; Alvarez and Lippi 2009).

This increased account use will also lead to a “mechanical” increase in our dependent variable, average balance, because beneficiaries will be leaving a portion of their transfer in the account for a longer period of time. For example, the 27% who make two withdrawals with the card withdraw during the first withdrawal 71% of the total amount withdrawn over the bimester (which might be less than the total deposited if they intend to save some), then return on average 9 days later to make a second withdrawal of the remaining 29% of the total they withdraw over the period. For these 9 days, 29% of the amount they withdrew over the bimester (and hence did not save) is nevertheless captured in the balance; we call the effect of this on the average balance over the period the “mechanical effect.” Furthermore, even for those who make one withdrawal of the entire transfer, the average balance will be positive if they wait some number of days after receiving the deposit before withdrawing it.\footnote{From discussions with Bansefi and Oportunidades officials, a portion of the lapse of time that the transfer remains in the account before being withdrawn can be explained by an administrative feature of the program: Oportunidades provides transfer recipients with a calendar of dates when their accounts will be credited; however, Oportunidades often transfers the deposits to the Bansefi accounts several days prior to the due date for two reasons. First, it does so in order to avoid backlogs and ensure that all accounts have been credited by the announced calendar dates. Second, it transfers funds in bulk (all on the same day for a particular locality) but staggers the calendar dates on which Oportunidades beneficiaries are told their funds will be made available, in order to avoid congestion at ATMs and Bansefi branches.}

We compute the mechanical effect for each account in each bimester using data on the amounts of and timing between deposits and withdrawals during each bimester, as described in detail in Appendix B, and subtract this from the average balance to create a variable we call “net balance.”

Figure 3.10 shows the \( \phi_k \) coefficients from (3.1), using net balance (i.e., average balance minus mechanical effect) as the dependent variable. The debit cards lead to an increase in the stock of savings, with net balances in the account tending to increase over time with the debit card. In Wave 1, there is a marked delay of about one year before beneficiaries start using the account to save. As expected, after subtracting out the mechanical effect from average balances, the treatment effect is smaller in magnitude, reaching about 900 pesos after two years with the card, compared to 1400 pesos in the average balances specification.

### 3.5 Effect of Debit Cards on Marginal Propensity to Save

To measure the propensity to save, we control for the amount received in transfers each period. This is important since there is a large amount of variation in transfers received within accounts over time, as well as between accounts. The variation within an account over time can be explained by local elections in certain localities, compliance with program
conditions, payment amounts varying depending on the time of year, and family structure.\footnote{When there is an election, Oportunidades has to give the transfer in advance, so that there is no payment close to the election month. In practice, this means that beneficiaries receive no payment in the bimester of the election and an additional payment toward the end of the preceding bimester. If a family does not comply with program conditions such as school attendance and health check-ups, the payment is suspended, but if the family returns to complying with the conditions, the missed payment is added into a future payment. As an example of payments varying by time of year, the program includes a school component that is not paid during the summer, and a school supplies component that is only paid during one bimester out of the year. Changes in family structure affect the transfer amount because one child might age into or out of the program, for example.}

In the spirit of asset accumulation models, we assume that savings in period $t$ is a function of assets in period $t - 1$, income in period $t$, (time-invariant) individual preferences, and period-specific shocks such as changes to prices:

$$Savings_{it} = f(Assets_{i,t-1}, Income_{it}, \lambda_i, \delta_i).$$

(3.2)

Linearizing $f$, separating assets into the savings balance in the Bansefi account and other assets, and separating income into Oportunidades transfer income and other income gives

$$Savings_{it} = \lambda_i + \delta_i + \beta \text{Net Balance}_{i,t-1} + \kappa \text{Other Assets}_{i,t-1}$$

$$+ \gamma \text{Transfers}_{it} + \xi \text{Other Income}_{it} + \varepsilon_{it},$$

(3.3)

where $\varepsilon_{it}$ captures period-specific idiosyncratic shocks. Our administrative data from Bansefi only include transfers and balances, but not other income and other assets; after removing these terms from (3.3), each household’s average other income and average assets over time are captured by the fixed effect $\lambda_i$, while idiosyncratic changes in these variables over time add noise in the error term. Our measure of savings at time $t$ is the difference in net balance between time $t$ and time $t - 1$; we thus have

$$Net \ Balance_{it} - Net \ Balance_{i,t-1} = \lambda_i + \delta_i + \beta \text{Net Balance}_{i,t-1} + \gamma \text{Transfers}_{it} + \varepsilon_{it},$$

(3.4)

where $\gamma$ gives the marginal propensity to save out of transfer income. Since transfers are on average about 20\% of total income in our sample, dividing our estimates by five gives a rough approximation of the marginal propensity to save out of total income. Grouping terms in (3.4) gives

$$Net \ Balance_{it} = \lambda_i + \delta_i + \theta \text{Net Balance}_{i,t-1} + \gamma \text{Transfers}_{it} + \varepsilon_{it},$$

(3.5)

where $\theta = 1 + \beta$; then to estimate the effect of receiving a debit card on the marginal propensity to save out of transfers and allow this effect to change over time, we estimate

$$Net \ Balance_{it} = \lambda_i + \delta_i + \theta \text{Net Balance}_{i,t-1} + \sum_{k=2}^{12} \alpha_k T_{j(i)}(k = t)$$

$$+ \sum_{k=2}^{12} \gamma_k \text{Transfers}_{it} \times I(k = t) + \sum_{k=2}^{12} \psi_k \text{Transfers}_{it} \times T_{j(i)} \times I(k = t)$$

(3.6)
winsorizing both net balance and transfers at the 95th percentile.

As is well-known, however, fixed effects panel data models with a lagged dependent variable (also known as dynamic panel data models) are biased and inconsistent (Nickell 1981). We thus use the system generalized method of moments (GMM) estimator proposed by Arellano and Bover (1995) and Blundell and Bond (1998), which is consistent for fixed $T$, large $N$ (as we have here) and performs well in Monte Carlo simulations, especially for large $N$ (Blundell, Bond, and Windmeijer 2001). The two-step system GMM estimator also appears to perform better than (Kiviet 1995, Kiviet 1999) and (Bruno 2005) least squares dependent variable correction methods when $N$ is large (Bun and Kiviet 2006). The effect of the debit card on the marginal propensity to save out of transfer income in bimester $k$ is $\alpha_k/\mu_k + \psi_k$, where $\mu_k$ is average transfers in bimester $k$; Figure 3.11 plots the $\alpha_k/\mu_k + \psi_k$ estimates along with their confidence intervals. Standard errors of the parameters in (3.6) are clustered at the bank branch level and corrected for finite sample bias following Windmeijer (2005); the formula for the variance of $\alpha_k/\mu_k + \psi_k$ is then approximated using the delta method. As before, we estimate (3.6) separately for wave 1 and wave 2.\footnote{Following the best reporting practices outlined in Roodman (2009b), the details of our two-step system GMM estimation are as follows. Lagged balance is used as an endogenous GMM-style instrument; because bias can increase in finite samples as $T$ increases (since this leads to more lags and, hence, more instruments: see Ziliak 1997; Roodman 2009a), to reduce the number of instruments we only use one lag of $\text{Balance}_{i,t-1}$ as an instrument. Because $\text{Transfers}_{it}$ is predetermined but not strictly exogenous, variables on the right hand side of (3.6) interacted with $\text{Transfers}_{it}$ are valid instruments in the system’s equation in levels, but not the equation in differences; as a result, we include time dummies and all interaction terms on the right-hand side of (3.6) as IV-style instruments in the system’s equation in levels, and time dummies and interaction terms excluding those interacted with $\text{Transfers}_{it}$ in the equation in differences. These specification choices result in a total instrument count of 70. Because our panel does not include gaps, we use first differencing—as in Blundell and Bond (1998)—rather than the sample-maximizing forward orthogonal deviations—as in Arellano and Bover (1995)—to eliminate fixed effects in the transformed equation to be estimated.}

In Figure 3.11, the marginal propensity to save out of the transfer is not significantly different between the treatment and control prior to receiving the card, and we observe a delayed effect after receiving the card: in wave 1, the effect remains statistically insignificant from 0 for the first three 4-month periods after receiving the card, while in wave 2 it is insignificant from 0 for the first two periods after they receive the card. The MPS then increases over time and, in wave 1 where we have more post-treatment data, increases substantially over the two years with the card. After one year with the card (in the November 2010–February 2011 period), account-holders save 26.8% of their transfer, which—using household survey data merged with administrative data from Oportunidades on bimonthly transfers to determine the proportion of total income coming from transfers—equals about 5.4% of total income. After close to two years (in the July–October 2011 period), it equals 39.1% of the transfer, or 7.9% of total income. In wave 2, the MPS increases sooner, reaching 10.8% of the transfer or 2.2% of total income after between 6 months and one year with the card.
3.6 Mechanisms

Why do we see a delayed savings effect after receiving the debit card, and why does the marginal propensity to save out of the transfer gradually increase with time? We conjecture that learning is at play and explore three kinds of learning: operational learning (i.e., learning how to use the technology), learning the program rules (specifically, that the program will not drop beneficiaries who accumulate savings), and learning to trust (that the bank is a safe place to save). The first involves knowledge of how to use the debit card and ATM, memorizing the cards PIN, etc. The second involves learning that the program will not use accumulated savings as a signal that the family is actually not poor enough to be receiving Oportunidades benefits. These first two explanations were conjectured by Oportunidades program officials when we shared our initial results from the administrative Bansefi data. The third involves learning that the risk of getting the money “stolen in the form of hidden fees, operational errors, or nefarious behavior by the bank is lower than initially believed. We find evidence that beneficiaries use the card to check their account balances, and that it thus provides them with a technology to monitor bank behavior, ensure that their money is not disappearing, and subsequently build trust in the bank.

We first use data from the ENCASDU, a survey that directly asks beneficiaries “Do you leave part of the monetary support from Oportunidades in your bank account?” and, if the response is no: “Why don’t you keep part of the monetary support from Oportunidades in your Bansefi bank account? The second question includes pre-written responses and an open-ended response. An example of an answer coded as lack of knowledge is “They didn’t explain the process for saving. An example of an answer coded as fear of being dropped from the program is “Because if I save in that account, they can drop me from the program.” An example of an answer coded as lack of trust is “Because if I don’t take out all the money I can lose what remains in the bank. The ENCASDU surveyed 8788 Oportunidades beneficiary households across rural, semi-urban, and urban areas; of these, the 1674 that received Oportunidades benefits in savings accounts tied to debit cards at the time of the survey make up our sample.

