Supplemental Appendix

Downstream Spillovers in Value Added Tax Enforcement: Evidence from China

Haonan Li
School of Public Finance and Taxation
Southwestern University of Finance and Economics
haonanli@swufe.edu.cn

Xuan Wang
National School of Development and
China Center for Economic Research
Peking University
xuanwang@nsd.pku.edu.cn

A Land Use Policy in China **A-2**

B Sources of Fake Invoices and Business Relevance Across VAT Systems **A-3**
B.1 Common Sources of Fake Invoices in VAT Systems A-3
B.1.1 Fake Invoices in Standard VAT Systems A-3
B.1.2 Policy Created Weak Verification Nodes A-5
B.1.3 The European Farmer Flat-Rate Scheme A-6
B.2 Business Relevance Requirements for Input Credits A-6

C Conceptual Framework: Optimal Tax Evasion Strategies of Firms **A-8**
C.1 Model Setup A-8
C.2 Under Standard VAT Regime A-9
C.3 The Presumptive VAT Reform A-12

D Measuring Suspicious Invoice Use **A-14**

E Voluntary VAT Registration **A-17**

F Alternative Treatment Assignments for Downstream Spillovers **A-18**
F.1 Treatment Assignment Using Transaction-Level VAT Invoice Data A-18
F.2 Cleaner Control Groups A-20

G Substitution to Other Sources of Fake Invoices **A-22**

H Sensitivity Analysis of Welfare Calculation **A-25**
H.1 Administrative Cost Sensitivity A-25
H.2 Exit-Related VAT and Production Losses A-26

I Supplementary Figures and Tables **A-27**

A Land Use Policy in China

The main reasons for the VAT treatment in China's agro-processing industry are the country's land use policy and the structure of its agricultural sector. China's agriculture is dominated by small-scale farming, driven by land use regulations and family-based production. Under the Household Responsibility System, adopted in the late 1970s, land ownership remained with the state or rural collectives, while individual households were allowed to cultivate small plots. This system resulted in fragmented and small-scale farming operations, where agricultural production remains reliant on smallholders and is largely self-sufficient. Most farmers do not operate as formal businesses, which means they do not issue VAT-compliant invoices.

Furthermore, fragmented land ownership complicates large-scale farming. Farmland allocated to households is often non-contiguous, making consolidation and scalability difficult. Although reforms have eased land transfer constraints in recent years, progress remains slow. In summary, the prevalence of small, informal farms in China's agricultural sector makes it difficult for agro-processing firms to obtain VAT-compliant invoices from farmers.

B Sources of Fake Invoices and Business Relevance Across VAT Systems

This appendix provides additional institutional background for the external validity discussion in the main text. Section I and the introduction explain the basic mechanism. The downstream propagation mechanism arises from two features of the VAT chain. First, input tax credit supporting documentation is weakly verified at certain nodes in the VAT chain. This loophole creates the supply of fake invoices and an invoice market for downstream firms. Second, downstream firms need input credit claims to appear consistent with their own business activities, which creates demand for invoice laundering. In this laundering process, intermediaries repackage low cost fake invoices into business consistent input purchases. Here we provide additional examples showing that these two features are not unique to China but arise in VAT systems more broadly. Therefore, downstream propagation may occur more widely across invoice credit VAT systems in other countries.

B.1 Common Sources of Fake Invoices in VAT Systems

B.1.1 Fake Invoices in Standard VAT Systems

We begin with fake invoice opportunities that can arise in standard invoice credit VAT systems. One common example is "flying invoices," which the main text discusses as a source of tradeable VAT documentation. The key point is that an invoice can have tax value beyond its role as a transaction record. In a standard VAT system, an invoice allows the buyer to claim input credits. If the original buyer has little use for the documentation, while another firm can use it to reduce VAT remittances or claim refunds, the invoice creates a tax benefit that can be traded or repackaged.

A simple example illustrates this valuation gap. Suppose an invoice records a purchase of value X and the VAT rate is τ. A VAT-registered firm with sufficient output tax liability may be willing to pay up to roughly τX for documentation that supports an input credit claim. By contrast, a final consumer, an unregistered buyer, or a buyer already in an excess credit position may place little value on the same documentation. This difference in valuation creates room for an intermediary market in invoices.

This is the basic mechanism through which flying invoices arise. In practice, they can come from at least two sources. The first source starts with invoices issued in business-to-business transactions but not fully used by the original business buyer. For example, a manufacturer may sell goods and issue a VAT invoice to a retailer. If the buyer already has enough input credits, has low output tax liability, or is unregistered, the invoice has little marginal value to that buyer. By contrast, the same invoice may be valuable to another VAT firm that needs deductible input documentation. This creates scope for the invoice to be diverted to another firm through intermediaries.

The second source arises at the retail stage. Final consumers usually cannot claim input credits and often do not request VAT invoices. This leaves retailers with room to issue invoices that are not demanded by the actual buyers. A retailer can then transfer these invoices to VAT firms that need input credit documentation. Because the underlying sale would have been reported anyway, the retailer can earn an extra fee while remitting roughly the same VAT. Because these patterns can be embedded in legal transaction records, it is difficult for tax authorities to detect them through invoice matching alone, and they are common concerns in VAT systems around the world.

Invoice mills provide another source of tradeable documentation. These entities issue invoices without corresponding real activity or without remitting the associated output tax, and may disappear before the tax authority can verify the underlying transactions. Carousel fraud in the European VAT system provides a related example. In such schemes, invoices can support downstream input credit or refund claims even when the corresponding upstream output tax is not ultimately remitted, often because a missing trader disappears before paying the tax (Keen and Smith, 2006). Although the institutional details differ from our setting, the common feature is that invoices may become detached from seller side remittance and real transaction activity. These examples illustrate a general detection problem. Invoice verification can confirm that a document exists and appears consistent across records, but it cannot by itself establish that the documented transaction is genuine or economically coherent. This means that tradeable invoices can be widely available even in standard VAT systems.

B.1.2 Policy Created Weak Verification Nodes

A second source of fake invoices comes from policy design. VAT systems often adopt simplified rules when strict seller side verification is administratively costly, especially when suppliers are small, numerous, outside the ordinary VAT system, and difficult to integrate into the standard invoice credit chain. These policies are not designed as evasion devices. They are usually introduced to reduce compliance costs or keep the VAT system neutral when the standard invoice credit mechanism is difficult to apply.

The enforcement relevant feature is that these policies may preserve some buyer side VAT treatment while relaxing the usual transaction-level link between seller remittance and buyer credit. Second-hand goods regimes provide a common example. The special feature of second-hand goods regimes is that second-hand goods are often bought from sellers who are not regular VAT taxpayers, such as private consumers or unregistered sellers. Because these sellers do not charge and remit output VAT in the ordinary way, the purchase does not naturally create a standard VAT invoice that links the seller's tax payment to the buyer's input credit. Even so, some VAT systems allow registered buyers to receive input tax relief or similar credit treatment for qualifying second-hand goods purchases, as long as they meet the relevant documentation requirements. Australia, New Zealand, and several Southern African countries, including South Africa, Botswana, Zimbabwe, and Namibia, have adopted such rules. These regimes differ from China's pre-reform agricultural self-billing rule, but they illustrate a similar administrative tradeoff. The tax system may allow buyers to receive input credit treatment for purchases from sellers who do not issue standard VAT invoices. This weakens the usual transaction-level link between seller remittance and buyer credit.

China's pre-reform agricultural self-billing rule is the empirical setting we study because it provides a particularly transparent version of this problem. Agro-processing firms purchased from smallholder farmers who were outside the standard VAT system and generally could not issue VAT invoices. Allowing processors to self-bill agricultural inputs reduced compliance and administrative costs, but it also gave buyers unusual control over the documentation supporting input credit claims. The 2012 presumptive tax reform sharply restricted this source of weakly verified credits, which allows us to identify how an enforcement shock at

such a node propagates downstream.

B.1.3 The European Farmer Flat-Rate Scheme

The European farmer flat-rate scheme illustrates the same agricultural administrability problem in a mature VAT system. As Cnossen (2018) explains, when VAT was introduced in Europe, many farmers were small and kept limited accounts. Fully integrating all farmers into the ordinary invoice credit chain would have imposed substantial compliance and administrative costs. At the same time, simply exempting farmers would leave input VAT embedded in farm prices, rather than being credited through the ordinary credit mechanism.

The farmer flat-rate scheme addresses this trade-off through a standardized compensation mechanism. Eligible farmers may charge a flat-rate compensation set by the government on sales to taxable purchasers. The compensation is intended to approximate the VAT paid on farmers' own inputs, and the taxable purchaser may deduct the corresponding amount from its VAT on sales (Cnossen, 2018).

This feature is closely related to our setting. In both settings, small agricultural suppliers are not fully integrated into the standard invoice credit chain, while downstream agro-processing firms retain some form of buyer side VAT treatment. This relaxes seller side verification relative to the textbook VAT benchmark, where the buyer's credit is backed by the seller's VAT remittance for the same transaction. Cnossen (2018) notes that the flat-rate scheme may invite farmers and processing firms to collude by artificially increasing farmers' sales. If the compensation is calculated as a percentage of reported sales, overstating the transaction value raises both the farmer's compensation and the purchaser's deduction, without generating an additional VAT liability on the seller side.

B.2 Business Relevance Requirements for Input Credits

In invoice credit VAT systems, input credit claims generally require more than a formally valid invoice: purchases must be connected to the taxpayer's taxable sales or normal business activities. This requirement limits the direct use of invoices from unrelated sectors. Table B.1 shows that such business relevance requirements are common across VAT systems.

Table B.1: Business Relevance Requirements for VAT Input Credits

Country/Region	Business Relevance Requirement	Source
China	The tax authority requires the underlying transaction to be genuine, to have a reasonable commercial purpose, and to be consistent with the taxpayer's actual business operations.	State Taxation Administration
United Kingdom	HMRC requires purchased goods or services to have a direct and immediate link with the taxpayer's taxable supplies. Without such a link, input VAT is generally not deductible.	HM Revenue & Customs
European Union	Under the EU VAT Directive, input VAT is deductible only when goods or services are used for the purposes of taxed transactions. The case-law-based standard also requires a direct and immediate link with taxable transactions.	European Union VAT Directive
Canada	CRA requires purchases and expenses to be related to the taxpayer's commercial activities. It also states that a purchase or expense must be reasonable in quality, nature, and cost in relation to the nature of the business.	Canada Revenue Agency
Singapore	IRAS requires input tax to be attributable to taxable supplies, meaning that the purchase should form a cost component of the firm's taxable business activity.	Inland Revenue Authority of Singapore
Australia	Input tax credits are allowed only for business related acquisitions, meaning that the purchase must be sufficiently connected to the taxpayer's business activities.	Australian Taxation Office

Note: This table summarizes official rules or guidance on the business relevance requirement for claiming VAT input credits.

Although China's pre-reform agricultural self-billing rule is uncommon, the downstream spillover mechanism is more general. Fake invoices can arise from multiple sources, but downstream firms need invoices that appear connected to their business activities. These conditions create a role for invoice laundering through intermediaries or transaction chains that repackage upstream fake credits into invoices consistent with downstream buyers' business activities. By disrupting the supply of fake invoices that supports this process, upstream enforcement can generate downstream compliance spillovers along the VAT chain.

C Conceptual Framework: Optimal Tax Evasion Strategies of Firms

In this section, we characterize the optimal tax evasion strategies and input choices of agro-processing firms under a standard VAT system and a presumptive VAT regime. Firms take product and invoice prices as given, choose their production level, and decide how much to evade and whether to sell fake invoices to downstream buyers. The framework describes the upstream supply of fake invoices that initiates the downstream propagation mechanism.

C.1 Model Setup

Consider a continuum of firms that differ in productivity A, distributed over $[\underline{A}, \overline{A}]$. A firm produces output $y = Af(x)$ using input $x \geq 0$, where $f(0) = 0$, $f' > 0$, $f'' < 0$, and $\lim_{x\to\infty} f'(x) = 0$. The output price is normalized to 1, while the input price is $p_x > 0$. Prices and invoice values are measured on a tax-inclusive basis where applicable, and $\tau \in (0, 1)$ is the corresponding tax fraction. Domestic firms pay VAT on actual taxable value added, $z(x) = Af(x) - \theta p_x x$, where $\theta \in [0, 1]$ is the share of deductible inputs. We consider firms with nonnegative actual taxable value added. In the absence of evasion, the firm's after-tax profit is:

$$R_{NE} = y - p_x x - \tau(y - \theta p_x x) \tag{C.1}$$

Fake Invoices Firms may use fake invoices to overreport deductible input value by $e \geq 0$, generating additional input tax credits of τe. All firms face the same evasion cost $g(e)$, with $g(0) = g'(0) = 0$, $g''(e) > 0$, and $\lim_{e\to\infty} g'(e) = \infty$. These conditions ensure a finite optimal evasion choice. Under a standard VAT, as illustrated in Figure 1, two uses of fake invoices are possible:

(1) **Self-Use of Fake Invoices:** The firm uses additional input credits to reduce its VAT liability. Until current VAT payable is fully offset, the marginal benefit per unit of fictitious input value is τ. Excess credits can be carried forward with discounted value $\rho\tau$ per unit, where $\rho \in [0, 1)$ for domestic firms.

(2) **Selling Fake Invoices:** The firm issues invoices for fictitious sales and earns a fee of $p_r\tau$ per unit of invoice value, where $p_r \in (0,1)$ is the fee as a fraction of the associated tax credit. It uses self-billed fake input invoices to offset the output tax liability generated by these sales.

Since $\rho < 1$ and $p_r < 1$, firms allocate fake credits to their own current liability before selling invoices or carrying credits forward. Once this liability is fully offset, selling is preferred when $p_r > \rho$; otherwise, the firm carries the excess credits forward. The resulting benefit depends on both evasion and actual taxable value added:

$$E(e,x) = \begin{cases} \tau e & if \quad e \le z(x) \\ \tau z(x) + p_r\tau(e - z(x)) & if \quad e > z(x), p_r > \rho \\ \tau z(x) + \rho\tau(e - z(x)) & if \quad e > z(x), p_r \le \rho \end{cases} \tag{C.2}$$

Exporters For zero-rated exporters, additional input tax credits can be refunded immediately and in full. The marginal benefit of using fake invoices to claim refunds is therefore τ, corresponding to $\rho = 1$, which exceeds the return from selling them, $p_r\tau$. Exporters thus prefer to use fake invoices to inflate their own refund claims rather than resell them. At an interior optimum, their evasion choice satisfies $g'(e) = \tau$.

C.2 Under Standard VAT Regime

Under the standard VAT, a domestic firm chooses input x and evasion e to maximize after-tax profit:

$$\pi^s(x,e) = y - p_x x - \tau(y - \theta p_x x) + E(e,x) - g(e). \tag{C.3}$$

Let $q = \max\{p_r, \rho\} < 1$. The firm's first-order conditions account for both the direct tax consequences of production and the alternative value of fake credits.

Optimal Input Choice Let λ^s denote the effective marginal tax on actual taxable value added. At the optimum, $\lambda^s = \tau$ when $e^* < z$, $\lambda^s = q\tau$ when $e^* > z$, and $\lambda^s = g'(z)$ when $e^* = z$, with $z = z(x^*)$. For an interior input choice, the first-order condition is:

$$\frac{Af'(x^*)}{p_x} = \frac{1 - \theta\lambda^s}{1 - \lambda^s} = 1 + \frac{\lambda^s(1-\theta)}{1-\lambda^s} \;\equiv\; 1 + \tau_E, \tag{C.4}$$

where $\tau_E = \lambda^s(1-\theta)/(1-\lambda^s)$ is the effective production wedge under standard VAT. For firms that have not fully offset their own liability, this reduces to $\tau(1-\theta)/(1-\tau)$. For firms with excess credits, it becomes $q\tau(1-\theta)/(1-q\tau)$: using credits to cover additional output tax forgoes their resale income or carry-forward value. A higher marginal wedge discourages input use when $\theta < 1$. At a zero-input optimum, the corresponding marginal-profit condition is nonpositive.

Optimal Evasion Conditional on $z = z(x^*)$, optimal evasion satisfies:

$$\begin{cases} g'(e^*) = \tau, & e^* < z, \\ q\tau \leq g'(z) \leq \tau, & e^* = z, \\ g'(e^*) = q\tau, & e^* > z. \end{cases} \tag{C.5}$$

The middle case allows the firm to stop exactly where its current VAT liability reaches zero. The two domestic invoice-use strategies and the exporter case are illustrated in Figure C.1:

(1) **Scenario One** (Panel (a)): Firms choose $e^* \leq z$ and use fake invoices only for their own deductions. This group includes firms that stop exactly at $e^* = z$, fully offsetting their liability without selling any surplus invoices.

