Tax Evasion and Productivity*

Hans Martinez†

Abstract

Value-added taxes (VAT) allow firms to deduct the taxes paid on purchases from the taxes collected on sales, creating an incentive for firms to reduce their tax bill by overreporting input purchases. The fundamental issue in detecting input overreporting stems from two unobservables, true inputs and productivity. Because firms differ in productivity, a firm reporting high input use, for a given level of output, may be either overreporting or simply less productive. To distinguish between overreporting and productivity, I use a structural production-function approach and a benchmark group, a subset of firms that report inputs correctly. Assuming the benchmark firms share the same technology as the others, I estimate the technology's parameters from them. With the estimates, I compare the inputs that potentially tax-evading firms report with the true inputs they would use. I apply the method to a Colombian manufacturing survey from 1981 to 1991. During this period, the Colombian sales tax worked as a VAT and the data include a subset of firms that, due to government and market scrutiny, reported inputs correctly.

My approach yields a simple and robust statistical test to detect input overreporting. Under the null hypothesis of correct reporting, the test is agnostic about the specification of both the probability of detection and the cost of overreporting. Among the 5 largest industries, the test rejects the null hypothesis in all 4 industries subject to the sales tax, and, as expected, it fails to reject in the industry exempt from the sales tax. Depending on the industry, firms overreport on average between 4.5 and 23.1 percent of their true inputs.

The magnitude of detected overreporting matters for methods of estimating production functions that exploit a flexible input to identify productivity (Levinsohn and Petrin 2003; Doraszelski and Jaumandreu 2013; Ackerberg et al. 2015; Gandhi et al. 2020). These methods typically rely on materials, the input that firms overreport. I show that failing to correct for overreporting yields larger output elasticities of materials, and more dispersed productivity distributions. Using the production-function parameters, I estimate a structural model of input overreporting. The model predicts that at the current rate, a 1 percent increase in the tax rate lowers net sales-tax revenue by about 1.7 percent of gross sales-tax revenue. Undetected overreporting costs the government 4.7 percent of the gross sales-tax revenue from potentially tax-evading firms.

JEL Classification: H26, D22, L60, O47
Keywords: Tax Evasion, Cost Overreporting, Production Function Estimation, Productivity

*This version: November 12, 2026; latest here. I thank Salvador Navarro, David Rivers, and Timothy Conley for helpful comments. All errors are my own.

†Department of Economics, Western University. Email: hmarti33@uwo.ca. Website: hansmartinez.com.

Introduction

Tax evasion through overreporting of input costs arises when firms inflate their reported purchases to claim additional deductions against their value-added tax (VAT) liabilities. Despite its relevance, we still do not know much about firms' cost overreporting. Tax authorities around the world document that this kind of evasion, often through fictitious invoices, is costly and widespread (OECD 2017). The losses amount to between 1.2 and 2.7 percent of total tax revenue in Chile, Colombia, Ecuador, Poland, and Slovakia, and about a ninth of the entire VAT compliance gap in Colombia and Slovakia (OECD 2017; Carrillo et al. 2022; CASE and IHS 2018; Longinotti 2022). Because some of these figures are tax administrations' own estimates, and all are conditional on detected evasion, true losses are plausibly larger.

Measuring firms' cost overreporting has been limited to studies with transaction-level tax records from Ecuador and Mexico, which show that firms acquire fake invoices to overreport costs (Carrillo et al. 2022; Zumaya et al. 2021). Tax records, however, are confidential, available for few countries, and accessible only to selected researchers. Without them, economists infer evasion from validation samples, comparing an indirect measure of true income between a truth-reporting group and a group that may misreport (Pissarides and Weber 1989; Gorodnichenko et al. 2009). This approach, however, was built for individuals. For firms, the natural indirect measure, output relative to inputs, also depends on an additional unobservable, productivity: a firm reporting high input use, for a given level of output, may be either overreporting or simply less productive.

I ask whether input overreporting can be detected and measured from standard production data, and how large its costs are. I study these questions with a Colombian manufacturing survey from 1981 to 1991. During this period, the Colombian sales tax worked as a VAT, and the data include a subset of firms that, due to government and market scrutiny, reported inputs correctly. I find that overreporting can be detected and measured from these data alone, that it is common and large, and that it is costly in two ways: for the government, which credits fictitious purchases, and for productivity measurement, in particular for proxy-variable methods that rely on the flexible input, the one firms overreport: materials.

To distinguish between overreporting and productivity, I use a structural production-function approach and a benchmark group of firms that report inputs correctly. In the absence of evasion, the first-order condition for materials shows that the firm's materials share of revenue equals the output elasticity of materials, a parameter of the technology common to its industry. Government and market scrutiny keep corporations from overreporting. Corporations are thus the benchmark that identifies the common technology in each industry and the distribution of measurement error in output. Among unincorporated firms, deviations from the common technology reveal overreporting up to measurement error in output. Deconvolution then separates the distribution of overreporting from that of the measurement error. Finally, employing an entropy-based method (Schennach 2014), I estimate a structural model in which firms choose inputs and overreporting to maximize expected

after-tax profits. In the model, firms choose how much to overreport by weighing the tax evaded against the probability of detection and a cost of evasion that varies with productivity.

Overreporting can be detected with a simple and robust statistical test. Under the null hypothesis of correct reporting, the test is agnostic about the specification of both the probability of detection and the cost of overreporting, and its size holds under any common production technology. A rejection is also unambiguous because no firm gains by underreporting deductible inputs. Among the 20 largest industries, the test detects overreporting in 9 of the 18 industries liable for the sales tax, which together produce about two fifths of manufacturing output, including 4 of the 5 largest. It detects no overreporting in either of the 2 food industries, which were exempt from the sales tax.

Overreporting is large and responds to the tax rate that rewards it. By deconvolution, I recover the distribution of overreporting in each industry. In the five largest industries where the test rejects at the 1 percent level and the deconvolution applies, unincorporated firms overreport on average between 4.5 and 23.1 percent of their true materials, depending on the industry. Overreporting also rose with the tax rate. Between 1983 and 1987, after Colombia raised the sales-tax rate for most manufacturing industries, overreporting by unincorporated firms in the affected industries rose significantly, by about 8.5 percentage points of true materials. The response implies a midpoint elasticity of overreporting with respect to the tax rate of 5.9.

Overreporting is costly for the government, and the cost grows with the tax rate. In the structural model, undetected overreporting costs 4.7 percent of the gross sales-tax revenue from unincorporated firms in the industries where overreporting is detected. A 1 percent increase in the rate on purchases lowers net sales-tax revenue by about 1.7 percent of gross sales-tax revenue. Because net sales-tax revenue is negative at the observed rate (the sales-tax refunds on claimed deductions exceed the gross sales-tax revenue from sales), a cut of about 17 percent in the rate on purchases would make it break even, and a cut of about 28 percent or more would make it significantly positive, at the 5 percent level.

Overreporting is also costly for productivity measurement. Proxy-variable methods that rely on a flexible input typically define it as materials (Levinsohn and Petrin 2003; Doraszelski and Jaumandreu 2013; Ackerberg et al. 2015; Gandhi et al. 2020). But this is the input firms overreport. In the five industries of the deconvolution, the uncorrected estimates show output elasticities of materials that are consistently larger, and productivity distributions that are more dispersed and less persistent, than the corrected ones. In the median of the five industries, the ratio of the 90th to the 10th percentile is 66 percent larger and the half-life of productivity shocks is 63 percent shorter.

Relation to the literature

These results relate to three strands of public economics and to one of productivity measurement. First, I provide measures of firms' cost overreporting without relying on confidential tax records.

Most evidence on firm evasion focuses on underreported revenue (Slemrod 2019), yet the cost side matters. Firms offset enforcement on revenue by raising reported costs (Carrillo et al. 2017; Asatryan and Peichl 2017), and paper trails make transactions between firms harder to misreport than sales to final consumers (Pomeranz 2015); therefore, firms that sell to other firms have little room to misreport on the sales side and more on the cost side. Survey-based measures of firm evasion rely on firms' self-reported share of sales declared to the tax authority (Dabla-Norris et al. 2019), whose truthfulness cannot be verified. My approach needs only reported inputs and output, and is thus equally applicable to standard firm-level surveys, administrative records, and confidential tax records.

Second, I extend the validation-sample approach from individuals to firms. In its classic application, the self-employed report less income relative to their food consumption than employees, whose income is reported by their employers (Pissarides and Weber 1989), and the same comparison has measured how evasion responds to a tax reform (Gorodnichenko et al. 2009). I use the production function to separate overreporting from productivity, with corporations as the truth-reporting group. Unlike the comparisons for individuals, my approach does not require the two groups to be alike in size or productivity: corporations are much larger than unincorporated firms, and the production function accounts for the differences in productivity behind this gap.

Third, the paper shows how tax-motivated overreporting of inputs biases production-function estimates. Proxy-variable methods rely on a flexible input, typically materials, either inverting its demand (Levinsohn and Petrin 2003; Ackerberg et al. 2015) or using its first-order condition (Gandhi et al. 2020). These methods treat reported inputs as true. Work on mismeasured inputs has focused on classical measurement error (mean zero and independent of the true input), mostly in capital (Collard-Wexler and De Loecker 2020). Overreporting is neither: firms choose it based on their own productivity. Furthermore, it affects exactly the input on which these methods rely. Correcting for it matters for any analysis built on measured productivity differences across firms.

Finally, I estimate how a particular form of evasion by firms responds to the tax rate, a response that the literature on taxes and revenue has mostly studied through individuals' reported income. Traditionally, the elasticity of personal taxable income has combined all the different responses to the rate, evasion included, into a single statistic (Feldstein 1999; Slemrod and Kopczuk 2002; Saez et al. 2012). However, when part of the cost of evasion is a transfer rather than a resource cost, this elasticity alone no longer measures the efficiency cost of taxation (Chetty 2009). Measuring that cost requires knowing how evasion itself responds to the rate, separately from real activity. Direct evidence on this response is scarcer, and comes from the gap between households' consumption and reported income (Gorodnichenko et al. 2009) and, for firms, from bunching at kinks and notches in the tax schedule (Best et al. 2015; Almunia and Lopez-Rodriguez 2018). I provide evidence for firms' overreporting of inputs under a sales tax that credits purchases, from both a statutory rate change and a structural model that separates overreporting from productivity. The structural model

also speaks to the theory of optimal commodity taxation under evasion by estimating how claimed deductions respond to the rate (Cremer and Gahvari 1993).

The rest of the paper follows the argument. Section 1 describes the Colombian sales tax, which credited purchases, and explains why Colombian corporations can serve as truth-tellers. Section 2 shows how the tax rate generates incentives for a firm to overreport materials, and how its optimal choice depends on the probability of detection and the cost of evasion. Section 3 shows how overreporting can be detected and measured from production data alone. Then, Section 4 demonstrates how failing to correct for overreporting biases production-function and productivity estimates. Section 5 leverages the 1983 reform to document that overreporting responds to the tax rate. Section 6 estimates the firm's evasion decision and the cost of undetected overreporting, and shows that a 1 percent increase in the rate on purchases lowers net sales-tax revenue by about 1.7 percent of gross sales-tax revenue. Section 7, as tradition demands, concludes.

1 Setting and Data

Colombian manufacturing in 1981–1991 offers a unique setting to study tax evasion by overreporting deductible inputs. First, a sales tax that worked as a value-added tax and rewarded overreporting materials: manufacturers could reduce the tax they owed on their sales by claiming fictitious deductions (Section 1.2). Second, a truth-reporting group of firms: corporations, which government and market scrutiny kept from evading (Section 1.3). Third, variation in the tax rate: the 1983 reform raised the sales-tax rate for most industries but left food products exempt (Section 1.4). Finally, firm-level records of output and inputs, together with legal form and taxes paid (Section 1.1). Because the approach uses the production function, it applies to any data with firms' output and inputs, such as surveys, administrative data, or tax records.

1.1 Data

I use the Colombian Annual Survey of Manufacturing (EAM), a panel of manufacturing plants with more than 10 employees from 1981 to 1991.[1] The EAM records, for each plant, the output and inputs that a production function needs, the plant's juridical organization, and the sales tax it charged on its sales and paid on its purchases.[2] The panel is well known: it has been used in seminal work on export dynamics, reallocation, and productivity, and to estimate production functions (Roberts and Tybout 1997; Das et al. 2007; Eslava et al. 2004; Gandhi et al. 2020). Similar to these studies,

[1]The survey's unit of observation is the plant; I refer to plants and firms interchangeably.

[2]The year in the panel is the activity (reference) year. DANE runs the EAM each year with the economic year immediately preceding collection as the reference period (Departamento Administrativo Nacional de Estadística (DANE) 1995, 40): the survey for year x is collected during $x + 1$ and labelled x (e.g. the 1992–1994 surveys were collected in 1993–1995, and the 1992 file values assets at December 1992; Departamento Administrativo Nacional de Estadística (DANE) (2018)).

Table 1: Summary statistics, manufacturing firms in Colombia (1981–1991).

Variable	Missing (%)	Mean	SD	Q1	Median	Q3
Output (log)	0.00	10.53	1.78	9.18	10.28	11.69
Capital (log)	0.00	9.03	2.02	7.69	8.94	10.34
Labour (log)	0.00	3.84	1.26	2.89	3.68	4.64
Raw materials, as reported (log)	0.57	9.61	2.05	8.24	9.46	10.96
Sales-tax rate on sales	0.06	0.066	0.049	0.007	0.070	0.100
Sales-tax rate on purchases	0.64	0.074	0.059	0.027	0.077	0.105

Firm-years with finite output, capital, labour and materials; corporations, LLCs, partnerships and proprietorships. Rate on sales: sales tax paid on sales over sales. Rate on purchases: sales tax paid on purchases over raw materials. Values above 1 are treated as data errors and counted as missing. Output, capital and raw materials are in logs of 1981 pesos; labour is in logs of employee-years.

I observe inputs as firms reported them; in contrast, I ask how the tax incentives of the period, including the 1983 reform, shaped those reports. The working sample has 40,508 firm-years from 5,915 plants in 28 industries, covering the four juridical organizations described in Section 1.3 (Table 1).

1.2 The sales tax: a credit on purchases

The Colombian sales tax worked as a value-added tax (VAT): manufacturers credited the tax paid on their purchases against the tax charged on their sales. The tax originally targeted the manufacturing sector, on finished goods and imports. From 1966, manufacturers could claim credit on the taxes paid on inputs *incorporated* into the product, and from 1974 on any purchase made by the firm, except the acquisition of capital goods; the credits worked through a system of refunds (Perry and Orozco de Triana 1990, 180).

The credit rewards overreporting materials. An artificial increase in the reported value of materials raises the credit, and so lowers the firm's tax liability, by the tax rate paid on purchases, peso per peso. Capital goods could not be credited and labour was not a taxed purchase, so I focus on raw materials consumed.[3]

Weak enforcement invited firms to collect this reward. Tax evasion flourished in Colombia during the sample period (Sánchez Torres and Gutiérrez Sourdis 1994), while the tax authority audited inefficiently, strained under its workload, and contended with legal loopholes (Perry and Orozco de Triana 1990).

[3] I exclude the other intermediates. Energy, fuels, and repair and maintenance are small, with medians of 1.0 percent, 0.3 percent, and 0.3 percent of gross output, and fuels and repair and maintenance are reported as zero in 8 and 15 percent of firm-years. Services, a median of 8.5 percent of gross output, entered the tax only gradually: the 1974 reform extended it to some services, and the 1983 reform to others, such as hotels, computing, maintenance, and rentals (McLure 1989, 65).

Food products (industries 311 and 312) offered almost no reward to overreporting, because they were exempt from the sales tax from its introduction (Perry and Orozco de Triana 1990, 181). Exempt firms charged no tax on their sales and, because the credits worked through a system of refunds (Perry and Orozco de Triana 1990, 180), could claim back the tax they paid on purchases, but they paid almost none: the median rate on purchases was 0.2 percent in industry 311 and 0.6 percent in industry 312, against 9.2 percent in the liable industries. A fictitious peso of materials therefore saved these firms almost nothing. Section 5 compares them with the liable industries.

1.3 Juridical organizations: corporations as truth reporters

I treat corporations as truthful reporters of materials because two kinds of scrutiny restrained them from overreporting: government and market. First, the Superintendent of Corporations closely supervised corporations and required each to have an auditor, according to the OAS/IDB Joint Tax Program's *Fiscal Survey of Colombia* (1965, 27–30); LLCs and proprietorships faced only the tax authority, whose audits were weak. Second, tradable shares subject corporations to market discipline, and the market penalizes aggressive tax behaviour even when it is legal. In the United States, stock prices fall, on average, when news of a corporation's involvement in tax shelters becomes public, although sheltering is avoidance rather than evasion (Hanlon and Slemrod 2009; Slemrod 2019). If the market penalizes legal avoidance, it should penalize illegal overreporting at least as much. Evidence from Ecuador agrees: the tax evaded by claiming fictitious purchases amounted to 1.7 percent for corporations and 11.5 percent for sole proprietorships (Carrillo et al. 2022). If corporations nonetheless overreport, my estimates understate the overreporting of unincorporated firms; they are a lower bound.

The data distinguish several juridical organizations. Four of them account for 98 percent of firm-years and 95 percent of real gross output[4]: corporations, limited liability companies (LLCs), partnerships, and proprietorships (Table 2). I define them following the *Fiscal Survey of Colombia* (1965, 27–30) and DANE's documentation of the EAM (2018, 12). Corporations (*sociedades anónimas*) are associations of capital, the counterpart of the US corporation: at least five shareholders, tradable shares of equal value, and liability limited to the capital contributed (1965; 2018). An LLC (*sociedad de responsabilidad limitada*) has two or more partners, not exceeding 20 (1965) or, by the 1990s, 25 (2018), whose stakes (*cuotas*) cannot be traded and whose liability is limited to the capital contributed; it is an association of persons, not of capital, and is not subject to the Superintendent of Corporations unless more than one-third of it is owned by a corporation (1965). Partnerships are associations of two or more persons in which the partners, or at least the managing partners, are jointly, severally, and unlimitedly liable for the partnership's operations (2018). They are a small

[4]The other organizations in the data are excluded: stock partnerships (*sociedades en comandita por acciones*), whose shares are tradable and which are taxed as corporations (Joint Tax Program of the Organization of American States and the Inter-American Development Bank 1965), and non-profit or state organizations, such as cooperatives, community enterprises, and state industrial enterprises, whose proceeds come from taxes, fees, or special contributions.

Table 2: Juridical organization, manufacturing firms in Colombia (1981–1991).

J. Org.	Firm-years	Plants	% of firm-years
Proprietorship	5,343	1,226	13.2
LLC	25,456	3,927	62.8
Partnership	1,425	319	3.5
Corporation	8,284	1,003	20.5

A plant that changes juridical organization is counted once in each.

Table 3: Income tax, liability, capital, and owners by juridical organization (1974–1983).

Organization	Income tax (company)	Income tax (owners)	Liability	Capital	Owners
Corporation	40%	On dividends received	Limited to capital contributed	Tradable shares	$N \geq 5$
LLC	20%	On all profits	Limited to capital contributed	Non-tradable stakes (*cuotas*)	$2 \leq N \leq 20$ (25)
Partnership	20%	On all profits	Unlimited[a]	Not a capital association	$N \geq 2$
Proprietorship	–	Individual schedule	Unlimited	Owner's own assets	$N = 1$

Rules in force from the 1974 reform until Ley 9 de 1983, which cut the LLC rate to 18%. Company: entity-level income-tax rate. Owners: individual income tax; proprietors paid only the graduated individual schedule, with a top rate of 56%. The limit on LLC partners is 20 in the 1965 source and 25 in the 1992–1994 survey documentation. Sources: tax rates, McLure (1989), p. 67, and Perry and Cárdenas (1986), vol. 1, pp. 23 and 36; liability, capital and owners, Fiscal Survey of Colombia (1965), pp. 27–30, and DANE (2018), p. 12.

[a] All partners in general partnerships; the managing partners in ordinary limited partnerships.

group (3.5 percent of firm-years), so I pool them with the other unincorporated firms. Proprietorships are individuals who run the business in their own name, without a separate legal personality (2018).[5]

Because of the income tax, incentives to overreport varied within industries by juridical organization (Table 3). Firms could reduce their taxable profits and, with them, the income tax by deducting the cost of fictitious materials. Between 1974 and 1983, partnerships and LLCs paid a rate of 20 percent, and their owners paid individual income tax again on their share of the profits, whether distributed or not (McLure 1989, 67; Perry and Cárdenas 1986, 1:23). Proprietors paid only the individual income tax, on a graduated schedule with a top rate of 56 percent (Perry and Cárdenas 1986, 1:36). Corporations paid a rate of 40 percent between 1974 and 1986, and their shareholders paid individual income tax again only on the dividends they received (McLure 1989, 67; Perry and Cárdenas 1986, 1:23). Even though corporations had incentives to overreport too, the government and market scrutiny described above restrained them.

[5]The *empresa unipersonal*, a single-owner legal entity, was introduced only in 1995 (Congreso de la República de Colombia 1995, art. 71).

1.4 The 1983 reform

The 1983 reform altered the reward to overreporting across and within industries. Across industries, the reform raised the sales-tax rate in most liable industries and left exempt industries untouched. Within industries, it cut the income-tax rates by different amounts for each juridical organization: most for proprietorships, less for LLCs, and not at all for corporations. For unincorporated firms in liable industries, the two changes pulled in opposite directions: the higher sales tax raised the reward to overreporting, and the lower income tax reduced it. I exploit this variation in Section 5 to test whether overreporting responded to its reward.

On the sales tax, the reform combined the items previously taxed at 6 and 15 percent into a single 10 percent bracket (Perry and Orozco de Triana 1990, 182), while the exempt industries remained exempt. The typical taxed firm was in the 6 percent bracket, so the reform raised its rate: between 1983 and 1985, the median rate on sales rose from 6 to 10 percent, and the median rate on purchases, the rate at which each fictitious peso of materials was credited, from 6.6 to 9.9 percent (Table 4).

The new rate, set by Decreto 3541 of 29 December 1983 under the extraordinary powers of Ley 9 de 1983, applied from 1 April 1984 (Presidencia de la República de Colombia 1983; Corte Constitucional de Colombia 2016; Perry and Orozco de Triana 1990), so 1983 is the last full pre-reform year and 1984 a partial one. The observed rates match this timing (Table 4). The median rate paid on sales is 6 percent in 1981–1983, 10 percent from 1985, and 12 percent in 1991, after Ley 49 de 1990 raised the general rate to 12 percent (Congreso de la República de Colombia 1990). In 1984 it is 9 percent, exactly what three months at 6 percent and nine months at 10 percent give ($0.25 \times 6\% + 0.75 \times 10\% = 9\%$).

On the income tax, Ley 9 de 1983 cut the individual schedule, faced by proprietorships, by 4.6 percentage points on average, with the top rate falling from 56 to 49 percent, and the LLC rate from 20 to 18 percent (Perry and Cárdenas 1986, 1:36).

2 A Model of Tax Evasion through Input Overreporting

Materials are the input that is both flexible and deductible. Because materials are flexible, the firm's first-order condition (FOC) equates the share of materials over sales to the output elasticity of materials and output measurement error. If the elasticity were known, the gap between the reported share of materials over sales and the elasticity would measure overreporting, up to the measurement error. This is the materials condition.

Because materials are deductible, every overreported peso lowers a firm's tax bill by the sales-tax rate on purchases. But overreporting is costly. A firm that overreports weighs this benefit against the probability of detection and the cost of overreporting, which include accounting and opportunity costs. This trade-off is the evasion condition.

Table 4: Observed sales-tax rates by survey year, on sales (τ_S) and on purchases (τ_P), against the statutory general rate in force. Rows in bold are the years a reform took effect: 1984 (Decreto 3541 de 1983, from 1 April) and 1991 (Ley 49 de 1990, from 1 January). Firms with a rate strictly between 0 and 50 percent.

Year	Statutory general rate	N	Median τ_S	Mean τ_S	P75 τ_S	Median τ_P
1981	15% basic, 6% preferential	4,399	6.0%	7.8%	9.8%	6.7%
1982	15% basic, 6% preferential	4,098	6.0%	7.7%	9.5%	6.1%
1983	15% basic, 6% preferential	3,273	6.0%	7.7%	9.5%	6.6%
1984	**10% from 1 April**	**3,005**	**9.0%**	**8.2%**	**10.0%**	**8.7%**
1985	10%	2,781	10.0%	9.0%	10.0%	9.9%
1986	10%	2,623	10.0%	9.0%	10.0%	9.8%
1987	10%	2,485	10.0%	8.9%	10.0%	9.9%
1988	10%	2,360	10.0%	8.9%	10.0%	9.8%
1989	10%	2,265	10.0%	8.9%	10.0%	9.7%
1990	10%	2,176	10.0%	8.8%	10.0%	9.7%
1991	**12% from 1 January**	**2,063**	**12.0%**	**10.3%**	**12.0%**	**11.5%**

τ_S: sales tax charged on sales over sales. τ_P: sales tax paid on purchases over raw materials.

This section derives both conditions, which the rest of the paper builds on. The materials condition (Equation 4) underlies the test, the deconvolution, and the production-function estimates (Section 3 to Section 4). The evasion condition (Equation 5) underlies the analysis of the 1983 reform and the counterfactual (Section 5, Section 6).

2.1 Setup

Price-taking firms maximize expected after-tax profits. They share a common technology but differ in productivity. A firm i in year t produces output Y_{it} with capital K_{it}, labour L_{it}, and materials M_{it},

$$Y_{it} = K_{it}^{\alpha_K} L_{it}^{\alpha_L} M_{it}^{\beta} \exp(\omega_{it}). \tag{1}$$

As in the production-function literature, materials are the flexible input, which the firm adjusts freely each year, while capital and labour are predetermined or dynamic. Productivity, ω_{it}, is known to the firm when it chooses its inputs, so inputs respond to it, which creates the simultaneity problem in estimating production functions (Griliches and Mairesse 1995).

The firm sells its output at price P_t and owes sales tax at rate τ_S on its sales. It buys materials at price ρ_t and pays a sales tax at rate τ_P, but the firm can deduct the tax paid on them from the tax it owes on sales.

2.2 Expected profit

The firm might choose to increase its reported material consumption by e_{it} units to claim additional credits on materials. Firms report $M^*_{it} = M_{it} + e_{it}$. Capital goods and labour expenses cannot be deducted from sales-tax liabilities.

Overreporting is detected with a probability $q(e_{it})$, which increases with the amount overreported. A firm that does not overreport has nothing to hide, and thus its probability of detection is zero, $q(0) = 0$. A firm that is caught repays the credit it claimed on e_{it} plus a fine, a multiple ϕ of that credit.

Evaders pay a cost of evasion $\kappa_{it}(e_{it}, \omega_{it})$. The cost is increasing in the overreported amount, is zero for a firm that does not overreport, $\kappa_{it}(0, \omega_{it}) = 0$, and is paid whether or not the firm is caught. $\kappa_{it}(e_{it}, \omega_{it})$ represents direct costs of overreporting, such as the price of fake invoices, and opportunity costs: any resource reallocated from a productive activity to tax evasion, such as the manager's time.

The cost of evasion depends on productivity, which lets the model match the observed pattern in recent evidence. In Ecuador, the share of firms' deductions backed by receipts from fake firms rises with firm revenue over most of the size distribution but drops sharply at the very top (Carrillo et al. 2022, 6). The probability of detection could not explain this pattern. If allowed to vary with size, it would imply that larger firms overreport more because their probability of detection is lower. This does not make economic sense. Larger firms are more visible to the tax authority, so their probability of detection is, if anything, at least as high as that of smaller firms. Only the drop at the top could come from a higher probability of detection.

In the model, the cost of evasion produces the whole pattern because it is a better candidate to explain the pattern. More productive firms are larger and have more resources to allocate to evasion, and thus their cost of evasion is lower. For the most productive firms, however, the cost rises relative to other tools: they can turn to legal avoidance instead of overreporting, which is illegal (Carrillo et al. 2022, 6). The largest firms also have more to lose if caught, and evading requires more employees to collude (Almunia and Lopez-Rodriguez 2018). A cost of evasion that falls with productivity over most of the distribution and rises at the top captures both forces. Therefore, I assume a cost of evasion convex in productivity.

Since underreporting inputs would increase the tax liability, no firm has an incentive to do so. Hence $e_{it} \geq 0$.

The optimization problem of the firm is then,

$$\begin{aligned} \max_{M,e,K,L} \ \mathbb{E}[\pi_{it}] = (1-\tau_S) P_t K_{it}^{\alpha_K} L_{it}^{\alpha_L} M_{it}^{\beta} \exp(\omega_{it}) - (1-\tau_P)\rho_t M_{it} - r_t K_{it} - w_t L_{it} \\ + \tau_P \rho_t e_{it}[1 - (1+\phi) q(e_{it})] - \kappa_{it}(e_{it}, \omega_{it}). \end{aligned} \tag{2}$$

The first term is revenue net of the tax on sales; the second, the cost of true materials net of the credit on them; the third and fourth, the rental cost of capital at rate r_t and the wage bill at wage w_t. The fifth term is the credit on overreported materials net of the expected repayment and fine, and the last is the cost of evasion.