We estimate

\[ y_i = \alpha + \gamma I(\text{Card} \leq 6 \text{ months})_i + u_i, \]  

(3.7)

where three regressions are run in which the dependent variable \( y_i = 1 \) if the beneficiary reports not saving due to (i) a lack of knowledge, (ii) fear they will be dropped from the program, or (iii) lack of trust. We estimate the unconditional probability, i.e. beneficiaries who report saving are included in the regression with \( y_i = 0 \). The unconditional probability is the more relevant measure; instead using the conditional probability (only including those who save in the regression) would mean that the delayed effect we have observed of debit cards on savings could drive the result. Standard errors are clustered at the locality level. We test the null hypothesis \( \gamma = 0 \), where a rejection of the null would imply that the dependent variable we are testing—which is related to either learning to use the technology, learning program rules, or learning to trust the bank—changes over time with the card. Although this
survey is cross-sectional, we exploit the variation in time with the debit card, exogenously determined by the staggered locality-level rollout of the cards.

Figure 3.12a and Table 3.2a show the results. The first thing to note is that lack of knowledge and fear of being dropped from the program after saving are rarely cited as reasons for not saving (combined, less than 4 percent of the sample who have had the card for less than 6 months do not save for these reasons), while lack of trust is cited by 27 percent of those who have had the card for less than 6 months. Second, the proportion who report not saving due to a lack of knowledge does not change over time; in contrast, trust increases gradually with experience: beneficiaries with more than 6 months with the card are 36 percent less likely to report not saving due to low trust than those with less than 6 months with the card.

Next, we explore mechanisms behind operational learning and learning to trust the bank using the 2012 Payment Methods Survey. The survey includes a number of questions related to operational learning: “What have been the main problems you have had with the ATM?; “In general, does someone help you use the ATM?; “Do you know your PIN by heart?; “Did they tell you that with the card you have a Bansefi savings account? It also includes a question on balance checking (“In the last bimester, how many times did you check your balance?), which is a mechanism that beneficiaries could use to build trust in the bank once they have a debit card. The Payment Methods Survey included 5381 households, drawn by stratified (by payment method and locality) random sampling from all Oportunidades beneficiaries; of these, our sample is made up of the 1641 who received their benefits on debit cards tied to savings accounts.

We again use specification (3.7), with $y_{it}$ equal to: (i) the self-reported number of balance checks over the past bimester; (ii) the self-reported number of balance checks over the past bimester without withdrawing any money, constructed as the total number of balance checks minus the number of withdrawals; and dummies if the respondent reports (iii) it is hard to use the ATM; (iv) she gets help using the ATM; (v) she knows her PIN; (vi) she knows she can save in the account. Because this survey was conducted in 2012, those with the card for at least 6 months now include both wave 1 and wave 2, while beneficiaries in the localities we treat as control localities throughout this paper make up the group with cards for less than 6 months.

Figure 3.12b and Table 3.2b show the results. Both the number of balance checks and number of balance checks without withdrawing decrease over time with the card. Making trips to the ATM specifically to check the account balance (i.e., making a balance check without withdrawing any money) decreases by 36 percent after six months compared to the first 6 months (from an average of 0.53 balance checks without withdrawing to an average of 0.34), while most measures that indicate knowledge of how to use the technology do not change over time: the proportions who report it is hard to use the ATM (around 10 percent), that they get help using the ATM (55 percent), and that they know they can save in the account (32 percent) do not change, although there is a statistically significant increase in the proportion who know their PINs (from 49 to 58 percent).

Finally, we use the administrative transactions data from the 342,709 Bansefi accounts,
which include the date, time, and fee charged for each balance check at an ATM for each account, to investigate whether the mechanism that appears to be driving the increase in trust—balance checks which clients use to monitor and, over time, build trust in the bank—holds true in the administrative data; the increased power we have from a large number of observations in the administrative data allows us to take a more granular look at balance checks over time. Note that balance checks at a Bansefi branch are possible both before and after receiving the debit card. Nevertheless, if the distance to the nearest bank branch is high, the debit cards provide a technology that greatly reduces the cost of balance checking (by enabling clients to check their balances at the closest ATM of any bank). Since balance checks at a Bansefi branch are not charged a fee—unlike balance checks at ATMs in Mexico—we do not observe them in our data, which is why average balance checks (at ATMs) in the graph begin after debit card receipt.

Figure 3.13 plots the number of times people check their balance per bimester, with vertical lines indicating the timing of card receipt. Again due to the shifting of some payments to the end of the previous bimester (which might affect the bimester timing of balance checks), we continue grouping adjacent bimesters into four-month periods. We observe that the number of balance checks per bimester is initially around 2.5 checks on average in wave 1 and 1.5 checks on average in wave 2, but in both waves decreases during each four-month period after beneficiaries receive the card, consistent with the trust-building hypothesis. This is consistent with the learning to trust hypothesis: through checking her balance, the client learns that her money is still in the account and updates downward her belief about the risk of losing money. With simple Bayesian learning, balance checking has decreasing marginal benefit and she therefore checks her balance less over time.\footnote{The operational learning hypothesis makes a different prediction regarding the evolution of balance checks over time: since it becomes easier—less costly—for a beneficiary to check her balance as she learns to use the technology (e.g., by memorizing her PIN or learning how to use the ATM), if anything we might expect her to check her balance more over time if operational learning were the mechanism at play.}

3.7 Increase in Overall Savings vs. Substitution

The increase in formal Bansefi account savings might come at the expense of other types of savings that the household is already conducting, in such a way that total savings is not affected. The question of whether the observed increase in Bansefi savings crowds out other saving is relevant not only if one is concerned with total household savings, but also to understand the mechanics through which the effect on formal savings is operating and as a first step towards thinking about the broader welfare implications of providing a formal savings account with a debit card.

Does the provision of the debit card and the resulting increase in formal savings represent an increase in overall savings, or is it merely a substitution from other forms of saving? To address this question, we use Oportunidades ENCELURB panel survey, conducted in four waves during the years 2002, 2003, 2004 and November 2009 to 2010. This survey is
conducted by Oportunidades and has comprehensive modules on consumption, income, and assets for 6272 households in urban and semi-urban areas. Of the 6272 households in the post-treatment wave of ENCELURB, 2951 live in urban areas and, according to administrative data provided by Oportunidades and merged with the survey, are Oportunidades beneficiaries when interviewed in the post-treatment wave and receive their benefits in a savings account (with or without a debit card); this is the sample used in our analysis, except in the placebo tests described in Section 3.9.

As before, we use a DID strategy where we examine changes in consumption, savings, and income across beneficiaries, exploiting the differential timing of debit card receipt. Because the ENCELURB was conducted after wave 1 localities had received cards but before wave 2 or control localities had received cards, we compare those with cards (wave 1) to those who have not yet received cards (waves 2 and control), respectively referring to them as “treatment” and “control” in this section of the paper. The identification assumption is that in the absence of the debit card, treatment and control groups would have experienced similar changes in consumption, income, and assets. We formally test for parallel trends in Section 3.9, and since we indeed find that trends were parallel prior to treatment, we now test whether there was an increase in savings, which we construct as income minus consumption from the income and consumption modules of ENCELURB. We estimate

$$y_{it} = \lambda_i + \delta_t + \gamma D_{j(i)t} + \nu_{it}$$

(3.8)

separately for five dependent variables: consumption, income, savings (constructed as income minus consumption), purchase of durables, and an asset index. All variables except assets are measured in pesos per month, $i$ indexes households, and $t$ indexes survey years. As before, all variables are winsorized at the 5% level (i.e., the 95th percentile, and also the 5th percentile in the case of variables that do not have a lower bound of 0) to avoid results driven by outliers. The asset index dependent variable is constructed as the first principal component of dummy variables indicating ownership of the assets that are included in all rounds of the survey questionnaire: car, truck, motorcycle, TV, video or DVD player, radio, washer, gas stove and refrigerator. Time-invariant differences in household observables and unobservables are captured by the household fixed effect $\lambda_i$, common time shocks are captured by the time fixed effects $\delta_t$, and $D_{j(i)t} = 1$ if locality $j$ in which household $i$ lived prior to treatment has received debit cards by time $t$; i.e., in the notation used in specifications (3.1) and (3.6), $D_{j(i)t} \equiv T_{j(i)} \times I(t = 2009–10)$. We use the locality of residence prior to treatment to avoid capturing migration effects in our estimation and estimate cluster-robust standard errors at the locality level.

---

13The 2002, 2003, and 2004 waves had around 17,000 households, but due to budget constraints the number of localities was cut for the 2009-2010 wave. The consumption, income, and assets modules of Oportunidades’ analogous survey for rural areas have been used by Angelucci and De Giorgi (2009), Attanasio et al. (2013), Janvry et al. (2015), Gertler, Martinez, and Rubio-Codina (2012), and Hoddinott and Skoufias (2004), while these modules from the ENCELURB have been used by Angelucci and Attanasio (2013) and Behrman et al. (2012).
If the increase in formal savings constitutes an increase in total savings then we expect $\gamma > 0$ for total savings (defined as income minus consumption), and if we observe $\gamma = 0$ for income we expect $\gamma < 0$ for consumption. If there is no substitution of savings from assets (and if they are not using the formal savings accounts to save up for assets, at least in the short run), we expect $\gamma = 0$ for the purchase of durables (which measures a flow) and the asset index (which measures a stock). This is indeed what we find. Figure 3.15 shows that consumption decreased by about 130 pesos on average (statistically significant at the 10% level). Meanwhile, there is no effect on income; we also test the difference in the coefficients of consumption and income using a stacked regression (which is equivalent to seemingly unrelated regression when the same regressors are used in each equation, as is the case here); although both are noisily measured, the difference in the coefficients is significant at the 10% level ($p = 0.072$). Purchase of durables and the stock of assets do not change, ruling out a crowding out of these forms of saving. The increase in savings, measured as income minus consumption—which, although a crude measure of savings, is commonly used (e.g., Dynan, Skinner, and Zeldes 2004)—is estimated at slightly more than 200 pesos, and is significant at the 5% level.

