

Firms as Tax Collectors*

Pablo Garriga[†]

Dario Tortarolo[‡]

Job Market Paper

This version: July 14, 2022

[Most recent version here](#)

Abstract

We study the implications of delegating tax collection duties to firms. We exploit a major reform to the withholding regime of the turnover tax in the City of Buenos Aires, where several large firms were appointed to act as collection agents (CAs) based on a predetermined revenue threshold. Combining rich firm-to-firm administrative data with quasi-experimental methods, we show that: (i) firms newly appointed as CAs do not change their reported business activity, (ii) firms with preexisting commercial ties to CAs increase their reported sales, and (iii) the government collects more tax revenue. The analysis of a subsequent reform that reduced third-party tax collection shows that firms respond symmetrically by reducing their reported sales. Our findings suggest that reforms to tax administrations can have a considerable impact when it comes to raising revenue and building tax capacity.

*We would like to thank John Friedman, Pierre Bachas, Anne Brockmeyer and Kenneth Chay for research guidance and very fruitful discussions. We are also thankful to David Agrawal, Youssef Benzarti, Lorenzo Casaburi, Jonathan Weigel, and the attendees at the Brown University Applied Micro Lunch, “NBER Business Taxation in a Federal System”, and “6th Zurich Conference in Public Finance in Developing Countries” for their insightful comments and suggestions. We are grateful to Hilario Iñiguez and Octavio Pera of AGIP for helping us to understand the data and for making this work possible. The views expressed are those of the authors alone and do not reflect the views of AGIP.

[†]Brown University, pablo_garriga@brown.edu, web page: <https://pablogarriga.github.io>

[‡]University of Nottingham and IFS, dario.tortarolo@nottingham.ac.uk

1 Introduction

Governments in developing countries struggle to raise revenue and build tax capacity. Increasing taxes and mitigating noncompliance are the most common policy measures considered by authorities. Yet achieving large-scale capacity often requires fundamental transformations of tax administrations.¹ The question of how taxes should be collected is usually at the forefront of tax authorities' debates. Among the possible methods, withholding systems—in which third parties collect and remit taxes owed by related parties—have historically played a central role in easing the burden of tax administration.

Tax withholding systems are ubiquitous. Part of the reason for the income tax's global success, for example, is that employers withhold at source from their employees (Besley and Persson, 2014). Similarly, the built-in withholding mechanism of the value added tax (VAT) has been shown to enhance compliance (Pomeranz, 2015; Waseem, 2020). Withholding of business income is particularly common in developing countries where governments lack the resources and capacity to accurately measure and tax firm activity (Slemrod, 2008; Slemrod and Velayudhan, 2018; Brockmeyer and Hernandez, 2019). In this case, financial institutions and large companies collect taxes owed by other firms in their commercial network. The use of this collection device for indirect taxes has surged over the last decade (Figure 1), but little is known about its implications.

This paper asks what the effects are of delegating tax collection duties to firms. To answer this, we analyze an unprecedented expansion of a withholding scheme to collect the turnover tax in the City of Buenos Aires, Argentina. Withholding entails changes to two main aspects of tax collection: who remits the tax and when it is paid. Under the standard filing procedure, firms are required to report their monthly sales and then apply the corresponding tax rate to determine the tax owed. Under the withholding scheme, some large firms are appointed as collection agents (CAs) and collect part of the tax in advance from their commercial partners whenever there is a purchase or a sale. While in a first-best world the point of collection of the tax would be irrelevant, in the real world changes in the collection mechanism can matter for compliance (see Slemrod, 2008; Kopczuk, Marion, et al., 2016). Our goal is to provide insight into the implications of withholding for both the collecting party and its linked taxpayers.

We bring new evidence to the discussion by exploiting a reform in 2016 that dramatically expanded the turnover tax withholding scheme. Under this reform, firms were

¹Such administrative reforms have proven highly effective according to recent research on Indonesia by Basri et al. (2021).

appointed as CAs according to size: firms whose sales in the preceding year were greater than AR\$60 million (the 97th percentile) were automatically appointed as CAs. As a consequence of this change, the number of CAs more than doubled from one month to the next (Figure 2). Appointed firms then started to withhold the tax from their commercial partners. This implied an increase in the share of taxpayers' tax liability collected indirectly at source by CAs in lieu of direct payments to the tax authority. This setting allows us to estimate both the *direct effects* of expanding the withholding scheme on the firms tasked with collecting taxes and the *indirect effects* on those that now face an increase in the amount of taxes withheld at source. To our knowledge, we are one of the first to study the full rollout of a policy of this kind; existing papers analyzing withholding of business income taxes omit the effects on the withholding agents, who bear the administrative burden of the task.

Another novel feature of our paper concerns the data. Our empirical analysis relies on two administrative sources processed by the tax administration: (1) monthly turnover tax declarations for nearly the entire universe of firms operating in the City of Buenos Aires; (2) transaction-level details of purchases and sales, provided by firms acting as CAs. These tax records allow us to trace out the commercial linkages between firms and build a business-to-business database, which is a key feature for our analysis. By combining the two data sources, we obtain a monthly panel of 180,000 firms and 1.6 million client-supplier pairs spanning 2015 to 2020. We combine these data with quasi-experimental methods that leverage the exogenous variation created by the reform to estimate direct and indirect effects.

Our empirical analysis begins by documenting the implementation of the program and the consequent expansion of the withholding net to provide some insights into how the policy actually worked in practice. We present graphical evidence suggesting that newly appointed firms quickly took up their roles and started collecting the tax from clients and suppliers very much like other preexisting CAs. We then proceed to analyze whether appointing firms as CAs affects their economic activity. We use a regression discontinuity design that exploits the discrete change in the probability of acting as a CA caused by the 2016 reform, where appointments were decided by applying the AR\$60 million cutoff rule to 2015 sales.² We estimate whether CAs change their reported sales by comparing firms on either side of the cutoff. Our evidence shows no visible discontinuity in reported sales in the years 2017, 2018, and 2019, suggesting that CAs' business activity was not

²We show that the rule was followed closely but with some exceptions: firms belonging to some specific industries were either left out or included in the collection scheme regardless of their 2015 sales. Nonetheless, the setting allows us to estimate a fuzzy regression discontinuity design.

affected by their new tax collecting role. We attribute this result to the fact that appointed firms were among the largest—and presumably most formal—firms in the economy, with highly streamlined tax-filing practices, such that collecting taxes from partners and remitting them to the tax administration did not imply an increased burden or a change in enforcement perceptions.

Next, we analyze the response of taxpayers to withholding. We focus on firms that regularly trade with newly appointed CAs and, as a consequence of the reform, experience an increase in the amount of tax collected through withholding. We use a difference-in-differences methodology that relies on the fact that some firms were more exposed to the reform than others. Exposure is defined as the presence of new CAs in the network of commercial partners of a given firm. Firms in the control group are only linked to pre-existing CAs and do not experience any change in the way their tax is collected; firms in the treatment group are linked to newly appointed CAs, and thus some of their direct payments to the tax authority are withheld in advance after the reform goes into effect. The first-stage results show that both the number of linked CAs and the share of taxes withheld at source increase sharply among firms more exposed to new CAs, confirming that the reform only had bite for the treated group. We then assess the response of taxpayers' self-reported sales. Our estimates show that the reported sales of both groups were evolving in parallel trends before the reform. Hence, our control group provides a credible counterfactual of the trend that treated firms would have followed absent the reform. After the increase in exposure to new CAs, however, sales for the treated group rise by 5.8 percentage points. We also show an increase in tax collected from treated firms after the reform that is similar in magnitude. The findings are robust to alternative specifications that allow for more flexible definitions of treatment and control groups. One potential interpretation of these results is that withheld taxes act as a lower bound on self-reported income: if firms were underreporting their sales before the reform, having a larger part of the tax remitted by a CA may have forced them to increase the amount they must report.

In the last part of the analysis, we leverage a policy change in September 2018 that went in the opposite direction to the one studied so far and that affected a subgroup of firms linked to CAs. We analyze a temporary reduction in tax withholding on financial transactions that targeted firms whose sales in 2017 were below AR\$10 million. While, for firms below the threshold, banks no longer withheld taxes on deposits received from other firms, firms above the threshold remained subject to bank withholding. This allows us to shed light on whether firms respond symmetrically to changes in withholding by reporting lower sales when withholding decreases. Using a difference-in-discontinuity

design, we show that firms below the cutoff experienced a reduction in the share of taxes withheld at source and responded by reducing their reported sales. That is, the same set of firms that report more sales when they become subject to withholding by CAs in 2016 exhibit a reduction in their reported sales when the bank withholding decreases in 2018. This result strikes us as remarkable and could suggest that a reduction in the coverage of third-party information (that is, banks providing information to the tax authority on the amount of money going in and out of businesses' bank accounts) gives firms more freedom to misreport their sales and thus to reduce their tax liability.

Taken together, our findings suggest that withholding by large firms can be an effective tax collection tool. Appointing firms to collect taxes does not seem to hurt their activity directly. And substituting direct tax payments with withholding at source increases reported sales of firms linked to newly appointed CAs, which, in turn, increases total tax revenue collected by the government.

This paper builds on several strands of the public finance and development literature. First, it contributes to the literature on tax compliance and enforcement (Allingham and Sandmo, 1972; Kleven, Knudsen, et al., 2011; Slemrod and Yitzhaki, 2002; Slemrod, 2019). In particular, we build on recent papers that emphasize two critical aspects of modern tax systems: how firms play a key role as fiscal intermediaries and how tax compliance critically depends on automatic features such as withholding (Kleven, Kreiner, et al., 2016; Slemrod, 2008).

Second, we contribute to burgeoning empirical research on the role that withholding can play in improving tax compliance of firms. In Costa Rica, Brockmeyer and Hernandez (2019) exploit changes in firm-specific withholding rates from credit and debit card companies who withhold a fraction of retailers' sales. They find an increase in sales tax remittance and revenue for firms facing higher withholding rates. The mechanisms behind this are default payment and increased enforcement perceptions. Meanwhile, Waseem (2020) studies the roll out of the VAT in Pakistan. Manufacturing firms, already in the tax net, increase their reported revenue when the tax is extended upstream to the energy sector and their inputs become subject to withholding. Waseem explains this behavior by arguing that the withheld amount acts as a lower bound on self-reported sales.

Our work is closely related to these two papers but builds on their work in five key aspects. First, our scope is somewhat broader as the reform affects firms across sectors as well as upstream and downstream the supply chain. This is an important aspect, as tax compliance usually varies along the supply chain and across sectors (Pomeranz, 2015). Our results thus reflect more closely what the average effect of withholding might be

in the economy. Second, the appointment rule based on pre-reform sales allows us to gauge the *direct* effects that CAs might experience by collecting taxes on behalf of the tax authority. Such direct effects have been largely overlooked by the literature thus far. Third, our research design is also very appealing to study *indirect* effects of withholding on firms linked to newly appointed CAs. The sharp and massive appointment of CAs is an arguably exogenous event from the point of view of their trade partners.³ Fourth, we leverage two policy reforms going in opposite directions, allowing us to document the effects on compliance of the implementation and removal of withholding. Fifth, we focus on the turnover tax, the largest subnational revenue source, in a context where a federal VAT with its self-enforcing mechanisms—as described by Waseem (2020)—was already in place. Our findings thus suggest that businesses might follow disparate compliance strategies when reporting their sales in subnational and federal tax returns.

Lastly, and on a broader level, we contribute to the growing literature on optimal tax administration and the importance of tax collection under imperfect and weak enforcement (Keen and Slemrod, 2017). For instance, for property tax collection, recent research has analyzed performance pay (Khan et al., 2016), the trade-offs of appointing city chiefs or state agents as collectors (Balán et al., 2022), and better assignment of collectors to teammates and households (Bergeron, Bessone, et al., 2021). In the context of business taxation, however, the evidence on administrative reforms is more scant. Recent work has studied, for example, monitoring effects of large taxpayer offices on compliance (Almunia and Lopez-Rodriguez, 2018; Basri et al., 2021). As low- and middle-income countries develop, the barriers to implementing more sophisticated tax collection systems decrease and appointing large firms to collect taxes from trade partners in advance could become a feasible tool that raises more revenue, as our work shows.

The paper is organized as follows. Section 2 provides details on the tax and the general implications of the reform. Section 2.2 discusses our data sources. Section 3 describes the conceptual framework we use to explain the empirical results. Section 4 documents the expansion of the withholding scheme and its direct effect on CAs while Sections 5 and 6 study the indirect effect on taxpayers. Section 7 concludes.

³This includes, among other cases, firms that go from being not withheld at all to being withheld at source for the first time.

2 Context, Data, and Reforms

2.1 Subnational Turnover Taxes in Argentina

The turnover tax (TT) is a subnational tax levied on firm revenue with no deductions for costs. It applies to all transactions taking place in the supply chain—that is, business-to-business and business-to-consumer transactions. This implies that it creates “cascading effects” in which final goods are taxed multiple times throughout production and thus incentivizes vertical integration. Despite its distortive effects, the tax is simple to calculate and administrate relative to other taxes, which explains why it is used in all of the twenty-four jurisdictions in Argentina.⁴ In fact, it constitutes the main subnational revenue source, accounting for about 75% of total tax collected in the City of Buenos Aires.

How is the TT collected? The TT is filed monthly and it can be paid directly by taxpayers (standard method) or collected indirectly by third parties (withholding method). Under direct payment, taxpayers log on to the tax administration’s website at the end of the month, self-report the gross income accrued in each activity over that month; the system applies the corresponding tax rate and generates a payment coupon. Under the withholding method, the tax administration can designate firms, banks, or credit and debit card companies to operate as collection agents (CAs). In the case of CA firms, every time a client purchases inputs from a CA or a supplier sells goods or services to a CA, part of the invoice amount is withheld by the CA and remitted to the tax authority. Hence, this are regular firms that have the extra task of withholding from their trading partners. In the case of CA banks, every time a payment is deposited in the bank account, the bank withholds part of it and remits the funds to the tax authority. In the case of CA card companies, every time a transaction is made with a debit or a credit card, the card company withholds part of it. In all these cases, taxpayers still have to file taxes at the end of each month but the withheld amount constitutes a credit in favor of the taxpayer who can discount it from the tax liability before making any payments.⁵

⁴For recent research about such distortions see Bilicka et al. (2022) and Hansen et al. (2021). Hansen et al. (2021) show that the tax is re-emerging in the US, where nine states are currently using it. They attribute this to the following facts: (i) the large tax base (the TT taxes services, while retail sales taxes do not) can generate large revenues with low rates, (ii) retail sales taxes share its main drawback of taxing inputs, and (iii) its broad tax base limits firms’ tax evasion opportunities.

⁵For more details about the TT and the way it is collected, see Appendix B.

2.2 Administrative data

We combine two data sets derived from information collected by the Tax Administration Office in the City of Buenos Aires (AGIP). The first data source consists of monthly tax declarations for nearly the entire universe of firms operating in the city in the period 2015-2020. These declarations are filed electronically through an account each taxpayer has with the tax administration. We observe all line items completed in a standard filing procedure (anonymized taxpayer identifier; sales;⁶ tax liability; amount of tax withheld from different sources; outstanding credits or debits) and a series of firm characteristics used to determine the corresponding tax and withholding rates (firm type: small or large; industry: 995 six-digit codes; location: local or out of province).

To construct our second data set, we leverage a unique feature of our setting: firms acting as CAs are required to file a supplementary “invoice summary” that lists all their purchases and sales in a given month. The line items included in this form are: the taxpayer identifiers of both parties involved in the transaction, total value of the transaction, the amount withheld, and whether the transaction was a purchase or a sale. We use these filings to trace commercial linkages between CAs and their commercial partners. A limitation of this data source is that trade linkages are only observed when a firm becomes a CA—that is, when it starts filing the supplementary form that records transactions. This implies that for some firms we observe linkages since the beginning of our sample in 2015, while for firms linked to CAs appointed under the reform, we observe linkages only at or after November 2016.⁷ Combining the two sources of data, we obtain a panel of 180,000 firms and 1.6 million client-supplier pairs spanning 2015 to 2020.

Table 1 presents summary statistics for the period prior to the reform. Panel A describes the full sample, which contains over 180,000 firms, roughly 5 percent of them are CAs. Overall, the firm-size distribution in terms of revenue is right-skewed. CAs are orders of magnitude larger than regular firms which corroborates the enrollment heuristics followed by the government of having large firms act as CAs. Panel B corresponds to the estimating sample used in Sections 5 and 6. When evaluating the response of taxpayers, we restrict the sample of firms to those that file taxes regularly (we drop firms that report positive sales in less than 50 percent of the periods).⁸

⁶Technically, the tax base is “gross income”—which is the amount that a business earns from the sale of goods or services, before the deduction of expenses. For simplicity, we refer to the tax base as “sales.”

⁷Figure E.1 illustrates the process by which we construct our full data set and stresses the sequence in which the information is revealed.

⁸We provide further insight about the data in Appendix D, where we document the variation used in the paper in more detail and without imposing any sample restrictions.

Our main focus is on withholding by firm CAs. However, we also analyze a policy reform that temporarily eliminated the withholding by banks on a subset of firms. In Section 3, we provide more details and describe the practical implications of the tax collection mechanisms—here, instead, we focus on describing the expansion of the withholding scheme by CAs, the main source of variation in our analysis.

2.3 A Sharp Expansion in the Net of Tax Collectors

Withholding by CAs was introduced in the City of Buenos Aires in 1985; between 1985 and 2016 some large firms were appointed as CAs on a case-by-case basis. These firms had to withhold part of the taxes owed by their commercial partners whenever there was a purchase or a sale. In July 2016, a resolution was passed with the intention of expanding the set of firms acting as tax collectors.⁹ It established a new appointment criterion based on sales: firms whose annual sales in 2015 were greater than AR\$60 million would be automatically enrolled as CAs.¹⁰ Newly appointed firms were notified in October and started their collection duties in November 2016. Importantly, all the firms who were operating as collectors before this reform, remained in the tax collection scheme regardless of their revenue and, since this was a one-time reform, no new appointments were made from 2017 through 2020. For clarity, we summarize the timeline in the diagram below.