2.3 First-order conditions

2.3.1 The materials condition

Taking the FOC of Equation 2 with respect to M_{it}, we have,

$$(1 - \tau_S)P_t \, \beta K_{it}^{\alpha_K} L_{it}^{\alpha_L} M_{it}^{\beta-1} \exp(\omega_{it}) = (1 - \tau_P)\rho_t. \tag{3}$$

Equation 3 looks familiar except for the tax wedge: the left-hand side is the marginal revenue product of materials and the right-hand side is the price of materials, both net of taxes.

The taxes matter when $\tau_P \neq \tau_S$. By changing either rate, the government changes the price of materials relative to output, and with it the firm's choice of materials. If $(1 - \tau_P)/(1 - \tau_S) > 1$, materials become more expensive relative to output than without taxes, and the firm uses less of them; if the ratio is below one, as when the rate on purchases exceeds the rate on sales, materials become cheaper and the firm uses more of them.

The econometrician observes output, including measurement error ε_{it}, as a function of inputs. Multiplying Equation 3 by M_{it}, dividing by the value of observed output, $P_t Y_{it}$, and taking logs gives the log share of materials in output, net of sales taxes:

$$\ln\left(\frac{(1 - \tau_P)\rho_t M_{it}}{(1 - \tau_S)P_t Y_{it}}\right) = \ln\beta - \varepsilon_{it}. \tag{4}$$

With a single rate, the tax terms cancel and the gross share, $\rho_t M_{it}/P_t Y_{it}$, satisfies the same equation as the no-tax optimization condition. The two shares differ little in the data: the log ratio $\ln\left((1 - \tau_S)/(1 - \tau_P)\right)$ has a standard deviation of 0.05 across firm-years, against 0.34 for the log gross share.

Productivity does not appear in the share. If we knew $\ln\beta$, a deviation of the reported share from it could separate overreporting from productivity. Furthermore, if we knew the moments or the distribution of ε_{it}, then we could estimate moments and properties of the distribution of overreporting. Section 3 builds on Equation 4 to identify β and to test for overreporting. Section 3.6 employs deconvolution techniques to learn about the distribution of overreporting by removing ε_{it}.

2.3.2 The evasion condition

The FOC with respect to e_{it} shows that,

$$\tau_P \rho_t [1 - (1 + \phi)(q(e_{it}) + q'(e_{it}) e_{it})] = \frac{\partial \kappa_{it}(e_{it}, \omega_{it})}{\partial e_{it}}. \tag{5}$$

To better understand the incentives to overreport, Equation 5 can be rewritten in terms of the elasticity of the probability of detection with respect to overreporting, $\xi_{q,e} = q'(e_{it}) e_{it} / q(e_{it})$:

$$\tau_P \rho_t [1 - (1 + \phi)(1 + \xi_{q,e})\, q(e_{it})] = \frac{\partial \kappa_{it}(e_{it}, \omega_{it})}{\partial e_{it}}. \tag{6}$$

The left-hand side is the marginal benefit of overreporting, and the right-hand side is its marginal cost. The marginal benefit decreases with the fine, ϕ, the probability of detection, $q(e_{it})$, and the elasticity of the probability of detection, $\xi_{q,e}$.

A higher tax rate on purchases raises overreporting. A higher rate on purchases raises the credit on every fictitious peso, whether or not the firm is caught, so the firm overreports more. By the implicit function theorem, $\partial e^*_{it} / \partial \tau_P > 0$ for any probability of detection and cost of evasion that satisfy the second-order condition (Section B). This sign is the prediction that Section 5 tests with the 1983 reform, and $e^*_{it}(\tau_P)$ is the object that Section 6 estimates.

Some firms choose not to overreport at all. For them, the marginal cost of evasion exceeds the marginal benefit at any positive level of overreporting, so Equation 5 does not hold with equality; instead, at $e_{it} = 0$,

$$\tau_P \rho_t \leq \frac{\partial \kappa_{it}(0, \omega_{it})}{\partial e_{it}}. \tag{7}$$

Firms that pay no tax on their purchases ($\tau_P = 0$) are always in this group.

The model allows the two rates to differ, as they do in the data. Firms sell and buy goods taxed at different rates, and the two rates differ by more than one percentage point in 56 percent of firm-years. But the two rates might be equal, $\tau_S = \tau_P = \tau$ (Section B.1). In that case, the FOC of materials would not have a tax wedge and the single rate would replace τ_P in Equation 5.

2.3.3 Corporations

Corporations are corner firms by assumption. They share the technology of the other firms and may differ in productivity, but audited statements and scrutiny by government and markets keep them from overreporting (Section 1.3). I assume that Equation 7 always holds for them. For other firms, a corner solution follows from κ_{it}, ω_{it}, and τ_P. For corporations, $e_{it} = 0$ is a constraint, and I use

them only as the truth-reporting benchmark. Corporations may still overreport a little: in Ecuador, the tax evaded through fake receipts amounted to 1.7 percent for corporations, against 11.5 percent for sole proprietorships (Carrillo et al. 2022, 6). If Colombian corporations also overreport a little, my estimates understate the overreporting of unincorporated firms and are a lower bound.

2.4 Relation to other work

The model departs from standard production-function methods that depend on the flexible input (Levinsohn and Petrin 2003; Doraszelski and Jaumandreu 2013; Ackerberg et al. 2015; Gandhi et al. 2020) in two ways. First, true materials are unobserved. Second, firms pay sales taxes. This matters because, if the rate on sales differs from the rate on purchases, the taxes change the price of materials relative to output. Hence, the government can indirectly influence how much material the firm uses and thus real output (Section 2.3.1).

Regarding the firm's evasion decision in the tradition of Allingham and Sandmo (1972), Almunia and Lopez-Rodriguez (2018) sit close to the model presented in this section. I depart from them in three ways. First, the firm overreports a deductible input under a sales tax that credits purchases, rather than underreporting revenue under a tax on profits. Second, the probability of detection does not jump when the firm's reported revenue crosses a threshold. Colombia had no discontinuity in the probability of detection we could exploit. Third, the cost of evasion varies with the firm's productivity, as explained before in this section (Section 2.2).

I focus on overreported inputs rather than underreported sales because underreporting sales is harder for firms that sell to other firms, as Colombian manufacturers mostly did. First, sales between firms leave a thicker paper trail than sales to consumers. Consumers do not always ask for invoices, whereas firms require invoices so they can deduct their purchases (Slemrod 2019, 925–27). Second, seller and buyer firms face opposing incentives, which renders coordination between them unlikely: the seller gains from underreporting the sale, but the buyer gains from overreporting the purchase. In a 1984 survey, Colombian manufacturers sold 13 percent of their production directly to the public, about 20 percent to retailers, and most of the rest to wholesalers (Perry and Orozco de Triana 1990, 183–84, table 16-3). Focusing on Colombian manufacturers between 1981 and 1991 therefore lets me study input overreporting largely separately from sales underreporting.

3 Identifying Tax Evasion

In this section, I show how to test for overreporting and recover its distribution. Equation 4 shows that the share of materials equals the output elasticity of materials and the output error. Importantly, under Cobb-Douglas, the output elasticity of materials is constant and does not depend on productivity. In general, the output elasticity of the flexible input of a Hicks-neutral production function does not depend on productivity conditional on inputs. This is what allows me to distinguish between

overreporting and productivity.

Since corporations report materials truthfully and share a common technology with unincorporated firms, the output elasticity of materials can be recovered from corporations, even though corporations and unincorporated firms are heterogeneous in productivity. The difference between the materials share of gross output of unincorporated firms and the output elasticity of materials then recovers overreporting up to the output measurement error.

The difference forms an observed residual whose expected value can be tested. Under the null hypothesis of correct reporting, the expected value of this residual is zero. Under the alternative, it is positive.

The distribution of the unincorporated firms' residual is a convolution of the distributions of overreporting and measurement error, and can be used to recover the distribution of overreporting. Because the output measurement error is independent of juridical organization and corporations do not overreport, I can recover the distribution of the measurement error from the residual of corporations. Deconvolution requires the addends to be independent and the distribution of one of them to be known. Measurement error is independent of overreporting, and its distribution is known. It follows that the distribution of overreporting is identified.

The method can be applied to any dataset containing firm-level inputs and output, and a benchmark group. Firm-level surveys, administrative firm records, or confidential transaction-level invoice data usually contain both.

I find that among the 20 largest manufacturing industries in Colombia during the sample period, the test rejects correct reporting in 9 of the 18 industries liable for the sales tax and in neither exempt industry. In the 5 largest industries where the test rejects at the 1 percent level and the deconvolution applies, unincorporated firms overreport on average between 4.5 and 23.1 percent of their true materials.

3.1 The reported materials share

Regardless of juridical organization, firms in an industry share the Cobb-Douglas technology of Equation 1 but may differ in productivity. The econometrician observes output with measurement error, $Y_{it} = M_{it}^{\beta} K_{it}^{\alpha_K} L_{it}^{\alpha_L} \exp(\omega_{it} + \varepsilon_{it})$, as well as capital, labour, and prices of intermediate inputs and output, and the sales-tax rates on sales and on purchases, $\{Y_{it}, K_{it}, L_{it}, \rho_t, P_t, \tau_S, \tau_P\}$. And she observes only overreported intermediate inputs, $M_{it}^{*} = M_{it} + e_{it}$.

Let s_{it} denote the log share of reported materials in output, net of sales taxes, $s_{it} \equiv \ln\left((1 - \tau_P)\rho_t M_{it}^{*}/(1 - \tau_S)P_t Y_{it}\right)$.[6] To relate it to the materials condition (Equation 4), which holds for true materials, I subtract and add the log share of true materials:

[6] In the single-tax model (Section B.1) the tax terms cancel and s_{it} is the gross share, $\ln(\rho_t M_{it}^{*}/P_t Y_{it})$.

$$
\begin{aligned}
s_{it} &= \ln\left(\frac{(1-\tau_P)\rho_t M^*_{it}}{(1-\tau_S)P_t Y_{it}}\right) - \ln\left(\frac{(1-\tau_P)\rho_t M_{it}}{(1-\tau_S)P_t Y_{it}}\right) + \ln\left(\frac{(1-\tau_P)\rho_t M_{it}}{(1-\tau_S)P_t Y_{it}}\right) \\
&= \ln\left(\frac{M^*_{it}}{M_{it}}\right) + \ln\beta - \varepsilon_{it}.
\end{aligned}
$$

The last line uses Equation 4 to replace the log share of true materials. The remaining term is the log ratio of reported to true materials. Let $u_{it} \equiv \ln(M^*_{it}/M_{it}) = \ln(1 + e_{it}/M_{it})$ denote this log overreporting ratio. Then

$$
s_{it} = \ln\beta - \varepsilon_{it} + u_{it}. \tag{8}
$$

3.2 Identifying the output elasticity of materials

Assumption 1 (Truthful reporters)**.** Corporations do not overreport inputs (Section 1.3).

For corporations, $u_{it} = 0$, and Equation 8 becomes $s_{it} = \ln\beta - \varepsilon_{it}$.

In addition, I impose the following assumption.

Assumption 2 (Independence)**.** The measurement error in output, ε_{it}, is independent of overreporting u_{it}, of true materials M_{it}, and of the firm's juridical organization, with $\mathbb{E}[\varepsilon_{it}] = 0$.

Taking expectations among corporations recovers the output elasticity of materials, because measurement error has mean zero and is independent of juridical organization (Assumption 2):

$$
\ln\beta = \mathbb{E}\left[s_{it} \mid \text{corporation}\right]. \tag{9}
$$

With $\ln\beta$ known, the corporations' share also recovers their measurement error, $\varepsilon_{it} = \ln\beta - s_{it}$, and with it its distribution, f_ε.

3.3 The observed residual

Having recovered the output elasticity of materials, β, I can form the following variable from observed data for every firm. Subtracting $\ln\beta$ from both sides of Equation 8 gives

$$
\mathscr{V}_{it} \equiv s_{it} - \ln\beta = u_{it} - \varepsilon_{it}. \tag{10}
$$

The observed residual recovers overreporting up to measurement error, which is independent of overreporting. Among unincorporated firms, its expected value is the expected overreporting, $\mathbb{E}[\mathscr{V}_{it} \mid \text{unincorporated}] = \mathbb{E}[u_{it} \mid \text{unincorporated}]$, because measurement error has mean zero in

every juridical organization (Assumption 2). Since firms never underreport ($e_{it} \geq 0$, Section 2), $u_{it} \geq 0$, and this expected value is zero only if no unincorporated firm overreports.

3.4 Testing for overreporting

Equation 10 suggests a test for overreporting. Let $\mu_{\mathscr{V}} \equiv \mathbb{E}[\mathscr{V}_{it} \mid \text{unincorporated}]$. The null hypothesis is that unincorporated firms do not overreport, $H_0 : \mu_{\mathscr{V}} = 0$, and the alternative is that they do, $H_1 : \mu_{\mathscr{V}} > 0$. The test is one-sided because no firm has an incentive to underreport deductible inputs. Underreporting would raise a firm's tax liability and decrease after-tax profits (Section 2). A rejection can therefore only point to overreporting.

To test the null hypothesis, I build a confidence set for $\mu_{\mathscr{V}}$ by test inversion. The test detects overreporting where the confidence set is non-empty and excludes zero. Test inversion consists of testing every candidate value μ_0 on a grid over $[0, 1]$ and keeping the values the test does not reject. These values form a confidence set for $\mu_{\mathscr{V}}$. Inverting a test at the 5 percent level gives a 95 percent confidence set (Casella and Berger 2002, sec. 9.2.1).

The test described above implicitly takes β from corporations as the true parameter. I make the assumption explicit by using two moments, one for each group. Let C_{it} equal one for corporations and zero otherwise. If β is the output elasticity of materials and $\mu_0 \geq 0$ is the expected residual of unincorporated firms,

$$\mathbb{E}[C_{it}(s_{it} - \ln \beta)] = 0, \qquad \mathbb{E}[(1 - C_{it})(s_{it} - \ln \beta - \mu_0)] = 0. \tag{11}$$

The first moment restates Equation 9; the second states that the residual of unincorporated firms (Equation 10) has mean $\mu_0 \in [0, 1]$.

Let $\hat{g}(\ln \beta, \mu_0)$ stack the sample means of the two moments over the n firm-years of the industry. The test statistic is

$$TS(\ln \beta, \mu_0) = n\, \hat{g}(\ln \beta, \mu_0)' \, \hat{\Omega}(\ln \beta, \mu_0)^{-1} \, \hat{g}(\ln \beta, \mu_0),$$

where $\hat{\Omega}$ is the covariance matrix of the moments, clustered by plant and re-estimated at each candidate rather than fixed at a preliminary estimate (Anderson and Rubin 1949; Stock and Wright 2000). Under the null hypothesis, the statistic is asymptotically chi-squared with d_g degrees of freedom.[7]

[7] I evaluate the statistic at $\ln \hat{\beta}$, the corporations' mean log share. That is, I take the corporations' estimate as the true value. The first moment is exactly zero at this value, and the statistic tests only the second: whether the residual of unincorporated firms has mean μ_0. Hence $d_g = 1$. Section D reports a conservative version of the test with $d_g = 2$, and an alternative bootstrap test that carries the sampling error of $\ln \hat{\beta}$.

Table 5: Test for overreporting of materials by industry. Mean of $\mathcal{V}_{it}$ among unincorporated firms and 95 percent confidence set for its expectation (sharp test), with $\ln \beta$ estimated from corporations.

Industry	$\hat{\mu}$	95% region, sharp
311 Food products	0.024	$(\cdot,\ 0.048]$
312 Other food products	−0.018	$(\cdot,\ 0.076]$
313 Beverages	0.390	$[0.310,\ 0.470]$
321 Textiles	0.151	$[0.115,\ 0.188]$
322 Wearing apparel	0.224	$[0.195,\ 0.253]$
323 Leather products	−0.192	†
324 Footwear	0.095	$[0.059,\ 0.131]$
331 Wood products	0.250	$[0.175,\ 0.325]$
332 Furniture	−0.007	$(\cdot,\ 0.059]$
341 Paper products	−0.066	$(\cdot,\ 0.004]$
342 Printing and publishing	0.074	$[0.036,\ 0.112]$
351 Industrial chemicals	0.165	$[0.001,\ 0.329]$
352 Other chemicals	0.057	$[0.002,\ 0.111]$
356 Plastic products	−0.107	†
369 Non-metallic minerals	0.193	$[0.109,\ 0.277]$
381 Metal products	−0.005	$(\cdot,\ 0.025]$
382 Non-electrical machinery	−0.014	$(\cdot,\ 0.026]$
383 Electrical machinery	0.034	$(\cdot,\ 0.089]$
384 Transport equipment	−0.129	†
390 Other manufacturing	0.057	$(\cdot,\ 0.115]$

$\hat{\mu}$: mean of $\mathcal{V}_{it} = \ln(\rho_t M^*_{it}/P_t Y_{it}) - \ln \hat{D}$, with the log materials share net of sales taxes and $\ln \hat{D}$ the mean among corporations. 95% test-inversion region for $\mu = E[\mathcal{V}_{it}]$ over the grid $\mu \in [0, 1]$ (under the model $\mu \geq 0$), taking $\ln \hat{D}$ as the true value; plant-clustered, efficiently weighted moments; sharp, $\chi^2_{1,0.95} = 3.84$ critical value (the corporations' moment holds exactly at $\ln \hat{D}$). Where the region excludes 0, the test rejects no overreporting. $(\cdot,\ b]$: $\mu = 0$ is not rejected. † Empty region: every $\mu \in [0, 1]$ is rejected.

The test does not depend on the functional forms of the detection probability and the cost of evasion, nor on the Cobb-Douglas form of the technology. A truthful firm faces no detection and incurs no cost of evasion, $q(0, \cdot) = 0$ and $\kappa(0, \cdot) = 0$; under the null hypothesis, every firm's materials condition is therefore the one without evasion. For any technology that corporations and unincorporated firms share, the residual is then $-\varepsilon_{it}$ under the null, and under the alternative its expected value is positive as long as output is concave in materials. With a second flexible input that is not deductible, a monotone output elasticity of that input replaces concavity, but the sign of the residual under the alternative is then not known in advance. Section D derives both results.

3.5 Testing results

I test for overreporting the 20 largest manufacturing industries by the number of firm-years.

The test detects overreporting in 9 of the 18 industries liable for the sales tax (Table 5): Beverages (313)[8], Textiles (321), Wearing apparel (322), Footwear (324), Wood products (331), Printing and publishing (342), Non-metallic minerals (369) and, narrowly, Industrial chemicals (351) and Other chemicals (352). In the first seven, the test also rejects at the 1 percent level. The nine industries include 4 of the 5 largest by output and together produce 41 percent of manufacturing output.[9]

The test does not detect overreporting in either industry exempt from the sales tax, Food products (311) and Other food products (312). This is expected. Firms in these industries had no incentives to evade sales taxes (Section 1.2). The result aligns with the model, which predicts little overreporting where the reward is small.

In the other nine liable industries, the test fails to reject correct reporting. The confidence set either does not exclude zero or is empty (no candidate in $[0, 1]$ is accepted). As with any statistical test, a failure to reject is inconclusive.

3.6 Deconvolving tax evasion

We can do better, however. The test in Section 3 detects whether unincorporated firms overreport on average, but not how overreporting is spread across firms. Deconvolution recovers the whole distribution of overreporting by removing the output measurement error from the distribution of the observed residual.

The distribution matters because it shows whether overreporting in a particular industry is concentrated in a few aggressive firms or spread across most firms. The two patterns respond differently to the tax rate. A firm overreports only if the credit on the first fictitious peso exceeds the marginal cost of evasion; otherwise it reports truthfully (Equation 7). Among the firms that overreport, the amount differs because the cost of evasion depends on productivity. The test rejection can therefore arise from many firms that each overreport a little or from a few firms that overreport a lot while the rest report truthfully. When the tax rate changes, a firm that already overreports adjusts the overreporting amount (intensive margin), while a truthful firm starts to overreport only if the rate rises enough for the credit to exceed its marginal cost of evasion (extensive margin). The distribution shows how many firms sit near the truthful corner, where a change in the rate can make them start or

[8] In Beverages (313), the mean residual of unincorporated firms, 0.39, is the largest of any tested industry. It partly reflects differences in products between the two groups. Corporations produce mostly beer and soft drinks; unincorporated firms produce wine, spirits and soft drinks. Beer corporations report a low materials share, which lowers the benchmark. Within the same product (four-digit industry), the mean residual falls to 0.23 but remains positive. The mean is also higher because the sample excludes state enterprises (Section 1.3). In this industry they are distilleries, and the departments held a monopoly on producing aguardiente and rum (Ocampo and Perry 1983, 251–54). Their materials share is lower than the corporations'. With them and the other excluded organizations in the sample, the mean residual is 0.23, and the test still rejects correct reporting at the 1 percent level.

[9] The rejections are unlikely to come from testing many industries. If no industry overreported, a test at the 5 percent level across the 20 industries would reject in about one of them by chance, and the probability of nine or more rejections would be about 2×10^{-7}.

stop overreporting.

In the five largest industries where the test rejects at the 1 percent level and the deconvolution applies, unincorporated firms overreport on average between 4.5 and 23.1 percent of their true materials, depending on the industry. Moreover, most firms overreport. In every industry, the 10th percentile of the distribution lies between 3.2 and 18.0 percent of true materials.

3.6.1 The convolution

By Assumption 2, measurement error is independent of overreporting and of juridical organization. Its distribution, f_ε, is therefore the same for corporations and unincorporated firms and known from corporations, whose residual is $-\varepsilon_{it}$ (Section 3.2).

Because $\mathscr{V}_{it}$ is the sum of two independent terms, u_{it} and $-\varepsilon_{it}$, the density of the observed residual among unincorporated firms is the convolution of the densities of overreporting and of measurement error (Meister 2009):

$$f_{\mathscr{V}}(\mathscr{V}) = \int f_u(\mathscr{V} + \varepsilon) f_\varepsilon(\varepsilon)\,\mathrm{d}\varepsilon, \tag{12}$$

where $f_{\mathscr{V}}$ and f_u denote the densities of $\mathscr{V}_{it}$ and u_{it} among unincorporated firms. The data give $f_{\mathscr{V}}$, corporations give f_ε, and f_u is the only unknown. Deconvolution solves Equation 12 for f_u. Section G describes the estimator, a spline density fitted by penalized maximum likelihood.

Deconvolution works here because, under Cobb-Douglas, the elasticity is constant: the log ratio of reported to true materials enters the residual as an additive term, and no other term in the residual depends on the firm's materials or other inputs. In this setting, deconvolution is a simple and elegant tool: it needs only reported shares, the corporations' residuals and Equation 12.

Production functions with an elasticity that varies with materials do not rule out recovering overreporting, but they require either far more data for deconvolution or a different method, such as the one in Section 6. With an elasticity $G(M_{it}, Z_{it})$ that depends on materials and on the other inputs, $Z_{it} = (K_{it}, L_{it})$, the residual would be the reported share minus the log elasticity at reported materials. Corporations identify this elasticity because they report true materials. The residual's overreporting term, $\ln[M^*_{it}/(M^*_{it} - e_{it})] - \ln[G(M^*_{it}, Z_{it})/G(M^*_{it} - e_{it}, Z_{it})]$, would depend on reported materials and the other inputs as well as on overreporting. If measurement error were also independent of the other inputs, deconvolving the residual conditional on (M^*_{it}, Z_{it}) would recover the density of this term. A density transformation would then give the density of overreporting, provided the term increases with e_{it}. Each value of the conditioning inputs, however, would need its own deconvolution, which would demand far more data than the single deconvolution under Cobb-Douglas.

Table 6: Distribution of the overreporting ratio e/M by industry, unincorporated firms.

Industry	Mean	SD	P10	Median	P90
321 Textiles	16.2%	5.3%	9.5%	15.8%	23.4%
322 Wearing apparel	23.1%	3.9%	18.0%	23.1%	28.2%
324 Footwear	10.5%	5.8%	4.4%	9.3%	18.2%
342 Printing and publishing	4.5%	7.2%	3.2%	4.0%	5.1%
369 Non-metallic minerals	12.4%	2.3%	9.5%	12.2%	15.4%

Overreporting ratio $x = e/M$: overreported materials as a share of true materials. Moments and quantiles of the density $f_x(y) = f_u(\ln(1+y))/(1+y)$, obtained by transforming the deconvolved density of $u = \ln(1 + e/M)$ (penalized B-spline deconvolution, log materials share net of sales taxes). Sample: unincorporated firms that pay sales tax on purchases. Industries: the five largest, by share of output, where overreporting is detected at the 1% level and the variance condition is satisfied.

3.6.2 The overreporting ratio

The deconvolution recovers the density of the log ratio u_{it}. To read overreporting as a share of true materials, I transform it into the overreporting ratio, $x_{it} \equiv e_{it}/M_{it} = \exp(u_{it}) - 1$. Because x_{it} is a monotonic, differentiable function of u_{it}, its density follows from a density transformation (Hogg et al. 2019, Theorem 1.7.1),

$$f_x(y) = f_u(\ln(1+y))\left|\frac{1}{1+y}\right|, \qquad y \geq 0. \tag{13}$$

Table 6 shows the distribution of the overreporting ratio in the five largest industries, by share of output, that meet two conditions: the test rejects at the 1 percent level, and the residual is at least as dispersed among unincorporated firms as among corporations, which the deconvolution requires.[10] Section C reports all nine industries that enter the structural estimation.

Overreporting is widespread within industries but differs markedly across them. The mean overreporting ratio ranges from 4.5 percent of true materials in Printing and publishing (342) to 23.1 percent in Wearing apparel (322). In every industry the median of the density is different from zero. Most unincorporated firms overreport, and few report close to truthfully.

The spread differs across industries as well. In Wearing apparel and Non-metallic minerals (369), firms overreport similar shares of their materials. The 10th and 90th percentiles of these industries lie 6 to 10 percentage points apart. Footwear (324) and Textiles (321) have a long right tail. The 90th percentile is about twice the median in Footwear and one and a half times the median in Textiles.

[10] Because overreporting and measurement error are independent (Assumption 2), the variance of the residual among unincorporated firms equals the variance of overreporting plus the variance of measurement error. It therefore cannot fall below the corporations' variance, which is the variance of measurement error alone. I treat the condition as failed when fewer than 5 percent of bootstrap replications, which resample plants, give a ratio of variances of at least 1. Beverages (313) fails it: its unincorporated variance is 0.69 times the corporations', and the ratio reaches 1 in only 3.6 percent of replications.

The model attributes these differences within an industry to the cost of evasion, which varies with productivity and with an idiosyncratic shock (Section 2).

4 Production Function Parameters and Productivity

Materials overreporting biases the production-function and productivity estimates of proxy-variable methods that rely on a flexible input (Levinsohn and Petrin 2003; Doraszelski and Jaumandreu 2013; Ackerberg et al. 2015; Gandhi et al. 2020). These methods typically use materials as the flexible input. For example, the output elasticity of materials can be recovered from the materials log share of output (Gandhi et al. 2020). Since reported materials are inflated, the estimate of the elasticity is too large. This bias then permeates the rest of the production-function parameters and productivity.

In this section, I show how to recover the remaining parameters of the production function, the elasticities of capital and labour, and productivity net of overreporting.

Failing to correct for overreporting leads to a larger output elasticity of materials and more dispersed and less persistent productivity. Uncorrected estimates overstate the ratio of the 90th to the 10th percentile of productivity by 66 percent and understate the half-life of productivity shocks by 63 percent, in the median of the five industries.

4.1 Identification

In the literature, it is not uncommon to assume that productivity follows a Markov process. That is,

$$\omega_{it} = g(\omega_{it-1}) + \eta_{it} \tag{14}$$

Let $\ln X_{it} = x_{it}$. With the estimate of β and the observed residual $\mathscr{V}_{it}$ at hand, I can form the following observed variable,

$$\begin{aligned} \mathscr{W}_{it} &\equiv y_{it} - \beta(m^*_{it} - \mathscr{V}_{it}) \\ &= y_{it} - \beta(m_{it} + u_{it} - u_{it} + \varepsilon_{it}) \\ &= y_{it} - \beta m_{it} - \beta\varepsilon_{it} \\ &= \alpha_K k_{it} + \alpha_L l_{it} + \omega_{it} + (1-\beta)\varepsilon_{it} \end{aligned} \tag{15}$$

Here, intuitively, I am trading unobservables, the unobserved tax evasion for the measurement error.

By replacing ω_{it} from Equation 15 in Equation 14, I get

$$\mathscr{W}_{it} = (1-\beta)\varepsilon_{it} + \alpha_K k_{it} + \alpha_L l_{it} + g(\mathscr{W}_{it-1} - (1-\beta)\varepsilon_{it-1} - \alpha_K k_{it-1} - \alpha_L l_{it-1}) + \eta_{it} \quad (16)$$

If productivity follos an AR(1) process $g(\xi) = \gamma_0 + \gamma_1 \xi$, I can use instruments Z to learn α and g, such that,

$$\mathbb{E}[\eta_{it} + (1-\beta)\varepsilon_{it} + \gamma_1(1-\beta)\varepsilon_{it-1}|Z] = 0. \quad (17)$$

Once the parameters of the production function are identified, I can use the following observed variable to recover ω_{it},

$$\begin{aligned} \widetilde{\mathscr{W}}_{it} &\equiv \mathscr{W}_{it} - \alpha_K k_{it} - \alpha_L l_{it} \\ &= \omega_{it} + (1-\beta)\varepsilon_{it} \end{aligned} \quad (18)$$

As the instrument, I use $\widetilde{\mathscr{W}}_{it-2}$ lagged two years because it is correlated with past productivity but independent of the innovation, η_{it}, and of the two measurement errors in Equation 17.