The results in Figure 3.15 are from our main specification where we winsorize the dependent variable at 5 percent (specifically, at the 95th percentile, as well as the 5th percentile if the variable does not have a lower bound of 0). Appendix Table 3.6, columns 1–3 show that the effects are robust to using the raw data without winsorizing, winsorizing at 1 percent, or—as in our main specification—winsorizing at 5 percent (we follow Kast and Pomeranz (2014) and others who show the robustness of results to these three possibilities). They are also robust to including baseline characteristics interacted with time fixed effects, as well as municipality-specific time effects, both to control for specific time trends more flexibly (Appendix Table 3.6 columns 4 and 5).

These results mean that total savings—not just formal savings—increase, and that at least a substantial portion of this increase in being funded by lower consumption. A back of the envelope calculation reveals that the magnitude of the increase in monthly savings from this household survey is in line with the average increase of savings in the account from the administrative data: from the propensity to save specification, after 1 year, beneficiaries who received cards in wave 1 save 26.8% of their transfer more than the control group. Using ENCELURB, transfers are, on average, 20.2% of income for the treatment group, implying that the savings effect in the Bansefi administrative data is about 5.4% of income. The effect for savings (income minus consumption) in the ENCELURB household survey data shown

---

14The household characteristics interacted with time fixed effects in this robustness check are measured at baseline and include characteristics of the household head (whether the household head worked, a quadratic polynomial in years of schooling, and a quadratic polynomial in age), whether the household has a bank account, variables used to measure poverty by Oportunidades (the proportion of household members with health insurance, the proportion aged 15 or older that are illiterate, the proportion aged 6 to 14 that do not atted school, the proportion aged 15 or older with incomplete primary education, and the proportion aged 15 to 29 with less than 9 years of schooling), and dwelling characteristics (dirt floor, no bathroom, no water, no sewage, number of occupants per room).
in Figure 13 equates to 4.9% of income. Taken at face value, this suggests that most of the increase in savings in the account is new saving. This result is consistent with other studies where formal savings products were offered, which found that the increased savings in these products did not crowd out other forms of saving (Ashraf, Karlan, and Yin 2015; Dupas and Robinson 2013a; Kast, Meier, and Pomeranz 2012).

### 3.8 Does Money Burn a Hole in Your Pocket?

Because the accounts pay no interest, but there was clearly an unmet demand for savings among program beneficiaries, we explore why they were not able to save before (for example, under the mattress). Since the results in Figure 3.15 show that the debit card induces higher total savings through decreased consumption, we might expect that it influences different components of consumption differentially. We thus examine the proportion of income spent on several categories of consumption goods, listed in order of average budget share: meat, dairy, and produce; tortillas and cereals; health and education; transportation; fats and sweets (junk food, fats, soda); temptation goods (where we group the three most frequently cited temptation goods in Banerjee and Mullainathan (2010): alcohol, tobacco, and sugar); and entertainment. We use the proportion of total income spent on each consumption category, rather than the level of consumption in that category, because individual × time-specific shocks to income, which we expect to be passed through as shocks to various consumption categories, would otherwise add noise to the estimation through the error term; we use total income rather than total consumption in the denominator because, from the results in Figure 3.15, total income does not change differentially between the treatment and control groups.

We estimate a DID specification with household and year fixed effects and standard errors clustered at the locality level; specifically, for each consumption category $g$,

$$\frac{\text{Consumption}_{git}}{\text{Income}_{it}} = \lambda_{gi} + \delta_{gt} + \gamma_{g} D_{j(i)t} + \nu_{git},$$  

where $\text{Consumption}_{git}$ is monthly consumption of goods category $g$ by household $i$ at time $t$ (in pesos) and $\text{Income}_{it}$ is total monthly income of household $i$ at time $t$ (in pesos). We find that the only category of goods in which the treatment group shows a statistically significant decrease relative to the control is that of temptation goods (Figure 3.16).\textsuperscript{15} More specifically, the treatment group reduces the proportion of income it spends on temptation goods by 20% relative to the control group. Nevertheless, as the thick horizontal bars in Figure 3.16 show, and as we would expect, only a small portion of income is spent on temptation goods at baseline (3%). As a result, the statistically significant reduction in spending on temptation goods is significant at the 10% rather than 5% level ($p = 0.023, q = 0.086$).

\textsuperscript{15}The whiskers in Figure 3.16 show the 95% confidence intervals when no adjustment is made for multiple hypothesis testing. After adjusting for multiple hypothesis testing using the sharpened false discovery rate (Benjamini, Krieger, and Yekutieli 2006; Anderson 2008), the result for temptation goods is significant at the 10% rather than 5% level ($p = 0.023, q = 0.086$).
goods only explains 16% of the total effect of the debit cards on consumption. We refrain from attempting to determine what categories explain the remaining 84% of the effect since the results for other spending categories are not statistically significant.

Although our grouping of temptation goods is based on the goods most frequently mentioned by Banerjee and Mullainathan (2010), it could be viewed as arbitrary (and, indeed, we do not find a decrease in the grouping of fats and sweets—junk food, fats, and soda—which could also be classified as temptation goods); we thus look separately at each item in the temptation good category, and find a statistically significant decrease in consumption of alcohol and sugar, but not of tobacco.

We interpret this result as evidence that it is difficult to save informally due to self-control problems, and that these problems can be partially solved by access to a formal savings account (but that low indirect transaction costs and trust in the bank are necessary conditions for these formal savings accounts to be used). This finding is consistent with the demand for commitment savings devices (e.g., Ashraf, Karlan, and Yin 2006; Bryan, Karlan, and Nelson 2010) if the savings accounts without debit cards, which could be used as an even stronger commitment due to the high indirect cost of accessing savings, would have been too strong of a commitment (since strong commitment devices have low take-up and low use relative to weak commitment devices: see Karlan and Linden 2014; Laibson 2015), or if the bank accounts were merely not trusted prior to being able to cheaply monitor them with debit cards. Under either explanation, trust appears to have been a necessary condition for formal saving, given the delayed savings effect and self-reported reasons for not saving initially.

The self-control problems that prevent the poor from saving prior to having access to a trusted formal savings account could result directly from an asset-based poverty trap, as in Bernheim, Ray, and Yeltekin (2015), a model that is consistent with the empirical finding that microcredit decreases temptation good consumption (Angelucci, Karlan, and Zinman 2015; Augsburg et al. 2015; Banerjee et al. 2015). Alternatively, it is possible that the self-control problems stem from the timing of access to the money: Carvalho, Meier, and Wang (2016), using exogenous variation in the timing of an experiment relative to payday, find that those who are more financially constrained behave in a more present-biased way. If the beneficiary withdraws her money and attempts to save at home, she has easy access to it throughout the two month period (including access to the portion she intended to save rather than spend that period); toward the end of the period she is likely to be more financially constrained and thus behave in a more present-biased way. On the contrary, if she trusts the bank and decides to save in her Bansefi account, she makes her saving decision when initially withdrawing benefits, when she is less financially constrained due to having recently received the Oportunidades payment.

It is also possible that saving money informally is difficult because the beneficiary lacks control over her husband or partner’s access to money saved at home, and the husband has different (perhaps more present-biased) time preferences. Anderson and Baland (2002, p. 963) present evidence that participation in rotating savings and credit associations “is a strategy a wife employs to protect her savings against claims by her husband for immediate
consumption.” Consistent with differing preferences between spouses, Rubalcava, Teruel, and Thomas (2009) find that Oportunidades income (paid directly to the wife and viewed as the wife’s income) tends to be spent more on investments in the future than other income does. When spouses have differing time preferences (even if neither is present-biased), the collective decision making of the household becomes present-biased (Jackson and Yariv 2014), making soft commitment devices such as bank accounts more attractive.

We thus test whether debit cards lead to an increase in overall savings because cash saved at home is susceptible to the husband taking or requesting the money, whereas it is difficult for him to access the savings in a formal bank account. Since single beneficiaries (i.e., beneficiaries who are not living with a spouse or partner) would not be affected by this barrier, a first pass to exploring whether beneficiaries’ lack of control over their husbands’ access to the money is a barrier to saving informally is to test whether a single woman responds differently to the card than a woman who is living with her spouse or partner. Of course, a significant result would need to be interpreted with the caveat that single and married women have many other differences that might interact with the effect of a debit card on savings. We estimate

\[ \text{Savings}_{it} = \lambda_i + \delta_t + \gamma D_{j(i)t} + \xi D_{j(i)t} \times H_i + \sum_{k \in K} \zeta_k H_i \times I(t = k) + \nu_{it}, \quad (3.10) \]

where \( \text{Savings}_{it} \) is again constructed as income minus consumption and winsorized at the 5th and 9th percentiles, \( H_i \) is a time-invariant measure of heterogeneity, and \( K = \{2003, 2004, 2009–10\} \) (dropping 2002 to avoid collinearity with the household fixed effects). The \( H_i \times I(t = k) \) terms thus allow the evolution of savings over time to vary with \( H_i \) even in the absence of treatment. In this case, \( H_i = 1 \) if the beneficiary is single in the post-treatment survey wave (since marriage should not be endogenously affected by receiving the debit card and we do not want the effect to be driven by beneficiaries whose marital status changes between pre- and post-treatment).

---

16 On the other hand, debit cards might lead the husband to have higher access to the money, especially in households where the woman has low bargaining power. The high indirect transaction costs of a bank account without a card, and the requirement that the card holder herself appear at the bank branch to withdraw money, could make control over the husband’s access to the money easier without a debit card. Indeed, in Schaner (forthcoming), women with low bargaining power are hurt by receiving debit cards because they lose control over the money. As we have already seen, however, the Bansefi bank accounts were not being used to save prior to receiving the debit cards. Even so, the beneficiary might be able to hide the money at home but unable to prevent her husband from taking and using her card to withdraw money. In our survey data, however, only 4% of beneficiaries report that their spouse sometimes withdraws money from the account.

17 Although it is easy to identify whether the Oportunidades beneficiary is married and living with her spouse in ENCELURB, it is difficult to determine with certainty whether unmarried beneficiaries nevertheless live with a partner. We thus include beneficiaries living in the same household as another adult who is not the household head’s child or grandchild in our “non-single” group. Using this definition, the non-single group of beneficiaries is made up of 95% married women and 5% who are not married but living in the same household as another adult.

