Producer Services in China

Xing Hu* Ming Li† Hongsong Zhang‡

This Draft: October 12, 2026

Abstract

Manufacturing firms outsource producer services from external suppliers, yet micro-evidence on the impact of service outsourcing on firm production remains limited. Using value-added tax invoice records on Chinese manufacturing firms from 2017 to 2018, we identify firms' service outsourcing and document a correlation between purchased services and firm growth. We then develop a make-or-buy model and a two-regime production-function estimation to evaluate the impact of producer services as an input on firm production, while recovering the technology of in-house service production. The estimates imply modest increasing returns in final-good production and decreasing returns in in-house service production. Relative to the final-good technology, the in-house service production is more capital-intensive. We then exploit China's 2018 VAT credit refund as a quasi-experiment to show that high-rate service purchases increased with pre-reform credit exposure along both the extensive and intensive margins, accompanied by higher post-reform revenues, profit rates, and local supplier sales. Our policy evaluations indicate that the refund policy generates $1.41 billion in output gains, while fully exempting service purchases from VAT tied to firms' stranded credits would yield $1.77 billion.

Keywords: Service inputs; Production outsourcing; Production function estimation; Value-added tax; Input tax credits

JEL Codes: D22, D24, H25, L23, L60

*HKU Business School, The University of Hong Kong. Email: huxing@connect.hku.hk.
†Department of Economics, The Chinese University of Hong Kong. Email: mingli001@cuhk.edu.hk.
‡HKU Business School, The University of Hong Kong. Email: hszhang@hku.hk.

1 Introduction

Manufacturing firms could outsource producer services and benefit from specialization. In the process of production, a firm can perform tasks such as design, processing, repair, transport, and administration using its own internal factors, or it can contract them out to external suppliers. While this make-or-buy decision can enhance firm productivity through matching and specialization, convincing micro-level evidence on the impact of service outsourcing on firm production remains limited. A primary obstacle is empirical: conventional production data record labor, materials, and capital without disentangling the service tasks embedded within intermediate expenditure. As a result, purchased services are typically pooled with materials or observed only for the selected subsample of firms that actively buy them.

These data limitations leave two key questions unanswered. First, how much do purchased services add to firm output? Second, how do tax policies affect the choice between buying services and making them in house? Answering the first question requires a basic distinction: zero purchases do not mean zero service input. Non-purchasing firms often perform the same tasks with their own workers and equipment. Treating them as zero-input users confounds the make-or-buy choice with input intensity, while dropping them creates a severe selection bias. Neither approach recovers the in-house technology. The second question centers on tax frictions under a value-added tax (VAT). While input VAT is neutral under immediate deduction, firms with stranded credits bear a positive shadow cost on taxed market purchases. Internal labor, by contrast, generates no VAT. This asymmetry creates an effective tax penalty on outsourcing that varies with the firm's credit balance, category tax rates, and local supplier access. Addressing both margins is essential for measuring the true returns to services and evaluating how tax policies influence firms' choice.

China provides a useful setting in which to study these forces. The service share of aggregate value added rose substantially during the past two decades, and producer services in China grows even more than other major economies. At the same time, China's VAT invoice system records the buyer, seller, value, tax rate, and item code for business-to-business transactions. We merge these invoices with Corporate Income Tax (CIT) settlement records containing employment, assets, revenue, and industry. The resulting data reveal both the services purchased from other firms and the factors used within the purchasing firm. They also allow us to classify service inputs into production, research and technical, demand-related, administrative, transport, and other producer services.

Three facts motivate the empirical analysis. First, service purchasing is common. The share of firm-quarters with positive service purchases rose from 67.8 percent in 2017 to 75.3 percent

in 2018, creating a sizable extensive margin within the panel. Second, service expenditure varies considerably across categories. Services account for 4.82 percent of aggregate variable costs. Production services dominate spending for typical firms, while transport services represent a larger share among large firms. Third, purchasing status tracks production patterns. In raw data, buyers operate at larger scale. Within the same firm across quarters, purchasing periods exhibit lower labor, materials, and capital use per unit of output. These patterns indicate that service purchasing reflects an endogenous choice, precluding standard estimation on the buyer subsample alone.

Based on the facts, we develop a make-or-buy model in which service tasks are essential inputs. A firm can purchase these producer-service tasks or produce them using its own labor, materials, and capital. Local service supply lowers both the price index and the fixed contracting cost. The VAT system alters the effective purchase price through the interaction between category tax rates and the shadow cost of excess credits. These shifters shift the outsourcing threshold without directly altering technology. The model delivers the instruments, their expected signs, the conditioning set for the control function, and a joint estimating equation across both purchasing regimes. Estimating production functions on active buyers introduces selection on unobserved productivity, as highlighted by Olley and Pakes (1996). Our two-regime estimator jointly models buyers and non-buyers, imposing the productivity process on the shared latent state.

The 2018 refund of excess VAT credits provides complementary reduced-form variation. Document No. 70 lowered the shadow cost of input credits for eligible manufacturing industries. Because effective price wedges scale with statutory rates, the model predicts stronger purchasing responses for high-rate categories, such as processing, repair, and transport. We test this prediction by comparing high- and low-rate services within the same firm-quarter across varying pre-reform credit exposures. This design isolates policy-driven price responses from the production-function parameters. We subsequently combine both sets of estimates to calculate counterfactual changes in output.

The empirical analysis yields three main findings. First, accounting for the extensive margin substantially raises the estimated output elasticity of services. A pooled control function using $\ln(1+S)$ yields a service elasticity of 0.0099. By contrast, our two-regime estimator yields a conditional elasticity of 0.0374 and recovers the parameters of in-house service production. Relative to final-good production, in-house service tasks are more capital-intensive: column (2) of Table 4 reports $\beta_k = 0.1553$ compared to $\alpha_K = 0.0205$, with $\beta_l = 0.0735$ and $\beta_m = 0.6211$. Returns to scale in final-good production are 1.0629.

Second, the credit refund generated larger increases in high-rate service purchases among firms with higher pre-reform credit exposure. The triple-difference coefficient combining both margins is 0.0269. The corresponding estimate is 0.0238 for spending conditional on purchase, alongside a 0.0090 increase in purchasing probability. The policy also improved downstream manufacturer performance and stimulated local supplier sales: the user revenue coefficient is 0.0325, and the local supplier sales coefficient is 0.2995. These estimates support the effective-price mechanism and demonstrate the transmission of buyer-side tax relief to the local service market.

Third, policy simulations indicate measurable aggregate output gains. Applying the observed credit-refund policy at each firm's measured exposure increases service purchases by 2.20 percent after accounting for local-supply feedback, yielding $1.41 billion in output gains. Fully exempting purchased services from VAT tied to each firm's stranded-credit exposure raises service purchases by 2.83 percent and output by $1.77 billion.

This paper makes three contributions. The first is to the literature on firm boundaries and service outsourcing. Property-rights theories show how incomplete contracts and residual control shape the make-or-buy margin (Grossman and Hart, 1986; Hart and Moore, 1990; Grossman and Helpman, 2002; Antràs, 2003; Antràs and Helpman, 2004). Relational-contracting models highlight repeated interactions in task allocation (Baker et al., 2002). Empirical studies document domestic outsourcing of support and logistics tasks (Abraham and Taylor, 1996; Goldschmidt and Schmieder, 2017), link thicker local markets to lower contracting frictions (Ono, 2007), and relate supplier proximity to firm performance (Bernard et al., 2019). Sourcing studies examine the location and ownership of production stages (Grossman and Rossi-Hansberg, 2008; Fort, 2017; Alfaro et al., 2019), with recent work focusing on market-access services under export complexity (Berlingieri and Pisch, 2026). Evidence that specialized intermediate inputs improve productivity (Amiti and Konings, 2007; Halpern et al., 2015) supports treating purchased services as distinct inputs. We contribute by tracking domestic buyer–supplier transactions across six service categories and linking both purchasing margins directly to internal factor use.

The second contribution is to production-function estimation. Control-function methods use flexible inputs to proxy for productivity (Levinsohn and Petrin, 2003; Ackerberg et al., 2015; Wooldridge, 2009), while selection models incorporate endogenous participation into the productivity process (Olley and Pakes, 1996). Dynamic panel methods provide an alternative route using predetermined instruments (Blundell and Bond, 2000). Identification is challenging in gross-output settings with multiple flexible inputs (Gandhi et al., 2020), and revenue measures combine physical productivity with markups (Klette and Griliches, 1996;

De Loecker, 2011; De Loecker et al., 2016; Foster et al., 2008). Flexible-input conditions also facilitate markup recovery (De Loecker and Warzynski, 2012; De Loecker et al., 2020). Endogenous productivity models show how directed investment enters the Markov process (Doraszelski and Jaumandreu, 2013), and technology-regime studies demonstrate parameter heterogeneity across operating modes (Collard-Wexler and De Loecker, 2015; Syverson, 2004). In our setting, materials and purchased services are both flexible inputs. Services enter production only when firms cross an endogenous outsourcing threshold. The two-regime estimator integrates this threshold into the control function, exploits excluded variation from effective tax prices and local supply, and identifies in-house technology alongside the conditional elasticity of purchased services.

The third contribution is to the empirical literature on VAT frictions and tax policy. Theoretical analyses highlight the self-enforcing nature of the VAT alongside the distortions caused by incomplete refunds and rate disparities (Keen and Lockwood, 2010; Pomeranz, 2015; Naritomi, 2019). Micro-level studies show that VAT structure and enforcement affect firm reporting, capital accumulation, and trade flows (Carrillo et al., 2017; Liu and Mao, 2019; Chandra and Long, 2013; Benzarti and Carloni, 2019; Bachas et al., 2024). Prior evaluations of China's 2018 reform focus on total factor productivity (Guo and Li, 2023; Liu et al., 2024), labor income shares (Yang and Si, 2025), innovation efficiency (Huang et al., 2025), environmental investment (Chen et al., 2024), green strategy (Liao and Li, 2026), and digital transformation (Li et al., 2026a). Rich administrative tax data also allow precise identification of firm responses to tax incentives (Chen et al., 2021). Our study focuses directly on service purchasing as the primary outcome. The cross-category design separates extensive and intensive margin responses, tests the effective-price channel, and tracks market spillovers to local suppliers. Combining these empirical responses with production estimates allows quantitative evaluation of broader tax exemptions under transparent partial-equilibrium assumptions.