To recover the distribution of productivity net of measurement error, I deconvolve $\widetilde{\mathscr{W}}_{it} = \omega_{it} + (1-\beta)\varepsilon_{it}$ with the corporations' distribution of measurement error, as in Section 3.6. Section C.1 provides more details of the estimation for both the production function and productivity.

4.2 Production function and productivity estimates

The estimates cover the five industries of Section 3.6. Section C reports all nine industries of the structural estimation.

As expected, the uncorrected estimates in Table 7 display larger output elasticities of materials in all five industries. The uncorrected elasticities of capital and labour are higher in some industries, lower in others, and similar in the rest. This is expected, as capital and labour now absorb a different variation from output that was being captured by materials before.

Failing to correct for overreporting raises dispersion and reduces persistence in productivity. The 90th-to-10th-percentile ratio of $\exp(\omega_{it})$ is 1.70 to 2.30 with the correction, against 2.38 to 6.86 for GNR. In the median of the five industries, the uncorrected ratio is 66 percent larger (range 23 to 291 percent; Table 8).

With respect to persistence, the AR(1) coefficient of ω_{it} is 0.93 to 0.96 with the correction, against 0.81 to 0.89 for GNR. The half-life of a productivity shock, $\ln 0.5 / \ln \hat{\gamma}_1$, is therefore 9 to 15 years with the correction and 3 to 6 years without it. In the median of the five industries, the uncorrected

Table 7: Production function estimates by method: evasion-corrected with instrument $\tilde{\mathcal{W}}_{it-2}$, and uncorrected GNR and OLS. Below each corrected estimate, the 95 percent test-inversion region (conservative regions in Table 23). Other instruments: Table 26.

		Corrected, $\tilde{\mathcal{W}}_{it-2}$		GNR			OLS		
Industry	$\hat{\beta}$	$\hat{\alpha}_K$	$\hat{\alpha}_L$	$\hat{\beta}$	$\hat{\alpha}_K$	$\hat{\alpha}_L$	$\hat{\beta}$	$\hat{\alpha}_K$	$\hat{\alpha}_L$
321	0.385	0.300	0.349	0.435	0.225	0.340	0.681	0.093	0.274
	[0.357, 0.416]	[0.210, 0.375]	[0.230, 0.485]						
322	0.335	0.182	0.392	0.417	0.175	0.343	0.578	0.047	0.397
	[0.288, 0.391]	[0.085, 0.265]	[0.320, 0.470]						
324	0.412	0.119	0.427	0.451	0.131	0.406	0.705	0.013	0.300
	[0.365, 0.465]	[−0.215, 0.270]	[0.190, 0.570]						
342	0.333	0.292	0.408	0.355	0.221	0.476	0.578	0.066	0.430
	[0.294, 0.377]	[0.215, 0.360]	[0.305, 0.530]						
369	0.244	0.333	0.549	0.279	0.254	0.557	0.495	0.128	0.519
	[0.207, 0.288]	[0.210, 0.485]	[0.210, 0.740]						

Evasion-corrected: $\hat{\beta}$ is the corporations' first-stage estimate (log materials share net of sales taxes); below it, its 95% test-inversion region from the corporations' share moment (χ^2_1). Point estimates of (α_K, α_L): the minimum of the test-inversion statistic, with the covariance of the moments re-estimated at each candidate, the statistic whose regions are shown. Instrument $\tilde{\mathcal{W}}_{it-2}$ (exactly identified); below the estimates, the projections of the 95% test-inversion region for (α_K, α_L) (χ^2_2, 0.005 grid over $[0, 1]^2$, extended beyond it where the region reaches its edge). GNR and OLS: uncorrected estimates on the same sample (GNR's share regression on the log materials share net of sales taxes). Industries: the five largest, by share of output, where overreporting is detected at the 1% level and the variance of the unincorporated firms' residual is at least the corporations'.

half-life is 63 percent shorter (range 37 to 67 percent; Table 8).[11]

5 Fiscal Policy Evaluation: The 1983 Reform

Firms are known to overreport costs when tax enforcement tightens, which raises the expected cost of evading, but less is known about the benefit side: whether they overreport more when the tax rate, which sets how much each peso of fictitious inputs saves, goes up. I investigate this question and find evidence that input overreporting rises when the tax rate increases. In 1983, Colombia raised the sales-tax rate in most manufacturing industries while others remained exempt. Overreporting by unincorporated firms rose after the reform, reaching about 8 percent of true materials by 1987, and the increase was concentrated in the industries where the rate went up.

Whether firms evade more when tax rates rise matters for policymakers. If evasion rises with the rate, a tax increase raises less revenue than its mechanical effect suggests, and evasion affects the marginal efficiency cost of raising tax rates (Slemrod 2019, 907). The concern is especially relevant for sales and value-added taxes, whose credit on purchases is a natural target of fictitious deductions

[11] In the literature, productivity is defined in levels as $\exp(\omega_{it} + \varepsilon_{it})$, with ε_{it} an ex-post output shock. Here, ε_{it} is measurement error in output, and productivity is $\exp(\omega_{it})$. The uncorrected estimates in Table 8 use the same convention. Both compare ω_{it} with ω_{it}. The published productivity ratios of Gandhi et al. (2020), however, are not directly comparable with these.

Table 8: Productivity corrected for tax evasion (deconvolved, instrument $\tilde{\mathscr{W}}_{it-2}$) and uncorrected (GNR), by industry.

Industry	Method	Mean	SD	Median	75/25	80/20	90/10	95/5	Skewness	$\hat{\gamma}_1$
321	Corrected	16.5	6.7	15.2	1.41	1.56	2.30	3.39	-0.04	0.95
	GNR	21.2	11.0	19.0	1.82	2.15	3.38	5.31	-0.34	0.87
322	Corrected	57.8	15.9	57.6	1.42	1.58	2.16	2.73	-0.48	0.94
	GNR	34.9	16.5	33.3	1.80	2.13	3.59	5.65	-0.70	0.88
324	Corrected	42.4	11.9	40.8	1.32	1.44	1.94	2.49	0.28	0.96
	GNR	28.8	10.4	27.5	1.53	1.74	2.38	3.08	-0.17	0.87
342	Corrected	19.1	4.0	19.0	1.25	1.33	1.70	2.13	-0.30	0.93
	GNR	24.2	11.5	22.2	1.69	1.92	2.88	4.14	-0.20	0.81
369	Corrected	19.0	5.6	18.6	1.30	1.39	1.75	2.41	-0.65	0.93
	GNR	32.8	22.9	28.3	2.69	3.33	6.86	10.95	-0.49	0.89

Productivity in levels, $\exp(\omega_{it})$; percentile ratios as in Gandhi, Navarro and Rivers (2020, Table 3); skewness of ω_{it}; $\hat{\gamma}_1$: persistence, the AR(1) coefficient of ω_{it} (corrected: IV estimate within the production function step; GNR: OLS on firm-level ω_{it}). Corrected: moments of the deconvolved density of ω (penalized B-spline deconvolution of $\widetilde{W}_{it} = \omega_{it} + (1-\beta)\,\varepsilon_{it}$), with the production function estimates of instrument $\tilde{\mathscr{W}}_{it-2}$. GNR: firm-level ω_{it} from the uncorrected GNR (2020) Cobb-Douglas estimation, with the measurement error removed. Log materials share net of sales taxes, juridical organization codes 6–9 excluded. Industries: the five largest industries, by share of output, where overreporting is detected at the 1 percent level and the variance of the unincorporated firms' residual is at least the corporations'.

(Carrillo et al. 2022). Yet recent empirical work says more about enforcement than about tax rates.

The 1983 reform changed statutory rates rather than enforcement, which makes it a rare setting in which to study this question. Governments have agreed to randomize letters, audits and information reporting, but not tax rates (Slemrod 2019), and among the studies Slemrod reviews, the evidence on tax rates comes mostly from misreported imports (Fisman and Wei 2004; Handley and Moore 2017). The same literature shows that firms adjust reported costs: when tax authorities pushed firms to report more revenue, Ecuadorian firms raised reported costs by 96 and Armenian firms 70–80 cents per dollar of additional revenue (Carrillo et al. 2017; Asatryan and Peichl 2017).

In the model, the sales tax rate shifts the marginal benefit of overreporting materials: a higher rate on purchases, τ_P, raises the credit claimed on each peso of fictitious inputs (Section 2). Although the model has no income tax, the same logic extends to it: overreported inputs also lower reported profits and, thus, the income-tax base. An increase in either rate therefore raises the incentive to overreport inputs, and a decrease lowers it.

The 1983 fiscal policy changed incentives *across* industries and *within* industries. Across industries, the sales-tax (ST) rate increased from 6 to 10 percent for most ST-liable industries, while ST-exempt industries, food products, remained exempt (Section 1.4). Within industries, the income tax changed by juridical organization (JO), and in the opposite direction. *Ley 9 de 1983* cut individual income-tax rates by 4.6 percentage points on average across income brackets relative to the 1982 schedule, and

Table 9: Average individual income-tax rates by taxable income, 1982 schedule (Decreto 2809) vs. the 1983 schedule that stayed (Decreto 397, Ley 9 de 1983), transcribed from Perry and Cárdenas (1986), Cuadro III.1.

Taxable income (thousand 1982 pesos)	1982 schedule	1983 schedule	Change (pp)	Change (%)
200	5.14%	0.24%	-4.90	-95.3
300	10.03%	6.92%	-3.11	-31.0
400	15.65%	10.14%	-5.51	-35.2
500	17.45%	12.80%	-4.65	-26.6
600	21.01%	15.83%	-5.18	-24.7
800	26.28%	21.03%	-5.25	-20.0
1,000	30.06%	24.85%	-5.21	-17.3
1,500	35.61%	31.46%	-4.15	-11.7
2,000	39.03%	35.30%	-3.73	-9.6
Average			**-4.63**	**-22.0**[a]

Average income-tax rates for an individual declarant with three dependants. 1982: Decreto 2809; 1983: Decreto 397 (Ley 9 de 1983). Source: Perry and Cárdenas (1986), vol. 1, Cuadro III.1.

[a] Excluding the lowest bracket, whose rate falls to almost zero (-95.3%).

thus proprietorships faced lower rates (Perry and Cárdenas 1986). LLCs received a smaller cut, from 20 to 18 percent (Perry and Cárdenas 1986). Overreporting inputs lowers both taxes, through a larger credit on purchases and a smaller income-tax base, so the two changes pull the incentive to overreport in opposite directions. In ST-liable industries, the net incentive is therefore ambiguous for both: proprietorships were compensated for much of the sales-tax increase through the income tax, while LLCs were insufficiently compensated. In ST-exempt industries, the income-tax cut alone lowered the incentive for proprietorships and changed it little for LLCs, whose cut was small.

Because the income-tax cut differed between proprietorships and LLCs, the reform's incentives also varied along the size distribution. Proprietorships have a single owner and are the smallest. Tax authorities find sole proprietorships and small businesses hard to penetrate, and evasion in them is more sustainable (Slemrod 2019, 940). At the median, LLCs have about two and a half times proprietorships' labour and three times their gross output. LLCs have several owners and non-tradable shares. In Spain, firms bunch just below the revenue threshold above which the Large Taxpayers' Unit monitors them more closely (Almunia and Lopez-Rodriguez 2018; Slemrod 2019, 925–26). At the top, corporations are several times larger than LLCs in capital, labour and gross output, and their shares are tradable. In addition, the government keeps corporations under close supervision and requires them to have an auditor. Lastly, since evidence suggests stock prices of public companies fall when news of their involvement in tax shelters becomes public (Hanlon and Slemrod 2009; Slemrod 2019, 909), corporations might have lower incentives to evade. In this chapter corporations serve as the benchmark of non-evaders, on the assumption that audited

Table 10: Implied marginal income-tax rates by income interval, 1982 vs. 1983 schedules, computed from the average rates in Table 9 as the change in tax over the change in income between consecutive points. Each figure is an average over the interval.

Taxable income interval (thousand 1982 pesos)	1982 schedule	1983 schedule	Change (pp)
200–300	19.8%	20.3%	+0.5
300–400	32.5%	19.8%	-12.7
400–500	24.6%	23.4%	-1.2
500–600	38.8%	31.0%	-7.8
600–800	42.1%	36.6%	-5.5
800–1,000	45.2%	40.1%	-5.0
1,000–1,500	46.7%	44.7%	-2.0
1,500–2,000	49.3%	46.8%	-2.5

Implied marginal rates: the change in tax over the change in income between consecutive points of Perry and Cárdenas (1986), Cuadro III.1 (average rates). Each figure is an average over the interval.

statements, third-party reporting and the market's reaction keep them from overreporting.

5.1 Empirical approach

From the materials FOC, for corporations, the truth-reporters, i in industry j at year t, the log share of materials' cost over sales is

$$s^C_{ijt} = \ln \beta_j - \varepsilon_{ijt} \tag{19}$$

where $s_{ijt} \equiv \ln\left((1-\tau_{P,ijt})\rho_t M^*_{ijt}/((1-\tau_{S,ijt})P_t Y_{ijt})\right)$ is the log share of materials' cost over sales, both net of sales taxes (Section E), β_j is the output elasticity of materials of industry j, and ε_{ijt} is the output error term. For unincorporated firms, the evaders,

$$s^N_{ijt} = \ln \beta_j + u_{ijt} - \varepsilon_{ijt} \tag{20}$$

where $u_{ijt} = \ln(1 + e_{ijt}/M_{ijt})$ is the overreporting of materials' cost due to tax evasion.

Let D^N_{ijt} be a dummy indicating whether the firm is an unincorporated firm. Then,

$$\begin{aligned}(1-D^N_{ijt})s_{ijt} + D^N_{ijt}s_{ijt} &= (1-D^N_{ijt})(\ln\beta_j - \varepsilon_{ijt}) + D^N_{ijt}(\ln\beta_j + u_{ijt} - \varepsilon_{ijt}) \\ s_{ijt} &= \ln\beta_j + D^N_{ijt}u_{ijt} - \varepsilon_{ijt}\end{aligned} \tag{21}$$

Taking expectations,

$$E[s_{ijt}|D^N = 0, j] = \ln \beta_j, \qquad E[s_{ijt}|D^N = 1, j] - E[s_{ijt}|D^N = 0, j] = \mu \tag{22}$$

identifies the average overreporting level, $\mu \equiv E[u_{ijt} \mid D^N_{ijt} = 1]$, a population expectation; its estimate is $\hat{\mu}$.

In the case of the fiscal policy changes, I am interested in the changes in average tax evasion after the 1983 fiscal reform. Taking the difference of the expectations between year t' and $t = 1983$,

$$\begin{aligned} s_{ijt'} - s_{ij,1983} &= D^N_{ijt}(u_{ijt'} - u_{ij,1983}) - (\varepsilon_{ijt'} - \varepsilon_{ij,1983}) \\ E[s_{ijt'} - s_{ij,1983}|D^N = 1] &= \mu_{t'} - \mu_{1983} \equiv \Delta\mu_{t'} \end{aligned} \tag{23}$$

where μ_t is the annual average tax evasion of year t, and $\Delta\mu_t$ is the change in average tax evasion between year t and 1983, the last full year before the reform took effect (Section 1.4).

The identifying assumptions are explicit in Equation 21 and Equation 23: (i) corporations do not overreport (Equation 19); (ii) firms within an industry share a common Cobb-Douglas technology, with a constant output elasticity of materials β_j (Equation 19 and Equation 20); and (iii) the measurement error ε_{ijt} is independent of evasion, true materials and juridical organization, with $E[\varepsilon_{ijt} \mid D^N_{ijt}, t] = 0$.

Even though Equation 23 resembles a traditional difference-in-differences design, the result does not require parallel trends for identification. Under (i)-(iii), Equation 21 and Equation 23 show that the difference between unincorporated firms and corporations in year t is the average overreporting itself, μ_t, and changes in μ_t over time, including before 1983, are changes in evasion.

To estimate Equation 21, I use the following regression specification,

$$s_{ijt} = \ln \beta_j + \sum_{t=1981}^{1991} \mu_t D^N_{ijt} \mathbb{1}[\text{year} = t] + \zeta_{ijt} \tag{24}$$

where μ_t is the average tax evasion of unincorporated firms in year t; $\ln \beta_j$ enter as industry fixed effects, estimated jointly with the μ_t, so no first-stage estimate enters the regression and its standard errors need no bootstrap or generated-regressor correction; and $\zeta_{ijt} = \nu_{ijt} - \varepsilon_{ijt}$, where $u_{ijt} = \mu_t + \nu_{ijt}$ and ν_{ijt} is the idiosyncratic component of tax evasion that is not captured by the year dummies. Likewise, I estimate Equation 23 with the following regression,

$$s_{ijt} = \ln \beta_j + \mu D^N_{ijt} + \sum_{t=1981}^{1991} \mathbb{1}[t \neq 1983] \Delta\mu_t D^N_{ijt} \mathbb{1}[\text{year} = t] + \zeta_{ijt} \tag{25}$$

where $\Delta\mu_t$ is the difference in tax evasion of unincorporated firms between year t and 1983, the last

full year before the reform took effect, and μ is the average tax evasion of unincorporated firms in 1983. The outcome is the share net of sales taxes, so the estimates exclude the mechanical effect of the rate change on the share; Section E estimates that effect.

The 1983 fiscal reform changed incentives to evade taxes over time *differently* for firms in ST-liable and ST-exempt industries. Most ST-liable industries faced an ST rate *increase* from 6 to 10 percent, while ST-exempt industries *remained exempt* before and after the reform. Note that firms in *all* industries are subject to the income tax, which also generates incentives to evade taxes by overreporting input costs: lower reported profits due to higher input costs reduce the tax base, and thus the tax liability. Therefore, the income tax and the ST generate tax evasion incentives for firms in ST-liable industries, while only the income tax does for firms in ST-exempt industries. For unincorporated firms,

$$\begin{aligned} s^N_{ijt} &= \ln \beta_j + u^{ST}_{ijt} + u^{IT}_{ijt} - \varepsilon_{ijt} \quad j \in \text{ST-Liable} \\ s^N_{ijt} &= \ln \beta_j + u^{IT}_{ijt} - \varepsilon_{ijt} \qquad\quad j \in \text{ST-Exempt} \end{aligned} \tag{26}$$

where u^{ST}_{ijt} and u^{IT}_{ijt} stand for the overreporting due to the ST and the income tax (IT), respectively. Let D^E_{ijt} be a dummy indicating whether the industry is ST-exempt and D^L_{ijt} a dummy indicating whether the industry is ST-liable, such that $D^E_{ijt} + D^L_{ijt} = 1$. Combining Equation 19 with Equation 26,

$$s_{ijt} = \ln \beta_j + D^N_{ijt} D^E_{ijt}\, u^{IT}_{ijt} + D^N_{ijt} D^L_{ijt} (u^{ST}_{ijt} + u^{IT}_{ijt}) - \varepsilon_{ijt} \tag{27}$$

Taking expectations,

$$\begin{aligned} E[s_{ijt} | D^N = 1, D^E = 1, j] - \ln \beta_j &= \mu^{IT}_E \\ E[s_{ijt} | D^N = 1, D^L = 1, j] - \ln \beta_j &= \mu^{ST}_L + \mu^{IT}_L \equiv \mu^{ST+IT}_L \end{aligned}$$

identifies the average overreporting by firms in ST-exempt industries, μ^{IT}_E, and in ST-liable industries, μ^{ST+IT}_L. If I assume $\mu^{IT}_E = \mu^{IT}_L$, I can also identify the overreporting due to ST, μ^{ST}_L. Since income-tax rates depend on the juridical organization and not on the industry, I assume that the income tax generates the same incentives to evade taxes across industries, whether they are exempt of ST or not. Because I aggregate different juridical organizations in exempt and liable industries into unincorporated firms, this assumption might break if the composition of juridical organizations is non-trivially different between exempt and liable industries. The regression versions of Equation 24 and Equation 25 interact the year dummies with $D^N_{ijt} D^E_{ijt}$ and $D^N_{ijt} D^L_{ijt}$, giving one path per industry type, $\mu_{E,t}$ and $\mu_{L,t}$ (and $\Delta\mu_{E,t}$, $\Delta\mu_{L,t}$ relative to 1983). Splitting unincorporated firms further by JO, into LLCs and proprietorships, gives four such paths.

5.1.1 Changes in true materials cannot explain the pattern

Since $u_{ijt} = \ln(1 + e_{ijt}/M_{ijt})$, the empirical exercise remains valid as long as changes in s are not explained by changes in M (or ε), but only by e or e/M: changes in the proxy should be explained by changes to the incentives to adjust overreporting, not by firms adjusting their optimal choice of true inputs.

The alternative argument is that changes in true inputs explain the increase in the overreporting share. Suppose e is unchanged. Then, after the reform, unincorporated firms would have to lower their true inputs, by either an increase in relative prices or productivity shocks. Productivity shocks are unrelated to tax rates, so this is unlikely. According to the FOC, taxes do not show in the optimal choice of firms. Besides, the increase after the reform is gradual, while the reform made only a one-time change. So this alternative argument does not hold. The evasion channel is more direct: an increase in the sales tax rate leads to an increase in the overreporting share because a higher tax rate increases the marginal benefit of overreporting.

5.2 Results: all unincorporated firms

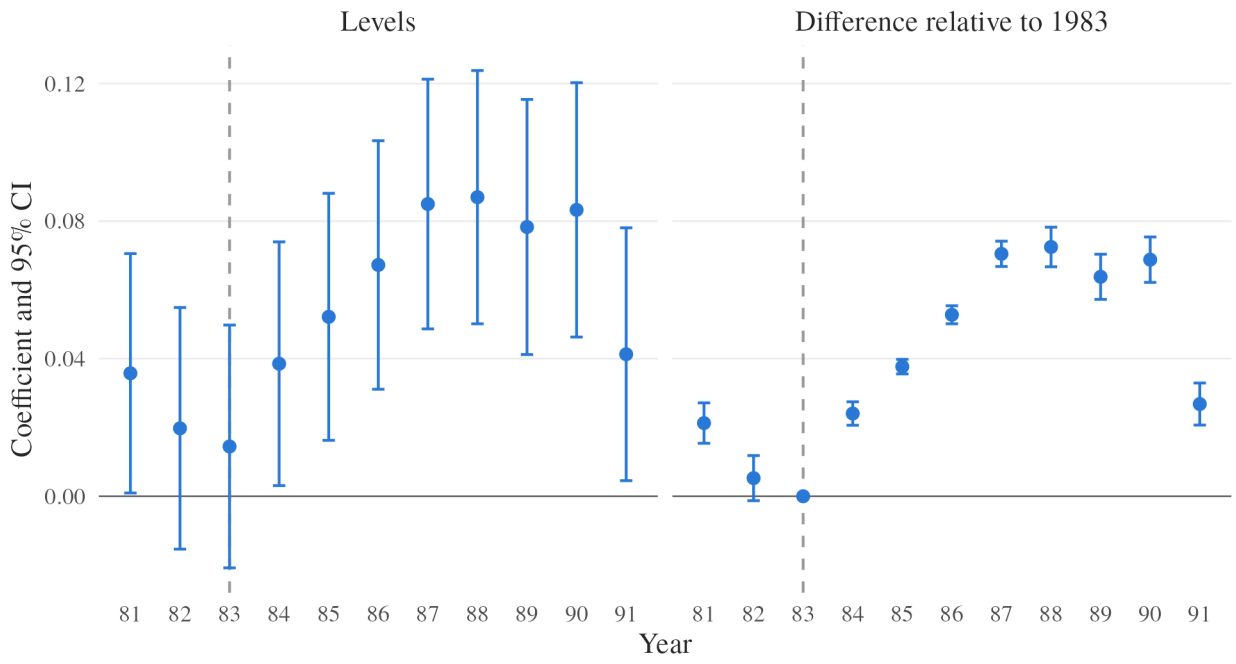


Figure 1: Overreporting of unincorporated firms relative to corporations by year, 1981–1991, all industries. Points are coefficient estimates with 95 percent confidence intervals; standard errors two-way clustered by plant and year; industry fixed effects. Left panel: levels, $\hat{\mu}_t$; right panel: differences with respect to 1983, $\Delta\hat{\mu}_t$. The dashed line marks 1983, the last full year before the reform took effect.

Overreporting rose once the reform raised the benefit of overreporting in most industries: after

Table 11: Overreporting of unincorporated firms relative to corporations, all industries. Coefficients behind the figure, by year: levels ($\hat{\mu}_t$) and differences with respect to 1983 ($\Delta\hat{\mu}_t$).

Year	Level	Diff. to 1983
1981	0.036* (0.018)	0.021*** (0.003)
1982	0.020 (0.018)	0.005 (0.003)
1983	0.014 (0.018)	–
1984	0.039* (0.018)	0.024*** (0.002)
1985	0.052** (0.018)	0.038*** (0.001)
1986	0.067*** (0.018)	0.053*** (0.001)
1987	0.085*** (0.019)	0.070*** (0.002)
1988	0.087*** (0.019)	0.072*** (0.003)
1989	0.078*** (0.019)	0.064*** (0.003)
1990	0.083*** (0.019)	0.069*** (0.003)
1991	0.041* (0.019)	0.027*** (0.003)

Coefficients with standard errors in parentheses, two-way clustered by plant and year. Industry fixed effects in all columns. Level: $\hat{\mu}_t$; Diff.: $\Delta\hat{\mu}_t = \hat{\mu}_t - \hat{\mu}_{1983}$ (1983 is the reference, –). * p<0.1, ** p<0.05, *** p<0.01. Observations: 39,495.

the sales-tax increase took effect, input overreporting by unincorporated firms increased gradually (Figure 1, Table 11). Input overreporting reached an average of about 9 percent of true materials in 1987, where it stabilized until 1990.[12]

The standard errors of the levels and of the differences over time look very different. In levels, the large standard errors suggest that tax evasion is widely heterogeneous across firms. This heterogeneity may come from differences in the tax rate due to firm characteristics (industry, JO, and exports), and in the cost of evasion across firms. The differences over time, however, are tight for a mechanical reason. Every year's level $\hat{\mu}_t$ is estimated against the same baseline, so estimation error in that baseline enters every year with the same sign, and the level estimates are strongly positively correlated. Since $\mathrm{Var}(\hat{\mu}_t - \hat{\mu}_{1983}) = \mathrm{Var}(\hat{\mu}_t) + \mathrm{Var}(\hat{\mu}_{1983}) - 2\,\mathrm{Cov}(\hat{\mu}_t, \hat{\mu}_{1983})$, this shared error cancels in the differences. The correlation affects inference only, not the consistency of the estimates: the regressors are group-by-year dummies, and $E[\zeta_{ijt} \mid D^N_{ijt}, \text{year}] = 0$ because ε is independent of (e, M) and ν_{ijt} is, by construction, the deviation from the group-year mean. Standard errors are two-way clustered by plant and year.

5.2.1 Overreporting before the reform

Corporations are a clean comparison group that will not react to the policy changes, because the assumption is that these firms do not evade taxes by overreporting. Thus, I can interpret the estimates

[12]This path closely follows estimates of sales-tax evasion computed from Colombia's national accounts (Sánchez Torres and Gutiérrez Sourdis 1994, 107). In these estimates, the uncollected share of potential revenue from the sales tax on domestic transactions rose from 58 percent in 1984 to about 69 percent in 1987, stayed between 68 and 71 percent until 1990, and fell to 65 percent in 1991, the same year overreporting fell (Figure 9).

Table 12: Overreporting of unincorporated firms in ST-exempt and ST-liable industries. Coefficients behind the figure, by year: levels ($\hat{\mu}_t$) and differences with respect to 1983 ($\Delta\hat{\mu}_t$).

	ST-exempt		ST-liable	
Year	Level	Diff.	Level	Diff.
1981	0.001 (0.027)	-0.016** (0.006)	0.044* (0.021)	0.029*** (0.003)
1982	0.003 (0.028)	-0.015* (0.008)	0.025 (0.021)	0.010** (0.004)
1983	0.017 (0.028)	–	0.015 (0.021)	–
1984	0.024 (0.028)	0.007 (0.004)	0.043* (0.021)	0.028*** (0.002)
1985	0.022 (0.030)	0.005 (0.003)	0.060** (0.021)	0.045*** (0.001)
1986	0.017 (0.029)	0.000 (0.002)	0.079*** (0.021)	0.064*** (0.002)
1987	0.034 (0.029)	0.017*** (0.004)	0.097*** (0.021)	0.082*** (0.002)
1988	0.023 (0.031)	0.006 (0.009)	0.102*** (0.022)	0.087*** (0.003)
1989	0.036 (0.031)	0.019* (0.009)	0.089*** (0.022)	0.073*** (0.004)
1990	0.024 (0.031)	0.007 (0.009)	0.097*** (0.022)	0.082*** (0.004)
1991	0.001 (0.032)	-0.016 (0.009)	0.051** (0.022)	0.036*** (0.003)

Coefficients with standard errors in parentheses, two-way clustered by plant and year. Industry fixed effects in all columns. Level: $\hat{\mu}_t$; Diff.: $\Delta\hat{\mu}_t = \hat{\mu}_t - \hat{\mu}_{1983}$ (1983 is the reference, –). * p<0.1, ** p<0.05, *** p<0.01. Observations: 39,495.

as the change in overreporting relative to 1983, the last full year before the reform took effect.