18 The sign and statistical significance of the point estimates on \( \xi \) and \( \gamma \) are the same, and magnitudes
Chapter 3: Banking on Trust

101

If the husband or partner’s access to money is a barrier to saving for women living with a husband (or other adult) but not for single women, we expect \( \gamma > 0, \gamma + \xi = 0, \) and \( \xi < 0 \). Table 3.3 column 1 shows that we do find \( \gamma > 0 \) and cannot reject \( \gamma + \xi = 0, \) but—although the point estimate of \( \xi \) is 168 pesos (fairly close to the average treatment effect from Table 3.6, column 3)—it is not statistically significant from 0, so we cannot reject \( \xi = 0 \).

If a lack of control over the husband’s access to money is indeed a barrier to saving, we would also expect treatment effect heterogeneity among women who do live with a husband or partner based on their bargaining power in the household. To test for this, we proxy for baseline female bargaining power using four questions asked only in the first wave of the survey on who makes the primary decisions in the household: whether to take their children to the doctor if they are sick, whether the children have to attend school, whether to buy them new clothes when needed, and “important decisions that affect the household members (transport, moving, changing jobs).” We code these questions as +1 if a woman makes the decision, 0 if spouses make them jointly, and −1 if a man makes the decision, then following Kling, Liebman, and Katz (2007), standardize the variables to each have a mean of 0 and standard deviation of 1 and average them to create a summary measure of female bargaining power. We estimate (3.10) on the subset of women living with a spouse (or other adult), with \( H_i \) as this summary measure of baseline female bargaining power. Our hypothesis that women with high bargaining power could already exercise control over money saved in the home, and thus should not have as large of a treatment effect as women with low bargaining power prior to receiving the card, would mean that \( \xi < 0 \).

The results of this test are shown in Table 3.3, column 2.\(^{19}\) Indeed, we find \( \xi < 0 \), significant at the 10% level. A one standard deviation decrease in baseline female bargaining power translates to an increase of about 198 pesos in the savings effect of the debit card, nearly as large as the average treatment effect in the full sample. This suggests that a woman with low bargaining power at baseline (and hence less control over money saved informally) receives a larger benefit from the card because it enables her to build trust in the bank and subsequently save in the account, which is out of reach of her husband. A woman with high bargaining power at baseline, on the other hand, was already able to prevent her husband from spending informal savings prior to receiving the card, and thus receives a lower benefit from the card. As a result, among women who are married or living with a partner, the savings effect caused by the debit card is higher for those with low baseline bargaining power.

Although it is common to measure bargaining power using decision-making questions similar to those used here (e.g., Ashraf, Karlan, and Yin 2010; Antman 2014), bargaining power can alternatively be measured based on differences in the husband and wife’s age,

---

\(^{19}\) Of the 2951 households in our sample, the Oportunidades beneficiary lives with a spouse or partner in 2098 (71%). Of these 2098, only 1625 are included in the regression for column 2; the difference of 473 households is because 93 were not included in the 2002 wave of the survey (the only wave to ask these bargaining power questions), while 380 were included but refused to answer one or more of the bargaining power questions.
education, literacy, and income. We now test the robustness of our result to measuring bargaining power using these variables, following the method used by Schaner (forthcoming). Specifically, we construct a Kling, Liebman, and Katz (2007) summary measure of proxied bargaining power based on differences between spouses in these four variables at baseline. We standardize age, years of schooling, a dummy measuring literacy, and the total income earned by the beneficiary or her partner to each have a mean of 0 and standard deviation of 1. We then subtract the value of each variable for the husband or partner from the value for the beneficiary to create an indicator that is increasing in the woman’s relative “power” in that dimension, then average across the four indicators. We again standardize the resulting bargaining power proxy so that the regression coefficient can be interpreted as the effect of a one standard deviation in female bargaining power on savings. The sample is again restricted to women who are married or living with a partner.\textsuperscript{20} The coefficient under this alternative specification, shown in Table 3.3, column 3, is slightly lower than in the first specification (−163 pesos compared to −198) and is not statistically significant. Because the significance of our prior result is marginal and not robust to alternative measures of bargaining power (although, reassuringly, point estimates retain their sign and remain similar in magnitude), and because bargaining power could be correlated with unobservables that explain the heterogeneous treatment effect, we interpret the results based on differences in baseline bargaining power as merely suggestive.

Another potential barrier to saving informally is that money saved at home could be in demand from friends and relatives. It is obvious from Baland, Guirkinger, and Mali (2011) and Jakiela and Ozier (2016) that the desire to conceal money in a savings account to avoid demands from others extends beyond one’s spouse to friends and relatives. Ideally, we would test whether transfers from the household to other households decreased after receiving the card; the question on transfers of money to other households was not included in the post-wave survey, however. We thus estimate (3.10) with \( H_i \) as a dummy variable equal to 1 if the household reported giving money to other households at baseline (specifically, in any of the three pre-treatment waves). Because those with higher demands for money from friends and relatives are more likely to have \( H_i = 1 \), if this is a barrier to saving informally we expect \( \xi > 0 \). The results of this test, shown in Table 3.3, column 4, are inconclusive: although the point estimate on the interaction term is large, at 354 pesos, the standard error is very large, and the effect is statistically insignificant from 0. It is worth noting that only 7\% of the sample has \( H_i = 1 \). This suggests that demands for money from relatives and friends might be a barrier to saving informally, but—if so—that this barrier only affects a small fraction of Oportunidades recipients.

\textsuperscript{20}The sample includes 141 households less than in the previous specification due to households with missing values for years of schooling of the beneficiary or her partner.
3.9 Robustness

Internal Validity Checks

The identifying assumption for (3.1) and (3.6) is that the beneficiaries that received the debit card in waves 1 and 2 would have had the same trend in average balances and marginal propensity to save as the control group in the absence of treatment. While the assumption is inherently untestable, its plausibility was confirmed by two sets of results presented sections 3.2 and 3.4. First, although the rollout was not random, most means between treatment and control do not have statistically significant differences; there is a difference, however, in population, transfer amount, and percent of the transfer withdrawn. (For percent of the transfer withdrawn, the F-test of equality between the three means is rejected, and a test of equality of wave 1 and the control is rejected, but the test of equality between wave 2 and the control is not rejected.) More important, average balances follow parallel pre-treatment trends in wave 1 and the control prior to wave 1 receiving debit cards, and in wave 2 and the control prior to wave 2 receiving debit cards: this can be seen visually in Figure 3.6 and is formally tested in Section 3.4. The similarity of savings in the treatment and control groups before treatment contrasts sharply with the diverging trends after debit cards are received. The fact that results comparing the control to two waves receiving debit cards in different years are similar suggests this is not an artefact of a shock in a particular month or year.

Similarly, the identifying assumption for the household survey panel data results on savings, income, consumption, purchase of durables, and ownership of assets in (3.8) is that these variables would have followed parallel trends in the absence of treatment. Figure 3.14 shows these parallel trends graphically for the pre-treatment rounds of the survey. In addition, because there are many years between the last pre-treatment ENCELURB survey year (2004) and the year of treatment (2009), we supplement the ENCELURB parallel trends tests with tests using data from the 2004–2008 rounds of the ENIGH, a national income and expenditure survey used for Mexico’s official poverty measurement. This is a repeated cross-section survey conducted in even years (but additionally conducted in 2005) that sampled between 20,000 and 30,000 households during each year in this time frame.

Although the publicly available version of the survey does not include each household’s locality code, which determines whether the household lives in a treatment or control locality, we obtained the locality codes for sampled households from Mexico’s National Institute of Statistics and Geography. Although Oportunidades receipt is reported in the survey, there is a large discrepancy between the number of beneficiaries according to the survey (after expansion factors are applied) and the number in national accounts (Scott 2014), a problem also common in developing countries (Meyer, Mok, and Sullivan 2015), so to have sufficient power for our test we restrict the analysis to the poorest 20 percent of surveyed households to proxy for Oportunidades recipients, rather than use self-reported Oportunidades receipt. Again, the parallel trends can be clearly seen visually in Figure 3.14.
Alternative Explanations and Placebo Tests

We have argued that the card allows beneficiaries to build trust in the bank by monitoring the banks activity through balance checks. We now explore alternative explanations for the observed delayed savings effect and increasing marginal propensity to save over time. First, it could be that accumulating time with the savings account, rather than with the card, drives the increase over time. Second, while the hypothesis that debit cards increased trust through bank monitoring is demand-driven, the effect could be supply-driven if banks optimally responded to the increased debit card concentration by opening up more ATMs or bank branches in those localities; if such an expansion were gradual, it could explain the delayed savings effect and increasing marginal propensity to save over time. Third, the effect might be driven by locality-specific shocks unrelated to the debit cards. Fourth, the debit cards could merely make savings more salient, as in Akbas et al. (2015), by giving beneficiaries a reminder (in the form of an object carried with them) of their savings intentions.

There are a number of reasons that it is unlikely that the effects are driven by experience with the savings account leading beneficiaries to learn the benefits of saving, rather than time with the debit card itself. First, both treatment and control accounts are accumulating time with their savings accounts simultaneously. Second, because the savings accounts were mainly rolled out between 2002 and 2004 (Figure 3.4), most beneficiaries had already accumulated several years with the account by 2009, when treatment begins. Indeed, the median date of account opening in our 342,709 accounts is October 18, 2004, and less than 5 percent of accounts had existed for less than two years when they received debit cards. Third, our results from Section 3.5 include account fixed effects, so any time-invariant effect of having the account for a longer period of time would be absorbed. Fourth, to test for a time-varying effect of having the account for a longer period of time, we test whether results vary when we run the analysis separately for two groups: those who have had the account for more vs. less time. We use the median date of account opening to split the accounts into these two groups, and find that results are very similar. Appendix Figure 3.18 shows the equivalent of Figure 3.11 separately for older accounts (panels (a) and (b)), opened before the median date of October 18, 2004, and younger accounts (panels (c) and (d)) opened on or after that date.

