Impact of the GST on Corporate Tax Evasion: Evidence from Indian Tax Records*

Sumit Agarwal† Shashwat Alok‡ Shiv Dixit§ Tejaswi Velayudhan¶

December 6, 2022

Abstract

In July 2017, India replaced its fragmented indirect tax structure with a nationwide Goods and Services Tax (GST). The change in the tax regime induced an increase in third-party reporting and introduced an enforcement notch. In this paper, we study the impact of the GST on tax compliance by businesses. We document that firms reported higher revenues and costs in response to the tax change. However, these effects can be attributed to both increased efficiency and greater tax compliance. To unpack the impact of the regime change on evasion, we focus on financial statement fraud. We find that while the GST reduced revenue underreporting, it also prompted evasion on other (non-verifiable) margins. We propose a novel technique to detect cost overreporting that exploits variations in the composition of firm inputs around tax exemption thresholds. We use our proposed method to show that firms overreported their wage bills in response to the regime change. We develop a structural model that is consistent with these empirical patterns and use it to study several ways of increasing tax revenues.

Keywords: GST, India, tax evasion

JEL Classification Numbers: D22, H25, H26

*We thank Alan Auerbach, Jim Poterba, and Krishnamurthy Subramanian for their valuable comments. Shubham Roy provided excellent research assistance. We are grateful to the Ministry of Finance, Government of India for sharing administrative data.

†National University of Singapore; ushakri@yahoo.com

‡Indian School of Business; shashwat_alok@isb.edu

§Indian School of Business; shiv_dixit@isb.edu

¶UC Irvine; tvelayud@uci.edu

1 Introduction

Limited tax capacity is a key challenge in many developing countries. A growing literature emphasizes the role of tax structures that encourage information sharing, whether it be directly with federal authorities or indirectly through third parties, in curbing tax evasion. Such structures can raise tax revenues by making economic transactions more transparent and reducing income underreporting. However, plugging one hole in the tax revenue bucket may open another. More specifically, improvements in the verification of final and intermediate goods may not increase tax collections if firms make offsetting adjustments by misreporting less verifiable items. In this paper, we investigate the presence of such evasion shifting in the context of corporate taxation in India.

Corporate tax rates in India are among the highest in the world. As emphasized by Gutmann (1977) and Clotfelter (1983), a large tax burden implies a high return on evasion. Furthermore, tax norms have predominantly varied considerably across Indian states, with the lack of harmonization impeding tax enforcement. It is unsurprising, then, that financial statement fraud has been a cause of concern in India, where many firms hide sales or inflate expenses.

In July 2017, India replaced its fragmented indirect tax structure with a single indirect tax that applied across the country–the Goods and Services Tax (GST). A key objective of the regime change was to address low tax compliance levels using third-party reporting and invoice matching. A centralized electronic filing system was established through which various state and federal tax authorities could share information, including the details of consolidated financial statements of companies registered in the country. We study the extent to which this regime change has been successful in promoting tax compliance by businesses using a novel administrative dataset of corporate tax returns of Indian firms.

This paper is divided into three parts. In the first part of the paper, we examine the impact of the GST on revenues and expenses reported by firms. At the aggregate level, revenue and costs both decreased post-GST. However, these aggregate-level trends cannot be construed as causal evidence. To identify the impact of the GST on firm revenues and expenses, we use a difference-in-differences (DD) design. By comparing the response of firms that supplied goods that were exempt from GST to those that were not, we show that the regime change had a statistically significant and positive impact on firm revenues and costs. However, these results do not provide an insight into the underlying mechanism; They can be rationalized by increased production efficiency stemming from the reduction in border tariffs and the harmonization of the commodity tax code, or by greater tax compliance stemming from increased enforcement under the new regime. To ascertain the effect of the GST on tax compliance, we focus on two margins of financial statement fraud: revenue underreporting and cost overreporting.

In the second part of the paper, we show that revenue underreporting decreased in the GST era using techniques from the bunching literature. We follow Saez (2010), Kleven and Waseem (2013), and Chetty et al. (2011) to quantify the extent of revenue underreporting at indirect tax exemption thresholds. We find that the extent of bunching in the distribution of firm revenues around GST exemption thresholds reduced by 30 percent due to the regime change. Moreover, we find that the extent of underreporting in the GST era is highly correlated with the return to evasion, i.e., the tax rate.[1]

In the third part of the paper, we propose a new method to detect cost overreporting and argue that this margin of evasion increased after the implementation of GST for a subset of firms. To identify financial statement fraud, the existing literature has focused on the input-output share (Johnson et al., 1997). This approach assumes a stable relationship between some physical input and output. Deviations from the fixed input-output ratio are then construed as income underreporting. However, firms can evade taxes by overreporting deductible expenses as well. Carrillo et al. (2017) find that even though third-party reporting increases revenue reporting, it also increases reported costs, so the tax burden of the firm remains unchanged; Presumably, because sales are easier to monitor than costs.[2] In such a scenario, inferences based on the deviations of the input-output ratio may be incorrect.

However, detecting cost overreporting is challenging. A potential problem with identification in Carrillo et al. (2017) is that increased costs could have been used to finance increased production. That is, a growing firm can lead to biased estimates. To overcome this challenge, we propose a novel technique to detect cost overreporting in which we identify evaders using variations in the composition of firm expenses. Our identification strategy leverages two institutional details. First, firms with revenues below INR 20 lakhs were exempted from filing every month a form detailing all outward supplies made, input tax credit claimed, tax liability ascertained, and taxes paid. Thus, the operations of firms below the exemption threshold were more opaque to the tax administrators than those of others. Second, under the new tax regime, taxpayer reports were verified against third-party information using invoice-matching. Moreover, the previous indirect tax regime did not permit input tax credit (ITC) claims for several taxes, such as the Central Sales Tax, the Entry Tax, and the Luxury Tax. But under GST, there is no restriction on claiming ITCs. Thus, monitoring external costs (such as the cost of intermediate goods) became more effective than monitoring internal

[1] This contributes to Fisman and Wei (2004), who identify the response of evasion to product-specific import tariffs and value-added tax in China by comparing data on exports and imports.

[2] Slemrod et al. (2017) find a similar result for sole proprietorships in the U.S. Another related paper is Handley and Moore (2017), who show that reported (deductible) transport costs vary positively with tariff rates, which they attribute to misreporting.

costs (such as wages). In India, about two-thirds of wage employees are casual workers who do not receive social security benefits and are compensated daily (ILO, 2018a). The wages of these workers lack a paper trail and lie outside the purview of tax authorities. We argue that the GST prompted firms to overreport such non-verifiable costs.

To illustrate the mechanism, we present a simple model in which the government has imperfect visibility of firm expenses. In the model, firms have an incentive to overreport non-verifiable costs to reduce their corporate tax burden. We show that firms overreport their non-verifiable costs by more below the tax exemption thresholds that allow firms to fly under tax authorities' radar. To test this hypothesis, we employ a sharp regression discontinuity (RD) design where we examine jumps in the ratio of non-verifiable expenses to total expenses around GST exemption thresholds. In our baseline sample, which focuses on the period 2017-18, we detect a negative jump at the exemption threshold: the ratio of non-verifiable expense to total expenses decreases by 11.8 percent when revenues are above the tax exemption threshold relative to when they are below. Our RD estimates are most pronounced in FY2018, i.e., which is the first fiscal year following the implementation of GST. We do not find any evidence of cost overreporting before the implementation of GST.

To provide further evidence on our proposed mechanism, we exploit variation in the visibility of intermediate inputs relative to labor inputs. In the value-added tax (VAT) system operational before GST, state governments issued way bills for cross-border shipments, which provided details of the goods being transported. Since the VAT was an origin-based tax, and the rates varied across states, checking consignments and tax payment at borders was the norm. The GST, in contrast, is a destination-based tax on final consumption, and the tax rates are uniform across the country. Thus, under the GST, way bills were no longer required to ascertain the ultimate tax liability. Consequently, states abolished border checkposts after the rollout of GST in July 2017. Nonetheless, state governments still wanted to track goods moving in and out of their borders to prevent tax evasion. However, due to disruptions in the implementation of the new system, which was rolled out in a phased manner starting in June 2018, there was a blackout period for about a year where cross-border transactions were relatively less visible than before. As a result, to reduce their corporate tax bill, firms that procured a larger portion of their intermediate inputs from out-of-state had a larger incentive to overreport these costs (or, equivalently, a lower incentive to overreport their labor costs relative to other expenses). This prediction is borne out in the data as well.

Our identification strategy relies on detecting discontinuities in the behavior of firms around tax exemption thresholds. These firms are relatively small by construction. Nonetheless, one can infer the impact of the GST on large firms using the underlying characteristics of evading firms. Service providers primarily drive our baseline RD results. These firms tend to be labor-intensive with a substantial wage bill. In contrast, we do not find any evidence of cost overreporting for non-service providers, or for capital-intensive firms. Since large firms tend to be capital intensive, our results imply that the regime change is unlikely to have increased financial statement fraud in such firms. Thus, overall, our results suggest that the GST had a positive impact on tax compliance by businesses in India, barring a set of relatively small labor-intensive firms that responded by overreporting their costs.

Our reduced-form analysis can detect the presence of cost overreporting but not its magnitude. To compute the underlying level of cost overreporting implied by our RD estimates, we structurally estimate our model. We develop a simulated method of moments estimation approach that is consistent with results from reduced-form estimators. Notably, our model accounts for endogenous responses in costs and revenues. Our results suggest that firms around the tax exemption threshold (i.e., those with revenues lower than INR 40 lakhs above the threshold) overreported their costs by 9.2 percent.

Lastly, we use the estimated model to study how the authorities can best increase tax revenues. A notable counterfactual experiment that we run is reducing the tax deductibility of costs, which may seem like a natural solution to the problem of cost overreporting. Our baseline model features a pure profit tax, similar to that implemented in India. Best et al. (2015) argue that increasing the tax deductibility of costs can lower evasion as it reduces the marginal benefit of evasion without affecting its marginal cost. We show that the presence of an enforcement notch can overturn this result. While firms above the tax exemption threshold are exposed to the mechanism outlined in Best et al. (2015) and thereby evade less, firms below the exemption threshold that face a lower probability of being audited evade more. We show that increasing the tax deductibility of costs increases cost overreporting by firms below the tax exemption threshold by more than the decrease above the threshold. Hence, on net, reducing the tax deductibility of costs would have the unintended effect of decreasing tax revenues.[3]

Related Literature. Our paper contributes to the literature that examines the effect of third-party reporting on tax compliance. Much of the previous work has focused on how third-party information is essential for tax collection in developed countries (Kleven et al., 2016). The literature that studies the importance of information and third-party reporting for effective taxation in developing countries is primarily theoretical (Kopczuk and Slemrod, 2006; Gordon and Li, 2009), and empirical evidence in this domain is thin. Due to cross-

[3]We present these results with a noteworthy caveat. In the extreme case of a pure turnover tax, firms below the tax exemption threshold have no incentive to overreport their costs. Nevertheless, Diamond and Mirrlees (1971) show that a turnover tax is inefficient. Thus, in principle, a turnover tax could be welfare-enhancing if the shadow price of public funds is large relative to the efficiency cost associated with adopting a turnover tax.

country differences in enforcement and informational constraints, the impact of third-party reporting on compliance in developing economies could be substantially different from that in the developed world. There is strong evidence that third-party information increases the reporting of firm revenues considerably in developed countries (see, for instance, Pomeranz (2015) and Almunia and Lopez-Rodriguez (2018)). The reduction in revenue underreporting may be more pronounced in emerging economies as they tend to have a larger pool of self-employed workers and informal firms. Better detection may stimulate the entry of these agents into the formal sector and thereby raise tax revenues. Indeed, we find that Indian firms increased reported revenues considerably with the advent of third-party verification in the GST era. On the other hand, the effectiveness of third-party reporting in developing economies may be limited as firms may respond by making offsetting adjustments. Specifically, in response to increased detection of sales, firms may inflate costs that are harder to verify using third-party reports. Though Carrillo et al. (2017) provide suggestive evidence along these lines using data from Ecuador, there is virtually no causal evidence on the impact on third-party reporting on cost overreporting. To the best of our knowledge, this study is the first to provide such evidence.

To examine the differential ease of evasion under profit versus turnover taxation, Best et al. (2015) also present a model in which firms can overreport costs. Nevertheless, they admit that cost evasion is not crucial for their empirical or conceptual results. Their bunching estimates identify approximately the aggregate evasion reduction when switching from profit to turnover taxation, i.e., their estimation procedure cannot separate cost from output evasion. Bachas and Soto (2021) rely on a framework similar to Best et al. (2015) to compute the elasticity of taxable corporate profit with respect to the tax rate using Costa Rican data. While their reduced-form approach only provides a lower bound to cost elasticity, they structurally decompose the observed reduction in profits into a decrease in reported revenue and an increase in reported cost. Their approach recovers the extent of cost overreporting around notches in taxable income, but not around enforcement notches. Our approach, in contrast, can speak to how cost overreporting responds to enforcement intensity.

2 Institutional Background

2.1 Recent Developments in the Indirect Tax Regime

The GST is an indirect tax levied at each step in the production process and refunded to all parties in the production chain barring the final consumer. Before the implementation of GST, India had a fragmented VAT system.

There are substantial differences in tax administration in the pre- and post-GST era. In the previous system, companies were paying taxes at different production stages and were also being taxed separately by various government authorities. This often led to double taxation, i.e., goods were being taxed at the factory gate and the retail store. This led to cascading effects as sales taxes were levied on the gross value without any input credits. This was partially addressed in the previous regime via input credits for Central Value Added Taxes (CENVAT). The state VATs, however, were separate from the CENVAT, and these taxes were not allowed to be credited against each other. This issue was aggravated by the fact that some states levied draconian entry taxes that restricted the free movement of goods around the country.

In addition, the variation in tax norms across states created a complex tax system. Under the GST, all state and central taxes were subsumed under a single integrated system, which was easier to navigate. Furthermore, under the previous regime, the exemption threshold for VAT ranged from INR 5-10 lakhs depending on the state, and was INR 10 lakhs under The Finance Act (1994).[4] When the GST was rolled out, the threshold limit for registration had been kept at INR 20 lakhs for both goods and services. Importantly, all registered firms were required to routinely self-declare details of all outward supplies made, input tax credit claimed, tax liability ascertained, and taxes paid. Firms with revenues below INR 20 lakhs were exempted from filing these forms. It is reasonable to assume that the increased frequency of exchange of detailed financial information also increased the probability of detecting evasive responses. Thus, the differential information requirement on either side of the tax exemption threshold essentially introduced an enforcement notch.