(2) **Scenario Two** (Panel (b)): Optimal evasion exceeds actual taxable value added ($e^* > z$). Firms either carry excess credits forward or sell fake invoices, depending on which yields a higher marginal return. When $p_r > \rho$, firms reduce their own VAT remittance by τz and sell invoices with fictitious sales value $e^* - z$.

(3) **Scenario Three** (Panel (c)): Exporters whose additional input credits are refunded immediately and in full face $\rho = 1 > p_r$. They use fake invoices to inflate their own export tax refunds rather than reselling them.

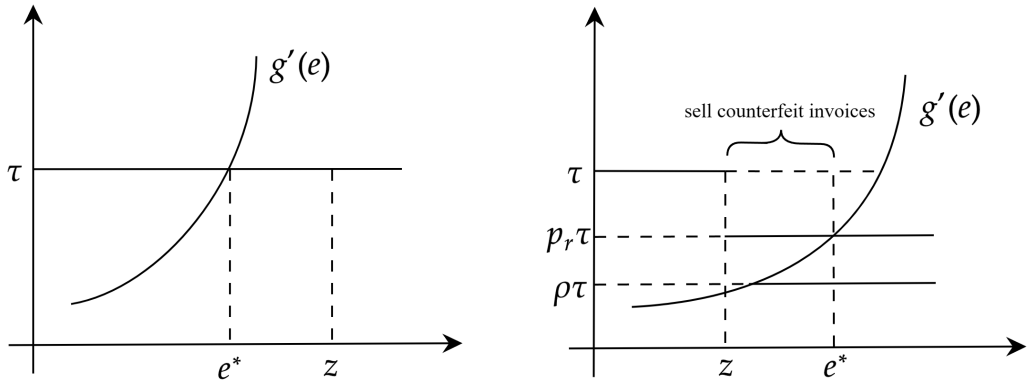
Implications for Firm Size and Evasion Strategies For the firm-size comparison, assume that optimal real output and actual taxable value added increase with productivity, and consider the case $p_r > \rho$. Define $\bar{e} = (g')^{-1}(p_r\tau)$. Firms with $z < \bar{e}$ choose $e^* = \bar{e}$,

fully offset their own liability, and sell the surplus $\bar{e} - z$. Firms with $z \geq \bar{e}$ use fake invoices only for their own deductions. With common evasion costs, smaller firms are therefore more likely to sell fake invoices. Zero VAT remittance is necessary for selling in this framework but is not sufficient, since firms at $e^* = z$ sell no surplus. An invoice mill is the limiting case $A = 0$ and $x = 0$, with no real production and $z = 0$; its earnings come entirely from issuing and selling fake invoices.

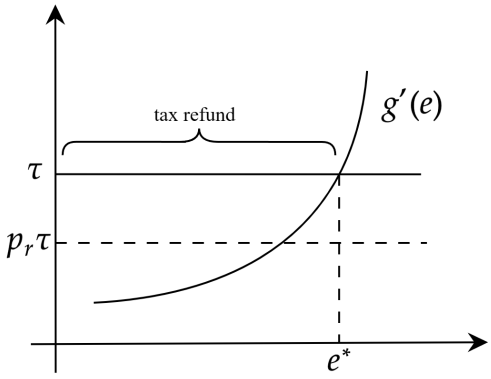
Figure C.1: Strategies for Using Fake Invoices in Different Scenarios

(a) Scenario One: Self-Bill Fake Invoice to Reduce Firm's Own VAT Liability when $e^* \leq z$

(b) Scenario Two: Sell Invoice to Downstream Firms when $e^* > z$ and $p_r > \rho$



(c) Scenario Three: Exporters Claim Refund



Note: These figures illustrate the firm's uses of fake invoices. Panel (a) shows firms that use fake invoices only for their own deductions, with $e^* \leq z$. This group also includes the boundary choice $e^* = z$. Panel (b) shows firms with $e^* > z$ and $p_r > \rho$, which sell invoices with fictitious sales value $e^* - z$ to downstream firms. Panel (c) shows exporters whose additional input credits are refunded immediately and in full; these firms use fake invoices to inflate their own refund claims.

C.3 The Presumptive VAT Reform

A core feature of the presumptive VAT reform is that it replaces invoice-based deductions with a presumptive deductible amount, μy, where $\mu \in [0, 1)$ is the input-output ratio determined by the tax authorities. Overreporting input purchases no longer increases the firm's deductions. For each unit of fictitious sales, the additional output tax is τ, while any additional presumptive credit is at most $\mu\tau$. We consider a post-reform invoice price $0 < p_r^p < 1 - \mu$, so the selling fee is below the additional net tax liability even before selling costs. Selling fake invoices is then unprofitable. The firm's remaining profit function is:

$$\pi^p(x, e) = y - p_x x - \tau(1 - \mu)y - g(e). \tag{C.6}$$

Optimal Evasion Since fictitious input purchases no longer reduce tax liabilities and $g(e)$ is minimized at zero, the optimal choice is $e^p = 0$. Eliminating the pre-reform fictitious purchases reduces the reporting component of input tax credits by τe^*. For an invoice seller, eliminating fictitious sales also reduces the reporting component of output tax by $\tau(e^* - z(x^*))$. Total changes in reported taxes additionally reflect real production and the change in legitimate deductions. With an unchanged downward-sloping demand curve and incomplete replacement by other suppliers, the contraction in aggregate fake-invoice supply raises its market price and reduces downstream use, provided the resulting price remains below the profitability threshold above.

Optimal Input Choice Under the presumptive VAT system, the first-order condition for an interior input choice becomes:

$$\frac{Af'(x^p)}{p_x} = 1 + \frac{\tau(1 - \mu)}{1 - \tau(1 - \mu)} \equiv 1 + \tau_E^p, \tag{C.7}$$

where $\tau_E^p = \tau(1 - \mu)/[1 - \tau(1 - \mu)]$. The presumptive tax introduces an additional production distortion when this wedge exceeds the firm's pre-reform wedge τ_E in Equation (C.4). A higher μ lowers τ_E^p and mitigates this distortion.

Change in VAT Remittance Following the reform, the firm's VAT remittance changes due to: (1) changes in real production, (2) the mismatch between the presumptive deduction ratio and actual input ratio, and (3) reduced fake invoice activity.

Let $\Delta T = T^p - T^s$ denote the change in current-period VAT remittances for a domestic firm. Before the reform, $T^s = \tau \max\{z(x^*) - e^*, 0\}$; after the reform, $T^p = \tau(1-\mu)Af(x^p)$. The change can be decomposed as:

$$\Delta T = \underbrace{\tau\big(z(x^p) - z(x^*)\big)}_{\text{Real Response}} + \underbrace{\tau\big(\theta p_x x^p - \mu A f(x^p)\big)}_{\text{Deduction Ratio Bias}} + \underbrace{\tau \min\{e^*, z(x^*)\}}_{\text{Reporting Response}}. \tag{C.8}$$

The first component captures the change in actual taxable value added, $z(x^p) - z(x^*)$. When $\tau_E^p > \tau_E$, the higher production wedge reduces input use and output, lowering this component. A higher μ mitigates the production distortion, since $\partial \tau_E^p / \partial \mu < 0$.

The second component compares the deductions allowed under the two regimes at the post-reform input choice. For $x^p > 0$, define $\theta_p = \mu A f(x^p)/(p_x x^p)$; this component then equals $\tau p_x x^p(\theta - \theta_p)$. A more generous presumptive deduction, $\theta_p > \theta$, lowers VAT revenue, while $\theta_p < \theta$ raises it. Increasing μ reduces the production wedge but also lowers revenue through this deduction margin, holding production fixed. The expression in Equation (C.8) also applies at $x^p = 0$, when the deduction-bias term is zero.

The final component is the recovery of current-period VAT previously offset through fake input claims. For firms with $e^* \leq z(x^*)$, this reporting component is τe^*. For firms with $e^* > z(x^*)$, it is capped at $\tau z(x^*)$. Excess credits used to support fictitious sales or carried forward do not further reduce the supplying firm's current-period remittance. Their withdrawal can nevertheless affect downstream firms or future tax payments.

Whether the presumptive tax reform raises VAT remittance depends on the reporting gain relative to any revenue losses from production distortions and more generous presumptive deductions.

D Measuring Suspicious Invoice Use

This section describes how we identify the use of suspicious invoices in the administrative data. Following Li and Wang (2026), we identify suspicious invoice use by comparing VAT deductible inputs reported in VAT returns with total costs reported in corporate income tax (CIT) returns. VAT deductible inputs have a narrower scope and stricter documentation requirements. They generally require compliant VAT invoices and exclude items such as wages and interest. CIT total costs cover a broader set of expenses and can be supported by a wider range of accounting documents. Therefore, under truthful reporting, VAT deductible inputs should not exceed CIT total costs. We define an indicator for suspicious invoice use equal to one when VAT deductible inputs exceed CIT total costs.

Institutional features make this comparison informative in our setting. VAT and CIT were handled by separate arms of the tax authority, information sharing across tax types was costly, and systematic data integration began only after 2018. Cross-checking was therefore limited during our sample period, allowing discrepancies between VAT and CIT reports to persist. Li and Wang (2026) document these barriers and the late integration of tax data in China. Their descriptive evidence shows weak coordination across taxes. Only a small share of VAT evasion cases result in CIT penalties, and only a small share of CIT evasion cases result in VAT penalties. Even for the Urban Maintenance and Construction Tax, which is mechanically tied to VAT, penalties are not universal. Similar gaps exist in other settings. Almunia et al. (2024) report large mismatches between customs and VAT records in Uganda.

Some discrepancies between VAT deductible inputs and CIT total costs may arise even without tax evasion, which may affect the identification of suspicious invoice use. As shown in Table D.1, these discrepancies mainly come from three sources, including differences in the timing of deduction recognition, documentation requirements, and the scope of deductible items. The indicator may overstate suspicious invoice use when VAT deduction occurs before the corresponding CIT cost is recognized, as in the case of fixed assets or inventory. By contrast, the indicator becomes more conservative when invoicing is delayed relative to accrual-based cost recognition. In addition, differences in documentation requirements and deductible scope generally make the indicator more conservative, because CIT total costs

include some items that are not deductible for VAT purposes. These timing differences should be interpreted as accounting mismatches within a given period rather than permanent gaps. Over a longer horizon, earlier VAT deduction may be offset by lower VAT deductible inputs relative to CIT costs in later periods.

Table D.1: Discrepancies between VAT Deductible Inputs and CIT Total Costs

Source of Discrepancy	VAT Treatment	CIT Treatment	Implication for Fake Invoice Indicator
Source (i): Timing of Deduction Recognition	Invoice date	Accrual basis	(1) Invoice date < accrual-based recognition date: may overstate fake invoice use. (e.g., fixed assets, inventory turnover lag). (2) Invoice date > accrual-based recognition date: tends to make the indicator more conservative. (e.g., delayed invoicing).
Source (ii): Documentation Requirements	Requires compliant invoices	Can rely on broader supporting documents	Tends to make the indicator more conservative.
Source (iii): Scope of Deductible Items	Narrower, specifically defined deductible items	Broader deductible items (e.g., deductible business entertainment expenses, interest, wages)	Tends to make the indicator more conservative.

We first alleviate concerns that the indicator may overstate suspicious invoice use because of timing differences, especially for fixed assets and inventory. In these cases, VAT deductible inputs may temporarily exceed CIT total costs even without suspicious invoice use.

For inventory, VAT can be deducted when materials are purchased and a VAT invoice is issued, whereas the corresponding CIT cost is recognized only later when the inventory is sold. This timing gap may temporarily push VAT deductible inputs above CIT total costs even without fake invoices. To mitigate this concern, we restrict the sample to firms with higher inventory turnover, for which the lag between input purchases and accrual-based cost recognition should be shorter. Figure A.16 shows that the results remain similar, suggesting that our estimates are unlikely to be driven by inventory-related timing mismatches.

For fixed assets, VAT deductions are claimed in full in the year of acquisition, whereas the corresponding CIT cost is recognized gradually through depreciation. This timing difference may mechanically push VAT deductible inputs above CIT total costs even without fake

invoices. To mitigate this concern, we adjust total costs using Equation (D.1), where *Depr* denotes depreciation, *Fixed* denotes fixed asset investment, and *Adjusted Cost* is total cost after adjustment. This adjustment aligns the treatment of fixed asset investment more closely with the VAT deduction concept. In Figure A.17, Adj. I reports estimates based on this fixed asset correction, which remain close to the baseline estimates.

$$Adjusted\,Cost = Total\,Cost - Depr + Fixed \tag{D.1}$$

To further mitigate concerns about unobserved factors that may mechanically push VAT deductible inputs above CIT total costs, we also apply stricter thresholds for the suspicious invoice indicator. Specifically, we require the ratio of deductible inputs to total costs to exceed 102%, 104%, 106%, 108%, and 110%, respectively. Figure A.15 shows that the results remain similar.

We next consider factors that may make the indicator more conservative. Although these factors do not mechanically overstate suspicious invoice use, they may still introduce measurement noise and reduce the precision of the indicator. CIT total costs include several expenses that are not deductible for VAT purposes, such as business entertainment expenses, interest, and wages. Their inclusion makes it less likely that VAT deductible inputs exceed CIT total costs. To examine whether these components affect our results, we progressively exclude them from the total costs. In Figure A.17, Adj. II further excludes business entertainment expenses, Adj. III additionally excludes interest, and Adj. IV further excludes wages. The estimates remain close to the baseline estimates across these alternative definitions.

In short, comparing VAT deductible inputs with CIT total costs gives us a simple and transparent way to detect suspicious invoice use. Because several accounting differences tend to make this measure conservative, the indicator is more likely to understate than overstate the prevalence of suspicious invoices.

E Voluntary VAT Registration

The voluntary registration of small firms into the VAT system offers an additional perspective for understanding the role of fake invoices. Small firms below the VAT registration threshold are allowed to register voluntarily. Small firms opt into the VAT system because they can issue invoices to other VAT taxpayers or the government. The presumptive tax reform does not alter the incentives for VAT registration, except for those related to the use of fake invoices. It provides us with a good opportunity to present more evidence of how firms exploit loopholes.

We use the industry–province–year level voluntary registration rate as the outcome variable, defined as the proportion of below-threshold firms that registered as VAT taxpayers. Panel (a) of Figure A.4 shows that the reform reduces voluntary VAT registration among agro-processing firms below the official threshold by 16.4 percentage points. Panel (b) shows a similar decline among downstream wholesale and retail firms. Table A.6 further shows no significant changes in real production among below-threshold agro-processing firms. Taken together, these findings suggest that some small firms voluntarily registered as VAT taxpayers to take advantage of opportunities to use fake invoices. This observation again highlights that imperfectly enforced VAT can even be worse than the alternative sales tax.

F Alternative Treatment Assignments for Downstream Spillovers

This section presents additional robustness checks that use alternative downstream treatment assignments for the spillover analysis. In the baseline analysis, we map each reformed agro-processing industry to closely related wholesale and retail industries using official industry codes and descriptions. This approach provides a transparent baseline assignment, but it is necessarily coarser than the ideal exposure measure. A wholesale or retail firm may sell multiple agro-product categories while being registered under a single official industry code. As a result, the baseline mapping may not fully capture the firm's input linkages to all upstream agro-processing industries.

This concern can affect the estimation in two related ways. First, some industries classified as controls may in fact be partially exposed to the reform because firms in these industries purchase from reformed agro-processing industries or closely connected intermediaries. Second, some treated downstream industries may have been partially exposed before their assigned treatment year if they purchased from another upstream agro-processing industry that had already been reformed. In both cases, part of the true response would appear in the comparison group or in the pre-treatment period. This would be more likely to understate the true downstream spillover than to overstate it.

To address these concerns, we construct alternative downstream assignments. We use transaction-level VAT invoice data to measure observed input linkages between upstream agro-processing industries and downstream wholesale and retail industries.

F.1 Treatment Assignment Using Transaction-Level VAT Invoice Data

We use transaction-level VAT invoice data to construct a data-driven measure of downstream exposure. These data allow us to observe input linkages between upstream agro-processing industries and downstream wholesale and retail industries. Because of confidentiality restrictions, the invoice data are de-identified and do not contain firm names. We aggregate the linkage information to the four-digit industry level and calculate, for each downstream

industry, the share of agro-inputs sourced from each upstream agro-processing industry.

The main limitation of these invoice data is that they are available only for 2017, after the reform period. So we do not use them to replace the baseline treatment assignment. Instead, we use them as a validation exercise. This limitation is less severe in our paper because post-reform invoice flows are likely to be less contaminated by fake or laundered invoices, and they are closer to genuine commercial relationships. Before the reform, invoice flows could partly reflect the invoice laundering behavior that we studied. After the self-billing loophole was closed, observed VAT linkages are more likely to capture economically meaningful input-output relationships.