A second possible explanation for the increase in savings over time is that banks gradually expanded complementary infrastructure in localities where treated beneficiaries live. Depending on the costs of each branch and ATM machine, this could be a profit-maximizing response to the increase in the number of debit card holders in treated localities. The increasing marginal propensity to save over time could be the result of the staggered expansion of this infrastructure, not increased trust. If this is so, then the increase in savings would have to be reinterpreted not only as the effect of debit cards but of the expansion of the whole enabling technology. Using quarterly data for each municipality on the number of bank branches and ATMs for Bansefi and all other banks, we test if there was indeed a contemporaneous expansion of infrastructure and if this was correlated geographically with Oportunidades debit card expansion or with savings in our accounts.
We first test for a relationship between the rollout of ATM cards and a supply-side expansion of banking infrastructure (ATMs and bank branches)\(^{21}\) by estimating:

\[
y_{mt} = \lambda_m + \delta_t + \sum_{k=-4}^{4} \beta_k D_{m,t+k} + \varepsilon_{jt},
\]

where \(y_{mt}\) is the number of total ATMs, total bank branches, Bansefi ATMs, or Bansefi branches in municipality \(m\) in quarter \(t\) and \(D_{mt}\) equals one if at least one locality in municipality \(m\) has Oportunidades debit cards in quarter \(t\). We include one year (four quarters) of lags and one year of leads to test for a relationship between bank the debit card rollout and bank infrastructure. For this test, we use data on the number of ATMs and bank branches by bank by municipality by quarter from the Comisin Nacional Bancaria y de Valores (CNBV), from the last quarter of 2008 through the last quarter of 2013 (since the rollout was from 2009 to 2012, when what we refer to as control group localities received debit cards). We separately test whether lags of credit card receipt predict banking infrastructure (i.e., whether there is a supply-side response to the rollout of debit cards) by testing \(\beta_{-4} = \cdots = \beta_{-1} = 0\) and whether leads of credit card receipt predict banking infrastructure (i.e., whether debit cards were first rolled out in municipalities with a recent expansion of banking infrastructure) by testing \(\beta_1 = \cdots = \beta_4 = 0\). We find evidence of neither relationship, failing to reject the null hypothesis of each test for each of the four dependent variables (Table 3.4).

To rule out locality-specific shocks that could be driving the savings effect, as opposed to the effect being driven by the debit cards, we perform a placebo test using poor non-Oportunidades households in the treated vs. control localities in the ENCELURB data. The ENCELURB initially included households deemed potentially eligible for the Oportunidades program as it was expanded to urban areas; some households did not become beneficiaries (either they were deemed ineligible or did not take up the program). Because these non-beneficiaries were “potentially eligible” for the program to be included in the survey, they are similarly (though not quite as) poor compared to the Oportunidades beneficiaries who make up our main sample. Because they did not receive debit cards during the rollout, due to not being Oportunidades beneficiaries, these individuals in treatment and control localities serve as a good placebo test for locality-level shocks. The results are presented in Figure 3.17a. The DID estimates on consumption, income, and savings are all insignificant from 0, although due to the low number of non-Oportunidades beneficiaries in ENCELURB (382 households), the estimates are very noisy. Nevertheless, it is comforting that the point estimates are substantially close to 0 relative to the coefficients from our main sample, and the coefficients for consumption and savings actually have the opposite sign as the coefficients from the main regression (shown again in panel (c) for comparison). This suggests that, although the noisy placebo estimates’ 95% confidence intervals do include the point estimates from our main sample, locality-level shocks do not explain the observed results.

\(^{21}\)We do not test an expansion of point of service (POS) payment terminals because the data on POS terminals by municipality does not begin until 2011, toward the end of our study period.
Finally, we test for a salience effect of the cards themselves, where the card—which a beneficiary might carry with her in a wallet or purse—serves as a salient reminder of her savings goals. In some localities, beneficiaries received their benefits through Bansefi but did not have access to a Bansefi savings account (and thus had to withdraw all of their money each pay period at a Bansefi branch); in these localities, the government decree requiring all beneficiaries to receive benefits through a plastic card led to receiving benefits on a pre-paid card, still without access to a savings account. Again using ENCELURB, we find that in localities without savings accounts that switched to a pre-paid card prior to the last round of the survey compared to localities without savings accounts that did not switch prior, there was no differential effect on consumption, income, or savings. These estimates are again noisier than the results from the main sample (here we have 2300 households), but the DID coefficient from the placebo consumption regression is statistically significant at the 10% level from the coefficient from the corresponding consumption regression in the full sample.

**Alternative Barriers to Saving Informally**

An alternative potential barrier to saving informally is that the money risks being stolen if saved at home. An anticipated reduction in crime was one of the primary motivations for the change to debit cards; in the U.S., changing the payment method of cash welfare payments to debit cards caused a significant decline in burglary, assault, and larceny (Wright et al. 2014). In developing countries, risk of theft has been anecdotally reported as a reason for not saving at home by cash transfer recipients in the Dominican Republic (Center for Effective Global Action 2015), and is pointed out as a potential mechanism in Malawi by Brune et al. (2016).

To test this hypothesis, we test whether high-crime municipalities—where saving informally would be more difficult due to risk of theft—have a higher treatment effect. Specifically, we use three measures of crime at the municipal level: crimes recorded, thefts recorded, and homicides. Crimes and thefts recorded are based on reported crimes and thus suffer from reporting bias; nevertheless, total thefts is the variable that likely best represents the risks people face if saving informally. These variables come from Mexico’s Executive Secretariat of the National Public Safety System. Homicides come from national vitality records and are thus less susceptible to bias. All crime rates are for the 2008 calendar year, immediately prior to the rollout of debit cards. For each of these three measures of crime, we estimate (3.10) where $H_i$ (or, more precisely, $H_{m(i)}$) is either the crime rate per 100,000 in the municipality $m$ in which household $i$ lives (Table 3.5, columns 1, 3, and 5) or a dummy variable equal to 1 if the crime rate is greater than the median, with the median calculated based on the municipal crime rates faced by each beneficiary household in our sample (columns 2, 4, and 6). If risk of theft is a barrier to saving, we expect $\xi > 0$, i.e., there was a higher savings effect from the debit cards in localities with higher crime and thus greater risk that informal savings are stolen. In all specifications, the point estimate on the interaction term is statistically insignificant from 0; in many, it is close to 0; and in the majority, the sign
is negative, contrary to our hypothesis. This suggests that risk of theft is not a barrier to saving informally in our context.

### 3.10 Conclusion

Although trust in financial institutions is by no means a sufficient condition to enable the poor to save, our findings suggest that it is a necessary condition. A lack of trust in banks could explain why a number of studies have found modest effects of offering savings accounts to the poor, even when these accounts have no fees or minimum balance requirements. Debit cards, a simple technology with high scale-up potential, provided beneficiaries of Mexico’s large-scale cash transfer program Oportunidades with a mechanism to monitor banks by checking their balances at any bank’s ATM; once beneficiaries built trust in banks, they began to save and their marginal propensity to save increased over time. We find that the observed increase in formal savings represents an increase in overall savings rather than a substitution from other forms of saving, and that beneficiaries reduce consumption of temptation goods, suggesting that saving informally is difficult and the use of financial institutions to save helps solve self-control problems.

The size of the savings effect, at 5% of income after one year with the debit card and 10% after two years, is larger than that of studies on various savings interventions such as subsidizing bank fees, increasing interest rates, and providing commitment savings devices. As a result, interventions that enable account holders to monitor banks and increase their trust in financial institutions may be a promising avenue to enable the poor to save in the formal financial sector. Debit cards and other forms of mobile money, which are simple, scalable technologies that are gaining traction in many developing countries, could thus be a highly effective means of increasing financial inclusion among millions of government cash transfer recipients worldwide.
Figure 3.1: Low Trust in Banks Around the World

Notes: \(N = 82,587\) individuals in 60 countries. Low trust in banks is defined as “not very much confidence” or “none at all” for the item “banks” in response to the following question: “I am going to name a number of organizations. For each one, could you tell me how much confidence you have in them: is it a great deal of confidence, quite a lot of confidence, not very much confidence or none at all?” Countries are divided into quintiles, with quintile cut-offs rounded to the nearest percentage point in the legend. Darker shades indicate countries with a higher percent of the population reporting low trust in banks.

Figure 3.2: Low Trust in Banks by Education Level in Mexico

Notes: \(N = 1993\) individuals. Low trust in banks is defined as “not very much confidence” or “none at all” for the item “banks” in response to the following question: “I am going to name a number of organizations. For each one, could you tell me how much confidence you have in them: is it a great deal of confidence, quite a lot of confidence, not very much confidence or none at all?”
Chapter 3: Banking on Trust

Figure 3.3: Cross-Country Comparison of Trust in Banks and Saving in Financial Institutions

Sources: World Values Survey (WVS), Wave 6 (2010–2014); Global Findex; World Development Indicators (WDI).
Notes: $N = 56$ countries. The y-axis plots residuals from a regression of the proportion that save in financial institutions (from Global Findex) against controls (average age, education, and perceived income decile from WVS, GDP per capita and growth of GDP per capita from WDI). The x-axis plots residuals from a regression against the same controls of the proportion that respond “a great deal of confidence” or “quite a lot of confidence” in response to the WVS question “I am going to name a number of organizations. For each one, could you tell me how much confidence you have in them: is it a great deal of confidence, quite a lot of confidence, not very much confidence or none at all?” The solid line shows a kernel-weighted local polynomial regression, while the gray area shows its 95% confidence interval.
Figure 3.4: Timing of Roll-out and Data

(a) Administrative Bank Account Data

(b) Household Survey Data

Source: Number of Oportunidades bank accounts with cards and without cards by bimester is from administrative data provided by Oportunidades.
Chapter 3: Banking on Trust

Figure 3.5: Geographic Coverage and Expansion of Debit Cards

Sources: Administrative data from Oportunidades on timing of debit card receipt by locality and shape files from INEGI.
Notes: \( N = 275 \) localities (44 in control, 143 in wave 1, 88 in wave 2). The area of each urban locality included in the study is shaded according to its wave of treatment. Urban localities that were not included in the Oportunidades program at baseline or were included in the program but did not pay beneficiaries through Bansefi savings accounts are not included in the figure or in our study.

Figure 3.6: Evolution of Average Balances

Sources: Administrative data from Bansefi on average account balances by bimester and timing of card receipt.
Notes: \( N = 5,834,468 \) account-bimester observations from 343,204 accounts. Average balances are winsorized at the 95th percentile.
Chapter 3: Banking on Trust

Figure 3.7: Difference between Treatment and Control in Average Balances

(a) Wave 1 vs. Control
(b) Wave 2 vs. Control

Sources: Administrative data from Bansefi on average account balances by bimester and timing of card receipt.
Notes: (a) \( N = 2,023,862 \) from 171,441 accounts. (b) \( N = 3,086,749 \) from 270,046 accounts. The figure plots \( \phi_k \) from (3.1).