The new system also aims to improve tax compliance using invoice matching. Both central and state GST administrations are bestowed with the power to audit. To facilitate enforcement, a common nation-wide IT backbone—the GST Network (GSTN)—has been put in place through which all transactions are required to be filed. In most countries, enforcement is achieved using operational audits. In contrast, in the Indian GST, compliance is accomplished via 100% matching of all the invoices in the case firms with a turnover of above INR 75 lakhs via the GSTN.

Another reason to expect better tax compliance in the GST era is the increased utilization of the ITC. A growing literature argues that verifying taxpayer reports against third-party information is critical for tax collection. Furthermore, voluntary registration makes value added taxes unique amongst all the major taxes. Voluntary registration refers to a situation where a firm registers for the GST even if it is below the turnover threshold, and thus is not required to do so. This is likely to occur when a firm has large purchases of intermediate inputs. In this case, it may be profitable to voluntarily register for the GST so the firm can

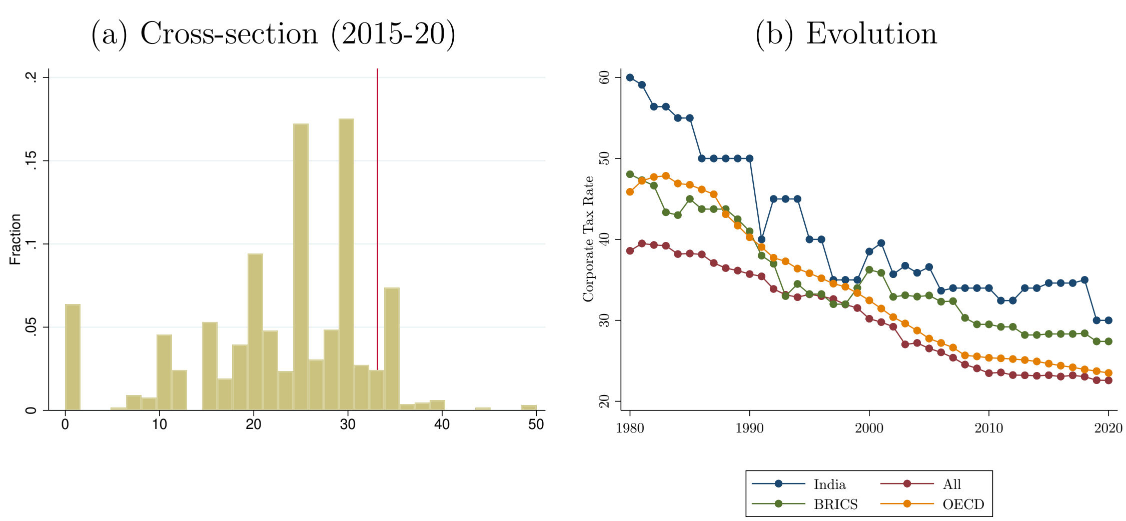
[4] However, companies producing a set of listed products were earlier exempt from paying central excise duties under the small-scale industry exemption which let off firms with revenue of up to INR 1.5 crores.

claim back input tax, while passing some or most of the burden of the output tax on to the purchaser.

2.2 Predominance of High Corporate Taxation

Prior to the corporate tax cut in September 2019, which is toward the end of our sample, the tax burden on firms in India was substantial. Domestic companies with an annual turnover up to INR 250 crores were taxed at a rate 25 percent, while those above this threshold were taxed at 30 percent. Foreign companies were taxed at a higher rate at 40 percent. Companies with a turnover of more than a crore faced graduated surcharges, while those below were exempt.

Corporate tax rates in India are among the highest in the world. This can be seen clearly in Panel (a) of Figure 1, which depicts the distribution of statutory corporate tax rates over the sample period that we study, i.e., 2015-2020. This is not a recent phenomenon. Panel (b) of Figure 1 compares the evolution of statutory corporate tax rates in India with the rest of the world. In 1980, corporate tax rates around the world averaged 40.1 percent, while in India, they were 60 percent. Despite the secular downward trend in corporate tax rates over the past four decades, corporate taxes in India have consistently been larger than international comparators in both the developed and developing world. In 2020, the OECD and BRICS had average statutory rates of 23.5 and 27.4 percent respectively, which were considerably lower than that in India.



Source: Tax Foundation Database.

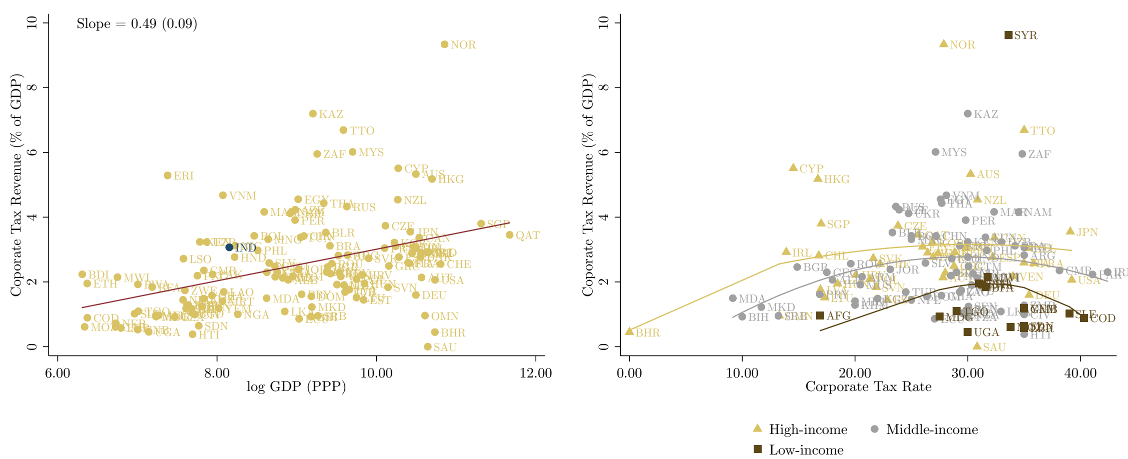
Notes: In panel (a), we only include observations in the period 2015-2020. The red line marks the average corporate tax rate in India over this period.

Figure 1: Statutory Corporate Tax Rates

2.3 Prevalence of Financial Statement Fraud

Developing countries collect substantially less corporate tax revenue as a share of their gross domestic product than developed countries (Figure 2(a)). While it is plausible that lower corporate tax revenues in India stem from the fact that statutory rates are relatively high in the country, it could also be due to tax evasion (Gordon and Li, 2009; Bachas and Soto, 2021). Figure 2(b) shows that the cross-country distribution of corporate tax revenues cannot be completely explained by differences in tax rates. While the existence of a Laffer curve for corporate taxes is prevalent within each income group, conditional on corporate tax rates, high-income countries tend to raise about twice as much corporate tax revenue as a share of gross domestic product than low-income countries. This pattern suggests that the importance of tax evasion in explaining variations in corporate tax revenues across countries may be of first-order.

(a) Tax Revenues on GDP (b) Tax Revenues on Tax Rates



Source: Tax Foundation Database, UNU-WIDER Government Revenue Dataset, World Bank World Development Indicators Database.

Notes: Observations reflect averages over the period 2000-2015. Countries with less than one million in population are excluded. In panel (b), we estimate Laffer curves (across income groups according to the World Bank classification) by fitting a fractional polynomial of degree two.

Figure 2: Corporate Tax Revenues

Even survey evidence suggests that financial statement fraud is pervasive in India, whereby a substantial amount of firms hide sales or inflate expenses, often via ghost employees and fictitious contracts. Deloitte India Fraud survey report tracks sentiments on the topic of corporate fraud in India. In 2016, three surveys–which focused on large (domestic and multinational) companies, small and medium enterprises, and working professionals–were conducted. In their large companies survey, 10 percent of the respondents experienced financial misreporting over the last two years. In the small and medium companies survey, 21 percent of survey respondents experienced financial misreporting over the previous two years. In their working professionals survey, 40 percent of the respondents suspected their organization had experienced financial statement fraud. A majority of the respondents also felt that corporate tax fraud is likely to rise in the coming years (Deloitte, 2016).

There are at least two reasons for the prevalence of financial statement fraud in India. First, the country's high level of direct and indirect tax rates imply a high return to evasion. Second, tax exemption thresholds prompt businesses to underreport their revenues. Like in most countries around the world that use VATs, there is a minimum registration threshold based on annual turnover below which companies in India do not need to register for VAT. As VAT rates are often quite high, this incentivizes firms to underreport their reported turnover to avoid registering for the VAT. This allows firms to avoid bearing the compliance cost associated with registration and enables them to fly under the tax authorities' radar by withholding detialed information regarding their inputs. Since tax authorities cannot construct a comprehensive picture of firms' costs through third-party information, avoiding VAT registration permits firms to continue to reduce their tax liability by overreporting costs.

3 Data and Descriptive Statistics

Our primary data source is administrative tax records from the Ministry of Corporate Affairs (MCA), Government of India. As emphasized by Card et al. (2010), administrative data have far fewer problems with attrition, non-response, and measurement error than traditional survey data sources. We rely on a random sample of all registered companies in India. Our sample ranges from 2015 to 2020, and covers 21,538 firms.

We match stock and flow data from firm balance sheets and profit & loss accounts. We rely on tax filings under three umbrellas: AOC-4, IND-AS, and XBRL. Form AOC-4 is used to file the financial statements for each financial year with the Registrar of Companies (ROC). Non-bank financial corporations (NBFCs) are required to comply with Indian Accounting Standards (IND-AS). Barring a few exceptions, the following set of companies report their financials in the XBRL form: (i) companies listed with any Stock Exchange(s) in India and their Indian subsidiaries; (ii) companies having paid-up capital of INR 5 crores or above; (iii) companies having turnover of INR 100 crores or above; (iv) companies required to prepare their financial statements as per Companies Rules (2015). Companies in banking, insurance, power sector, NBFCs, and housing finance companies are exempted from XBRL filing.

We merge the administrative data with survey data from the Annual Survey of Industries (ASI), conducted by the National Statistical Office, Ministry of Statistics & Programme Implementation, Government of India. The ASI allows us to assess the dependence of the overreporting of labor costs on differences in the visibility of intermediate inputs relative to labor inputs. In particular, we proxy the relative change in the visibility of intermediate goods after the implementation of GST using the percentage of inputs sourced from out-of-state. After the implementation of GST, intermediate goods were not scrutinized at state borders, and thus the net effect on the visibility of these goods due to the regime change was subdued for those firms that sourced more inputs from out-of-state. We make the following set of assumptions to calculate input procurement ratios. First, we exclude fuel inputs from the calculation because they do not come under the purview of GST. Second, we exclude items classified as "others" in the ASI as they cannot be mapped to standardized product classifications. Third, ASI provides data on inputs at the manufacturing unit level and not the product level. Thus, we do not know the mix of inputs used to generate each output for firms supplying multiple products. For these firms, we assume that all recorded inputs are used to generate a specific product in proportion to the respective product's share of total output. Lastly, we assume that a manufacturer will only look to procure a good from outside the state in which it is located if the in-state production of that good fails to meet the in-state demand.

Table 1 reports summary statistics for the firms used in the analysis. The last column of the table tests for differences in group means of firms in the pre- and post-GST period. We see a marked decline in firm revenues and costs post the implementation of GST. Average firm revenues reduced from INR 261.0 crores to INR 95.3 crores, while average expenses reduced from INR 246.7 crores to INR 91.9 crores. As we argue below, these differences do not capture the causal impact of the GST.

Figure 3 depicts the evolution of the cross-sectional mean of revenues and costs for firms in our sample. Both revenues and costs decreased in the GST era when we consider the entire sample. However, when we focus on firms around the exemption threshold, i.e., firms with average revenues above INR 19 lakhs and below INR 21 lakhs over the sample period, we see a very different picture. Firms around the exemption threshold report losses on average. Furthermore, these firms reported higher costs when the GST was implemented.

4 Impact of the GST on Reported Revenues and Costs

Figure 3 showed that both revenues and costs decreased in the GST era. However, one cannot attribute this decrease to the tax change since it coincided with a period of relatively anemic growth that could be due to other policies (such as Demonetization) and negative shocks from home and abroad. Indeed, we find the opposite when we control for these confounding conditions using a DD strategy. Despite the decreasing trend in mean revenues and costs, we find that the GST positively impacted both variables. This is consistent with two hypotheses. The first hypothesis is that the GST led to firm growth in the sample of

Table 1: Summary Statistics

	Observations (1) Total	All (2) Mean (S.D.)	Pre-GST (3) Mean (S.E.)	Post-GST (4) Mean (S.E.)	Difference (4) - (3) Est. (S.E.)
P&L a/c variables					
Total revenue	87,309	206.81	261.00	95.27	-166.73
		(3079.63)	(15.53)	(2.38)	(15.71)
Total expenses	87,309	196.10	247.65	91.94	-155.71
		(2897.06)	(14.62)	(1.93)	(14.75)
Employee compensation	87,302	14.45	18.14	6.99	-11.14
		(239.82)	(1.21)	(0.17)	(1.22)
Cost of materials	87,302	96.84	120.23	49.60	-70.63
		(1523.63)	(7.70)	(0.71)	(7.73)
Power and fuel costs	54,142	4.22	5.94	0.76	-5.19
		(354.28)	(2.28)	(0.02)	(2.28)
Finance cost	87,302	7.45	8.99	4.34	-4.64
		(179.79)	(0.67)	(1.25)	(1.42)
Insurance expenses	54,142	0.06	0.07	0.03	-0.04
		(1.40)	(0.01)	(< 0.01)	(0.01)
Auditing expenses	54,142	0.01	0.01	0.01	≈ 0
		(0.08)	(< 0.01)	(< 0.01)	(< 0.01)
Balance sheet variables					
Total assets	93,325	446.25	387.52	549.15	161.64
		(6577.09)	(22.40)	(44.39)	(49.72)
Borrowings	93,320	81.26	76.51	89.59	13.08
		(1743.96)	(5.70)	(12.12)	(13.40)
Equity	93,320	146.16	25.74	357.15	331.42
		(22295.51)	(1.42)	(200.84)	(200.85)
Cash	93,320	15.73	16.39	14.59	-1.80
		(205.09)	(0.90)	(0.96)	(1.32)

Notes: All variables are denominated in INR Crore. The pre- and post-GST periods correspond to the years 2015-2017 and 2018-2020 respectively. Columns 2 report means and standard deviations in parentheses. Column 5 reports differences of group means between columns 3 and 4 with standard errors in parentheses.

impacted firms; This increase in production justifies both higher revenues and costs. An alternative explanation is that the GST led to better income monitoring, which increased reported firm revenues without a substantial change in production. In order to reduce their corporate tax burden, firms reported higher costs, which were harder to monitor. In Section 6, we provide evidence on the latter hypothesis.