The key step in the invoice based assignment is to decide which upstream industries should be treated as a downstream industry's main suppliers. The VAT invoice data shows how much of the downstream agro-input purchases is covered under alternative cutoffs for defining main suppliers. The tradeoff is that a lower cutoff defines main suppliers more broadly and includes many minor suppliers, and a higher cutoff defines main suppliers more strictly and keeps only the largest suppliers. The data shows that lowering the cutoff below 20% adds little additional coverage, suggesting that the additional links are mostly weak. By contrast, raising the cutoff above 20% sharply reduces coverage, suggesting that economically important supplier relationships begin to be excluded. We therefore use 20% as the baseline cutoff because it separates major upstream suppliers from weaker input links in a transparent, data-driven way.

This cutoff is also consistent with the concentration of input sourcing in the invoice data. On average, the largest upstream source accounts for more than 40% of a downstream industry's agro-input purchases, while the second-largest source accounts for around 20%. Thus, an upstream agro-processing industry with an input share above 20% is typically one of the downstream industry's main suppliers, rather than a minor trading partner.

Using this rule, we classify a downstream industry as exposed to the reform once one of its main upstream agro-processing suppliers becomes treated. In the baseline invoice based assignment, an upstream agro-processing industry is defined as a main supplier if it accounts for more than 20% of the downstream industry's agro-input purchases. Figure A.8 reports the downstream spillover estimates using this invoice based assignment. The results are

similar to the baseline estimates. Downstream firms assigned to the treatment group remit more VAT after the reform, mainly because they claim fewer input tax credits and are less likely to use suspicious invoices. Figure A.9 further shows that the results do not rely on the 20% cutoff. The estimates remain similar when we vary the cutoff over a wide range.

F.2 Cleaner Control Groups

A remaining concern is that some control industries may be indirectly treated if industry classifications do not fully capture firms' business scope, or if VAT enforcement shocks propagate through more than one layer of the supply chain. We address this concern by constructing progressively cleaner control groups.

We use the transaction-level VAT invoice data to construct data-driven cleaner control groups. Direct exposure measures how strongly a downstream industry is directly connected to treated upstream agro-processing industries. We define direct exposure as the input share from treated upstream agro-processing industries. Next, we construct a first-order indirect exposure measure that captures connections through one additional layer of the supply chain. A control industry may not purchase directly from treated upstream agro-processing industries, but it may purchase from other wholesale and retail industries that are themselves directly exposed to treated agro-processing industries. For each downstream industry j, we calculate indirect exposure as:

$$Indirect\,Exposure_j = \sum_i s_{ji} \times Direct\,Exposure_i \tag{F.1}$$

where i and j are both wholesale and retail industries, s_{ji} is the input share from seller industry i to buyer industry j, and $Direct\,Exposure_i$ is seller industry i's direct exposure to treated upstream agro-processing industries.

We focus on this two-layer measure for two reasons. First, it is the most directly relevant margin for the mechanism in our paper. The invoice laundering channel requires invoices that look consistent with the buyer's business, and this consistency is strongest along immediate or closely connected product chains. More distant links would require a longer sequence of transactions and additional assumptions about how documentation is repackaged across industries. Second, because our invoice linkage data are de-identified and aggregated to the four-digit industry level, extending the exposure measure to many higher-order links

would quickly make the network dense. This would effectively classify all industries as indirectly exposed and leave no usable control group. Such an exercise would therefore be less transparent and harder to interpret. For this reason, we use the two-layer measure as a conservative robustness check rather than as a full model of all possible network propagation. If some higher-order exposure remains in the control group, it should move control industries in the same direction as treated downstream industries. This would make the baseline estimates more likely to understate the true downstream spillover than to overstate it.

We combine direct and indirect exposure into a total exposure measure:

$$Total\,Exposure_j = Direct\,Exposure_j + Indirect\,Exposure_j \tag{F.2}$$

This total exposure measure captures both direct connections to treated upstream agro-processing industries and indirect connections through one additional layer of the supply chain. We exclude industries with high total exposure from the control group. As shown in Figure A.10, the results remain similar after this restriction.

Across these alternative assignments and cleaner control groups, the downstream results remain similar. These checks support the interpretation that the baseline mapping is a noisy but informative proxy for downstream exposure. They also alleviate the concern that the baseline spillover estimates are driven by indirectly affected industries remaining in the control group.

G Substitution to Other Sources of Fake Invoices

A natural concern is that downstream firms may substitute toward other sources of fake invoices after the reform restricts access to invoices originating from treated upstream agro-processing firms. If such a substitution were substantial, the reform would mainly shift evasion to other invoice sources rather than reduce evasion itself. We therefore examine whether there is evidence of substantial substitution after the reform.

We first use real world cases to select plausible substitution channels. We summarize the expenditure categories that appear most frequently in publicly documented fake invoice cases from court records. Agricultural products account for a substantial share, but the documented cases also involve other categories, including transportation services, construction services, precious metals, advertising services, logistics auxiliary services, and processing and repair services.

Guided by the case evidence above, we first examine expenditure categories that are observable in the data and could plausibly be used for substitution. If affected downstream firms responded to the loss of agro-processing invoices by substituting toward these alternative invoice sources, we would expect them to report higher purchases or deductible inputs in these categories after the reform. In the main analysis, we find no significant increase in transportation service or processing service expenditures after the reform, as shown in Tables 3 and A.13. We further examine other observable categories that frequently appear in documented fake invoice cases, including construction services, advertising services, logistics auxiliary services, and repair services. We again find no significant increases in these expenditures. These results provide no evidence of substantial substitution toward these observable alternative expenditure channels.

We next test whether industries that could serve as alternative fake invoice suppliers show patterns consistent with downstream firms shifting fake invoice demand away from treated agro-processing firms. First, some alternative fake invoice channels, such as precious metals, may not be directly observable in downstream expenditure categories. Second, substitution may also be reflected in the behavior of industries that could supply alternative fake invoices, so supplier side patterns provide an additional way to check for this type of substitution.

If downstream firms substituted away from treated agro-processing firms toward untreated agro-processing industries, those untreated industries should show patterns consistent with expanded invoice supply after the reform. Specifically, reported output tax should increase as more invoices are issued to buyers, and input tax credit may also increase as firms offset the associated output tax liabilities.

We examine two types of potential substitute supplier industries. The first is the untreated agro-processing industries. A natural substitution margin is that, after the self-billing loophole was closed in treated agro-processing industries, downstream firms may have shifted demand for fake invoices toward untreated agro-processing industries where the same self-billing arrangement remained available. If this occurred, untreated agro-processing industries should exhibit increases in reported output tax and possibly input tax credit. To test this possibility, we compare untreated agro-processing industries with firms in the four adjacent two-digit sectors and estimate event-study specifications for input tax credit and output tax. The result does not show similar patterns of increased supply of suspicious invoices.

The second type consists of non-agro industries that may serve as alternative fake invoice channels. Guided by the case evidence, we focus on the transportation industry, the business services industry, and the precious metals processing industry. The business services industry includes categories such as advertising services and logistics auxiliary services. Transportation and business service expenses are common and plausible for many firms, while the underlying service transactions may be difficult to verify. The precious metals processing industry is also consistent with the logic of fake invoice transactions because gold-related transactions are high value, standardized, and relatively easy to transport. These features may make it easier to separate invoice flows from underlying goods flows and to disguise transactions without real economic substance.

Within each of these industries, we identify firms that are more likely to act as invoice sellers using two pre-reform indicators also used in the main analysis. The first captures a lack of real economic activity and is measured by electricity consumption relative to sales. The second captures limited VAT remittance and is measured by VAT remittance relative to assets. For each measure, we compare firms with zero values to firms above the median of the same measure. If fake invoice supply expanded after the reform, firms with zero

pre-reform values should display larger increases in input tax credit or output tax relative to firms above the median. We test this prediction using event-study specifications for the input tax credit and the output tax for the transportation industry, the business services industry, and the precious metals processing industry, respectively. Across these industries, we do not observe clear post-reform increases in either input tax credit or output tax among firms more likely to act as invoice sellers. These results provide no evidence that the reform induced a systematic expansion of invoice supply activity in these alternative industries.

For brevity, we do not report the results of these additional substitution tests, which are available from the authors upon request. Taken together, the evidence does not support substantial substitution toward other fake invoice sources. Affected downstream firms do not increase expenditures in categories commonly associated with fake invoices, and industries that could supply alternative fake invoices do not show patterns consistent with expanded invoice supply. At a minimum, the increase in downstream VAT remittances implies that any remaining substitution was not large enough to offset the reform's overall effect. These findings support the interpretation that the downstream spillover estimates reflect a reduction in credits supported by fake invoices rather than a mere reallocation of evasion to other invoice sources.

H Sensitivity Analysis of Welfare Calculation

This appendix examines the sensitivity of the welfare calculation in Section VI to alternative assumptions about administrative costs and exit-related VAT revenue and production losses. The main text reports a benchmark calculation in which the presumptive tax reform generates large revenue gains relative to administrative and production distortion costs. Because administrative costs and exit-related losses are difficult to measure precisely, we evaluate how the welfare conclusion changes under alternative assumptions. Even under very conservative assumptions, the additional administrative and exit-related costs remain small relative to the reform's revenue gains.

H.1 Administrative Cost Sensitivity

Our benchmark administrative cost calculation assumes that tax personnel assigned to the presumptive tax program devote 30% of their time to reform-related tasks, such as setting and updating deduction ratios. We also allocate the relevant tax authority personnel to the reform using the reformed industries' share of total output. In the benchmark, this share is 2.3%. Together, these assumptions imply a benchmark administrative cost of approximately \$4.1 million.

We conduct four sensitivity checks for the administrative cost calculation. First, we vary the assumed share of staff time devoted to the reform. The benchmark assumes that relevant tax personnel devote 30% of their time to reform-related tasks. We allow this share to vary over a wide range and recompute the threshold value of v required for $\Delta W > 0$. As shown by the black solid line in Panel (a) of Figure A.18, the welfare conclusion remains very similar.

Second, we replace the benchmark staff allocation proxy with alternative measures. The benchmark allocates relevant tax personnel to the reform using the reformed industries' share of total output. We replace this proxy with two alternatives, the VAT revenue share of reformed industries and the share of treated province-industry pairs among all province-industry pairs. As shown by the blue short-dashed and red long-dashed lines in Panel (a) of Figure A.18, the results remain very similar across these alternative allocation proxies.

Third, we allow for additional fixed administrative costs. The benchmark calculation focuses on the wage cost of tax officials, but implementing the reform may also involve fixed

costs not captured by staff time. One example is the purchase of third-party assessment services to help determine deduction ratios. To approximate this component, we use the average value of government procurement projects for similar assessment services, about $3.72 million. As shown by the orange dashed line in Panel (a) of Figure A.18, adding this component does not materially change the welfare conclusion.

Fourth, we allow the reform to reduce administrative costs. Under the pre-reform self-billing system, tax authorities had to verify firms' purchase records and audit potentially fabricated transactions. By replacing self-billed input claims with standardized deduction ratios, the reform may have reduced the cost of invoice verification and monitoring. We therefore allow the net change in administrative costs to be negative. For example, a value of -30% corresponds to administrative savings equal to 30% of the relevant wage bill. When administrative costs decline, the threshold value of v required for $\Delta W > 0$ becomes lower.

H.2 Exit-Related VAT and Production Losses

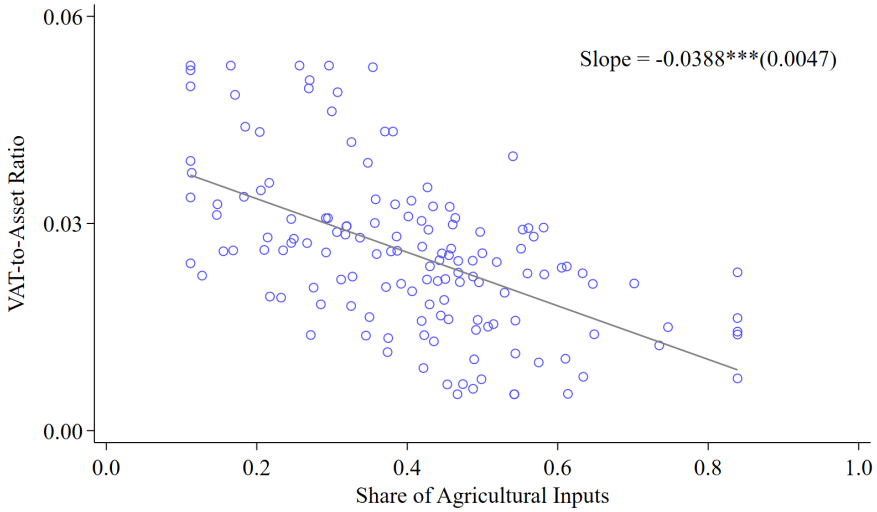
We next examine how the welfare calculation changes under alternative assumptions about the VAT revenue and production losses associated with reform-induced firm exits. In the main text, we treat all estimated exit-related VAT revenue and production losses as fully realized. This is conservative because some exiting firms may have relied more on fake invoices than on real production. If so, their exit would not imply an equal loss of true economic activity.

To address this concern, we vary the fraction of estimated exit-related VAT revenue and production losses that is treated as actual losses. The benchmark calculation includes the full estimated losses. We then consider alternatives in which only part of these losses is realized, allowing for the possibility that some exiting firms relied more on evasion activities than genuine production.

As shown in Panel (b) of Figure A.18, when only part of the estimated exit-related losses is realized, the implied welfare gain becomes larger. Thus, including the full estimated losses makes the benchmark conservative.

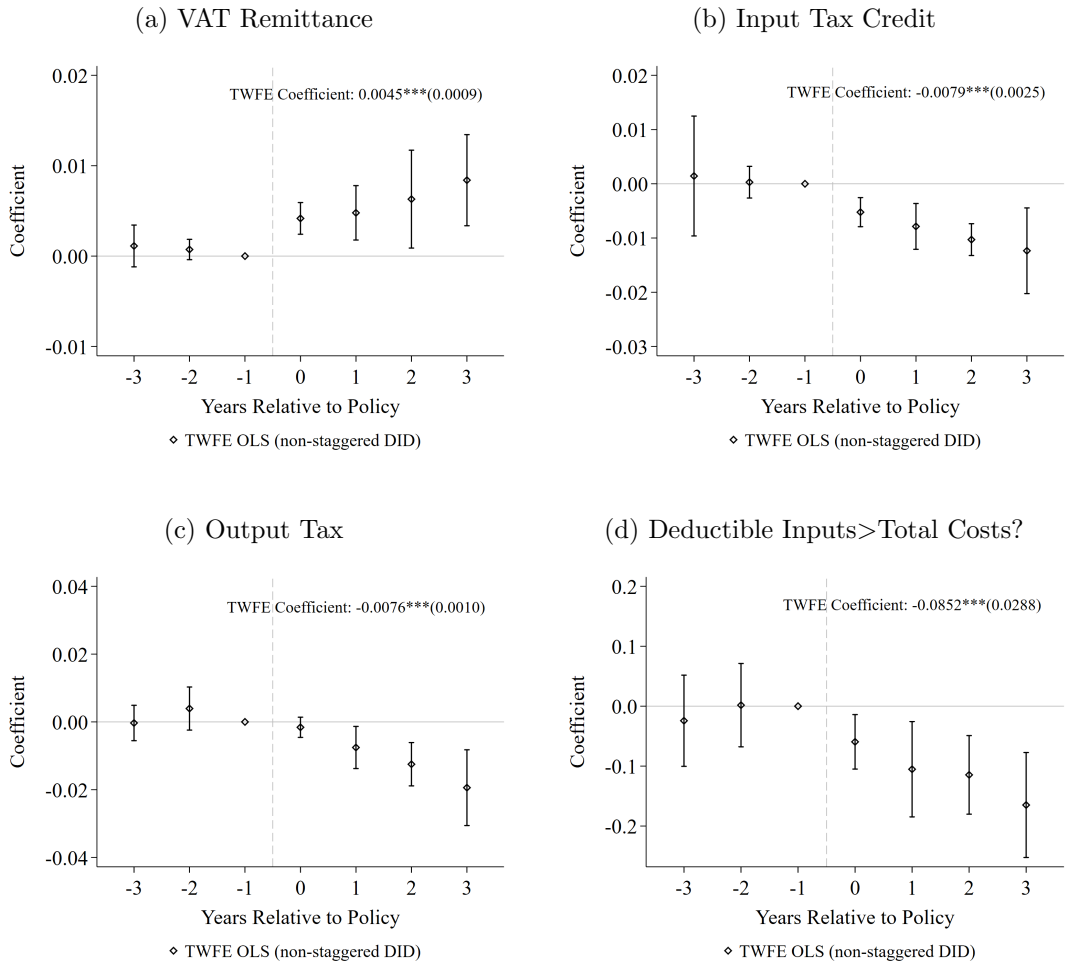
I Supplementary Figures and Tables

Figure A.1: Share of Agricultural Inputs and VAT-to-Asset Ratio



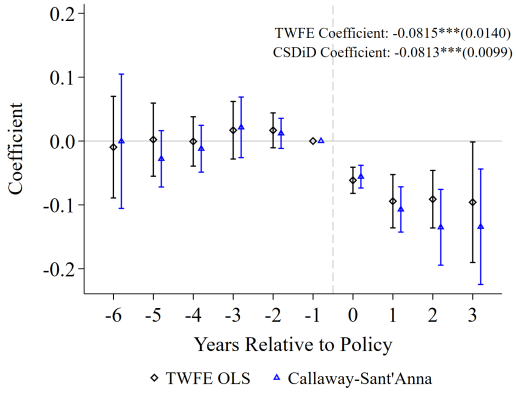
Note: The figure plots the correlation between the industry-level share of agricultural inputs and the VAT-to-asset ratio across agro-processing industries. The agricultural input share is defined as the ratio of agricultural inputs to total deductible inputs, while the VAT-to-asset ratio is VAT remittance divided by total assets in the preceding year. Both variables are computed as four-digit industry averages of firm-level ratios and are winsorized at the 1st and 99th percentiles. The reported slope coefficient is estimated from a simple linear regression of the VAT-to-asset ratio on the agricultural input share at the industry level, with robust standard errors.