The rest of the paper proceeds as follows. Section 2 describes the administrative data and institutional background. Section 3 documents service suppliers, usage, industry composition, and input patterns. Section 4 develops and estimates the make-or-buy model. Section 5 studies the VAT credit refund and presents the counterfactual exercises. Section 6 concludes.

2 Data and Institutional Background

2.1 Data

Our empirical analysis focuses on the contributions of producer services in the production, which requires firm-level records to separate service tasks from materials. The analysis mainly uses two datasets from administrative sources.

Value-added tax invoices records. The first source is the universe of VAT invoices issued in mainland China, collected by the State Taxation Administration. The same records underlie Li et al. (2026b), who use them to study the structure of firm-to-firm production networks; we use the 2017–2018 window and exploit a different dimension of the records. Each invoice names the selling firm, the buying firm, and the value of the transaction, and carries an item code describing what was sold, together with quantity, unit price, and the statutory VAT rate applied. According to the tax law of mainland China, a buyer cannot deduct input VAT without a special invoice, so both sides of a business-to-business transaction have an incentive to have it issued and recorded, and misreporting carries legal penalties. These invoice and crediting rules follow the Provisional VAT Regulations and the rules that completed the nationwide VAT transition (State Council of the People's Republic of China, 2017; Ministry of Finance and State Administration of Taxation, 2016). Sales to final consumers generate no deduction and are largely absent, which is immaterial here because our unit of analysis is an input purchase. Small-scale taxpayers below the reporting threshold are exempt, so the smallest firms are underrepresented on both sides of the market.

What makes these records usable for a make-or-buy question is the item code. We aggregate invoices to the buyer $\times$ quarter level and sort each line into materials, energy, equipment, or one of six service categories — production services (S^P), research and technical services (S^R), demand-related services (S^D), administrative services (S^A), transport (S^T), and other services (S^O) — using the code prefix. The classification follows the economic question testing whether the purchased item is a task the buyer could instead have carried out with its own factors. Contract processing and repair is assigned to S^P, since subcontracting a processing step is precisely the margin the paper is about.

Corporate income tax settlement records. The invoice data describe purchases but say nothing about the firm's own factors. For those we use the annual settlement records of the corporate income tax, which cover firms subject to annual filing and report employment, total assets, sales, and a detailed industry code. Relative to the survey panels conventionally

used for Chinese manufacturing firms (e.g., Brandt et al., 2017), these records extend below the above-scale survey threshold and cover firms outside manufacturing. Because they are filed under a tax obligation and checked by the tax authority, they are also less exposed to the recall and format problems of survey instruments. The cost is frequency: the settlement is annual, so employment and the asset stock enter our quarterly panel as within-year anchors, and we discuss in Section 4 what that implies for which input elasticities can be identified.

The dependent variable in the production-function estimation is quarterly revenue deflated by an industry price index. Revenue, materials expenditure, and purchased services are constructed from quarterly invoices. Capital is based on the annual asset stock, using within-year depreciation backcasting together with quarterly capital equipment rentals observed in the invoices; labor combines annual corporate income tax headcount with quarterly labor dispatch observed in the invoices to obtain quarterly-level labor and capital inputs. Materials, capital, and services enter as the deflated indexes used in the production-function data.

We merge the two sources by firm identifier and restrict attention to manufacturing firms observed between 2017Q1 and 2018Q4, which gives a firm $\times$ quarter panel of 4,183,493 observations on 948,394 distinct firms, of which roughly 730,000 to 740,000 are present in any given year. The production-function sample in Table 4 is smaller, at 2,874,030 observations on 738,068 firms, because the productivity process requires a lagged observation and the local-supply instrument requires the firm's city to be identified. Table 1 reports summary statistics for the estimation sample. Three features are worth noting before the analysis. The typical firm is small: median sales are \$0.30 million and median employment is 15. Materials dominate the cost structure, at 66.7 percent of sales at the median. And purchased services are a small expenditure item — 0.6 percent of sales at the median, 1.5 percent among firms that buy — but one that three quarters of firm-quarters record.

2.2 Institutional Background

China's aggregate production structure shifted markedly toward services between 1998 and 2018. Panel A of Figure 1 shows a sustained increase in the service share of GDP alongside a decline in the industry share. The broader service sector, however, includes consumer-facing activities that are not inputs into manufacturing. Panel B therefore isolates producer services, defined as information and communication, finance, professional and technical services, administrative support, and transportation and storage. China's producer-service share also rose, but remained below that of the major advanced economies shown in the figure. This gap motivates an analysis of the firms that use such services and the frictions that govern their purchases.

Table 1: Summary Statistics

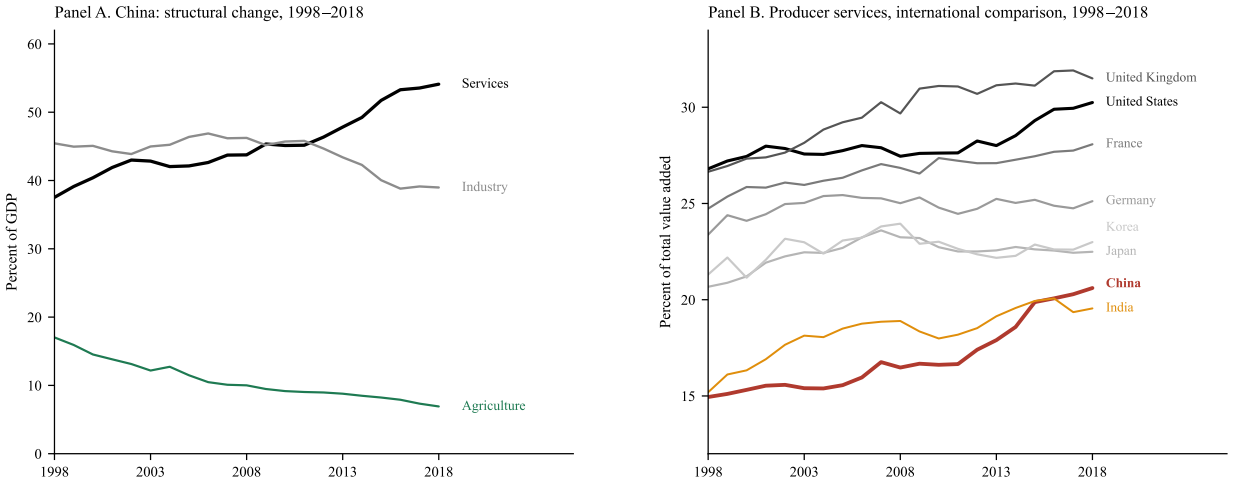
Statistics	Median	Mean	SD	IQR	IDR
Panel A: Output and own factors					
Sales (million RMB)	2.018	6.029	12.200	4.947	14.329
Materials expenditure (million RMB)	1.306	4.039	8.488	3.268	9.572
Capital (million RMB)	4.964	22.085	85.691	13.553	45.987
Employment (headcount)	15	37.095	63.592	32	87
Panel B: Purchased service inputs					
Service purchases, six categories (million RMB)	0.013	0.209	0.663	0.106	0.492
of which dispatched labor (million RMB)	0.000	0.003	0.093	0.000	0.000
Panel C: Input intensities and the outsourcing margin					
Materials / sales (percent)	66.739	72.747	36.612	42.538	84.711
Service purchases / sales (percent)	0.626	3.167	6.945	3.136	8.928
conditional on buying (percent)	1.473	4.252	7.756	4.318	11.206
Buys services, D_{it}	1	0.745	0.436	1	1
Firms buying services	625,669 (84.77%)				
Total number of firms	738,068				
Firm-quarters buying services	2,140,540 (74.48%)				
Observations (estimation sample)	2,874,030				
Observations (full quarterly panel)	4,183,493				
Firms (full quarterly panel)	948,394				

Note: The sample is the estimation sample of Section 4.3, which is the full 2017–2018 quarterly panel of manufacturing firms less observations without a lagged value for the productivity process and observations whose city cannot be identified; the last block reports the size of the full panel for comparison. IQR is the interquartile range and IDR the interdecile range ($p_{90} - p_{10}$). Monetary values are nominal RMB. The six service categories are production services, research and technical services, demand-related services, administrative services, transport, and other services; a firm counts as buying if it records positive spending in any of them in that quarter.

VAT credits and the 2018 refund. China's VAT is collected on sales while allowing general taxpayers to deduct VAT paid on eligible inputs. In its textbook form, the credit-invoice design supports self-enforcement and production neutrality, but incomplete refunds and multiple rates can restore cascading (Keen and Lockwood, 2010; Pomeranz, 2015). When current output VAT is large enough to absorb input credits immediately, an additional taxed purchase is a pass-through. When input VAT exceeds output VAT, however, the unused balance is carried forward, and delay gives that purchase a positive shadow cost. This feature is particularly relevant for service outsourcing because purchased services generate input VAT, whereas wages used for the corresponding in-house task do not. Statutory rates also differ across service categories. The rate schedule and its 2018 adjustment follow Caishui No. 36 and No. 32 (Ministry of Finance and State Administration of Taxation, 2016, 2018a). During our sample, processing and repair were taxed at 17 or 16 percent, transport at 11 or 10 percent, and modern services at 6 percent.