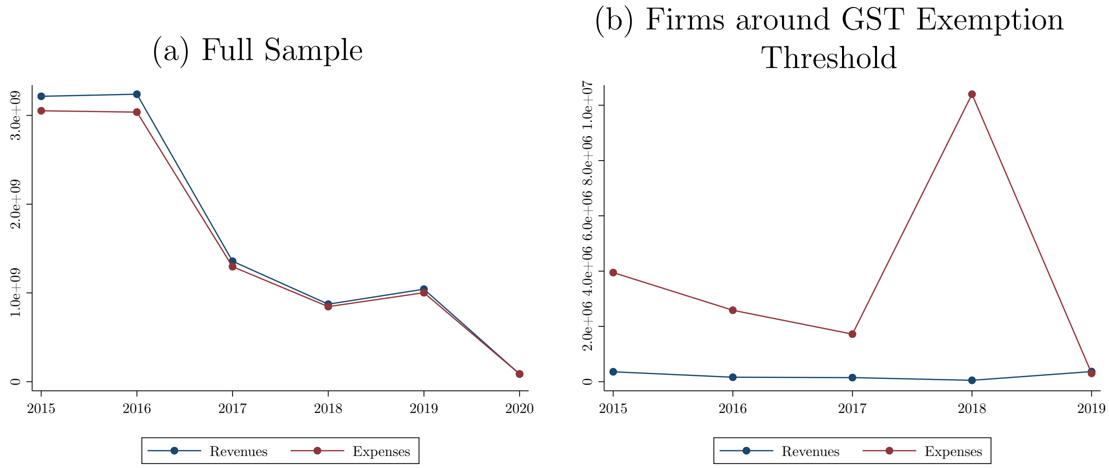


Figure 3: Evolution of Key Variables

4.1 Empirical Strategy

A DD estimation strategy follows naturally from the GST policy variation. While GST subsumes most indirect taxes from the previous VAT regime, certain levies have been kept out of its purview.[5] We construct the treatment variable as follows:

$$\text{GST}_{i,t} = \mathbb{1}_{p_i \in \text{Non-exempted goods}} \times \mathbb{1}_{r_{i,t} > \text{ INR 20 lakhs}} \times \mathbb{1}_{t > 2017},$$

where i denotes a firm, t denotes the tax filing year, p_i denotes the HSN code of the product supplied by firm i, and $r_{i,t}$ denotes the revenue of firm i in period t. The DD strategy is implemented using the regression framework:

$$y_{i,t} = \beta \text{GST}_{i,t} + \phi_i + \psi_t + \epsilon_{i,t},$$

where $y_{i,t}$ denotes outcome variables for firm i in period t, ϕ_i denotes firm fixed effects, and ψ_t denotes year fixed effects. Our primary object of interest is β, which captures the average treatment effect (ATE).

Our treatment group comprises of firms supplying goods that were not exempt from GST and had a turnover above INR 20 lakhs as these firms were required to bear the compliance cost associated with filing GST returns. Our control group comprises of firms supplying goods exempt from GST or had a turnover below INR 20 lakhs or both. Note that the control group in our baseline empirical strategy includes two sets of firms that were partially treated. The first set of firms supply products exempt from GST but have revenues above INR 20 lakhs, due to which these firms still have to bear GST compliance costs. The second

[5]See Notification No. 2/2017-Central Tax (Rate) issued by the Ministry of Finance of the Government of India on June 28, 2017.

set of firms have revenues lower than INR 20 lakhs but supply products that are not exempt from GST; changes in the tax structure itself may have induced a behavioral response in such firms. Thus, including these two sets of firms in our analysis can bias the treatment effects. To purge the control group of such partially treated firms, we also consider a truncated sample where we exclude these firms.

4.2 Validity of the DD Design

To check if the DD strategy is successful in this context, we need to ensure that the counterfactual trends in treatment and control states are identical. To test this assumption, we follow Autor (2003). Specifically, we run:

$$y_{i,t} = \sum_{l=-3}^{0} \beta_l \text{GST}_{i,t}(t = 2018 + l) + \phi_i + \psi_t + \epsilon_{i,t}.$$

A test of the parallel trends assumption is $\beta_l = 0 \ \forall l < 0$, i.e., the coefficients on all leads of the treatment should be zero. Figure 6 reports the results, which verify the DD assumption.

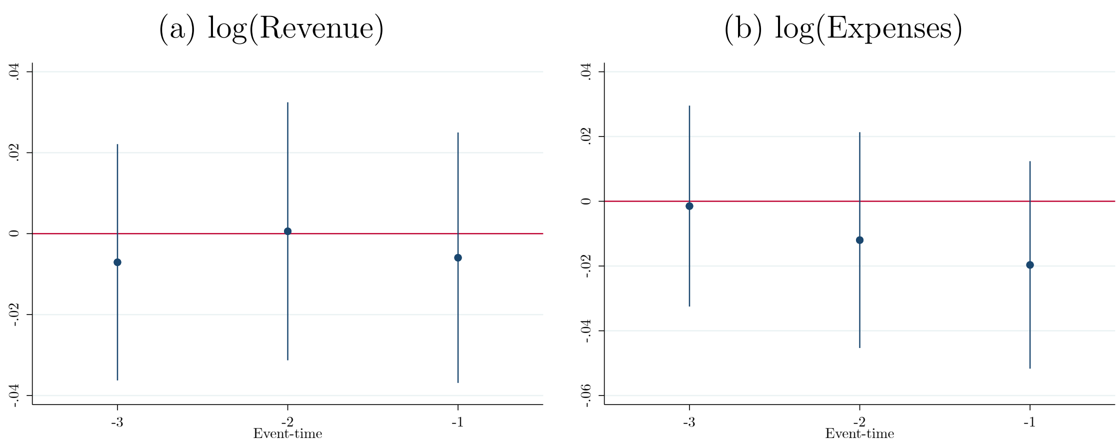


Figure 4: Testing the Parallel Trend Assumption

4.3 Results

Despite the negative correlations presented in Table 1, our DD estimates show that both revenues and costs grew after the implementation of GST; see Table 2. We find that relative to other firms, fully treated firms increased reported revenues and expenses by 140.3 percent and 98.6 percent, respectively. Next, we focus on the truncated sample that excludes partially treated firms from our control group, i.e., we exclude observations where firms have revenues above INR 20 lakhs and supply tax-exempted products and observations where firms have revenues below INR 20 lakhs and supply non-tax-exempted products. This allows us to contrast the experience of fully treated firms (i.e., those that have revenues above INR

20 lakhs and supply non-exempted goods) to fully untreated firms (i.e., only those firms that have revenues below INR 20 lakhs and supply exempted goods). Using our truncated sample, we find that the GST increased reported revenues by 245.1 percent and reported expenses by 194.4 percent.

Table 2: Effect of Treatment on Firm Revenues and Costs

Sample of firms	Full sample		Truncated sample	
Outcome variable	log(Revenue)	log(Expenses)	log(Revenue)	log(Expenses)
GST	1.403***	0.986***	2.451***	1.944***
	(0.0701)	(0.0525)	(0.411)	(0.352)
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Observations	85842	86866	81799	81793

Notes: In the truncated sample, we exclude observations satisfying $r_{i,t}$ < INR 20 lakhs and $p_i \in$ Non-exempted goods and observations satisfying $r_{i,t} \geq$ INR 20 lakhs and $p_i \in$ Exempted goods. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

5 Revenue Underreporting

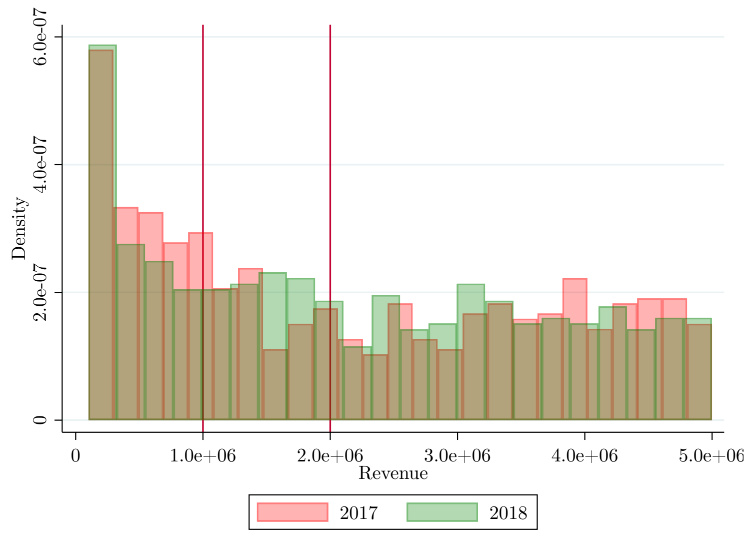
A firm with relatively modest input costs may find it desirable to restrict the scale of its operations or misreport sales. In this case, bunching occurs, where a firm keeps its reported taxable turnover just below the registration threshold. Figure 5 plots the histogram of firm revenues in the tax filing year prior and post the implementation of the GST. In the old tax, firms with a turnover of below INR 10 lakhs were exempt from GST. In the new regime, this exemption threshold was increased to INR 20 lakhs. In response, we see that the excess mass of firms below the INR 10 lakhs threshold moves closer to the new threshold of INR 20 lakhs.

5.1 Empirical Strategy

We follow Saez (2010), Kleven and Waseem (2013), and Chetty et al. (2011) to quantify the extent of bunching at indirect tax exemption thresholds. In particular, we denote the number of firms in a discrete revenue bin j by c_j, and the midpoint of revenue in the respective bin as r_j. We estimate the counterfactual density by fitting a p degree polynomial to these counts, excluding observations in a range $[r_L, r_U]$ around the exemption threshold T:

$$c_j = \sum_{i=0}^{p} \beta_i (r_j)^i + \sum_{i=r_L}^{r_U} \gamma_i \mathbb{1}[r_j = i] + \epsilon_j.$$

The excess number of firms who locate near the kink relative to the counterfactual density is given by $B = \sum_{j=r_L}^{r_H} (c_j - \hat{c}_j)$, where $\hat{c}_j = \sum_{i=0}^{p} \hat{\beta}_i (r_j)^i$ denotes the estimated counterfactual



Notes: We restrict attention to firms with revenues above INR 1 lakhs and below INR 50 lakhs. The VAT exemption threshold was changed from INR 10 lakhs to INR 20 lakhs under the new tax regime.

Figure 5: Distribution of Revenues

density. We define the excess mass around the bunch point relative to the average density of the counterfactual revenue distribution between r_L and r_H:

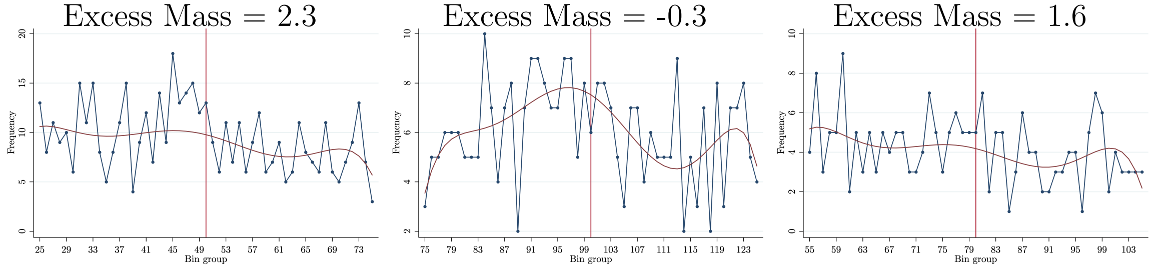
$$b = \frac{B}{\sum_{j=r_L}^{r_H} \hat{c}_j/(r_H - r_L)}.$$

5.2 Results

Figure 6 depicts the results. There is substantial excess mass (2.3) around the INR 10 lakhs bunch point prior to the implementation of GST. This is consistent with existing theoretical and empirical evidence on bunching below VAT exemption thresholds (Keen and Mintz, 2004; Liu et al., 2019). Moreover, there is negative excess mass (-0.3) around INR 20 lakhs in the pre-GST period. After the policy change, however, we see a large increase in the excess mass around the INR 20 lakhs bunch point. In addition, the excess mass around the new threshold is about 30 percent lower than that around the old threshold, which suggests that third-party reporting under the GST substantially reduced revenue underreporting.

Next, we examine if the extent of bunching is dependent on the level of GST rates. Much of the extant literature focuses on bunching at a notch of the personal income tax schedule; see, for instance, Kleven and Waseem (2013) or Chetty et al. (2011). An analysis of the effect of tax rates on the extent of bunching is challenging in the context of personal income taxes since both the tax rate and the exemption eligibility are dependent on earnings. While Chetty et al. (2011) argue that larger kinks generate larger tax elasticities, their results could be confounded by the fact that individuals with different earnings profiles behave differently. In contrast, in the context of indirect taxes, tax rates are dependent on

(a) $t \leq 2017$, $\bar{r}$ = INR 10L Excess Mass = 2.3 (b) $t \leq 2017$, $\bar{r}$ = INR 20L Excess Mass = -0.3 (c) $t > 2017$, $\bar{r}$ = INR 20L Excess Mass = 1.6



Notes: These figures show the revenue distribution around the GST exemption (demarcated by the vertical red lines) for firms between between 2015-2019. The series shown in dots is a histogram of revenues. Each point shows the number of observations in a INR 20,000 bin for panels (a) and (b), and a INR 25,000 bin for panel (c). The solid line beneath the empirical distribution is a sixth-degree polynomial fitted to the empirical distribution excluding 25 bins above and below the cutoff. Leftmost (rightmost) bin in bunching windows is 6 (1) bins below (above) the bunch point. Firms with revenues below INR 1 lakh and above INR 50 lakhs are discarded. In panel (a), we consider observations on or before 2017 around the threshold of INR 10 lakhs. In panel (b), we consider observations on or before 2017 around the threshold of INR 20 lakhs. In panel (c), we consider observations after 2017 around the threshold of INR 20 lakhs.

Figure 6: Revenue Bunching at GST Exemption Thresholds

product type, while the exemption threshold is dependent on firm revenues. We exploit this variation in tax rates around the exemption threshold to estimate the effect of tax rates on bunching. In particular, we investigate if the excess mass in the revenue distribution around the GST exemption threshold is higher for firms selling products that are subject higher than average GST rates. Table 3 reports the results, which suggest that there is more revenue underreporting in the subsample of firms facing higher than average GST rates. This finding is robust to alternative measures of the counterfactual density.[6]

Table 3: Bunching at Exemption Thresholds and GST Rate

	Degree 4	Degree 5	Degree 6
Full Sample	2.031	2.026	1.696
Below average GST	0.718	0.718	0.855
Above average GST	3.002	3.002	2.234

Notes: This table the excess mass of the revenue distribution at the bunch point INR 20 lakhs relative to the counterfactual density using a polynomial of degrees four, five, and six. The sample is restricted to observation after 2017 in which revenues are larger than INR 1 lakh and less than INR 50 lakhs. The bin-width is set to INR 1.5 lakhs. The bunching window spans one bin to the left and right of the bunch point. When fitting the polynomial, we consider four bins to the left and to the right of the bunch point.