Figure 3.8: Withdrawal to Deposit Ratio per Bimester

Sources: Administrative data from Bansefi on transactions by quarter and timing of card receipt.
Notes: \( N = 14,594,799 \) transactions from 343,204 accounts.
Chapter 3: Banking on Trust

Figure 3.9: Distribution of Withdrawals

(a) Wave 1 vs. Control
(b) Wave 2 vs. Control

Sources: Administrative data from Bansefi on transactions by bimester and timing of card receipt.
Notes: $N = 14,594,799$ transactions from 343,204 accounts.
Figure 3.10: Difference between Treatment and Control in Net Balances

(a) Wave 1 vs. Control
(b) Wave 2 vs. Control

Sources: Administrative data from Bansefi on average account balances by bimester, timing and amount of transfer payments, timing and amount of withdrawals, and timing of card receipt.

Notes: (a) $N = 2,023,862$ from 171,441 accounts. (b) $N = 3,086,749$ from 270,046 accounts. Net balances refer to average balances minus the mechanical effect on average balance of leaving a portion of the deposit in the account for a certain number of days before withdrawing it. The figure plots $\phi_k$ from (3.1).
Figure 3.11: Difference between Treatment and Control in Marginal Propensity to Save Out of Transfer

Sources: Administrative data from Bansefi on average account balances by bimester, timing and amount of transfer payments, timing and amount of withdrawals, and timing of card receipt.

Notes: (a) $N = 1,852,416$ from 171,441 accounts. (b) $N = 2,816,671$ from 270,046 accounts. (Total number of observations does not include the $t = 1$ observations, which are not included in the regressions but are used to generate $y_{ij,t-1}$ for $t = 2$ observations.) The figure plots $\alpha_k / \mu_k + \psi_k$ from (3.6) estimated by Blundell and Bond (1998) two-step system GMM, where $\mu_k$ is average transfers in period $k$. Average balances and transfer amounts are winsorized at the 95th percentile within the treatment and control groups and within each time period. The variance of $\frac{\alpha_k}{\mu_k} + \psi_k$ is estimated using the delta method.
Figure 3.12: Trust and Knowledge Over Time with the ATM Card

(a) ENCASDU (2010)

(b) Payment Methods Survey (2012)

Sources: ENCASDU 2010 and Payment Methods Survey 2012.
Notes: (a) $N = 1674$. (b) $N = 1617$, or less in some regressions if there were respondents who reported “don’t know” or refused to respond (see Table 3.2 for number of observations in each regression). Balance checks are measured over the past bimester. Bars for “debit card < 6 months” are colored light blue in (a) because at the time of ENCASDU 2010, those with the card 6 months or less were in wave 2 localities; bars for “debit card < 6 months” are colored orange in (b) because at the time of Payment Methods Survey 2012, those with the card 6 months or less were in control localities.
Figure 3.13: Balance Checks (Administrative Data)

Source: Administrative transactions data from Bansefi.
Notes: Number of balance checks per account tied to a debit card. Prior to receiving the card it was possible to check balances at Bansefi branches only, and balance checks at Bansefi branches are not recorded in our transactions data because they are free of charge.

Figure 3.14: Parallel Pre-Treatment Trends in Household Survey Data

(a) ENCELURB: Pre-Treatment Consumption (pesos/week)
(b) ENIGH: Pre-Treatment Consumption (pesos/week)
Figure 3.15: Effect of the debit card on consumption, income, total savings, purchase of durables, and assets

Sources: ENCELURB panel survey combined with administrative data on timing of card receipt and transfer payment histories for each surveyed beneficiary household.
Notes: \( N = 11,275 \) (number of households = 2951).
Figure 3.16: Effect of the debit card on different categories of consumption

- Meat, dairy, produce
- Tortillas and cereals
- Health and education
- Transportation
- Junk food, fats, soda
- Alcohol, tobacco, sugar
- Entertainment

Percent change in proportion of income spent on...

Sources: ENCELURB panel survey combined with administrative data on timing of card receipt and transfer payment histories for each surveyed beneficiary household.

Notes: \( N = 11,275 \) (number of households = 2951). Each plotted coefficient is from a separate regression using (3.9), and shows the percent change in the proportion of income spent on that category of consumption. In other words, the graph plots \( \gamma_g/\mu_g \), where \( \mu_g \) is the mean proportion of income spent on consumption category \( g \) by the control group at baseline. Categories are sorted in descending order of the percent of income spent on each consumption category at baseline, i.e. \( 100\mu_g \), which is shown by the thick horizontal bars. The whiskers show 95% confidence intervals with no adjustment for multiple hypothesis testing. After adjusting for multiple hypothesis testing using the sharpened false discovery rate (Benjamini, Krieger, and Yekutieli 2006; Anderson 2008), the result for the “alcohol, tobacco, and sugar” category is significant at the 10% rather than 5% level (\( p = 0.023, q = 0.086 \)).
Figure 3.17: Placebo Tests

(a) Placebo 1: Poor Non-Beneficiaries

(b) Placebo 2: Pre-Paid Cards

(c) Original Estimates (For Comparison)

Sources: ENCELURB panel survey combined with administrative data on timing of card receipt and transfer payment histories for each surveyed beneficiary household.

Notes: (a) $N = 1415$ (number of households = 382); (b) $N = 8862$ (number of households = 2300); (c) $N = 11,275$ (number of households = 2951).
### Chapter 3: Banking on Trust

#### Table 3.1: Comparison of Baseline Means

<table>
<thead>
<tr>
<th>Variable</th>
<th>Control (Wave 2)</th>
<th>Wave 1</th>
<th>wave 2</th>
<th>Diff. W1–C</th>
<th>Diff. W2–C</th>
<th>F-test p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Locality-level data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log population</td>
<td>10.57</td>
<td>11.18</td>
<td>11.48</td>
<td>0.60***</td>
<td>0.91***</td>
<td>0.00***</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.10)</td>
<td>(0.16)</td>
<td>(0.14)</td>
<td>(0.19)</td>
<td></td>
</tr>
<tr>
<td>Bansefi branches per 100,000</td>
<td>1.27</td>
<td>1.23</td>
<td>1.58</td>
<td>−0.03</td>
<td>0.32</td>
<td>0.411</td>
</tr>
<tr>
<td></td>
<td>(0.28)</td>
<td>(0.13)</td>
<td>(0.23)</td>
<td>(0.30)</td>
<td>(0.36)</td>
<td></td>
</tr>
<tr>
<td>% HHs in poverty</td>
<td>15.93</td>
<td>13.20</td>
<td>12.23</td>
<td>−2.73</td>
<td>−3.71*</td>
<td>0.177</td>
</tr>
<tr>
<td></td>
<td>(1.67)</td>
<td>(0.75)</td>
<td>(1.09)</td>
<td>(1.82)</td>
<td>(1.99)</td>
<td></td>
</tr>
<tr>
<td>Occupants per room</td>
<td>1.18</td>
<td>1.11</td>
<td>1.12</td>
<td>−0.07</td>
<td>−0.06</td>
<td>0.260</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td></td>
</tr>
<tr>
<td>Number of localities</td>
<td>44</td>
<td>143</td>
<td>88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel B: Administrative bank account data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average balance</td>
<td>581.25</td>
<td>670.32</td>
<td>614.29</td>
<td>89.07</td>
<td>33.05</td>
<td>0.112</td>
</tr>
<tr>
<td></td>
<td>(12.46)</td>
<td>(56.24)</td>
<td>(21.26)</td>
<td>(55.33)</td>
<td>(23.95)</td>
<td></td>
</tr>
<tr>
<td>Number of deposits</td>
<td>1.06</td>
<td>1.05</td>
<td>1.06</td>
<td>−0.02</td>
<td>−0.01</td>
<td>0.907</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.04)</td>
<td>(0.03)</td>
<td>(0.04)</td>
<td>(0.03)</td>
<td></td>
</tr>
<tr>
<td>Size of transfer</td>
<td>1506.55</td>
<td>1809.50</td>
<td>1761.26</td>
<td>302.96***</td>
<td>254.71***</td>
<td>0.000***</td>
</tr>
<tr>
<td></td>
<td>(12.73)</td>
<td>(20.16)</td>
<td>(17.47)</td>
<td>(23.67)</td>
<td>(21.15)</td>
<td></td>
</tr>
<tr>
<td>Number of withdrawals</td>
<td>1.03</td>
<td>1.01</td>
<td>1.02</td>
<td>−0.01</td>
<td>−0.01</td>
<td>0.757</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.02)</td>
<td></td>
</tr>
<tr>
<td>Percent withdrawn</td>
<td>98.56</td>
<td>97.50</td>
<td>99.64</td>
<td>−1.06**</td>
<td>1.08</td>
<td>0.021**</td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(0.45)</td>
<td>(0.71)</td>
<td>(0.46)</td>
<td>(0.72)</td>
<td></td>
</tr>
<tr>
<td>Years with account</td>
<td>5.31</td>
<td>5.49</td>
<td>5.21</td>
<td>0.17</td>
<td>−0.10</td>
<td>0.510</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.15)</td>
<td>(0.25)</td>
<td>(0.17)</td>
<td>(0.26)</td>
<td></td>
</tr>
<tr>
<td>Number of accounts</td>
<td>97,922</td>
<td>73,070</td>
<td>171,717</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sources: Census (2005), Bansefi branch locations (2008), poverty estimates from Oportunidades (based on 2005 Census), timing of card receipt by locality from Oportunidades, and administrative data from Bansefi.