The immediate implication of the reform was a sharp increase in the number of CAs. Figure 2 plots the number of firms acting as CAs over time.¹¹ Its pattern exactly matches the timeline described above: the number of active CAs gradually increased in the period prior to the reform; between October 2016 and November 2016, however, the number of CAs more than doubled (it increased from 3,656 to 8,664 from one month to the next), and remained mostly unchanged afterward (the slight decrease is explained by firm exit).

⁹AGIP resolutions number 364, 421, and 486 from 2016. Figure E.2 shows a newspaper article describing the reform as a major tax change in the jurisdiction.

¹⁰This corresponds to roughly USD 6 million in 2015 affecting the top 3% of firms of the sales distribution.

¹¹Figure E.3 shows an example of the register of withholding agents published monthly by the tax administration. We digitized these files and constructed a database which we merged with our main data. The list of CAs is quite diverse in terms of sector with coffee producers, pharmaceutical firms, insurance companies, footwear manufacturers, and construction wholesalers, among others.

These newly appointed CAs began to withhold the tax from their commercial partners which increased the share of taxpayers' tax liability collected indirectly at source by CAs *in lieu* of direct payments.

2.4 A Decrease in Withholding by Banks

We analyze another policy change that took place during the period of our study that—unlike the reform described so far—reduced exposure to withholding for a group of firms. In September 2018, the city's government suspended the withholding by banks on the financial transactions of firms whose sales in 2017 were below AR\$10 million. This measure was part of a stimulus plan targeted at small and mid-sized enterprises during a national economic crisis. Although this suspension was initially meant to last six months, it was extended for another six months in March 2019. We use this second reform to the city's tax collection mechanism as a quasi-experiment to test whether a reduction in the share of taxes collected via withholding leads to a fall in reported tax liability.

3 Conceptual Framework

In this section we describe how tax collection works in practice and provide some insights on how changing the tax collection mechanism might affect firm behavior. Figure 4 summarizes the main features of the direct-payment and withholding mechanisms in a simplified setting with four agents: supplier, retailer, consumer, and tax administration. The supplier sells inputs to the retailer, which pays X for the goods. In turn, the retailer takes the inputs and turns them into final goods, which are sold to the consumer, who pays Y for them. As described earlier, all firms along the supply chain are liable for the tax. In our setting, the supplier faces a tax rate τ_S and the retailer faces a tax rate τ_R , while the withholding rate (a fraction of τ) is represented by $\alpha \in (0, 1)$.

Under the direct-payment mechanism, taxpayers self-report their sales to the tax authority at the end of the month. The supplier reports X , and the total tax liability for the period is $\tau_S X$. Likewise, the retailer reports Y such that its liability is $\tau_R Y$.

Under the withholding scheme, firms acting as CAs are in charge of withholding a part of the total tax liability from their commercial partners and remitting this amount to the tax administration. This task is performed in addition to their own duties as taxpayers to remit their own tax payments.

Withholding may be performed by either the seller or the buyer. When the seller acts as CA, it adds the withheld tax to the total sale amount. In our setting, this is represented by the supplier charging an extra amount for the sale to the retailer, $X(1 + \alpha\tau_R)$, which is then remitted to the tax administration (along with the seller's own taxes). When the retailer's tax is due, it will only pay the outstanding balance that remains after deducting the remitted amount from the total tax liability, $\tau_R Y - \alpha\tau_R X$. When the buyer acts as CA, it subtracts the withheld tax from the total purchase amount. In our setting, the retailer pays $X(1 - \alpha\tau_S)$ to the supplier and remits $\alpha\tau_S X$ to the tax administration. In this case, the supplier is entitled to deduct $\alpha\tau_S X$ from its tax bill $\tau_S X$ before making the final payment.

Withholding through CAs changes two main aspects of tax collection: *who* remits and *when* the tax is paid. Under the direct-payment mechanism, the tax is filed by the liable party on the due date. In contrast, under the withholding mechanism, the tax is paid by a third party in advance; the tax is withheld at source and remitted on a regular basis. These changes carry associated implications for both the collecting party and the taxpayer.

For CAs, appointment leads to closer scrutiny and increased penalties if they do not fulfill their role as CAs. Thus, the appointment itself might trigger changes in enforcement perceptions that lead to an increase in reported sales for firms that were previously underreporting (see Brockmeyer and Hernandez, 2019). Additionally, withholding is an administrative burden, as it requires CAs to keep a register of transactions with trading partners and to file an additional form detailing such transactions. As pointed out by Slemrod (2008), this burden might be offset by the "cash-flow benefit" of withholding: CAs may hold onto the funds they withhold from the time of collection until the time of remittance, giving them the equivalent of an interest-free loan. However, in our setting the magnitude of these two contrasting forces and the changes in enforcement perceptions are likely to be small for three reasons. First, the collection process is highly streamlined: when a firm is appointed CA, it is provided with a tax collection software that automatically incorporates withholding into its transactions.¹² Second, firms must transfer the withheld funds every month following a pre-specified schedule that only allows for a small window of time between the transaction and the remittance deadline during which they could dispose of the funds. Third, the authority appoints very large firms as CAs (the top 3%) for which evasion rates are usually close to zero (e.g., Kopczuk and Slemrod, 2006; Best, Shah, et al., 2021). Firms with such scale tend to truthfully report their gross sales (the tax base in our setting) and, at best, might engage in another type of creative accounting (e.g., Alstadsæter et al., 2019).

¹²The software [e-ARCIBA](#) constitutes the source from which our business-to-business data are extracted.

For linked firms, on the other hand, there are three implications of switching to the withholding collection mechanism. First, withholding enables each business-to-business transaction to be recorded in two places, generating third-party information. This may alter enforcement perceptions since the information could be used by the subnational tax administration to detect non-complying firms. Note, however, that the self-enforcing properties of a coexistent federal VAT somehow limits the extent of this channel, at least for VAT-registered firms whose transactions were already leaving a paper trail. Any response driven by this channel would thus require firms to be reporting sales differently to the federal and the subnational agency. Presumably, sales are lower in the City of Buenos Aires if firms believe that VAT information is not shared across agencies.

Second, the withheld amount itself may act as a lower bound on self-reported sales. If a firm reports sales such that the tax owed is lower than withholdings, the resulting difference will accumulate as a credit. This behavior is usually regarded as an indicator of underreporting by the tax authority and may lead to future audits if it persists through time. For this reason, taxpayers will try to avoid reporting sales such that their tax liability falls below the withheld amount. Hence, the withheld amount defines a lower bound on reported sales. Third, if withholding implies that a firm subject to it ends up facing a higher effective tax rate or closer scrutiny, the firm may decide—whenever possible—to switch trading partners to non-CAs to minimize its tax burden. Thus, over time, we might expect some trade links to break, or become less frequent, or decrease in volume. This reasoning goes in line with Gadenne et al. (2020) who document how tax systems can distort firm-to-firm trade decisions in markets where VAT and non-VAT-paying firms co-exist leading to segmentation of supplier networks.¹³

The first two implications can be further rationalized using the framework of firm tax evasion under third-party reporting developed by Carrillo et al. (2017), Brockmeyer and Hernandez (2019), and Waseem (2020). These papers primarily consider settings in which there is a VAT or sales tax, which allow for cost deduction. Our setting is slightly less complex because the TT does not allow for any type of deduction. Consider a firm that is already operating with CAs and thus subject to some withholding. If evasion costs are sufficiently low, it will underreport sales up to the point at which tax owed is completely offset by withheld funds. When a new trading partner is appointed CA and additional transactions become subject to withholding, the firm can only avoid falling into negative liability territory by increasing reported sales. Therefore, we should expect to see an increase in reported sales for firms more exposed to withholding. We should also expect to

¹³See also Pessina (2020) for a change in VAT remittance among firms trading with government entities.

see an upward shift in the distribution of tax liability as firms try to bunch close to the lower bound set by withheld funds. As shown in Pomeranz (2015), we would expect the response of reported sales to be stronger in sectors in which informality is high or at the retail stage in which the cost of evasion is lower.

4 Expansion of the Withholding Scheme

We begin the empirical analysis by documenting the expansion of the withholding scheme. Specifically, we describe the implementation of the appointment rule, the expansion of the withholding net, and the response of newly appointed CAs as they take on their new role.

We use a series of figures to provide evidence of whether the appointment rule worked in practice and how quickly the new CAs started withholding from linked firms; we find that take-up was high and firms responded quickly. Next, we analyze the response of appointed firms. Here, we use a regression discontinuity design (RD) that exploits the sharp change in the probability of acting as a CA based on 2015 sales (before the reform) and we compare the evolution of post-reform sales for firms close to the cutoff to test whether CAs change their behavior. If enforcement perceptions changed and if firms were underreporting sales prior to the reform, we should expect to see a response in sales for firms beyond the threshold. Our results suggest that CAs' business activity evolved similarly across the threshold over time. We interpret this as evidence that the policy did not cause any significant changes in their business structure.

Implementation of the appointment rule. Figure 5 shows the 2015 sales distribution for the whole sample of firms. It is immediately apparent from its inspection that firms targeted by the reform were among the largest in the economy: the AR\$60 million cutoff roughly corresponds to the 97th percentile of the firm-size distribution. The bottom panel zooms in around the cutoff and displays no discontinuities or bunching around it, which confirms that firms were not able to manipulate the assignment variable used to by the authority to appoint firms into the withholding scheme. This is a promising fact that suggests that we can use firms with sales in 2015 close to AR\$60 million to explore the direct effect of being appointed as a CA.

According to the appointment rule, firms above the AR\$60 million threshold should have been appointed to act as CA. Figure 6 provides evidence of how binding this rule was. On the horizontal axis we take the running variable and split it into equally spaced bins of size AR\$10 million; the vertical axis shows the probability of treatment (that is, of

being appointed as a CA) for firms in each of these bins. Under perfect compliance the probability of treatment should be zero for firms below the cutoff and one for firms above. However, the figure shows that some firms were enrolled as CAs despite being below and that not all firms above it were enrolled.

Expansion of the withholding net. Appointment of CAs by itself may not lead to any changes in tax collection if the firms selected for the task do not fulfill their duties. To ensure compliance, the tax administration closely supervises their activities and has set up fines for those who omit to withhold or delay the remittance of funds. The next figures are intended to show that appointed firms rapidly take up their new duties. To do so, we compare CAs appointed in November of 2016 with a group of preexisting CAs appointed in 2012 who we use as a benchmark of what we would expect to be a “normal” behavior from firms that are already established in their roles as CAs.

Figure 7a analyzes the extensive margin of collection duties: for each month it shows the proportion of CAs in each group that are withholding from their commercial partners. It takes roughly a year for newly appointed CAs to fully take on their role which matches the behavior documented in Figure 3 for the aggregate outcomes. Nevertheless, their starting point is very high: in the first month after appointment, about 75 percent of new CAs are actively withholding from other firms. Figure 7b, in turn, shows the intensive margin of collection duties: for each month it shows the average number of firms whose taxes are withheld by old CAs (light blue line) and new CAs (dark blue line). On average, newly appointed CAs trade with more than 50 firms, either purchasing inputs from them or selling goods to them. This is remarkably stable during the period of analysis. In each transaction, the new CAs now withhold part of the tax on behalf of their trade partners as an advance payment. Note also that new CAs withhold from fewer firms than old CAs. This might be because they are smaller or younger than preexisting CAs. Taken together, these elements provide evidence that the policy worked in practice. The final step in our description of the expansion of the withholding scheme is to assess whether it had any direct effect on newly appointed CAs.

4.1 Macro Evidence of the Reform

A striking feature of the reforms analyzed in this paper is that the implied changes in tax collection can be visualized in aggregate time series. Such changes speak by themselves of the relevance of such reforms.

Interestingly, this change in the collection system, although neutral in theory, ended

up increasing tax revenue for the government. Figure 3 provides macro evidence of the reform using aggregate and public time series data from the tax authority.¹⁴ The first two panels show the growth in the relative importance of withholding by CAs: both the share of TT collected through withholding and the amount of revenue raised by CAs experienced a sharp increase right at the time of the reform. The series keep increasing during 2017, suggesting some sluggish adjustment of newly appointed firms (we confirm this in the analysis of the micro data).

The third panel in Figure 3 shows the aggregate effect on revenue. It compares tax revenue in the City of Buenos Aires with that in the Province of Buenos Aires. These are the largest jurisdictions in the country in terms of economic activity, and their economic performances track each other closely, making each one a natural benchmark for the other. Specifically, the figure shows the percentage difference in tax revenue between the two jurisdictions relative to October 2016. This is analogous to running a difference-in-differences regression and, therefore, captures the aggregate causal effect of appointing firms as CAs (that is, it captures both direct and indirect effects). The graph shows that the City of Buenos Aires, relative to the province, was on a downward tax revenue trend, which is reverted right after the reform in November 2016. By 2019, the city's tax revenue was 20% higher than the province in the baseline period. The macro evidence thus shows that the reform had a large impact. Our goal in the remainder of the paper is to flesh out the direct and indirect effects across taxpayers.

4.2 Direct Effects on Newly Appointed Collection Agents

The empirical exercise that we perform in this section consists in testing whether firms respond to their enrollment as CAs by looking at their reported sales in the periods after the appointment event. To do so, we use a fuzzy RD framework focusing on firms whose sales in 2015 were close to AR\$60 million.

Recall that the reform to the withholding scheme established that firms whose sales in 2015 were at least AR\$60 million would be automatically enrolled as CAs. In Figure 6 we showed that, although assignment was not deterministic, there was a discrete increase in the probability of treatment at the threshold. In addition, Figure 5 showed that there were no discontinuities around the threshold of the running variable used to determine appointment into the CA program, suggesting that firms were not able to manipulate

¹⁴We find it striking that the effects of the reform were large enough to be visible in aggregate time series. These figures are in fact the reason that motivated us to pursue this project.

their entry into the regime. Moreover, the resolution that set the guidelines for the reform was enacted in July 2016, leaving no possibility for firms to adjust their 2015 sales. All these facts provide a strong case in favor of enrollment status being as good as randomly assigned. In the absence of any changes, we should expect reported sales in the periods after the appointment event to evolve smoothly around the cutoff. Therefore, finding any jumps in the outcome of interest for firms close to the boundary would be suggestive evidence of a causal effect of the change in appointment status.

We identify the response of firms to enrollment as CAs by running regressions of the following form:

$$y_i = \alpha + \beta \cdot 1(R_i \geq c) + \gamma \cdot (R_i - c) + \varepsilon_i. \quad (1)$$

Here y_i denotes the outcome of interest for firm i for any given year, $c = \text{AR\$}60$ million is the cutoff set by the reform, and R_i is the running variable, 2015 sales. The coefficient of interest capturing the effect of the discontinuity at c is β . In the first stage, we show that the probability of a firm above the threshold being appointed as a CA in November 2016 changes discretely around the cutoff. We then analyze whether firms adjust their (log) reported sales in response to enrollment.

Table 2 reports the results for the first stage and the 2SLS estimates from a fuzzy RD specification. Column 1 shows the first-stage change in the probability of a firm above the threshold being appointed as a CA in November 2016 (the coefficient is analogous to the size of the jump in the probability of treatment shown in Figure 6). Columns 2 through 5 correspond to the effect on reported annual sales (these coefficients are represented graphically in Figure 8). Column 2 excludes the two treated months of 2016, November and December, and thus the results of this column can be interpreted as a placebo test, as the column only contains pre-reform sales. Importantly, none of these estimates are statistically significant, suggesting that new CAs did not respond by changing their reported sales. We attribute this result to the fact that appointed firms were among the largest—and presumably most formal—firms in the economy, so tax collection duties did not saddle them with significant compliance costs to them, nor did they change their perception of tax enforcement.

The results shown in this section suggest that firms that became CAs did not change their behavior in response to enrollment. If tax collection duties did not induce any large changes in the business structure of newly appointed CAs, then it is likely that there were no other fundamental changes in their interactions with trading partners. In our view, this helps to reduce the concerns about potential confounding factors in the analysis that follows were we study how taxpayers react when their business partners become CAs.

5 How Do Firms Respond to an Increase in Withholding?

In this section we study the indirect implications of expanding the withholding scheme. We ask how do firms who were previously paying the tax through direct payments respond to an increase in withholding by CAs? The ideal experiment to answer this question would be to randomly allocate CAs throughout the economy and compare the response in reported sales of firms that are commercially linked to these CAs to those that are not. While we cannot randomly assign tax collection duties to firms, we can leverage on the design of the reform and use the built-in variation in exposure to CAs to identify a causal effect of withholding on reported sales. More specifically, we rely on the following quasi-experiment: for regular taxpayers, the expansion of the withholding scheme and the consequent increase in the number CAs meant that it became more likely to be subject to withholding by CAs in their commercial network and, therefore, tax collection became less reliant on direct payments of the tax as it was replaced by withholding at source.

The key econometric challenge that must be tackled to precisely estimate the effect of withholding on self-reported sales is to distinguish it from other time-varying shocks that might have taken place around the time of the reform. To address this issue we implement a difference-in-differences (DD) design relying on taxpayers' differential exposure to newly-appointed CAs. Since firms can have different commercial partners, those that happened to be connected to firms above the size threshold became more exposed to CAs than others. While the choice of trade partners itself is not exogenous, firms certainly could not influence whether their partners were appointed as CA or not. Hence, the increase in exposure to CAs was arguably exogenous to taxpayers.

We use trade linkages to construct two distinct groups of firms: those that are connected to newly appointed CAs (our treatment group, T) and those that are not (our control group, C). Firms in T have commercial ties with CAs appointed in November 2016 and therefore experience an increase in the number of CAs in their network at the time of the reform. Firms in C only trade with preexisting CAs and do not experience a change in the number of CAs in their network of commercial partners. Note that firms in T might also have commercial ties with preexisting CAs; the key difference between T and C is the presence of linkages to newly appointed CAs.