6 Cost Overreporting

In this section, we propose a novel approach to detect cost overreporting and argue that this margin of evasion increased after the implementation of GST.

[6]These results complement the analysis of Fisman and Wei (2004), who use a different approach but obtain similar conclusions. They measure evasion using discrepancies in export and import data, and find that extent of underreporting is highly correlated with the tax rate.

6.1 Conceptual Framework

We begin by showing that expenditure shares of inputs are orthogonal to revenues across the firm productivity distribution under fairly general conditions. We then present a model of cost overreporting to show that firms' incentives to inflate non-verifiable costs introduce discontinuities in expenditure shares of inputs around enforcement notches.

6.1.1 Constant Input Shares under Full Verification

There is a continuum of firms, indexed by i. Firm i can access the following technology

$$F_i(\{x_{ij}\}) \equiv z_i \prod_j (x_{ij})^{\alpha_{ij}},$$

where x_{ij} denotes input j of firm i, $\alpha_{ij} > 0\ \forall j$ and $\sum_j \alpha_{ij} \leq 1$. Firm i chooses $\{x_{ij}\}_j$ to maximize

$$z_i \prod_j (x_{ij})^{\alpha_{ij}} - \sum_j w_j x_{ij},$$

where w_j denotes the price of acquiring input j. We focus on equilibrium outcomes where there is free entry and exit due to which firms earn zero profits:

$$z_i \prod_j (x_{ij})^{\alpha_{ij}} - \sum_j w_j x_{ij} = 0\ \forall i.$$

In any equilibrium, firms equate $\forall j$ marginal revenues $(z_i \alpha_{ij} (x_{ij})^{\alpha_{ij}-1} \prod_{k \neq j} (x_{ik})^{\alpha_{ik}})$ with marginal costs (w_j), which implies

$$z_i \prod_j (x_{ij})^{\alpha_{ij}} = \frac{w_j x_{ij}}{\alpha_{ij}} \quad \forall j.$$

Combining the above expression with the zero profit condition, we find that expenditure shares of inputs are only dependent on model primitives: $\frac{w_j x_{ij}}{\sum_j w_j x_{ij}} = \alpha_{ij}$. Note that in equilibrium, expenditure shares of inputs do not respond to changes in firm productivity or revenues:

$$\frac{w_j x_{ij}}{\sum_j w_j x_{ij}} \perp z_i \prod_j (x_{ij})^{\alpha_{ij}}\ \forall i, j.$$

To see the implications of the above result more clearly, consider a setting where the only difference between firms' production technologies is their productivities. That is, α_{ij}'s do not vary across i, but z_i's do. In this case, even though firm revenues may vary across the firm type distribution, expenditure shares of inputs will be identical across firms. Thus, if the authorities can perfectly verify firm expenditures, we should observe identical expenditure shares across the distribution of firm revenues. This logic is the basis for the following proposition.

Proposition 1. *If $\alpha_{ij} = \alpha_j \ \forall i$ and inputs are verifiable, then reported expenditure shares of inputs are independent of revenues.*

6.1.2 Discontinuities in Reported Input Shares under Partial Verification

When inputs are partially verifiable, firms may have an incentive to inflate costs in order to reduce their corporate income tax burden. It is useful to analyze this cost reporting decision separately from production decisions in a stylized setting with both verifiable and non-verifiable input costs, which we denote by c^v and c^n respectively. The optimal cost reporting decision of a firm can be derived as the solution to the following problem:

$$\min_{\hat{c}^n} \mathbb{1}(r \geq c^v + \hat{c}^n)\tau(r - c^v - \hat{c}^n) + \chi\mathbb{P}(r \geq T)(\hat{c}^n - c^n)^2/2,$$

where r denotes the revenue of the firm, and $\hat{c}^n$ denotes reported non-verifiable costs. The scaler τ captures the corporate tax rate, and T denotes the GST exemption threshold. The indicator $\mathbb{1}(r \geq c^v + \hat{c}^n)$ rules out subsidies for reported losses.

All taxpayers registered under GST (i.e., those with reported revenues above the INR 20 lakhs exemption threshold) are required to file every month a form detailing all outward supplies made, input tax credit claimed, tax liability ascertained, and taxes paid. This provides tax authorities with rich information regarding firm revenues and expenses, increasing the probability of detecting financial statement fraud. We capture this by assuming that the audit probability, $\mathbb{P}(r \geq T)$, is positive when $r \geq T$ and zero otherwise. Moreover, the authorities may choose to allocate more audit officers to investigate the operations of firms that appear to be larger based on tax filings. We capture such size-dependent enforcement by allowing the audit probability to be conditional on reported revenues.[7] We model the penalty for cost misreporting using a quadratic cost function. In particular, we assume that the fine for misreporting costs is given by $\chi(\hat{c}^n - c^n)^2/2$, where $\chi > 0$. For ease of exposition, we set $\chi = 1$ for now. To obtain a better fit to the data, we allow this parameter to be free later when we structurally estimate the model.

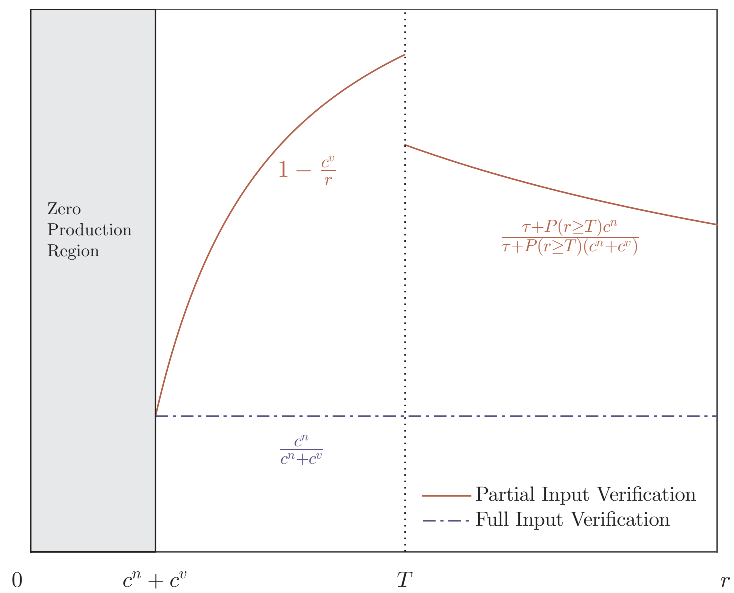
Given (r, c^v), the optimal report of non-verifiable costs when revenues are below the exemption threshold is $\hat{c}^n(r < T) = r - c^v$. When revenues are above the exemption threshold, in contrast, the optimal reporting choice of the firm is $\hat{c}^n(r \geq T) = \frac{\tau}{\mathbb{P}(r \geq T)} + c^n$. Notice that $\hat{c}^n(r < T) > \hat{c}^n(r \geq T)$ if $\pi > \frac{\tau}{\mathbb{P}(r \geq T)}$. That is, firms have a larger incentive to inflate non-verifiable costs below the GST threshold if the fraud detection probability is

[7] To save notation, we suppress this dependence in the exposition.

sufficiently high. We summarize this discussion in the following proposition.

Proposition 2. *Suppose a positive fraction of inputs is non-verifiable, and the fraud detection probability is sufficiently high and increasing in revenues. Then the reported expenditure share of non-verifiable inputs is increasing in revenues below the GST exemption threshold, exhibits a negative jump at the threshold, and is decreasing in revenues above the threshold.*

Figure 7 graphically illustrates these results.



Notes: We assume $P(r \geq T) = r/\bar{r}$ and use the following parameterization for the numerical illustration: $(c^n, c^v, T, \tau, \bar{r}) = (0.1, 0.4, 1.5, 0.5, 3)$.

Figure 7: Reported Input Shares as a Function of Revenues

6.2 Empirical Strategy

We use a sharp RD design where we examine jumps in the ratio of non-verifiable expenses to total expenses around tax exemption thresholds. The RD design (Hahn et al., 2001, Imbens and Lemieux, 2008) exploits a discontinuity in the treatment assignment to identify a causal effect. It can be used when treatment assignment is determined on the basis of a cutoff score, s, on an observed forcing variable. The forcing variable in this design is excess revenue relative to the time-specific tax exemption thresholds. We construct a score for the treatment as follows:

$$s_{i,t} \equiv \begin{cases} r_{i,t} - \text{INR 10 lakhs} & t \leq 2017 \\ r_{i,t} - \text{INR 20 lakhs} & t > 2017 \end{cases}.$$

We consider the following specification for estimating the RD treatment effect:

$$y_{i,t} = \alpha + \beta \mathbb{1}(s_{i,t} \geq 0) + f(r_{i,t}) + \epsilon_{i,t} \quad \forall r_{i,t} \in (0, \text{INR 50 lakhs}),$$

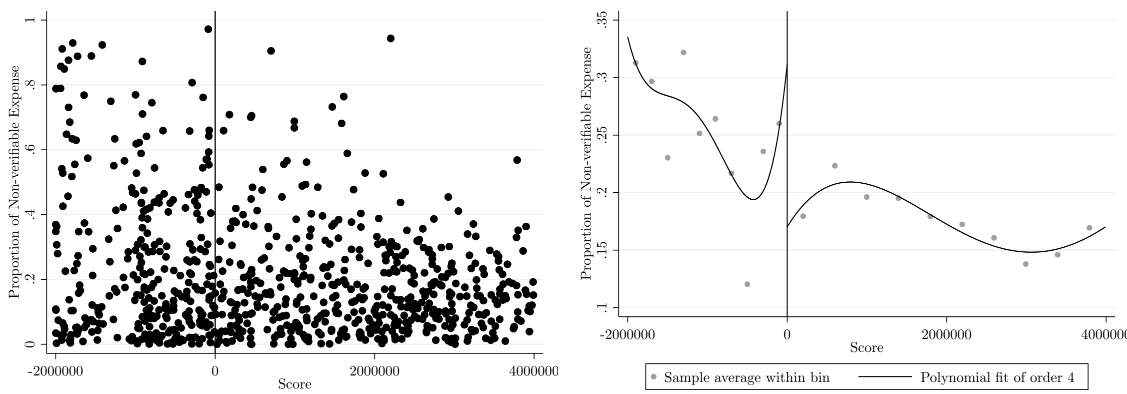
where $y_{i,t}$ is the outcome variable (i.e., non-verifiable expense ratio) and f is continuous function. As discussed above, our main identifying assumption is that the production technologies of firms with revenues around the exemption thresholds are similar, barring differences in productivity, i.e., $\alpha_{ij} = \alpha_j \ \forall i : r_{i,t} \in (0, \text{INR 50 lakhs}) \ \forall t$.

We proxy non-verifiable costs using employee benefits expenses, which includes all forms of compensation given by an enterprise in exchange for service rendered by employees. Wage employment is composed of salaried and casual wage employment. The tax authorities can track wages of salaried workers by matching contributions to the Employees' Provident Fund Scheme, which is obligatory for all firms with at least 20 employees. However, about 62 percent of wage employees are casual workers–consisting mainly of people from poorer households engaged in irregular work–compensated on a daily basis (ILO, 2018a). Casual wages lie outside the purview of tax authorities and are in stark contrast to expenditures on intermediate goods (which can be tracked by invoice matching) and salaried wages (which can be tracked using compulsory pension payments). We focus on firm-year observations with revenues less than INR 50 lakhs, and discard outliers that are respectively below and above the 5th and 95th percentile of observations of non-verifiable expenses and total expenses. We also drop observations that feature non-verifiable expense ratios larger than one. Our baseline analysis is focused on the years 2017 and 2018.

6.3 Results

Panel (a) of Figure 8 depicts the relationship between the non-verifiable expense ratio and the score. Though this plot is helpful to detect outliers, its effectiveness for visualizing the RD design is limited. In panel (b) of Figure 8, we plot the binned outcome means shown against the score, adding a fourth-order global polynomial fit estimated separately for treated and control observations. The global fit reveals that the observed regression function seems to be non-linear. Thus, Proposition 1 suggests that all input costs are not verifiable. Furthermore, the global fit is in line with Proposition 2. In particular, the function is (locally) increasing to the left of the exemption threshold. The plot also reveals a negative jump at the cutoff: the average non-verifiable expense ratio is higher in firm-year observations right below the exemption threshold than in those above. Lastly, the function seems to be decreasing for a large subset of the domain above the threshold.

Table 4 reports the RD estimates using a linear estimator with a triangular kernel and MSE-optimal bandwidth. In panel A of the table, we present RD estimates without accounting for the set of covariates. The results indicate that within the bandwidth, the ratio of non-verifiable expense to total expenses decreases by 11.8 percent when revenues are above the tax exemption threshold. The estimated effect of the treatment is not only large, but



(a) Scatterplot (b) RD Plot

Figure 8: Effect of Exemption Threshold on Non-verifiable Expense Ratio

also statistically significant at conventional levels. The estimated change in the non-verifiable expense ratio remains negative, but becomes imprecise, when we consider the full sample period. In the final two columns, we restrict attention to the sample of firms that supply products that are not exempt from GST; Our results are robust to this change. In addition, our point estimates also do not change much when we control for the set of covariates. Our results suggest that firms with revenues around the GST exemption responded to the tax change by overreporting their costs.

Table 4: Exemption Threshold and Non-verifiable Expense Ratio

Sample of firms	Full sample		Non-exempted firms	
Sample period	2017-18	2015-20	2017-18	2015-20
Panel A: Excluding covariates				
RD estimate	-0.11784**	-0.04573	-0.1185**	-0.0486
	(0.05407)	(0.03104)	(0.05438)	(0.03133)
	[804]	[2094]	[790]	[2043]
Panel B: Including covariates				
RD estimate	-0.11225**	-0.02982	-0.11411**	-.03578
	(0.0523)	(0.03405)	(0.05257)	(0.03324)
	[717]	[1855]	[703]	[1807]

Notes: To compute the RD estimates, we use a linear estimator with a triangular kernel and MSE-optimal bandwidth. Standard errors are reported in parenthesis and are clustered at the firm level. Observations are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5 reports additional RD estimates for different bandwidths and kernel functions. Variation in kernels is ordered by column and bandwidths by row. Our baseline results are derived using the MSE-optimal bandwidth. It is well known that as the bandwidth increases, the bias of the local estimator increases, and its variance decreases. In the second row of the table, we increase the bandwidth by 50% of the MSE-optimal bandwidth and find a similar estimate with smaller standard errors. We also investigate the sensitivity of our estimates to the choice of kernel functions. In addition to the triangular kernel, we use Epanechnikov and uniform kernels to construct the estimators. The estimates are broadly consistent with our baseline empirical findings in that they are negative and statistically significant.