Notes: W1 = wave 1, W2 = wave 2, C = control, Diff. = difference. For the administrative data from Bansefi, baseline is defined as January 2009 to October 2009 (prior to any accounts receiving cards in the data from Bansefi).
Table 3.2: Trust and Knowledge Over Time with the ATM Card

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Has card ≤ 6 months</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: ENCASDU Survey (2010): Doesn’t save in Bansefi due to . . .</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of knowledge</td>
<td>0.017***</td>
<td>−0.003</td>
<td>1,674</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.010)</td>
<td></td>
</tr>
<tr>
<td>Fear of ineligibility</td>
<td>0.019***</td>
<td>0.015</td>
<td>1,674</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.015)</td>
<td></td>
</tr>
<tr>
<td>Lack of trust</td>
<td>0.175***</td>
<td>0.095**</td>
<td>1,674</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.044)</td>
<td></td>
</tr>
</tbody>
</table>

| **Panel B: Payment Methods Survey (2012)** |      |                     |     |
| Lack of trust                                           |      |                     |     |
| Times checked balance                                   | 1.146*** | 0.251** | 1,493 |
|                                                       | (0.039) | (0.105)            |     |
| Times checked balance without withdrawing               | 0.336*** | 0.190** | 1,490 |
|                                                       | (0.035) | (0.093)            |     |
| Lack of knowledge                                       |      |                     |     |
| Hard to use ATM                                         | 0.106*** | 0.002 | 1,617 |
|                                                       | (0.013) | (0.025)            |     |
| Gets help using ATM                                      | 0.498*** | 0.050 | 1,612 |
|                                                       | (0.023) | (0.048)            |     |
| Knows PIN                                                | 0.575*** | −0.085** | 1,609 |
|                                                       | (0.017) | (0.034)            |     |
| Knows can save in account                               | 0.353*** | −0.034 | 1,617 |
|                                                       | (0.023) | (0.046)            |     |

Notes: * indicates statistical significance at \( p < 0.1 \), ** \( p < 0.05 \), and *** \( p < 0.01 \). Standard errors are clustered at the locality level. The “Mean” column shows the mean for those who have had the card for more than six months; the “Has card ≤ 6 months” column shows the regression coefficient on a dummy for those who have had the debit card for six months or fewer (i.e., the difference relative to the mean column). The precise questions on trust and knowledge are as follows. In the ENCASDU, the questions are “Do you leave part of the monetary support from Oportunidades in your bank account?” and, if the response is no, “Why don’t you keep part of the monetary support from Oportunidades in your Bansefi bank account?” The regressions presented here are not conditional on saving; those who report yes to the first question are coded with trust and knowledge dependent variables of 0 and included in the regressions. The second question includes pre-written responses and an open-ended response (“other; specify”; 4% of the sample in this table responded using the open-ended option); both pre-written and open-ended responses were coded as lack of knowledge, fear of ineligibility, lack of trust, or another explanation for not saving. In the Payment Methods Survey, each regression comes from a different survey question. These questions are: (1) Times checked balance: “In the last bimester, how many times did you consult your balance?” (2) Times checked balance without withdrawing: created by subtracting “In the last bimester, how many times did you withdraw money from the ATM?” from (1); (3) Hard to use ATM: responded “The ATM is difficult to use” (pre-written response) or a similar open-ended response to the question “What have been the main problems you have had with the ATM? [Wait for a response and record up to three of the options]”; (4) Gets help using ATM: “In general, does someone help you use the ATM?”; (5) Knows PIN: “Do you know your PIN by heart?”; (6) Knows can save in account: “Did they tell you that with the card you have a Bansefi savings account?”
Table 3.3: Other Barriers to Saving Informally

<table>
<thead>
<tr>
<th>Dependent variable: savings</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Has card at $t$</td>
<td>277.82**</td>
<td>241.59*</td>
<td>284.15*</td>
<td>215.15*</td>
</tr>
<tr>
<td></td>
<td>(126.94)</td>
<td>(137.05)</td>
<td>(147.29)</td>
<td>(110.05)</td>
</tr>
<tr>
<td>Has card at $t \times$ single</td>
<td>−168.32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(176.08)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Has card at $t \times$ baseline female bargaining power</td>
<td></td>
<td>−198.07*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(based on self-reported decision making)</td>
<td></td>
<td>(115.56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Has card at $t \times$ baseline female bargaining power</td>
<td></td>
<td>−163.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(based on age, education, literacy, income differences)</td>
<td></td>
<td>(133.21)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Has card at $t \times$ household gave money to others at baseline</td>
<td></td>
<td></td>
<td>354.82</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(419.10)</td>
<td></td>
</tr>
<tr>
<td>Number of households</td>
<td>2,951</td>
<td>1,625</td>
<td>1,484</td>
<td>2,951</td>
</tr>
<tr>
<td>Number of observations</td>
<td>11,275</td>
<td>6,300</td>
<td>5,778</td>
<td>11,275</td>
</tr>
<tr>
<td>Subsample</td>
<td>All</td>
<td>Not single*</td>
<td>Not single*</td>
<td>All</td>
</tr>
<tr>
<td>Time fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Household fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Winsorized</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
</tr>
</tbody>
</table>

Notes: *Not single refers to beneficiaries who live with a spouse (95% of the group) or at least one other adult (5% of the group). * indicates statistical significance at $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$. Standard errors are clustered at the locality level, using pre-treatment (2004) locality. Dependant variable is savings, constructed as income minus consumption and measured in pesos per month. “Baseline female bargaining power” uses questions only included in the 2002 wave of the survey on who decides (i) whether to whether to take their children to the doctor if they are sick, (ii) whether the children have to attend school, (iii) whether to buy them new clothes when needed, and (iv) “important decisions that affect the household members (transport, moving, changing jobs).” The measure is constructed by coding the responses to these four questions as 1 if a woman makes the decision, 0 if they make the decision jointly, and –1 if a man makes the decision, then the responses from the multiple questions are standardized and averaged following Kling, Liebman, and Katz (2007). “Household gave money to others at baseline” is a dummy variable equal to 1 if the household reported making transfers to others in any of the pre-treatment waves of the survey.
Table 3.4: Supply-Side Response

<table>
<thead>
<tr>
<th></th>
<th>Total ATMs</th>
<th>Total Branches</th>
<th>Bansefi ATMs</th>
<th>Bansefi Branches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current quarter</td>
<td>-1.52</td>
<td>0.03</td>
<td>-0.01</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(4.14)</td>
<td>(0.30)</td>
<td>(0.01)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>1 quarter lag</td>
<td>0.01</td>
<td>0.02</td>
<td>-0.02</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(4.11)</td>
<td>(0.34)</td>
<td>(0.01)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>2 quarter lag</td>
<td>-10.83</td>
<td>0.08</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(5.64)</td>
<td>(0.26)</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>3 quarter lag</td>
<td>-5.42</td>
<td>0.08</td>
<td>-0.03</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(2.98)</td>
<td>(0.26)</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>4 quarter lag</td>
<td>-0.74</td>
<td>0.42</td>
<td>0.00</td>
<td>-0.03</td>
</tr>
<tr>
<td></td>
<td>(5.97)</td>
<td>(0.50)</td>
<td>(0.01)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>1 quarter lead</td>
<td>-1.10</td>
<td>-0.12</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(3.66)</td>
<td>(0.36)</td>
<td>(0.00)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>2 quarter lead</td>
<td>-6.09</td>
<td>0.25</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(4.90)</td>
<td>(0.34)</td>
<td>(0.02)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>3 quarter lead</td>
<td>-7.84</td>
<td>0.25</td>
<td>-0.03</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(8.00)</td>
<td>(0.65)</td>
<td>(0.01)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>4 quarter lead</td>
<td>7.58</td>
<td>0.59</td>
<td>-0.01</td>
<td>-0.06</td>
</tr>
<tr>
<td></td>
<td>(10.32)</td>
<td>(0.94)</td>
<td>(0.03)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Mean control group</td>
<td>198.29</td>
<td>36.87</td>
<td>0.49</td>
<td>1.41</td>
</tr>
<tr>
<td>F-test of lags</td>
<td>1.26</td>
<td>0.20</td>
<td>0.68</td>
<td>0.96</td>
</tr>
<tr>
<td>[p-value]</td>
<td>[0.29]</td>
<td>[0.94]</td>
<td>[0.61]</td>
<td>[0.43]</td>
</tr>
<tr>
<td>F-test of leads</td>
<td>0.69</td>
<td>0.44</td>
<td>0.79</td>
<td>0.67</td>
</tr>
<tr>
<td>[p-value]</td>
<td>[0.60]</td>
<td>[0.78]</td>
<td>[0.53]</td>
<td>[0.62]</td>
</tr>
</tbody>
</table>

Municipality fixed effects Yes Yes Yes Yes
Quarter fixed effects Yes Yes Yes Yes

Notes: * indicates statistical significance at $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$. The table shows $\beta_k$ from

$$y_{jt} = \lambda_j + \delta_t + \sum_{k=-4}^{4} \beta_k D_{j,t+k} + \varepsilon_{jt}$$

where $y_{jt}$ is the number of ATMs or bank branches of any bank or of Bansefi in municipality $j$ during quarter $t$, $D_{jt} = 1$ if municipality $j$ has at least one locality with Oportunidades debit cards in quarter $t$. The F-test of lags tests $\beta_{-4} = \cdots = \beta_{-1} = 0$; the F-test of leads tests $\beta_{1} = \cdots = \beta_{4} = 0$. 
# Chapter 3: Banking on Trust

## Table 3.5: Crime as a Barrier to Saving Informally

<table>
<thead>
<tr>
<th>Dependent variable: savings</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Has card at $t$</td>
<td>223.76*</td>
<td>189.63</td>
<td>153.84</td>
<td>295.98**</td>
<td>268.20*</td>
<td>151.25</td>
</tr>
<tr>
<td></td>
<td>(117.48)</td>
<td>(115.45)</td>
<td>(117.85)</td>
<td>(123.30)</td>
<td>(116.74)</td>
<td>(135.26)</td>
</tr>
<tr>
<td>Has card at $t \times$ municipal crimes per 100,000</td>
<td>−0.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Has card at $t \times$ above median crimes per 100,000</td>
<td>−69.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(206.45)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Has card at $t \times$ municipal thefts per 100,000</td>
<td>−0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Has card at $t \times$ above median thefts per 100,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>59.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(201.25)</td>
</tr>
<tr>
<td>Has card at $t \times$ municipal homicides per 100,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>−6.92</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(4.46)</td>
</tr>
<tr>
<td>Has card at $t \times$ above median homicides per 100,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>−165.77</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(217.16)</td>
</tr>
</tbody>
</table>

| Number of households        | 2,951     | 2,951     | 2,951     | 2,951     | 2,951     | 2,951     |
| Number of observations      | 11,275    | 11,275    | 11,275    | 11,275    | 11,275    | 11,275    |
| Time fixed effects          | Yes       | Yes       | Yes       | Yes       | Yes       | Yes       |
| Household fixed effects     | Yes       | Yes       | Yes       | Yes       | Yes       | Yes       |
| Winsorized                  | 5%        | 5%        | 5%        | 5%        | 5%        | 5%        |

Notes: * indicates statistical significance at $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$. Standard errors are clustered at the locality level, using pre-treatment (2004) locality. Dependant variable is savings, constructed as income minus consumption and measured in pesos per month.