Figure A.2: Dynamic Effects of the Centrally Determined 2012 National Phase Reform



Note: These figures present the dynamic effects of the centrally determined 2012 national reform phase. The treatment group is restricted to industries covered by the 2012 national rollout, whose timing and treated industries were determined by the central government rather than by provincial governments. Panels (a)–(d) report the effects on VAT remittance, input tax credit, output tax, and the use of suspicious invoices, respectively. The figure reports event-study estimates and 95% confidence intervals from the dynamic TWFE model in Equation (2), estimated using OLS. The omitted period is 2011. The figure also reports the average effects from the TWFE specification in Equation (1). Dependent variables in Panels (a)–(c) are normalized with the firm's total assets in the preceding year. In Panel (d), we use deductible inputs exceeding total costs as a signal of suspicious invoice use (Li and Wang, 2026). All dependent variables are winsorized at the 1st and 99th percentiles. All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level.

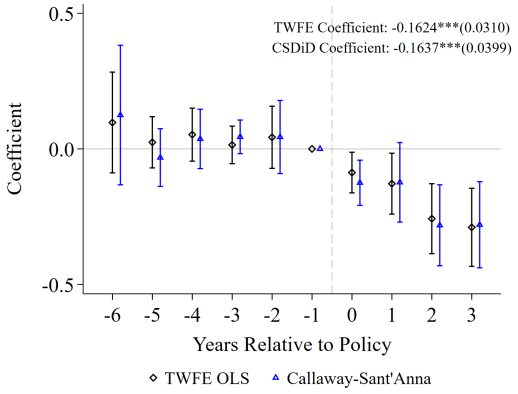
Figure A.3: Impact of the Presumptive Tax Reform on Agricultural Inputs



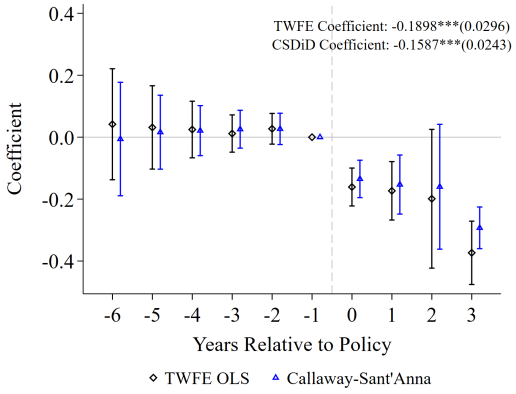
Note: This figure shows the impact of the presumptive tax reform on agricultural inputs. The dependent variable is normalized with the firm's total assets in the preceding year and winsorized at the 1st and 99th percentiles. The figure overlays the event-study estimates and 95% confidence intervals using two different estimators: the dynamic TWFE model in Equation (2), estimated using OLS (black diamonds), and the estimator of Callaway and Sant'Anna (2021) (blue triangles). The figure also reports the average effects from both estimators. All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level.

Figure A.4: Impact of the Presumptive Tax Reform on Voluntary VAT Registration

(a) Agro-Processing Firms

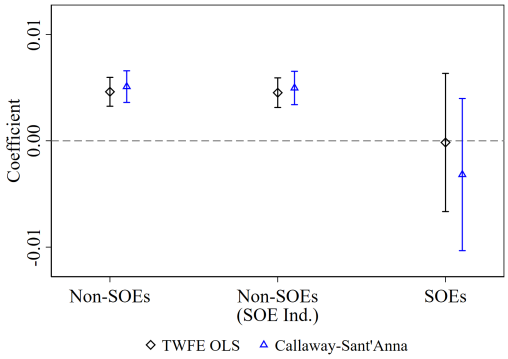


(b) Wholesale and Retail Firms



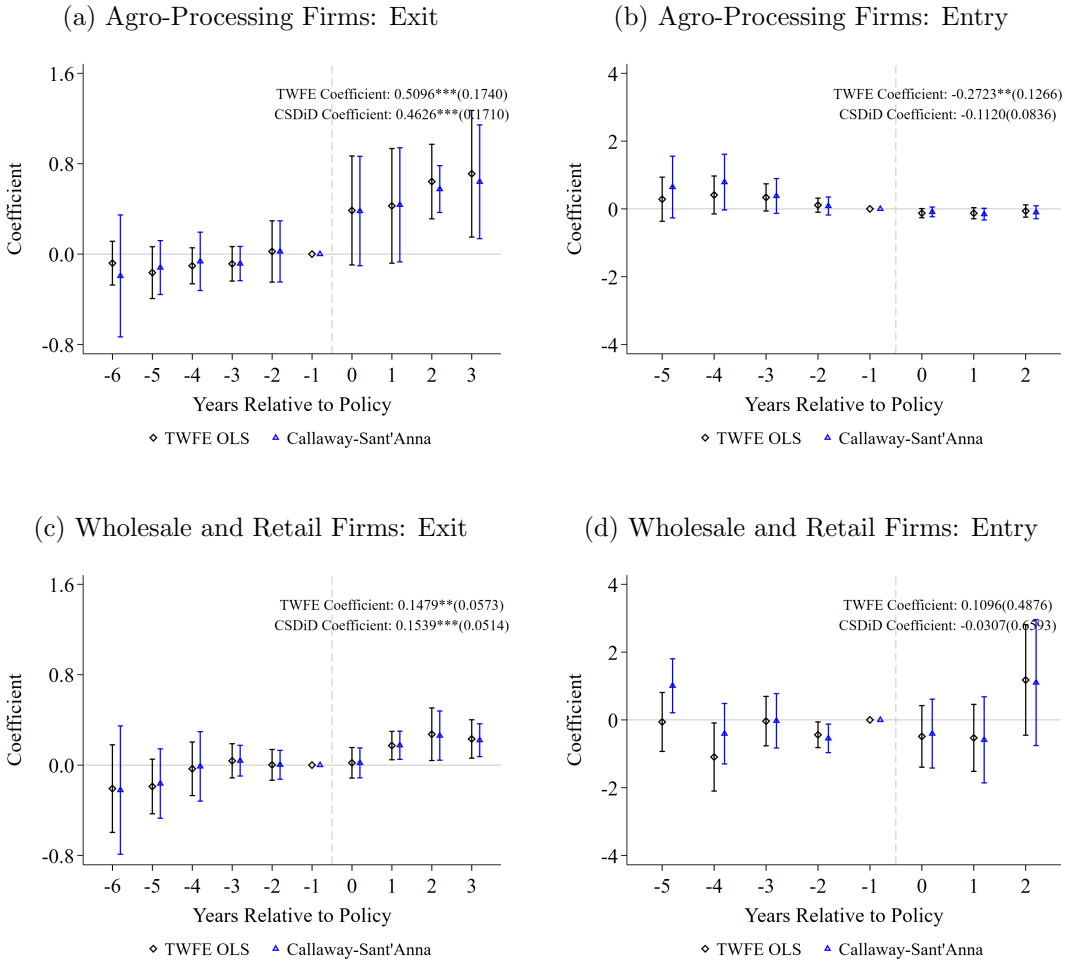
Note: These figures present the impact of the presumptive tax reform on voluntary VAT registration. The dependent variable is the industry–province level share of below-threshold firms that voluntarily register for VAT. Panel (a) presents results for the agro-processing firms, while Panel (b) shows the effects for the downstream wholesale and retail firms. The figures overlay the event-study estimates and 95% confidence intervals using two different estimators: the dynamic TWFE model in Equation (3), estimated using OLS (black diamonds), and the estimator of Callaway and Sant'Anna (2021) (blue triangles). The figures also report the average effects from both estimators. All specifications include industry-by-province and province-by-year fixed effects. Standard errors are clustered at the four-digit industry level.

Figure A.5: Estimates of Deduction Ratio Bias



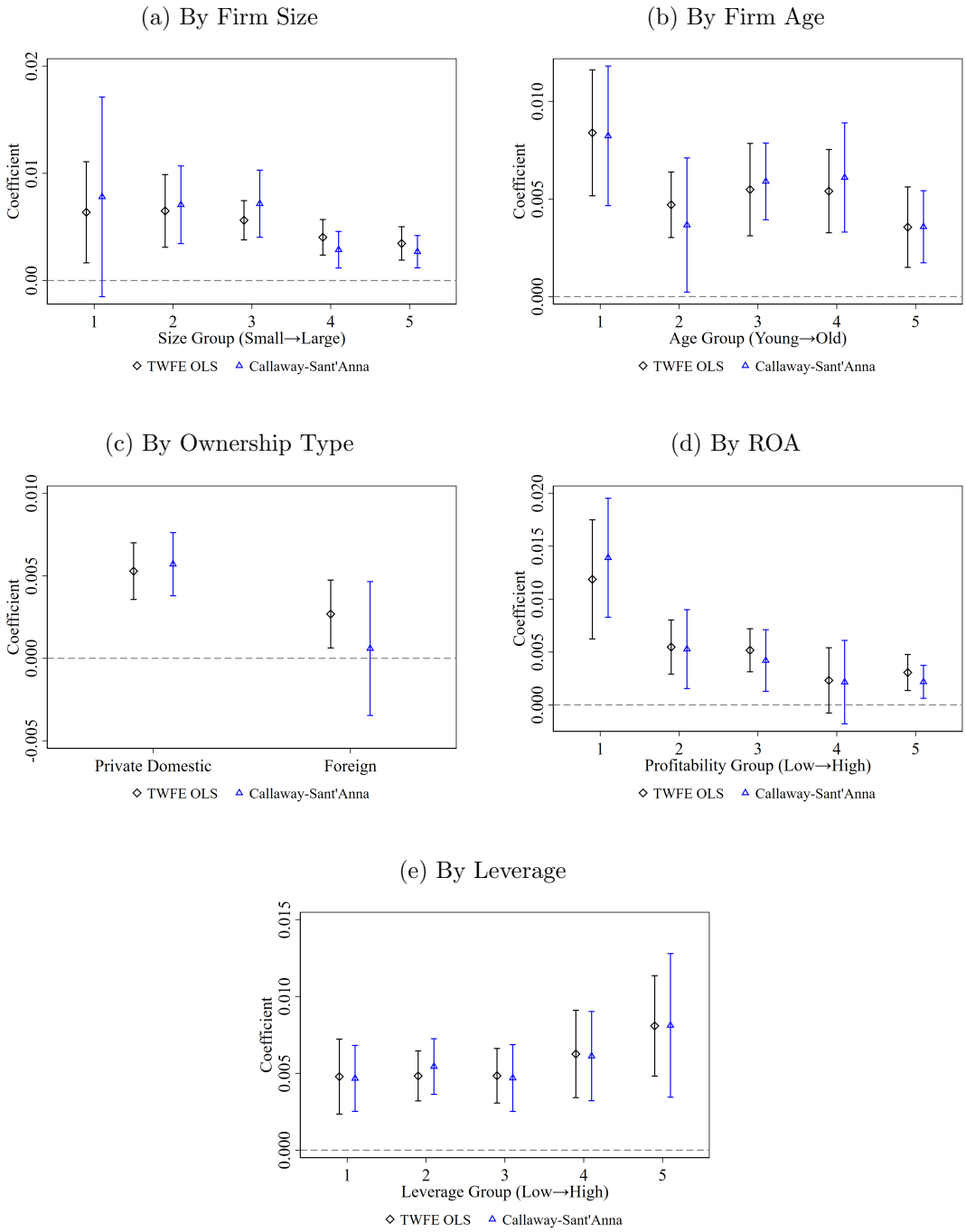
Note: The figure shows the heterogeneous impact of the reform on VAT remittance, which helps to estimate the deduction ratio bias. SOEs are used as a low-evasion reference group because they have minimal incentives to evade taxes and exhibit no significant real production responses to the reform. The middle group, Non-SOEs (SOE Ind.), restricts the non-SOE sample to industries in which SOEs are observed, so that the comparison is less affected by differences in industry composition. The SOE estimates help us assess whether deduction ratio bias is large enough to explain the increase in VAT remittance. The figure overlays the estimates and 95% confidence intervals using two different estimators: the TWFE model in Equation (1), estimated using OLS (black diamonds), and the estimator of Callaway and Sant'Anna (2021) (blue triangles). All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level.

Figure A.6: Policy Effects on Firm Entry and Exit



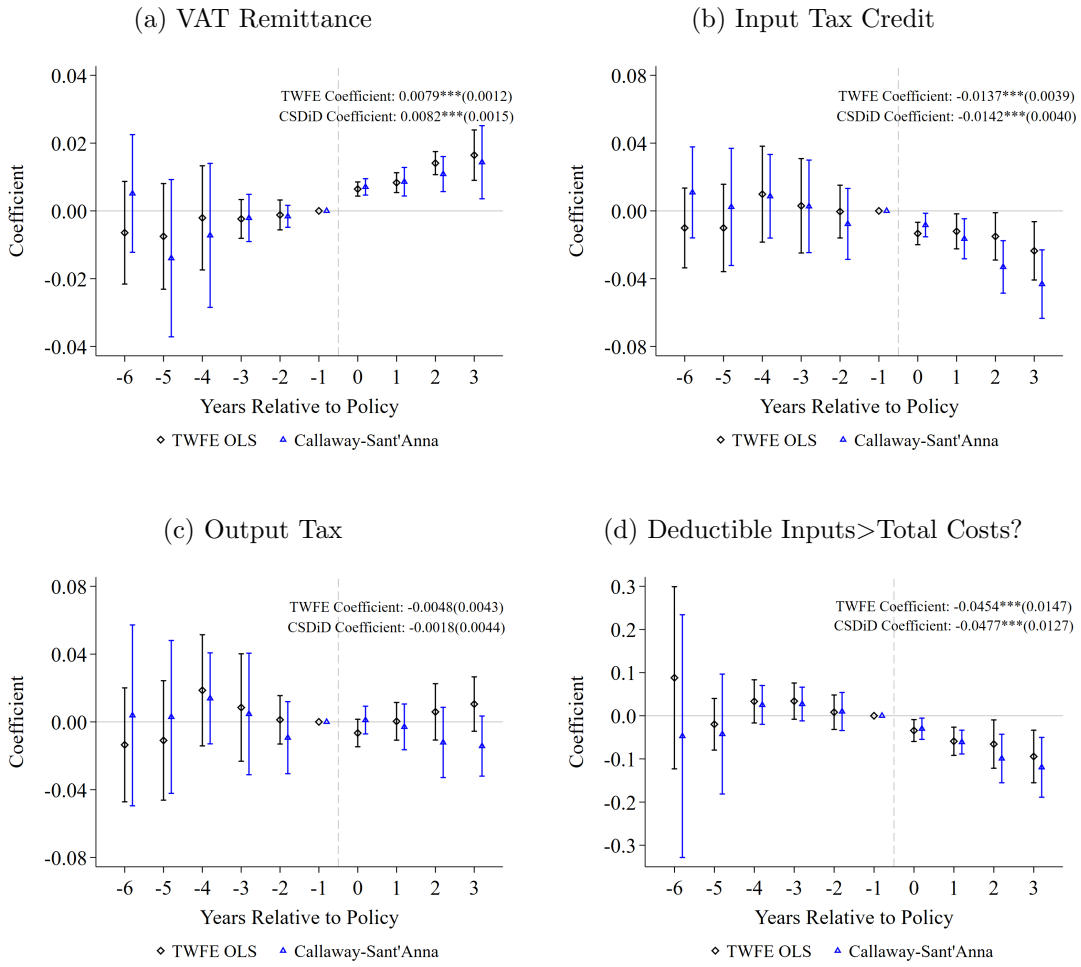
Note: These figures show the effects of the presumptive tax reform on firm exit and entry. Panels (a) and (b) report the direct effects on agro-processing firms, while Panels (c) and (d) report the spillover effects on downstream wholesale and retail firms. Because the data, in principle, include only firms that have been registered for more than one year, entry effects are estimated from period -5 to $+2$. The figures overlay the event-study estimates and 95% confidence intervals from two estimators: the dynamic TWFE model in Equation (3), estimated using OLS (black diamonds), and the estimator of Callaway and Sant'Anna (2021) (blue triangles). The figures also report the average effects from both estimators. All specifications include industry-by-province and province-by-year fixed effects. Standard errors are clustered at the four-digit industry level.