In an ideal setting the two groups would be constructed according to commercial linkages observed in a baseline period prior to the reform. However, due to the way the transaction-level data is reported, we do not observe a firm's commercial linkages unless one of its business partners becomes a CA. For firms connected to new CAs, this implies

that the first observed transaction is in November 2016 at earliest. To circumvent this issue, we assume that trade links remain stable for a six-month period and, therefore, any purchase or sale observed up to six months after the reform is informative about pre-existing commercial relationships prior to the reform. Then, we classify into T any firm that transacts with a newly appointed CA between November 2016 and April 2017 and classify into C those that only transact with preexisting CAs during the same period. We test the robustness of this approach and find that results do not vary considerably after changing the bandwidth of months used to construct the two groups.

Our DD design compares firms that experienced an increase in the number of CAs they relative to firms that did not. Under the assumption that the two groups would have behaved similarly in the absence of the policy change, their comparison will reveal the causal effect of an increase in the share of tax collected through withholding on taxpayer behavior. Coefficient β in the following regression captures this effect:

$$y_{it} = \beta \cdot Treat_i \cdot Post_t + \theta_i + \gamma_t + \varepsilon_{it}. \quad (2)$$

Here y_{it} is the outcome for firm i in period t , $Treat_i$ is an indicator for firms in the treatment group, $Post_t$ is an indicator for time periods after November 2016 (the reform date), θ_i are firm fixed effects, and γ_t are time-period fixed effects. We focus on the following outcomes. To gauge the first-stage results of the reform for taxpayers, we look at the number of CAs (extensive margin) and the share of tax withheld (intensive margin). The behavioral response is captured by self-reported sales, and tax liability.¹⁵ Although the data are provided at the monthly level, we aggregate them to the quarterly level in order to avoid some issues that are present in the most granular definition, such as the fact that some firms report zero sales in a given month and that large seasonal fluctuations in economic activity are related to the combination of the holiday season and the end of the fiscal year. In doing so, we define quarters relative to the reform (that is, the quarter of the reform, $t = 0$, corresponds to the months of November 2016 through January 2017). We restrict the time frame of our regressions to all available quarters from the pre-period (starting November 2015) to the last quarter of 2019 (ending January 2020, before the COVID-19 pandemic began to take a toll on economic activity). Having defined the time variable in this way, we construct a balanced panel of firms that report positive sales in at least half of the periods in which they are present in the data.

¹⁵Recall that the TT does not allow for deduction of costs, then tax liability is equal to the tax rate multiplied by the tax base (reported sales). Nevertheless, analyzing both reported sales and tax liability will allow to cross-validate our results.

We complement the DD estimates with the following event-study analyses:

$$y_{it} = \sum_{t=-q}^{-2} \delta_t \cdot D_{it} + \sum_{t=0}^m \beta_t \cdot D_{it} + \theta_i + \varepsilon_{it}. \quad (3)$$

Here D_{it} is an event-study indicator for each quarter relative to the baseline period, August-October 2016 ($t = -1$). The coefficients δ_t and β_t estimate the effect on the outcome of interest for each period relative to the baseline, net of firm fixed effects. This specification provides compelling visual evidence of the validity of the identification assumption, as the coefficients δ_t reveal pre-reform trends.

5.1 Empirical Results

Figures 9 and 10 show the event study estimates. The figures on the left show results for the levels regressions for T and C, we plot the coefficients for the time dummies plus the constant for each group. The figures on the right display the δ_t and β_t coefficients of equation 3. In addition, Table 3 Panel A reports the DD estimates from equation 2.

Figure 9 provides evidence on the first-stage of the reform: the change in the number of commercial partners acting as CAs and in the share of due taxes withheld at source. The first panel shows that the number of CAs connected to firms in T rose by an average of 3.5 in the months after the reform, while remaining unchanged in C, and that this post-reform gap persists throughout the study period. The second panel shows that the share of tax withheld by CAs exhibit a similar pattern, rising by almost 20 percentage points in the T right after the reform while remaining practically unchanged for C. Since the two groups track each other closely in the year prior to the reform and diverge immediately after, the evidence is consistent with quasi-random assignment of the new CAs.

Figure 10 shows the analogous plots for the response of self reported sales to the increase in withholding. The estimates show a similar pattern to those documented in the first-stage figure. Both groups have nearly identical outcomes in the pre-reform period, but they diverge markedly right after. Initially sales become 3.5 percentage points higher in T relative to C, with a second discrete jump two years after (which is likely due to the effects of the second reform that took place in September 2018). Overall, the estimates suggest that sales for firms in T outgrew those of C by 5.8 percentage points. Is the effect compatible with a real increase in sales or rather an increase in reported sales? Since the data does not provide information on volume of sales, it is impossible to provide a definitive answer. Nevertheless, we interpret the evidence as suggestive of a reporting effect:

the sharpness of the response precisely at event time, after nearly identical trends prior to the reform is consistent with firms correcting their tax filings to match the withheld amounts collected and remitted by CAs.

The final step in this part of our analysis is to evaluate the robustness of our results alternative specifications. Recall that, to assign firms into T and C, we assumed that trade links are stable six months after the reform—that is, any transaction observed between November 2016 and April 2017 reflects the commercial linkages that existed in the pre-period. One concern with this assumption is that our choice of months can affect the composition of T and C which, in turn, might have an impact on the estimated outcomes.

Table 4 investigates the robustness of the regression in Table 3 Panel A with respect to the number of months used to estimate baseline trade linkages. The first column replicates Panel A of Table 3, while the remaining columns repeat the specification changing only the number of months used to estimate these linkages: from one-month, where trade links are constructed based on transactions observed right at the time of the reform, to twelve-months, where we use one year of transaction data to construct trade links between taxpayers. The estimates remain stable across specifications, suggesting that the results do not depend on the months chosen to construct the trade linkages.

6 How Do Firms Respond to a Decrease in Withholding?

We analyze another policy change that took place during the period of our study that—unlike the reform described so far—reduced exposure to withholding for a group of firms. In September 2018 the city’s government exempted withholding by banks on the financial transactions of firms whose sales in 2017 were below AR\$10 million. This policy was part of a stimulus plan targeted at small and mid-sized enterprises during a national economic crisis. The policy measure was initially meant to last six months, although it was later extended for another six months. We use this second reform to the city’s tax collection mechanism as a quasi-experiment to test whether a reduction in the share of taxes collected via withholding leads to a fall in reported tax liability. Specifically, this reduction in withholding, combined with the threshold used to determine eligibility, allows us to implement a difference-in-discontinuity design that uses firms’ sales in 2017 to construct treatment and control groups and then compare the groups’ outcomes before and after the policy change.

Figure 11 plots the assignment rule for the policy. We take bins of firm sales in 2017 and

calculate the average share of tax withheld by banks in each bin for August and September 2018 (before and after the waiver, respectively). The share of taxes withheld by banks has no relation to the running variable in the month before the waiver. In the following month, however, there exists a decisive downward shift in the bank withholding share for firms just below the AR\$10 million threshold relative to those just above. Conditional on a firm's sales in 2017, being granted an exemption from bank withholding is as good as randomly assigned for those that are close to the cutoff, therefore, this setting allows us to address any potential endogeneity concerns that remained from the previous analysis.

Moving on to the identification of firms' responses, recall from Section 3 that noncompliant firms will try to underreport sales up to the point at which the tax liability offsets any withheld funds. If the coverage of withholding decreases, then the firm will be able to underreport more. We use a similar DD research design to the one used in Section 5 to test this hypothesis. In this case, we compare firms on either side of the waiver eligibility cutoff: firms with 2017 sales between AR\$5 million and AR\$10 million are part of the treatment group, while firms with 2017 sales between AR\$10 million and AR\$20 million are part of the control group.

Figures 12 and 13, and Table 3 panel B display the results for this approach. Figure 12 shows the first-stage results for the share of taxes withheld by banks. The share of tax collected through bank withholding falls sharply right at the time of implementation for T relative to C. During the four-quarter period that preceded the waiver, the contribution of bank withholding to turnover taxes is very similar in the treatment and control groups and displays identical patterns. After the waiver, it falls from above 20 percent to just under 10 percent for the treated firms while remaining largely unchanged for the control firms; and this effect persists throughout the length of the waiver. Figure 13a shows that there is a corresponding 2.5 percentage point decline in the reported sales of the treatment group relative to the control group after the waiver, even as the two groups track each other exactly prior to the waiver. We interpret these findings as follows: the policy led to a reduction in the coverage of third-party information for firms waived from bank withholding which gave firms more freedom to misreport their sales, allowing them to reduce their tax liability.

7 Concluding Remarks

In this paper, we studied the effects of delegating tax collection duties to firms. We analyzed the expansion of the withholding scheme used to collect the TT in the City of

Buenos Aires, Argentina, which entailed an increase in the share of taxpayers' tax liability collected indirectly at source by CAs in lieu of direct payments to the tax authority. The setting allowed us to estimate both the direct effects of expanding the withholding scheme on the firms that were tasked with collecting taxes and the indirect effects on those that now faced an increase in the amount of taxes being withheld at source.

Our results show that relying on large firms to withhold taxes from their commercial partners is an effective tax collection tool for two reasons. First, appointing firms to collect taxes did not hurt their activity. Second, substituting direct payments for withholding at source increased reported sales, which, in turn, increased tax revenue. Taken together, our findings bolster the argument of recent research by Basri et al. (2021) showing that administrative reforms can be highly effective in raising revenue and building tax capacity. Moreover, we believe that our results could have important implications for the way countries determine how to collect taxes. At the initial stages of development countries might lack the administrative and enforcement capacity to use firms as tax collectors but, as countries develop and improve their IT systems, substituting direct means of payment with indirect withholding schemes that rely on firms seems to be the right move. Our findings and the nature of the studied reforms can thus provide guidance to other middle- and low-income countries on ways to determine who the right tax collector is as a function of the level of development.

However promising our results might be, we also consider that more research is needed in order to understand the broader implications of relying on firms to collect taxes. One clear limitation of this tax collection device is the fact that not all firms will be good candidates to perform as CAs. The question that remains open is: what are the limits to appointing firms as CAs? For sufficiently small firms, the burden associated with this task will be larger than the gains faced by the tax administration.

Another aspect that must be analyzed in further detail relates to overwithholding. We stated that a taxpayer reporting a liability lower than its withheld tax was is usually regarded as an indicator of underreporting by the tax administration and that taxpayers would try to avoid this. In practice, some taxpayers may end up being overwithheld even if they are not underreporting. Anecdotal evidence suggests that tax administrations have incentives to set higher withholding rates from taxpayers that are based in other jurisdictions and, therefore, these taxpayers are more likely to build up credit in their favor. Since credits are generally hard to reclaim, taxpayers could face liquidity constraints—ultimately hampering their growth.

References

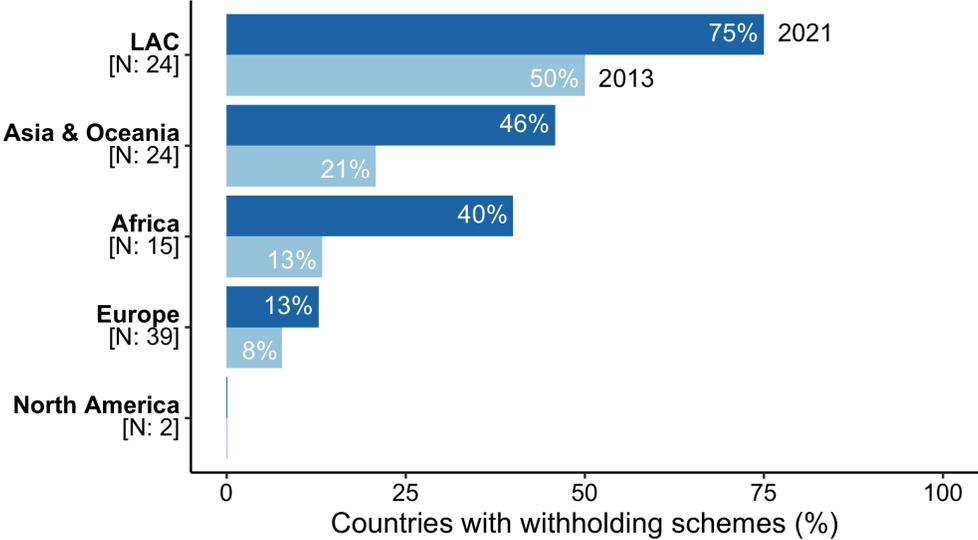
- Allingham, M. G. and A. Sandmo (1972). "Income tax evasion: a theoretical analysis". *Journal of Public Economics* 1 (3), pp. 323–338.
- Almunia, M. and D. Lopez-Rodriguez (2018). "Under the radar: The effects of monitoring firms on tax compliance". *American Economic Journal: Economic Policy* 10 (1), pp. 1–38.
- Alstadsæter, A., N. Johannesen, and G. Zucman (2019). "Tax Evasion and Inequality". *American Economic Review* 109 (6), pp. 2073–2103.
- Balán, P., A. Bergeron, G. Tourek, and J. L. Weigel (2022). "Local Elites as State Capacity: How City Chiefs Use Local Information to Increase Tax Compliance in the Democratic Republic of the Congo". *American Economic Review* 112 (3), pp. 762–97.
- Basri, M. C., M. Felix, R. Hanna, and B. A. Olken (2021). "Tax Administration versus Tax Rates: Evidence from Corporate Taxation in Indonesia". *American Economic Review* 111 (12), pp. 3827–71.
- Bergeron, A., P. Bessone, J. K. Kabeya, G. Tourek, and J. Weigel (2021). "Optimal Assignment of Bureaucrats: Evidence From Randomly Assigned Tax Collectors in The DRC".
- Bergeron, A., G. Tourek, and J. Weigel (2021). "The State Capacity Ceiling on Tax Rates: Evidence from Randomized Tax Abatements in the DRC".
- Besley, T. and T. Persson (2014). "Why do developing countries tax so little?" *Journal of economic perspectives* 28 (4), pp. 99–120.
- Best, M., J. Shah, and M. Waseem (2021). *Detection Without Deterrence: Long-Run Effects of Tax Audit on Firm Behavior*. Tech. rep.
- Best, M. C., A. Brockmeyer, H. J. Kleven, J. Spinnewijn, and M. Waseem (2015). "Production versus revenue efficiency with limited tax capacity: theory and evidence from Pakistan". *Journal of Political Economy* 123 (6), pp. 1311–1355.
- Bilicka, K. A., X. Hou, and J. Xing (Jan. 2022). *How Distortive are Turnover Taxes? Evidence from Replacing Turnover Tax with VAT*. Working Paper 29650. National Bureau of Economic Research.
- Brockmeyer, A. and M. Hernandez (2019). "Taxation, Information and Withholding: Evidence From Costa Rica". *The World Bank*.
- Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik (2017). "Rdrobust: Software for Regression-discontinuity Designs". *The Stata Journal* 17 (2), pp. 372–404.
- Carrillo, P., D. Pomeranz, and M. Singhal (2017). "Dodging the taxman: Firm misreporting and limits to tax enforcement". *American Economic Journal: Applied Economics* 9 (2), pp. 144–64.

- Gadenne, L., T. K. Nandi, and R. Rathelot (2020). “Taxation and Supplier Networks: Evidence from India”.
- Gerard, F., J. Naritomi, A. Seibold, and B. Zulian (2019). “Two-Tier Tax Systems and Firms: Evidence from Brazil”.
- Hansen, B., K. S. Miller, and C. Weber (2021). “Vertical integration and production inefficiency in the presence of a gross receipts tax”.
- Keen, M. and J. Slemrod (2017). “Optimal tax administration”. *Journal of Public Economics* 152, pp. 133–142.
- Khan, A. Q., A. I. Khwaja, and B. A. Olken (2016). “Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors”. *The Quarterly Journal of Economics* 131 (1), pp. 219–271.
- Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011). “Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark”. *Econometrica* 79 (3), pp. 651–692.
- Kleven, H. J., C. T. Kreiner, and E. Saez (2016). “Why can modern governments tax so much? An agency model of firms as fiscal intermediaries”. *Economica* 83 (330), pp. 219–246.
- Kopczuk, W., J. Marion, E. Muehlegger, and J. Slemrod (2016). “Does tax-collection invariance hold? evasion and the pass-through of state diesel taxes”. *American Economic Journal: Economic Policy* 8 (2), pp. 251–86.
- Kopczuk, W. and J. Slemrod (2006). “Putting firms into optimal tax theory”. *American Economic Review* 96 (2), pp. 130–134.
- Musgrave, R. A. (1969). “Cost-benefit analysis and the theory of public finance”. *Journal of Economic Literature* 7 (3), pp. 797–806.
- Naritomi, J. (2019). “Consumers as tax auditors”. *American Economic Review* 109 (9), pp. 3031–72.
- Pessina, L. (2020). “Who writes the check to the government does matter: Evidence from firm-to-firm links”. *Job Market Paper*.
- Pomeranz, D. (2015). “No taxation without information: Deterrence and self-enforcement in the value added tax”. *American Economic Review* 105 (8), pp. 2539–69.
- Risch, M. (2021). “Does Taxing Business Owners Affect Employees? Evidence from a Change in the Top Marginal Tax Rate”.
- Slemrod, J. (2008). “Does it matter who writes the check to the government? The economics of tax remittance”. *National Tax Journal*, pp. 251–275.
- Slemrod, J. (2019). “Tax compliance and enforcement”. *Journal of Economic Literature* 57 (4), pp. 904–54.

- Slemrod, J. and S. Yitzhaki (2002). "Tax avoidance, evasion, and administration". *Handbook of public economics*. Vol. 3. Elsevier, pp. 1423–1470.
- Slemrod, J. B. and T. Velayudhan (2018). "Do Firms Remit at Least 85 Percent of Tax Everywhere?" *Journal of Tax Administration* 4 (1).
- Waseem, M. (2020). "The role of withholding in the self-enforcement of a value-added tax: Evidence from pakistan". *The Review of Economics and Statistics*, pp. 1–44.