Caishui [2018] No. 70 introduced refunds of excess credits for firms in 18 eligible industries

Figure 1: The Service Sector in China and in International Comparison



Sources: World Bank, World Development Indicators (Panel A); OECD Trade in Value Added, 2025 edition, drawn from the Inter-Country Input-Output tables, at current USD (Panel B).
Note: Panel A plots value added by broad sector as a share of GDP. Panel B measures producer services as ISIC Rev.4 divisions J (information and communication), K (financial and insurance), M (professional, scientific and technical), N (administrative and support) and H (transportation and storage), expressed as a share of total value added.

in the second half of 2018. The policy did not change the technology for producing a service task, but it reduced the shadow cost of input VAT for exposed firms. The difference in statutory rates implies a stronger effective-price change for high-rate categories. Section 4.2 formalizes this mechanism, and Section 5 estimates the corresponding reduced-form response of service purchases (Ministry of Finance and State Administration of Taxation, 2018b).

3 Producer Services in Chinese Manufacturing

To motivate our theoretical model and empirical strategy, this section establishes four stylized facts regarding the supply, adoption, industry intensity, and factor-saving patterns of producer services across Chinese manufacturing firms.

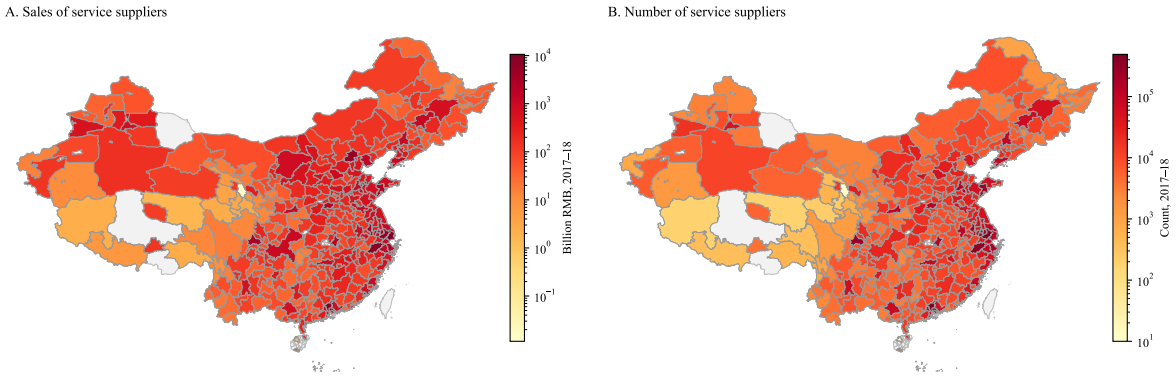
3.1 Geographic Distribution of Suppliers

Producer-service suppliers exhibit pronounced geographic concentration across mainland China. Figure 2 maps the spatial distribution of service providers, plotting total supplier sales in Panel A and the number of distinct suppliers in Panel B across prefecture-level cities over the 2017–2018 period. Both dimensions highlight extreme spatial clustering, with the vast majority of service activity concentrated in coastal economic centers and core metropolitan

regions (notably the Yangtze River Delta, the Pearl River Delta, and the Beijing–Tianjin–Hebei corridor), while interior regions display markedly sparser supplier networks.

Although sales volume and supplier counts share similar spatial footprints, they capture economically distinct margins of local service supply: total sales reflect the joint outcome of supplier entry and operational scale, whereas the provider count measures the depth and thickness of the local supplier pool. Because producer-service transactions often involve task-specific customization, frequent monitoring, and relationship-specific investments, thicker local markets reduce search frictions, lower contracting costs, and mitigate hold-up risks. This geographic variation establishes that firms' access to external services is heavily conditioned by their local environment, motivating our use of the standardized local service location quotient (LS_{ct}) as an excluded supply shifter in Section 4.

Figure 2: Geographic Distribution of Producer-Service Suppliers



Note: Panel A maps total sales by service suppliers, and Panel B maps the number of suppliers. Both measures aggregate the six service categories over 2017–2018.

3.2 Adoption and Spending

Our second finding is that while producer service purchasing is widespread, participation is fluid and displays substantial churn across quarters. Table 2 reports service adoption rates at both annual and quarterly frequencies. At the annual level, service purchasing is pervasive: the share of manufacturing firms purchasing at least one producer service rose from 78.74 percent in 2017 to 84.30 percent in 2018. However, examining the same firms at the quarterly frequency reveals that participation is intermittent, with quarterly adoption rates standing at 67.78 percent in 2017 and 75.29 percent in 2018. The persistent 9-to-11 percentage point wedge between annual and quarterly participation demonstrates that service sourcing is not a permanent, static organizational trait; rather, a substantial fraction

of manufacturers regularly move back and forth across the make-or-buy margin over time. This dynamic participation underscores the empirical necessity of explicitly modeling the extensive margin (D_{it}) rather than discarding non-purchasing quarters.

Table 2: Share of Firms Buying Service Inputs

	Firm × year		Firm × quarter	
	Frequency	Percentage	Frequency	Percentage
Year: 2017				
Buying service inputs	573,589	78.74%	1,414,558	67.78%
Not buying service inputs	154,861	21.26%	672,359	32.22%
Total	*728,450*	*100%*	*2,086,917*	*100%*
Year: 2018				
Buying service inputs	624,548	84.30%	1,578,541	75.29%
Not buying service inputs	116,321	15.70%	518,035	24.71%
Total	*740,869*	*100%*	*2,096,576*	*100%*

Note: A firm counts as buying service inputs if it records positive spending in any of the six service categories. The two panels differ in the unit of observation. At the firm × year level a firm counts as a buyer if it bought in any quarter of the year, so the participation rate is mechanically higher than at the firm × quarter level, where each quarter is counted separately. Both are reported because the two answer different questions: whether a firm uses purchased services at all, and how regularly it does so. The gap between them — eleven percentage points in 2017 and nine in 2018 — indicates that a substantial share of buyers purchase in some quarters but not others.

Table 3 turns to the intensive margin, reporting the expenditure share of purchased services in variable costs across three distinct averaging concepts. Across all six categories, purchased services constitute 4.55 percent of variable costs for the average firm-quarter (unweighted mean) and 4.82 percent when weighted by expenditure. While services represent a modest expenditure share relative to raw materials, this aggregate figure masks significant heterogeneity across functional categories. Production services (contract processing, assembly, and repair) dominate outsourcing budgets, representing an unweighted share of 2.79 percent across all firms and expanding to 5.04 percent conditional on buying. In contrast, transport services exhibit a much larger aggregate share (1.39 percent) than unweighted share (0.85 percent), reflecting disproportionate procurement by large-scale manufacturers. Specialized services, such as R&D, technical consulting, and demand-related activities, feature low unconditional adoption rates but takes sizable share (e.g., 1.85 percent of variable costs for demand services) among active users.

Table 3: Expenditure Share in Variable Cost of Service Input Categories

	Unweighted firm mean	Aggregate share	Mean among buyers
All services	4.55%	4.82%	12.78%
By category			
Production services	2.79%	2.63%	5.04%
Transport services	0.85%	1.39%	2.34%
Administrative services	0.24%	0.25%	1.26%
Demand services	0.11%	0.08%	1.85%
R&D and technical services	0.10%	0.11%	1.23%
Other services	0.46%	0.36%	1.06%

Note 1: The denominator is variable cost, defined as materials plus energy plus the six service categories; capital expenditure and leasing are excluded.

Note 2: The three columns are three different averages of the same underlying quantity and answer different questions. The unweighted firm mean averages S_b/variable cost across firm-quarters, giving every firm equal weight and counting non-buyers as zero: it describes the typical firm. The aggregate share is total spending on category b divided by total variable cost, so large firms dominate: it describes where manufacturing input spending as a whole goes. The mean among buyers conditions on positive spending: it describes how much a firm commits once it decides to buy at all.

Note 3: The three differ in informative ways. Transport is 0.85% of the typical firm's variable cost but 1.39% in the aggregate, because large firms buy proportionally more of it. Demand services are rarely bought but account for 1.85% of variable cost among firms that use them. Production services have the largest unweighted and conditional shares and are the only category that is both widely purchased and large for those who purchase.