Table 5: Alternative RD Specifications

Kernel function	Triangular	Epanechnikov	Uniform
h_{MSE}	-0.11765**	-0.1155*	-0.10869*
	(0.05382)	(0.05709)	(0.05948)
$h_{MSE} \times 1.5$	-0.09651**	-0.09156**	-0.0479***
	(0.0444)	(0.04335)	(0.04119)

Notes: Standard errors are clustered at the firm level. h_{MSE} denotes the MSE-optimal bandwidth. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In Table 6, we present RD estimates corresponding to subsamples in which we restrict attention to observations pre, post and during the implementation of GST. The results show that cost overeporting increased when the GST was implemented. However, these effects were not persistent.

Table 6: Heterogeneity in RD Estimates (by Year)

Sample of firms	Full sample			Non-exempted firms		
Sample period	2017	2018	2019	2017	2018	2019
RD estimate	-0.02086	-0.14432*	0.02027	-0.01731	-0.14523*	0.01998
	(0.10645)	(0.0845)	(0.07444)	(0.10736)	(0.0847)	(0.07201)
	[451]	[353]	[290]	[441]	[349]	[280]

Notes: To compute the RD estimates, we use a linear estimator with a triangular kernel and MSE-optimal bandwidth. Standard errors are reported in parenthesis and are clustered at the firm level. Observations are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.4 Validity of the RD Design

6.4.1 Continuity of the Score Density

Sorting or bunching of firms along the threshold can bias the RD results. An underlying assumption in the RD design that the number of treated observations just above the cutoff should be approximately similar to the number of control observations below it. This could be a result of firms not having the ability to precisely manipulate the value of the score that they receive. Even if firms actively attempt to affect their score by underreporting revenues, in the absence of precise manipulation, random change can place roughly the same amount of firms on either side of the cutoff, leading to a continuous probability density function when the score is continuously distributed. RD applications where there is an abrupt change in the number of observations at the cutoff tend to be less credible.

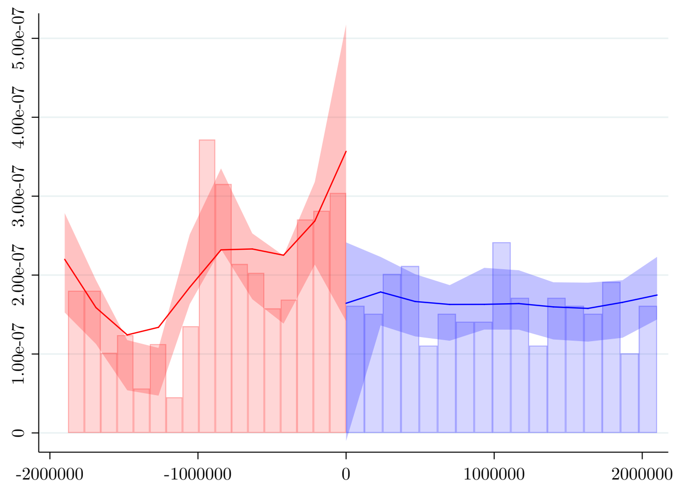


Figure 9: Estimated Density of Running Variable

We follow Cattaneo et al. (2017) to test that the density of the running variable is continuous at the cutoff. The null hypothesis is that there is continuity of the density functions for control and treatment units at the cutoff. Therefore, failing to reject implies that there is no statistical evidence of manipulation at the cutoff, and offers evidence supporting the validity of the RD design. We fail to reject the null hypothesis of no difference in the density of treated and control observations at the cutoff at the 5% significance level. Figure 9 provides a graphical representation of the continuity in density test approach, exhibiting both a histogram of the data and the actual density estimate with shaded 95% confidence intervals. The test proposed from Cattaneo et al. (2017) descends from McCrary (2008), but requires fewer choices of tuning parameters, and takes fuller advantage of local polynomial regression. For robustness, we also check for discontinuities using McCrary's version of the test; The bandwidth and binsize for the test were chosen based on the algorithm outlined in Section III.B of McCrary (2008). Again, we fail to reject the null hypothesis that the discontinuity is zero at the 5% significance level.

6.4.2 Covariate Balance

In Table 7, we inspect the control variables at the discontinuity. Each row presents RD estimates at the exemption threshold using a linear estimator with a triangular kernel. Our first set of results, reported in panel A of the table, employ MSE-optimal bandwidths. All 95% robust confidence intervals contain zero, with p-values ranging from 0.354 to 0.846. In other words, we do not find any evidence of discontinuous jumps at the cutoff. For falsification purposes, it may be more appropriate to use CER-optimal bandwidths as we are mostly interested in inference and not the point estimates. CER-optimal bandwidths leads to lower size distortions than tests implemented using the MSE-optimal bandwidths.

In panel B of the table, we show that using CER-optimal bandwidths does not alter our conclusions.

Table 7: Testing Balance of Covariates around GST Exemption Thresholds

Variable	Optimal Bandwidth	RD Estimator	p-value	Confidence Interval
Panel A: MSE-optimal bandwidth				
Total assets	7.91×10^5	67.82	0.354	[-92.98, 259.90]
Borrowings	3.88×10^5	-111.72	0.846	[-1110.22,1354.81]
Equity	6.52×10^5	-51.89	0.561	[-213.03, 115.52]
Cash	3.73×10^5	404.79	0.346	[-2405.13, 6857.65]
Panel B: CER-optimal bandwidth				
Total assets	2.71×10^5	181.28	0.438	[-1761.30, 4068.60]
Borrowings	2.81×10^5	-460.03	0.424	[–1162.55, 489.21]
Equity	4.73×10^5	-60.41	0.569	[-260.92,143.54]
Cash	5.74×10^5	93.41	0.348	[-110.63, 314.20]

Notes: All variables are denominated in INR Crore. To compute the RD estimates, we use a linear estimator with a triangular kernel. Standard errors are reported in parenthesis and are clustered at the firm level.

6.4.3 Sensitivity to Observations near the Exemption Threshold

In this section, we investigate the possibility of firms manipulating their revenues (i.e., the score in our RD design). If systematic manipulation of score values has occurred, it is natural to assume that the units closest to the cutoff are those most likely to have engaged in manipulation. To test if this is the case, we use a donut-hole approach and investigate the sensitivity of our RD estimates to the response of firms with revenues close to the GST exemption threshold. In Table 8, we exclude observations with $\mid r_{i,t} \mid < \Lambda$ and recompute MSE-optimal bandwidths for $\Lambda \in \{5000, 10000, 15000\}$. The conclusion of our baseline analysis remains largely unchanged as both the original and the new estimated effects are significant at 10% level.

Table 8: RD Estimation for the Donut-Hole Approach

Donut Hole Radius (in INR)	0	5000	10000	15000
RD estimate	-0.11784**	-0.11844**	-0.12252*	-0.13009*
	(0.05407)	(0.05906)	(0.06646)	(0.06905)
	[804]	[803]	[800]	[799]

Notes: To compute the RD estimates, we use a linear estimator with a triangular kernel and MSE-optimal bandwidth. Standard errors are reported in parenthesis and are clustered at the firm level. Observations are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.4.4 Placebo Tests

Another illustrating validity test for the RD design is to compare our baseline results with placebos outcomes. In Table 9, we examine discontinuity in the non-verifiable expense ratio at the INR 20 lakhs cutoff for the pre-GST period, and at the INR 10 lakhs cutoff for the post-GST period. Since the tax exemption thresholds were revised from INR 10 lakhs to INR 20 lakhs under the GST, we should not see any discontinuous pattern at the respective cutoffs for these placebo outcomes. The point estimates for both placebo outcomes are insignificant at conventional levels.

Table 9: Placebo Tests: Alternative Exemption Thresholds

Sample period	Pre-GST	Post-GST
Running variable	$r_{i,t}$ − INR 20 lakhs	$r_{i,t}$ − INR 10 lakhs
RD estimate	-0.04156	0.11846
	(0.06463)	(0.17682)
	[1451]	[643]

Notes: To compute the RD estimates, we use a linear estimator with a triangular kernel and MSE-optimal bandwidth. Standard errors are clustered at the firm level. Observations are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In Table 10, we examine discontinuity in the share of verifiable expenses at the tax exemption thresholds. We consider several types of expenses: cost of materials, power and fuel costs, finance costs for debt servicing, insurance expenses, and payments to auditors. The greater visibility of these expenses stems from the fact that they are due to a third party, and, thus, evasion on these margins would be more costly as it would require firms to collude. The 90 percent confidence intervals for all these placebo outcomes capture zero, further supporting our identification strategy.

Table 10: Placebo Tests: Share of Verifiable Expenses

	Cost of Materials	Power and Fuel Costs	Finance Cost	Insurance Expenses	Auditing Expenses
RD estimate	0.04216	0.06197	-0.03953	0.00309	0.00027
	(0.10593)	(0.04357)	(0.03769)	(0.00277)	(0.00254)
	[788]	[755]	[805]	[755]	[755]

Notes: To compute the RD estimates, we use a linear estimator with a triangular kernel and MSE-optimal bandwidth. Standard errors are clustered at the firm level. Observations are reported in square brackets.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.5 Evidence on the Mechanism

Our mechanism for evasion shifting relies on differential visibility of intermediate goods and labor inputs. We test our proposed mechanism in this section by exploiting variation in the visibility of intermediate goods.

Pre-GST, intermediate goods were scrutinized at state borders, which made it harder for firms to hide these expenses. After the GST, border tax checkposts were abolished to facilitate the movement of goods. This reduced the visibility of intermediate goods sourced by a firm from states other than the one where it is located, even though the net impact of the GST on the visibility of these goods may have well been positive due to the advent of invoice matching. Furthermore, the larger the fraction of inputs sourced from out-of-state (which we refer to as the out-of-state input procurement ratio), the larger the marginal decrease in the visibility of these goods due to lack of scrutiny at state borders. Thus, to reduce their corporate tax bill, firms with higher out-of-state input procurement ratios have a larger incentive to overreport the cost of their intermediate goods, and a lower incentive to overreport their labor costs relative to other expenses.

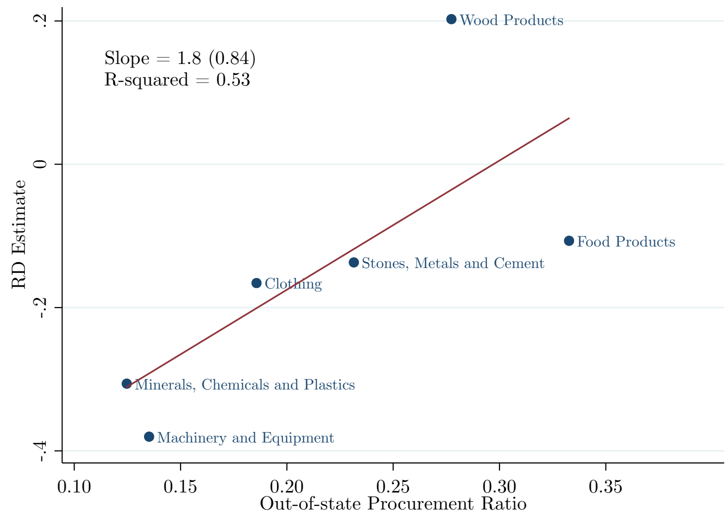
Table 11 provides evidence in support of our proposed mechanism. Post-GST, firms with higher than average out-of-state input procurement ratios did not shift to overreporting labor costs relative to other costs. However, we detect a discontinuity in the non-verifiable expense ratio for firms with below average out-of-state input procurement ratios, which goes in the expected direction. Moreover, we do not find any evidence of discontinuities in the non-verifiable expense ratio for either set of firms before the GST was implemented. This suggests that our empirical results are driven by variations in the extent of visibility of certain inputs relative to others, which is consistent with our theory.

Table 11: Heterogeneity in RD Estimates (by Out-of-state Input Procurement)

Sample period	Post-GST		Pre-GST	
Sample of firms	High OPR	Low OPR	High OPR	Low OPR
RD estimate	0.10754	-0.28084**	-0.00281	0.84957
	(0.14725)	(0.14217)	(0.05847)	(0.58084)
	[122]	[153]	[126]	[186]

Notes: High (low) out-of-state input procurement ratio (OPR) refers to the subsample of firms with OPR greater than or equal to (less than) mean values. To compute the RD estimates, we use a linear estimator with a triangular kernel and MSE-optimal bandwidth. Standard errors are clustered at the firm level. Observations are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

These findings are echoed in Figure 10, which breaks down the sample of firms based on standardized (output) product classifications. These estimates should be taken with a grain of salt as they lack precision at this level of disaggregation. Nevertheless, the figure shows that firms that supply products whose inputs were primarily sourced from the state where the products were manufactured were more disposed to overreporting their labor costs relative to other expenses after the GST was implemented. Firms that fall into this category are those supplying machinery, equipment, minerals, chemicals and plastics. This is not surprising given the fact that these firms belong to heavy industries and thus a key factor determining the location of their plants is the distance from their primary input source. Contrastingly, firms supplying food products, which procure a substantial share of their inputs from out-of-state, exhibit a smaller negative jump in their reported wage bills relative to other costs.



Notes: We restrict attention to the post-GST period when computing the RD estimates. We define industry classifications as follows: "Food and Animal Products" includes firms with codes for Indian Trade Clarification based on Harmonized System (ITC-HS) below 2500; "Mineral, Chemicals and Plastics" includes firms with ITC-HS codes $\geq$ 2500 and $<$ 4100; "Wood Products" includes firms with ITC-HS codes $\geq$ 4400 and $<$ 5000; "Clothing" includes with ITC-HS codes $\geq$ 5000 and $<$ 6800; "Stones, Metals and Cement" includes firms with ITC-HS codes $\geq$ 6800 and $<$ 8400; "Machinery and Equipment" includes firms with ITC-HS codes $\geq$ 8400 and $<$ 9400. Since the ASI records data on only manufacturing firms, we discard service sector firms from our analysis (i.e., those with ITC-HS codes $\geq$ 9900). We also exclude firms classified as supplying hides & skins (i.e., those with ITC-HS codes $\geq$ 4100 and $<$ 4400), miscellaneous products (i.e., those with ITC-HS codes $\geq$ 9400 and $<$ 9700), works of art (i.e., those with ITC-HS codes $\geq$ 9700 and $<$ 9900) as we do not have enough observations to perform bandwidth calculations for our RD estimates.