The policy changes of the time were the result of a public debate following the 1981 report of the Bird–Wiesner Mission, a technical mission on Colombia's intergovernmental finances led by Richard Bird (international tax expert) and Eduardo Wiesner (Colombian economist) that recommended better tax administration to reduce evasion (Junguito and Rincón 2004; Bolaños Bolaños 2019).[13] The reform was enacted at the end of December 1983 and took effect on 1 April 1984 (Section 1.4). These changes aimed to reduce the fiscal deficit without resorting to money financing and to reactivate investment through tax incentives, while broadening the tax base and simplifying collection (Sánchez Torres and Gutiérrez Sourdis 1994).

A plausible scenario is that firms could have reduced their overreporting during the public debate as a way to signal compliance and avoid further tax increases, and then reverted their behaviour after the reform was enacted. If this is the case, then part of the increase in overreporting after 1983 is a return to previous levels of tax evasion. More data points before 1981 would help to clarify this issue.

5.3 Results: ST-liable vs. ST-exempt industries

The increase is concentrated where the sales tax, and with it the benefit of overreporting, went up. After the fiscal reform of 1983, unincorporated firms in ST-liable industries increased tax evasion

[13]The mission also recommended stronger collection tools, wider withholding, and more auditing (Bolaños Bolaños 2019, 723).

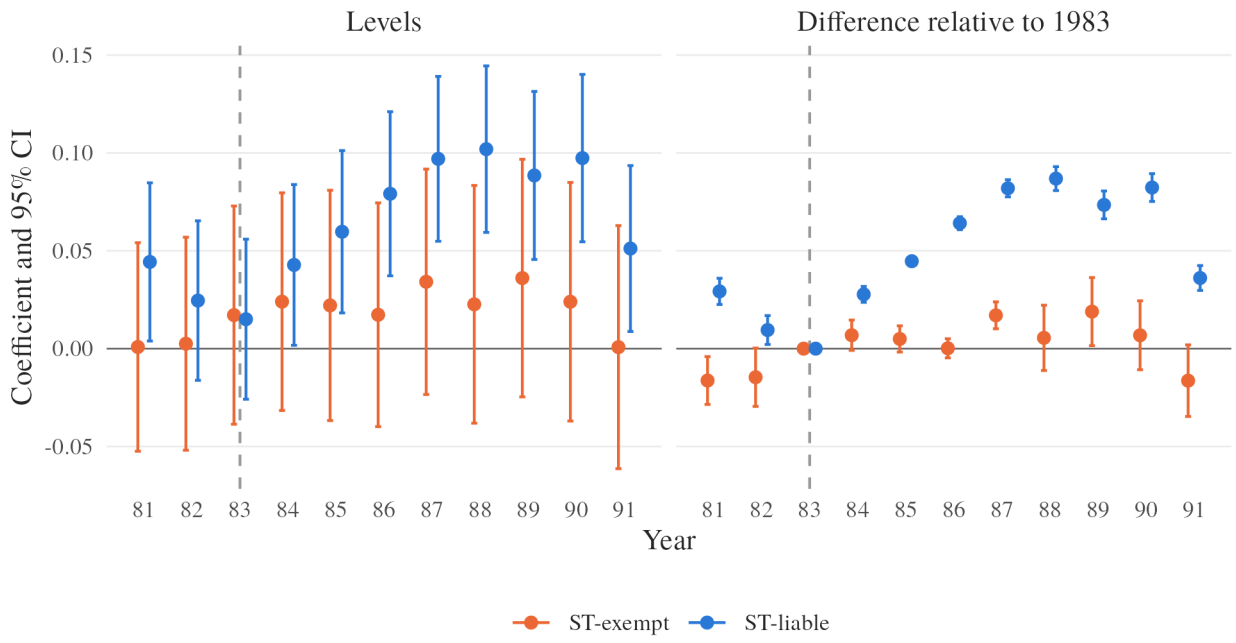


Figure 2: Overreporting of unincorporated firms relative to corporations by year, 1981–1991, in ST-exempt and ST-liable industries. Points are coefficient estimates with 95 percent confidence intervals; standard errors two-way clustered by plant and year; industry fixed effects. Left panel: levels, $\hat{\mu}_t$; right panel: differences with respect to 1983, $\Delta\hat{\mu}_t$. The dashed line marks 1983, the last full year before the reform took effect.

significantly, while those in ST-exempt industries increased it only slightly (Figure 2, Table 12). In ST-liable industries, overreporting rose gradually until 1987 and stabilized until 1990. In ST-exempt industries, the increase was small, at most two percentage points (1987 and 1989). The increase is *larger* for ST-liable firms, the most affected by the reform.

This comparison isolates the sales tax as a shifter of the marginal benefit: ST-liable industries faced the rate increase, while ST-exempt industries did not. Firms in ST-liable industries faced a higher sales tax, only partly offset by lower income-tax rates, while firms in ST-exempt industries faced the income-tax cut alone. The larger response in ST-liable industries is consistent with the sales-tax increase dominating. The slight increase in ST-exempt industries, driven by proprietorships, is unexpected; Section 5.4 discusses two possible explanations.

5.4 Results: income-tax incentives, LLCs vs. proprietorships

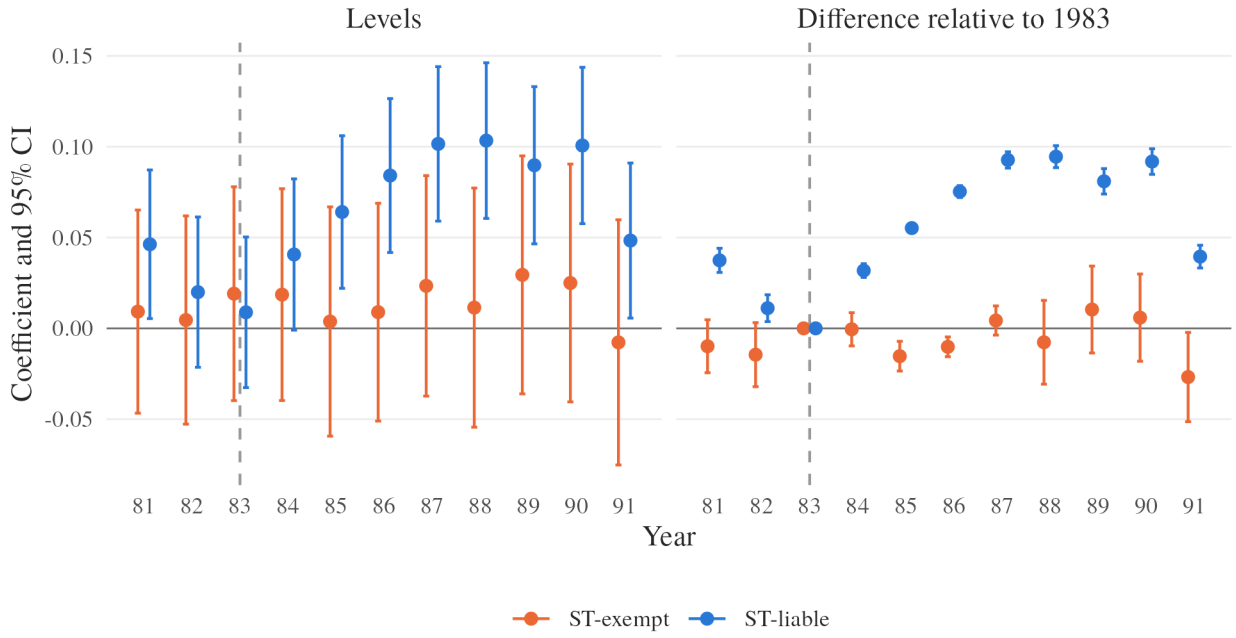


Figure 3: Overreporting of LLCs relative to corporations by year, 1981–1991, in ST-exempt and ST-liable industries. Points are coefficient estimates with 95 percent confidence intervals; standard errors two-way clustered by plant and year; industry fixed effects. Left panel: levels, $\hat{\mu}_t$; right panel: differences with respect to 1983, $\Delta\hat{\mu}_t$. The dashed line marks 1983, the last full year before the reform took effect.

Within industries, the income-tax cut lowered the benefit of overreporting, by more for proprietorships than for LLCs. LLCs in ST-liable industries show a significant increase in tax evasion, while LLCs in exempt industries show changes of about one percentage point or less (Figure 3, Table 13). LLCs in exempt industries mainly faced a decrease in their entity-level income tax. This decrease would

Table 13: Overreporting of LLCs in ST-exempt and ST-liable industries. Coefficients behind the figure, by year: levels ($\hat{\mu}_t$) and differences with respect to 1983 ($\Delta\hat{\mu}_t$).

	ST-exempt		ST-liable	
Year	Level	Diff.	Level	Diff.
1981	0.009 (0.029)	-0.010 (0.007)	0.046* (0.021)	0.037*** (0.003)
1982	0.005 (0.029)	-0.014 (0.009)	0.020 (0.021)	0.011** (0.004)
1983	0.019 (0.030)	–	0.009 (0.021)	–
1984	0.019 (0.030)	-0.001 (0.005)	0.041* (0.021)	0.032*** (0.002)
1985	0.004 (0.032)	-0.015*** (0.004)	0.064** (0.021)	0.055*** (0.001)
1986	0.009 (0.031)	-0.010*** (0.003)	0.084*** (0.022)	0.075*** (0.002)
1987	0.023 (0.031)	0.004 (0.004)	0.102*** (0.022)	0.093*** (0.002)
1988	0.011 (0.034)	-0.008 (0.012)	0.103*** (0.022)	0.095*** (0.003)
1989	0.029 (0.033)	0.010 (0.012)	0.090*** (0.022)	0.081*** (0.004)
1990	0.025 (0.033)	0.006 (0.012)	0.101*** (0.022)	0.092*** (0.004)
1991	-0.008 (0.034)	-0.027* (0.013)	0.048* (0.022)	0.039*** (0.003)

Coefficients with standard errors in parentheses, two-way clustered by plant and year. Industry fixed effects in all columns. Level: $\hat{\mu}_t$; Diff.: $\Delta\hat{\mu}_t = \hat{\mu}_t - \hat{\mu}_{1983}$ (1983 is the reference, –). * p<0.1, ** p<0.05, *** p<0.01. Observations: 39,495.

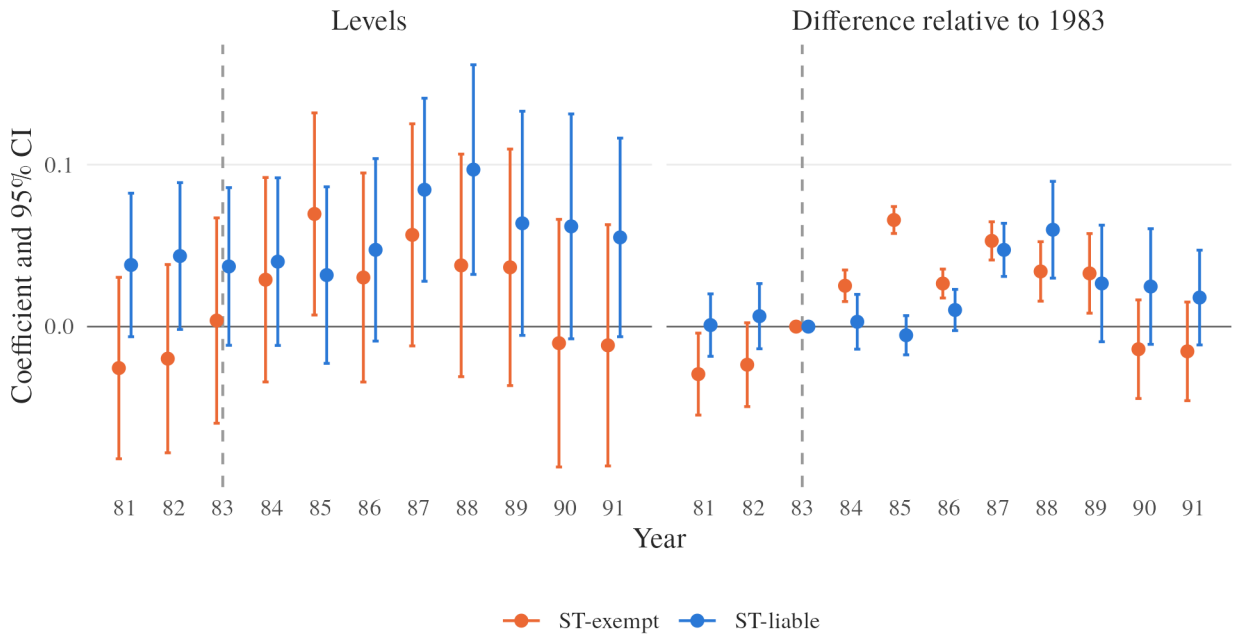


Figure 4: Overreporting of proprietorships relative to corporations by year, 1981–1991, in ST-exempt and ST-liable industries. Points are coefficient estimates with 95 percent confidence intervals; standard errors two-way clustered by plant and year; industry fixed effects. Left panel: levels, $\hat{\mu}_t$; right panel: differences with respect to 1983, $\Delta\hat{\mu}_t$. The dashed line marks 1983, the last full year before the reform took effect.

Table 14: Overreporting of proprietorships in ST-exempt and ST-liable industries. Coefficients behind the figure, by year: levels ($\hat{\mu}_t$) and differences with respect to 1983 ($\Delta\hat{\mu}_t$).

	ST-exempt		ST-liable	
Year	Level	Diff.	Level	Diff.
1981	-0.026 (0.029)	-0.029** (0.013)	0.038 (0.023)	0.001 (0.010)
1982	-0.020 (0.030)	-0.024 (0.013)	0.044* (0.023)	0.006 (0.010)
1983	0.004 (0.032)	–	0.037 (0.025)	–
1984	0.029 (0.032)	0.025*** (0.005)	0.040 (0.026)	0.003 (0.009)
1985	0.070* (0.032)	0.066*** (0.004)	0.032 (0.028)	-0.005 (0.006)
1986	0.030 (0.033)	0.027*** (0.005)	0.047 (0.029)	0.010 (0.006)
1987	0.057 (0.035)	0.053*** (0.006)	0.085** (0.029)	0.047*** (0.008)
1988	0.038 (0.035)	0.034*** (0.009)	0.097** (0.033)	0.060*** (0.015)
1989	0.037 (0.037)	0.033** (0.013)	0.064 (0.035)	0.027 (0.018)
1990	-0.010 (0.039)	-0.014 (0.016)	0.062 (0.035)	0.025 (0.018)
1991	-0.012 (0.038)	-0.015 (0.016)	0.055 (0.031)	0.018 (0.015)

Coefficients with standard errors in parentheses, two-way clustered by plant and year. Industry fixed effects in all columns. Level: $\hat{\mu}_t$; Diff.: $\Delta\hat{\mu}_t = \hat{\mu}_t - \hat{\mu}_{1983}$ (1983 is the reference, –). * p<0.1, ** p<0.05, *** p<0.01. Observations: 39,495.

lead to lower incentives to evade taxes. However, since the reduction is small, it is not unreasonable to find only small changes in tax evasion for these industries. In ST-liable industries, LLCs faced two opposing effects: the increase in sales tax rates, which would increase incentives to evade taxes, and the income-tax cut, which would decrease incentives to evade taxes. The evidence suggests that the increase in sales tax rates dominates the effect of the income-tax cut, leading to a significant increase in tax evasion for LLCs in ST-liable industries.

Proprietorships show a more heterogeneous response (Figure 4, Table 14). In ST-liable industries, their overreporting barely changed in 1984–1986, consistent with the income-tax cut compensating them for much of the sales-tax increase; it rose from 1987, peaking in 1988. In ST-exempt industries, their overreporting rose after 1983, even though they faced only the income-tax cut. This result is unexpected, and I offer two possible explanations. First, the incentive to overreport depends on the marginal tax rate, not the average one, and the 1983 cut in marginal rates was much smaller than the cut in average rates at both ends of the income distribution: close to zero at low incomes and about 2 percentage points at high incomes (Table 10). Second, the 1982–1983 changes stopped individuals from fully deducting their financial expenses (Ocampo and Perry 1983), which gave proprietorships, but not LLCs, a reason to replace a lost deduction with overreported inputs.

Proprietorships' standard errors are also several times those of LLCs, which suggests that they responded less uniformly. This is plausible: proprietorships range from family workshops to young entrepreneurial firms, with varying degrees of process standardization, record-keeping, while LLCs tend to be larger firms with more established processes, and their responses look more homogeneous.

5.5 Overreporting responds more than proportionally to the tax rate

In ST-liable industries, overreporting rose by about 6 percent for each 1 percent rise in the sales tax rate on purchases, τ_P.

Between 1983 and 1987, the average τ_P of unincorporated firms in ST-liable industries rose from 8.2 to 10.0 percent (the median rose from 7.1 to 10.1 percent), and their overreporting rose from 2.7 to 11.3 percent of true materials, about 8.5 percentage points. The midpoint elasticity of overreporting with respect to τ_P is 5.9, with a 95 percent confidence set of 4.2 to 8.5. The set rejects elasticities of one or less ($p < 0.001$).

I use the midpoint elasticity because overreporting was close to zero before the reform. A percentage change measured from a base that close to zero is unstable: the usual elasticity is 25, but its bootstrap interval runs from -260 to 301. The confidence set instead inverts a test of each candidate elasticity, with variance clustered by plant. Section F defines the elasticity and the test, and reports the results for each year from 1984 to 1988.

The elasticity describes the response to the reform as a whole. The same years include the income-tax cut and the gradual phase-in of the response, so it is not the elasticity of overreporting to τ_P with everything else held fixed, which is what the counterfactual in Section 6 measures.

5.6 Takeaways and link to the structural model

Input overreporting rises with the marginal benefit of overreporting. After the 1983 reform raised the sales tax, overreporting by unincorporated firms increased, reaching about 9 percent of true materials by 1987, and the increase was concentrated in ST-liable industries, where the benefit rose. Within those industries it rose most among LLCs, whose small income-tax cut left most of the sales-tax increase in place, while proprietorships, compensated by a larger income-tax cut, barely responded at first. The one result that does not fit, the rise among proprietorships in ST-exempt industries, has two possible explanations in the reform's other provisions (Section 5.4).

This evidence supports the model's key comparative static, $\partial e^* / \partial \tau_P > 0$ (Section 2), which the structural estimation cannot test by regression: overreporting and the tax rate are both built from the same reported materials, so a regression of one on the other is mechanically biased. It also supports the identification of the detection probability in Section 6, which rests on the same source of variation: changes in the sales-tax benefit of overreporting move evasion.

6 The Counterfactual: Sales-Tax Revenue and the Rate on Purchases

How does net sales-tax revenue change as the government changes the sales-tax rate on purchases? As the rate increases, the incentive to claim additional fictitious deductions increases, but more overreporting raises the probability of detection. Therefore, firms face a trade-off. Conversely, a decrease in the rate reduces the incentive to overreport. Firms decrease their overreporting (intensive margin), and some stop overreporting altogether (extensive margin). An increase in the rate also raises firms' use of true materials because it lowers the net-of-tax price of materials relative to output. Hence, firms buy more materials and produce more. Larger output raises gross sales-tax revenue, but more materials raise sales-tax refunds because claimed deductions rise. How does net sales-tax revenue change at the margin for a change in the rate? How much of that change comes from firms overreporting more, and how much from firms buying more materials?

I study these questions using the estimates of the structural objects from the firms' tax evasion model. With the parameters at hand, I construct a partial-equilibrium revenue curve, the expected net sales-tax revenue as a function of the rate on purchases. Net sales-tax revenue is the difference between gross sales-tax revenue, the tax collected on firms' sales, and sales-tax refunds, the credit the government pays on the deductions firms claim for their purchases, true and fictitious.

Since overreporting affects government revenue through claimed deductions, I decompose the response of claimed deductions into two parts: the *materials response* of firms buying more or fewer materials and the *overreporting response* of firms overreporting more or less.

I find that, among the firms in the nine industries suspected of overreporting (Section 3), claimed deductions respond elastically to changes in the rate on purchases, and the response is more elastic to increases than to cuts. A 1 percent increase in the rate raises claimed deductions by about 1.39 percent. The materials response accounts for 0.18 points and the overreporting response for 0.21 points. A 1 percent cut lowers them by about 1.32 percent. The materials response accounts for 0.18 points and the overreporting response for 0.14 points.

Net sales-tax revenue falls as the rate rises. A 1 percent increase in the rate lowers net sales-tax revenue by about 1.7 percent of gross sales-tax revenue. At the current rate, the net sales-tax revenue is negative. A cut of about 17 percent would make net sales-tax revenue break even, and a cut of about 28 percent or more would make net sales-tax revenue significantly positive, at the 5 percent level. Undetected overreporting costs the government 4.7 percent of gross sales-tax revenue from these firms.

6.1 Government Revenue and Claimed Deductions

Given the firm's optimization problem, the government's expected net sales-tax revenue from firm i is,

$$R_i(\Delta; M_i) = \tau_{S,it} P_t Y_{it}(\Delta) - \tilde{\tau}_{P,it}[M_{it}(\Delta) + (1 - q(\tilde{e}_i(\Delta)))\tilde{e}_i(\Delta)] \tag{28}$$

where $\tilde{\tau}_{P,it} \equiv (1 + \Delta)\tau_{P,it}$ is the counterfactual tax rate on purchases. The first term in Equation 28 is gross sales-tax revenue, $\tau_{S,it} P_t Y_{it}(\Delta)$. The second term is claimed deductions: the sales-tax refunds the government pays at the new rate, $\tilde{\tau}_{P,it}$, on the firm's true materials, $M_{it}(\Delta)$, plus the counterfactual undetected overreporting, $(1 - q(\tilde{e}_i(\Delta)))\tilde{e}_i(\Delta)$.

A change in the rate on purchases moves net sales-tax revenue through three channels. First, every peso the firm claims is credited at the new rate. Second, a higher rate raises the marginal benefit of overreporting in the evasion condition (Equation 5), leading the firm to overreport more. Third, the rate on purchases changes the net-of-tax price of materials relative to output (Section 2.3.1). A higher rate yields materials cheaper, therefore the firm buys more materials, claims the credit on more purchases, and produces more output, which raises gross sales-tax revenue. With a higher rate, the first two channels lower net sales-tax revenue. The third lowers net sales-tax revenue through refunds and raises it through gross sales-tax revenue.

The following analysis takes government enforcement effort and tax-evasion penalties as constant. In general, governments can choose the level of enforcement effort, and therefore they can affect the probability of detection $q(\cdot)$. Governments might also collect additional income by penalizing firms caught evading. During the sample period, the Colombian tax administration struggled with inefficiency, as it lacked budget and control over its staff, and it audited little. Evasion penalties were also ineffective. Some penalties were trivial and others too large to be applied (Perry and Cárdenas 1986, 2:143–146).

As a partial-equilibrium analysis, the firm's capital, labour and productivity, the prices of output and materials, the rate on sales, as well as the output measurement error and the overreporting cost shocks are held fixed. True materials and overreporting respond to the new rate. The analysis lets output and gross sales-tax revenue respond to changes in true materials.

6.2 Estimating the model

6.2.1 Setup

Using estimates of the production function parameters $(\beta, \alpha_K, \alpha_L)$ and constructed residuals, $\mathscr{V}_{it}$, $\tilde{\mathscr{W}}_{it}$ from Section 3.6 and Section 4, the objective is to recover the parameters of the probability of detection and the cost of evasion, $(q(e), \kappa_{it}(e, \omega))$.

I assume a concave probability of detection and a cost of evasion that is linear in evasion and convex in productivity (Section 2):

$$q(e_{it}) = \left(e_{it}\lambda_{0,jt}\right)^{\lambda_1}, \qquad 0 < \lambda_1 < 1 \tag{29}$$

$$\kappa_{it}(e_{it}, \omega_{it}) = e_{it}\mathbf{e}^{\delta_0 - \delta_1\omega_{it} + \delta_2\omega_{it}^2 + \psi_{it}} \tag{30}$$

where $\lambda_{0,jt}$ is a scaling parameter and $0 < \lambda_1 < 1$ defines the curvature of the probability of detection. The scaling parameter brings industries consuming different inputs and thus input levels to a relative scale within the industry and comparable across industries at each t. In this way, a common probability of detection across industries can be found. Put differently, $\lambda_{0,jt}$ sets the level of detection and λ_1 its curvature. With $0 < \lambda_1 < 1$ and a cost linear in e, the second-order condition holds for every firm (Section B).

Let $x_{it} \equiv \lambda_{0,jt}e_{it}$. Then $q + q'e = (1 + \lambda_1)x^{\lambda_1}$, and the FOC with respect to tax evasion for evaders (interior solution) becomes,

$$\tau_{P,it}\rho_t(1 - (1 + \lambda_1)x_{it}^{\lambda_1}) = \mathbf{e}^{\delta_0 - \delta_1\omega_{it} + \delta_2\omega_{it}^2 + \psi_{it}}$$

Let

$$h(e_{it}; \lambda_1, \lambda_{0,jt}) \equiv \ln(\tau_{P,it}\rho_t) + \ln\left(1 - (1 + \lambda_1)x_{it}^{\lambda_1}\right), \qquad x_{it} \in (0, c_{\lambda_1}), \quad c_{\lambda_1} = (1 + \lambda_1)^{-1/\lambda_1}$$

Taking logs, the evasion FOC becomes,

$$\psi_{it} = h(e_{it}; \lambda_1, \lambda_{0,jt}) - \delta_0 + \delta_1\omega_{it} - \delta_2\omega_{it}^2 \tag{31}$$

The ceiling c_{λ_1} is the largest relative overreporting the FOC can rationalize; there the detection probability is $1/(1 + \lambda_1)$. The parameters of interest are $\theta = (\lambda_1, \lambda_{0,jt}, \delta_0, \delta_1, \delta_2)$.

6.2.2 Estimation

Using the firm-level constructed variables from Section 3.6 and Section 4, I cast the estimation problem in terms of a single unobservable, true materials M_{it},[14]

$$e(M) = M_{it}^* - M, \qquad \varepsilon(M) = \ln(M_{it}^*/M) - \mathscr{V}_{it}, \qquad \omega(M) = \tilde{\mathscr{W}}_{it} - (1 - \beta)\varepsilon(M)$$

[14]Recasting in terms of M follows the convention of *unobservable* variables in Schennach (2014). Error terms, such as ε and ω in our framework, are not considered unobservable variables.

Therefore, I rewrite Equation 31 as,

$$\psi(M) \;=\; h(e(M)) - \delta_0 + \delta_1 \omega(M) - \delta_2 \omega(M)^2 \tag{32}$$

In addition to depending on a single unobservable, M_{it}, the setup presents two non-standard features. First, it is nonlinear and nonseparable in M through the log term in h. Second, the domain of M depends on the parameters being estimated. The support of the latent variable depends on $(\lambda_1, \lambda_{0,jt})$: the FOC ceiling restricts true materials to $M \in (\max\{0, M^*_{it} - c_{\lambda_1}/\lambda_{0,jt}\}, M^*_{it})$.

The method *Entropic Latent Variable Integration via Simulation* (ELVIS) (Schennach 2014) addresses both nonlinearity and nonseparability, and a parameter-dependent support of the unobserved variable.

Let f_Z denote the density of the observed data Z_{it}, and let $f_{M|Z}$ denote a candidate conditional density of M_{it} given Z_{it}, so that $f_{M|Z}\, f_Z$ is the implied joint density of (M_{it}, Z_{it}).[15] The question is whether some choice of $f_{M|Z}$ makes $E_{f_{M|Z} f_Z}[g(M, Z, \theta)] = 0$.

Theorem 2.1 of Schennach (2014) shows the abstract *existence* problem of $f_{M|Z}$ is equivalent to a more concrete optimization problem:

$$\inf_{f_{M|Z}} \left\| E_{f_{M|Z} f_Z}[g(M, Z, \theta)] \right\| = 0 \quad \Longleftrightarrow \quad \inf_{\gamma \in \mathbb{R}^{d_g}} \left\| E_{f_Z}[\tilde{g}(Z, \theta, \gamma)] \right\| = 0 \tag{33}$$

where

$$\tilde{g}(Z, \theta, \gamma) \equiv \frac{\int g(M, Z, \theta) \exp\left(\gamma' g(M, Z, \theta)\right) f^0_{M|Z}(M|Z)\, dM}{\int \exp\left(\gamma' g(M, Z, \theta)\right) f^0_{M|Z}(M|Z)\, dM} \tag{34}$$

Among the conditional distributions of M given Z that satisfy the moment conditions, ELVIS selects the one with maximum entropy relative to an arbitrary reference density $f^0_{M|Z}$. In the optimization problem, $\gamma \in \mathbb{R}^{d_g}$ is the Lagrange multiplier vector on the moment constraints. $f^0_{M|Z}$ only needs to be positive wherever the true conditional density of M is. Any choice of $f^0_{M|Z}$ with the right support gives the same $\hat{\theta}$ (Remark 2.3, Schennach 2014). In this setting, this allows me to avoid specifying f_ψ or f_ε or searching over the space of all conditional distributions of M. Instead, I solve a finite optimization problem.