Figure A.7: Additional Heterogeneity by Firm Characteristics



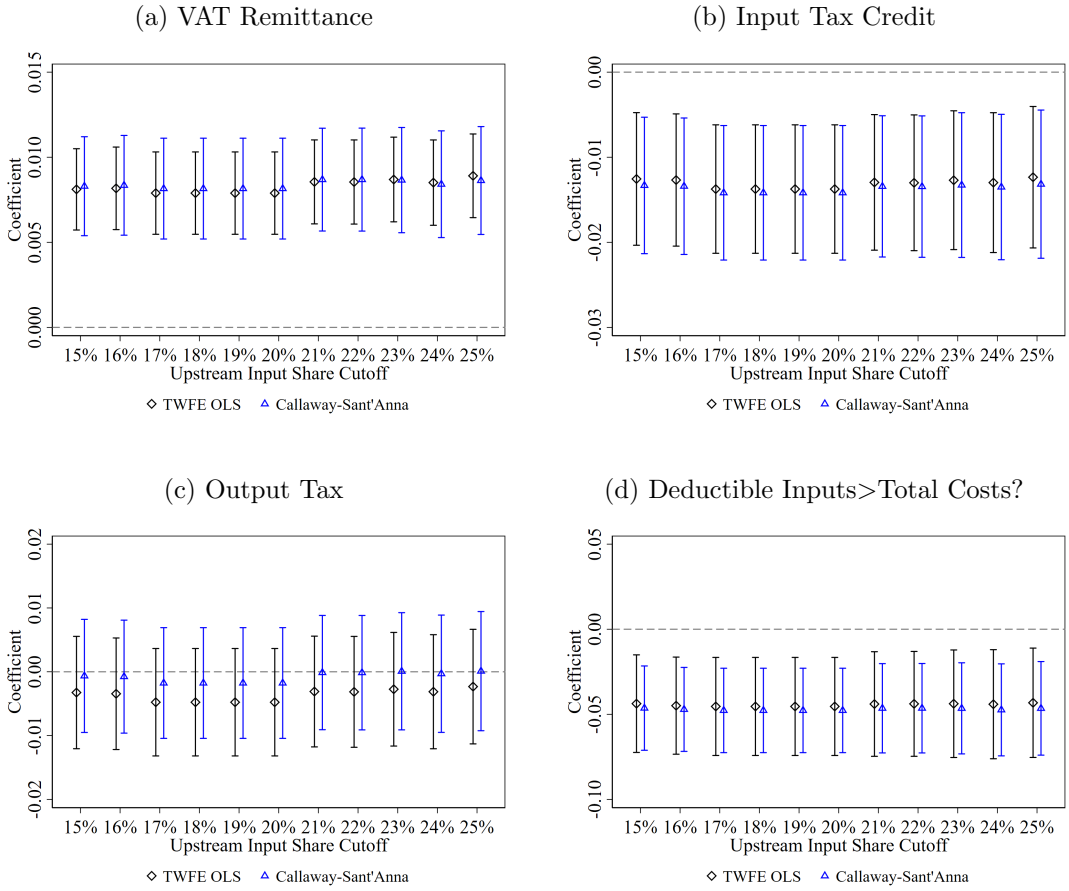
Note: These figures show additional heterogeneity in the effect of the presumptive tax reform on VAT remittance across pre-reform firm characteristics. Panels (a)–(e) report results by pre-reform firm size, firm age, ownership type, ROA, and leverage, respectively. Firm size is measured by total assets, firm age by years since registration, ROA by net profit over lagged total assets, and leverage by debt over lagged total assets. For Panels (a), (b), (d), and (e), firms are divided into five groups based on quintiles of the corresponding pre-reform firm characteristic. Panel (c) reports estimates separately for private domestic and foreign-invested firms. The figures report average effects estimated using two estimators: the TWFE model in Equation (1), estimated using OLS (black diamonds), and the estimator of Callaway and Sant'Anna (2021) (blue triangles). All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level.

Figure A.8: Downstream Treatment Assignment Using Transaction-Level Data



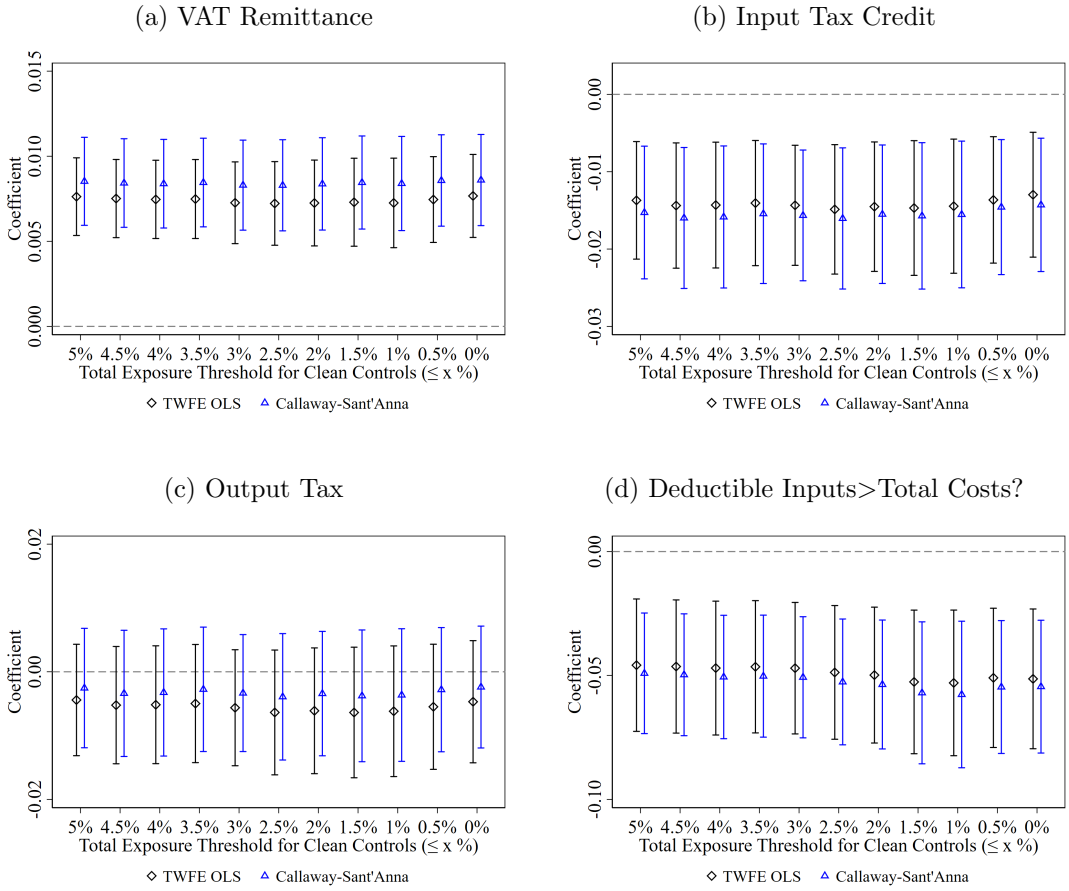
Note: These figures show the downstream spillover effects of the presumptive tax reform using transaction-level VAT linkage data to define downstream exposure. A downstream industry is classified as treated if at least one treated agro-processing industry accounts for more than 20% of its relevant upstream inputs. Panels (a)–(d) report the effects on VAT remittance, input tax credit, output tax, and the use of suspicious invoices, respectively. Dependent variables in Panels (a)–(c) are normalized with the firm's total assets in the preceding year. In Panel (d), we use deductible inputs exceeding total costs as a signal of suspicious invoice use (Li and Wang, 2026). All dependent variables are winsorized at the 1st and 99th percentiles. The figures overlay the event-study estimates and 95% confidence intervals using two different estimators: the dynamic TWFE model in Equation (2), estimated using OLS (black diamonds), and the estimator of Callaway and Sant'Anna (2021) (blue triangles). All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level.

Figure A.9: Downstream Treatment Assignment Using Transaction-Level Data with Different Cutoffs



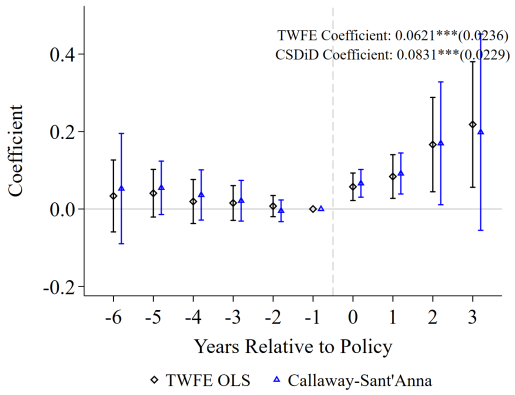
Note: These figures show the downstream spillover effects of the presumptive tax reform under alternative cutoffs for the treatment definition. A downstream industry is classified as treated if at least one treated agro-processing industry accounts for a share above the cutoff shown on the horizontal axis. Panels (a)–(d) report the effects on VAT remittance, input tax credit, output tax, and the use of suspicious invoices, respectively. Dependent variables in Panels (a)–(c) are normalized with the firm's total assets in the preceding year. In Panel (d), we use deductible inputs exceeding total costs as a signal of suspicious invoice use (Li and Wang, 2026). All dependent variables are winsorized at the 1st and 99th percentiles. The figures report average effects estimated using two different estimators: the TWFE model in Equation (1), estimated using OLS (black diamonds), and the estimator of Callaway and Sant'Anna (2021) (blue triangles). All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level.

Figure A.10: Downstream Spillover Effects Using Control Groups with Low Total Exposure



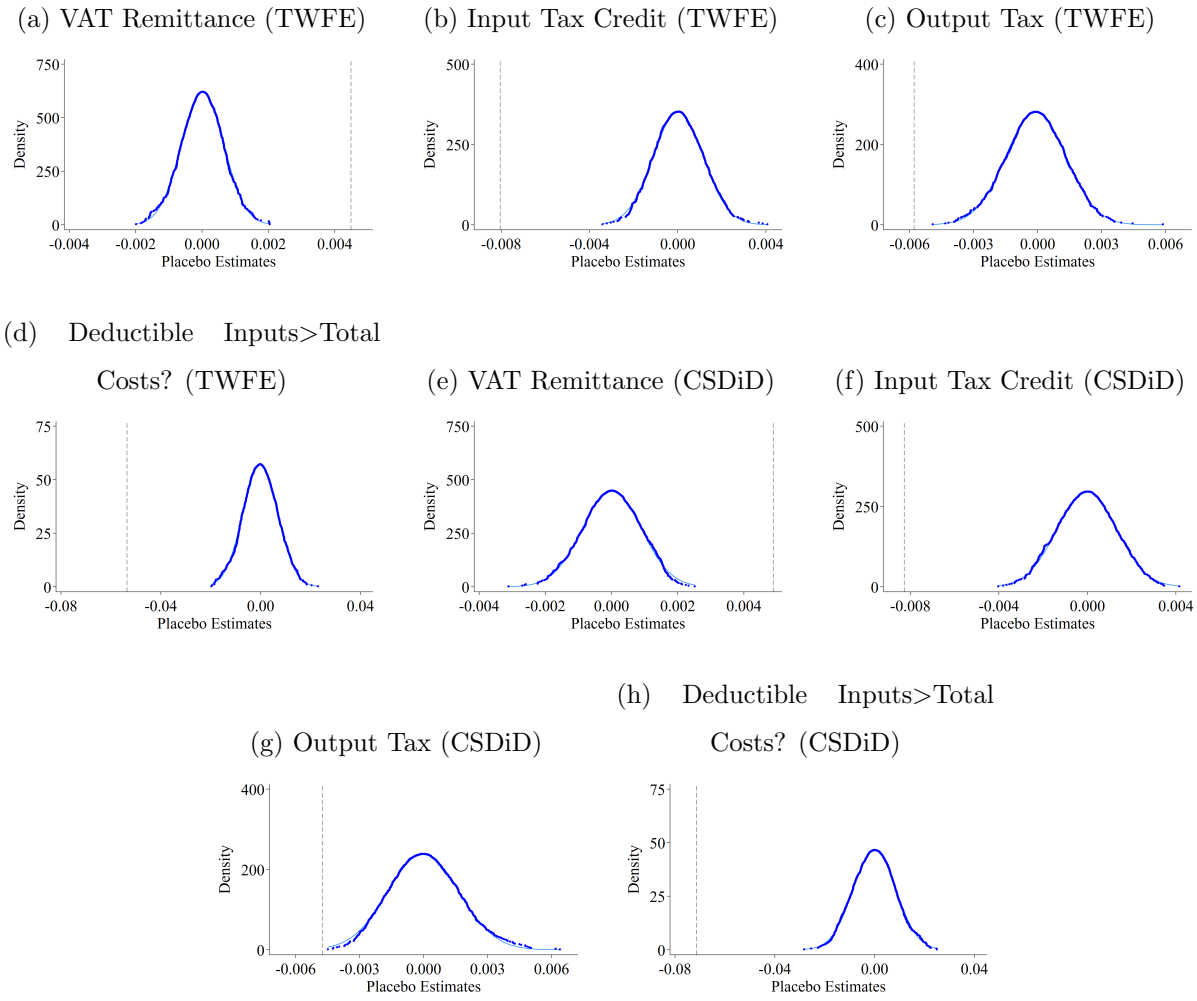
Note: These figures show the downstream spillover effects of the presumptive tax reform using cleaner control groups based on total exposure. Total exposure is defined as the sum of direct exposure and indirect exposure. Direct exposure measures input linkages to treated upstream agro-processing industries, while indirect exposure measures exposure transmitted through one additional wholesale and retail layer. We exclude industries with high total exposure from the control group. Panels (a)–(d) report the effects on VAT remittance, input tax credit, output tax, and the use of suspicious invoices, respectively. Dependent variables in Panels (a)–(c) are normalized with the firm's total assets in the preceding year. In Panel (d), we use deductible inputs exceeding total costs as a signal of suspicious invoice use (Li and Wang, 2026). All dependent variables are winsorized at the 1st and 99th percentiles. The figures report average effects estimated using two different estimators: the TWFE model in Equation (1), estimated using OLS (black diamonds), and the estimator of Callaway and Sant'Anna (2021) (blue triangles). All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level.

Figure A.11: Impact of the Presumptive Tax Reform on Export Product Prices



Note: The figure shows the impact of the presumptive tax reform on export product prices. We use customs data at the firm–product–destination-country–year level for estimation. The dependent variable is the logarithm of firms' export prices. The figure overlays the event-study estimates and 95% confidence intervals using two different estimators: the dynamic TWFE model (black diamonds), and the estimator of Callaway and Sant'Anna (2021) (blue triangles). The figures also report the average effects from both estimators. All specifications include firm-product-destination-country and year fixed effects. Standard errors are clustered at the four-digit industry level.