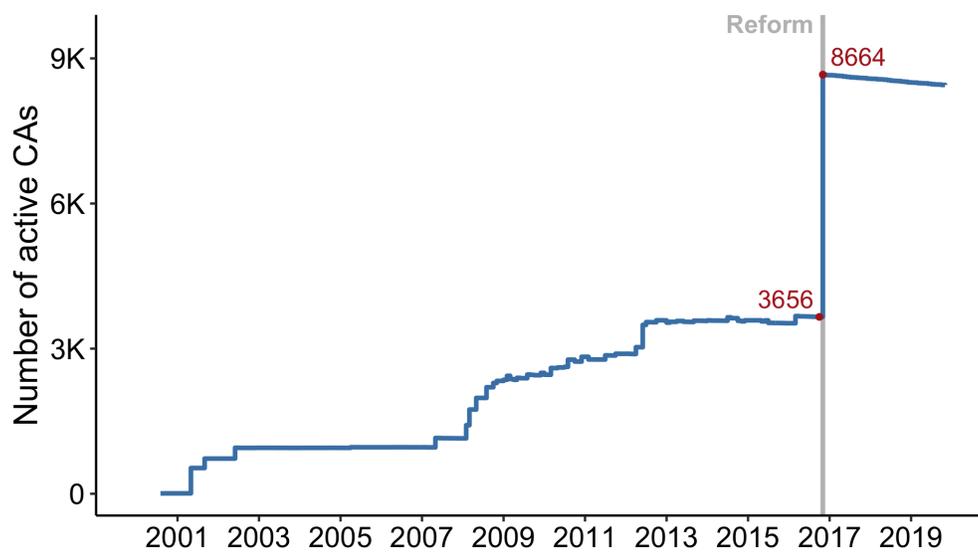
Figures

Figure 1: Withholding of indirect taxes by region



Notes: This figure shows the prevalence of tax withholding by region and its change over time. It groups countries by their geographical region and calculates the share of countries classified as having a withholding scheme for indirect taxes in 2013 (light blue bar) and 2021 (dark blue bar). The data comes from “Worldwide VAT, GST and Sales Tax Guide” published by the Ernst & Young which we processed using text-analysis techniques described in Appendix C.

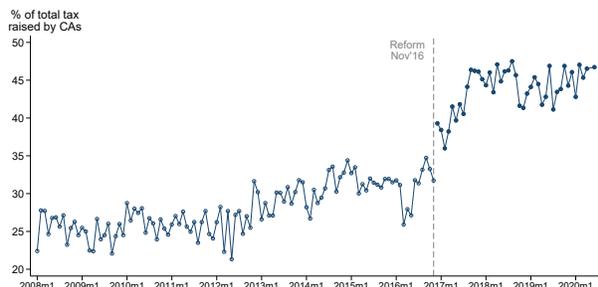
Figure 2: Number of firms acting as collection agents



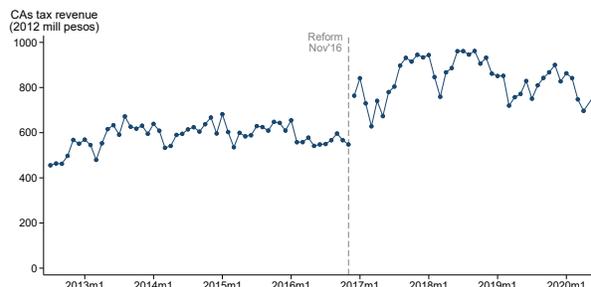
Notes: This figure shows the number of firms that are part of the collection agent (CA) register at any point in time. It was constructed using the monthly CA register published by the tax administration as public information. By comparing successive months, we obtain the entry or exit status for each firm, independent of whether it is observed in our administrative tax records. See Figure E.3 for more details on how we scraped the information. The gray vertical line corresponds to November 2016, the date at which the expansion of the withholding scheme entered into effect. Highlighted in red are the number of CAs in October 2016 and November 2016.

Figure 3: Macro evidence that documents the expansion of withholding

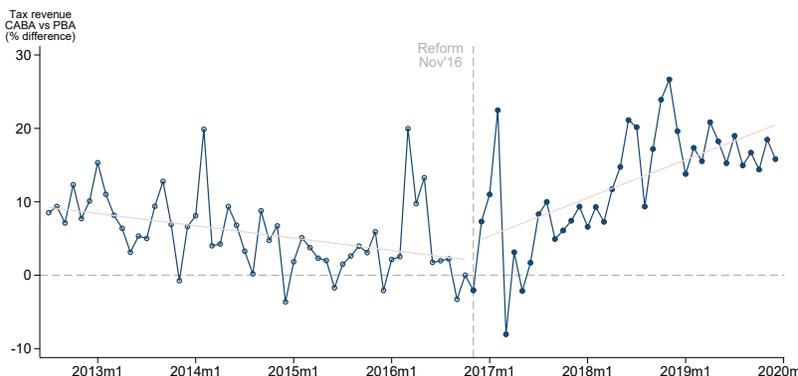
(a) Share of tax withheld by CAs



(b) Tax revenue raised by CAs



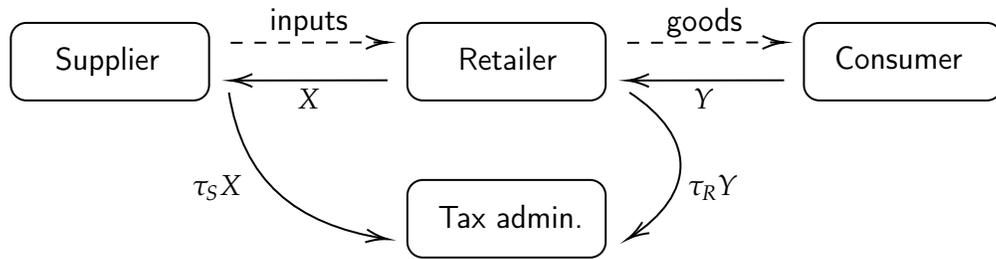
(c) Tax revenue: City of Buenos Aires vs. Province of Buenos Aires



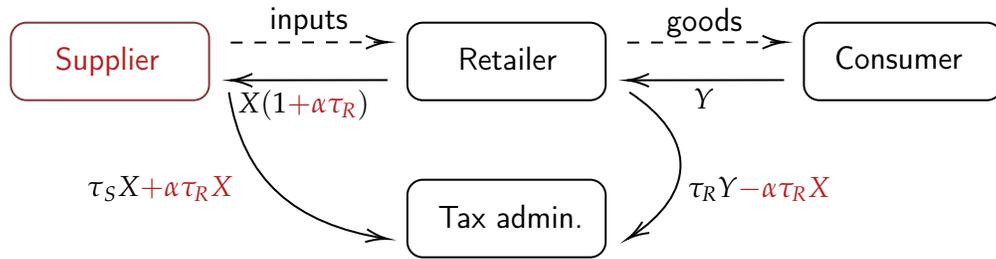
Notes: This figure documents the expansion of the withholding scheme in the City of Buenos Aires by plotting the evolution of aggregate outcomes over time. Panel (a) shows the share of total turnover tax revenue raised via withholding by collection agents (CAs). Panel (b) shows the levels of revenue raised by CAs in real terms. Panel (c) shows the percentage difference in turnover tax revenue between City of Buenos Aires and the Province of Buenos Aires relative to October 2016, the month before the reform. The graph includes a separate linear fit before and after the reform. The dashed vertical lines indicate the timing of the reform. Source: Own elaboration based on aggregate data from “Dirección General de Estadística y Censos (Ministerio de Hacienda y Finanzas GCBA).”

Figure 4: Tax collection mechanisms

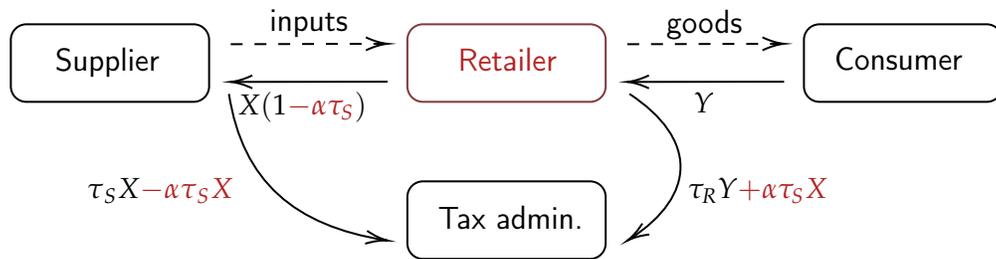
Direct payment



Withholding by seller (supplier)

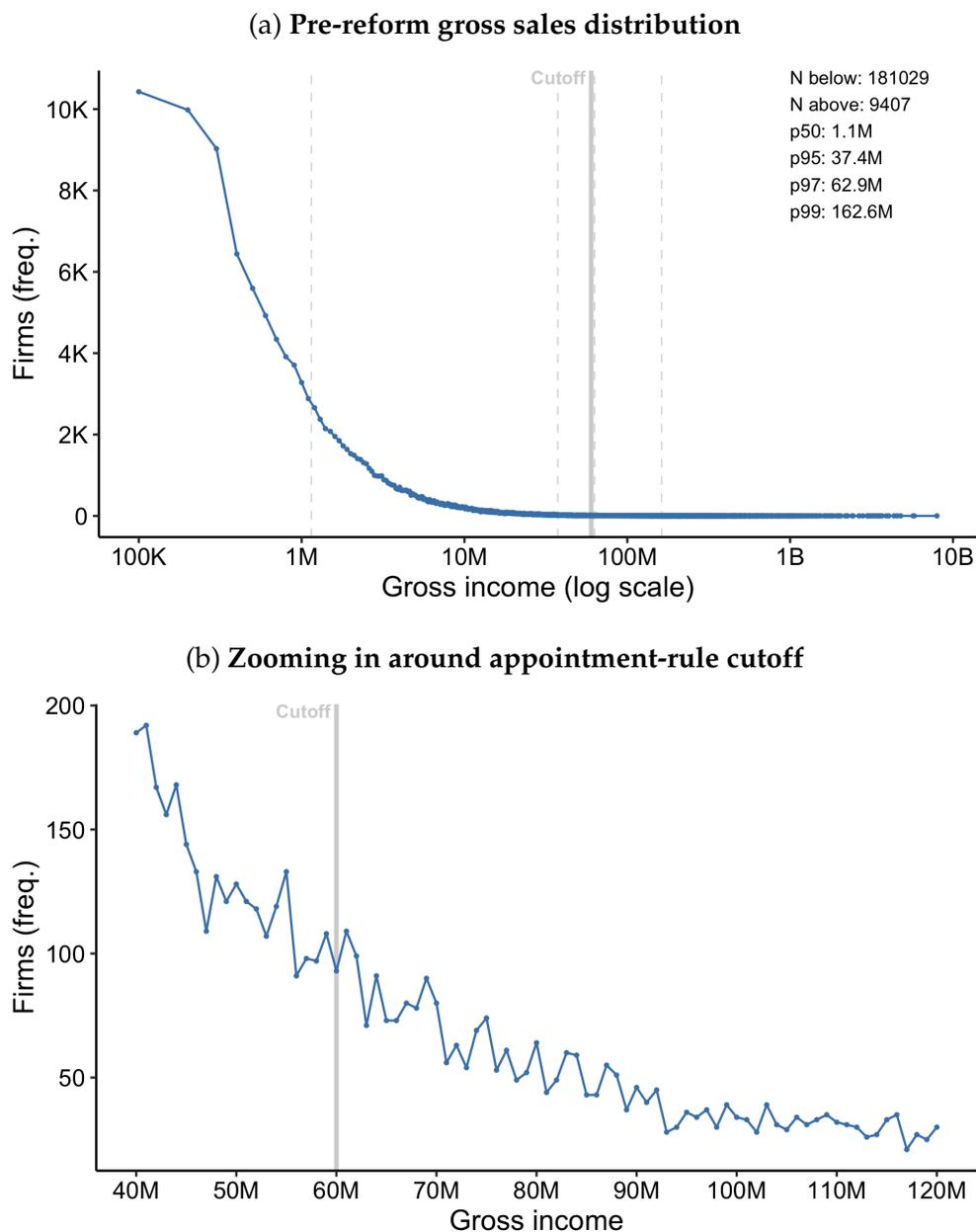


Withholding by buyer (retailer)



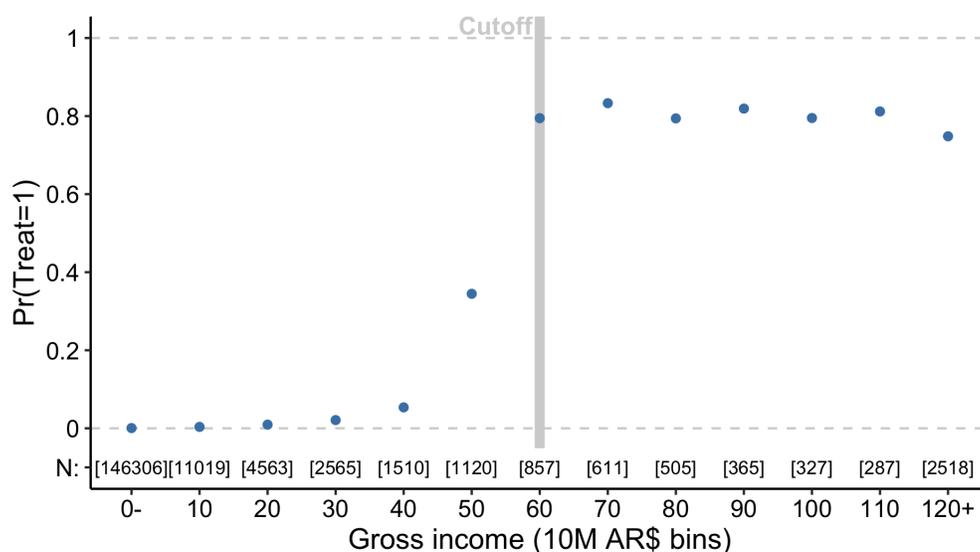
Notes: This figure summarizes the main features of the direct-payment and withholding tax collection mechanisms. The supplier and retailer have sales that are liable for the turnover tax. The first diagram represents a setting where the supplier and retailer self report their sales to the tax administration. The other two diagrams depict scenarios where either the supplier or the retailer act as a collection agent (CA). In each of these cases the party acting as CA is highlighted in red as are the amounts withheld and remitted by the CA.

Figure 5: Firm size distribution



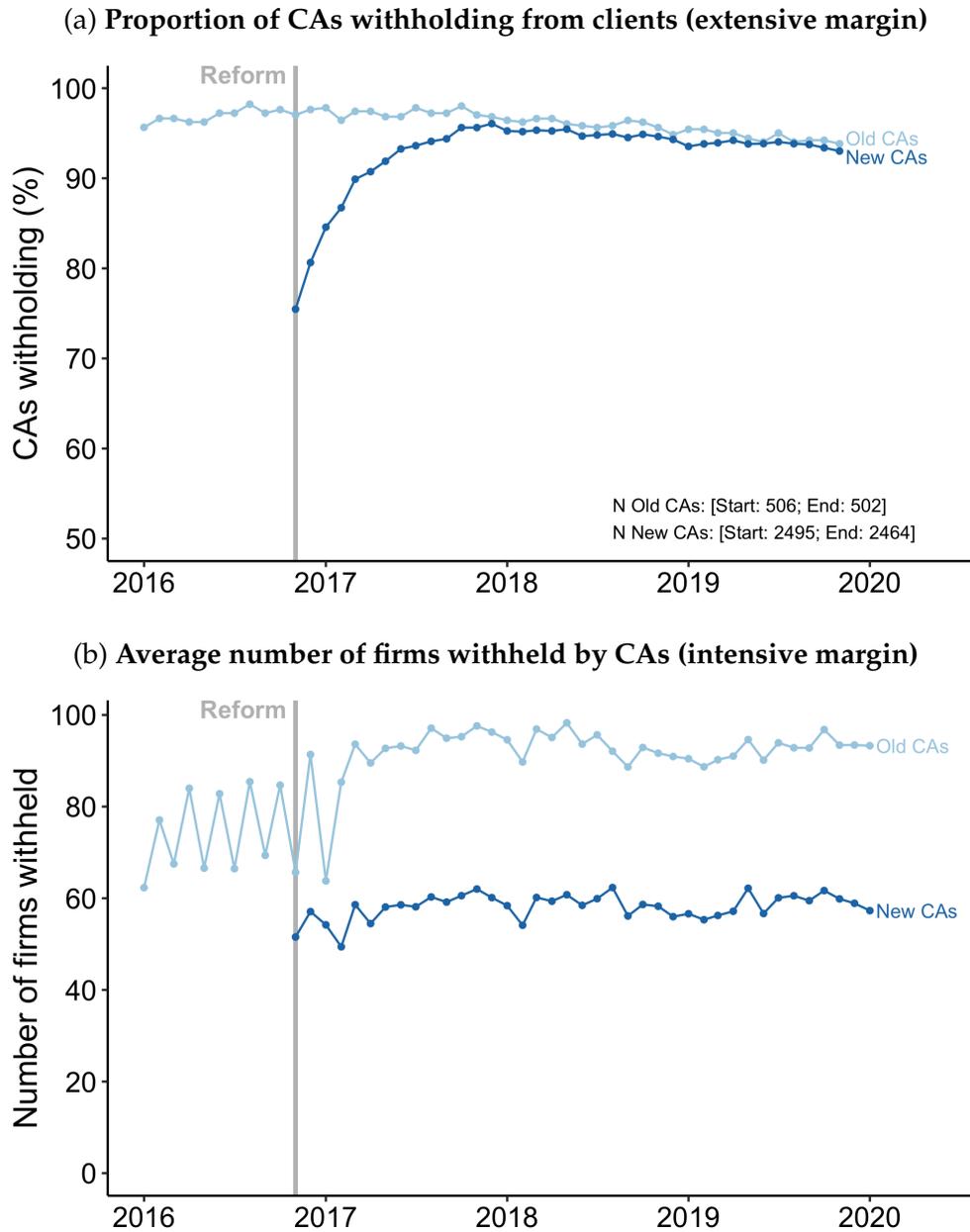
Notes: This figure shows the firm size distribution according to firm’s gross income (sales) in 2015, the same variable that was used to determine the appointment of firms into the collection agent regime in 2016. The top panel shows the distribution for firms that reported sales equal or greater than size AR\$100,000, using bins of size AR\$100,000 (~US\$8,000). The bottom panel focuses on firms that reported sales between AR\$40 million and AR\$120 million, using bins of size AR\$1 million (~US\$80,000). The solid gray lines labelled as “Cutoff” indicate the location of income threshold used to determine appointment; the dashed lines indicate p50, p95, p97, and p99.