The true conditional distribution of M satisfies the moment condition at the true θ_0. Theorem 2.1

[15] Schennach (2014) states the result for general probability measures. I write it with densities because, in this application, M_{it} is continuous on the known support $(\max\{0, M^*_{it} - c_{\lambda_1}/\lambda_{0,jt}\}, M^*_{it})$; every candidate conditional distribution therefore has a density there. The reference measure follows Schennach (2014)'s Proposition 2.1: a density proportional to $\exp(-\|D^{-1}(g(M) - g(M^*))\|^2)$ times the uniform on the support, with fixed row scales D, which keeps the tilted density proper for every γ.

guarantees some γ also solves the right-hand side of Equation 33 at θ_0. Searching over θ for the value where this finite-dimensional problem has a solution therefore recovers θ_0, regardless of the chosen $f^0_{M|Z}$, without requiring knowledge of the true conditional distribution of M.

Although the entropy-maximizing distribution satisfies $E[g(M, Z, \theta_0)] = 0$, it should not be treated as the actual distribution of M, since many conditional distributions of M given Z may satisfy the moment condition (Schennach 2022, 1250). This matters for the counterfactual question. Since expected counterfactual revenue depends on the latent M_{it}, it would be tempting to average it over draws from such a distribution.

Instead, I include the counterfactual object as an auxiliary parameter with its own moment condition and estimate it jointly with θ and γ. The counterfactual question is equivalent to testing if a candidate value of the counterfactual quantity is compatible with the distribution of M conditional on Z such that its auxiliary moment and the rest of the moment conditions are satisfied.

For estimation, I set $\lambda_{0,jt} = 1/(\bar{\lambda}\bar{M}_{j,t-1})$ where $\bar{M}_{j,t-1}$ is the mean reported materials of firm i's industry j in the previous year. The scale makes detection depend on how much a firm overreports relative to a typical firm in its industry, not on the peso amount. $\bar{\lambda}$ sets the level of detection and λ_1 its curvature. With $0 < \lambda_1 < 1$ and a cost linear in e, the second-order condition holds for every firm (Section B).

I use the model's independence assumptions to set up unconditional moment conditions: ε is independent of $(M, e, \omega, \psi, \tau_P)$ and ψ is independent of $(\omega, M, \tau_{P,it})$, with $E[\varepsilon] = E[\psi] = 0$. They imply the following set of moments,

$$g(M, Z, \theta) = \begin{bmatrix} \psi(M, Z) \\ \psi(M, Z) \ln M \\ \psi(M, Z)\omega(M, Z) \\ \psi(M, Z)\omega(M, Z)^2 \\ \varepsilon(M, Z)\psi(M, Z) \\ \varepsilon(M, Z)\frac{\partial h(e)}{\partial \lambda_1} \\ \psi(M, Z) \ln \tau_{P,it} \\ \varepsilon(M, Z)\frac{\partial h(e)}{\partial \bar{\lambda}} \\ \varepsilon_j(M, Z) \end{bmatrix}$$

Note that the cost shock is independent of the tax rate, the benefit shifter, such that variation in the rate is absorbed through evasion. The industry rows set $E[\varepsilon|j] = 0$, such that in each industry the model's mean overreporting matches the data's mean $\mathscr{V}$, $E[u|j] = E[\mathscr{V}_j]$, from Section 3.6. Section A.1 describes how these moments are implemented in estimation.

Table 15: Model's parameters estimates at the operating point,identified from a grid search over $(\lambda_1, \bar{\lambda})$ profiling over $(\delta_0, \delta_1, \delta_2)$ and γ. Interior firms in the nine industries where the test rejects. The test statistic TS_n fails to reject at the 5 percent significance level the null hypothesis that $E[g(M, Z, \hat{\theta})] = 0$.

	Estimate
λ_1	0.75
$\bar{\lambda}$	0.50
$\hat{\delta}_0$	43.13
$\hat{\delta}_1$	27.15
$\hat{\delta}_2$	3.80
$\hat{\omega}^* = \hat{\delta}_1/2\hat{\delta}_2$	3.57
Firm-years (n)	12,050
max $\lvert\hat{\gamma}\rvert$	195.4
TS_n ($R = 1000$)	23.69
TS_n ($R = 4000$)	24.74

R = draws per firm.

6.2.3 Testing and inference

The moments are combined into a single test statistic through the continuously updated estimator (CUE) objective $\hat{L}_n(\theta)$ of Schennach (2014)'s Definition F.1. Her Theorem F.1 shows that, at any θ satisfying the moment condition, $TS_n(\theta) = 2n\hat{L}_n(\theta)$ is asymptotically bounded above by a $\chi^2_{d_g}$ random variable. Comparing $TS_n(\theta)$ with the $\chi^2_{d_g}$ critical value therefore gives a conservative test, whose size is at most its nominal level.

To obtain the confidence set of a counterfactual quantity $\theta_T(\Delta)$ (Section 6.3), I invert the test, as in the counterfactual exercise of Aguiar and Kashaev (2021, App. F). The confidence set collects the candidate values that the conservative test does not reject, $\{\theta_T : TS(\Delta, \theta_T) \leq \chi^2_{d_g,1-\alpha}\}$, and the point estimate is the candidate with the lowest statistic (Section A.4).

I searched over a $(\lambda_1, \bar{\lambda})$ grid, profiling over $(\delta_0, \delta_1, \delta_2)$ and γ (Section A) and found an operating point. The test statistic TS_n fails to reject at the 5 percent significance level the null hypothesis that $E[g(M, X, \hat{\theta})] = 0$. The estimation sample has 12,050 firm-years of unincorporated firms in the nine industries where the test in Section 3 rejects.[16] Details of the implementation, including the exact objective function can be found in Section A. The next section discusses the operating point for the parameters of interest.

6.2.4 Model parameter estimates

Table 15 displays the operating point found by grid search over $(\lambda_1, \bar{\lambda})$. The estimates indicate that the cost function is decreasing below $\hat{\omega}^*$ and convex in productivity, since $\hat{\delta}_1 = 27.15 > 0$ and $\hat{\delta}_2 = 3.80 > 0$. This implies that the cost of evasion falls with productivity up to $\hat{\omega}^* = 3.57$ and rises after it. $\hat{\omega}^*$ lies at about the 57th percentile of productivity in the estimation sample.

The implied mean detection probability is 6.8 percent (95 percent set 5.7 to 11.0 percent). The median is 2.0 percent, 80th percentile 18.5 percent, 90th percentile 27 percent, and maximum 57 percent. The quantiles are forward-simulated at the operating point. The mean is estimated with its own moment.

The model's mean overreporting $E[u]$ is within 0.02 of the data's $E[\mathscr{V}]$ in seven of the nine industries. 37 percent of interior firms overreport by at least 5 percent ($u \geq 0.05$).

The level of detection $\bar{\lambda}$ is identified only through the model. Note that the data come from a firm-level survey and do not contain audit records, but, under the model, they are not needed. The model identifies $\bar{\lambda}$ through the evasion condition (Equation 5). Because the cost shock ψ is independent of the tax rate, the probability of detection must offset changes in the benefit of evading, $\tau_P \rho_t$. If detection is too low to offset them, the changes show up in ψ, and the test rejects that level. Higher values of $\bar{\lambda}$ mean lower detection. The test rejects the higher values I tried and fails to reject only near $\bar{\lambda} = 0.5$.

Having said this, a cautious reader would find comfort knowing that relative detection risk, a firm's detection probability relative to its industry's mean, barely moves across grid points. Relative measures, such as elasticities and shares of gross sales-tax revenue, should depend less on the scaling parameter, $\bar{\lambda}$, than levels, such as the detection probabilities, the claimed deductions and net sales-tax revenue in Section 6.4. Such a reader could therefore lean towards the relative measures.

6.3 Counterfactual design

The objective is to estimate how the counterfactual overreporting response to a change in the tax rate moves net sales-tax revenue through fictitious deductions. I also consider changes in output and claimed deductions on true purchases due to changes in the counterfactual true materials decision of firms. The counterfactual overreporting and the counterfactual true materials, however, depend on the same latent M_i as the moments of Section 6.2.2. Since only M_i^* is observed, these quantities cannot be computed firm by firm from the data.

Instead, I cast each counterfactual as a moment condition and add it to the set of moments in Section 6.2.2. For example, net sales-tax revenue $R_i(\Delta; M_i)$ in Equation 28 becomes the moment $E[R_i(\Delta; M_i) - \theta_R(\Delta)] = 0$, where $\theta_R(\Delta)$ is its mean over the interior firms. In general, for a quantity

[16] The sample is the one of (Section 3.6)

$T_i(\Delta; M_i)$ whose mean is the auxiliary parameter $\theta_T(\Delta)$, I cast it as $E[T_i(\Delta; M_i) - \theta_T(\Delta)] = 0$. The confidence set of $\theta_T(\Delta)$ comes from inverting the test (Section 6.2.3).

Next, I show how the counterfactual overreporting and the counterfactual true materials follow from the model, and how both depend on M_i. Holding the firm's productivity ω_i and cost shock ψ_i fixed, the evasion condition at the new rate gives the counterfactual overreporting in closed form. With $B_i \equiv 1 - (1 + \lambda_1)x_i^{\lambda_1}$ at the baseline, where x_i is the same as in Section 6.2.1,

$$x_i'(\Delta)^{\lambda_1} = \max\left\{0,\ \frac{1 - B_i/(1+\Delta)}{1+\lambda_1}\right\}, \qquad \tilde{e}_i(\Delta) = x_i'(\Delta)/\lambda_{0,jt} \tag{35}$$

At the current rate, every interior firm overreports: $x_i > 0$ and $B_i < 1$. Only a cut in the rate can make a firm stop, when $B_i \geq 1 + \Delta$.

The counterfactual overreporting responds to the rate only through $B_i/(1 + \Delta)$. By the evasion condition, B_i is the firm's marginal cost of evasion relative to the benefit of evading, $\tau_{P,it}\rho_t$. A change in the rate scales the benefit by $1 + \Delta$ but leaves the marginal cost unchanged, because the marginal cost depends only on ω_i and ψ_i, which are held fixed. The scaling parameter $\lambda_{0,jt}$ stays fixed too, as it is a lagged industry mean, which the firm's choice does not move.

The counterfactual overreporting depends on the latent M_i through B_i, which contains current overreporting, $x_i = \lambda_{0,jt}(M_i^* - M_i)$. Each draw of M_i therefore gives its own B_i and its own counterfactual overreporting.

True materials respond through the materials condition (Equation 3). With capital, labour, productivity, prices and the rate on sales held fixed, I solve it for counterfactual and current true materials, $M_{it}(\Delta)$ and M_{it}. The ratio of the two depends only on the rate on purchases:

$$r_{it}(\Delta) \equiv \frac{M_{it}(\Delta)}{M_{it}} = \left(\frac{1 - (1+\Delta)\tau_{P,it}}{1 - \tau_{P,it}}\right)^{-\frac{1}{1-\beta}} \tag{36}$$

The factor $r_{it}(\Delta)$ depends only on $\tau_{P,it}$, Δ and β, not on the firm's true materials. The response grows with β, the output elasticity of materials, through the exponent $1/(1 - \beta)$. Across interior firms, true materials rise on average by 4.8 percent with a 30 percent increase.

The counterfactual true materials depend on the latent M_i through current true materials, which the factor scales: $M_{it}(\Delta) = r_{it}(\Delta)M_{it}$. Each draw of M_i therefore gives its own counterfactual true materials, but the factor is the same for every draw.

I report two summaries. The first traces the revenue curve: net sales-tax revenue on a grid of changes in the rate, Δ, from -30 to $+30$ percent. The second measures the response near the current rate through arc elasticities of net sales-tax revenue and of claimed deductions at $\Delta = \pm 1$, ± 1.5 and ± 2 percent.

6.3.1 Net sales-tax revenue and claimed deductions

Substituting the counterfactual overreporting (Equation 35) and the counterfactual true materials (Equation 36) into Equation 28 gives net sales-tax revenue at the new rate for each draw of M_i:

$$R_i(\Delta; M_i) = r_{it}(\Delta)^{\beta}\, \tau_{S,it} P_t Y_{it} - \tilde{\tau}_{P,it}\big[r_{it}(\Delta) M_{it} + (1 - q(\tilde{e}_i(\Delta)))\tilde{e}_i(\Delta)\big].$$

Its mean is the auxiliary parameter $\theta_R(\Delta)$. Gross sales-tax revenue scales by $r_{it}(\Delta)^{\beta}$ (Section A.4). Claimed deductions split into deductions on true materials and deductions on undetected overreporting.

Totals for the whole economy multiply each interior mean by the number of interior firm-years and add the corner firms.[17] The corner firms fall into two groups: (a) firms that pay no sales tax on their purchases ($\tau_P = 0$; 3,123 firm-years) get no credit at any rate and have no incentive to start overreporting, and (b) firms in the remaining 19 industries (16,154 firm-years). Firms in (b) are treated as truthful, $M = M^*$ and $e = 0$. The test fails to reject correct reporting for them, or their industry is small and they were not tested for overreporting, and the ex-ante selection rule keeps them out of the structural estimation. The claimed deductions of all corner firms are $(1+\Delta)\tau_{P,it} r_{it}(\Delta) M^*_{it}$, and their gross sales-tax revenue scales by $r_{it}(\Delta)^{\beta}$. Neither needs a draw of M_i. Corner firms add gross sales-tax revenue but no loss to undetected overreporting. The reported loss share for the whole economy is therefore a lower bound.

At the current rate, mean net sales-tax revenue from the interior firms is negative, -391 real pesos per firm-year (95 percent set $[-605, -174]$). Even if the interior firms reported truthfully ($M_i = M_i^*$; Section 6.5), mean net sales-tax revenue would be negative, with a 95 percent set of $[-590, -42]$ real pesos per firm-year.

6.3.2 Elasticities

I build each elasticity from a paired difference between the new and the current rate, computed on the same draws of M_i. In this way, noise common to both rates cancels in the difference.

For mean net sales-tax revenue, $R(\Delta)$, I report two elasticities. The first expresses the change as a share of gross sales-tax revenue, $[R(\Delta) - R(0)]/(\Delta\, G(0))$, where $G(0)$ is mean gross sales-tax revenue at the current rate (1,547 real pesos per firm-year). The second is the relative percentage magnitude elasticity, $[R(\Delta) - R(0)]/(\Delta\, |R(0)|)$, with $|R(0)| = 391$ real pesos per firm-year.

For the mean arc elasticity of claimed deductions, I decompose its response into two: the materials response and the overreporting response. With mean claimed deductions per firm-year $C(\Delta) =$

[17] The top 0.5 percent of interior firms by reported materials (61 firm-years), trimmed from the estimation, enter with the same formula.

Table 16: Net sales-tax revenue per firm-year (real pesos) from the interior firms, by change Δ in the rate on purchases, true materials responding: estimates and 95 percent sets. The materials-channel columns give the mean change in true materials and in gross sales-tax revenue.

	Net sales-tax revenue		Materials channel	
Δ	Estimate	95% set	True M	Gross revenue
−30%	264	[49, 480]	−4.4%	−28
−27.5%	214	[2, 426]	−4.0%	−26
−20%	59	[−166, 279]	−2.9%	−19
−10%	−159	[−372, 51]	−1.5%	−10
−5%	−272	[−485, −64]	−0.8%	−5
0	−391	[−605, −174]	0	0
+5%	−528	[−740, −309]	+0.8%	5
+10%	−668	[−882, −468]	+1.5%	10
+20%	−965	[−1,180, −733]	+3.2%	20
+30%	−1,270	[−1,545, −1,013]	+4.8%	31

ELVIS interior firms (12,050 firm-years). Net sales-tax revenue = gross sales-tax revenue minus expected sales-tax refunds. 95% set: conservative test, $TS \leq \chi^2_{18,.95} = 28.87$. True materials respond: the materials first-order condition with the two tax rates, capital, labour and productivity fixed. Gross revenue: mean change in gross sales-tax revenue from the output response.

$$E\big[\tilde{\tau}_{P,it}\big(r_{it}(\Delta)M_{it} + (1 - q(\tilde{e}_i(\Delta)))\tilde{e}_i(\Delta)\big)\big],$$

$$\frac{C(\Delta) - C(0)}{\Delta\, C(0)} = 1 + \frac{E\big[\tilde{\tau}_{P,it}(r_{it}(\Delta) - 1)M_{it}\big]}{\Delta\, C(0)} + \frac{E\big[\tilde{\tau}_{P,it}\big((1 - q(\tilde{e}_i(\Delta)))\tilde{e}_i(\Delta) - (1 - q(e_i))e_i\big)\big]}{\Delta\, C(0)},$$

where the second term on the right-hand side is the materials response and the third, the overreporting response. The mechanical effect, 1, applies the new rate to unchanged claims. I estimate the materials response and the overreporting response each with its own moment and auxiliary parameter.

6.4 Results

6.4.1 Net sales-tax revenue and claimed deductions

A 30 percent cut in the rate on purchases would turn net sales-tax revenue from negative to positive (Figure 5, Table 16). Net sales-tax revenue rises steadily as the rate falls. The estimate breaks even at a cut of about 17 percent, and for cuts of about 28 percent or more the 95 percent set lies entirely above zero ($[2, 426]$ at -27.5 percent). Raising the rate pushes net sales-tax revenue further below zero. A 20 percent increase lowers it from -391 to -965 real pesos per firm-year, and the 95 percent sets of the two levels do not overlap.

For the whole economy, a cut in the rate also turns net sales-tax revenue positive, but the cut must be larger (Figure 6). Summed over every firm-year in the sample, net sales-tax revenue is -15.9 million real pesos at the current rate (95 percent set $[-18.4, -13.2]$). The estimate breaks even at a

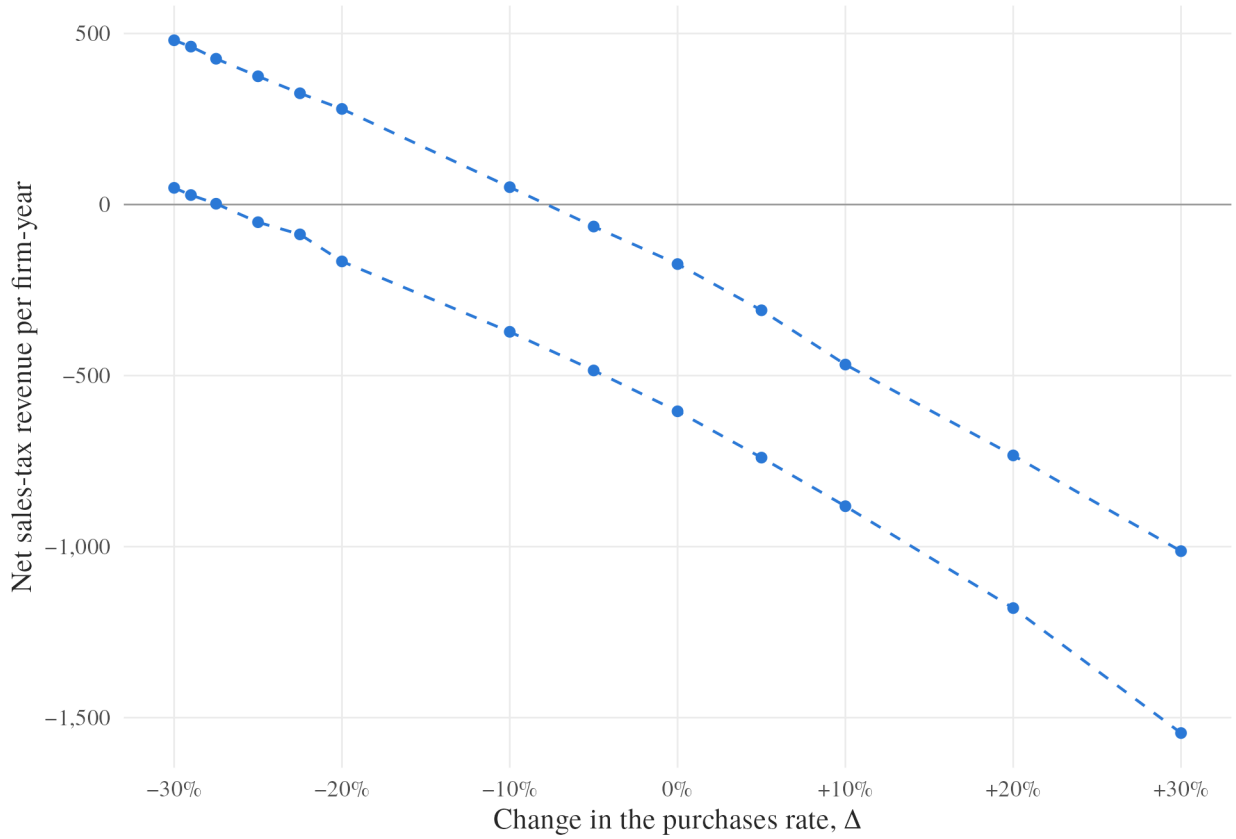


Figure 5: 95 percent sets (conservative test) for net sales-tax revenue per firm-year (real pesos) from the interior firms, by change Δ in the rate on purchases, true materials responding. Dashed lines join the lower bounds and the upper bounds of the sets at each Δ evaluated.

cut of about 24 percent, and a 30 percent cut raises it to 3.6 million ([1.0, 6.2]). The whole economy needs a larger cut than the interior firms because of the corner firms in the other 19 industries. Their claimed deductions, 32.1 million real pesos, exceed their gross sales-tax revenue by 9.4 million, more than half of the negative total. These firms do not overreport: their claims respond to the rate only mechanically and through materials.

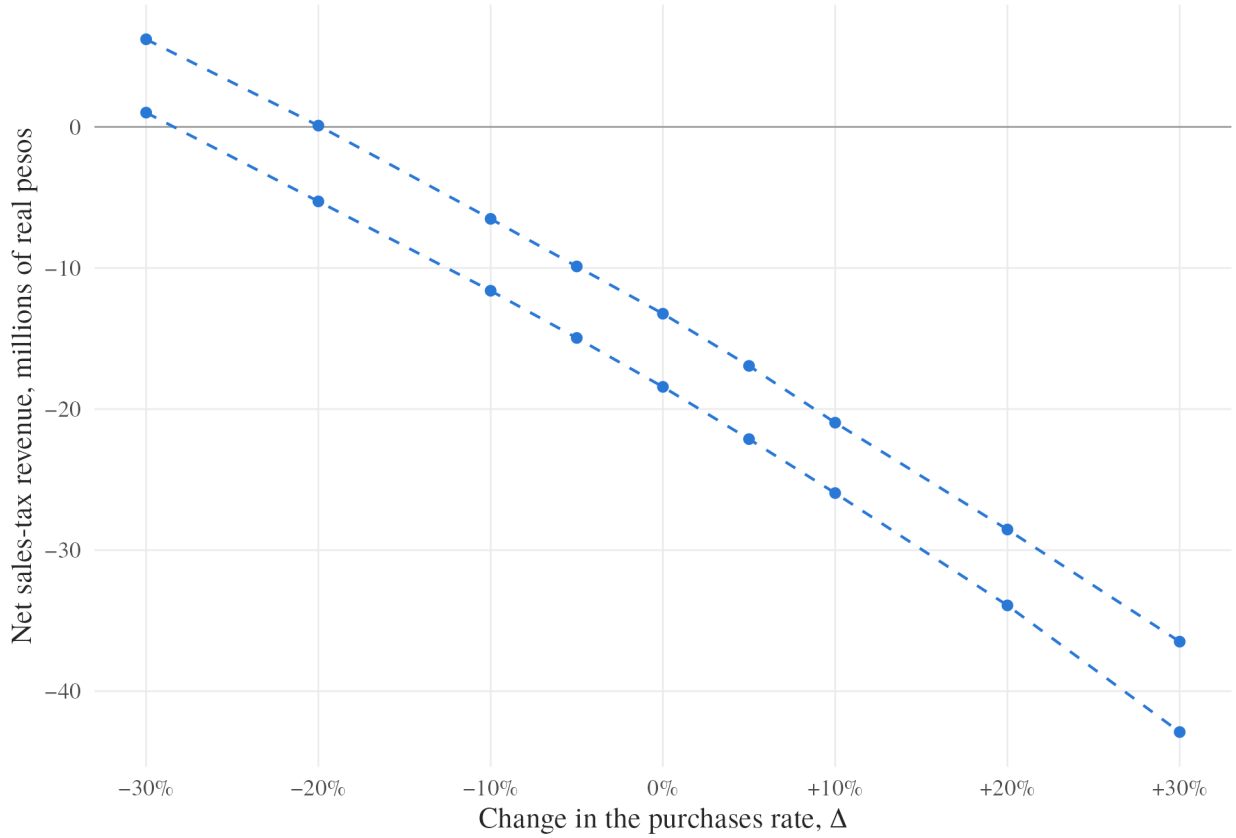


Figure 6: Net sales-tax revenue in the whole economy, total over every firm-year in the sample, 1981–1991 (millions of real pesos), by change Δ in the rate on purchases, true materials responding. The total adds the interior firms, the trimmed firms and the corner firms. Dashed lines join the lower bounds and the upper bounds of the 95 percent sets (conservative test) at each Δ evaluated; the sets carry only the uncertainty about the interior firms.

6.4.2 Elasticities

Near the current rate, a 1 percent increase in the rate lowers net sales-tax revenue by about 1.7 percent of gross sales-tax revenue, or 26 real pesos per firm-year (Table 17). The elasticity is -1.68 at +1 percent, with a 95 percent set of $[-1.91, -1.47]$, and -1.60 at -1 percent ($[-1.81, -1.38]$). The relative percentage magnitude elasticity is -6.63 at +1 percent ($[-7.53, -5.81]$) and -6.32 at -1 percent ($[-7.16, -5.48]$).

Claimed purchase deductions respond more than one for one to the rate on purchases, because firms

buy more materials and overreport more when the rate rises (Table 17, Figure 7). The elasticity of claimed deductions with respect to τ_P is 1.39 at +1 percent. Of this elasticity, 1 is mechanical: the same claimed purchases earn a 1 percent higher credit. The materials response adds 0.18 (95 percent set $[0.15, 0.21]$) and the overreporting response adds 0.21 ($[0.19, 0.23]$). At -1 percent the elasticity is 1.32: the materials response is again 0.18 ($[0.15, 0.21]$), but the overreporting response falls to 0.14 ($[0.12, 0.16]$).

A back-of-envelope calculation shows that, of the 26 real pesos per firm-year that a 1 percent increase withdraws from net sales-tax revenue, the materials response accounts for 10 percent and the overreporting response for 16 percent. The materials share is net of the extra gross sales-tax revenue: buying more materials raises output and the sales tax on it, but the extra credit on those materials is about 3.6 times larger. For a 1 percent cut, the shares are 79 for mechanical, 10 for materials, and 11 percent for overreporting response.

The overreporting response is asymmetric, larger for increases than for cuts. The asymmetry comes from the incentives generated by the tax, which creates a floor on overreporting, $e_i \geq 0$. A firm never underreports, because fewer deductions would increase its tax liability and decrease its after-tax profits. This matters for firms at the margin of overreporting, firms that would jump in or out of evasion depending on the rate. These firms are not trivial. In the estimated distribution of overreporting, about a third of the mass lies below a detection probability of 0.6 percent. These firms play the role of the extensive margin. A 1 percent increase raises their overreporting from almost nothing, and a 1 percent cut drives them out of evasion. They account for about 86 percent of the gap between the responses to increases and to cuts at ± 1 percent. Firms operating at the intensive margin, firms that keep overreporting and only adjust how much, respond almost equally in both directions. This is the case for firms with a detection probability of 2 percent or more.

6.5 Revenue lost to undetected overreporting

The same machinery measures what undetected overreporting costs at the current rate. The counterfactual is potential net sales-tax revenue. Potential net sales-tax revenue is the revenue the government would collect without overreporting, $P_i = \tau_{S,it} P_t Y_{it} - \tau_{P,it} M_{it}$. The loss is the credit paid on undetected overreporting, $L_i = P_i - R_i(0; M_i) = \tau_{P,it}(1 - q(e_i))e_i$.

I measure the loss as a percentage of gross sales-tax revenue, $G(0)$. The ratio of the mean loss to mean gross sales-tax revenue is cast as a moment with its own auxiliary parameter. The base of the ratio, gross sales-tax revenue, is observed and positive, whereas potential revenue can be negative and true credits are latent.

Undetected overreporting costs the government 4.7 percent of gross sales-tax revenue from the interior firms (95 percent set $[2.8, 7.1]$) and at least 1.5 percent ($[1.1, 3.1]$) across all firms in the sample (Table 18), a lower bound for the reasons in Section 6.3.1. The order of magnitude matches

Table 17: Arc elasticities of claimed purchase deductions and of net sales-tax revenue with respect to τ_P, at $\Delta = \pm 1, \pm 1.5, \pm 2$ percent, true materials responding. Claimed deductions are decomposed into the materials response and the overreporting response. 95 percent conservative-test sets.