## Appendix A: Additional Figures and Tables

Table 3.6: Change in Savings and Assets After Receiving Card

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumption</td>
<td>–178.76**</td>
<td>–144.01*</td>
<td>–132.90*</td>
<td>–251.28**</td>
<td>–149.74**</td>
</tr>
<tr>
<td></td>
<td>(85.60)</td>
<td>(74.76)</td>
<td>(67.30)</td>
<td>(116.00)</td>
<td>(68.58)</td>
</tr>
<tr>
<td>Income</td>
<td>85.91</td>
<td>106.55</td>
<td>72.16</td>
<td>40.93</td>
<td>69.71</td>
</tr>
<tr>
<td></td>
<td>(158.58)</td>
<td>(138.41)</td>
<td>(123.88)</td>
<td>(137.48)</td>
<td>(122.05)</td>
</tr>
<tr>
<td>P-value Consumption vs. Income</td>
<td>[0.055]*</td>
<td>[0.033]**</td>
<td>[0.071]*</td>
<td>[0.008]**</td>
<td>[0.053]*</td>
</tr>
<tr>
<td>Savings = Income – Consumption</td>
<td>264.66*</td>
<td>241.30**</td>
<td>215.98**</td>
<td>285.42**</td>
<td>234.59**</td>
</tr>
<tr>
<td></td>
<td>(134.64)</td>
<td>(114.94)</td>
<td>(103.50)</td>
<td>(118.95)</td>
<td>(104.74)</td>
</tr>
<tr>
<td>Purchase of durables</td>
<td>5.94</td>
<td>6.22</td>
<td>7.99</td>
<td>6.55</td>
<td>6.91</td>
</tr>
<tr>
<td></td>
<td>(12.55)</td>
<td>(8.52)</td>
<td>(4.82)</td>
<td>(6.78)</td>
<td>(4.55)</td>
</tr>
<tr>
<td>Asset index</td>
<td>0.04</td>
<td>0.04</td>
<td>0.06</td>
<td>–0.07</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Number of households</td>
<td>2,951</td>
<td>2,951</td>
<td>2,951</td>
<td>2,951</td>
<td>2,938</td>
</tr>
<tr>
<td>Number of observations</td>
<td>11,275</td>
<td>11,275</td>
<td>11,275</td>
<td>11,275</td>
<td>11,243</td>
</tr>
<tr>
<td>Time fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Household fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Municipality × time fixed effects</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Household characteristics × time</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Winsorized</td>
<td>No</td>
<td>1%</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
</tr>
</tbody>
</table>

Notes: * indicates statistical significance at $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$. Standard errors are clustered at the locality level, using pre-treatment (2004) locality. Dependant variables are measured in pesos per month, with the exception of the asset index. Asset index is the first principal component of assets that are included in both the early (2002, 2003, 2004) and post-treatment (2009–2010) versions of the survey: car, truck, motorcycle, television, video or DVD player, radio or stereo, washer, gas stove, and refrigerator. Household characteristics are measured at baseline (2004, or for households that were not included in the 2004 wave, 2003). They include characteristics of the household head (working status, a quadratic polynomial in years of schooling, and a quadratic polynomial in age), whether anyone in the household has a bank account, a number of characteristics used by the Mexican government to target social programs (the proportion of household members with access to health insurance, the proportion age 15 and older that are illiterate, the proportion ages 6-14 that do not attend school, the proportion 15 and older with incomplete primary education, the proportion ages 15-29 with less than 9 years of schooling), and dwelling characteristics (dirt floors, no bathroom, no piped water, no sewage, and number of occupants per room). The number of households in column (5) is slightly lower because 13 households have missing values for one of the household characteristics included (interacted with time fixed effects) in that specification.
Figure 3.18: Difference between Treatment and Control in Marginal Propensity to Save Out of Transfer, Separated by Time with Account

(a) Wave 1 vs. Control (Older Accounts)

(b) Wave 2 vs. Control (Older Accounts)

(c) Wave 1 vs. Control (Younger Accounts)

(d) Wave 2 vs. Control (Younger Accounts)

Sources: Administrative data from Bansefi on average account balances by bimester, transfer payments, and timing of card receipt.

Notes: (a) $N = 743,776$ from 99,362 accounts; (b) $N = 905,335$ from 118,228 accounts; (c) $N = 455,172$ from 79,511 accounts; (d) $N = 1,088,677$ from 157,717 accounts. See the notes to Figure 3.11 for the specification. Accounts are split into older accounts and younger accounts based on the median account opening date, which is October 18, 2004.
Appendix B: Mechanical Effect

This appendix explains the computation of the mechanical effect for every pattern of deposit and withdrawal that occurs in a bimester. The mechanical effect is defined as the contribution to average balances from the transit of the Opportunidades transfers on recipients’ accounts, and interpreted as the part of average balances that does not represent net savings in the bimester.

Table 3.7: Computation of Mechanical Effect

<table>
<thead>
<tr>
<th>Pattern</th>
<th>% Total</th>
<th>Conditions</th>
<th>Mechanical Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A. Regular patterns: single deposit into account in the bimester</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) DW 73.4%</td>
<td>( W \leq D )</td>
<td>( W &gt; D )</td>
<td>( \text{lapse}_{DW} \ast \text{Withdrawal} )</td>
</tr>
<tr>
<td>(2) DWW 9.1%</td>
<td>( W_i \geq D )</td>
<td>( W_i &lt; D ) &amp; ( W_i + W_2 \geq D )</td>
<td>( \text{lapse}_{DW} \ast \text{Withdrawal}<em>1 + \text{lapse}</em>{DW_2} \ast (\text{Deposit} - \text{Withdrawal}_1) )</td>
</tr>
<tr>
<td>(3) DWW 1.7%</td>
<td>( W_i \geq D )</td>
<td>( W_i &lt; D ) &amp; ( W_i + W_2 \geq D )</td>
<td>( \text{lapse}_{DW} \ast \text{Withdrawal}<em>1 + \text{lapse}</em>{DW_2} \ast (\text{Deposit} - \text{Withdrawal}_1) )</td>
</tr>
<tr>
<td>( W_i + W_2 &lt; D ) &amp; ( W_1 + W_2 + W_3 \geq D )</td>
<td>( \text{lapse}_{DW} \ast \text{Withdrawal}<em>1 + \text{lapse}</em>{DW_2} \ast \text{Withdrawal}<em>2 ) + ( \text{lapse}</em>{DW_3} \ast (\text{Deposit} - \text{Withdrawal}_1 - \text{Withdrawal}_2) )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel B. Irregular patterns: multiple deposits into account in the bimester</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) DDWW 3.1%</td>
<td>( W_1 \leq D_i ) &amp; ( W_2 \leq D_2 )</td>
<td>( W_1 &gt; D_i ) &amp; ( W_2 \leq D_2 )</td>
<td>( \text{lapse}_{D_i W_1} \ast \text{Withdrawal}<em>1 + \text{lapse}</em>{D_i W_2} \ast \text{Withdrawal}_2 )</td>
</tr>
<tr>
<td>( W_1 &gt; D_i ) &amp; ( W_2 \leq D_2 )</td>
<td>( W_1 &gt; D_i ) &amp; ( W_2 &gt; D_2 )</td>
<td>( \text{lapse}_{D_i W_1} \ast \text{Withdrawal}<em>1 + \text{lapse}</em>{D_i W_2} \ast \text{Deposit}_2 )</td>
<td></td>
</tr>
<tr>
<td>(5) DWD 3.0%</td>
<td>( W \leq D )</td>
<td>( W &gt; D )</td>
<td>( \text{lapse}_{DW} \ast \text{Withdrawal} )</td>
</tr>
<tr>
<td>( \text{lapse}_{DW} \ast \text{Deposit} )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6) DDW 2.7%</td>
<td>( W \geq D_i + D_2 )</td>
<td>( W &lt; D_i + D_2 ) &amp; ( W \leq D_2 )</td>
<td>( \text{lapse}_{D W} \ast \text{Deposit}<em>1 + \text{lapse}</em>{D W} \ast \text{Deposit}_2 )</td>
</tr>
<tr>
<td>( W &lt; D_2 )</td>
<td>( \text{lapse}_{D W} \ast (\text{Withdrawal} - \text{Deposit}<em>1) + \text{lapse}</em>{D W} \ast \text{Deposit}_2 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7) DWDW 1.6%</td>
<td>( W_i \leq D_i ) &amp; ( W_2 \leq D_2 )</td>
<td>( W_i &gt; D_i ) &amp; ( W_2 \leq D_2 )</td>
<td>( \text{lapse}_{D W_1} \ast \text{Withdrawal}<em>1 + \text{lapse}</em>{D W_2} \ast \text{Withdrawal}_2 )</td>
</tr>
<tr>
<td>( W_i &gt; D_i ) &amp; ( W_2 &gt; D_2 )</td>
<td>( \text{lapse}_{D W_1} \ast \text{Deposit}<em>1 + \text{lapse}</em>{D W_2} \ast \text{Deposit}_2 )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( D_i \) indicate the \( i^{th} \) deposit and \( W_j \) indicate the \( j^{th} \) withdrawal within a bimester. \( \text{lapse}_{D W} \) measures the number of days between the \( i^{th} \) deposit and the \( j^{th} \) withdrawal, divided by the number of days in the bimester. The patterns listed here represent 95% of all bimonthly patterns, but (up to the first four transactions per bimester of) all patterns have been coded to obtain an estimate of the mechanical effect.
Bibliography


IMF. “Medición del Incumplimiento y del Gasto Tributario en Costa Rica”. In: (2012).


[125] Dina Pomeranz. “No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax”. In: *American Economic Review* 8 (2015), pp. 2539–2569. ISSN: 0002-8282. DOI: 10.1257/aer.20130393. URL: http://qut.summon.serialssolutions.com/2.0.0/link/0/eLVHCXwY2BQADaSjdJSU40tkpNTzJNU1LTDBMN0o1TgLv9yKqZrKYBtSae4mxYuSSYmA7FXQrMrDNb2w0xsCbfDnlcC3iuWIsGgkGhungKqvyBFqoTIX0zxCQTo8SktFRD40RGm9rEFA.