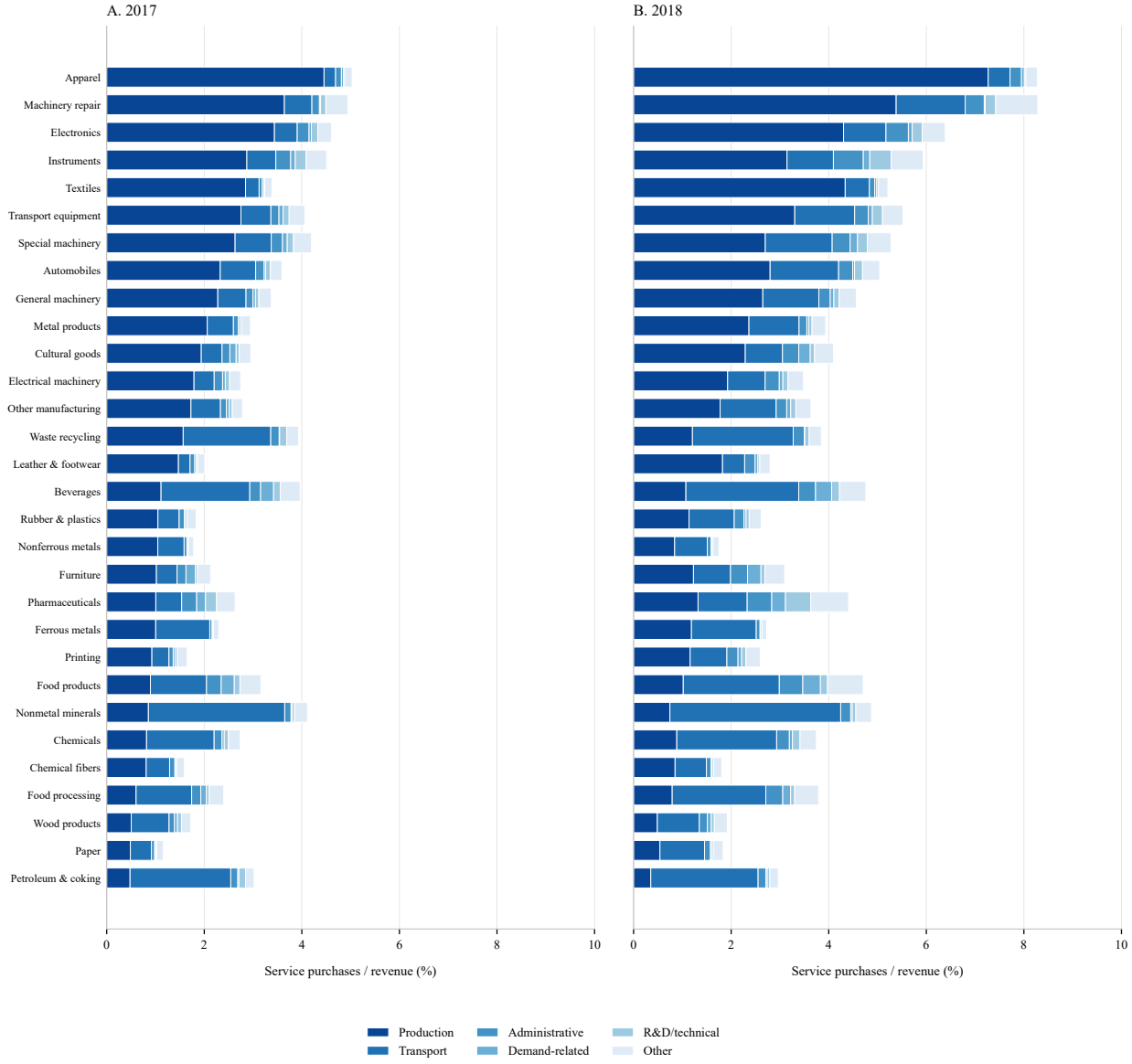
3.3 Industry Composition

The intensity of service input use varies systematically across two-digit manufacturing sectors. Figure 3 ranks industries by their production-service expenditure as a share of revenue in 2017 (Panel A) and applies the identical ranking to 2018 (Panel B). The cross-industry distribution exhibits a clear technological gradient: downstream assembly and discrete manufacturing industries—such as apparel, machinery repair, electronics, optical instruments, and transport equipment—exhibit the highest service intensities, with total service purchases reaching 6 to 8 percent of industry revenue. Conversely, heavy process industries (such as petroleum and coking, paper, and chemicals) allocate substantially smaller revenue shares to external services.

Crucially, this cross-industry ranking is exceptionally stable across the two sample years, indicating that the propensity to outsource intermediate service tasks is governed by persis-

tent, industry-specific production technologies rather than idiosyncratic, transitory demand shocks. At the same time, the varying combinations of production, transport, administrative, and technical services across sectors demonstrate that different manufacturing processes rely on distinct service tasks.

Figure 3: Service Input Intensity by Manufacturing Industry

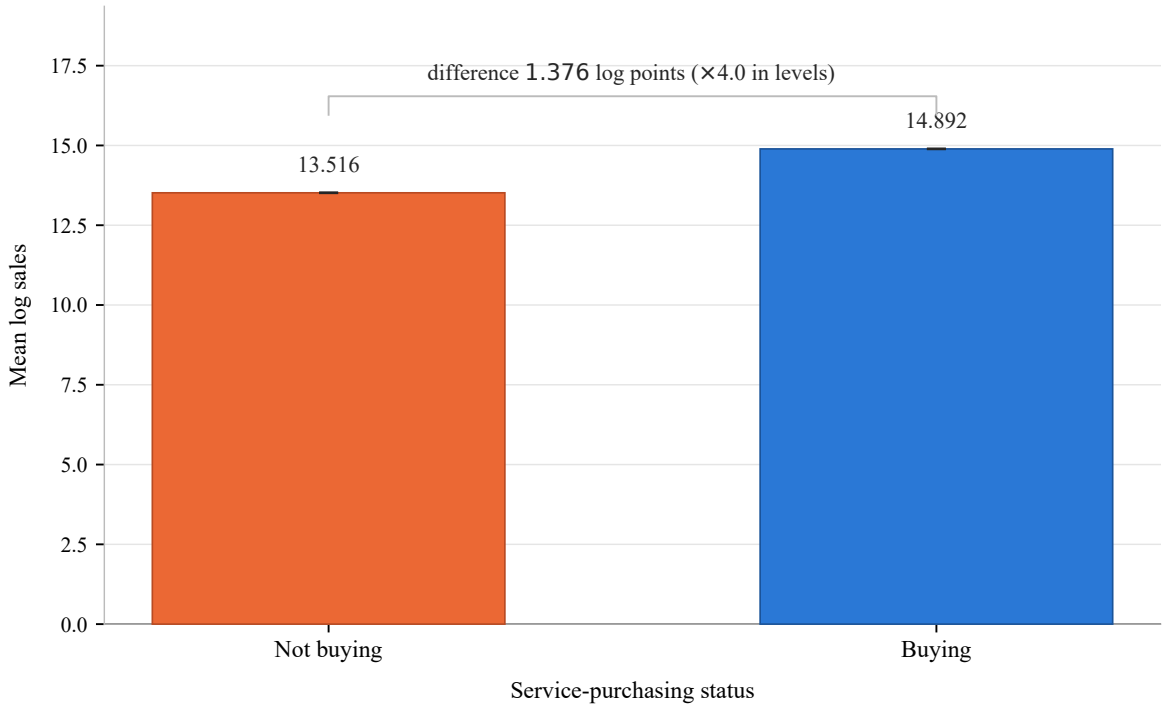


Note: The bars report purchases of services as a share of industry revenue and decompose those purchases into the current six service categories. Industries are ordered by production services as a share of industry revenue in 2017. Industry names are abbreviated from the two-digit manufacturing classification.

3.4 Firm Size and Input Use

A salient pattern in the micro data is that service-purchasing firms are substantially larger across all operational dimensions. As documented in Figure 4, firms that purchase services exhibit mean log sales that are 1.376 log points higher than those of non-purchasing firms—a fourfold difference in levels. This scale advantage is equally evident in internal factor employment: buyers operate with 1.114 log points more labor and 1.429 log points more capital. The proportional scale of these gaps across sales, employment, and capital indicates that service adoption strongly correlates with overall firm productivity and size, consistent with models where only larger firms generate sufficient volume to amortize the fixed contracting costs of external procurement. Consequently, standard production-function estimators restricted solely to active buyers would suffer from severe selection bias on unobserved productivity.

Figure 4: Firm Size by Service-Purchasing Status



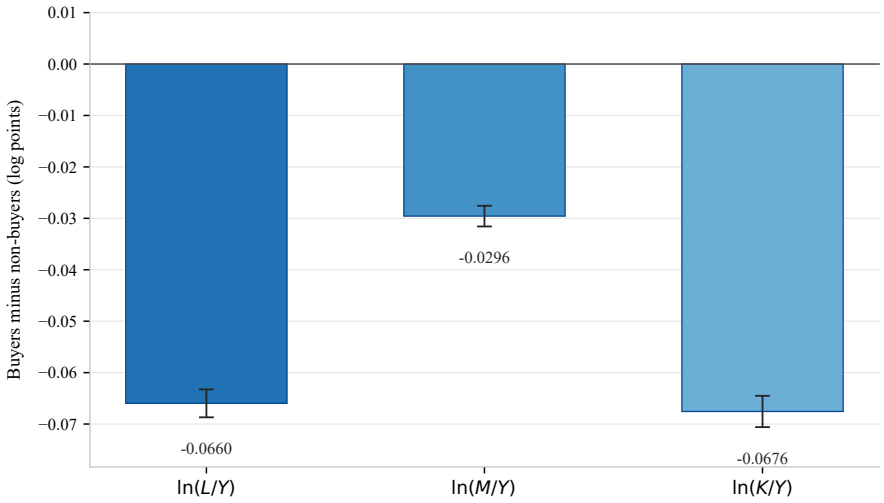
Note: Bars are group means of log sales with 95% confidence intervals from standard errors clustered at the firm; with 733,490 and 2,140,540 observations the intervals are narrower than the line width. The vertical axis starts at zero, so the visual step between the bars understates the difference: 1.376 log points is a factor of 4.0 in levels, and the bracket reports both. The same comparison for employment and capital gives gaps of 1.114 and 1.429 log points. The differences are descriptive: purchasing status is a choice, and these are not causal effects of buying on size.

While cross-sectional comparisons reflect firm sorting by scale, analyzing within-firm input

adjustments provides direct evidence on factor substitution. Figure 5 compares input-output ratios across purchasing and non-purchasing quarters within the same firm, absorbing firm fixed effects and industry-by-quarter shocks. When a firm switches from in-house provision to purchasing services, its internal factor use per unit of output drops significantly across all inputs: log labor per unit of output ($\ln(L/Y)$) declines by 0.0660 log points, materials ($\ln(M/Y)$) fall by 0.0296 log points, and capital ($\ln(K/Y)$) drops by 0.0676 log points.

The uniform negative sign across all three inputs confirms that purchased producer services substitute for internal tasks that the firm would otherwise have to perform using its own labor, materials, and equipment. Furthermore, the fact that internal capital and labor experience substantially larger reductions than materials suggests that in-house service tasks are particularly factor-intensive. While these within-firm associations remain descriptive because the purchasing choice responds endogenously to unobserved productivity shocks, they provide direct micro-level empirical grounding for our structural framework in Section 4, where in-house service production requires internal factor expenditure that can be relieved through external market sourcing.