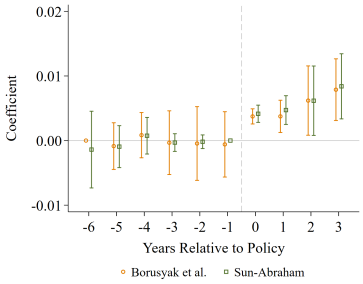
Figure A.12: Permutation Test with Randomly Assigned Reform Years



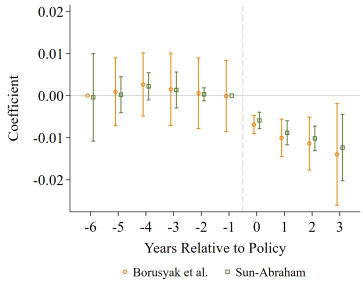
Note: These figures report permutation tests with randomly assigned reform years. In each simulation, we randomly permute the actual reform years across firms while keeping the actual distribution of treatment cohorts unchanged. We then construct a placebo reform indicator based on the assigned placebo reform year and re-estimate the baseline specification. This procedure is repeated 1,000 times, yielding the placebo distribution plotted in each panel. The dashed vertical line marks the actual estimate from the baseline specification. Panels (a)–(d) report results using the TWFE estimator, and Panels (e)–(h) report results using the CSDiD estimator. VAT remittance, input tax credit, and output tax are normalized with the firm's total assets in the preceding year and winsorized at the 1st and 99th percentiles. The suspicious invoice indicator is defined as an indicator for deductible inputs exceeding total costs. All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level.

Figure A.13: Main Results Using Alternative Robust Dynamic Estimators

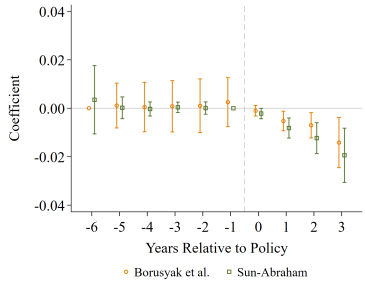
(a) Direct VAT Remittance



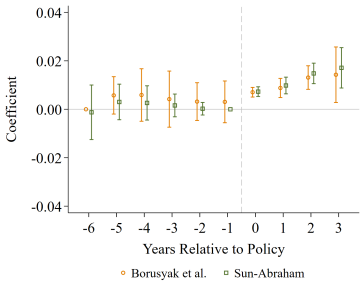
(b) Direct Input Tax Credit



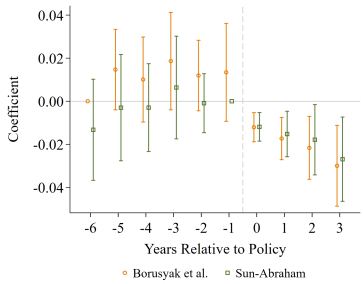
(c) Direct Output Tax



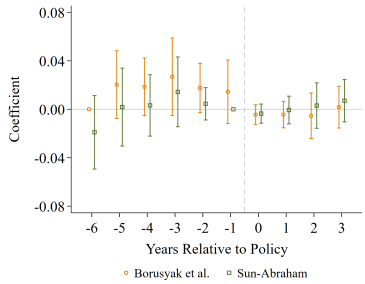
(d) Spillover - Downstream VAT Remittance



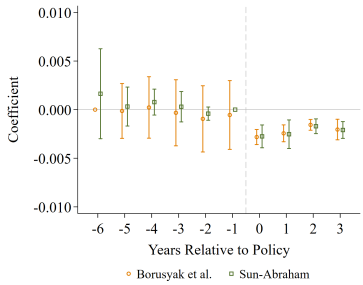
(e) Spillover - Downstream Input Tax Credit



(f) Spillover - Downstream Output Tax



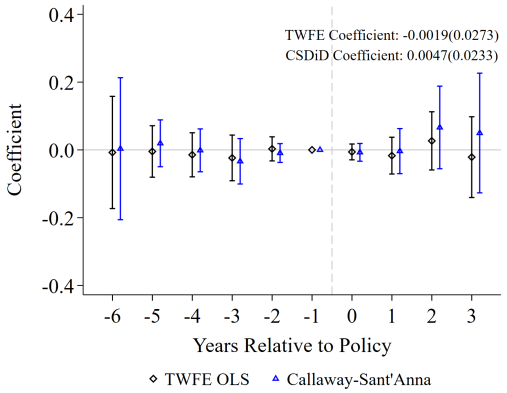
(g) Spillover - Export Stage Tax Refund



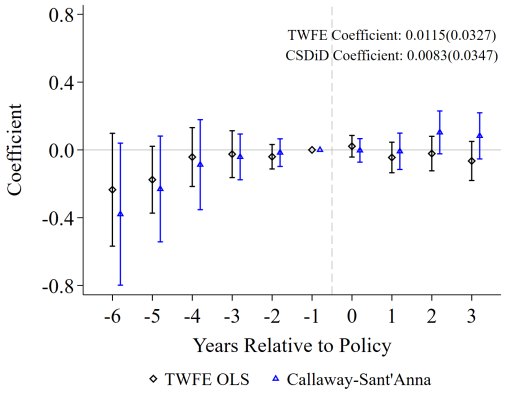
Note: These figures show the direct effects (Panels (a)–(c)), the spillover effects on wholesale and retail firms (Panels (d)–(f)), and the spillover effects on the export stage (Panel (g)) using alternative robust dynamic estimators. The figures overlay the event-study estimates and 95% confidence intervals using two alternative estimators: the estimator of Borusyak, Jaravel and Spiess (2024) (orange circles) and the estimator of Sun and Abraham (2021) (green squares). Dependent variables in Panels (a)–(g) are normalized with the firm's total assets in the preceding year and winsorized at the 1st and 99th percentiles. All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level.

Figure A.14: Impact of the Presumptive Tax Reform on Firms' Total Assets

(a) Agro-Processing Firms



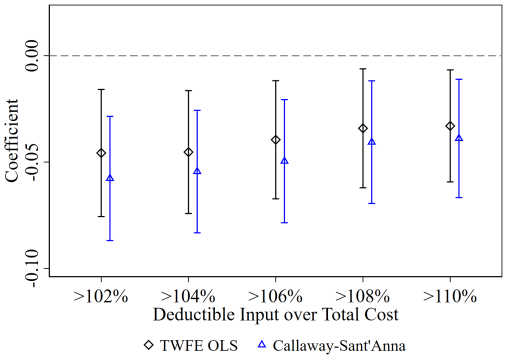
(b) Wholesale and Retail Firms



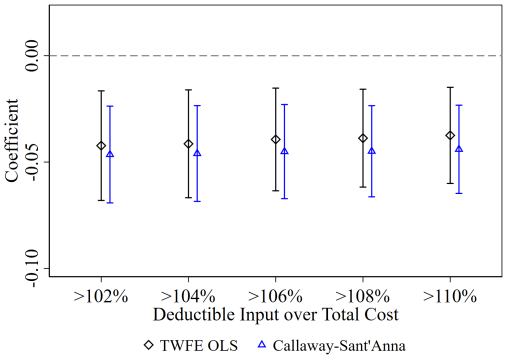
Note: These figures examine the impact of the presumptive tax reform on the logarithm of firms' lagged total assets, which serve as the denominator for the main outcome variables. Panel (a) presents results for the agro-processing firms, while Panel (b) shows the effects for the downstream wholesale and retail firms. The figures overlay the event-study estimates and 95% confidence intervals using two different estimators: the dynamic TWFE model in Equation (2), estimated using OLS (black diamonds), and the estimator of Callaway and Sant'Anna (2021) (blue triangles). The figures also report the average effects from both estimators. All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level.

Figure A.15: Alternative Thresholds for the Suspicious Invoice Indicator

(a) Agro-Processing Firms

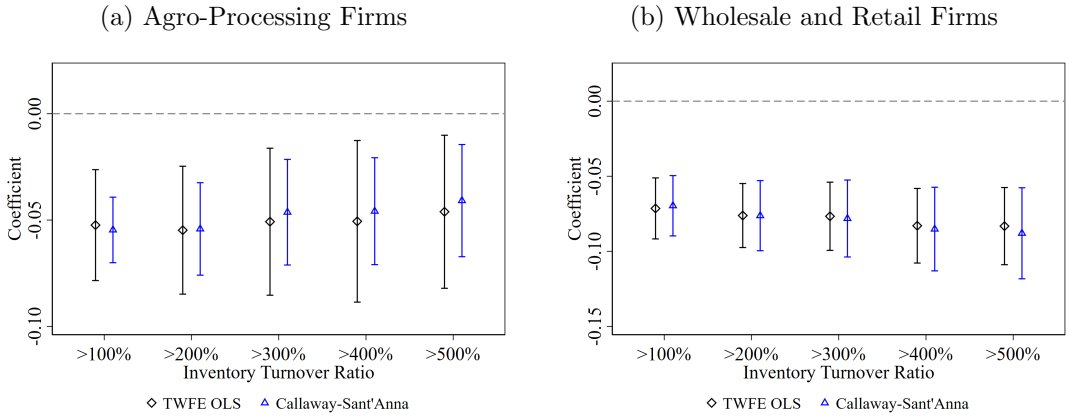


(b) Wholesale and Retail Firms



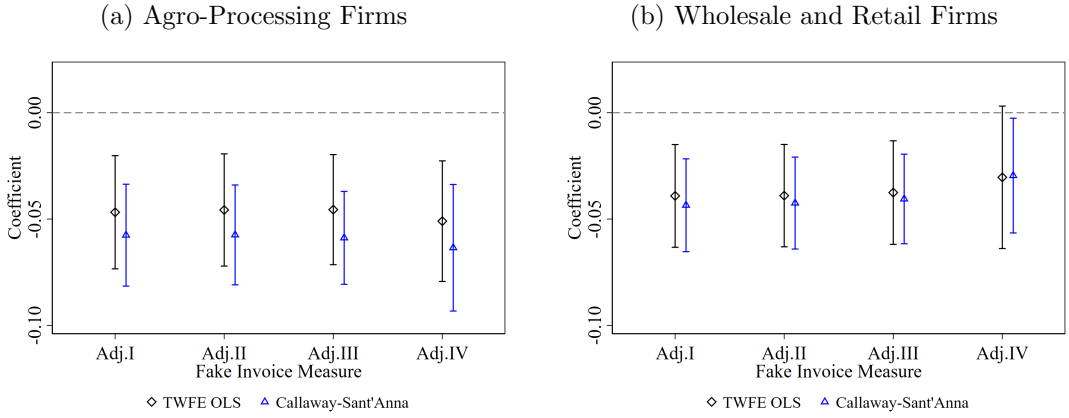
Note: These figures examine the sensitivity of the suspicious invoice results to stricter threshold definitions. We define the suspicious invoice indicator using thresholds of 102%, 104%, 106%, 108%, and 110%, respectively. Panel (a) reports results for agro-processing firms, and Panel (b) reports results for downstream wholesale and retail firms. The figure overlays estimates and 95% confidence intervals using two different estimators: the TWFE model in Equation (1), estimated using OLS (black diamonds), and the estimator of Callaway and Sant'Anna (2021) (blue triangles). All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level.

Figure A.16: Effects on Suspicious Invoice Use Among Firms with Higher Inventory Turnover



Note: These figures examine whether the suspicious invoice results are sensitive to timing differences between VAT invoice recognition and accrual based cost recognition under the CIT. We re-estimate the reform effect on the suspicious invoice indicator in subsamples of firms with progressively higher pre-reform inventory turnover. Specifically, we restrict the sample to firms whose pre-reform inventory turnover exceeds 100%, 200%, 300%, 400%, and 500%, respectively. Panel (a) reports results for agro-processing firms, and Panel (b) reports results for downstream wholesale and retail firms. The figure overlays estimates and 95% confidence intervals using two different estimators: the TWFE model in Equation (1), estimated using OLS (black diamonds), and the estimator of Callaway and Sant'Anna (2021) (blue triangles). All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level.

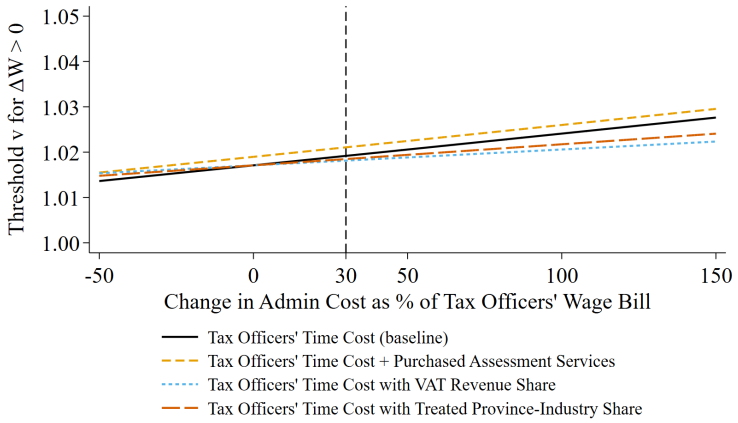
Figure A.17: Measures of Suspicious Invoice Use with Alternative Total Cost Definitions



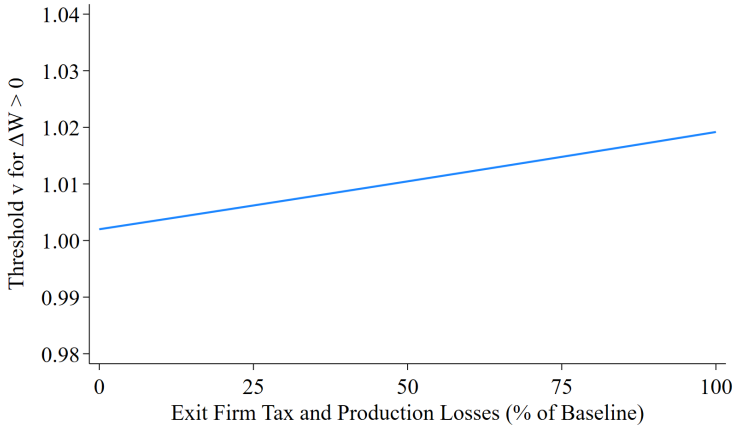
Note: These figures examine whether the suspicious invoice results are sensitive to alternative definitions of total costs. We re-estimate the reform effect on the suspicious invoice indicator using progressively adjusted total cost measures that move the CIT cost concept closer to the VAT deduction concept. Adj. I applies the fixed asset adjustment in Equation (D.1); Adj. II further excludes business entertainment expenses; Adj. III additionally excludes interest, and Adj. IV further excludes wages. Panel (a) reports results for agro-processing firms, and Panel (b) reports results for downstream wholesale and retail firms. The figure overlays estimates and 95% confidence intervals using two different estimators: the TWFE model in Equation (1), estimated using OLS (black diamonds), and the estimator of Callaway and Sant'Anna (2021) (blue triangles). All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level.

Figure A.18: Sensitivity of the Threshold v to Administrative Costs and Exit Firm Losses

(a) Administrative Cost



(b) Loss of Exit Firms



Note: Panel (a) plots the threshold value of v required for $\Delta W > 0$ under alternative assumptions about administrative costs. The black solid line varies the share of staff time devoted to the reform. The blue short-dashed and red long-dashed lines use alternative staff allocation proxies. The orange dashed line adds fixed administrative costs based on government procurement values for similar assessment services. The vertical dashed line marks the benchmark assumption that tax officers devote 30% of their time to the reform. Panel (b) shows how the threshold value of v changes under different assumptions about the extent to which estimated exit-related losses represent actual economic and fiscal losses. A value of 100% assumes that all estimated losses are realized, while lower values allow for the possibility that some of these losses reflect the disappearance of tax evasion rather than real economic activity.

Table A.1: Firm Behavior and Tax Base under Different Regimes

Production Stage and Reported Value Added	Farmers (Informal Sector)	Vegetable and Fruit Processing Firms	Vegetable and Fruit Retailers	Total VAT Base
Standard VAT system (with no evasion)	50	100–50	200–100	150
Standard VAT system (with evasion)	50	100–50–30 (overreported input credit to reduce VAT) +45 (selling fake invoice) –45 (overreported input credit for invoice laundering)	200–100–45 (fake invoice)	75
Presumptive Tax	50	100–55	200–100	145

Notes: The table illustrates how firm reporting behavior affects the VAT base under alternative regimes. Column (1) lists tax regimes. Columns (2)–(4) present the reported value added of farmers, processing firms, and retailers, respectively, and Column (5) provides the total reported VAT base. Under the "with evasion" scenario, processing firms overreport input credits to reduce their own VAT liability (-30), sell fake invoices ($+45$), and claim self-billed input tax credits (-45) to offset the additional output tax liability from fake invoice sales, while retailers use fake invoices to understate VAT (-45). The presumptive tax curtails such practices and increases the VAT base to 145.