Figure 6: Probability of appointment as collection agent



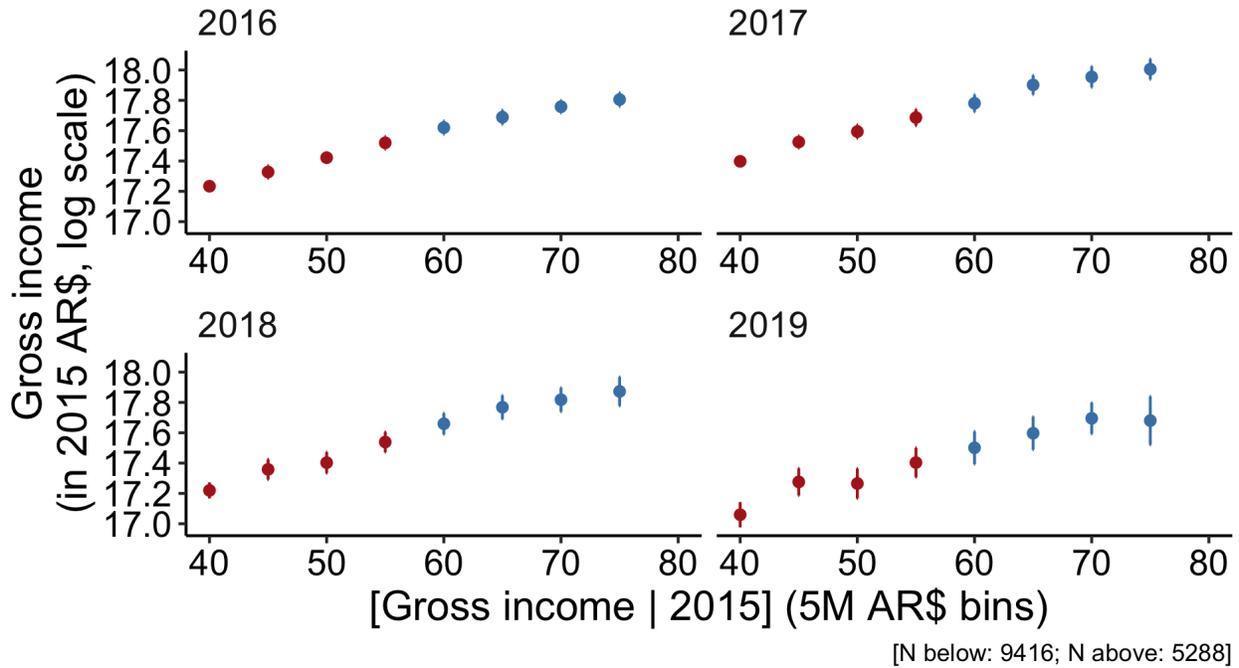
Notes: The reform enacted in November 2016 stated that firms whose gross income in 2015 was greater than AR\$60 million would be automatically appointed as collection agents (CAs). The figure shows the probability of appointment as CA by 2015 gross income bin, where each bin is of size size AR\$10 million. “Cutoff” indicates the location of the appointment threshold. The numbers in brackets at the bottom of the plot indicate the total number of firms in each bin.

Figure 7: Withholding by collection agents: extensive and intensive margins



Notes: This figure describes the behavior of newly appointed collection agents (CAs), appointed in November 2016, by comparing them to a group of old CAs, appointed in 2012. Panel (a) plots the proportion of new and old CAs that are withholding from commercial partners in each month. Panel (b) shows the average number of firms that get withheld from new and old CAs in each month. The gray vertical line corresponds to November 2016, the date at which the expansion of the withholding scheme entered into effect and, therefore, the first month of activity for new CAs.

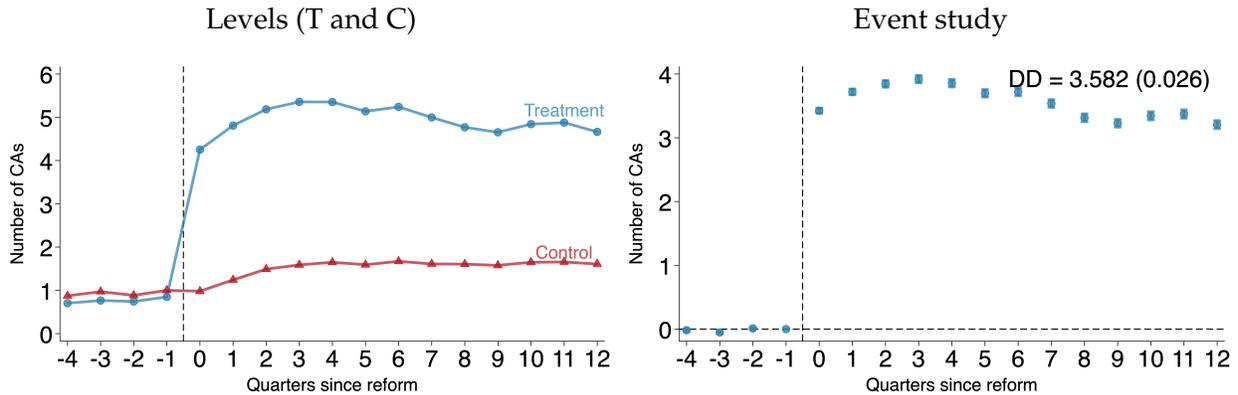
Figure 8: Gross income of newly appointed collection agents vs. non-eligible firms



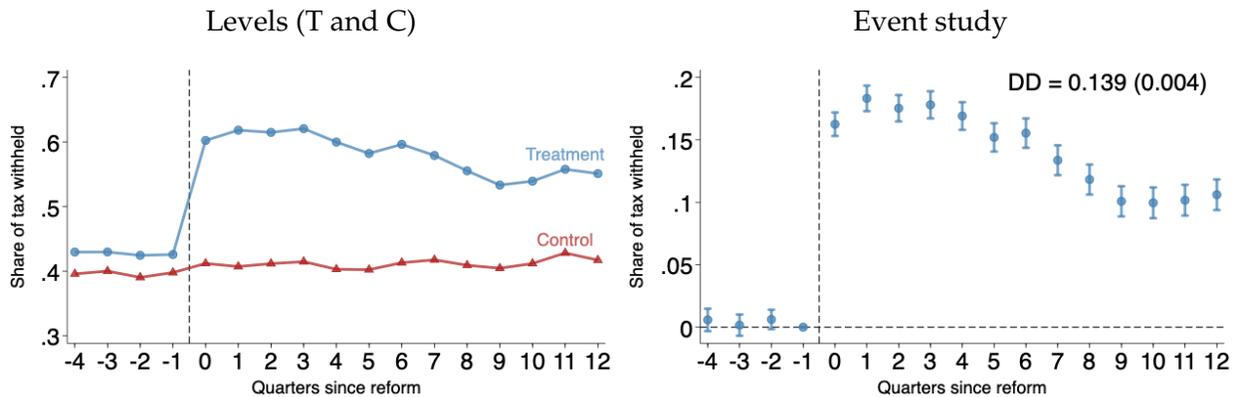
Notes: This figure shows the reform’s impact on annual sales for firms whose gross income (sales) in 2015 were in the vicinity of the appointment threshold. Each panel displays log annual sales (measured in 2015 pesos) for eight equally spaced bins (width AR\$5 million) of the running variable (2015 sales). The blue dots represent firms above the appointment threshold, eligible to become CAs. The red dots correspond to firms below the appointment threshold, non-eligible firms. The top left panel corresponds to 2016 and excludes the two treated months (November and December), hence, it can be interpreted as a placebo test because it only reflects pre-reform sales. The remaining panels show annual sales for 2017, 2018, and 2019. The sample is a balanced panel of firms. We exclude outliers (growth rate below percentile 1 or above percentile 99) from our calculations.

Figure 9: Event Study plots: first stage

(a) Number of collection agents



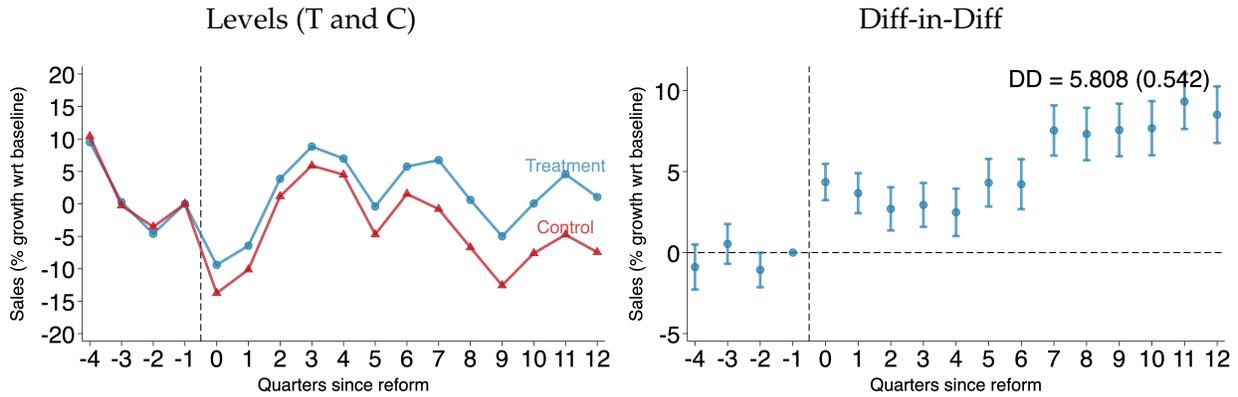
(b) Share of tax withheld by CAs



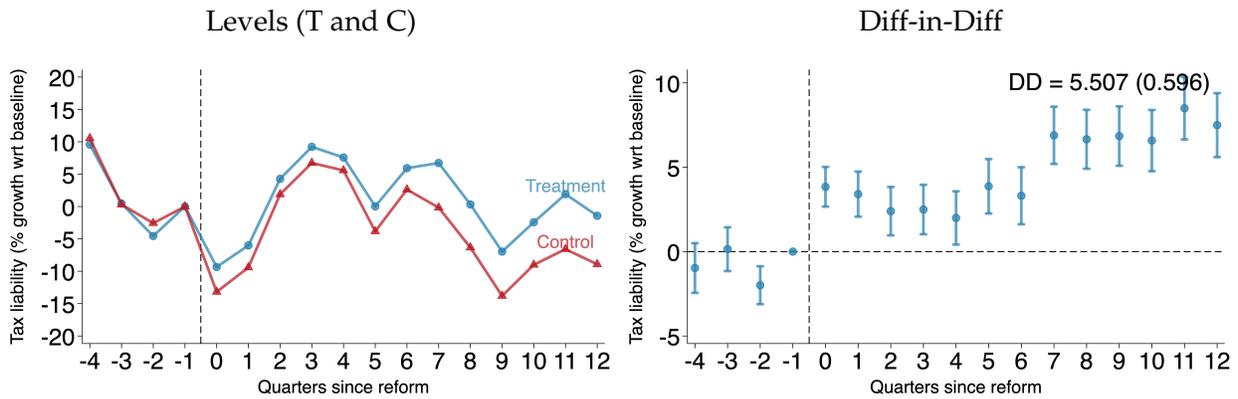
Notes: This figure compares the first stage outcomes for firms connected to newly appointed collection agents (treatment group, T) and firms that are not connected to newly appointed collection agents (control group, C). Each panel shows the regression estimates for a different outcome. The plots on the left correspond to the regression of the outcome of interest on the full set of firm and quarter fixed effects, dropping the dummy for the baseline period (run separately for the two groups of firms). These plots show the coefficients on the time dummies plus the constant, where each coefficient represents the average outcome within-firm relative to the baseline. The plots on the right correspond to the event study regression, equation 3, where each coefficient represents the relative difference between the two groups in the given quarter and the vertical spikes correspond to the 95% confidence interval. The DD estimates found in Table 3 are displayed at the top right corner of these figures. All regression specifications contain a balanced panel of firms that filed taxes in at least six months every year. The dashed vertical line indicates the timing of the reform.

Figure 10: Event Study plots: behavioral response

(a) Self reported sales

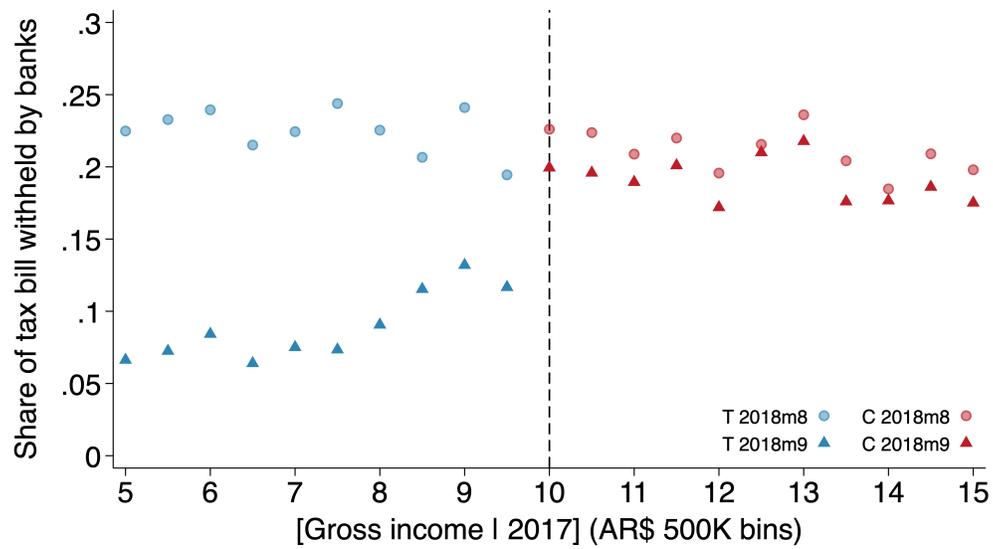


(b) Tax liability



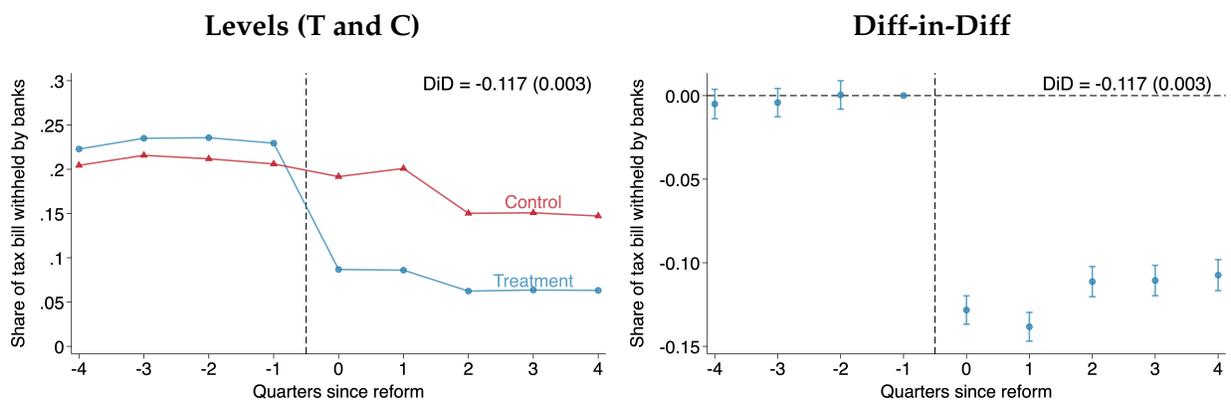
Notes: This figure compares the response on self reported sales and tax liability for firms connected to newly appointed collection agents (treatment group, T) and firms that are not connected to newly appointed collection agents (control group, C). Each panel shows the regression estimates for a different outcome. The plots on the left correspond to the regression of the outcome of interest on the full set of firm and quarter fixed effects, dropping the dummy for the baseline period (run separately for the two groups of firms). These plots show the coefficients on the time dummies plus the constant, where each coefficient represents the average outcome within-firm relative to the baseline. The plots on the right correspond to the event study regression, equation 3, where each coefficient represents the relative difference between the two groups in the given quarter and the vertical spikes correspond to the 95% confidence interval. The DD estimates found in Table 3 are displayed at the top right corner of these figures. All regression specifications contain a balanced panel of firms that filed taxes in at least six months every year. The dashed vertical line indicates the timing of the reform.

Figure 11: Tax withheld by banks



Notes: In September 2018 withholding by banks was waived for firms whose sales in 2017 were below AR\$10 million. This figure takes bins of firm sales in 2017 (the policy’s running variable) and calculates the average share of tax withheld by banks in each bin. Bins below the threshold are colored in blue; bins above the threshold are colored in red. The circles correspond to August 2018, the month before the policy change, and the triangles correspond to September 2018, when the policy change took effect.