Figure 5: Input–Output Ratios and Service Purchasing



Note: The bars report the coefficient on the service-purchasing indicator from separate regressions of $\ln(L/Y)$, $\ln(M/Y)$, and $\ln(K/Y)$. All specifications absorb firm and two-digit-industry-by-quarter fixed effects. Standard errors are clustered by firm; whiskers are 95% confidence intervals. The underlying sample contains 2,874,030 firm-quarter observations from 738,068 firms and 217 industry-by-quarter cells; singleton fixed-effect groups are omitted in estimation.

4 Empirical Model

This section develops a parsimonious empirical model of service sourcing. The model has two roles. First, it connects the observed zero-purchase decision to the production function, so that firms that buy services and firms that make the task in house can be estimated jointly. Second, it shows how local service supply and the shadow cost of excess input VAT shift purchasing without shifting technology, which motivates the instruments and the counterfactual mapping.

4.1 Setup

Time t is a quarter, i indexes manufacturing firms, and $x_{it} = (l_{it}, m_{it}, k_{it})$ denotes log labor, materials, and capital. The dependent variable q_{it} is industry-deflated revenue. Let $D_{it} = \mathbf{1}(S^B_{it} > 0)$ indicate a positive purchase of producer services and let $s_{it} = \ln S^B_{it}$ when $D_{it} = 1$.

Two production regimes. For firms that buy services, we write the log production function as

$$q_{it} = \alpha_L l_{it} + \alpha_M m_{it} + \alpha_K k_{it} + \alpha_S s_{it} + \omega_{it} + \varepsilon_{it}, \qquad D_{it} = 1, \tag{1}$$

where ω_{it} is productivity observed by the firm when it chooses inputs, $\alpha_S > 0$, and ε_{it} is an ex-post shock or measurement error. For firms that report no service purchase, the service task is not observed separately. We represent the task produced in house by

$$s^H_{it} = c_H + \beta_l l_{it} + \beta_m m_{it} + \beta_k k_{it} + \omega_{it}, \qquad D_{it} = 0. \tag{2}$$

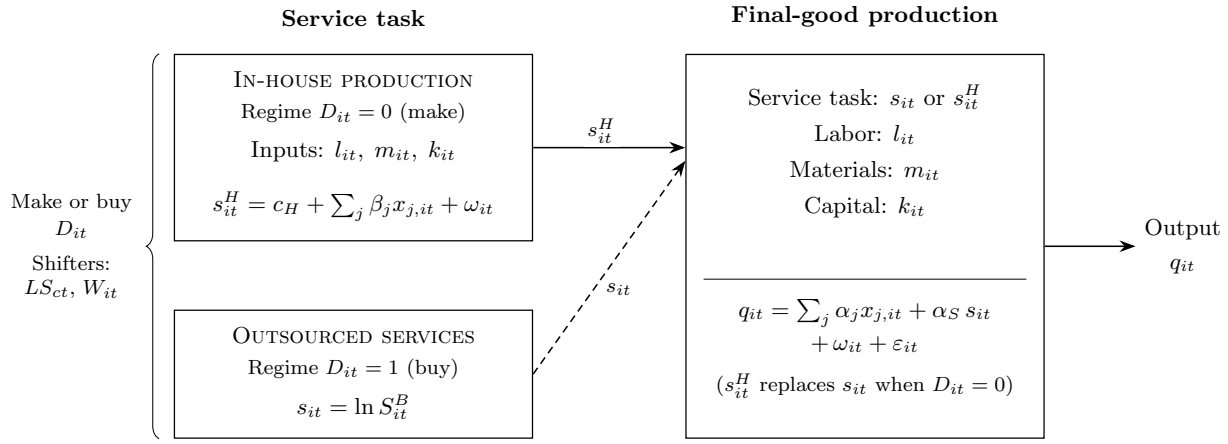
This equation is an empirical technology for the unobserved service task. The coefficients β_j are input elasticities. Substituting (2) into the production function for non-buyers and combining the two regimes gives

$$\begin{aligned} q_{it} = & \sum_{j \in \{L,M,K\}} \left[\alpha_j + (1 - D_{it})\theta_j\right] x_{j,it} + D_{it}\alpha_S s_{it} + (1 - D_{it})C_0 \\ & + \left[1 + (1 - D_{it})\alpha_S\right] \omega_{it} + \varepsilon_{it}, \end{aligned} \tag{3}$$

where $\theta_j \equiv \alpha_S \beta_j$ and $C_0 \equiv \alpha_S c_H$. Equation (3) is the estimating equation used below. $\alpha_L + \alpha_M + \alpha_K + \alpha_S$ measures returns to scale in final-good production, while $\beta_l + \beta_m + \beta_k$ measures returns to scale in the in-house service production.

Figure 6 summarizes this structure. The service task enters final-good production through one of two branches. A firm that buys uses the purchased quantity s_{it}, so its labor, materials, and capital enter output only through the final-good elasticities α_j. A firm that does not

Figure 6: Two-Regime Production Structure



Note: The figure illustrates Equations (1)–(3). The solid arrow denotes the in-house regime ($D_{it} = 0$), in which the service task s_{it}^H is produced from the firm's own labor, materials, and capital and is not observed separately. The dashed arrow denotes the purchasing regime ($D_{it} = 1$), in which the service task is measured by purchases $s_{it} = \ln S_{it}^B$ recorded on VAT invoices. Local supply LS_{ct} and the credit wedge W_{it} shift the make-or-buy choice and purchased quantities but are excluded from both technologies.

buy produces the task with the same three inputs. Each input then affects output twice: directly through α_j, and indirectly through the in-house mapping with weight $\theta_j = \alpha_S \beta_j$. The make-or-buy indicator D_{it} determines which branch applies. The shifters introduced in Section 4.2, local supply LS_{ct} and the credit wedge W_{it}, move firms across the branches and along the purchased-service margin, but enter neither technology. Productivity enters both technologies. For a non-buyer it therefore appears in output with weight $1 + \alpha_S$, which is why Section 4.3 imposes the productivity process on the common state ω_{it} rather than on the regime-dependent composite.

Equation (3) is a two-regime empirical approximation, not a claim that every buyer physically ceases all in-house service activity. The buyer branch treats any unobserved in-house component as part of the controlled latent state. If that component changes systematically with purchased services after conditioning on the state and the instruments, α_S should be interpreted as the conditional elasticity of industry-deflated revenue with respect to observed service purchases, rather than the elasticity with respect to the total service task. This interpretation is also the one used in the counterfactual calculation.

4.2 The Value-Added Tax, Credits, and Effective Prices

The credit wedge. A general taxpayer deducts input VAT from output VAT. A purchased service with a special invoice therefore generates a credit of $\tau_{bt} p_{S,ct}^b S_{it}^B$, where b denotes the

service category. Wages paid to the firm's own workers generate no input credit. When current output VAT is sufficient, the credit is used immediately and the statutory tax does not change the effective purchase price. When input VAT exceeds output VAT, the unused credit is carried forward, so delay or non-recovery gives the credit a positive shadow cost. These rules follow the Provisional VAT Regulations and Caishui No. 36 (State Council of the People's Republic of China, 2017; Ministry of Finance and State Administration of Taxation, 2016); the 2018 rate adjustment follows Caishui No. 32 (Ministry of Finance and State Administration of Taxation, 2018a). Caishui No. 70 introduced refunds of excess credits for firms in 18 eligible industries in 2018 (Ministry of Finance and State Administration of Taxation, 2018b).

Let $\lambda_{it} \in [0, 1]$ denote the present-value loss from one additional unit of input VAT. It equals zero when the credit is deducted immediately and is positive in a carryforward position. A simple representation is

$$\lambda_{it} = (1 - \psi_{it}) \left(1 - \mathbb{E}_t \delta^{T_i - t}\right)$$

for a firm with excess credits, where T_i is the date at which future output VAT absorbs the balance, δ is the discount factor, and ψ_{it} is the refunded fraction. We treat the opening credit position and the policy as predetermined at the quarterly purchase decision. The effective price of category b is then

$$\widetilde{p}^{b}_{S,ict} = p^{b}_{S,ct} \left(1 + \lambda_{it} \tau_{bt}\right), \qquad W_{ibt} \equiv \ln\left(1 + \lambda_{it} \tau_{bt}\right). \tag{4}$$

A firm with $\lambda_{it} = 0$ does not respond to the statutory rate. For a firm with excess credits, a refund lowers λ_{it}, and the resulting price change is larger for a category with a higher τ_{bt}. This cross-category contrast motivates the triple-difference design in Section 5.

Local supply and the sourcing decision. Producer services often require frequent interaction or a close match between the supplier and the buyer. We use LS_{ct}, the standardized seller-side service location quotient in city c, as a measure of market thickness. Rather than specifying supplier entry and matching in full, we summarize their price implication by

$$\ln p^{b}_{S,ct} = \delta_{bt} - \eta_b LS_{ct} + u^{p}_{bct}, \qquad \eta_b > 0. \tag{5}$$

This reduced-form relation captures both a lower price index and better matches in a thicker local market, consistent with evidence on local service outsourcing and supplier proximity (Ono, 2007; Bernard et al., 2019). Local supply may also reduce the fixed cost of finding and contracting with a supplier, as in models of specialized input matching and relational

contracts (Grossman and Helpman, 2002; Baker et al., 2002).

Combining (4) and (5) yields

$$\ln \widetilde{p}^{b}_{S,ict} = \delta_{bt} - \eta_b LS_{ct} + W_{ibt} + u^{p}_{bct}.$$

We therefore write the purchase decision and conditional service demand as

$$D_{it} = \mathbf{1}\left\{h(x_{it}, \omega_{it}) + \gamma_{LS} LS_{ct} - \gamma_W W_{it} + \nu_{it} \geq 0\right\}, \qquad \gamma_{LS}, \gamma_W > 0, \tag{6}$$

$$s_{it} = g(x_{it}, \omega_{it}, LS_{ct}, W_{it}), \qquad \frac{\partial g}{\partial LS} > 0, \quad \frac{\partial g}{\partial W} < 0, \qquad D_{it} = 1. \tag{7}$$

The function $h(\cdot)$ collects the relative value of buying rather than making and the fixed contracting cost. Equations (6) and (7) impose only the price and market-thickness monotonicities used in estimation. They also make the selection problem explicit: D_{it} and s_{it} depend on the same productivity state that enters Equation (3).