	Claimed deductions			Net sales-tax revenue	
Δ	Total	Materials	Overreporting	Share of gross revenue	Relative percentage magnitude
−2%	1.29	0.17 [0.15, 0.20]	0.12 [0.10, 0.14]	−1.56 [−1.77, −1.38]	−6.19 [−7.02, −5.45]
−1.5%	1.31	0.18 [0.15, 0.20]	0.13 [0.10, 0.16]	−1.58 [−1.77, −1.39]	−6.23 [−6.98, −5.49]
−1%	1.32	0.18 [0.15, 0.21]	0.14 [0.12, 0.16]	−1.60 [−1.81, −1.38]	−6.32 [−7.16, −5.48]
+1%	1.39	0.18 [0.15, 0.21]	0.21 [0.19, 0.23]	−1.68 [−1.91, −1.47]	−6.63 [−7.53, −5.81]
+1.5%	1.40	0.18 [0.16, 0.21]	0.22 [0.20, 0.24]	−1.70 [−1.90, −1.50]	−6.71 [−7.51, −5.93]
+2%	1.41	0.18 [0.15, 0.22]	0.23 [0.21, 0.24]	−1.71 [−1.93, −1.50]	−6.75 [−7.61, −5.94]

Arc elasticities with respect to the purchases rate, true materials responding, ELVIS interior firms; conservative 95% sets in brackets ($\chi^2_{18,.95} = 28.87$). Claimed deductions: $[C(\Delta) - C(0)]/(\Delta C(0)) = 1$ (mechanical) + materials response + overreporting response; the total is the sum of the point estimates. Net sales-tax revenue: share of gross revenue, $[R(\Delta) - R(0)]/(\Delta G(0))$, with mean gross sales-tax revenue $G(0) = 1{,}547$ real pesos per firm-year; relative percentage magnitude, $[R(\Delta) - R(0)]/(\Delta |R(0)|)$, with $|R(0)| = 391$.

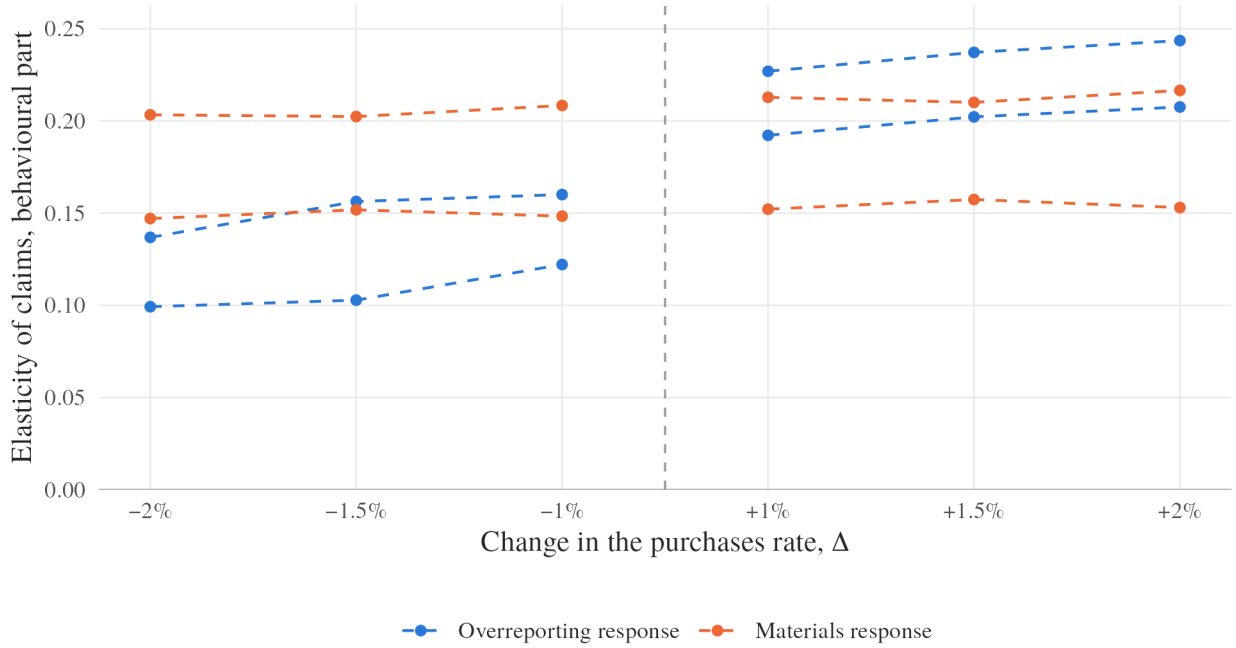


Figure 7: Behavioural part of the elasticity of claimed purchase deductions with respect to τ_P, by change Δ in the purchases rate: overreporting response (firms overreport more or less) and materials response (firms buy more or fewer materials), 95 percent conservative-test sets, true materials responding, interior firms. Dashed lines join the lower bounds and the upper bounds of the sets, separately for cuts and increases; the vertical line separates cuts from increases.

Table 18: Revenue lost to undetected overreporting at the current rate, as a share of gross sales-tax revenue: interior firms and all firms in the sample. 95 percent conservative-test sets.

	Estimate	95% set
Interior firms	4.7%	[2.8%, 7.1%]
All firms in the sample	1.5%	[1.1%, 3.1%]

Loss $L = \tau_P(1 - q(e))e$: the credit paid on undetected overreporting. Each share is the mean loss over mean gross sales-tax revenue. Interior firms: mean gross sales-tax revenue 1,547 real pesos per firm-year. All firms in the sample: every other firm's gross sales-tax revenue is added with a loss of zero, which makes the share a lower bound. $TS \leq \chi^2_{18,.95} = 28.87$.

what tax authorities have detected from input overreporting, 1.2 to 2.7 percent of total tax revenue in five countries (OECD 2017; Carrillo et al. 2022; CASE and IHS 2018; Longinotti 2022).

6.6 Takeaways

Using a structural approach, I show that, among the industries suspected of overreporting, this type of tax evasion costs 4.7 percent of potential gross sales-tax revenue. The rate elasticity of net sales-tax revenue, measured as percentage of gross sales-tax revenue, is about -1.7, and, measured as the relative percentage magnitude, -6.6.

7 Conclusion

This paper asks whether firms overreport deductible inputs more when the tax rate rises, and what their response does to sales-tax revenue. Answering these questions requires dealing with true inputs, which are unobserved: for a given output, a firm that reports high input use may be overreporting, or it may simply be less productive.

Firms overreport more when the rate rises. In Colombian manufacturing between 1981 and 1991, unincorporated firms overreported materials in about half of the industries liable for the sales tax and in neither exempt industry, and they overreported more after a reform raised the rate.

Firms' overreporting response makes raising the rate on purchases costlier for net sales-tax revenue. In the structural model, a higher rate on purchases raises claimed deductions by more than its mechanical effect, because firms both overreport more and buy more materials. Net sales-tax revenue therefore falls by more than the higher credit alone implies. At the current rate, undetected overreporting already costs the government a non-trivial part of the gross sales-tax revenue from these firms.

I propose a method to separate overreporting from productivity with production data alone. Because the method needs only reported inputs, output and a group of firms that report truthfully, it applies to firm surveys, administrative records and confidential tax records alike. The method also yields estimates of the production function and productivity that are corrected for overreporting.

For tax policy, the findings imply that a change in the rate on purchases moves net sales-tax revenue by more than its mechanical effect, and that the evasion response can be measured apart from the response of true purchases. For productivity measurement, the findings imply that estimates built on reported inputs should account for overreporting before they are used to compare firms.

References

Ackerberg, Daniel A, Kevin Caves, and Garth Frazer. 2015. "Identification Properties of Recent Production Function Estimators." *Econometrica* 83 (6): 2411–51. https://doi.org/10.3982/ECTA13408.

Aguiar, Victor H., and Nail Kashaev. 2021. "Stochastic Revealed Preferences with Measurement Error." *Review of Economic Studies* 88 (4): 2042–93. https://doi.org/10.1093/restud/rdaa067.

Allingham, Michael G., and Agnar Sandmo. 1972. "Income Tax Evasion: A Theoretical Analysis." *Journal of Public Economics* 1 (3-4): 323–38. https://doi.org/10.1016/0047-2727(72)90010-2.

Almunia, Miguel, and David Lopez-Rodriguez. 2018. "Under the Radar: The Effects of Monitoring Firms on Tax Compliance †." *American Economic Journal: Economic Policy* 10 (1): 1–38. https://doi.org/10.1257/pol.20160229.

Anderson, T. W., and Herman Rubin. 1949. "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations." *Annals of Mathematical Statistics* 20 (1): 46–63. https://doi.org/10.1214/aoms/1177730090.

Asatryan, Zareh, and Andreas Peichl. 2017. *Responses of Firms to Tax, Administrative and Accounting Rules: Evidence from Armenia*. CESifo Working Paper No. 6754. Center for Economic Studies and ifo Institute (CESifo).

Best, Michael Carlos, Anne Brockmeyer, Henrik Jacobsen Kleven, Johannes Spinnewijn, and Mazhar Waseem. 2015. "Production Versus Revenue Efficiency with Limited Tax Capacity: Theory and Evidence from Pakistan." *Journal of Political Economy* 123 (6): 1311–55. https://doi.org/10.1086/683849.

Bolaños Bolaños, Lucía del Carmen. 2019. "El impuesto de renta en Colombia: una mirada desde el gasto tributario." *Boletín Mexicano de Derecho Comparado* 52 (155): 721–56. https://doi.org/10.22201/iij.24484873e.2019.155.14946.

Carrillo, Paul, Dave Donaldson, Dina Pomeranz, and Monica Singhal. 2022. *Ghosting the Tax Authority: Fake Firms and Tax Fraud*. NBER. http://www.nber.org/papers/w30242.

Carrillo, Paul, Dina Pomeranz, and Monica Singhal. 2017. "Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement." *American Economic Journal: Applied Economics* 9 (2): 144–64. https://doi.org/10.1257/app.20140495.

CASE, and IHS. 2018. *Study and Reports on the VAT Gap in the EU-28 Member States: 2018 Final Report*. TAXUD/2015/CC/131. European Commission, Directorate General Taxation and Customs Union.

Casella, George, and Roger L. Berger. 2002. *Statistical Inference*. 2nd ed. Duxbury.

Chetty, Raj. 2009. "Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance." *American Economic Journal: Economic Policy* 1 (2): 31–52. https://doi.org/10.1257/pol.1.2.31.

Collard-Wexler, Allan, and Jan De Loecker. 2020. *Production Function Estimation and Capital Measurement Error*. NBER. http://www.nber.org/papers/w22437.

Congreso de la República de Colombia. 1990. *Ley 49 de 1990 (Diciembre 28)*. Diario Oficial No. 36.615, 30 de diciembre de 1990. Art. 26: general sales-tax rate of 12% from 1 January 1991. https://normograma.dian.gov.co/dian/compilacion/docs/ley_0049_1990.htm.

Congreso de la República de Colombia. 1995. *Ley 222 de 1995*. Diario Oficial No. 42.156, 20 de diciembre de 1995. https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=6739.

Corte Constitucional de Colombia. 2016. *Sentencia C-633 de 2016*. https://www.corteconstitucional.gov.co/relatoria/2016/C-633-16.htm.

Cremer, Helmuth, and Firouz Gahvari. 1993. "Tax Evasion and Optimal Commodity Taxation." *Journal of Public Economics* 50 (2): 261–75. https://doi.org/10.1016/0047-2727(93)90052-U.

Dabla-Norris, Era, Mark Gradstein, Fedor Miryugin, and Florian Misch. 2019. *Productivity and Tax Evasion*. CESifo. www.RePEc.org.

Das, Sanghamitra, Mark J. Roberts, and James R. Tybout. 2007. "Market Entry Costs, Producer Heterogeneity, and Export Dynamics." *Econometrica* 75 (3): 837–73.

Departamento Administrativo Nacional de Estadística (DANE). 1995. *Metodología Encuesta Anual Manufacturera – EAM*. DANE, Dirección de Metodología y Producción Estadística.

Departamento Administrativo Nacional de Estadística (DANE). 2018. *Colombia – Encuesta Anual Manufacturera – EAM – 1992 a 1994: Documentación DDI*. Archivo Nacional de Datos (ANDA), study DANE-DIMPE-EAM-1992-A-1994. http://microdatos.dane.gov.co/index.php.

Doraszelski, Ulrich, and Jordi Jaumandreu. 2013. "R&D and Productivity: Estimating Endogenous Productivity." *The Review of Economic Studies* 80 (4): 1338–83. https://doi.org/10.1093/restud/rdt011.

Eilers, Paul H. C., and Brian D. Marx. 1996. "Flexible Smoothing with b-Splines and Penalties." *Statistical Science* 11 (2): 89–121. https://doi.org/10.1214/ss/1038425655.

Eslava, Marcela, John Haltiwanger, Adriana Kugler, and Maurice Kugler. 2004. "The Effects of Structural Reforms on Productivity and Profitability Enhancing Reallocation: Evidence from Colombia." *Journal of Development Economics* 75 (2): 333–71. https://doi.org/10.1016/J.JDEVECO.2004.06.002.

Feldstein, Martin. 1999. "Tax Avoidance and the Deadweight Loss of the Income Tax." *Review of Economics and Statistics* 81 (4): 674–80. https://doi.org/10.1162/003465399558391.

Fisman, Raymond, and Shang-Jin Wei. 2004. "Tax Rates and Tax Evasion: Evidence from 'Missing Imports' in China." *Journal of Political Economy* 112 (2): 471–96. https://doi.org/10.1086/381476.

Gandhi, Amit, Salvador Navarro, and David A Rivers. 2020. "On the Identification of Gross Output Production Functions." *Journal of Political Economy* 128 (8): 2973–3016. https://doi.org/10.1086/707736.

Gorodnichenko, Yuriy, Jorge Martinez-Vazquez, and Klara Sabirianova Peter. 2009. "Myth and Reality of Flat Tax Reform: Micro Estimates of Tax Evasion Response and Welfare Effects in Russia." *Journal of Political Economy* 117 (3): 504–54. https://doi.org/10.1086/599760.

Griliches, Zvi, and Jacques Mairesse. 1995. "Production Functions: The Search for Identification." National Bureau of Economic Research, March.

Handley, Kyle, and Timothy J. Moore. 2017. "Tariffs, Transport Costs, and Tax Evasion." Unpublished manuscript.

Hanlon, Michelle, and Joel Slemrod. 2009. "What Does Tax Aggressiveness Signal? Evidence from Stock Price Reactions to News about Tax Shelter Involvement." *Journal of Public Economics* 93 (1-2): 126–41. https://doi.org/10.1016/j.jpubeco.2008.09.004.

Hogg, Robert V., Joseph W. McKean, and Allen T. Craig. 2019. *Introduction to Mathematical Statistics*. Eigth. Pearson.

Joint Tax Program of the Organization of American States and the Inter-American Development Bank. 1965. *Fiscal Survey of Colombia*. Johns Hopkins Press.

Junguito, Roberto, and Hernán Rincón. 2004. "La política fiscal en el siglo XX en Colombia: una visión global." *Coyuntura Económica*, December. https://www.repository.fedesarrollo.org.co/handle/11445/942.

Kang, Yicheng, and Peihua Qiu. 2021. "Nonparametric Deconvolution by Fourier Transformation and Other Related Approaches." In *Handbook of Measurement Error Models*, edited by Grace Y. Yi, Aurore Delaigle, and Paul Gustafson. Chapman and Hall/CRC. https://doi.org/10.1201/9781315101279-11.

Levinsohn, James, and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *Review of Economic Studies* 70 (2): 317–41. https://doi.org/10.1111/1467-937X.00246.

Longinotti, Fernando Peláez. 2022. *Efficiency and Tax Gap in Latin America and the Caribbean: Value Added Tax and Corporate Income Tax*. CIAT Working Paper No. 1. Inter-American Center of Tax Administrations (CIAT).

McLure, Charles E., Jr. 1989. "Analysis and Reform of the Colombian Tax System." In *Tax Reform in Developing Countries*, edited by Malcolm Gillis. Duke University Press. https://pdf.usaid.gov/pdf/_docs/pnabp729.pdf.

Meister, Alexander. 2009. *Deconvolution Problems in Nonparametric Statistics*. Vol. 193. Lecture Notes in Statistics. Springer. https://doi.org/10.1007/978-3-540-87557-4.

Ocampo, José Antonio, and Guillermo Perry. 1983. "La reforma fiscal, 1982 - 1983." *Coyuntura Económica* 13 (1): 215–64. https://ideas.repec.org/a/col/000438/014004.html.

OECD. 2017. *Technology Tools to Tackle Tax Evasion and Tax Fraud*. OECD. www.oecd.org.

Perry, Guillermo, and Mauricio Cárdenas. 1986. *Diez años de reformas tributarias en Colombia*. Fedesarrollo.

Perry, Guillermo, and Alba Lucía Orozco de Triana. 1990. "The VAT in Colombia." In *Value Added Taxation in Developing Countries*, edited by Malcolm Gillis, Carl S. Shoup, and Gerardo P. Sicat. The World Bank.

Pissarides, Christopher A., and Guglielmo Weber. 1989. "An Expenditure-Based Estimate of Britain's Black Economy." *Journal of Public Economics* 39 (1): 17–32. https://doi.org/10.1016/0047-2727(89)90052-2.

Pomeranz, Dina. 2015. "No Taxation Without Information: Deterrence and Self-Enforcement in the Value Added Tax." *American Economic Review* 105 (8): 2539–69. https://doi.org/10.1257/aer.20130393.

Presidencia de la República de Colombia. 1983. *Decreto 3541 de 1983 (Diciembre 29)*. Diario Oficial No. 36.452, 18 de enero de 1984. Issued under the extraordinary powers of art. 53 of Ley 9 de 1983. https://normograma.dian.gov.co/dian/compilacion/docs/decreto_3541_1983.htm.

Roberts, Mark J., and James R. Tybout. 1997. "The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs." *American Economic Review* 87 (4): 545–64.

Saez, Emmanuel, Joel Slemrod, and Seth H. Giertz. 2012. "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." *Journal of Economic Literature* 50 (1): 3–50. https://doi.org/10.1257/jel.50.1.3.

Sánchez Torres, Fabio, and Catalina Gutiérrez Sourdis. 1994. "Reformas tributarias en Colombia 1980–1992. Aspectos de equidad, eficiencia, y simplificación administrativa." *Coyuntura Económica* 24 (1): 97–110. http://hdl.handle.net/11445/2242.

Schennach, Susane M. 2014. "Entropic Latent Variable Integration via Simulation." *Econometrica* 82 (1): 345–85. https://doi.org/10.3982/ecta9748.

Schennach, Susanne M. 2022. "Measurement Systems." *Journal of Economic Literature* 60 (4): 1223–63. https://doi.org/10.1257/jel.20211355.

Slemrod, Joel. 2019. "Tax Compliance and Enforcement." *Journal of Economic Literature* 57 (4): 904–54. https://doi.org/10.1257/JEL.20181437.

Slemrod, Joel, and Wojciech Kopczuk. 2002. "The Optimal Elasticity of Taxable Income." *Journal of Public Economics* 84 (1): 91–112. https://doi.org/10.1016/S0047-2727(01)00095-0.

Stock, James H., and Jonathan H. Wright. 2000. "GMM with Weak Identification." *Econometrica* 68 (5): 1055–96. https://doi.org/10.1111/1468-0262.00151.

Zumaya, Martin, Rita Guerrero, Eduardo Islas, et al. 2021. *Identifying Tax Evasion in Mexico with Tools from Network Science and Machine Learning*. April. http://arxiv.org/abs/2104.13353.

Online Appendix

A ELVIS Derivation and Implementation

A.1 Objective and implementation

The moments in Section 6 follow from the model's independence assumptions. This section describes how they enter the estimator, the objective and its test, and the computation of the tilted moments.

Let $\hat{g}(\theta, \gamma) = n^{-1} \sum_{i=1}^{n} \tilde{g}(Z_i, \theta, \gamma)$ be the sample average of the tilted moment in Equation 34, and let $\hat{V}(\theta, \gamma)$ be the covariance of the tilted moments, clustered by plant. Following Schennach (2014)'s Definition F.1, the objective is the continuously updated quadratic form

$$\hat{L}_n(\theta) = \min_{\gamma \in \mathbb{R}^{d_g}} \frac{1}{2} \hat{g}(\theta, \gamma)' \, \hat{V}(\theta, \gamma)^{-} \, \hat{g}(\theta, \gamma),$$

where $\hat{V}^{-}$ is a generalized inverse. Schennach (2014) writes the objective as the supremum over γ of the negative of this quadratic form; minimizing the nonnegative form gives the same statistic. The generalized inverse keeps every eigen-direction of $\hat{V}$ with a positive eigenvalue, following Aguiar and Kashaev (2021). The multiplier γ is unbounded. Schennach (2014) and Aguiar and Kashaev (2021) both treat a diverging γ as an expected solution, not a numerical failure.

Schennach (2014)'s Theorem F.1 shows that, at any θ that satisfies the moment condition, $TS_n(\theta) = 2n\hat{L}_n(\theta)$ is asymptotically bounded above by a $\chi^2_{d_g}$ random variable. The test compares TS_n with the $\chi^2_{d_g}$ critical value directly, with no credit for the parameters profiled out or for γ. Because the counterfactual supports a policy claim, I use this conservative test (as in Aguiar and Kashaev 2021), whose asymptotic size is at most its nominal level. The seventeen moments are the eight rows of Section 6 plus one ε row for each of the nine industries, and the 95 percent critical value is $\chi^2_{17,.95} = 27.59$.

The integrals in Equation 34 have no closed form. I compute them by self-normalized importance sampling. For each firm, 1,000 draws of M on its support come from a mixture of three proposals:

uniform in M, log-uniform in M, and log-uniform in the distance to the support's lower edge. The draws are reweighted to the reference density of Section 6, whose row scales D are computed once and held fixed across every compared fit. The draws are also fixed across fits, which makes the objective smooth in (θ, γ) (Schennach 2014, 360).

For numerical stability, two rows are implemented as follows. First, the scores of h diverge at the FOC ceiling. With $x = \lambda_{0,jt} e$ and $B = 1 - (1 + \lambda_1) x^{\lambda_1}$, the scores are

$$\frac{\partial h}{\partial \lambda_1} = -\frac{x^{\lambda_1}\left[1 + (1 + \lambda_1)\ln x\right]}{B}, \qquad \frac{\partial h}{\partial \bar{\lambda}} = \frac{\lambda_1 (1 + \lambda_1)\, x^{\lambda_1}}{\bar{\lambda}\, B},$$

and both go to infinity as $B \to 0$, where a single extreme draw can dominate the weighting matrix. The rows therefore use the softsign transformation $s/(1 + |s|)$, bounded in $(-1, 1)$, with B floored at 10^{-6}. The transformed rows remain valid moments, because any function of e alone is independent of ε. I use the softsign for two reasons: the score in λ_1 changes sign, and the softsign decays polynomially, which keeps firms distinguishable across the range of e, whereas exp and tanh saturate. Second, the row for the independence of ψ and the purchases rate interacts ψ with $\ln \bar{\tau}_{P,jt}$, the log of the leave-one-out industry–year mean purchases rate, rather than with the firm's own rate.

At each grid point of the detection parameters (Section A.3), I estimate the cost parameters by Nelder–Mead, in two passes. Every grid point starts from the same validated point and the draws are common within a fit. Before the search begins, γ is solved for at the starting point.

A.2 Sample

I drop the top 0.5 percent of interior firm-years by reported materials: 61 firm-years, 22 of them in Textiles (321). These firms could not be rationalized by the model during estimation. Either these firms are extremely aggressive in evasion or there was an additional measurement error which the model did not include. The largest firms in Textiles report between 7 and 13 times their industry's lagged mean of reported materials, $\bar{M}_{j,t-1}$. At the operating point, the FOC ceiling caps their overreporting at about $0.24\,\bar{M}_{j,t-1}$, or $u \lesssim 0.02$, while their observed residual $\mathscr{V}_{it}$ is about 0.27. The model must then attribute almost all of their residual to measurement error in output. The trimmed firms enter the counterfactual without a model of their evasion: their claimed deductions respond to the rate only mechanically and through materials, $(1 + \Delta)\,\tau_{P,it}\, r_{it}(\Delta)\, M^{*}_{it}$.

A.3 Search over the detection parameters

The detection parameters $(\lambda_1, \bar{\lambda})$ are fixed on a grid, and at each grid point the cost parameters $(\delta_0, \delta_1, \delta_2)$ and γ are estimated (Figure 8). On a coarse grid, $\lambda_1 \in \{0.65, 0.70, 0.75\}$ and $\bar{\lambda} \in \{0.5, 4.5, 9\}$, only $(0.75, 0.5)$ passes the conservative test, with $TS = 23.7$; the next best point, $(0.70, 9)$, reaches 29.5. A fine grid around it, $\lambda_1 \in \{0.725, 0.75, 0.775\}$ and $\bar{\lambda} \in \{0.4, 0.5, 0.6\}$,

finds a second passing point, $(0.775, 0.5)$, with $TS = 22.0$ at 1,000 draws per firm and 23.8 at 4,000, against 23.7 and 24.7 at the operating point. At this second point, the generalized inverse keeps all 17 eigen-directions of $\hat{V}$, $(\hat{\delta}_0, \hat{\delta}_1, \hat{\delta}_2) = (42.2, 27.1, 3.87)$ and $\hat{\omega}^* = 3.50$. The difference in TS lies within the noise between refits, and the fit is less settled: the effective sample size per firm is smaller and γ is larger. I therefore keep $(0.75, 0.5)$ as the operating point. The passing region is a set as with all modeld with latent variables (Schennach 2014).

The the level of detection is identified under the model with $\bar{\lambda} \leq 1$. Larger values will interact with λ_1 however larger values would be difficult to justify. Relative detection risk on the other hand is nearly free of $\bar{\lambda}$: across industries, the relative risks at $\bar{\lambda} = 0.5$ and at $\bar{\lambda} = 9$ have a correlation of 0.87.

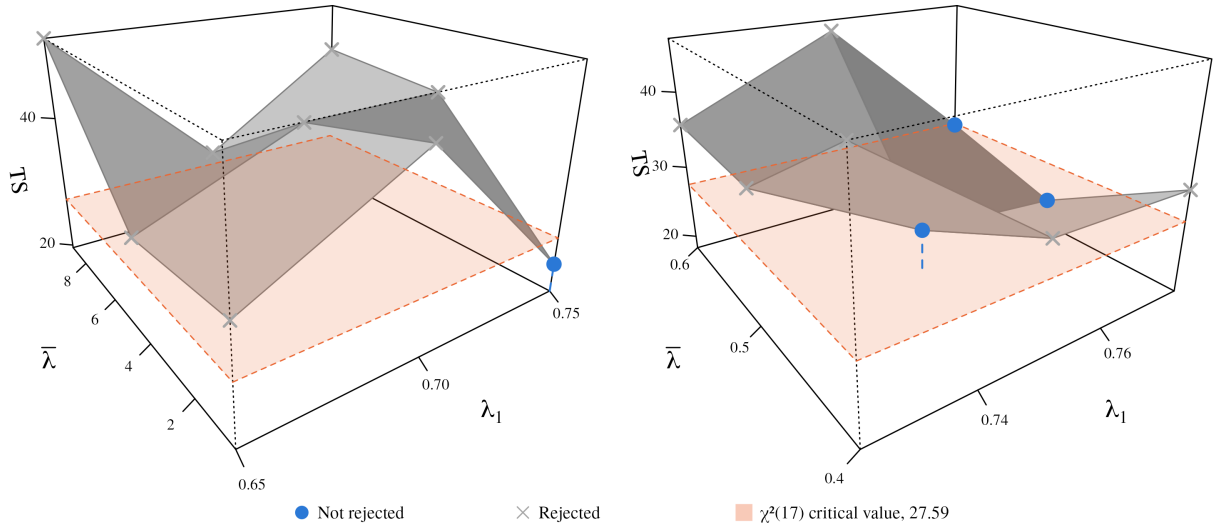


Figure 8: Conservative test statistic over the detection parameters. Left: coarse grid; right: fine grid around the operating point. Each cell is a separate fit with $(\delta_0, \delta_1, \delta_2)$ and γ estimated; height is TS_{cons}. Blue dots are not rejected, grey crosses are rejected; the plane is $\chi^2_{17,.95} = 27.59$.

A.4 Counterfactual implementation

This section states every formula the counterfactual computes, for one firm-year i and one candidate value M of its true materials. Materials are measured as real expenditure (nominal expenditure divided by p_{gdp}), so purchase credits are $\tau_{P,it}M$. The counterfactual purchases rate is $\tilde{\tau}_{P,it} = (1 + \Delta)\tau_{P,it}$; the rate on sales $\tau_{S,it}$ does not change.