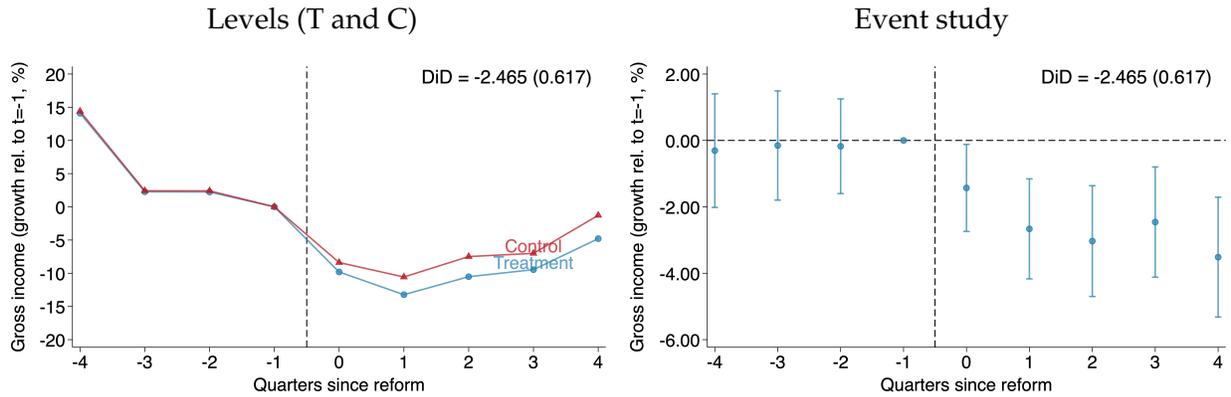
Figure 12: Event Study plots: Tax withheld by banks



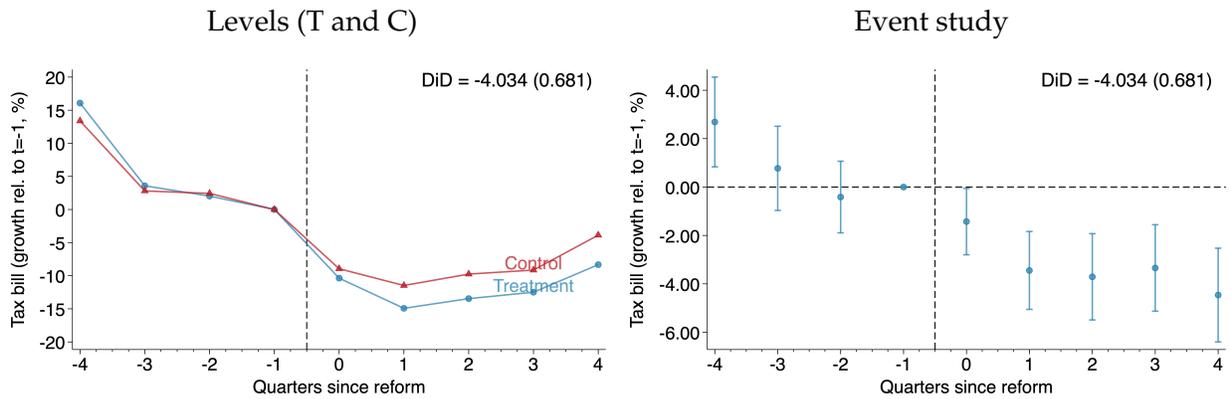
Notes: This figure compares the first stage outcome for firms below the 2017 sales threshold used to determine eligibility for the waiver on bank withholding (treatment group, T) and firms above the threshold (control group, C). The plot on the left corresponds to the regression of the outcome of interest on the full set of firm and quarter fixed effects, dropping the dummy for the baseline period (run separately for the two groups of firms). This plot shows the coefficients on the time dummies plus the constant, where each coefficient represents the average outcome within-firm relative to the baseline. The plot on the right corresponds to the event study regression, where each coefficient represents the relative difference between the two groups in the given quarter and the vertical spikes correspond to the 95% confidence interval. The DD estimates found in Table 3 are displayed at the top right corner of these figures. All regression specifications contain a balanced panel of firms that filed taxes in at least six months every year. The dashed vertical line indicates the timing of the reform.

Figure 13: Event Study plots: behavioral response

(a) Self reported sales



(b) Tax liability



Notes: This figure compares the response on self reported sales and tax liability for firms below the 2017 sales threshold used to determine eligibility for the waiver on bank withholding (treatment group, T) and firms above the threshold (control group, C). Each panel shows the regression estimates for a different outcome. The plots on the left correspond to the regression of the outcome of interest on the full set of firm and quarter fixed effects, dropping the dummy for the baseline period (run separately for the two groups of firms). These plots show the coefficients on the time dummies plus the constant, where each coefficient represents the average outcome within-firm relative to the baseline. The plots on the right correspond to the event study regression, where each coefficient represents the relative difference between the two groups in the given quarter and the vertical spikes correspond to the 95% confidence interval. The DD estimates found in Table 3 are displayed at the top right corner of these figures. All regression specifications contain a balanced panel of firms that filed taxes in at least six months every year. The dashed vertical line indicates the timing of the reform.

Tables

Table 1: Summary statistics

	No. firms	Gross revenue (1,000 of AR\$)				Tax liability (1,000 of AR\$)			
		p5	p50	mean	p95	p5	p50	mean	p95
Panel A: Full sample									
Firms	183,503	0	42	249	1,052	0	1	6	28
Collection agents	9,366	3	2,757	19,209	39,947	0	85	942	1,451
Panel B: Estimating sample									
Firms	80,208	7	123	415	1,664	0	3	11	45

Notes: This table reports summary statistics calculated in the period January–October 2016, before the expansion of the withholding regime. Gross revenue and tax liability are expressed in thousands of 2016 Argentinian pesos. Panel A corresponds to the full sample. Panel B corresponds to the sample used to estimate firm responses to withholding, as described in Sections 5 and 6.

Table 2: Regression discontinuity estimations

	First stage	Reported sales			
	(1)	2016 (2)	2017 (3)	2018 (4)	2019 (5)
RD estimate	0.323*** (0.054)	0.845 (0.588)	0.555 (0.621)	0.839 (0.571)	0.554 (0.559)
Number of firms	4,098	3,235	3,195	3,151	3,102

Notes: This table reports the first-stage and the 2SLS estimates from a fuzzy regression discontinuity (RD) specification. The running variable is annual gross income in 2015. The reform to the withholding system enacted in November 2016 stated that firms whose gross income in 2015 was greater than AR\$60 million (with some exceptions) would be automatically appointed as collection agents (CAs). Column 1 shows the first-stage change in the probability of a firm above the threshold being appointed as a CA in November 2016. This first stage is graphically shown in Figure 6. Columns 2 through 5 correspond to the logarithm of reported annual sales. Column 2 excludes the two treated months of 2016, November and December. The results of this column can be interpreted as the result of a placebo test, as they only reflect pre-reform sales. *, **, and *** denote statistical significance at the 10%, 5% and 1% levels. The regression is estimated for the range [40, 60) to the left of the cutoff, and [60, 80) to the right of the cutoff using a linear approximation and a triangular kernel. RD estimates are computed with the `rdrobust` routine from Calonico et al., 2017.

Table 3: Difference-in-differences estimations

	First stage (1)	Reported sales (2)	Tax revenue (3)
Panel A: Increase in withholding			
<i>Treat · Post</i>	0.139*** (0.004)	5.508*** (0.542)	5.507*** (0.596)
Observations	1,223,156	1,223,156	1,223,156
Number of firms	75,846	75,846	75,846
\bar{y}_{t-1}	11.623	965.219	33.437
Panel B: Decrease in withholding			
<i>Treat · Post</i>	-0.108*** (0.003)	-2.798*** (0.701)	-4.285*** (0.776)
Observations	142,913	142,913	142,913
Number of firms	16,006	16,006	16,006
\bar{y}_{t-1}	7.236	1156.487	33.105

Notes: This table reports the first-stage change in withholding and the reduced-form estimates from a difference-in-differences specification. Column 1 shows the first-stage results. For Panel A, this corresponds to the share of taxes withheld by CAs, while for Panel B it corresponds to the share of taxes withheld by banks. Columns 2 and 3 correspond to the gross income and the tax revenue growth relative to the baseline period (third quarter of 2016 and third quarter of 2018, respectively). The last row of each panel shows the average value of the variable reported in columns at baseline. Standard errors clustered at the firm level are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5% and 1% levels.

Table 4: Difference-in-differences estimations

	6-months (1)	1-month (2)	3-months (3)	9-months (4)	12-months (5)
First stage					
Treat · Post	0.139*** (0.004)	0.153*** (0.004)	0.144*** (0.004)	0.139*** (0.004)	0.139*** (0.004)
Reported sales					
Treat · Post	5.808*** (0.542)	4.404*** (0.489)	5.471*** (0.514)	5.938*** (0.560)	6.476*** (0.576)
Tax revenue					
Treat · Post	5.507*** (0.596)	3.994*** (0.527)	5.408*** (0.558)	5.504*** (0.621)	5.884*** (0.640)
Observations	1223156	903935	1103822	1301080	1356128
Number of firms	75846	55770	68339	80751	84181

Notes: This table investigates robustness of the regression in Table 3 Panel A with respect to the number of months used to estimate baseline trade linkages. Column 1 replicates Panel A of Table 3, columns 2 to 5 repeat the specification changing only the number of months used to estimate these linkages. The last row shows the number of observations and firms in each specification. Standard errors clustered at the firm level are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5% and 1% levels.

Appendixes

A Data construction

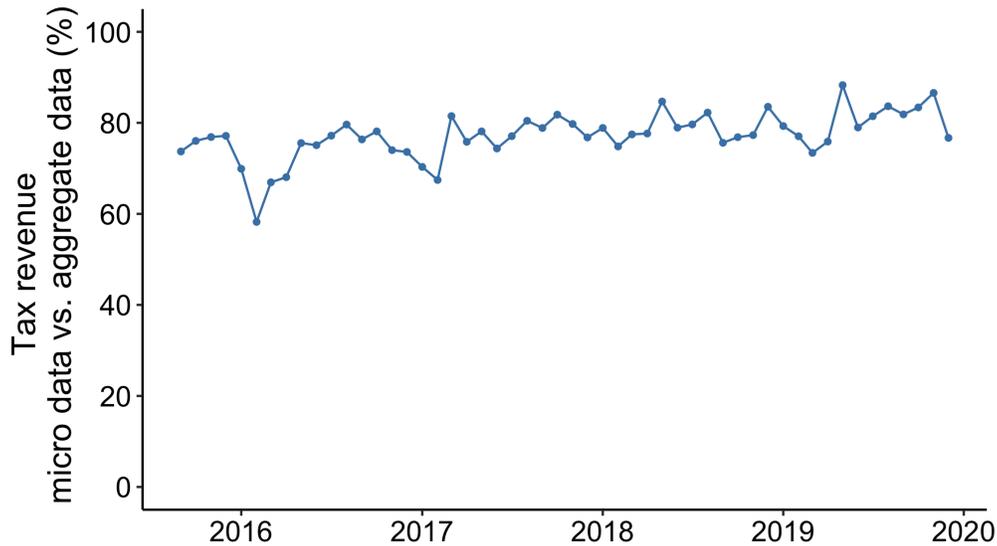
As stated in Section 2.2, our main data source consists of *nearly* the entire universe of firms operating in the City of Buenos Aires in the period 2015-2020. Obtaining the universe of taxpayers was infeasible due to computing power limitations from the tax administration. In this Section we describe the process undertaken for the construction of our data.

As one of the key components of our study was the use of commercial links between CAs and their commercial partners, we requested all transactions linked to CAs. However, the invoice data are defined at the CA-client-month level, and providing all line items proved to be too burdensome for the tax administration's servers. To circumvent this issue, we requested data from a random sample of CAs: out of the approximately 8000 CAs, we obtained data for 3000 CAs. So, in practice, our data consists of a sample of CAs and the universe of taxpayers linked to them.

The data construction process followed the steps outlined next. First, define the list of CAs on duty in 2017 and select 3000 CAs randomly—2500 newly-appointed in 2016, 500 appointed prior to 2016. Next, for this subset of CAs, take all their transaction-level declarations. Using these declarations we identify all customers who had a transaction with the CAs and build a list of their tax IDs. Finally, for all the IDs in our list plus all the IDs of our sample of CAs, we obtain the monthly tax filings.

How do our data compare to the universe of tax filings? Figure A.1 shows that the tax revenue collected from our sample is very close to the total revenue—turnover tax revenue calculated from the micro data amounts to approximately 80% of the total tax revenue throughout the whole period of our study. Therefore, although we only have a sample of taxpayers, in practice, they explain the majority of the commercial activity captured by the tax. The taxpayers that we do not observe are likely small firms who do not have transactions with CAs.

Figure A.1



Notes: This figure compares turnover tax receipts calculated using our randomly-selected sample from the micro data relative to the total tax revenue reported by the government in monthly reports. Source: Own elaboration based on aggregate data from “Dirección General de Estadística y Censos (Ministerio de Hacienda y Finanzas GCBA).”

B Turnover Tax: Further Details

The Turnover Tax (TT)—also known as gross receipts tax—is an indirect provincial tax imposed in each of the twenty-four jurisdictions in Argentina (twenty-three provinces and the City of Buenos Aires). It is levied on the gross sale of goods or services with no possibilities for deducting costs. It therefore applies to all the transactions taking place in the supply chain—that is, business-to-business and business-to-consumer transactions. This type of tax is simple to collect but creates “cascading effects” whereby final goods are taxed multiple times throughout production.¹ The TT represents the main source of own revenue in all the jurisdictions (about 75% of total tax receipts in the City of Buenos Aires).² Each jurisdiction has its own subnational tax laws and TT regulations regarding

¹In environments with limited enforcement capacity this type of tax provides limited opportunities for firms to engage in tax evasion relative to profit-based taxes like the VAT and corporate income tax (Best et al., 2015).

²The other taxes levied at the provincial level are the stamp tax, property tax, vehicle tax, and other minor duties.

the tax base, tax rates, apportionment rules, collection schemes, exemptions, and promotional regimes applicable to each activity.

Tax rates. The tax rates typically vary from 3.5% to 5% for the sale of goods and services depending on the taxpayer activity, annual turnover from the previous year, and location where the transaction takes place (inside versus outside the province). For example, in 2016, retailers with 2015 revenue less than forty-nine million pesos faced a 3% tax rate on total sales. This rate increased to 4% if their 2015 revenue was greater than forty-nine million pesos (that is, it operates as a notch rather than a kink). Importantly, firms selling out of province might be taxed at higher tax rates (as if there were internal customs). This type of firm is informally called a “foreign jurisdiction” firm. When filing taxes, multiactivity firms must discriminate their tax base by the types of activity they carry out and apply the corresponding tax rate in each case. For example, a large footwear manufacturer that also sells shoes to final consumers could be taxed at 1% for its manufacturing sales and 4% for its retail sales.

Local versus multiprovince firms. Taxpayers are classified as *local* when they only sell goods or services within the jurisdiction (for example, a corner shop) or *multiprovince* when they operate both inside and outside the jurisdiction (for example, a manufacturer from another province selling goods in the City of Buenos Aires). These firms are part of the so-called Multilateral Agreement.

Apportionment formula. For firms operating in multiple provinces, there exists an apportionment system to distribute the tax base across places. The formula is based on a sales factor and an expenditure factor determined at year $t-1$ (called “unified coefficient”). In particular, 50% of total sales are distributed according to the ratio of a taxpayer’s sales in the taxing province to its overall sales; and the other 50% are distributed according to the ratio of a taxpayer’s expenditures in the taxing province to its overall expenditure.

How is the TT collected. The TT can be paid directly by taxpayers (standard method) or collected indirectly by third parties (withholding method). In the standard method, taxpayers log on to the tax administration’s website at the end of the month, self-report the gross revenue accrued in each activity over that month; the system applies the corresponding tax rate and generates a payment coupon. Under the withholding method, the tax administration can designate firms, banks, or credit and debit card companies to operate as collection agents (CAs). In the case of firms as CAs, every time a taxpayer purchases inputs from a CA or sells goods or services to a CA, part of the invoice amount is withheld by the CA and remitted to the tax authority. In the case of banks, every time a payment is deposited in the bank account, the bank withholds part of it and remits the

funds to the tax authority. In the case of card companies, every time a transaction is made with a debit or a credit card, the card company withholds part of it. In all these cases, taxpayers still have to file taxes at the end of each month but the withheld amount constitutes a credit in favor of the taxpayer who can discount it from the tax liability before making any payments.

Withholding and reverse withholding. Withholding can operate both downstream and upstream. Upstream withholding is similar to that used for the collection of personal income tax in the US. In this case the payer in the transaction (called “retaining agent”) withholds a fraction of the total amount of the sale and remits it to the tax administration in concept of future tax payments from the payee. In the downstream case, the supplier (called “perceiving agent”) withholds from the payer by adding the withheld tax to the total sale and remits it to the tax administration in concept of future tax payments from the payer. The remitted funds are accumulated in the withheld taxpayer’s account and can be used to deduct from future outstanding tax liabilities. Importantly, in all these cases, the CA must report very detailed information to the tax authority: the unique identifiers of the firms it traded with, the total amount of each transaction, and the withheld amount. This information is automatically available for taxpayers when they file their monthly tax returns using the official processing software. If a CA fails to report a transaction, then taxpayers can report it themselves so that they can claim the advance payment back. When that happens, CAs can face severe penalties. Hence, CAs have strong incentives to perform their collection duties accurately.

Overwithholding. In practice, the sum of withheld amounts could be smaller or greater than the tax liability. On the one hand, not all transactions are subject to withholding and, generally, the withholding rate is lower than the statutory tax rate. This implies that taxpayers are usually able to compensate a fraction of the tax liability with the advance payments withheld at source. On the other hand, it could happen that the monthly withheld amount exceeds the monthly liability, thus generating a credit balance in favor of the taxpayer that can be rolled over for future tax payments. These funds cannot easily be withdrawn and are essentially an interest-free loan from taxpayers to the tax authority. The main reason why this could happen is that credit card companies and banks also act as CAs. So, a given transaction between two trading partners might be subject to withholding multiple times (for example, by a supplier, by a credit card company, and finally when the money is deposited in a bank account).