Firm Responses. The reduced-form policy estimate is converted into a service-purchase response without solving the firm's full optimization problem. Let $\widehat{\zeta}$ denote the estimated semi-elasticity of purchases with respect to a one-percentage-point reduction in the effective tax wedge per unit of firm exposure. For a policy exposure e_i, the category-level direct response is $\widehat{\zeta}(100\tau_b e_i)$. Aggregating the six service categories gives

$$\Delta \ln S^{CF}_i = \ln\left[\frac{\sum_b S_{ib} \exp\{\widehat{\zeta}(100\tau_b e_i)\}}{\sum_b S_{ib}}\right] + \Delta \ln S^{LS}_i, \tag{8}$$

where $\Delta \ln S^{LS}_i$ is the calibrated local-supply feedback. For the credit-refund benchmark, $e_i = \text{Treat70}_i \times \text{CreditExpo}_i$. For the service-VAT exemption, all six category rates are set to zero and $e_i = \text{CreditExpo}_i$, so a firm without stranded credits has no effective-price response. Applying the cross-category policy estimate to each category in this way requires the common-semi-elasticity assumption stated in Section 5.3. Holding other inputs, the sourcing regime, and general-equilibrium prices fixed, the production function maps purchases into output according to

$$\Delta \ln Q^{CF}_i \simeq \alpha_S \Delta \ln S^{CF}_i, \qquad \Delta Q^{CF}_i \simeq Q_i \alpha_S \Delta \ln S^{CF}_i. \tag{9}$$

These equations are the partial-equilibrium accounting implemented in Section 5.3.

4.3 Estimation

The purchase indicator and purchased quantity are endogenous in Equation (3). Estimating the buyer branch alone would select firms on ω_{it}, while coding non-purchases as an ordinary zero in $\ln(1+S)$ would force the extensive and intensive margins into the same regressor. We instead retain both regimes and estimate the first-stage control function

$$q_{it} = D_{it}\phi_1(l_{it}, k_{it}, m_{it}, s_{it}, LS_{ct}, W_{it}) + (1 - D_{it})\phi_0(l_{it}, k_{it}, m_{it}, LS_{ct}, W_{it}) + \varepsilon_{it}. \tag{10}$$

For buyers, service demand is inverted conditional on LS_{ct} and W_{it}. For non-buyers, who have no observed s_{it}, productivity is recovered from a regime-specific materials-demand inversion. Because this is a gross-output production function in which materials and purchased services are both flexible inputs, monotonicity of one flexible-input demand equation is not sufficient by itself. The additional demand shifters (LS_{ct}, W_{it}) provide the variation required by the maintained identification argument (Gandhi et al., 2020).

Excluded variation. The tax instrument aggregates category-level wedges with firms' fixed 2017 service-expenditure shares (ϖ_{ib}):

$$W_{it}^{IV} = \sum_b \varpi_{ib} \ln\left(1 + \widehat{\lambda}_{it}\tau_{bt}\right), \qquad \widehat{\lambda}_{it} = \text{CreditExpo}_i \left(1 - \psi\, \text{Treat70}_i \text{Post70}_t\right). \tag{11}$$

The instrument combines pre-reform exposure with statutory category rates, the 2018 rate adjustment, and the Document No. 70 refund. The baseline instrument vector also uses interactions of LS_{ct} with predetermined firm size and labor intensity, together with the high-rate by post-policy by exposure interaction. The exclusion restriction is that, conditional on the fixed effects and predetermined controls, these shifters affect purchases but not the innovation in firm productivity.

Holding the 2017 shares fixed makes them predetermined, but does not by itself make the resulting shift-share instrument exogenous (Goldsmith-Pinkham et al., 2020). Likewise, local supplier concentration may respond to local manufacturing conditions. The baseline therefore absorbs industry-by-quarter shocks, and specifications with city-by-quarter effects provide a sensitivity check. The identifying variation should be interpreted conditionally on these controls rather than as an unconditional consequence of using statutory rates or supplier counts.

Joint second stage. Given the first-stage estimate $\widehat{\phi}_{it}$, productivity for a candidate param-

eter vector Θ is

$$\omega_{it}(\Theta) = \frac{\widehat{\phi}_{it} - \sum_j \alpha_j x_{j,it} - D_{it}\alpha_S s_{it} - (1 - D_{it})\left(C_0 + \sum_j \theta_j x_{j,it}\right)}{1 + (1 - D_{it})\alpha_S}. \tag{12}$$

We impose the productivity process on this common state rather than on the regime-dependent composite term:

$$\omega_{it} = a_{g(i)t} + \rho\omega_{i,t-1} + \beta_{SR} SR_{it} + \xi_{it}, \qquad \mathbb{E}[Z_{it}\xi_{it}] = 0, \tag{13}$$

where $a_{g(i)t}$ denotes industry-by-quarter effects, SR_{it} is the R&D-intensity measure, and Z_{it} contains the excluded variables and predetermined inputs. The parameters θ_j and α_S are estimated jointly, after which $\beta_j = \theta_j/\alpha_S$. Since this division makes the β_j directions sensitive to the precision of α_S, we use firm-cluster block-bootstrap standard errors for the two-regime estimates.

The model illustrates that thicker local supply raises the probability and quantity of service purchases, the credit wedge lowers both margins for firms with stranded credits, and a refund produces a larger effective-price change in higher-rate categories. The production-function estimates then translate the observed input response into the partial-equilibrium output effects in Equation (9).

Table 4 reports the production-function estimates. Column (1) is a conventional single-regime control function that pools buyers and non-buyers and enters services as $\ln(1 + S_{it})$. Column (2) estimates the two regimes jointly, places industry-by-quarter intercepts in the process for ω. Column (3) reports the estimation results for firms buying the services in the continuous two quarters and gives a verification for the estimates in panel A of column (2).

The service coefficients in the columns (1) and (2) answer different questions. In column (1), $\alpha_S = 0.0099$ combines zero and positive purchases through $\ln(1 + S)$, forcing the extensive and intensive margins into one regressor. In column (2), $\ln S$ enters only in the buying regime, while D_{it} and the in-house technology account separately for the make-or-buy margin. The estimate $\alpha_S = 0.0374$, with a bootstrap standard error of 0.0002, is therefore the elasticity of industry-deflated revenue with respect to purchased services conditional on buying. The numerical gap between 0.0374 and 0.0099 illustrates that the specifications use different transformations and estimate different objects. Column (3) asks how much of that gap survives once the control function is estimated only on buyers and services enter as $\ln S_{it}$. The sample requires a purchase in both the current and the previous quarter, so that $\ln S_{it}$ and its lag are defined. The service elasticity is 0.0320, with a standard error of 0.0002,

which lies close to Panel A of column (2). The buyer-only control function keeps a common intercept in the process for ω, and its moments are computed on buyers alone. It recovers the purchased-service elasticity.

The remaining coefficients describe how observed inputs map into the service task when the firm does not buy. The final-good elasticities in column (2) are 0.2152 for labor, 0.7899 for materials, and 0.0205 for capital. The in-house mapping has lower elasticities with respect to labor and materials, $\beta_l = 0.0735$ and $\beta_m = 0.6211$, but a higher elasticity with respect to capital, $\beta_k = 0.1553$, than the corresponding final-good coefficients. Including purchased services, final-good returns to scale are increasing to scales (1.0629), while the in-house producer services production elasticities sum to 0.8498, indicating decreasing to scales. The capital elasticity of final-good production in column (2), Panel A (0.0205), is smaller than the capital elasticity in column (1) (0.1375), but in in-house producer-service production in column (2), Panel B, the capital elasticity is high (0.1553). This indicates that the technology for producer-service production is relatively more capital-intensive than final-good production, and that outsourcing producer services can effectively reduce a firm's own capital input.

Table 4: Production Function Estimates: Control Function and Two-Regime

	(1) Pooled Estimation	(2) Two-regime Estimation	(3) Pooled Estimation 1=(Buyers only)
Panel A: Final-good input elasticities			
α_L (labor)	0.1239***	0.2152***	0.1195***
	(0.0140)	(0.0141)	(0.0197)
α_M (materials)	0.8095***	0.7899***	0.7852***
	(0.0009)	(0.0010)	(0.0014)
α_K (capital)	0.1375***	0.0205***	0.0936***
	(0.0060)	(0.0072)	(0.0105)
α_S (services)	0.0099***	0.0374***	0.0320***
	(0.0001)	(0.0002)	(0.0002)
Returns to scale	1.0808	1.0629	1.0303
	(0.0115)	(0.0083)	(0.0158)
Panel B: In-house service technology			
β_l (labor)	—	0.0735**	—
		(0.0286)	
β_m (materials)	—	0.6211***	—
		(0.0144)	
β_k (capital)	—	0.1553***	—
		(0.0280)	
$\beta_l + \beta_m + \beta_k$	—	0.8498	—
		(0.0283)	
C_0 (allocation constant)	—	−0.0778***	—
		(0.0078)	
Panel C: Productivity process			
β^{ω}_{SR} (R&D service share)	2.9076***	1.7144***	2.1816***
	(0.0663)	(0.0601)	(0.0692)
ρ (persistence)	0.9901***	0.9719***	0.9896***
	(0.0001)	(0.0004)	(0.0001)
a_0	0.0146***	absorbed	0.0229***
	(0.0001)		(0.0002)
Service input measure	$\ln(1 + S_{it})$	$\ln S_{it}$, $D_{it} = \mathbf{1}(S_{it} > 0)$	$\ln S_{it}$
Sample	Pooled	Pooled	$S_{it} > 0$, $S_{i,t-1} > 0$
Observations	2,874,030	2,874,030	1,842,980
Firms	738,068	738,068	532,817

Note 1: Column (1) is the single-regime control function that pools firms that buy services with firms that do not. Local supply enters only the first-stage polynomial. The instruments are lagged labor, lagged materials, current capital, lagged services, and a constant. Its standard errors are numerical-Jacobian sandwich errors clustered at the firm. Column (2) is the baseline two-regime model Column (3) uses the column (1) estimator and the column (2) labor measure, replaces $\ln(1+S_{it})$ with $\ln S_{it}$, and keeps firm-quarters that purchase services in both t and $t-1$.