Figure 10: Variation in RD estimates by Out-of-state Input Procurement

Our results contribute to Almunia and Lopez-Rodriguez (2018), who also exploit notches in enforcement intensity to study the effects of tax enforcement policies in Spain. However, the nature of the variation in enforcement intensity they consider differs from what we study here. They use the percentage of sales made to final consumers (whose transactions are harder to cross-check against other information sources) in each sector as a proxy for the visibility of sales made by firms in that sector. Moreover, they focus on how enforcement intensity affects revenue underreporting. Thus, the final consumption ratio is a suitable proxy for enforcement in their framework as it captures variations in the visibility of firm revenues. In contrast, we are interested in how enforcement notches affect cost overreporting. Thus, we require an enforcement intensity measure that captures variation in the visibility of certain firm inputs relative to others.

6.6 Heterogeneity

6.6.1 Industry and Factor Intensity

In Table 12, we run our estimation procedure over subsamples of firms belonging to specific industries. We find that our baseline results are primarily driven by firms in the service sector. For service providers, the share of non-verifiable expenses jumps down by 12.8 percent at the tax exemption thresholds, which is statistically significant. For non-service providers, which includes agricultural and manufacturing firms, the 90 percent confidence intervals for the RD estimate capture zero. We see a similar picture when we split the sample based on differences in factor intensity. We find no evidence of cost overreporting for capital-intensive firms (i.e., those with a material cost to expense ratio of larger than average). We find that our baseline results are driven by labor-intensive firms, which seems natural since wages comprise a substantial share of such firms' expenses, so there is more room to evade.

Table 12: Heterogeneity in RD Estimates (by Industry and Factor Intensity)

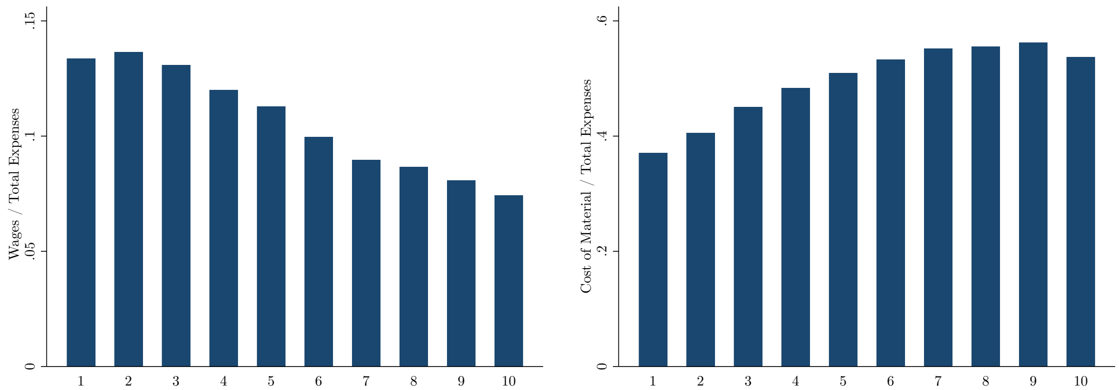
		by industry		by factor intensity	
Sample of firms	All	Service providers	Non-service providers	Labor-int. firms	Capital-int. firms
RD estimate	-0.11784**	-0.12772**	0.17855	-0.14471*	-0.02016
	(0.05407)	(0.06012)	(0.29219)	(0.07836)	(0.05398)
	[804]	[672]	[132]	[514]	[290]

Notes: Service providers are classified under heading numbers above 9900 as per the NPCS. Capital-intensive firms are firms that a ratio of the cost of materials to total expenses above average; labor-intensive firms are the residual. To compute the RD estimates, we use a linear estimator with a triangular kernel and MSE-optimal bandwidth. Standard errors are reported in parenthesis and are clustered at the firm level. Observations are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.6.2 Implications for Large Firms

Our identification strategy relies on the discontinuities in the behavior of firms around tax exemption thresholds; These firms are relatively small in size. However, one can extrapolate the impact of the GST on the compliance of large firms by examining the characteristics of small firms that overreport costs. In particular, we showed that our RD estimates for cost overreporting are driven by service providers that are typically labor intensive. In contrast, larger firms tend to be capital intensive. This can be seen clearly in Figure 11, which plots the binned means of wage bills and the cost of materials (as a share of total expenses) across revenue deciles. Moreover, we do not detect any fraud in reports concerning the cost of materials; see placebo test in the first column of Table 10. Thus, our results suggest that the regime change did not have an adverse impact on compliance among large firms.

(a) Labor Intensity (b) Capital Intensity



Notes: The x-axes measure revenue deciles.

Figure 11: Factor Intensity by Firm Size

6.7 Estimating the Level of Cost Overreporting

Marginal changes in cost overreporting at exemption thresholds can be directly obtained from our RD analysis. However, a reduced-form approach is not informative about the underlying *level* of cost overreporting. To compute this, we need to structurally estimate our model.

We compute the level of cost overreporting by

$$\Omega(c^v, c^n, \chi, \tau, T; P) \equiv \int_{c^n+c^v}^{T} (r - c^v)dr + \int_{T}^{\bar{r}} \left\{ \frac{\tau}{\chi \mathbb{P}(r \geq T)} + c^n \right\} dr - (\bar{r} - c^n - c^v)c^n,$$

where $\bar{r}$ denotes the upper bound of the revenue domain.[8] To see this, consider panel (a) of Figure 12, which compares reported and actual non-verifiable expenses over the revenue domain $[c^n + c^v, \bar{r}]$. Note that the first two terms of the above expression capture the area under non-verifiable expenses under hidden information, while the last term captures the area under non-verifiable expenses under complete information.

To estimate Ω, we use Generalized Method of Moments (GMM). Specifically, we first estimate firm expenses, $\{c^n, c^v\}$, and the penalty parameter, χ, using

$$(c^{n*}, c^{v*}, \chi^*) \in \text{argmin}_{(c^n, c^v, \chi) \in \mathbb{R}^3_+} \mathcal{D}(c^n, c^v, \chi \mid \tau, T)' W \mathcal{D}(c^n, c^v, \chi \mid \tau, T),$$

where W is a weighting matrix. Here $\mathcal{D}$ is a vector that collects distances between binned means of non-verifiable expense ratios that we obtain from our RD analysis and the corre-

[8]Here we implicitly assume that the distribution of firm revenues is uniform over $[c^n + c^v, \bar{r}]$. We relax this assumption in Section 6.7.2.

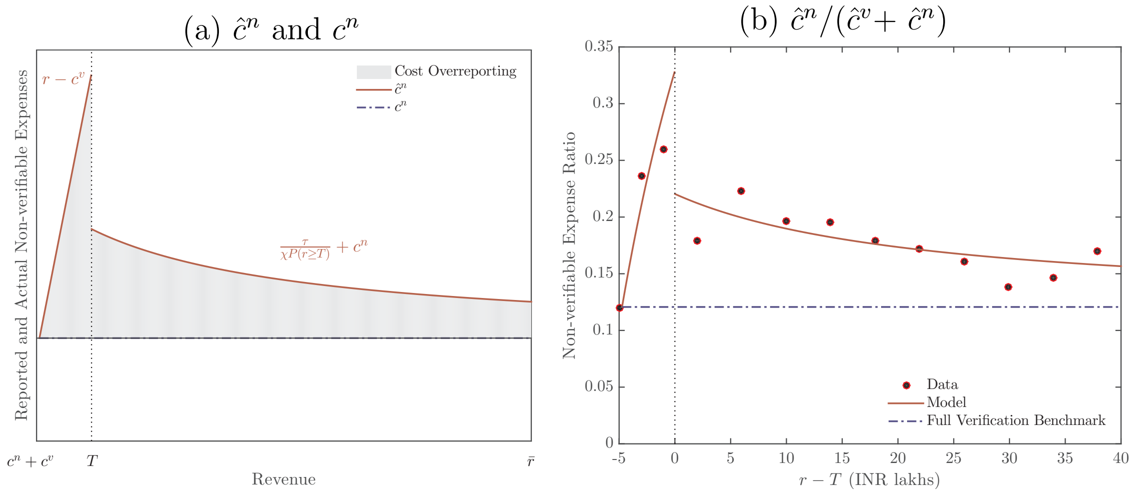


Figure 12: Estimating the Level of Cost Overreporting

sponding model moments. Specifically, the i^{th} entry of $\mathcal{D}$ is given by

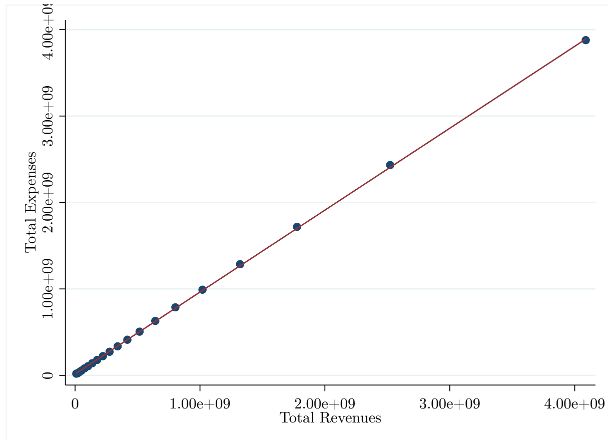
$$\mathcal{D}_i(c^n, c^v, \chi \mid \tau, T) \equiv \begin{cases} 1 - \frac{c^v}{\hat{r}_i} - \frac{\hat{c}_i^n}{\hat{c}_i^n + \hat{c}_i^v} & \text{if } \hat{r}_i < T \\ \frac{\tau + \chi \mathbb{P}(\hat{r}_i \geq T) c^n}{\tau + \chi \mathbb{P}(\hat{r}_i \geq T)(c^n + c^v)} - \frac{\hat{c}_i^n}{\hat{c}_i^n + \hat{c}_i^v} & \text{if } \hat{r}_i \geq T \end{cases},$$

where a hat over a variable denotes reported values. We map the running variable back to revenue bins by adjusting them upwards by INR 20 lakhs, and normalize the revenue bins by dividing by INR 10 lakhs. In the model, we assume that the audit probability is linear in revenue, i.e., $\mathbb{P}(r_i > T) = r_i/\bar{r}$. We set the upper bound of the revenue domain in the model, $\bar{r}$, to match the corresponding upper bound in the running variable in our RD analysis. We set the corporate tax rate and the GST exemption threshold to those observed in reality. Using this estimation procedure under the identity weighting matrix, we obtain $(c^{n*}, c^{v*}, \chi^*) = (0.18, 1.34, 3.84)$. Panel (b) of Figure 12 shows that our model does a good job in matching the data. We then use these results to estimate the underlying level of cost overreporting. Note that actual costs around the exemption threshold are given by $(\bar{r} - c^{n*} - c^{v*})(c^{n*} + c^{v*})$. Hence, costs were overreported by $\frac{\Omega(c^{v*}, c^{n*}, \chi^*, \tau, T; P)}{(\bar{r} - c^{n*} - c^{v*})(c^{n*} + c^{v*})} = 7.9$ percent by firms around the exemption threshold.

6.7.1 Endogenous Cost Structure

In the data, expenses are highly correlated with revenues; see Figure 13. In our baseline model of cost overreporting, however, we treat actual costs as fixed around the exemption threshold. Though our RD analysis does not rely on this assumption, it has important implications for the underlying level of cost overreporting. We relax this assumption in this section by allowing costs to endogenously vary with revenues.

Recall from our analysis of the production stage, input costs are a constant fraction of



Notes: We discard observations with negative revenues, as well as outliers that are respectively below and above the 5th and 95th percentile of observations of expenses.

Figure 13: Expenses and Revenues

revenues:

$$c^j = r\alpha_j \ \forall j \in \{n, v\}.$$

This implies that reported non-verifiable expenses are given by

$$\hat{c}^n = \begin{cases} r(1-\alpha_v) & \text{if } r < T \\ \frac{\tau}{\chi \mathbb{P}(r \geq T)} + r\alpha_n & \text{if } r \geq T \end{cases}.$$

Hence, the level of cost overreporting around the tax exemption threshold in this setting is given by

$$\int_0^T r(1-\alpha_v)dr + \int_T^{\bar{r}} \left\{ \frac{\tau}{\chi \mathbb{P}(r \geq T)} + r\alpha_n \right\} dr - \frac{\bar{r}^2 \alpha_n}{2}.$$

This is depicted in panel (a) of Figure 14.

We use a two-step estimation procedure to estimate the level of cost overreporting in this setting. First, note that when costs are treated as endogenous, then the reported non-verifiable expense ratio before the exemption threshold is equal to $(1-\alpha_v)$. We set $\alpha_v = 0.74$ to match the observation that the average reported non-verifiable expense ratio conditional on revenues being less than the exemption threshold is 0.26. We then estimate (α_n, χ) using a GMM estimation procedure analogous to our baseline approach that minimizes distances between binned means of the non-verifiable expense ratio around the exemption threshold in the data and the corresponding moments in the model; see panel (b) of Figure 14. This procedure yields $\alpha_n^* = 0.11$ and $\chi^* = 3.21$, which we use to estimate the underlying level of cost overreporting. Using this alternative method, we find that firms around the exemption threshold overreported costs by 5.2 percent. To see why this estimate is lower than our baseline estimate, note that here the marginal increase in the level of cost overreporting in

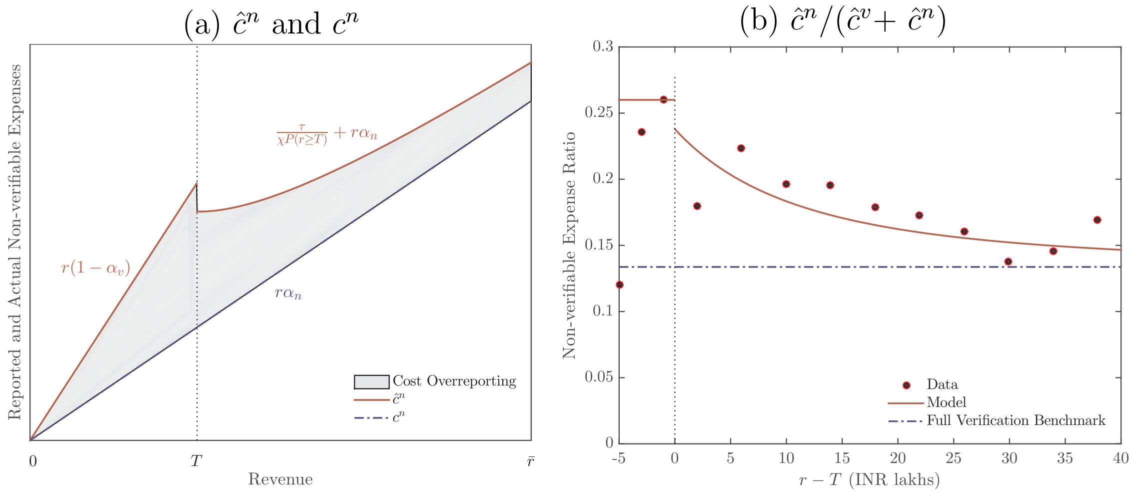


Figure 14: Estimating the Level of Cost Overreporting under Endogenous Cost Structure

response to a unit increase in revenue below the exemption threshold is $1 - \alpha_v - \alpha_n$, which is strictly lower than that in our baseline model. This is because verifiable costs rise with revenues when the former is endogenous, which reduces profits and thus the incentive to inflate non-verifiable costs.