The counterfactual holds fixed the detection and cost parameters at the operating point, including the detection scale $\lambda_{0,jt} = 1/(\bar{\lambda}\bar{M}_{j,t-1})$: it is a parameter of the detection function, set through the industry's lagged mean of reported materials, therefore it does not move with the firm's current

choices; capital and labour; productivity ω_{it}, the cost shock ψ_{it} and the output measurement error ε_{it}; and prices P_t, ρ_t.

A.4.1 Baseline objects at a candidate M

From the data Z_{it} and a candidate $M \in (\max\{0, M^*_{it} - c_{\lambda_1}/\lambda_{0,jt}\}, M^*_{it})$:

$$e = M^*_{it} - M, \qquad x = \lambda_{0,jt}\, e, \qquad B \equiv 1 - (1+\lambda_1)x^{\lambda_1}.$$

The baseline evasion FOC, $\tau_{P,it}\rho_t(1-(1+\lambda_1)x^{\lambda_1}) = \exp\{\delta_0 - \delta_1\omega + \delta_2\omega^2 + \psi\}$, says that B is the firm's marginal cost of evasion relative to its marginal benefit, $B = \exp\{\delta_0 - \delta_1\omega + \delta_2\omega^2 + \psi\}/(\tau_{P,it}\rho_t)$. Each candidate M therefore pins down the firm's type (ω, ψ), and the tilt integrates over candidates.

A.4.2 New evasion

At the new rate, the same marginal cost (same ω, ψ, linear cost in e) meets a marginal benefit scaled by $1+\Delta$:

$$\begin{aligned} &(1+\Delta)\tau_{P,it}\rho_t(1-(1+\lambda_1)x'^{\lambda_1}) = \tau_{P,it}\rho_t B \\ \Longrightarrow \quad &x'^{\lambda_1} = \max\left\{0,\ \frac{1 - B/(1+\Delta)}{1+\lambda_1}\right\}, \qquad e' = x'/\lambda_{0,jt}. \end{aligned} \tag{37}$$

Since $B/(1+\Delta) > 0$, $x'^{\lambda_1} < 1/(1+\lambda_1)$: the new evasion stays below the FOC ceiling c_{λ_1} for any Δ. A firm with $B \geq 1+\Delta$ stops overreporting.

e' does not depend on the response of M. The evasion FOC contains $\tau_{P,it}\rho_t$, the detection probability $q(e)$ through $\lambda_{0,jt}$ and λ_1, and the marginal cost $\exp\{\delta_0 - \delta_1\omega + \delta_2\omega^2 + \psi\}$. None of these is M: the detection parameter $\lambda_{0,jt}$ does not depend on the firm's materials, and the cost depends on productivity, not on the input level. Therefore, when M re-optimizes, the evasion FOC is unchanged and Equation 37 holds as is. The candidate M enters e' only through the baseline B, as the firm's type. (Symmetrically, the materials FOC below contains no e: $M \perp e \mid \omega$.)

A.4.3 Response of true materials

The materials FOC with the two rates (Section 2) is, with true output $Y = K^{\alpha_K}L^{\alpha_L}M^{\beta}\exp(\omega)$,

$$\begin{aligned} &(1-\tau_{S,it})P_t\,\beta K^{\alpha_K}L^{\alpha_L}M^{\beta-1}\exp(\omega) = (1-\tau_{P,it})\rho_t \\ \Longrightarrow \quad &M = \left[\frac{(1-\tau_{S,it})P_t\beta K^{\alpha_K}L^{\alpha_L}\exp(\omega)}{(1-\tau_{P,it})\rho_t}\right]^{\frac{1}{1-\beta}}. \end{aligned}$$

Taking the ratio of this solution at $\tilde{\tau}_{P,it}$ and at $\tau_{P,it}$, everything except the purchases rate cancels:

$$r_{it}(\Delta) \equiv \frac{M(\Delta)}{M} = \left(\frac{1 - (1+\Delta)\tau_{P,it}}{1 - \tau_{P,it}} \right)^{-\frac{1}{1-\beta}}. \tag{38}$$

r does not depend on M. The level of M is set by K, L, ω, P_t, ρ_t and $\tau_{S,it}$, all held fixed and cancel in the ratio. The ratio depends only on $\tau_{P,it}$, Δ and β. It is therefore the same for every candidate M of a firm: whatever the firm's true materials are, they scale by the same factor. With $\beta \approx 0.39$ and $\tau_P \approx 0.088$, the elasticity of M with respect to the purchases rate at $\Delta = 0$ is $\tau_P/[(1-\beta)(1-\tau_P)] \approx 0.16$. The new true materials are $M' = r_{it}(\Delta)M$, and reported materials become $M' + e'$.

Output and gross sales-tax revenue scale with M^β, holding K, L, ω and the measurement error ε fixed: $Y' = r_{it}^\beta Y$. The observed gross sales-tax revenue becomes $t1_{it}\, r_{it}^\beta$, where $t1_{it} = \tau_{S,it} \times$ nominal reported sales.

A.4.4 Claimed deductions and revenue

Claims are the credit on true purchases plus the credit on undetected overreporting:

$$C_i(\Delta; M) = (1+\Delta)\tau_{P,it}\Big[r_{it}(\Delta)\, M + (1 - q(e'))e' \Big], \qquad q(e') = x'^{\lambda_1}, \tag{39}$$

and net sales-tax revenue from the firm, in real terms, is

$$R_i(\Delta; M) = \frac{t1_{it}\, r_{it}(\Delta)^\beta}{p_{\text{gdp},t}} - C_i(\Delta; M). \tag{40}$$

With true materials held fixed (the robustness version), $r_{it} \equiv 1$: claims use M and the sales tax is the observed $t1_{it}/p_{\text{gdp},t}$.

A.4.5 Estimation

The target is the expected value per firm-year, in units of $S = n^{-1} \sum_i \tau_{P,it} M_{it}^*$ (the mean claimed credit in the data), entered as one more moment with its own parameter T:

$$g_{10}(M, Z_{it}; T) = \frac{R_i(\Delta; M)}{S} - T.$$

Only γ is free; θ stays at the operating point (Aguiar and Kashaev 2021, App. F). For each firm the

tilted moment is the importance-sampling average over draws $M^{(s)}$ of the latent,

$$\tilde{g}_{10,i} = \sum_s w_{is}(\gamma)\, g_{10}(M^{(s)}, Z_{it}; T), \qquad w_{is}(\gamma) = \frac{\pi_{is} \exp\left(\gamma' g(M^{(s)}, Z_{it})\right)}{\sum_{s'} \pi_{is'} \exp\left(\gamma' g(M^{(s')}, Z_{it})\right)},$$

where π_{is} is the importance weight of draw s with respect to the reference density $f^0_{M|Z}$ and g includes g_{10}. The draws are the same for every T and Δ. The sales-tax term $t1_{it} r_{it}^{\beta} / (p_{\text{gdp},t} S)$ is the same for every draw of firm i: it cancels in the weights and adds a firm-level constant to $\tilde{g}_{10,i}$. It still enters the variance $\hat{\Omega}$ across firms. Because the moment is linear in T, the profiled statistic is

$$TS(T) = 2n \min_{\gamma} \hat{L}_n(\hat{\theta}, \gamma; T),$$

$\hat{T}$ is its minimizer, and the reported set is the conservative one, $\{T : TS(T) \leq \chi^2_{d_g,.95}\}$ with $d_g = 18$ (the 17 estimation rows plus row 10). Row 10 is left out of the reference measure's penalty, since it is bounded for each firm. Because a failed inner solve can only raise TS, each γ problem is solved from several starting values (the operating γ, the solution at the previous T, and the best solution found so far), keeping the lowest; the set is the union of every T accepted along the way, located by bisection outward from $\hat{T}$. Revenue per firm-year is $\hat{T} \times S$, and the set's bounds scale the same way. The bases of the elasticities in Section 6.3 are treated as known.

A.4.6 Robustness: true materials fixed

The revenue curve keeps its shape when true materials stay at their current level (Table 19). At every Δ evaluated, the 95 percent sets with true materials fixed and with true materials responding overlap. With true materials fixed, the estimate breaks even at a cut of about 19 percent, against about 17 percent with true materials responding. Fixing true materials flattens the curve: from a 5 percent cut to a 5 percent increase, net sales-tax revenue falls by 227 real pesos per firm-year with true materials fixed and by 256 with true materials responding. The materials response steepens the curve because the extra credit on true materials exceeds the extra gross sales-tax revenue from the extra output (Section 6.4).

B The Single-Tax Model and the Second-Order Condition

This appendix presents the single-tax model that Section 2 nests, and derives the second-order condition (SOC) for the firm's problem, which translates into the curvature restriction on the detection function $q(e)$ used in Section 6.

Table 19: Net sales-tax revenue per firm-year (real pesos) from the interior firms, by change Δ in the rate on purchases, with true materials responding and with true materials fixed: estimates and 95 percent sets.

	True materials respond		True materials fixed	
Δ	Estimate	95% set	Estimate	95% set
−30%	264	[49, 480]	223	[0, 439]
−20%	59	[−166, 279]	23	[−180, 240]
−10%	−159	[−372, 51]	−183	[−395, 30]
−5%	−272	[−485, −64]	−285	[−486, −77]
0	−391	[−605, −174]	−391	[−605, −174]
+5%	−528	[−740, −309]	−512	[−709, −310]
+10%	−668	[−882, −468]	−640	[−840, −435]
+20%	−965	[−1,180, −733]	−899	[−1,126, −668]
+30%	−1,270	[−1,545, −1,013]	−1,163	[−1,386, −867]

ELVIS interior firms (12,050 firm-years). Net sales-tax revenue = gross sales-tax revenue minus expected sales-tax refunds. 95% set: conservative test, $TS \leq \chi^2_{18,.95} = 28.87$. True materials respond: the materials first-order condition with the two tax rates, capital, labour and productivity fixed.

B.1 The single-tax model

The single-tax model keeps every assumption of Section 2 except one: a single rate τ applies to both sales and purchases. The firm solves

$$\begin{aligned}\max_{K,L,M,e}\ \mathbb{E}[\pi_{it}] = (1-\tau)(P_t K_{it}^{\alpha_K} L_{it}^{\alpha_L} M_{it}^{\beta} \exp(\omega_{it}) - \rho_t M_{it}) - r_t K_{it} - w_t L_{it} \\ + \big[1 - q(e_{it})\big]\tau\rho_t e_{it} - q(e_{it})\phi\tau\rho_t e_{it} - \kappa_{it}(e_{it}, \omega_{it}),\end{aligned} \tag{41}$$

which is Equation 2 with $\tau_S = \tau_P = \tau$. The tax rate drops out of the materials FOC,

$$\beta K_{it}^{\alpha_K} L_{it}^{\alpha_L} M_{it}^{\beta-1} \exp(\omega_{it}) = \rho_t / P_t, \tag{42}$$

and the gross log share, $\ln(\rho_t M_{it}/P_t Y_{it}) = \ln\beta - \varepsilon_{it}$, takes the place of the net share in Equation 4. The evasion FOC is Equation 5 with τ in place of τ_P,

$$\tau\rho_t\big[1 - (1+\phi)(q(e_{it}) + q'(e_{it})e_{it})\big] = \frac{\partial\kappa_{it}(e_{it}, \omega_{it})}{\partial e_{it}}, \tag{43}$$

and firms for which $\tau\rho_t \leq \partial\kappa_{it}(0, \omega_{it})/\partial e_{it}$ do not overreport.

B.1.1 The elasticity of materials under the single-tax model

Under the single-tax model, the gross log share replaces the net share in the first stage. Table 20 compares the two estimates of β from the same corporations.

Table 20: Output elasticity of materials from the corporations' log materials share: two-tax model (net of sales taxes, the headline) and single-tax model (gross share), five industries of Section 4.

Industry	Corporations	Two-tax (net share)	Single-tax (gross share)	Difference
321	524	0.385	0.390	0.005
322	113	0.335	0.347	0.012
324	59	0.412	0.430	0.018
342	256	0.333	0.343	0.010
369	391	0.244	0.248	0.004

Output elasticity of materials, $\hat{\beta} = \exp(\mathbb{E}[s_{it} \mid \text{corporation}])$. Two-tax: log materials share net of sales taxes, $s_{it} = \ln[(\rho_t M_{it} - t_{2,it})/(P_t Y_{it} - t_{1,it})]$. Single-tax: the gross share $\ln(\rho_t M_{it}/P_t Y_{it})$. Same corporations (corporate firm-years in the production sample); difference = single-tax minus two-tax.

B.2 The second-order condition

B.2.1 Setup

Recall the firm's problem in the two-tax model, Equation 2. The firm chooses capital, labour, materials, and overreporting, but the analysis uses only two FOCs, for materials (Equation 3) and for overreporting (Equation 5). Capital and labour are predetermined or dynamic, and the model does not solve for them. A candidate solution is a strict local maximum if the Hessian of expected profit is negative definite at that point.

B.2.2 The evasion margin is separable

Overreporting is separable from the input choices. The probability of detection, the cost of evasion, and the credit on overreported materials do not depend on capital, labour, or materials: every cross-partial between e_{it} and the inputs $X_{it} = (K_{it}, L_{it}, M_{it})$ is zero. The Hessian of expected profit is therefore block-diagonal,

$$\nabla^2 \mathbb{E}[\pi_{it}] = \begin{bmatrix} \dfrac{\partial^2 \mathbb{E}[\pi_{it}]}{\partial e_{it}^2} & \mathbf{0}' \\ \mathbf{0} & \nabla_X^2 \mathbb{E}[\pi_{it}] \end{bmatrix}, \tag{44}$$

where $\nabla_X^2 \mathbb{E}[\pi_{it}]$ is the 3×3 Hessian with respect to the inputs and $\mathbf{0}$ is a 3×1 vector of zeros. Whatever the input choices require of that block, the evasion margin adds a single condition,

$$\frac{\partial^2 \mathbb{E}[\pi_{it}]}{\partial e_{it}^2} < 0,$$

The evasion margin and the inputs place no restrictions on each other. For materials, the only input whose FOC the analysis uses, the second-order condition holds because the output elasticity of

materials is below one, $0 < \beta < 1$:

$$\frac{\partial^2 \mathbb{E}[\pi_{it}]}{\partial M_{it}^2} = (1 - \tau_S) P_t \, \beta(\beta - 1) K_{it}^{\alpha_K} L_{it}^{\alpha_L} M_{it}^{\beta - 2} \exp(\omega_{it}) < 0.$$

The rest of this appendix analyzes the SOC for e_{it} on its own.

B.2.3 The scalar condition

Differentiating the evasion FOC (Equation 5) once more with respect to e_{it},

$$\frac{\partial^2 \mathbb{E}[\pi_{it}]}{\partial e_{it}^2} = -\tau_P \rho_t (1 + \phi) \Big(2q'(e_{it}) + e_{it} q''(e_{it}) \Big) - \frac{\partial^2 \kappa_{it}(e_{it}, \omega_{it})}{\partial e_{it}^2} < 0. \tag{45}$$

Rearranging, the SOC holds whenever

$$\tau_P \rho_t (1 + \phi) \Big(2q'(e_{it}) + e_{it} q''(e_{it}) \Big) + \frac{\partial^2 \kappa_{it}(e_{it}, \omega_{it})}{\partial e_{it}^2} > 0,$$

a sum of two terms. It is therefore sufficient, though not necessary, that the cost be convex in e_{it} and the curvature term strictly positive:

$$\frac{\partial^2 \kappa_{it}(e_{it}, \omega_{it})}{\partial e_{it}^2} \geq 0 \qquad \text{and} \qquad \tau_P \rho_t (1 + \phi) \Big(2q'(e_{it}) + e_{it} q''(e_{it}) \Big) > 0. \tag{46}$$

Under the headline cost function, $\kappa_{it}(e_{it}, \omega_{it}) = e_{it} \exp\{\delta_0 - \delta_1 \omega_{it} + \delta_2 \omega_{it}^2 + \psi_{it}\}$, which is linear in e_{it}, the first term is identically zero, $\partial^2 \kappa_{it} / \partial e_{it}^2 = 0$. The cost function therefore contributes nothing toward satisfying the SOC on its own; the entire burden falls on the second condition, the curvature of the detection function.

B.2.4 The curvature restriction

Since $\tau_P \rho_t (1 + \phi) > 0$, the second condition in Equation 46 reduces to $2q'(e_{it}) + e_{it} q''(e_{it}) > 0$. Economic theory imposes $q'(e_{it}) > 0$ (detection is increasing in overreporting) and $e_{it} > 0$ for an interior evader. Dividing through by $q'(e_{it}) e_{it} > 0$ gives

$$-\frac{q''(e_{it})}{q'(e_{it})} e_{it} < 2. \tag{47}$$

This is a restriction on the curvature of q alone: given the linear-in-e cost above, a firm's evasion cost never enters. It says q' may decline as e_{it} rises (a concave detection function), but not too fast

relative to its own level: the elasticity of the marginal detection probability with respect to e_{it} must exceed -2.

B.2.5 Application to the candidate detection functions

Table 21 applies Equation 47 to the headline detection function of Section 6 and to three alternative forms.

Table 21: Second-order condition by functional form of $q(\cdot)$

	$q(e)$	$q''(e)$	$-q''e/q'$	SOC holds when
Scaled power (headline)	$(\lambda_{0,jt}e)^{\lambda_1}$, $0 < \lambda_1 < 1$	$\lambda_1(\lambda_1 - 1)\lambda_{0,jt}^2(\lambda_{0,jt}e)^{\lambda_1-2}$	$1 - \lambda_1$	always $(0 < 1 - \lambda_1 < 1)$
Linear	$\min\{\lambda e, 1\}$	0	0	always
Exponential	$1 - e^{-\lambda e}$	$-\lambda^2 e^{-\lambda e}$	λe	$\lambda e < 2$
Convex	$\min\{(\lambda e)^k, 1\}$, $k > 1$	$k(k-1)\lambda^2(\lambda e)^{k-2}$	$1 - k$	always $(k > 1 \Rightarrow 1 - k < 0 < 2)$

Scaled power detection, $q(e) = (\lambda_{0,jt}e)^{\lambda_1}$, $0 < \lambda_1 < 1$, $\lambda_{0,jt} = 1/(\bar{\lambda}\bar{M}_{j,t-1})$ (headline specification, Section 6): with $x = \lambda_{0,jt}e$, $q'(e) = \lambda_1\lambda_{0,jt}x^{\lambda_1-1}$ and $q''(e) = \lambda_1(\lambda_1 - 1)\lambda_{0,jt}^2x^{\lambda_1-2}$, which give $-q''e/q' = 1 - \lambda_1 \in (0, 1)$. The SOC holds at every interior point, for any $0 < \lambda_1 < 1$ and any scale; the binding constraint is the FOC's own ceiling $x < (1 + \lambda_1)^{-1/\lambda_1}$.

Linear detection, $q(e) = \min\{\lambda e, 1\}$: $q''(e) = 0$ on the interior and $-q''e/q' = 0 < 2$. The SOC holds at every interior point, with no restriction on λe beyond the FOC's own ceiling ($\lambda e < 1/2$ with $\phi = 0$, from Equation 5).

Exponential detection, $q(e) = 1 - e^{-\lambda e}$: $q'(e) = \lambda e^{-\lambda e}$ and $q''(e) = -\lambda^2 e^{-\lambda e}$, which give $-q''e/q' = \lambda e$. The SOC requires $\lambda e < 2$, a genuine restriction, distinct from (and, for this functional form, looser than) the FOC's own ceiling $\lambda e < 1$ implied by Equation 5 with $\phi = 0$. Since $\lambda e < 1 \implies \lambda e < 2$, the FOC ceiling is the binding constraint in practice; the SOC is automatically satisfied wherever the FOC has an interior solution at all.

Convex detection, $q(e) = \min\{(\lambda e)^k, 1\}$, $k > 1$: $q'(e) = k\lambda(\lambda e)^{k-1}$ and $q''(e) = k(k-1)\lambda^2(\lambda e)^{k-2}$, which give $-q''e/q' = 1 - k$. Since $k > 1$, this is always negative and hence below 2: convex detection satisfies the SOC unconditionally, for any $k > 1$.

All four detection functions satisfy the second-order condition on the relevant part of their domain: scaled power, linear and convex unconditionally, and exponential provided $\lambda e < 2$, which that function's own FOC ceiling already implies. The SOC therefore does not discipline the choice among them; it is a modelling choice about the detection technology.

Table 22: Production-function estimates by industry, instrument $\tilde{\mathscr{W}}_{it-2}$, with 95 percent test-inversion regions (sharp test).

Industry	$\hat{\beta}$	β, sharp	$\hat{\alpha}_K$	α_K, sharp	$\hat{\alpha}_L$	α_L, sharp	n
313 Beverages	0.287	[0.269, 0.306]	0.406	[0.295, 0.575]	0.307	[−0.055, 0.520]	969
321 Textiles	0.385	[0.357, 0.416]	0.300	[0.210, 0.375]	0.349	[0.230, 0.485]	2,693
322 Wearing apparel	0.335	[0.288, 0.391]	0.182	[0.085, 0.265]	0.392	[0.320, 0.470]	4,247
324 Footwear	0.412	[0.365, 0.465]	0.119	[−0.215, 0.270]	0.427	[0.190, 0.570]	1,063
331 Wood products	0.324	[0.263, 0.398]	0.076	[−0.070, 0.200]	0.554	[0.400, 0.700]	951
342 Printing and publishing	0.333	[0.294, 0.377]	0.292	[0.215, 0.360]	0.408	[0.305, 0.530]	1,967
351 Industrial chemicals	0.340	[0.292, 0.396]	0.308	[−0.155, 0.495]	0.362	[0.005, 0.665]	759
352 Other chemicals	0.407	[0.379, 0.437]	0.325	[0.240, 0.420]	0.273	[0.165, 0.400]	2,205
369 Non-metallic minerals	0.244	[0.207, 0.288]	0.333	[0.210, 0.485]	0.549	[0.210, 0.740]	1,480

Output elasticities of materials ($\hat{\beta}$, from the first stage on corporations, log materials share net of sales taxes), capital ($\hat{\alpha}_K$) and labour ($\hat{\alpha}_L$) for the nine industries where overreporting is detected. Point estimates: the minimum of the test-inversion statistic, with the covariance of the moments re-estimated at each candidate (α_K, α_L). β, sharp: 95% test-inversion region for β from the corporations' share moment ($\chi^2_{1,0.95}$; the production-function moments are exactly identified given β and profiled). α, sharp: projections of the 95% test-inversion region ($\chi^2_{2,0.95}$, credit for profiling γ_0, γ_1) on a 0.005 grid over $[0, 1]^2$, extended beyond it where the region reaches its edge. n: observations in the first stage. Juridical organization codes 6–9 excluded.

The derivation carries over to the single-tax model (Section B.1) with τ in place of τ_P: the rate scales the curvature term but does not change its sign, and Equation 47 is the same condition in both models.

C Production-Function Estimates

The stage-2 estimation in Section 6 uses the nine industries where overreporting is detected. Table 22 lists the production-function estimates it uses, with instrument $\tilde{\mathscr{W}}_{it-2}$. Table 24 reports the distribution of the overreporting ratio for the same nine industries. Beverages (313) appears here but not in Section 3.6 or Section 4, because the variance of its unincorporated firms' residual is smaller than the corporations', which violates the condition the deconvolution requires. Its estimates are shown for reference only.

C.1 Estimation details

The production-function moments in Equation 17 hold for any instrument that is independent of the innovation η_{it} and of the measurement errors ε_{it} and ε_{it-1}. Table 25 lists the candidates by the residual they must be orthogonal to. With a single instrument Z, the vector $(1, Z, k_{it}, l_{it})$ interacted with the residual gives four moments for the four parameters $(\alpha_K, \alpha_L, \gamma_0, \gamma_1)$, and the system is exactly identified. The joint system interacts $(1, m^*_{it-1}, \tilde{\mathscr{W}}_{it-2}, k_{it}, l_{it})$ with the residual, which gives five moments for the same four parameters. In both systems, β stays at its first-stage value, estimated

Table 23: Production-function estimates by industry, instrument $\tilde{\mathscr{W}}_{it-2}$, with 95 percent test-inversion regions (conservative test).

Industry	$\hat{\beta}$	β, cons.	$\hat{\alpha}_K$	α_K, cons.	$\hat{\alpha}_L$	α_L, cons.	n
313 Beverages	0.287	[0.257, 0.321]	0.406	[0.095, 1.005]	0.307	(· , 0.665]	969
321 Textiles	0.385	[0.338, 0.440]	0.300	[0.185, 0.390]	0.349	[0.200, 0.525]	2,693
322 Wearing apparel	0.335	[0.258, 0.435]	0.182	[0.055, 0.285]	0.392	[0.305, 0.490]	4,247
324 Footwear	0.412	[0.335, 0.506]	0.119	[−0.350, 0.305]	0.427	[0.095, 0.605]	1,063
331 Wood products	0.324	[0.228, 0.459]	0.076	[−0.115, 0.240]	0.554	[0.350, 0.740]	951
342 Printing and publishing	0.333	[0.269, 0.412]	0.292	[0.190, 0.375]	0.408	[0.280, 0.570]	1,967
351 Industrial chemicals	0.340	[0.262, 0.442]	0.308	[−0.265, 0.565]	0.362	[−0.085, 0.805]	759
352 Other chemicals	0.407	[0.360, 0.460]	0.325	[0.215, 0.450]	0.273	[0.135, 0.445]	2,205
369 Non-metallic minerals	0.244	[0.184, 0.324]	0.333	[−0.145, 0.545]	0.549	[−0.235, 0.785]	1,480

Output elasticities of materials ($\hat{\beta}$, from the first stage on corporations, log materials share net of sales taxes), capital ($\hat{\alpha}_K$) and labour ($\hat{\alpha}_L$) for the nine industries where overreporting is detected. Point estimates: the minimum of the test-inversion statistic, with the covariance of the moments re-estimated at each candidate (α_K, α_L). β, cons.: 95% test-inversion region for β, conservative test ($\chi^2_{5,0.95}$, no credit for profiling). α, cons.: projections of the 95% conservative test-inversion region ($\chi^2_{4,0.95}$) on a 0.005 grid over $[0,1]^2$, extended beyond it where the region reaches its edge. (· or ·): the region is not bounded on that side within $[-1,3]^2$. n: observations in the first stage. Juridical organization codes 6–9 excluded.

Table 24: Distribution of the overreporting ratio e/M by industry, unincorporated firms, all nine stage-2 industries.

Industry	Mean	SD	P10	Median	P90
313 Beverages	48.0%	1.6%	46.0%	48.0%	50.1%
321 Textiles	16.2%	5.3%	9.5%	15.8%	23.4%
322 Wearing apparel	23.1%	3.9%	18.0%	23.1%	28.2%
324 Footwear	10.5%	5.8%	4.4%	9.3%	18.2%
331 Wood products	14.5%	7.3%	7.3%	12.8%	23.9%
342 Printing and publishing	4.5%	7.2%	3.2%	4.0%	5.1%
351 Industrial chemicals	13.2%	6.7%	6.9%	11.6%	21.4%
352 Other chemicals	6.4%	2.1%	4.2%	6.0%	9.0%
369 Non-metallic minerals	12.4%	2.3%	9.5%	12.2%	15.4%

Overreporting ratio $x = e/M$: overreported materials as a share of true materials. Moments and quantiles of the density $f_x(y) = f_u(\ln(1+y))/(1+y)$, obtained by transforming the deconvolved density of $u = \ln(1+e/M)$ (penalized B-spline deconvolution, log materials share net of sales taxes). Sample: unincorporated firms that pay sales tax on purchases. Industries: the nine where overreporting is detected.

from the corporations' log materials share net of sales taxes, with juridical organization codes 6 to 9 excluded.

Uncertainty comes from test inversion. At each candidate (α_K, α_L) on a 0.005 grid, the parameters (γ_0, γ_1) are profiled out in closed form, because the problem is linear in γ at fixed α. With a single instrument, the sharp test compares the statistic with χ^2_2, giving credit for profiling (γ_0, γ_1), and the conservative test compares it with χ^2_4; with the joint system, the critical values are χ^2_3 and χ^2_5. The grid covers $[0, 1]^2$ and extends beyond it in steps of 0.25 wherever a region reaches its edge, down to -1 and up to 3. An open bracket marks a region that is not bounded on that side within $[-1, 3]^2$, which happens only under m^*_{it-1} and for the conservative region for α_L in Beverages (313). The region for β comes from the corporations' share moment alone. With a single instrument, the production-function moments are exactly identified given β, and profiling them leaves the share moment unchanged.

The point estimate is the minimum of the same statistic whose regions are reported, with the covariance of the moments re-estimated at each candidate. A two-step estimator, which fixes the weighting matrix at a first-step point, understates the statistic when that point fits badly: in Beverages, the joint statistic is 1.62 at the two-step point and 9.96 with the covariance re-estimated there.