Table A.2: Positioning Relative to Related Studies

Paper	What the paper establishes	Main contrast	Our distinct margin
Pomeranz (2015)	VAT paper trails generate third-party information. When buyers have stronger incentives to request invoices, suppliers in turn report more sales, so enforcement can spill over upstream.	Direction of spillover and mechanism.	We study downstream propagation through a fake invoice supply channel. The reform restricts upstream firms' ability to provide commercially coherent invoices that downstream firms can use to support input credit claims.
Waseem (2022)	The upward extension of VAT triggers a withholding mechanism and leads downstream firms to report substantially higher sales.	Policy shock and mechanism of downstream response.	Downstream firms in our setting face no direct change in VAT coverage or withholding rules. Their response comes from an upstream restriction on fake invoice supply, which raises the cost of obtaining invoices.
Waseem (2023)	Measures the nature and extent of VAT noncompliance, including overclaimed refunds, undeclared sales, and invoice mills.	Measurement of evasion types.	Our paper focuses on invoice laundering as a specific form of VAT evasion. We show how fake invoices can be repackaged into commercially coherent invoices that downstream firms can use.
Carrillo et al. (2023)	Studies ghost firms and their client firms, documents the characteristics and patterns of ghost transactions, and evaluates enforcement targeted at ghost clients.	How fake invoices operate.	We show that fake invoice evasion is not always complete once invoices are produced. Business coherence requirements can lead firms to launder such invoices for downstream input credit claims.
Asatryan and Gomtsyan (2023)	Studies who benefits from VAT evasion using a retail enforcement reform and price pass through, with an emphasis on the incidence of evasion rents.	Incidence versus propagation.	We do not primarily estimate the incidence of evasion rents. We identify a downstream propagation mechanism in which tighter upstream enforcement increases downstream compliance.
Fan, Liu, Qian, and Wen (2024)	Studies invoice computerization and encryption in China, showing how stronger authentication and transaction matching directly affect compliance and firm behavior.	Technology versus rule change.	Their shock makes invoices harder to forge and easier to match. Our shock closes a self-billed input credit channel. We show that invoices can still support downstream noncompliance even when they pass computerized matching, if they are backed by fake credits and appear commercially coherent.

Table A.3: Province-Industry-Level Discrete Hazard Model for Reform Adoption

	(1)	(2)	(3)
Firm Characteristics			
Log Total Assets	0.0015		0.0016
	(0.0011)		(0.0011)
Exporter Share	-0.0011		-0.0009
	(0.0018)		(0.0017)
Fixed Assets Investment / Assets	0.0054		0.0030
	(0.0288)		(0.0295)
Wages / Assets	0.0054		-0.0030
	(0.0097)		(0.0062)
Debt / Assets	0.0077*		0.0069
	(0.0044)		(0.0041)
Tax and Suspicious Invoice Variables			
VAT Remittance / Assets		0.0247	0.0313
		(0.0357)	(0.0370)
Output Tax / Assets		-0.0102	-0.0085
		(0.0095)	(0.0103)
Input Tax Credit / Assets		0.0109	0.0149
		(0.0142)	(0.0150)
Deductible Inputs >Total Costs?		-0.0000	-0.0003
		(0.0025)	(0.0026)
Electricity / Sales		-0.0017	-0.0012
		(0.0018)	(0.0016)
Observations	15,114	15,114	15,114
Year FE	✓	✓	✓
Province FE	✓	✓	✓
Industry FE	✓	✓	✓

Note: This table reports province-industry-level discrete hazard models of reform adoption. The sample includes each province-industry cell up to and including the year it adopted the reform, and excludes post-adoption years. The dependent variable equals one in the year when a province-industry cell adopts the reform and zero otherwise. Column (1) includes lagged firm characteristics, Column (2) includes lagged tax and suspicious invoice variables, and Column (3) includes all lagged covariates. Year, province, and four-digit industry fixed effects are included in all specifications. Standard errors are clustered at the province level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.4: Pre-Reform Descriptive Comparison: Levels and Changes

	Levels			Annual Changes		
	Treatment	Control	T-C	Treatment	Control	T-C
Firm Characteristics						
Log Assets	9.344 (1.813)	8.438 (1.770)	0.906***	0.133 (0.658)	0.132 (0.687)	0.001
Agricultural Inputs / Assets	0.770 (0.989)	0.701 (1.045)	0.070***	-0.060 (0.367)	-0.064 (0.369)	0.004
Fixed Assets Investment / Assets	0.021 (0.048)	0.019 (0.047)	0.001***	-0.001 (0.056)	-0.001 (0.057)	0.001
Electricity / Assets	0.051 (0.164)	0.047 (0.167)	0.003*	-0.012 (0.197)	-0.011 (0.213)	-0.001
Wages / Assets	0.048 (0.088)	0.114 (0.178)	-0.066***	-0.000 (0.088)	-0.004 (0.145)	0.003
Debt / Assets	0.402 (0.234)	0.378 (0.245)	0.024***	-0.003 (0.235)	0.002 (0.234)	-0.005
VAT Payments and Compliance Measures						
VAT Remittance / Assets	0.023 (0.037)	0.032 (0.052)	-0.009***	-0.001 (0.033)	-0.000 (0.042)	-0.001
Taxable Sales / Assets	1.305 (1.320)	1.147 (1.338)	0.158***	-0.031 (0.573)	-0.033 (0.565)	0.002
Output Tax / Assets	0.141 (0.134)	0.129 (0.147)	0.012***	0.008 (0.069)	0.007 (0.075)	0.001
Deductible Inputs / Assets	1.490 (1.639)	1.182 (1.424)	0.308***	-0.023 (1.286)	-0.029 (0.878)	0.006
Input Tax Credit / Assets	0.135 (0.138)	0.117 (0.132)	0.018***	-0.005 (0.069)	-0.005 (0.062)	0.000
Deductible Inputs / Total Costs	0.896 (0.479)	0.773 (0.477)	0.123***	-0.014 (0.560)	-0.018 (0.513)	0.004
Deductible Inputs > Total Costs?	0.227 (0.419)	0.126 (0.332)	0.101***	-0.015 (0.526)	-0.010 (0.386)	-0.005
Electricity / Sales	0.066 (0.310)	0.088 (0.374)	-0.023***	-0.018 (0.385)	-0.012 (0.474)	-0.006

Note: This table compares pre-reform levels and annual changes between treated and control agro-processing firms. The table reports group means and standard deviations. The difference columns report the treatment mean minus the control mean. Differences may not exactly equal the displayed means because of rounding. *, **, and *** denote statistical significance from a t-test at the 10%, 5%, and 1% levels, respectively.

Table A.5: Direct Effects of the Presumptive Tax Reform (Sample with Agricultural Input Records)

	(1)	(2)	(3)
Dependent Variable	VAT Remittance	Input Tax Credit	Output Tax
Panel A. TWFE Estimator			
Policy	0.0043***	-0.0065***	-0.0055**
	(0.0012)	(0.0021)	(0.0023)
Observations	61,410	61,410	61,410
Adjusted R^2	0.61	0.89	0.85
Panel B. CSDiD Estimator			
Policy	0.0042***	-0.0078***	-0.0062***
	(0.0013)	(0.0013)	(0.0021)
Observations	61,410	61,410	61,410
Pre-Treat Mean of Dep. Var.	0.0226	0.1428	0.1493
Firm FE	✓	✓	✓
Year FE	✓	✓	✓

Note: The table shows the direct effects of the presumptive tax reform using the subsample with agricultural input records. All dependent variables are normalized with the firm's total assets in the preceding year and winsorized at the 1st and 99th percentiles. Panel A reports estimates from the TWFE model in Equation (1), estimated using OLS. Panel B reports the corresponding average effects using the estimator of Callaway and Sant'Anna (2021). All specifications include firm and year fixed effects. Pre-Treat Mean of Dep. Var. is the pre-reform mean of the dependent variable for treated firms. Standard errors are clustered at the four-digit industry level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.6: Real Production Responses of Firms Below the VAT Registration Threshold

	(1)	(2)	(3)	(4)
Panel A. Investments and Outsourcing Services				
Dependent Variable	Fixed Assets Investment	Intangible Assets Investment	Processing Service	Transportation Service
Policy (TWFE)	0.0011	0.0002	-0.0003	-0.0011
	(0.0260)	(0.0003)	(0.0046)	(0.0013)
Policy (CSDiD)	-0.0095	0.0004	0.0006	-0.0027**
	(0.0315)	(0.0005)	(0.0047)	(0.0014)
Observations	3,558	3,558	3,558	3,558
Panel B. Energy Consumption				
Dependent Variable	Electricity	Water	Coal	Oil
Policy (TWFE)	0.0696	0.5153	-0.0133	0.0010
	(0.1171)	(0.3465)	(0.0126)	(0.0033)
Policy (CSDiD)	0.0941	0.5094	-0.0112	0.0019
	(0.1209)	(0.3600)	(0.0134)	(0.0029)
Observations	3,558	3,558	3,558	3,558
Panel C. Other Firm Characteristics				
Dependent Variable	Wage	Debt	Depreciation	Interest
Policy (TWFE)	0.0114	0.0200	-0.0012	0.0001
	(0.0265)	(0.0566)	(0.0053)	(0.0003)
Policy (CSDiD)	0.0140	-0.0247	0.0024	-0.0001
	(0.0298)	(0.0471)	(0.0048)	(0.0003)
Observations	3,558	3,558	3,558	3,558
Firm FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓

Note: This table reports the real production responses of agro-processing firms with annual sales below the VAT registration threshold to the presumptive tax reform. Panel A reports investment and outsourced services, Panel B reports energy consumption, and Panel C reports other firm characteristics. All dependent variables are normalized with the firm's total assets in the preceding year and winsorized at the 1st and 99th percentiles. Policy (TWFE) reports estimates from the baseline specification in Equation (1), estimated using OLS. Policy (CSDiD) reports the estimator of Callaway and Sant'Anna (2021). All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. The unit of electricity is 10,000 kWh, and the unit of other energy consumption variables is tons.

Table A.7: Real Production Response: Heterogeneous Treatment Effects by Input Substitutability

	(1)	(2)	(3)	(4)
Panel A. Investments and Outsourcing Services				
Dependent Variable	Fixed Assets Investment	Intangible Assets Investment	Processing Service	Transportation Service
Policy (Hard)	0.0015	0.0007	-0.0006	-0.0008
	(0.0027)	(0.0006)	(0.0006)	(0.0009)
Policy (Easy)	-0.0008	-0.0001	0.0017	0.0004*
	(0.0019)	(0.0003)	(0.0010)	(0.0002)
Panel B. Energy Consumption				
Dependent Variable	Electricity	Water	Coal	Oil
Policy (Hard)	-0.0002	-0.2912	-0.0015	-0.0005
	(0.0083)	(0.1875)	(0.0071)	(0.0044)
Policy (Easy)	0.0025	0.2780	-0.0039	0.0031
	(0.0027)	(0.2813)	(0.0116)	(0.0020)
Panel C. Other Firm Characteristics				
Dependent Variable	Wage	Debt	Depreciation	Interest
Policy (Hard)	-0.0029	-0.0058	0.0005	0.0002
	(0.0037)	(0.0104)	(0.0009)	(0.0005)
Policy (Easy)	0.0042	0.0042	0.0007	-0.0001
	(0.0058)	(0.0062)	(0.0007)	(0.0003)
Firm FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓

Note: This table compares real production responses between hard-to-substitute and easy-to-substitute industries. Industries are classified as easy-to-substitute if the coefficient of variation of their pre-reform agricultural input shares is above the sample median, and as hard-to-substitute otherwise. The reported coefficients are estimated separately for the two groups using the baseline specification in Equation (1). All dependent variables are normalized with the firm's total assets in the preceding year and winsorized at the 1st and 99th percentiles. All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.8: Real Production Response: Producers vs. Likely Invoice Sellers (Electricity over Sales)

	(1)	(2)	(3)	(4)
Panel A. Investments and Outsourcing Services				
Dependent Variable	Fixed Assets Investment	Intangible Assets Investment	Processing Service	Transportation Service
Policy (Likely Legitimate Producers)	-0.0010	-0.0001	0.0015**	0.0002
	(0.0018)	(0.0003)	(0.0007)	(0.0003)
Policy (Likely Invoice Sellers)	0.0003	0.0008	0.0026*	0.0005
	(0.0030)	(0.0014)	(0.0015)	(0.0005)
Panel B. Energy Consumption				
Dependent Variable	Electricity	Water	Coal	Oil
Policy (Likely Legitimate Producers)	0.0034	0.1558	-0.0057	0.0025
	(0.0028)	(0.2610)	(0.0102)	(0.0019)
Policy (Likely Invoice Sellers)	0.0111	0.3664	0.0124	0.0040
	(0.0143)	(0.3702)	(0.0125)	(0.0044)
Panel C. Other Firm Characteristics				
Dependent Variable	Wage	Debt	Depreciation	Interest
Policy (Likely Legitimate Producers)	0.0050	0.0023	0.0007	-0.0000
	(0.0047)	(0.0066)	(0.0005)	(0.0003)
Policy (Likely Invoice Sellers)	0.0078	0.0120	0.0006	-0.0002
	(0.0049)	(0.0166)	(0.0019)	(0.0005)
Firm FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓

Note: This table compares real production responses between firms more likely to be legitimate producers and firms more likely to be invoice sellers. The classification is based on firms' pre-reform electricity consumption relative to sales. Firms with zero pre-reform electricity-to-sales ratios are classified as likely invoice sellers, while the remaining firms are classified as likely legitimate producers. The reported coefficients are estimated separately for the two groups using the baseline specification in Equation (1). All dependent variables are normalized with the firm's total assets in the preceding year and winsorized at the 1st and 99th percentiles. All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.9: Real Production Response: Producers vs. Likely Invoice Sellers (VAT over Assets)

	(1)	(2)	(3)	(4)
Panel A. Investments and Outsourcing Services				
Dependent Variable	Fixed Assets Investment	Intangible Assets Investment	Processing Service	Transportation Service
Policy (Likely Legitimate Producers)	-0.0008	-0.0001	0.0014*	0.0002
	(0.0018)	(0.0003)	(0.0007)	(0.0003)
Policy (Likely Invoice Sellers)	-0.0022	-0.0005	0.0041***	0.0003
	(0.0027)	(0.0013)	(0.0014)	(0.0006)
Panel B. Energy Consumption				
Dependent Variable	Electricity	Water	Coal	Oil
Policy (Likely Legitimate Producers)	0.0044	0.1790	-0.0045	0.0029
	(0.0027)	(0.2487)	(0.0092)	(0.0019)
Policy (Likely Invoice Sellers)	-0.0098	0.0948	-0.0016	-0.0028
	(0.0195)	(0.5238)	(0.0182)	(0.0057)
Panel C. Other Firm Characteristics				
Dependent Variable	Wage	Debt	Depreciation	Interest
Policy (Likely Legitimate Producers)	0.0049	0.0036	0.0008	-0.0000
	(0.0046)	(0.0061)	(0.0006)	(0.0003)
Policy (Likely Invoice Sellers)	0.0071	-0.0115	-0.0018	-0.0009
	(0.0084)	(0.0183)	(0.0016)	(0.0008)
Firm FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓

Note: This table compares real production responses between firms more likely to be legitimate producers and firms more likely to be invoice sellers. The classification is based on firms' pre-reform VAT remittance relative to assets. Firms with zero pre-reform VAT-to-asset ratios are classified as likely invoice sellers, while the remaining firms are classified as likely legitimate producers. The reported coefficients are estimated separately for the two groups using the baseline specification in Equation (1). All dependent variables are normalized with the firm's total assets in the preceding year and winsorized at the 1st and 99th percentiles. All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.10: Real Production Responses of State-Owned Enterprises

	(1)	(2)	(3)	(4)
Panel A. Investments and Outsourcing Services				
Dependent Variable	Fixed Assets Investment	Intangible Assets Investment	Processing Service	Transportation Service
Policy (TWFE)	-0.0011	-0.0018	0.0059*	0.0002
	(0.0057)	(0.0014)	(0.0034)	(0.0008)
Policy (CSDiD)	-0.0018	-0.0006	0.0028	-0.0017
	(0.0086)	(0.0004)	(0.0022)	(0.0013)
Observations	5,675	5,675	5,675	5,675
Panel B. Energy Consumption				
Dependent Variable	Electricity	Water	Coal	Oil
Policy (TWFE)	0.0031	0.2524	0.0017	0.0078
	(0.0172)	(0.3185)	(0.0092)	(0.0087)
Policy (CSDiD)	0.0083	0.5334	0.0125	0.0146
	(0.0168)	(0.3677)	(0.0120)	(0.0091)
Observations	5,675	5,675	5,675	5,675
Panel C. Other Firm Characteristics				
Dependent Variable	Wage	Debt	Depreciation	Interest
Policy (TWFE)	0.0104	-0.0222	0.0071	-0.0003
	(0.0097)	(0.0218)	(0.0055)	(0.0013)
Policy (CSDiD)	0.0099	-0.0231	0.0081*	-0.0001
	(0.0107)	(0.0160)	(0.0046)	(0.0018)
Observations	5,675	5,675	5,675	5,675
Firm FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓

Note: This table reports the real production responses of state-owned enterprises (SOEs) to the presumptive tax reform. Panel A reports investment and outsourced services, Panel B reports energy consumption, and Panel C reports other firm characteristics. All dependent variables are normalized with the firm's total assets in the preceding year and winsorized at the 1st and 99th percentiles. Policy (TWFE) reports estimates from the baseline specification in Equation (1), estimated using OLS. Policy (CSDiD) reports the estimator of Callaway and Sant'Anna (2021). All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. The unit of electricity is 10,000 kWh, and the unit of other energy consumption variables is tons.