Note 2: Standard errors are firm-cluster block-bootstrap standard errors.

Note 3: Columns (1), (2) and (3) do not estimate the same α_S. In column (1) the service input enters as $\ln(1 + S_{it})$ over the pooled sample, so α_S is the coefficient on a transformation that combines zero and positive purchases; it is not an elasticity with respect to S at zero. In columns (2) and (3) the service input enters as $\ln S_{it}$ and is interacted with the outsourcing indicator D_{it}, so α_S is the elasticity of industry-deflated revenue with respect to purchased services conditional on the firm buying.

* $p < .10$, ** $p < .05$, *** $p < .01$.

5 Value-added Tax Reform and Service Input

This section uses the 2018 credit refund policy as an external shock to estimate how the effective VAT wedge changes service purchases and firm outcomes.

5.1 The Value-added Tax Reform

Document No. 70 reduced the shadow cost of excess input credits for eligible industries (Ministry of Finance and State Administration of Taxation, 2018b). The model predicts that this relief should have a larger effect on high-rate services because the effective-price wedge $\ln(1+\lambda_{it}\tau_{bt})$ is increasing in the statutory rate. We let H_b indicate a high-rate categories (production services and transport services) and estimate this cross-category response at the firm level:

$$\begin{aligned} y_{ibt} = \beta_{\text{DDD}}\, H_b \times \text{Post}_t \times \text{FirmExpo}_i \\ + \delta\big(H_b \times \text{FirmExpo}_i \times t\big) + \mu_{it} + \nu_{bt} + \kappa_{ib} + \varepsilon_{ibt}. \end{aligned} \tag{14}$$

where b indexes service categories, $\text{FirmExpo}_i = \text{Treat70}_i \times \text{CreditExpo}_i$, and CreditExpo_i is measured before the reform. The empirical post period is 2018Q4. The fixed effects compare categories within the same firm and quarter while absorbing common category-specific time shocks and persistent firm-category differences, and the additional interaction allows a differential linear trend. The coefficient β_{DDD} is positive if high-rate purchases rise differentially after the refund among firms with greater exposure.

Table 5 reports the triple-difference estimates, decomposing the policy response into intensive and extensive margins. In Column (1), the combined-margins specification using $\ln(1+S_{ibt})$ yields a statistically significant coefficient of 0.0269 ($p<0.01$), indicating that manufacturers with higher pre-reform credit exposure differentially expanded their purchases of high-tax-rate services following the refund. Decomposing this aggregate response reveals that both margins played an active role. In Column (2), conditional on positive spending, high-rate service purchases among incumbent buyers grew by 2.38 percent per unit of exposure ($\beta=0.0238$, $p<0.01$), demonstrating a pronounced intensive-margin expansion. At the same time, the refund prompted new outsourcing decisions: the linear probability model in Column (3) shows a 0.90 percentage point increase in the likelihood of procuring high-rate services ($\beta=0.0090$, $p<0.01$). This extensive-margin entry is further corroborated by the conditional logit specification in Column (4) ($\beta=0.0421$, $p<0.01$). Taken together, these findings demonstrate that the reduction in the effective tax wedge operated along two distinct dimensions: it induced non-purchasing firms to cross the make-or-buy threshold into the market, while simultaneously scaling up service demand among existing buyers.

Table 5: VAT Credit Refund and Purchases of High-Rate Services

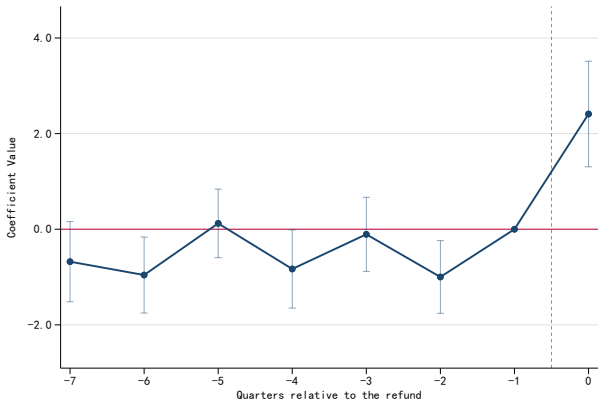
	(1) Combined Margins	(2) Intensive Margin	(3) Extensive Margin	(4) Extensive Margin
	$\ln(1+S_{ibt})$	$\ln S_{ibt}$	OLS	Logit
$\text{HighRate}_b \times \text{Post}_t \times \text{FirmExpo}_i$	0.0269***	0.0238***	0.0090***	0.0421***
	(0.0053)	(0.0080)	(0.0011)	(0.0083)
Firm × Quarter FE	YES	YES	YES	YES
Service Category × Quarter FE	YES	YES	YES	YES
Firm × Service Category FE	YES	YES	YES	NO
Observations	15,728,250	2,726,542	15,728,250	10,336,045
Adjusted R^2	0.706	0.603	0.599	
Pseudo R^2				0.505

Note: $\text{FirmExpo}_i = \text{Treat70}_i \times \text{CreditExpo}_i$, where CreditExpo_i is measured before the reform. All columns control for $\text{HighRate}_b \times \text{FirmExpo}_i$ interacted with a linear quarterly trend. Standard errors clustered at the industry level are in parentheses.

Column (4) is a conditional logit that conditions out firm-by-quarter effects and includes service-category-by-quarter effects. Firm-by-service-category effects cannot be included simultaneously in this estimator. Groups without within-group variation in purchasing status do not contribute to the conditional likelihood, which accounts for the smaller sample. The reported logit coefficient establishes direction and statistical significance but is not a probability marginal effect.

* $p < .10$, ** $p < .05$, *** $p < .01$

Figure 7: Event Study of the Credit Refund on Service Purchases



Note: The figure reports the event-study analogue of Equation (14), using the fixed effects, differential linear trend, and industry clustering in Table 5. The omitted quarter is 2018Q3. Bars are 95% confidence intervals, and units are log points × 100. The two margins are shown separately in Figure 8.

To validate the identifying assumption and trace the dynamic response, Figure 7 plots the event-study estimates corresponding to Equation (14), with 2018Q3 serving as the normalized baseline quarter. Two key features emerge from the dynamic pattern. First, the estimated coefficients in the pre-reform periods (from quarter −7 to −2) are statistically in-

distinguishable from zero and show basically no obvious pre-trend, confirming that exposed and unexposed firms were on parallel paths prior to the refund and ruling out obvious anticipation effects. Second, service purchases exhibit a sharp and immediate divergence precisely in the implementation quarter (2018Q4, quarter 0), jumping by over 2 percent for firms with higher credit exposure. Figure 8 in the appendix decomposes this dynamic trajectory into separate margins. Both the extensive probability of entering the service market and the intensive spending among incumbent buyers jump concurrently in 2018Q4, confirming that the policy reform triggered an immediate, synchronized adjustment across both margins.

5.2 Policy Impact

Direct effects on service users. The observed purchasing expansion should translate into tangible performance improvements if the VAT credit wedge previously distorted an economically productive input. Table 6 examines downstream manufacturer outcomes by regressing firm revenue and profitability on the interaction between pre-reform credit exposure and the post-refund period, absorbing firm and quarter fixed effects.

Table 6: Post-Refund Firm Outcomes by Pre-Reform Credit Exposure

	(1)	(2)	(3)	(4)
	ln Users' Revenue		Users' Profit Rate	
$\text{FirmExpo}_i \times \text{Post}_t$	0.0332***	0.0325***	0.0523***	0.0523***
	(0.0065)	(0.0067)	(0.0061)	(0.0061)
Capital Intensity $(\ln(K/L)_{it})$		YES		YES
Capital Stock $(\ln K_{it})$		YES		YES
Firm FE	YES	YES	YES	YES
Quarter FE	YES	YES	YES	YES
Observations	2,351,290	2,351,290	2,351,290	2,351,290
Adjusted R^2	0.899	0.901	0.254	0.254

Note: $\text{FirmExpo}_i = \text{Treat70}_i \times \text{CreditExpo}_i$, where credit exposure is measured before the reform, and Post_t equals one in 2018Q4. Standard errors clustered at the industry level are in parentheses.
* $p < .10$, ** $p < .05$, *** $p < .01$

The estimates confirm that easing the input tax friction yielded meaningful economic gains for exposed manufacturers. In Column (1), the baseline revenue elasticity is 0.0332 ($p < 0.01$). Controlling for contemporaneous capital intensity and capital stock in Column (2) leaves this estimate virtually unchanged at 0.0325 ($p < 0.01$), indicating an approximate 3.25 percent post-reform revenue expansion per unit of exposure. Similarly, Columns (3) and (4) show a robust and stable improvement in operating profitability: exposed firms experienced a 5.23 percentage point increase in their profit rate ($\beta = 0.0523$, $p < 0.01$), irrespective of internal capital adjustments. That these gains survive controls for internal

factor stocks indicates that the performance boost does not simply reflect parallel physical capital deepening; rather, it stems from the productivity and cost advantages of outsourcing specialized service tasks.

Indirect effects through local suppliers. Because producer services typically feature localized market interactions and contracting frictions, a buyer-side reduction in the effective service price should propagate upstream to stimulate the local supplier ecosystem. To capture this market-level spillover, Table 7 aggregates policy exposure to the prefecture-level city and assesses the post-refund growth in local service provider activity, absorbing city and quarter fixed effects.