The headline instrument is $\tilde{\mathscr{W}}_{it-2}$ alone. Table 26 compares it with m^*_{it-1} and with the joint system in the five industries of Section 4. Across all 28 industries, $\tilde{\mathscr{W}}_{it-2}$ gives no estimate at a bound of the parameter space and no empty region, and its regions are about as tight as the joint system's. The median width is 0.25 for α_K and 0.35 for α_L, against 0.23 and 0.40 for the joint system. On its own, m^*_{it-1} is a weak instrument: its regions for α_K reach 1 in 24 of the 28 industries. The joint sharp region is empty in 311, 312, 313 and 352, where the moments of the two instruments conflict. The untilded $\mathscr{W}_{it-2}$, which contains $\alpha_K k_{it-2} + \alpha_L l_{it-2}$, is not used.

GNR and OLS use the same sample and the same net-of-tax share. The GNR estimates come from an R port of the authors' Stata code, which reproduces it to four decimals.

The distribution of productivity in Section 4 comes from deconvolving $\tilde{\mathscr{W}}_{it} = \omega_{it} + (1 - \beta)\varepsilon_{it}$ with the corporations' distribution of measurement error, scaled by $1 - \beta$, using the estimator of Section G. The persistence of productivity is the coefficient γ_1 from regressing $\tilde{\mathscr{W}}_{it}$ on $\tilde{\mathscr{W}}_{it-1}$ with $\tilde{\mathscr{W}}_{it-2}$ as the instrument, which is independent of the measurement error in the lag, ε_{it-1}. Under GNR, it is the OLS coefficient of ω_{it} on ω_{it-1}.

Candidate instruments for the production-function moments of Section 4, by the residual they must be orthogonal to:

Table 26: Evasion-corrected production-function estimates under three instrument systems: $\tilde{\mathscr{W}}_{it-2}$ (the headline in Section 4), m^*_{it-1}, and both together (joint efficient GMM), with 95 percent test-inversion regions (sharp and conservative). Industries: the five of Section 4.

		$\tilde{\mathscr{W}}_{it-2}$		m^*_{it-1}		Joint efficient GMM	
Industry	$\hat{\beta}$	$\hat{\alpha}_K$	$\hat{\alpha}_L$	$\hat{\alpha}_K$	$\hat{\alpha}_L$	$\hat{\alpha}_K$	$\hat{\alpha}_L$
321	0.385	0.300	0.349	0.311	0.290	0.297	0.328
sharp	[0.357, 0.416]	[0.210, 0.375]	[0.230, 0.485]	[0.210, 0.470]	[0.145, 0.420]	[0.230, 0.355]	[0.240, 0.425]
cons.	[0.338, 0.440]	[0.185, 0.390]	[0.200, 0.525]	[0.175, ·)	[−0.690, 1.060]	[0.180, 0.375]	[0.205, 0.460]
322	0.335	0.182	0.392	0.167	0.394	0.161	0.373
sharp	[0.288, 0.391]	[0.085, 0.265]	[0.320, 0.470]	[0.090, 0.240]	[0.325, 0.465]	[0.075, 0.235]	[0.315, 0.430]
cons.	[0.258, 0.435]	[0.055, 0.285]	[0.305, 0.490]	[0.070, ·)	[0.025, 1.070]	[0.045, 0.260]	[0.295, 0.450]
324	0.412	0.119	0.427	0.710	0.470	0.123	0.433
sharp	[0.365, 0.465]	[−0.215, 0.270]	[0.190, 0.570]	[0.010, ·)	[0.130, 0.975]	[−0.265, 0.290]	[0.140, 0.570]
cons.	[0.335, 0.506]	[−0.350, 0.305]	[0.095, 0.605]	(·, ·)	[−0.150, 1.995]	[−0.350, 0.320]	[0.065, 0.600]
342	0.333	0.292	0.408	0.266	0.401	0.265	0.402
sharp	[0.294, 0.377]	[0.215, 0.360]	[0.305, 0.530]	[−0.400, ·)	[−0.570, 1.555]	[0.210, 0.310]	[0.335, 0.480]
cons.	[0.269, 0.412]	[0.190, 0.375]	[0.280, 0.570]	[−0.515, ·)	[−0.880, 1.730]	[0.180, 0.335]	[0.305, 0.525]
369	0.244	0.333	0.549	0.327	0.639	0.354	0.326
sharp	[0.207, 0.288]	[0.210, 0.485]	[0.210, 0.740]	[0.165, ·)	[−0.135, 1.575]	[−0.025, 0.450]	[0.045, 0.535]
cons.	[0.184, 0.324]	[−0.145, 0.545]	[−0.235, 0.785]	(·, ·)	[−0.430, ·)	[−0.200, 0.530]	[−0.270, 0.630]

$\hat{\beta}$ is the corporations' first-stage estimate (log materials share net of sales taxes); below it, its 95% test-inversion regions from the corporations' share moment: sharp (χ^2_1) and conservative (χ^2_5). Point estimates of (α_K, α_L): the minimum of the test-inversion statistic, with the covariance of the moments re-estimated at each candidate, the statistic whose regions are shown. $\tilde{\mathscr{W}}_{it-2}$ and m^*_{it-1}: one instrument each (exactly identified, statistic zero at the estimate); below them, the projections of the 95% test-inversion regions for (α_K, α_L) (0.005 grid over $[0, 1]^2$, extended beyond it where a region reaches its edge): sharp (χ^2_2, credit for profiling γ_0, γ_1) and conservative (χ^2_4). Joint efficient GMM: both instruments together; sharp (χ^2_3) and conservative (χ^2_5) regions. ($\cdot$ or $\cdot$): the region is not bounded on that side within $[-1, 3]^2$.

Table 25: Orthogonality by residuals

	Orthogonality
η_{it}	$k_{it}, l_{it}, \varepsilon_{it}, k_{it-1}, l_{it-1}, \varepsilon_{it-1}, \omega_{it-1}, \mathscr{W}_{it-1}, m_{it-1}, m^*_{it-1}$
ε_{it}	$k_{it}, l_{it}, \omega_{it}, m_{it}, m^*_{it}, \eta_{it}, k_{it-1}, l_{it-1}, \varepsilon_{it-1}, \omega_{it-1}, \mathscr{W}_{it-1}, m_{it-1}, m^*_{it-1}$
ε_{it-1}	$k_{it}, l_{it}, \omega_{it}, \varepsilon_{it}, m_{it}, m^*_{it}, \eta_{it}, k_{it-1}, l_{it-1}, \omega_{it-1}, \mathscr{W}_{it}, m_{it-1}, m^*_{it-1}$

D Robustness of the Test for Overreporting

Table 5 reports the sharp version of the test, which takes ln $\hat{\beta}$ from corporations as the true value. This appendix reports two alternatives and shows that the test does not depend on the functional forms of the detection probability, the cost of evasion or the technology (Section D.1). Table 27 gives the conservative test, which gives no credit for the corporations' moment. Table 28 gives a

Table 27: Test for overreporting of materials by industry, conservative test. Mean of $\mathcal{V}_{it}$ among unincorporated firms and 95 percent confidence set for its expectation, with $\ln \beta$ estimated from corporations.

Industry	$\hat{\mu}$	95% region, conservative
311 Food products	0.024	$(\cdot,\ 0.054]$
312 Other food products	−0.018	$(\cdot,\ 0.100]$
313 Beverages	0.390	$[0.287,\ 0.492]$
321 Textiles	0.151	$[0.105,\ 0.197]$
322 Wearing apparel	0.224	$[0.187,\ 0.261]$
323 Leather products	−0.192	†
324 Footwear	0.095	$[0.050,\ 0.140]$
331 Wood products	0.250	$[0.156,\ 0.344]$
332 Furniture	−0.007	$(\cdot,\ 0.075]$
341 Paper products	−0.066	$(\cdot,\ 0.022]$
342 Printing and publishing	0.074	$[0.026,\ 0.122]$
351 Industrial chemicals	0.165	$(\cdot,\ 0.375]$
352 Other chemicals	0.057	$(\cdot,\ 0.125]$
356 Plastic products	−0.107	†
369 Non-metallic minerals	0.193	$[0.087,\ 0.298]$
381 Metal products	−0.005	$(\cdot,\ 0.033]$
382 Non-electrical machinery	−0.014	$(\cdot,\ 0.036]$
383 Electrical machinery	0.034	$(\cdot,\ 0.103]$
384 Transport equipment	−0.129	†
390 Other manufacturing	0.057	$(\cdot,\ 0.130]$

$\hat{\mu}$: mean of $\mathcal{V}_{it} = \ln(\rho_t M^*_{it}/P_t Y_{it}) - \ln \hat{D}$, with the log materials share net of sales taxes and $\ln \hat{D}$ the mean among corporations. 95% test-inversion region for $\mu = E[\mathcal{V}_{it}]$ over the grid $\mu \in [0, 1]$ (under the model $\mu \geq 0$), taking $\ln \hat{D}$ as the true value; plant-clustered, efficiently weighted moments; conservative, $\chi^2_{2,0.95} = 5.99$ critical value (no credit for the corporations' moment). Where the region excludes 0, the test rejects no overreporting. $(\cdot,\ b]$: $\mu = 0$ is not rejected. † Empty region: every $\mu \in [0, 1]$ is rejected.

bootstrap test that re-estimates $\ln \hat{\beta}$ in every replication and thereby carries the sampling error of the first stage.

Both alternatives confirm most of the rejections of the sharp test. The conservative test (Table 27) rejects in seven of the nine industries where the sharp test rejects at the 5 percent level; in Industrial chemicals (351) and Other chemicals (352), its region includes zero. The bootstrap test (Table 28) rejects in seven industries at the 10 percent level or lower: Beverages (313), Textiles (321) and Wearing apparel (322) at 1 percent, Wood products (331) and Non-metallic minerals (369) at 5 percent, and Footwear (324) and Industrial chemicals (351) at 10 percent. It does not reject in Printing and publishing (342) or in Other chemicals (352). Neither alternative rejects in an industry where the sharp test does not.

Table 28: Bootstrap test for overreporting of materials by industry. Mean of $\mathcal{V}_{it}$ among unincorporated firms, with ln β re-estimated from corporations in every replication.

Industry	$\hat{\mu}$	95% CI
311 Food products	0.024	[−0.035, 0.081]
312 Other food products	−0.018	[−0.215, 0.148]
313 Beverages	0.390***	[0.288, 0.485]
321 Textiles	0.151***	[0.063, 0.236]
322 Wearing apparel	0.224***	[0.056, 0.374]
323 Leather products	−0.192	[−0.305, −0.087]
324 Footwear	0.095*	[−0.021, 0.232]
331 Wood products	0.250**	[0.014, 0.461]
332 Furniture	−0.007	[−0.277, 0.261]
341 Paper products	−0.066	[−0.190, 0.051]
342 Printing and publishing	0.074	[−0.061, 0.203]
351 Industrial chemicals	0.165*	[−0.050, 0.385]
352 Other chemicals	0.057	[−0.029, 0.150]
356 Plastic products	−0.107	[−0.195, −0.025]
369 Non-metallic minerals	0.193**	[0.005, 0.376]
381 Metal products	−0.005	[−0.095, 0.079]
382 Non-electrical machinery	−0.014	[−0.125, 0.089]
383 Electrical machinery	0.034	[−0.089, 0.161]
384 Transport equipment	−0.129	[−0.279, 0.021]
390 Other manufacturing	0.057	[−0.148, 0.282]

$\hat{\mu}$: mean of $\mathcal{V}_{it}$ among unincorporated firms, log materials share net of sales taxes, ln $\hat{D}$ the mean among corporations. Bootstrap with 2,000 replications: plants resampled separately within corporations and within unincorporated firms, with ln $\hat{D}$ re-estimated in every replication, so the test carries the first-stage sampling error. Basic (pivotal) 95% confidence intervals. *, **, *** denote significance at the 10%, 5%, and 1% levels (one-sided). The intervals are two-sided, so a one-sided rejection at 5% can coexist with an interval that includes zero.

D.1 The test under general detection, cost and technology

This section derives the results stated in Section 3.4: the test depends neither on the functional forms of the detection probability and the cost of evasion, nor on the Cobb-Douglas form of the technology.

Setup. Let output from the firm's perspective be $Y_{it} = G(K_{it}, L_{it}, M_{it}) \exp(\omega_{it})$, with $G_M > 0$, common to corporations and unincorporated firms in an industry. Let the detection probability and the cost of evasion depend on overreporting and on any other variables Z_{it}, possibly including materials: $q(e, Z)$ and $\kappa(e, Z)$. The firm's problem is

$$\max_{M,e,K,L} \mathbb{E}[\pi_{it}] = (1-\tau_S)P_t G(K_{it}, L_{it}, M_{it}) \exp(\omega_{it}) - (1-\tau_P)\rho_t M_{it} - r_t K_{it} - w_t L_{it} + \tau_P \rho_t e_{it}[1-(1+\phi)q(e_{it}, Z_{it})]$$

Under the null. A truthful firm faces no detection and incurs no cost of evasion: $q(0, Z) = 0$ and $\kappa(0, Z) = 0$ for every Z. At $e_{it} = 0$, the evasion terms are zero for every M_{it}, and so are their derivatives with respect to M_{it}. The materials condition is the one without evasion, $(1 - \tau_S)P_t G_M \exp(\omega_{it}) = (1 - \tau_P)\rho_t$. Multiplying by M_{it} and dividing by the value of observed output, $P_t Y_{it}$, gives

$$\ln\left(\frac{(1-\tau_P)\rho_t M_{it}}{(1-\tau_S)P_t Y_{it}}\right) = \ln D(K_{it}, L_{it}, M_{it}) - \varepsilon_{it}, \qquad D \equiv \frac{G_M M}{G}, \tag{48}$$

where D is the output elasticity of materials. Under Cobb-Douglas, $D = \beta$ and Equation 48 is Equation 4. Corporations identify D as a function of inputs, $\ln D(K, L, M) = \mathbb{E}[s_{it} \mid K, L, M, \text{corporation}]$, provided measurement error is also independent of capital and labour.

The residual. Evaluating the elasticity at each firm's reported materials, and adding and subtracting the log share of true materials, gives, suppressing K_{it} and L_{it},

$$\mathscr{V}_{it} \equiv s_{it} - \ln D(M^*_{it}) = \ln\left(\frac{M^*_{it}}{M_{it}}\right) - \ln\left(\frac{D(M^*_{it})}{D(M_{it})}\right) - \varepsilon_{it} = \ln\left(\frac{H(M^*_{it})}{H(M_{it})}\right) - \varepsilon_{it}, \tag{49}$$

where $H \equiv G/G_M$. Under the null, $M^*_{it} = M_{it}$ and $\mathscr{V}_{it} = -\varepsilon_{it}$ for any G. The test's size does not depend on the technology, and a rejection means that some firms overreport. Under the alternative, $M^*_{it} > M_{it}$ for overreporting firms, and $\ln\left(H(M^*_{it})/H(M_{it})\right)$ is positive whenever H increases in materials. Since $\partial H/\partial M = 1 - GG_{MM}/G_M^2$, this holds when output is concave in materials, $G_{MM} \leq 0$. The exact condition is weaker: $GG_{MM} < G_M^2$, i.e. $\ln G$ concave in materials. The expected residual is then positive and the test stays one-sided. Under Cobb-Douglas, $H = M/\beta$ and $\ln\left(H(M^*_{it})/H(M_{it})\right) = u_{it}$, as in Equation 10.

If q or κ depend on materials, evaders' true materials may depart from Equation 48 under the

alternative. This changes the power of the test, not its size: under the null no firm evades.

A second flexible input. Let N_{it} be a second flexible input that is not creditable against the sales tax, such as energy, fuels or services. Under the model, the firm gains nothing by overreporting it and reports it truthfully. With N_{it} in the technology, $G(K_{it}, L_{it}, M_{it}, N_{it})$, its log share, $s^N_{it} \equiv \ln\left(P_{N,t}N_{it}/(1-\tau_S)P_tY_{it}\right)$, satisfies $s^N_{it} = \ln D_N - \varepsilon_{it}$, with $D_N \equiv G_N N/G$. Evaluating D_N at reported materials,

$$\mathscr{V}^N_{it} \equiv s^N_{it} - \ln D_N(M^*_{it}) = -\ln\left(\frac{D_N(M^*_{it})}{D_N(M_{it})}\right) - \varepsilon_{it}.$$

Under the null, $\mathscr{V}^N_{it} = -\varepsilon_{it}$. Under the alternative, the first term differs from zero as long as D_N is strictly monotone in materials; concavity is no longer needed. Its sign depends on the technology: negative if D_N increases in materials, positive if it decreases. This version is a two-sided test, and a rejection still points only to overreporting.

Estimation. Beyond Cobb-Douglas, the elasticity varies with inputs. Estimating it from corporations and evaluating it at unincorporated firms' inputs requires a parametric form, such as translog, or overlap in inputs between the two groups; corporations are much larger. The test in Section 3 is the Cobb-Douglas case, $D = \beta$.

E The Tax Wedge in the Reform Estimates

A tax change can move reported materials shares mechanically, without any change in overreporting: prices and the tax paid on sales and purchases adjust to the new rates. In addition, a a missing tax-wedge in estimation could generate a omitted-variable bias. The two-tax model isolates the part of this adjustment that runs through the sales-tax rates. With different rates on sales, τ_S, and on purchases, τ_P, the gross materials share carries the wedge $\ln\left((1-\tau_S)/(1-\tau_P)\right)$, which depends on the rates themselves. The 1983 reform changed those rates in the liable industries. After the reform, a rise in the gross share could reflect either more overreporting or this wedge. The event study in Section 5 uses the share net of sales taxes, which removes the wedge.

Table 29 estimates the wedge's contribution directly, the same event study, with the gross minus the net share as the outcome. Its coefficients equal the gap between the event study on the gross share and on the net share. In liable industries, the wedge does not move before the reform. After it, the wedge rises by between 0.003 and 0.010 for unincorporated firms in liable industries, between a twentieth and a fifth of their response measured on the gross share. For proprietorships in 1984–1986, it rises by 0.011 to 0.013, between two fifths and all of their response in those years. A diff-in-diff study on the gross share would attribute this mechanical component to overreporting.

Table 29: The tax wedge in the event study of the 1983 reform: gross minus net-of-sales-tax log materials share, difference relative to 1983.

	Unincorporated firms			
Year	Liable	Exempt	LLCs, liable	Proprietorships, liable
1981	0.000 (0.001)	-0.003*** (0.001)	0.000 (0.001)	0.001 (0.003)
1982	0.001 (0.001)	0.002*** (0.001)	0.001 (0.001)	0.002 (0.003)
1983	–	–	–	–
1984	0.007*** (0.001)	0.002*** (0.001)	0.006*** (0.001)	0.013*** (0.003)
1985	0.008*** (0.001)	0.003*** (0.001)	0.007*** (0.001)	0.013*** (0.004)
1986	0.007*** (0.001)	0.003*** (0.001)	0.007*** (0.002)	0.011*** (0.004)
1987	0.006*** (0.001)	0.005*** (0.001)	0.005*** (0.001)	0.008** (0.003)
1988	0.004*** (0.001)	0.004*** (0.001)	0.003* (0.001)	0.009*** (0.003)
1989	0.006*** (0.001)	0.004*** (0.001)	0.005*** (0.001)	0.008** (0.004)
1990	0.003** (0.001)	0.004*** (0.001)	0.002 (0.001)	0.008** (0.004)
1991	0.010*** (0.001)	0.002*** (0.001)	0.010*** (0.002)	0.012*** (0.004)

Outcome: $\ln(\rho_t M^*_{it}/P_t Y_{it}) - \ln\left((1-\tau_{P,it})\rho_t M^*_{it}/(P_t Y_{it} - \tau_{S,it}\,\text{sales}_{it})\right)$. Same specification as the differences in the event study of the 1983 reform: difference relative to 1983 (reference, –), industry fixed effects, unincorporated firms (or LLCs, proprietorships) relative to corporations, in sales-tax-liable and exempt industries. Standard errors in parentheses, clustered by plant. * p<0.1, ** p<0.05, *** p<0.01. Observations: 39,495; juridical organization codes 6–9 excluded.

E.1 Aggregate estimates of sales-tax evasion

Figure 9 reproduces an estimate of the sales tax on domestic transactions collected as a percentage of its potential revenue, computed from Colombia's national accounts (Sánchez Torres and Gutiérrez Sourdis 1994, 107). The authors title the series "evasion of the domestic sales tax" (*evasión del IVA interno*); the uncollected share is 100 minus the plotted value. The series starts in 1984, the year the reform took effect. The uncollected share rose from 58 percent in 1984 to between 68 and 71 percent in 1987–1990 and fell to 65 percent in 1991. Overreporting by unincorporated firms follows the same path (Figure 1).

F The Elasticity of Overreporting to the Tax Rate

This appendix defines the elasticity of overreporting with respect to τ_P reported in Section 5.5, explains the test behind its confidence set, and reports the results by year.

F.1 Definition

Let μ_t be the average log ratio of reported to true materials of unincorporated firms in ST-liable industries in year t, and let $r_t = \exp(\mu_t) - 1$ be overreporting as a share of true materials. Let T_t be the average τ_P of the same firms. The midpoint elasticity between 1983 (subscript 0) and a later

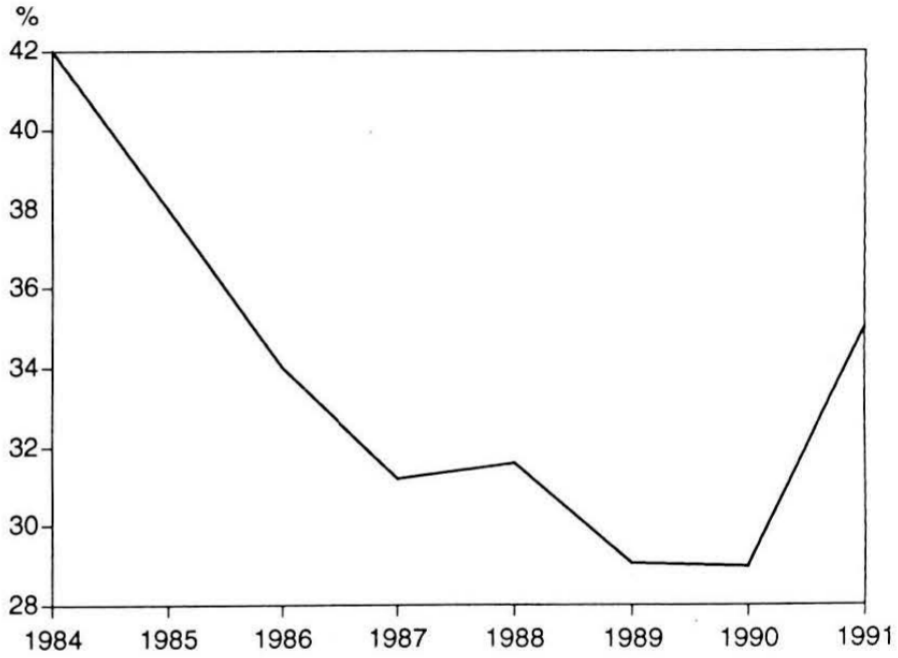


Figure 9: Sales tax on domestic transactions collected as a percentage of its potential revenue, Colombia, 1984–1991. The uncollected share, 100 minus the plotted value, measures evasion. Source: Sánchez Torres and Gutiérrez Sourdis (1994, Gráfico 4, p. 107), computed from the national accounts.

year (subscript 1) is

$$E = \frac{(r_1 - r_0)/\bar{r}}{(T_1 - T_0)/\bar{T}}, \qquad \bar{r} = \tfrac{1}{2}(r_0 + r_1), \quad \bar{T} = \tfrac{1}{2}(T_0 + T_1). \tag{50}$$

Both changes are divided by their midpoints, not by their 1983 values. Since r_0 is close to zero, a change measured from r_0 is very large and very noisy, while the midpoint changes the base only mildly. Because overreporting cannot be negative, $r_0, r_1 \geq 0$, and Equation 50 cannot exceed $2/d$, where $d = (T_1 - T_0)/\bar{T}$. The cap is 9.8 in 1987.

F.2 The test

The null hypothesis that the elasticity equals E_0 can be written without a division:

$$\theta(E_0) = (r_1 - r_0) - E_0\, d\, \bar{r} = 0. \tag{51}$$

The function θ depends on four means, $m = (\mu_0, \mu_1, T_0, T_1)$, each estimated from its own moment. As in the sharp version of the test for overreporting, I take the corporations' mean log share in each industry, $\ln \hat{\beta}_j$, as the true value. Then μ_t is the mean of $s_{ijt} - \ln \hat{\beta}_j$ and T_t the mean of τ_P among unincorporated firms in ST-liable industries in year t. The null asks whether these means, taken as the truth, imply an elasticity of E_0. The statistic is

$$TS(E_0) = \frac{\hat{\theta}(E_0)^2}{\nabla\theta(E_0)'\, \hat{V}\, \nabla\theta(E_0)},$$

compared with a χ^2_1 critical value. $\hat{V}$ is the variance of $\hat{m}$, estimated from centred influence scores summed within plant, with the divisor n, the same covariance estimator as in the test for overreporting, the production-function step and the structural estimator. The gradient depends on E_0, so the variance is re-evaluated at every candidate. The confidence set contains the values of E_0 that the test does not reject at 5 percent. Because the four means are just identified, no weighting matrix enters.

F.3 Results

Table 30 reports the elasticity for each year after 1983. The 1987 set, 4.2 to 8.5, lies below its cap, so the cap does not drive it. The elasticity is stable at roughly 4 to 7 once the response has phased in.

Table 30: Midpoint elasticity of overreporting with respect to τ_P, 1983 to each year, unincorporated firms in ST-liable industries. Overreporting r_t as a share of true materials, mean τ_P, elasticity from Equation 50, and the 95 percent test-inversion set, with variance clustered by plant.

Year	Overreporting, 1983	Overreporting, year	τ_P, 1983	τ_P, year	Elasticity	95% set
1984	2.7%	5.6%	8.2%	9.3%	5.1	[2.2, 10.4]
1985	2.7%	7.4%	8.2%	10.2%	4.1	[2.4, 6.9]
1986	2.7%	9.4%	8.2%	10.2%	5.0	[3.3, 7.5]
1987	2.7%	11.3%	8.2%	10.0%	5.9	[4.2, 8.5]
1988	2.7%	11.7%	8.2%	9.8%	6.6	[4.7, 9.5]

The 95% set collects the elasticities that a χ^2_1 test does not reject at 5%, taking the corporations' mean log share by industry, the overreporting levels and the mean rates as the truth.

G Estimating the Distribution of Overreporting

This appendix describes the estimator behind Section 3.6. I use a spline model of the density of overreporting, fitted to the observed residuals by penalized maximum likelihood, with the distribution of measurement error taken from corporations.

Following Kang and Qiu (2021), the density of overreporting is a logspline,

$$f_{u|\Theta}(u) = \exp(s(u;\theta) - C(\theta)), \quad L < u < R,$$

where $[L, R]$ is the support of f_u,

$$s(u;\theta) = \sum_{j=1}^{K-1} \theta_j B_j(u),$$

$\{B_j(u), j = 1, 2, \ldots, K\}$ is a sequence of cubic B-spline basis functions, and

$$C(\theta) = \log\left(\int_L^R \exp(s(u;\theta))\mathrm{d}u\right) < \infty.$$

Then $f_{u|\Theta}(u)$ is a positive density on (L, R), twice continuously differentiable. The last basis function is left out for identification: B-splines sum to one, $\sum_{j=1}^{K} B_j(u) = 1$ for all $u \in (L, R)$.

Because $\mathscr{V}_{it} = u_{it} - \varepsilon_{it}$, the log likelihood of the observed residuals is

$$
\begin{aligned}
l_{\mathscr{V}}(\theta) &= \sum_{i=1}^{n} \log\Big(\int f_{\varepsilon}(u - \mathscr{V}_i) f_{u|\theta}(u)\mathrm{d}u\Big) \\
&= \sum_{i=1}^{n} \log\Big(\int f_{\varepsilon}(u - \mathscr{V}_i) \exp(s(u;\theta))\mathrm{d}u\Big) - nC(\theta).
\end{aligned}
$$

To avoid overfitting, I subtract a roughness penalty from the log likelihood (Eilers and Marx 1996):

$$
\text{Penalty}(\theta) = \lambda \cdot \|\mathbf{D}\theta\|^2,
$$

where $\mathbf{D}$ takes third-order differences of adjacent coefficients. Each coefficient weights a B-spline that overlaps only its neighbours: large differences between adjacent coefficients make the log-density $s(u;\theta)$ wiggle, and the penalty keeps it smooth. The estimate $\hat{\theta}$ maximizes $l_{\mathscr{V}}(\theta) - \text{Penalty}(\theta)$.

The penalty on third-order differences leaves quadratic log-densities unpenalized: with cubic B-splines, it conserves the mean and variance of a density fitted directly to data for any λ, and the fit tends to a normal density as λ grows (Eilers and Marx 1996, 93, 99).

I estimate each industry separately. The density of measurement error, f_{ε}, is a kernel density of the corporations' residuals, with the Sheather–Jones plug-in bandwidth and exponential tails beyond the range of the data. The support of u is $[0, \max \mathscr{V} + 0.1]$, which imposes $u \geq 0$. The spline has eight interior knots at equally spaced percentiles of the non-negative values of $\mathscr{V}$, and the penalty weight is $\lambda = 10$. Adaptive 15-node Gauss–Legendre quadrature computes the integrals in the likelihood, and BFGS maximizes the penalized likelihood.