TT collection over time. Until the early 2000s, the tax was basically collected via direct monthly payments. From that point on, the tax administration started appointing firms to

act as CAs on an ad hoc basis—usually targeting the biggest firms in the economy. As the tax authority developed its IT system, the appointment of firms became more widespread and eventually expanded their net in a discrete way as was shown in Figure 2. That is precisely the main reform we exploit in our analysis.

Who becomes a CA? The 2016 reform we analyzed replaced the ad hoc way of appointing CAs with a simple revenue-based rule: firms whose 2015 revenue was greater than sixty million pesos were automatically enrolled as CAs (Resolution 364/AGIP/2016). These firms could act as both *perceiving agent*, when selling goods or services, and *retaining agent*, when purchasing goods or services.

When a firm is enrolled as a CA, it must start using a separate tax collection software that incorporates withholding into its sales. This software has an updated register of all firms that are liable to be withheld and also determines the rate that should be applied. The withholding rate is applied to the amount before taxes. Firms must transfer the withheld funds every fifteen days along with an annex containing the details of every transaction: tax identifier of the withheld party, total transaction amount, and amount perceived or retained. Anecdotally, given that appointed firms are large and formal, they do not find it very complicated to adapt to this new system.

Federal taxes. Firms are also subject to federal taxes. Depending on the firms' revenue, they have to file a monthly VAT return or be part of a simplified regime. The VAT base is total value added, which implies that firms can deduct purchases from their total revenue. On the other hand, firms in the simplified regime pay a fixed amount depending on the threshold determined by gross income, the same variable used to determine the tax rate and tax liability of the TT.

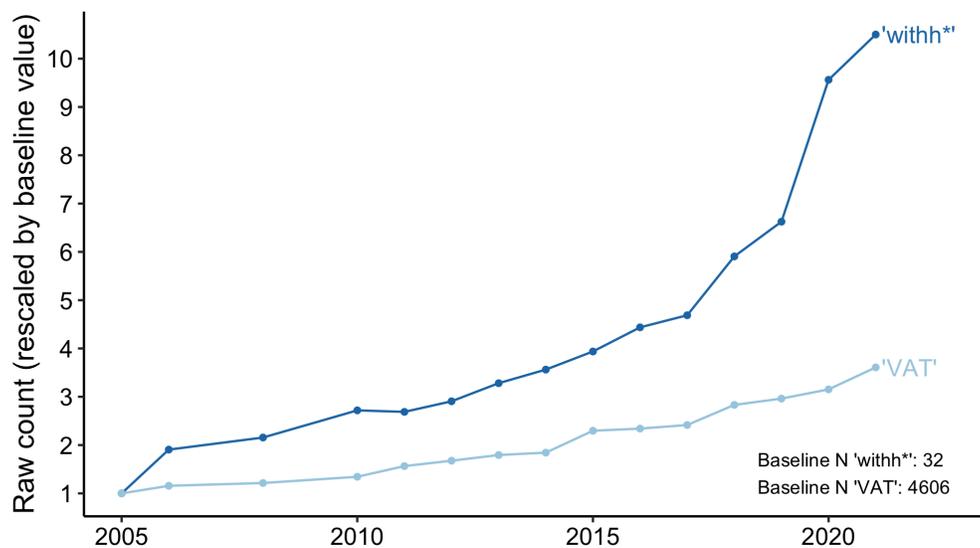
C Text Analysis Using Ernst & Young's Guide for Indirect Taxes

We explain in more detail the process followed to construct the database used in Figure 1. We started by downloading Ernst & Young's "Worldwide VAT, GST and Sales Tax Guide". We split each of these guides into country chapters where possible (2013 onward). Then, we created a text-analysis program to scrape information from each chapter. For each country in the guide and every year available, the program searched for terms related to "tax withholding." In particular, we searched for strings containing the term "withh*" (that is, "withholding," "withheld," "withhold," "withholdings," "withholder," etc.). The

program created a binary indicator equal to 1 if a country’s chapter contained any of these terms. As a benchmark, we repeated the text-analysis exercise by searching for terms related to the VAT, since this is also a widespread tax instrument.

Figure C.1 below shows the raw counts (scaled by the 2005 value) of the total number of matches per guide (that is, the sum across countries). The dark blue line corresponds to terms related to withholding and the light blue line to VAT. We observe a sharp increase in the number of matches over time. Moreover, the trend accelerates from 2018 onward. This suggests that withholding regimes are becoming more widespread. Moreover, from Figure 1, we can see that this is particularly pronounced in low- and middle-income countries, which usually have lower tax capacity. The results of our paper are encouraging and suggest that this might be a good development of tax systems.

Figure C.1: Number of matches per document: “withh*” vs. “VAT” (used as benchmark)



Notes: This figure summarizes the results from our text-analysis exercise. It plots the time series of the number of times that our program detects terms related to withholding (dark blue) or VAT (light blue). Each series is rescaled by the level in 2005. Source: Own elaboration based on Ernst & Young. Our R code is available upon request.

D Documenting the Reforms with Raw Data

What do the withholding reforms look like in the raw administrative data? In this section we document the variation used in the paper in more detail and without imposing any sample restrictions. That is, we simply open the databases as they were shared with us and in a couple of lines of code we compute summary statistics and time series that capture the nature of the identifying variation. The goal of this exercise is twofold. First, it adds *transparency* to the paper and the research process. Second, it is a way of showing the *aggregate relevance* of the analyzed reforms. That is, we consider *all* the firms in the data instead of only those better suited for our empirical strategies (for example, those linked to new versus old CAs, those above and below the relevant cutoffs). Reassuringly, even without imposing any sample restrictions, we document sharp changes in several withholding measures right at the time of the reforms. The facts that we document also help alleviate concerns about the quality of the data and the speed of implementation or take-up of the reforms.

We provide graphical evidence to shed light on the variation used in the paper. Recall that we exploit two reforms: an increase in withholding for firms linked to newly appointed CAs (November 2016) and a decrease in withholding from banks (September 2018). There is a third reform in January 2020 that further expanded the net of withholding agents by updating the parameters from the first reform (Resolution 296/19 AGIP). We decided to ignore this third reform in the main analysis of the paper due to its closeness to the COVID-19 pandemic. Nevertheless, for completeness we document its variation. In the next set of figures we denote these three reforms with vertical red lines.

Figure D.1 shows descriptive evidence of the relationship between total tax withholdings and tax liability for the period September 2015 to December 2020. This graph includes the zeros and thus captures both extensive- and intensive-margin changes in withholding. Panel (a) plots the ratio of total tax withholdings to tax liability. The numerator includes withholding from three sources: (i) commercial partners and card companies, (ii) banks, and (iii) the customs office. We report the *median* of the ratio to get a better sense of the distribution of firms and to prevent some large outliers from spoiling the figure. Relatedly, Panel (b) plots the fraction of firms whose monthly withholdings exceed the tax liability.³

³In practice, firms can be overwithheld when (i) they are not registered in the withholding register of a province and the commercial partner thus must apply a withholding rate that can be twice as large as the tax rate, or (ii) a lag occurs between the sales of a good or service and the month in which the bank or credit card company reports the withholding. For instance, it is common for firms to have substantial sales in a month but the money credited in the bank account in the following month, generating a time mismatch between the tax liability and the withheld amount.

Both panels exhibit discrete jumps right at the time of the reforms. The share of tax withholdings to tax liability increases by about 20 p.p. in November 2016, decreases by more than 20 p.p. in September 2018, and increases again by about 20 p.p. in January 2020. Likewise, the share of overwithheld firms increases from 23% to 30% in November 2016, decreases to 24% in September 2018, and increases again to 31% in January 2020.

The patterns from Figure D.1 mask substantial heterogeneity. We next decompose these patterns into the extensive and the intensive margin. This is reported in Figures D.2 and D.3, respectively. In addition to this, we unpack total withholding into its sources: (i) withholding by firms (for example, when trading with CAs or credit card companies) and (ii) withholding by banks (for example, when receiving deposits or making transfers).⁴ We also break the former into (i.a) withholding performed in the *sale of goods or services* and (i.b) reverse withholding made in the *purchase of inputs*. As explained below, we conclude that (i) the increase in withholding in Figure D.1 after the *first* reform is mostly explained by a combined extensive- and intensive-margin increase of reverse withholding in the purchase of inputs, and (ii) the sharp decrease in withholding after the *second* reform is entirely driven by an extensive-margin decrease in bank withholding.

Figure D.2 documents the extensive margin of withholding. We plot the share of firms whose taxes are withheld at source by type of withholding. Panel (a) suggests that although withholding is prevalent (blue dots), there is substantial variation across withholding types. Before the initial reform, about 60% of the firms were subject to withholding by firms (brown dots) and 50% by banks (green dots). Importantly, right after the first reform, the brown dots increase from 60% to 70% but the green dots do not change. This is reassuring, as the first reform only increased the number of firms operating as withholding agents—not banks. In contrast, after the second reform, which exempted a subgroup of firms from bank withholding, the green dots exhibit a sharp large decrease from 48% to 21% but the brown dots remain stable at 70%. When we further break down the withholding by firms (brown dots) in Panel (b), we observe a similar extensive-margin pattern for the withholding of sales (light blue) and purchases (light red) in the first reform.

Last, Figure D.3 documents the intensive margin of withholding over time. In this case, we plot the share of tax withholdings to tax liability for firms with positive withholding in each month separately. To capture responses at the upper end of the distribution, we compute the *average* rather than *median* and we exclude the largest outliers. Panel (a) focuses on the withholding by firms and Panel (b) on the withholding by banks. The top panel suggests that most of the response to the initial reform—the one that expanded the

⁴We omit withholdings by the customs office because it represents a tiny share of the total.

number of CAs—comes from withholding in purchases (light red dots) rather than in sales (light blue dots). Intuitively, when firms *purchase* intermediate inputs from upstream CAs, the invoice is *increased* by the amount of the withholding (it appears as an extra line in the receipt). When firms *sell* goods or services to downstream CAs, the invoice is *reduced* by the amount of the withholding (the seller receives a separate proof of payment for the withholding portion). With these two possibilities in mind, the figure might imply that the November 2016 reform likely appointed very large *upstream* CAs that were selling to downstream firms. This would explain the sharp increase of the light red dots and the relatively smooth pattern of the light blue dots. Finally, Panel (b) shows a smooth, flat evolution in the share of bank withholdings to tax liability. This is reassuring, as the September 2018 reform exempted a subgroup of firms *entirely* from bank withholdings. As a result, in this case we would expect to only observe changes in the extensive margin, which is confirmed by the green markers in Figures D.2a and D.3b.

Taken together, our facts suggest that the reforms we exploit in this paper operated as they ought to. The analysis also suggests that the reforms were substantially large, modified the way the turnover tax is collected and, ultimately, provide sharp identifying variation to study the implications of withholding from firms in a credible way.

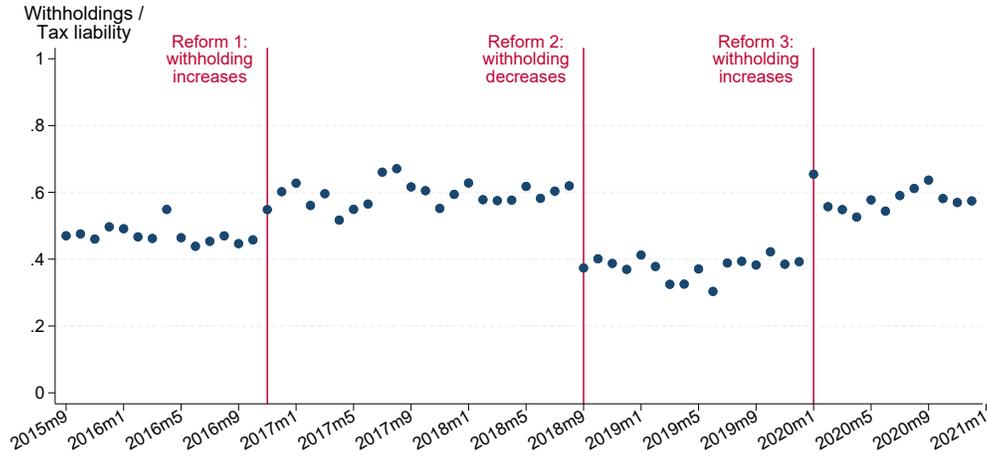
Table D.1: Statistics of withholding at baseline (September 2016)

	N Firms	Withheld firms	Withholdings / tax Liability				
		share	p5	p50	p75	p95	mean
Total withholding	132,897	0.806	0.029	0.434	0.969	2.643	0.770
<i>By type of withholding:</i>							
Sales to CAs	132,897	0.351	0	0	0.188	1.094	0.256
Purchases from CAs	132,897	0.514	0	0.002	0.116	0.923	0.214
Bank deposits	132,897	0.473	0	0	0.280	1.373	0.279

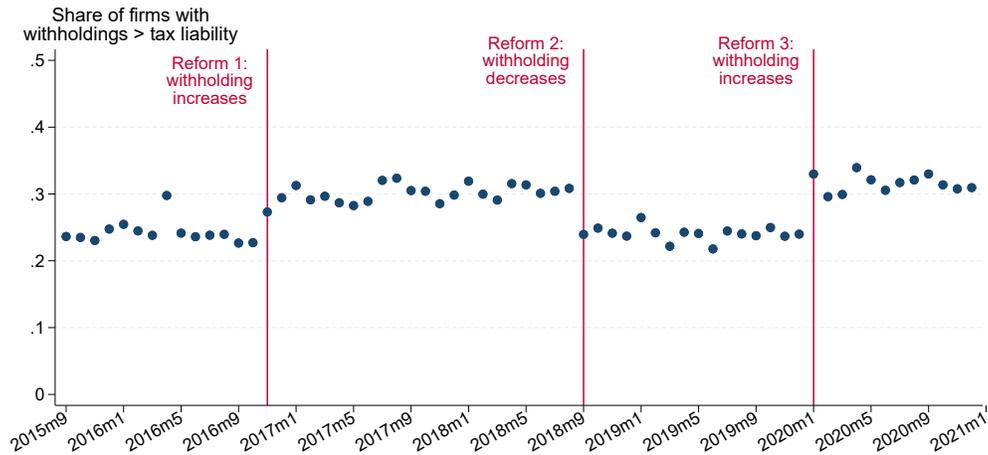
Notes: This table reports summary statistics of withholding calculated in September 2016, right before the expansion of the withholding regime. The first row shows statistics for total withholding. The last three rows break total withholdings down by type: those generated when selling to collection agents (CAs), those generated when purchasing intermediate goods from CAs, and those generated when receiving payments in a bank account. The first column shows the number of firms in the data. The second column reports the share of these firms that have positive withholdings. The last five columns show different moments of the share of tax withholdings to tax liability. Importantly, this ratio includes the zeros and thus captures both extensive and intensive margins.

Figure D.1: Tax withholdings versus tax liability

(a) Share of tax withholdings to tax liability



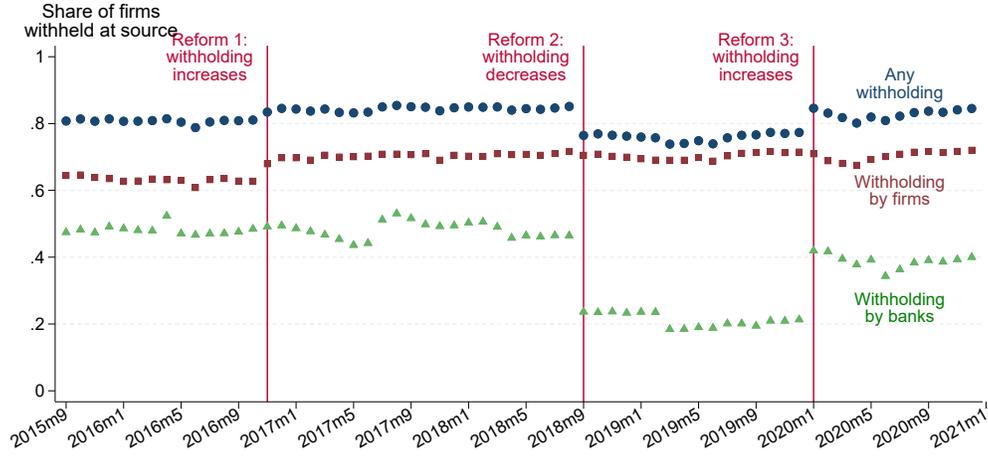
(b) Share of firms overwithheld



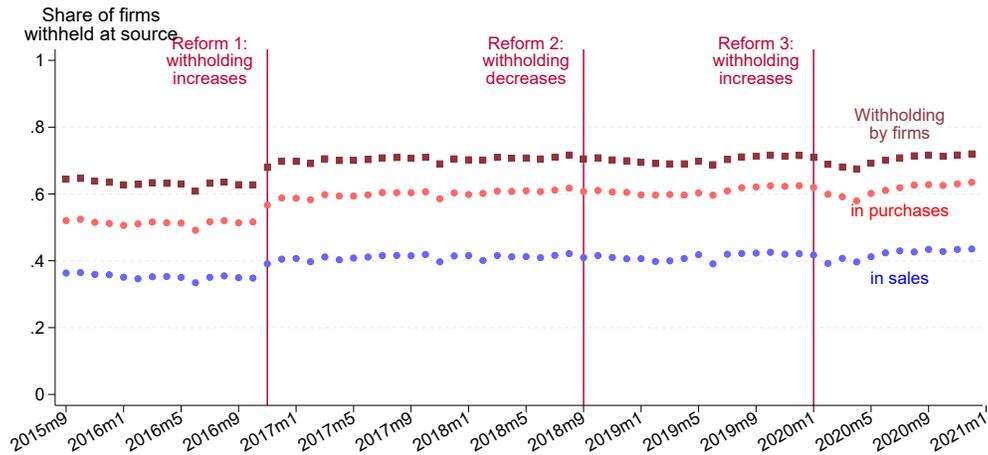
Notes: This figure shows descriptive evidence of the relationship between total tax withholdings and tax liability for the period September 2015 to December 2020. Panel (a) plots the ratio of total tax withholdings to tax liability. The numerator includes withholding from three sources: (i) commercial partners and card companies, (ii) banks, and (iii) the customs office. We report the *median* of the ratio to get a better sense of the distribution of firms and to avoid the influence of large outliers. Panel (b) plots the fraction of firms whose monthly withholdings exceed the tax liability. See the text for further details.