6.7.2 Endogenous Production Responses

In our baseline model of cost overreporting, firms' production decisions are divorced from their evasion decisions. In this section, we estimate the level of cost overreporting while taking into account endogenous production responses.

The problem of a firm with productivity $z \in \mathcal{Z}$ is to choose verifiable and non-verifiable inputs, $\{x_v, x_n\}$, and report non-verifiable costs, $\{\hat{c}^n\}$, to maximize:

$$\underbrace{z x_v^{\alpha_v} x_n^{\alpha_n} - w_v x_v - w_n x_n}_{\text{Profits}} - \underbrace{\tau[z x_v^{\alpha_v} x_n^{\alpha_n} - w_v x_v - \hat{c}^n]}_{\text{Tax Liability}} - \underbrace{\mathbb{1}(r \geq T)\chi z x_v^{\alpha_v} x_n^{\alpha_n} (\hat{c}^n - w_n x_n)^2/(2\bar{r})}_{\text{Evasion Costs}}.$$

Optimal allocations can be characterized by the following 2x2 system in $\{x_v, x_n\}$:

$$\begin{aligned} z x_v^{\alpha_v}\{\alpha_n x_n^{\alpha_n - 1}[1 - \tau - \frac{\tilde{\chi}}{2}(\hat{c}^n - w_n x_n)^2] + \tilde{\chi} w_n x_n^{\alpha_n}(\hat{c}^n - w_n x_n)\} &= w_n, \\ z \alpha_v x_v^{\alpha_v - 1} x_n^{\alpha_n}[1 - \tau - \frac{\tilde{\chi}}{2}(\hat{c}^n - w_n x_n)^2] &= (1 - \tau) w_v, \end{aligned}$$

where

$$\hat{c}^n = \begin{cases} z x_v^{\alpha_v} x_n^{\alpha_n}(1 - \alpha_v) & \text{if } \tilde{\chi} = 0 \\ \frac{\tau}{z x_v^{\alpha_v} x_n^{\alpha_n} \tilde{\chi}} + w_n x_n & \text{if } \tilde{\chi} > 0 \end{cases},$$

and $\tilde{\chi} \equiv \mathbb{1}(r \geq T)\frac{\chi}{\bar{r}}$. Let the solution to this problem under the assumptions $\tilde{\chi} = 0$ and

$\tilde{\chi} > 0$ be denoted by $\{x_v^*(z) \mid_{\tilde{\chi}=0}, x_n^*(z) \mid_{\tilde{\chi}=0}\}$ and $\{x_v^*(z) \mid_{\tilde{\chi}>0}, x_n^*(z) \mid_{\tilde{\chi}>0}\}$, respectively. Define $\forall z \in \mathcal{Z}$:

$$r^*(z) \equiv \begin{cases} r(z) \mid_{\tilde{\chi}=0} & \text{if } r(z) \mid_{\tilde{\chi}=0} < T \\ r(z) \mid_{\tilde{\chi}>0} & \text{if } r(z) \mid_{\tilde{\chi}>0} \geq T \end{cases},$$

where

$$\begin{aligned} r(z) \mid_{\tilde{\chi}=0} &\equiv z(x_v^*(z) \mid_{\tilde{\chi}=0})^{\alpha_v}(x_n^*(z) \mid_{\tilde{\chi}=0})^{\alpha_n}, \\ r(z) \mid_{\tilde{\chi}>0} &\equiv z(x_v^*(z) \mid_{\tilde{\chi}>0})^{\alpha_v}(x_n^*(z) \mid_{\tilde{\chi}>0})^{\alpha_n}. \end{aligned}$$

Reported and actual non-verifiable expenses are then given by:

$$\hat{c}^{n*}(z) = \begin{cases} r^*(z)(1-\alpha_v) & \text{if } r^*(z) < T \\ \frac{\tau}{r^*(z)\tilde{\chi}} + w_n x_n^*(z) \mid_{\tilde{\chi}>0} & \text{if } r^*(z) \geq T \end{cases},$$

and

$$c^{n*}(z) = \begin{cases} w_n x_n^*(z) \mid_{\tilde{\chi}=0} & \text{if } r^*(z) < T \\ w_n x_n^*(z) \mid_{\tilde{\chi}>0} & \text{if } r^*(z) \geq T \end{cases}.$$

We use the following approach to estimate the level of cost overreporting. We compute optimal allocations under the parameters estimated in Section 6.7.1, barring input prices, which we set to unity. We assume that z is uniformly distributed over a grid and compute $\{r^*(z), \hat{c}^{n*}(z), c^{n*}(z)\}$ $\forall z$. Notably, this procedure endogenously generates a revenue distribution, which relaxes the uniformity assumption we made in our baseline model of cost overreporting. Figure 15 shows that the model generates more small firms than large ones, as in the data.

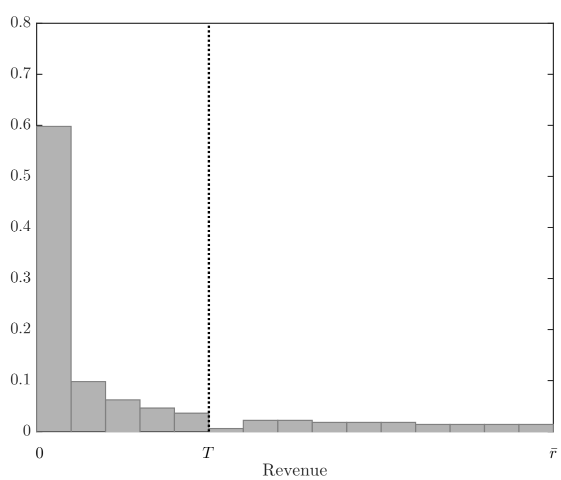
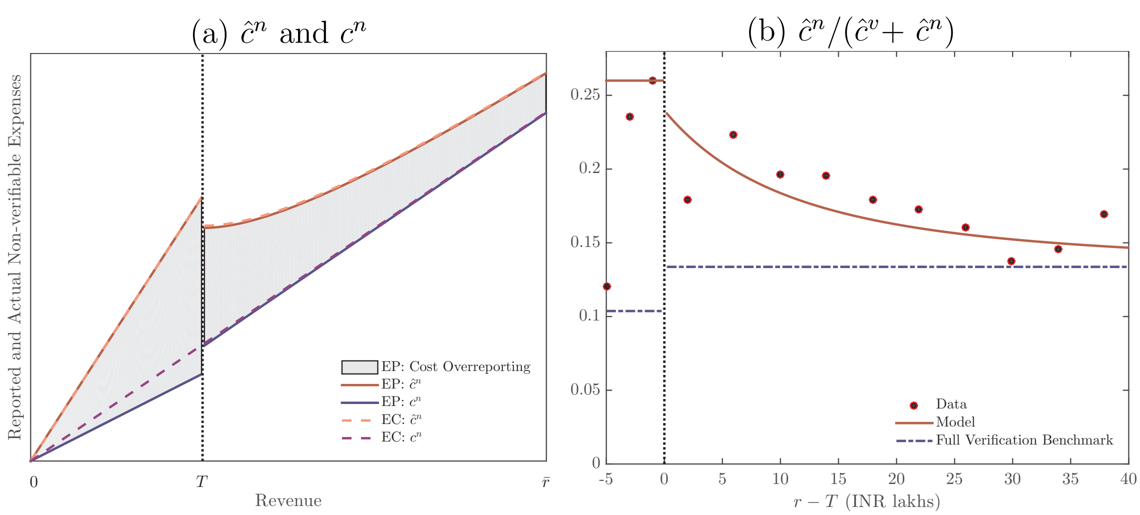


Figure 15: Simulated Distribution of Revenues

Panel (a) of Figure 16 plots simulated non-verifiable expenses and depicts the implied level of cost overreporting. Our estimate for the level of cost overreporting here is 9.2 percent, which is larger than the corresponding estimate in our baseline model. This difference is driven by real responses below the exemption threshold across the two settings. To see this, it is helpful to contrast optimal allocations here to those in the model outlined in Section 6.7.1 where the link between production and evasion decisions is severed. Note that reported non-verifiable costs below the exemption threshold are a constant fraction $(1 - \alpha_v)$ of revenues in both settings. However, when production decisions are endogenous, firms internalize that below the exemption threshold, the marginal benefit of hiring non-verifiable inputs (in terms of increased production) is taxed, unlike its marginal cost. This dichotomy reduces the demand for non-verifiable inputs and increases cost overreporting relative to the case in which revenues are exogenous. Another reason why our estimate for cost overreporting is larger in the model featuring endogenous production responses relates to the shape of the simulated distribution of firm revenues. Since cost overreporting by firms above the exemption threshold is similar across the two settings, the difference between the aggregate levels of cost overreporting would be smaller if the distribution of firm revenues was uniformly distributed instead of being positively skewed. Panel (b) of Figure 16 verifies that our model does a good job in matching reported non-verifiable expense ratios in the data.



Notes: Panel (a) contrasts non-verifiable costs in the model featuring endogenous production responses (EP) with those in the model where only the cost structure is endogenous (EC).

Figure 16: Estimating the Level of Cost Overreporting under Endogenous Production

6.8 Counterfactual Experiment: Partial Tax Deductibility of Costs

Note that firms are incentivized to overreport their costs precisely because they are tax deductible. It may then seem natural to lower the tax deductibility of expenses to stem evasion. This section argues that such a policy change may achieve the opposite of its intended effect.

Following Best et al. (2015) and Basri et al. (2021), we assume that proportion μ of costs are deductible from taxes. In particular, the firms' objective is given by

$$\mathbb{1}(r \geq c^v + \hat{c}^n)\tau(r - \mu[c^v + \hat{c}^n]) + \chi\mathbb{P}(r \geq T)(\hat{c}^n - c^n)^2/2.$$

Our benchmark estimation assumes $\mu = 1$, which captures a pure, nondistortionary profit tax akin to that implemented in India. A pure output tax can be captured by $\mu = 0$. When $0 < \mu < 1$, costs are only partially tax deductible. In this case, the firm will report the following non-verifiable expenses

$$\hat{c}^n = \begin{cases} \frac{r}{\mu} - c^v & \text{if } r < T \\ \frac{\tau\mu}{\chi\mathbb{P}(r\geq T)} + c^n & \text{if } r \geq T \end{cases}.$$

Using the Envelope Theorem, we can express the non-verifiable expense ratio as

$$\frac{\hat{c}^n}{\hat{c}^n + c^v} = \begin{cases} 1 - \frac{\mu c^v}{r} & \text{if } r < T \\ \frac{\mu\tau + \chi\mathbb{P}(r\geq T)c^v}{\mu\tau + \chi\mathbb{P}(r\geq T)(c^n + c^v)} & \text{if } r \geq T \end{cases}.$$

Note that the inclusion of partial tax deductibility of costs has important implications for the level of cost overreporting and its detection. To see this, observe that

$$\frac{\partial\{\frac{\hat{c}^n}{\hat{c}^n + c^v}\ |_{r<T}\}}{\partial\mu} < 0 \text{ and } \frac{\partial\{\frac{\hat{c}^n}{\hat{c}^n + c^v}\ |_{r\geq T}\}}{\partial\mu} > 0.$$

That is, relative to a nondistortionary profit tax, partial tax deductibility of costs exerts a force that induces firms to inflate their non-verifiable expenses by more (less) below (above) the tax exemption threshold. Thus, making costs less tax deductible increases the jump in the non-verifiable expense ratio at the tax exemption threshold, which makes cost overreporting easier to detect. This can be seen clearly in Figure 17, which depicts a series of counterfactuals where we progressively reduce the tax deductibility of costs, μ, keeping fixed the parameters estimated using our baseline model. Also note that, in general, reducing the tax deductibility of costs has an ambiguous effect on the level of cost overreporting. Nonetheless, our quantitative simulations suggest that its net impact on the level of cost overreporting is negative. That is, the rise in cost overreporting by firms below the exemption threshold outweighs the reduction in cost overreporting by firms above the threshold;

see Table 13.[9]

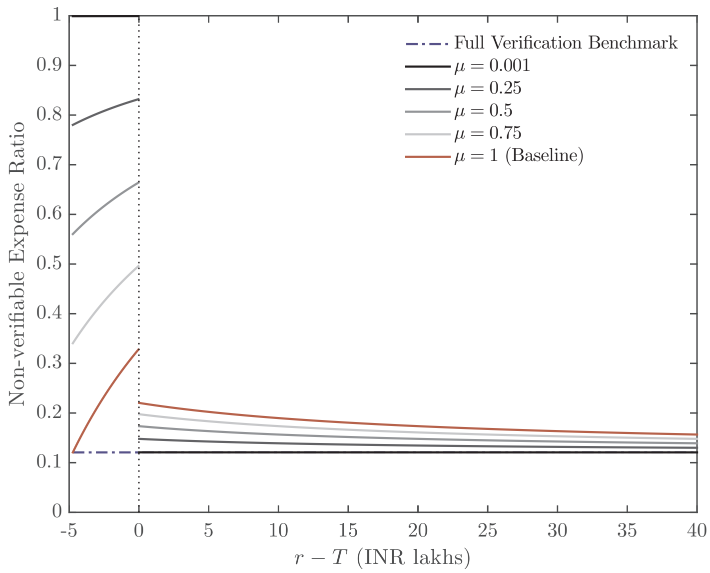


Figure 17: Effect of the Tax Deductibility of Costs on the Non-verifiable Expense Ratio

Table 13: Effect of the Tax Deductibility of Costs on the Level of Cost Overreporting

μ	0.001	0.25	0.5	0.75	1
$\frac{\Omega(c^{v*},c^{n*},\chi^*,\tau,T,\mu;P)}{(\bar{r}-c^{n*}-c^{v*})(c^{n*}+c^{v*})}$	122.2216	0.3990	0.1701	0.1042	0.0791

These results have important policy implications. Diamond and Mirrlees (1971) show that when tax enforcement is perfect, production efficiency can be achieved while taxing profits but not turnover. Best et al. (2015) argue that Diamond and Mirrlees' prescription is ill-suited to settings with limited tax capacity. They show that in developing economies with large informal sectors, moving toward a turnover tax (by reducing the tax deductibility of costs) may be desirable because it lowers evasion. This would also be true in our model if all firms faced a positive probability of being audited, an implicit assumption in Best et al. (2015). However, firms below the tax exemption threshold are never audited in our setting. We show that reducing the tax deductibility of costs increases evasion in this set of firms, which outweighs the reduction in evasion by firms above the tax threshold. Hence, in the presence of an enforcement notch, moving toward a turnover tax can increase aggregate evasion.