Table A.11: Heterogeneous Effects on Firm Exit by Pre-Reform Suspicious Invoice History

	(1) Suspicious	(2) Non-Suspicious	(3) Suspicious	(4) Non-Suspicious
Panel A. Direct Effect				
Policy (TWFE)	0.1933***	0.3163*		
	(0.0608)	(0.1656)		
Policy (CSDiD)			0.1793**	0.2834**
			(0.0862)	(0.1325)
Observations	19,656	19,656	19,656	19,656
Effect Normalized by Average Firm Count	0.112	0.026	0.104	0.023
Industry × Province FE	✓	✓	✓	✓
Province × Year FE	✓	✓	✓	✓
Panel B. Spillover Effect to Downstream				
Policy (TWFE)	0.0228**	0.1251**		
	(0.0100)	(0.0550)		
Policy (CSDiD)			0.0313***	0.1226***
			(0.0096)	(0.0457)
Observations	23,100	23,100	23,100	23,100
Effect Normalized by Average Firm Count	0.011	0.005	0.015	0.005
Industry × Province FE	✓	✓	✓	✓
Province × Year FE	✓	✓	✓	✓

Note: This table reports heterogeneous effects of the presumptive tax reform on firm exit by firms' pre-reform suspicious invoice history. The outcome variable is the number of exiting firms in each province-industry-year cell. Firms are classified as having pre-reform suspicious invoice history if they ever reported deductible inputs exceeding total costs pre-reform, our indicator of likely fake invoice use. Panel A reports results for agro-processing firms, and Panel B reports results for downstream wholesale and retail firms. Columns (1) and (2) report TWFE estimates from the market-level specification in Equation (3), estimated using OLS, for firms with and without pre-reform suspicious invoice history, respectively. Columns (3) and (4) report the corresponding CSDiD estimates. The dependent variable is the number of exiting firms, so the coefficients are not directly comparable across groups. We therefore normalize the estimated effect by dividing it by the average number of firms labeled by the suspicious invoice indicator in each province-industry-year cell. All specifications include industry-by-province fixed effects and province-by-year fixed effects. Standard errors are clustered at the four-digit industry level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.12: Effects on Potential Concealment of Sales

	(1)	(2)
Dependent Variable	Accounts Payable	Accounts Receivable
Panel A. TWFE Estimator		
Policy	-0.0006	0.0001
	(0.0008)	(0.0022)
Observations	268,410	268,410
Adjusted R^2	0.56	0.61
Panel B. CSDiD Estimator		
Policy	-0.0010	0.0014
	(0.0010)	(0.0011)
Observations	268,410	268,410
Pre-Treat Mean of Dep. Var.	0.0528	0.0695
Firm FE	✓	✓
Year FE	✓	✓

Note: This table examines the impact of the presumptive tax reform on accounts payable and accounts receivable, two measures related to potential concealment of sales. Both dependent variables are normalized with the firm's total assets in the preceding year and winsorized at the 1st and 99th percentiles. Panel A reports estimates from the baseline specification in Equation (1), estimated using OLS. Panel B reports the corresponding average effects using the estimator of Callaway and Sant'Anna (2021). All specifications include firm and year fixed effects. Pre-Treat Mean of Dep. Var. is the pre-reform mean of the dependent variable for treated firms. Standard errors are clustered at the four-digit industry level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.13: Real Production Response of Downstream Wholesale and Retail Firms

	(1)	(2)	(3)	(4)
Panel A. Investments and Outsourcing Services				
Dependent Variable	Fixed Assets Investment	Intangible Assets Investment	Processing Service	Transportation Service
Policy (TWFE)	0.0019	0.0000	0.0001	0.0006
	(0.0017)	(0.0000)	(0.0002)	(0.0006)
Policy (CSDiD)	0.0019	0.0000	0.0002	0.0006
	(0.0017)	(0.0000)	(0.0001)	(0.0008)
Observations	422,190	422,190	422,190	422,190
Panel B. Energy Consumption				
Dependent Variable	Electricity	Water	Coal	Oil
Policy (TWFE)	0.0003	0.0093	0.0008	-0.0005
	(0.0015)	(0.0114)	(0.0006)	(0.0006)
Policy (CSDiD)	0.0013	0.0096	0.0007	-0.0006
	(0.0013)	(0.0119)	(0.0006)	(0.0006)
Observations	422,190	422,190	422,190	422,190
Panel C. Other Firm Characteristics				
Dependent Variable	Wage	Debt	Depreciation	Interest
Policy (TWFE)	0.0066	0.0045	0.0002	-0.0001
	(0.0234)	(0.0203)	(0.0005)	(0.0001)
Policy (CSDiD)	0.0018	0.0020	0.0001	-0.0001
	(0.0200)	(0.0191)	(0.0005)	(0.0001)
Observations	422,190	422,190	422,190	422,190
Firm FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓

Note: This table reports the real production responses of downstream wholesale and retail firms to the presumptive tax reform. Panel A reports investment and outsourced services, Panel B reports energy consumption, and Panel C reports other firm characteristics. All dependent variables are normalized with the firm's total assets in the preceding year and winsorized at the 1st and 99th percentiles. Policy (TWFE) reports estimates from the baseline specification in Equation (1), estimated using OLS. Policy (CSDiD) reports the estimator of Callaway and Sant'Anna (2021). All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. The unit of electricity is 10,000 kWh, and the unit of other energy consumption variables is tons.

Table A.14: Real Production Responses of Exporting Firms

	(1)	(2)	(3)	(4)
Panel A. Investments and Outsourcing Services				
Dependent Variable	Fixed Assets Investment	Intangible Assets Investment	Processing Service	Transportation Service
Policy (TWFE)	0.0007	0.0012	0.0016	0.0006
	(0.0018)	(0.0009)	(0.0014)	(0.0005)
Policy (CSDiD)	-0.0006	-0.0003	0.0012	-0.0001
	(0.0021)	(0.0010)	(0.0016)	(0.0005)
Observations	84,847	84,847	84,847	84,847
Panel B. Energy Consumption				
Dependent Variable	Electricity	Water	Coal	Oil
Policy (TWFE)	-0.0037	0.2315	-0.0053	0.0032
	(0.0093)	(0.1481)	(0.0099)	(0.0033)
Policy (CSDiD)	-0.0104	0.1017	-0.0078	0.0053
	(0.0082)	(0.1618)	(0.0092)	(0.0040)
Observations	84,847	84,847	84,847	84,847
Panel C. Other Firm Characteristics				
Dependent Variable	Wage	Debt	Depreciation	Interest
Policy (TWFE)	-0.0037	-0.0003	0.0000	-0.0005
	(0.0056)	(0.0091)	(0.0010)	(0.0003)
Policy (CSDiD)	-0.0073	0.0043	0.0010	-0.0006
	(0.0056)	(0.0065)	(0.0007)	(0.0005)
Observations	84,847	84,847	84,847	84,847
Firm FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓

Note: This table presents the real production responses of exporting firms to the presumptive tax reform. Panel A reports investment and outsourced services, Panel B reports energy consumption, and Panel C reports other firm characteristics. All dependent variables are normalized with the firm's total assets in the preceding year and winsorized at the 1st and 99th percentiles. Policy (TWFE) reports estimates from the baseline specification in Equation (1), estimated using OLS. Policy (CSDiD) reports the estimator of Callaway and Sant'Anna (2021). All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. The unit of electricity is 10,000 kWh, and the unit of other energy consumption variables is tons.

Table A.15: Main Results Controlling for the Potential Effects of Golden Tax Project III

	(1) TWFE	(2) CSDiD
Panel A. Direct Effect		
VAT Remittance	0.0044***	0.0049***
	(0.0007)	(0.0007)
Input Tax Credit	-0.0081***	-0.0083***
	(0.0015)	(0.0010)
Output Tax	-0.0059***	-0.0047**
	(0.0012)	(0.0022)
Panel B. Spillover Effect to Downstream		
VAT Remittance	0.0080***	0.0091***
	(0.0012)	(0.0014)
Input Tax Credit	-0.0133***	-0.0141***
	(0.0038)	(0.0039)
Output Tax	-0.0033	-0.0007
	(0.0044)	(0.0044)
Panel C. Spillover Effect to Export		
Tax Refund	-0.0027***	-0.0027***
	(0.0005)	(0.0004)
Firm FE	✓	✓
Year FE	✓	✓

Note: This table reports the effects of the presumptive tax reform after controlling for potential confounding effects from the Golden Tax Project III (GTP III). We account for GTP III by including controls for its staggered provincial rollout during 2013–2015. Panel A reports the direct effects on agro-processing firms. Panel B reports the spillover effects on downstream wholesale and retail firms. Panel C reports the spillover effect on export tax refunds. All dependent variables are normalized with the firm's total assets in the preceding year and winsorized at the 1st and 99th percentiles. Column (1) reports estimates from the baseline specification in Equation (1), estimated using OLS with the GTP III rollout controls. Column (2) reports the estimator of Callaway and Sant'Anna (2021), with the same GTP III controls. All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.16: Robustness to Pre-Reform Covariates Interacted with Year Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)
VAT Remittance	0.0043***	0.0056***	0.0058***	0.0056***	0.0034***	0.0059***
	(0.0007)	(0.0013)	(0.0009)	(0.0008)	(0.0008)	(0.0007)
Input Tax Credit	-0.0111***	-0.0074***	-0.0076***	-0.0081***	-0.0117***	-0.0071***
	(0.0015)	(0.0024)	(0.0018)	(0.0015)	(0.0017)	(0.0014)
Output Tax	-0.0039***	-0.0104***	-0.0083***	-0.0078***	-0.0029**	-0.0089***
	(0.0015)	(0.0019)	(0.0013)	(0.0013)	(0.0014)	(0.0014)
Deductible Inputs>Total Costs?	-0.0420***	-0.0568***	-0.0596***	-0.0585***	-0.0430***	-0.0626***
	(0.0155)	(0.0198)	(0.0169)	(0.0164)	(0.0162)	(0.0160)
Firm FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Pre-reform cov. × Year FE	Log Assets	Ag. Inputs over Assets	Fixed Assets over Assets	Electricity over Assets	Wages over Assets	Debt over Assets

	(7)	(8)	(9)	(10)	(11)
VAT Remittance	0.0063***	0.0065***	0.0048***	0.0057***	0.0046***
	(0.0010)	(0.0010)	(0.0008)	(0.0008)	(0.0012)
Input Tax Credit	-0.0067***	-0.0052***	-0.0097***	-0.0080***	-0.0075***
	(0.0015)	(0.0014)	(0.0017)	(0.0015)	(0.0028)
Output Tax	-0.0104***	-0.0115***	-0.0080***	-0.0080***	-0.0102***
	(0.0014)	(0.0014)	(0.0013)	(0.0012)	(0.0016)
Deductible Inputs>Total Costs?	-0.0641***	-0.0701***	-0.0316**	-0.0587***	-0.0494***
	(0.0149)	(0.0141)	(0.0145)	(0.0167)	(0.0185)
Firm FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
Pre-reform cov. × Year FE	Sales over Assets	Deductible Inputs over Assets	Deductible Inputs over Costs	Electricity over Sales	All Covariates

Note: This table reports robustness checks that control for pre-reform covariates interacted with year fixed effects. The row labeled Pre-reform cov. × Year FE indicates which pre-reform covariate is interacted with year fixed effects in each column. Baseline VAT remittances, input tax credits, output taxes, and the suspicious invoice indicator are excluded because they are directly related to the main outcome variables. Columns (1)–(10) add one pre-reform covariate at a time, while Column (11) includes all pre-reform covariates interacted with year fixed effects. All specifications include firm and year fixed effects. Standard errors clustered at the four-digit industry level are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.17: Placebo Test with Other Taxes and Fees

	(1)	(2)	(3)
Panel A. Direct Effect			
Dependent Variable	PIT	Property Tax	Social Insurance
Policy (TWFE)	0.00003	-0.00002	-0.00040
	(0.00004)	(0.00002)	(0.00046)
Policy (CSDiD)	0.00004	-0.00003	0.00002
	(0.00006)	(0.00002)	(0.00041)
Observations	268,410	268,410	268,410
Panel B. Spillover Effect to Downstream			
Dependent Variable	PIT	Property Tax	Social Insurance
Policy (TWFE)	-0.00004	-0.00002	0.00194
	(0.00004)	(0.00002)	(0.00309)
Policy (CSDiD)	-0.00006	-0.00001	0.00155
	(0.00004)	(0.00001)	(0.00272)
Observations	422,190	422,190	422,190
Firm FE	✓	✓	✓
Year FE	✓	✓	✓

Note: This table reports placebo tests using other tax and fee outcomes. Panel A reports results for agro-processing firms, and Panel B reports results for downstream wholesale and retail firms. The dependent variables are personal income tax, property tax, and social insurance payments. All dependent variables are normalized with the firm's total assets in the preceding year and winsorized at the 1st and 99th percentiles. Policy (TWFE) reports estimates from the baseline specification in Equation (1), estimated using OLS. Policy (CSDiD) reports the estimator of Callaway and Sant'Anna (2021). All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.18: Main Results with Alternative Winsorizing Levels

	(1) TWFE	(2) CSDiD	(3) TWFE	(4) CSDiD
	98.5%		99.5%	
Panel A. Direct Effect				
VAT Remittance	0.0042***	0.0046***	0.0047***	0.0051***
	(0.0006)	(0.0007)	(0.0007)	(0.0007)
Input Tax Credit	-0.0079***	-0.0082***	-0.0083***	-0.0085***
	(0.0013)	(0.0010)	(0.0016)	(0.0010)
Output Tax	-0.0058***	-0.0048**	-0.0058***	-0.0048**
	(0.0011)	(0.0021)	(0.0012)	(0.0022)
Panel B. Spillover Effect to Downstream				
VAT Remittance	0.0076***	0.0088***	0.0084***	0.0093***
	(0.0010)	(0.0012)	(0.0012)	(0.0014)
Input Tax Credit	-0.0115***	-0.0132***	-0.0151***	-0.0161***
	(0.0032)	(0.0034)	(0.0044)	(0.0052)
Output Tax	-0.0005	0.0006	-0.0060	-0.0038
	(0.0037)	(0.0038)	(0.0053)	(0.0066)
Panel C. Spillover Effect to Export				
Tax Refund	-0.0025***	-0.0025***	-0.0029***	-0.0029***
	(0.0005)	(0.0004)	(0.0006)	(0.0004)

Note: This table reports the main results using alternative winsorizing levels. Columns (1) and (2) winsorize the dependent variables at the 1.5th and 98.5th percentiles, while Columns (3) and (4) winsorize them at the 0.5th and 99.5th percentiles. Panel A reports the direct effects on agro-processing firms. Panel B reports the spillover effects on downstream wholesale and retail firms. Panel C reports the spillover effect on export tax refunds. All dependent variables are normalized with the firm's total assets in the preceding year. Columns (1) and (3) report estimates from the baseline specification in Equation (1), estimated using OLS. Columns (2) and (4) report the estimator of Callaway and Sant'Anna (2021). All specifications include firm and year fixed effects. Standard errors are clustered at the four-digit industry level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

References

Almunia, Miguel, Jonas Hjort, Justine Knebelmann, and Lin Tian. 2024. "Strategic or Confused Firms? Evidence from "Missing" Transactions in Uganda." *Review of Economics and Statistics*, 106(1): 256–265.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. 2024. "Revisiting Event-Study Designs: Robust and Efficient Estimation." *The Review of Economic Studies*, 91(6): 3253–3285.

Callaway, Brantly, and Pedro H.C. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics*, 225(2): 200–230.

Cnossen, Sijbren. 2018. "VAT and Agriculture: Lessons from Europe." *International Tax and Public Finance*, 25: 519–551.

Keen, Michael, and Stephen Smith. 2006. "VAT Fraud and Evasion: What Do We Know and What Can Be Done?" *National Tax Journal*, 59(4): 861–887.

Li, Haonan, and Xuan Wang. 2026. "Firm Responses to Tax Audits: Regression Discontinuity Evidence from a Threshold-Based Audit Program." *The Economic Journal*, 136(674): 560–601.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics*, 225(2): 175–199.