Figure D.2: Tax withholdings: extensive margin

(a) Share of withheld firms per month



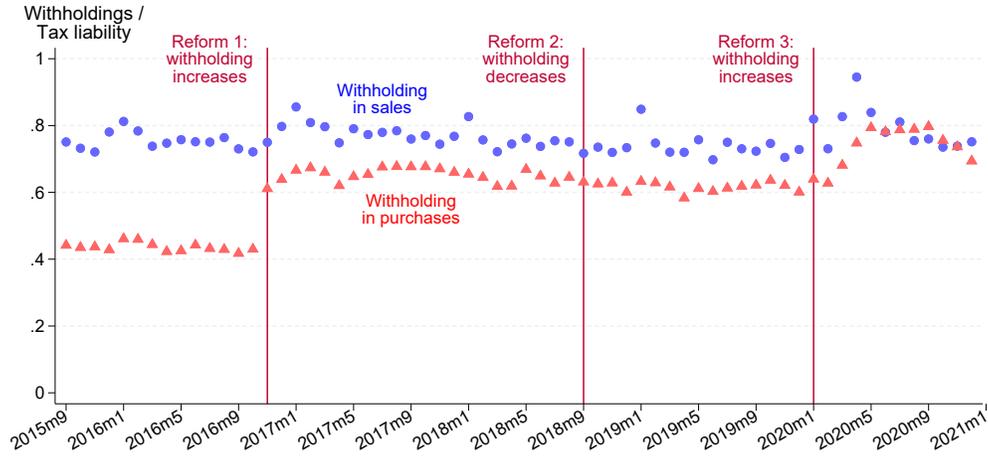
(b) Withholding in sales and purchases



Notes: This figure documents the extensive margin of withholding over time. It plots the share of firms withheld at source by type of withholding. The blue dots in the top panel correspond to firms with any withholding (that is, by collection agents, banks, or customs). The brown dots show withholding by collection agents. The green dots show withholding by banks. The bottom panel further breaks the withholding by firms (brown dots) from Panel (a) into withholding of sales (light blue) and withholding of purchases (light red). See the text for further details.

Figure D.3: Tax withholdings: intensive margin

(a) Withholding by commercial partners and credit card companies



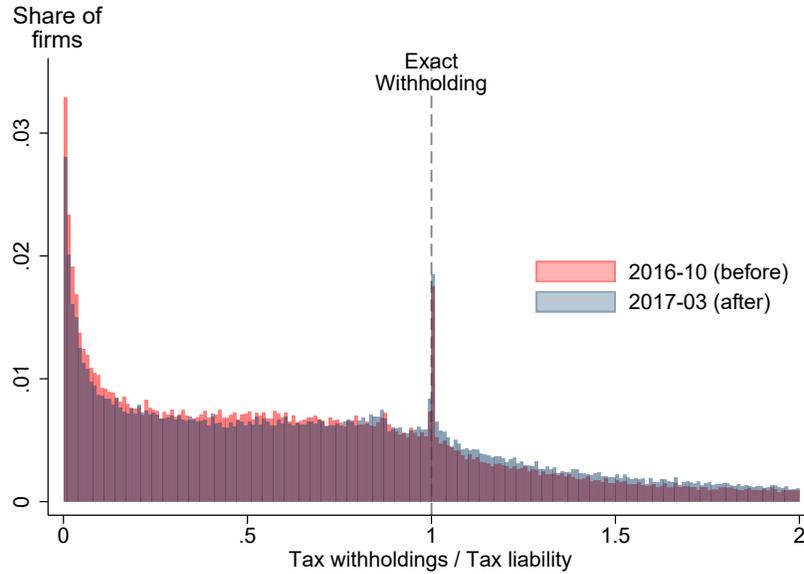
(b) Withholding from banks



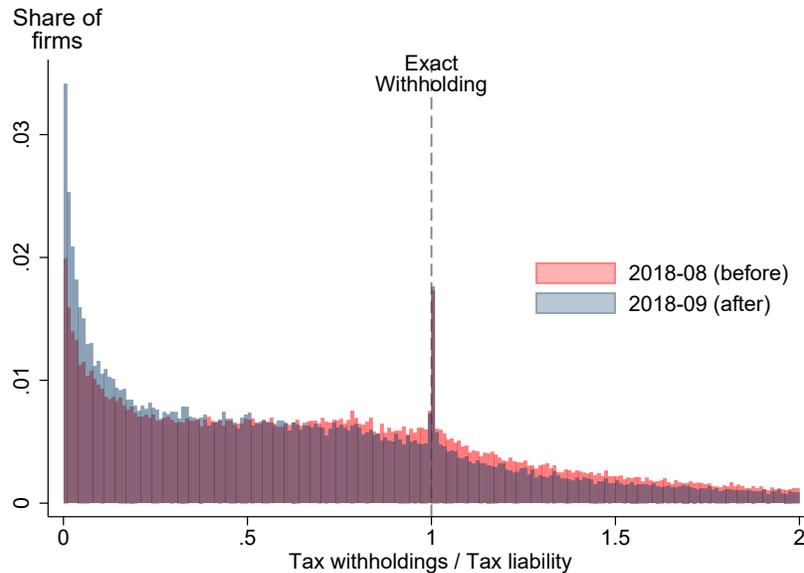
Notes: This figure documents the intensive margin of withholding over time. Panel (a) shows withholdings by firms and Panel (b) withholdings by banks. Each panel plots the share of tax withholdings to tax liability for firms with positive withholding in each month separately. To capture responses at the upper end of the distribution, we compute the *average* rather than *median* and we exclude the largest outliers. See the text for further details.

Figure D.4: Distribution of tax withholding to tax liability

(a) Before and after the increase in withholding by collection agents



(b) Before and after the decrease in withholding by banks

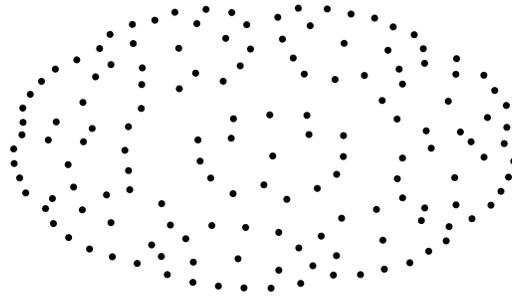


Notes: This figure shows the distribution of the ratio of total tax withholdings to tax liability. We use two hundred bins of 0.01 width. Panel (a) shows the density right before (pink) and after (blue) the increase in withholding by collection agents. Panel (b) shows the density right before (pink) and after (blue) the decrease in withholding by banks. In Panel (a) the density shifts to the right and in panel (b) it shifts to the left. Both panels exhibit bunching at the point at which withholding exactly matches the tax owed.

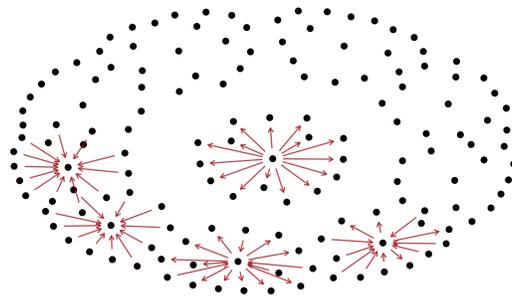
E Additional Figures and Tables

Figure E.1: Networks

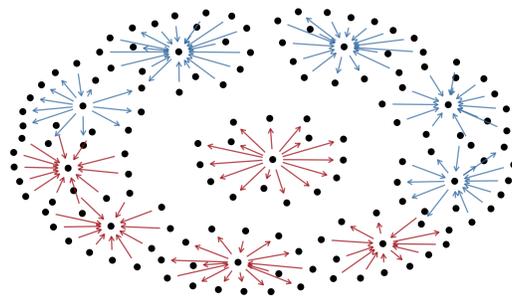
(a) No CA data



(b) Only old CA data



(c) Both old and new CA data



Notes: These figures illustrate the construction of the business-to-business linkages in our data set. Each node represents a real firm, and the connecting edges represent real linkages. The central nodes, from which the edges radiate, correspond to firms acting as collection agents (CAs) while the direction of the edges signifies whether the CA is a seller (outgoing) or a buyer (incoming) in the transaction. Notice that the same CA can simultaneously withhold sales and purchases and thus have arrows pointing in both directions.

Figure E.2: Relevance of the reform

ambito.com

Modificaciones en régimen de retenciones en CABA

EDICIÓN IMPRESA 11 Octubre 2016

Los agentes de retención en el régimen de recaudación reformulado serán designados de oficio cuando superen ciertas cifras de ingresos del año calendario anterior sin advertirse un mecanismo de reclamo.



Por [Jorge A. Carmona y Héctor A. Duguine](#)

El régimen recaudatorio vigente en la actualidad se haya dispuesto por Resolución N° 939/AGIP 2013, con las modificaciones introducidas por la Resolución N° 364/AGIP 2016 y la Resolución N° 421/AGIP 2016. **Las modificaciones introducidas, crean un típico régimen general de recaudación de retenciones y percepciones basado en gestión de alícuotas por Padrón de Contribuyentes, intentando emular el vigente en la provincia de Buenos Aires administrado por ARBA con retenciones y percepciones cruzadas, perdiendo los Agentes de Recaudación el beneficio de no ser retenidos y percibidos.** Hasta la modificación efectuada por la Resolución N° 364/AGIP 2016, en la CABA los agentes de recaudación eran en todos los casos designados taxativamente por el Fisco, a partir de ahora coexistirá ese régimen con el de padrones de alícuotas para el régimen general de retenciones y percepciones. Además de los nominados, deberán actuar en ese carácter los sujetos que desarrollen actividades en la Ciudad Autónoma de Buenos Aires que obtenido en el año calendario inmediato anterior ingresos por un monto a los \$60.000.000. **A diferencia de la provincia de Buenos Aires en donde el potencial agente debe inscribirse, en la CABA será designado de oficio cuando supere esa cifra. No advertimos un mecanismo expreso de reclamo.**

Con ese objeto deberán los ingresos gravados, exentos y no gravados correspondientes a todas jurisdicciones, netos de impuestos. Se considera que desarrollan actividades en la Autónoma de Buenos Aires aquellos sujetos que posean en esta jurisdicción agencias, representaciones, oficinas, locales y todo otro tipo de explotación, edificio, obra, depósito o similar y quienes se valgan para el ejercicio de su actividad en territorio de la Ciudad Autónoma de Buenos Aires, de los de comisionistas, corredores, consignatarios o martilleros. Existen una serie de sujetos excluidos de actuar como agentes de recaudación.

Notes: This figure shows an article from the *Ambito* newspaper that discusses the major expansion of the withholding scheme analyzed in this paper. The title is “Changes in the Withholding Regime in CABA (City of Buenos Aires),” and the text below reads “Withholding agents in the reformulated collection regime will be appointed ex officio when they exceed an income threshold from the previous calendar year and have no right to appeal the appointment.” Related newspapers also published articles with FAQs clarifying the details of the reform. For example, see [iProfesional](#). Source: [Ambito.com](#).

Figure E.3: Example of the register of withholding agents published monthly by the tax administration



NOMINA AGENTES DE RECAUDACION
MAYO 2017

ORDEN	AG	SV	CUIF	RAZON SOCIAL	NORMA	VEGENCIA
1248	15404	0	30711230188	CABLETECH ARGENTINA SA	RES. 364-AGIP-2016	01/11/2016
1249	8410	9	30573652084	CABLEVISION SA	RES. 1080-MHGC-07	02/05/2007
1250	9454	2	30500590692	CABOSCH S.A.	RES. 607-A-GIP-2009	01/10/2009
1251	15875	0	33504113049	CABOT ARGENTINA SAJC	RES. 364-AGIP-2016	01/11/2016
1252	1543	7	30537745742	CABRALES SOCIEDAD ANONIMA	Res. 1465-SHYF/01	01/09/2001
1253	9506	6	30708989245	CACHAY S.A.	RES. 50-AGIP-2010	01/03/2010
1254	11537	0	30507400228	CACHI YACO SA	RES. 364-AGIP-2016	01/11/2016
1255	13811	0	30700584131	CACHO SUSPENSION S.A.	RES. 364-AGIP-2016	01/11/2016
1256	15356	0	30710983107	CACIC SPORTS VISION SRL	RES. 364-AGIP-2016	01/11/2016
1257	9519	5	30544013544	CADENACI SA	RES. 50-AGIP-2010	01/03/2010
1258	9760	8	30578899002	CAFE OYAMBRE SRL	RES. 371-AGIP-2010	01/08/2010
1259	9761	1	30707732055	CAFE ROCAMORA S.R.L.	RES. 371-AGIP-2010	01/08/2010
1260	1507	9	30501077131	CAFES LA VIRGINIA S.A	Res. 430-SHYF/01	01/05/2001
1261	1544	0	30523456144	CAFES MUÑOZ SRL	RES. 371-AGIP-2010	01/08/2010
1262	9479	7	30504312212	CAGNOLI S.A.	RES. 675-AGIP-2009	01/12/2009
1263	14471	0	30708284250	CAIF COMPAÑIA ARGENTINA DE INVESTIGACIONES FARMACEUTICAS S.A.	RES. 364-AGIP-2016	01/11/2016
1264	15870	0	33500730639	CAIMARI	RES. 364-AGIP-2016	01/11/2016
1265	3749	6	30863205621	CAJA DE SEGUROS SA	Res. 430-SHYF/01	01/05/2001
1266	4749	2	30554475910	CAJA DE VALORES S.A.	Res. 1465-SHYF/01	01/09/2001
1267	3563	6	30701951251	CAJA NOTARIAL COMPLEMENTARIA DE SEGURIDAD SOCIAL	Res. 1465-SHYF/01	01/09/2001
1268	12739	0	30644396675	CALDENES SA	RES. 364-AGIP-2016	01/11/2016
1269	11340	0	30500039678	CALEDONIA ARGENTINA COMPAÑIA DE SEGUROS S.A.	RES. 364-AGIP-2016	01/11/2016
1270	13303	0	30683338121	CALENDER SA	RES. 364-AGIP-2016	01/11/2016
1271	9271	3	30555460526	CALERA BUENOS AIRES S.A.	RES. 568-AGIP-2008	01/02/2009
1272	9280	8	30709076902	CALERA CONGRESO SH D E JUAN M SELVA MARIANO F SELVA Y JUAN I SELVA	RES. 568-AGIP-2008	01/02/2009
1273	14607	0	30708555386	CALERA DE SAENZ P.E.A. S.A.	RES. 364-AGIP-2016	01/11/2016
1274	14152	0	30707587845	CALERA EL OMBU S.A	RES. 364-AGIP-2016	01/11/2016
1275	16050	0	33683632899	CALERA SAN JUSTO SA	RES. 364-AGIP-2016	01/11/2016
1276	12253	0	30596774640	CALFOR S.R.L.	RES. 364-AGIP-2016	01/11/2016
1277	10998	4	30710371527	CALFRAC WELL SERVICES (ARGENTINA) SA	RES. 353-AGIP-2014	01/07/2014
1278	15949	0	3359683679	CALICO S.A.	RES. 364-AGIP-2016	01/11/2016
1279	14658	0	30708687568	CALLMED SA	RES. 364-AGIP-2016	01/11/2016
1280	16136	0	33707528449	CALOREX S.A.	RES. 364-AGIP-2016	01/11/2016
1281	9089	6	30706046093	CALZADO LOPEZ TAIBO SRL	RES. 313-AGIP-2008	01/10/2008
1282	11908	0	30551587300	CALZADOS FERLU S.A.	RES. 364-AGIP-2016	01/11/2016
1283	9091	5	30690631721	CALZARTE SRL	RES. 313-AGIP-2008	01/10/2008
1284	12150	0	30585357657	CALZETTA NEUMATICOS SOCIEDAD ANONIMA	RES. 364-AGIP-2016	01/11/2016
1285	10800	4	33690589619	CALZIFICIO ALDAMA SA	RES. 592-AGIP-2012	01/10/2012
1286	9641	8	30700955008	CAMARA DE EMPRESAS DE CONTROL Y ADM. DE INFRACCIONES DE TRANS DE LA REP. ARG. ASOCIACION CIVIL	RES. 144-AGIP-2010	01/05/2010
1287	10323	5	30643748823	CAMARAS Y LUCES SA	RES. 259-AGIP-2012	01/06/2012
1288	15308	0	30710776063	CAMARIQAS SA	RES. 364-AGIP-2016	01/11/2016
1289	12621	0	30632829732	CAMASA SA	RES. 364-AGIP-2016	01/11/2016
1290	11392	0	30501948981	CAMBRE INDUSTRIAL COMERCIAL Y FINANCIERA SA	RES. 364-AGIP-2016	01/11/2016

Notes: This figure shows a page from the register of withholding agents published monthly by the tax administration. The full list, in particular, includes all the active withholding agents as of May 2017. This information is publicly available on the [AGIP website](#). Each row corresponds to a different firm. The fourth column shows the tax identifier, the fifth column shows the name of the company, the sixth column shows the resolution that appoints the firm to act as collection agent, and the last column shows the starting date. Not surprisingly, the majority of the dates correspond to November 1, 2016, the date at which the major reform we analyze was implemented. We downloaded all these monthly registers, digitized them, and constructed a database that we merged with our main administrative data. We also used this database to construct Figure 2. It is interesting to observe the mix of firms appointed to act as collection agents. In this list one can see coffee producers, pharmaceutical firms, insurance companies, footwear manufacturers, and construction-materials wholesalers, among others.