[9] In the augmented version of the model that features partial tax deductibility of costs, cost overreporting is computed using $\Omega(c^v, c^n, \chi, \tau, T, \mu; P) \equiv \int_{c^n+c^v}^{T} (\frac{r}{\mu} - c^v)dr + \int_T^{\bar{r}} \left\{ \frac{\tau\mu}{\chi\mathbb{P}(r \geq T)} + c^n \right\} dr - (\bar{r} - c^n - c^v)c^n$.

6.9 Elasticity of Corporate Tax Revenue: Tax Rates vs. Enforcement Intensity

In this section, we use our model to compare two natural approaches for increasing tax revenues: tax rates and enforcement intensity. On the one hand, if the elasticity of reported revenues and costs with respect to the tax rate is low, then simply increasing tax rates may be an effective way of increasing tax revenues. On the other hand, if the elasticity of the tax base with respect to the staff-to-taxpayer ratio is high, increasing enforcement intensity may be a more cost-effective approach.

In our model, corporate tax revenue (CTR) raised from firms generating revenue below $\bar{r}$ is given by

$$\text{CTR} = \int_0^{\bar{r}} \tau(r - \hat{c}^n - c^v)dr.$$

Since firms earning negative profits exit the market, we can decompose the above expression into

$$\text{CTR} = \int_{c^n+c^v}^{T} \tau(r - \hat{c}^n \mid_{r<T} - c^v)dr + \int_T^{\bar{r}} \tau(r - \hat{c}^n \mid_{r\geq T} - c^v)dr.$$

Using the Envelope Theorem, we can express tax revenues as

$$\text{CTR} = \tau(\bar{r} - T)\left[\frac{\bar{r}+T}{2} - \left(\frac{\tau}{\mathbb{P}(r \geq T)\chi} + c^n + c^v\right)\right].$$

The elasticity of CTR with respect to τ is given by

$$\mathcal{E}_{\text{CTR},\tau} = \frac{\partial \text{CTR}}{\partial \tau} \times \frac{\tau}{\text{CTR}} = \frac{\tau[(\bar{r}+T)/2 - 2\tau/(\mathbb{P}(r \geq T)\chi) - c^n - c^v]}{(\bar{r}+T)/2 - \tau/(\mathbb{P}(r \geq T)\chi) - c^n - c^v}.$$

Note that since $\tau \in [0,1]$, $\frac{(\bar{r}+T)/2 - 2\tau/(\mathbb{P}(r \geq T)\chi) - c^n - c^v}{(\bar{r}+T)/2 - \tau/(\mathbb{P}(r \geq T)\chi) - c^n - c^v} \leq 1$ and thus $\mathcal{E}_{\text{CTR},\tau} \leq 1$. Intuitively, the elasticity of CTR with respect to the tax rate is less than one because the tax base endogenously falls as the tax rate increases. It is worth noting that this behavioral response is purely due to evasion. In our calculations, we are implicitly assuming that the revenue distribution is uniform and inelastic to the tax rate, so in effect, we are switching off endogenous production responses. Thus, what we compute here serves as an upper bound for the elasticity of CTR with respect to the tax rate.

In our model, χ parametrizes not only the penalty for evasion, but also enforcement intensity. The elasticity of CTR with respect to χ is given by

$$\mathcal{E}_{\text{CTR},\chi} = \frac{\partial \text{CTR}}{\partial \chi} \times \frac{\chi}{\text{CTR}} = \frac{\tau}{[(\bar{r}+T)/2 - c^n - c^v]\chi\mathbb{P}(r \geq T) - \tau}.$$

Note that if $[(\bar{r}+T)/2 - c^n - c^v]\chi\mathbb{P}(r \geq T) > \tau$, then $\mathcal{E}_{\text{CTR},\chi} > 0$. Also note that $\mathcal{E}_{\text{CTR},\chi} > 1$

if $\chi < \frac{2\tau}{\mathbb{P}(r \geq T)[(\bar{r}+T)/2 - c^n - c^v]}$. Hence, if $\chi \in \left\{\frac{\tau}{\mathbb{P}(r \geq T)[(\bar{r}+T)/2 - c^n - c^v]}, \frac{2\tau}{\mathbb{P}(r \geq T)[(\bar{r}+T)/2 - c^n - c^v]}\right\}$, then

$$\mathcal{E}_{\mathrm{CTR},\chi} > 1 > \mathcal{E}_{\mathrm{CTR},\tau} > 0.$$

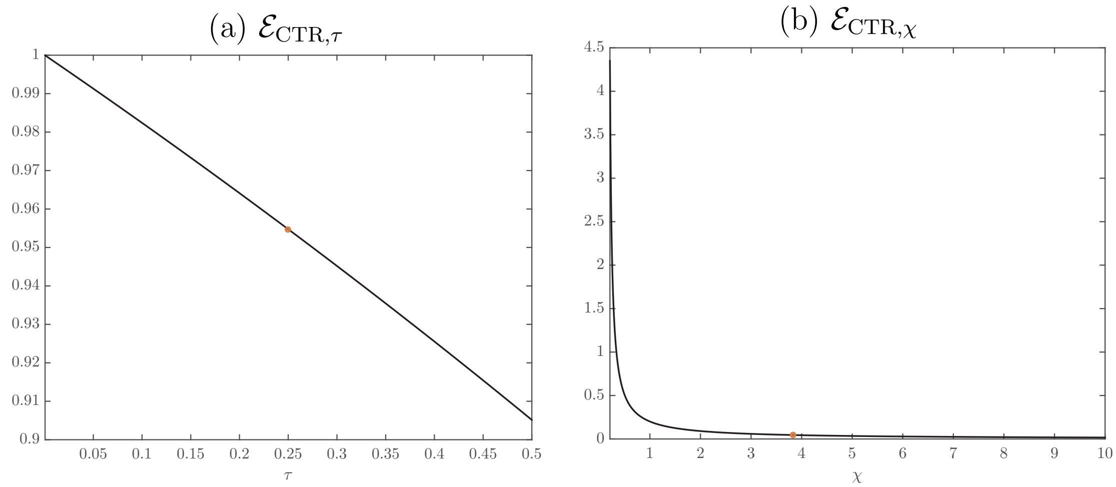
On the other hand, if χ is large enough, then $\mathcal{E}_{\mathrm{CTR},\chi} < \mathcal{E}_{\mathrm{CTR},\tau}$. To see this, note that $\mathcal{E}_{\mathrm{CTR},\chi}$ and $\mathcal{E}_{\mathrm{CTR},\tau}$ are continuous in χ, $\lim_{\chi\to\infty} \mathcal{E}_{\mathrm{CTR},\chi} = 0$ and $\lim_{\chi\to\infty} \mathcal{E}_{\mathrm{CTR},\tau} = \tau$. This also turns out to be the case in our quantitative analysis where we find that increasing the corporate tax rate by one percent has a larger impact on CTR than increasing the enforcement intensity by one percent (Figure 18). This is primarily because the prevailing level of enforcement intensity is estimated to be quite large to begin with, and thus marginal increases do not have a substantial impact on tax evasion. On the other hand, while an increase in the tax rate increases cost overreporting, the reduction in the tax base due to an increase in such evasion is not large enough to completely erode the increase in CTR stemming from an increase in the tax rate itself.

It is helpful to contrast our results to Basri et al. (2021), who show that tax revenues more than doubled when Indonesian tax authorities increased enforcement intensity by tripling staff-to-taxpayer ratios. Comparing this estimate with the elasticity of taxable income, they find that improved tax administration is equivalent to raising rates on affected firms by eight percentage points. As discussed above, the elasticity of taxable income depends on the tax rate and enforcement intensity, which vary across countries. Increasing enforcement intensity may be the preferred approach to raising tax revenues in countries where corporate tax rates are high and tax enforcement is weak, as in Indonesia. Alternatively, increasing tax rates may be more desirable in economies with lower tax rates and higher enforcement intensities.

7 Conclusion

Low corporate tax revenue is a significant challenge in many developing countries where evasion and informality limit tax capacity. This paper highlights the benefits and limitations of indirect tax administration in curbing corporate tax evasion by examining a recent regime change in India–the switch from a decentralized VAT to a unified GST. The regime change induced an increase in third-party reporting, which made the cost of materials more visible than other firm expenses to the authorities. It also created an enforcement notch at a turnover threshold beyond which firms were legally bound to share financial information with tax administrators routinely.

Using an administrative dataset of corporate tax returns of Indian firms over the period 2015-2020, we obtain three main results. First, firms reported higher revenues and costs after the implementation of GST. Second, much of the increase in reported revenue was



Notes: In panel (a), the red dot depicts the elasticity of CTR with respect to tax rates evaluated at 25 percent (i.e., the corporate tax rate in India during our sample period). In panel (b), the red dot depicts the elasticity of CTR with respect to enforcement intensity evaluated at the enforcement intensity estimated using GMM to match non-verifiable expense ratios in Indian data around the GST exemption threshold.

Figure 18: Elasticities of Corporate Tax Revenue

due to reduced income underreporting. Third, the regime change prompted labor-intensive firms to shift toward overreporting their wage bills that were relatively harder to verify. Our methodological innovation is to use variation in the relative visibility of specific inputs, in addition to notches in enforcement intensity, to detect the presence of cost overreporting.

The results of this paper have implications beyond India. More than 61 percent of the world's employed population works informally, and most of these workers reside in developing countries (ILO, 2018b). Less developed economies also tend to be more labor-intensive (Kaldor, 1957). Our results suggest that the presence of informal labor-intensive firms acts as a fly in the ointment of third-party reporting. Having a large pool of informal workers allows firms to inflate labor costs, which undercuts the rise in corporate taxable income that stems from the increased monitoring of final and intermediate goods.

References

Almunia, M. and D. Lopez-Rodriguez (2018): "Under the radar: The effects of monitoring firms on tax compliance," *American Economic Journal: Economic Policy*, 10, 1–38.

Autor, D. H. (2003): "Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing," *Journal of labor economics*, 21, 1–42.

Bachas, P. and M. Soto (2021): "Corporate Taxation under Weak Enforcement," *American Economic Journal: Economic Policy*, 13, 36–71.

Basri, M. C., M. Felix, R. Hanna, B. A. Olken, et al. (2021): "Tax Administration versus Tax Rates: Evidence from Corporate Taxation in Indonesia," *American Economic Review*, 111, 3827–3871.

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, 1311–1355.

Card, D., R. Chetty, M. S. Feldstein, and E. Saez (2010): "Expanding access to administrative data for research in the United States," *American Economic Association, Ten Years and Beyond: Economists Answer NSF's Call for Long-Term Research Agendas*.

Carrillo, P., D. Pomeranz, and M. Singhal (2017): "Dodging the taxman: Firm misreporting and limits to tax enforcement," *American Economic Journal: Applied Economics*, 9, 144–64.

Cattaneo, M. D., N. Idrobo, and R. Titiunik (2017): "A practical introduction to regression discontinuity designs," *Cambridge Elements: Quantitative and Computational Methods for Social Science-Cambridge University Press I*.

Chetty, R., J. N. Friedman, T. Olsen, and L. Pistaferri (2011): "Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records," *The quarterly journal of economics*, 126, 749–804.

Clotfelter, C. T. (1983): "Tax evasion and tax rates: An analysis of individual returns," *The review of economics and statistics*, 363–373.

Deloitte (2016): "India Fraud Survey: Edition II," https://www2.deloitte.com/in/en/pages/finance/articles/india-fraud-survey-edition-ii.html.

Diamond, P. A. and J. A. Mirrlees (1971): "Optimal taxation and public production I: Production efficiency," *The American economic review*, 61, 8–27.

Fisman, R. and S.-J. Wei (2004): "Tax rates and tax evasion: evidence from "missing imports" in China," *Journal of political Economy*, 112, 471–496.

Gordon, R. and W. Li (2009): "Tax structures in developing countries: Many puzzles and a possible explanation," *Journal of public Economics*, 93, 855–866.

Gutmann, P. M. (1977): "The subterranean economy," *Financial Analysts Journal*, 33, 26–27.

Hahn, J., P. Todd, and W. Van der Klaauw (2001): "Identification and estimation of treatment effects with a regression-discontinuity design," *Econometrica*, 69, 201–209.

Handley, K. and T. J. Moore (2017): "Tariffs, Transport Costs, and Tax Evasion," *Working Paper*.

ILO (2018a): "India wage report: Wage policies for decent work and inclusive growth," *India: International Labor Organization*.

——— (2018b): "Women and men in the informal economy: A statistical picture. Third edition," https://www.ilo.org/global/publications/books/WCMS_626831/lang--en/index.htm.

Imbens, G. W. and T. Lemieux (2008): "Regression discontinuity designs: A guide to practice," *Journal of econometrics*, 142, 615–635.

Johnson, S., D. Kaufmann, A. Shleifer, M. I. Goldman, and M. L. Weitzman (1997): "The unofficial economy in transition," *Brookings papers on economic activity*, 1997, 159–239.

Kaldor, N. (1957): "A model of economic growth," *The economic journal*, 67, 591–624.

Keen, M. and J. Mintz (2004): "The optimal threshold for a value-added tax," *Journal of Public Economics*, 88, 559–576.

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, 219–246.

Kleven, H. J. and M. Waseem (2013): "Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan," *The Quarterly Journal of Economics*, 128, 669–723.

Kopczuk, W. and J. Slemrod (2006): "Putting firms into optimal tax theory," *American Economic Review*, 96, 130–134.

Liu, L., B. Lockwood, M. Almunia, and E. H. Tam (2019): "VAT notches, voluntary registration, and bunching: Theory and UK evidence," *Review of Economics and Statistics*, 1–45.

McCrary, J. (2008): "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of econometrics*, 142, 698–714.

Pomeranz, D. (2015): "No taxation without information: Deterrence and self-enforcement in the value added tax," *American Economic Review*, 105, 2539–69.

Saez, E. (2010): "Do taxpayers bunch at kink points?" *American economic Journal: economic policy*, 2, 180–212.

Slemrod, J., B. Collins, J. L. Hoopes, D. Reck, and M. Sebastiani (2017): "Does credit-card information reporting improve small-business tax compliance?" *Journal of Public Economics*, 149, 1–19.