Table 7: Post-Refund Local Supplier Outcomes by City Exposure

	(1)	(2)	(3)	(4)
	ln Suppliers' Sales		ln Sales per Supplier	
$\text{CityExpo}_c \times \text{Post}_t$	0.2990***	0.2995***	0.3258***	0.3258***
	(0.0732)	(0.0743)	(0.0817)	(0.0820)
Local Manufacturing Firm Number		YES		YES
Local Manufacturing Revenue		YES		YES
Local Manufacturing Employment		YES		YES
City FE	YES	YES	YES	YES
Quarter FE	YES	YES	YES	YES
Observations	2,672	2,672	2,672	2,672
Adjusted R^2	0.982	0.982	0.837	0.838

Note: CityExpo_c is the 2017 revenue-weighted mean of firm exposure among manufacturing buyers in city c, and Post_t equals one in 2018Q4. Standard errors clustered at the prefecture-level city are in parentheses.
* $p < .10$, ** $p < .05$, *** $p < .01$

The results document substantial upstream demand spillovers. In Column (1), aggregate service supplier sales rise by 29.90 percent per unit of city exposure ($\beta = 0.2990$, $p < 0.01$). Crucially, conditioning on local manufacturing firm count, total revenue, and aggregate employment in Column (2) yields an almost identical estimate of 0.2995 ($p < 0.01$), confirming that this upstream surge is not an artifact of broad local industrial booms, but is specifically driven by manufacturing demand for service inputs. Columns (3) and (4) evaluate provider-level scale, reporting a 32.58 percent increase in sales per supplier ($\beta = 0.3258$, $p < 0.01$). The fact that sales per supplier expand at a comparable magnitude to total market sales reveals that the policy primarily scaled up incumbent service providers' operational capacity, rather than inducing a purely extensive-margin influx of small or newly registered entities. These market-level responses confirm the downstream-to-upstream transmission of tax relief and provide the empirical foundation for our counterfactual calibration of local-supply feedback.

5.3 Counterfactual Analysis

We consider two policy exercises. The first reproduces the observed credit-refund policy for eligible industries at each firm's measured exposure. The second sets the VAT rate on purchased services to zero for all firms, while allowing the effective price to fall only in proportion to each firm's own stranded-credit exposure. A firm without stranded credits therefore has no effective-price response in either exercise. Both calculations incorporate the calibrated local-supply adjustment and map service purchases into output using the service coefficient, $\alpha_S = 0.03736$, from column (2) of Table 4.

Table 8: Illustrative Partial-Equilibrium Policy Accounting

	(1) Credit-refund policy benchmark	(2) Service-VAT exemption
Policy coverage	Eligible industries	All firms
Effective exposure e_i	$\text{Treat70}_i \times \text{CreditExpo}_i$	CreditExpo_i
Service purchases: direct response	+1.46%	+2.08%
including local-supply feedback	+2.20%	+2.83%
Output gain (billion RMB)	**9.31**	**11.72**

Note 1: Column (1) applies the credit refund to listed industries at each firm's observed ratio of stranded input credits to revenue. Column (2) sets the VAT rate on all six purchased-service categories to zero and uses each firm's own observed credit exposure. A firm with no stranded credit has no effective-price response in either exercise.
Note 2: The purchase response is calibrated from the aggregate estimate 0.0269 in column (1) of Table 5. Dividing by the 8.92-percentage-point tax-rate difference and by the standard deviation of the underlying credit-exposure ratio gives a semi-elasticity of 0.07858 per percentage point and per unit of exposure. The current six-category service definition is used; "other services" are assigned the 6% rate. The local-supply response from Table 7 adds 0.00727 log point to service demand.
Note 3: Output gains use $\alpha_S = 0.03736$ from the two-regime model in column (2) of Table 4.

Table 8 summarizes the simulation results, highlighting two central economic takeaways. First, the actual 2018 refund policy produced substantial output gains by loosening input constraints in the treated industries. Direct price relief increased service purchases by 1.46 percent, while the induced deepening of the local supplier market provided an additional 0.74 percentage-point boost. Accounting for this local-supply feedback, aggregate service purchases expanded by 2.20 percent, generating an aggregate manufacturing output gain of 9.31 billion RMB (equivalent to $1.41 billion). Second, the broader service-VAT exemption illustrates the scope of unaddressed tax distortions in the wider economy. By extending relief beyond the 18 selected industries to all capital- or credit-constrained manufacturers, the direct purchasing response reaches 2.08 percent, and total purchases rise by 2.83 percent with supplier feedback. This universal exemption yields a total output expansion of 11.72 billion RMB ($1.77 billion)–a nearly 26 percent increase in output gains over the actual 2018 refund, indicating the potential of future tax policy reforms.

6 Conclusion

This paper investigates how tax frictions shape manufacturing firms' make-or-buy decisions and distort organizational boundaries. By merging administrative VAT invoices with corporate income tax records, we observe firm-to-firm service transactions alongside internal factor use across Chinese manufacturing. These data highlight an essential extensive margin: firms that do not purchase services often produce the corresponding tasks in house. Our two-regime production-function estimator incorporates this make-or-buy choice into the control-function framework, resolving the downward bias of standard pooled specifications. While a conventional pooled control function yields a service elasticity of 0.0099, our two-regime model yields a conditional elasticity of 0.0374. The estimation results conditional on firms buying services confirm this estimate. The structural estimates also uncover the parameters of in-house service provision. In-house service production is more capital-intensive than final-good production, exhibiting a capital elasticity of $\beta_k = 0.1553$ relative to $\alpha_K = 0.0205$. The estimates also indicate decreasing returns to scale in in-house service tasks alongside modest increasing returns of 1.0629 in final-good output.

The 2018 refund of excess VAT credits provides quasi-experimental variation that corroborates the model's effective-price mechanism. By lowering the shadow cost of stranded credits, the reform generated disproportionate purchasing expansions in high-rate service categories such as production and transport services. In our triple-difference specification, the combined elasticity across both margins is 0.0269. This response operates through both intensive spending growth, with a coefficient of 0.0238, and a 0.0090 increase in purchasing probability on the extensive margin. The policy also generated positive economic consequences across the production network. Exposed manufacturing buyers experienced post-reform revenue increases of 3.25 percent and a 5.23 percentage point improvement in profit rates. Furthermore, this buyer-side tax relief propagated to upstream markets, raising total local supplier sales by 29.95 percent and sales per supplier by 32.58 percent.

Counterfactual policy simulations quantify the aggregate output gains from reducing tax frictions on service outsourcing. Accounting for local-supply feedback, the observed 2018 refund policy expanded aggregate service purchases by 2.20 percent, generating $1.41 billion in output gains. A broader policy that fully exempts purchased services from VAT scaled to each firm's stranded credits increases service purchases by 2.83 percent, delivering $1.77 billion in output gains. These results demonstrate that input-credit usability directly governs firm specialization and intermediate input choices. Delayed refunds impose real shadow costs that alter the boundary of the firm. Separating make from buy is therefore essential for correctly evaluating production technology and designing neutral tax policy in modern

production networks.

References

Abraham, K. G. and Taylor, S. K. (1996). Firms' use of outside contractors: Theory and evidence. *Journal of Labor Economics*, 14(3):394–424.

Ackerberg, D. A., Caves, K., and Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6):2411–2451.

Alfaro, L., Antràs, P., Chor, D., and Conconi, P. (2019). Internalizing global value chains: A firm-level analysis. *Journal of Political Economy*, 127(2):508–559.

Amiti, M. and Konings, J. (2007). Trade liberalization, intermediate inputs, and productivity: Evidence from indonesia. *American Economic Review*, 97(5):1611–1638.

Antràs, P. (2003). Firms, contracts, and trade structure. *Quarterly Journal of Economics*, 118(4):1375–1418.

Antràs, P. and Helpman, E. (2004). Global sourcing. *Journal of Political Economy*, 112(3):552–580.

Bachas, P., Gadenne, L., and Jensen, A. (2024). Informality, consumption taxes, and redistribution. *Review of Economic Studies*, 91(5):2604–2634.

Baker, G., Gibbons, R., and Murphy, K. J. (2002). Relational contracts and the theory of the firm. *Quarterly Journal of Economics*, 117(1):39–84.

Benzarti, Y. and Carloni, D. (2019). Who really benefits from consumption tax cuts? evidence from a large VAT reform in france. *American Economic Journal: Economic Policy*, 11(1):38–63.

Berlingieri, G. and Pisch, F. (2026). Managing export complexity: The role of service outsourcing. *Review of Economics and Statistics*, pages 1–45. Advance online publication.

Bernard, A. B., Moxnes, A., and Saito, Y. U. (2019). Production networks, geography, and firm performance. *Journal of Political Economy*, 127(2):639–688.

Blundell, R. and Bond, S. (2000). GMM estimation with persistent panel data: An application to production functions. *Econometric Reviews*, 19(3):321–340.

Brandt, L., Van Biesebroeck, J., Wang, L., and Zhang, Y. (2017). WTO accession and performance of Chinese manufacturing firms. *American Economic Review*, 107(9):2784–2820.

Carrillo, P., Pomeranz, D., and Singhal, M. (2017). Dodging the taxman: Firm misreporting and limits to tax enforcement. *American Economic Journal: Applied Economics*, 9(2):144–164.

Chandra, P. and Long, C. (2013). VAT rebates and export performance in China: Firm-level evidence. *Journal of Public Economics*, 102:13–22.

Chen, S., Xiao, Y., and Zhang, Z. (2024). Value-added tax credit refund and environmental protection investment: Evidence from chinese heavy-polluting enterprises. *Journal of Cleaner Production*, 450:141945.

Chen, Z., Liu, Z., Suárez Serrato, J. C., and Xu, D. Y. (2021). Notching R&D investment with corporate income tax cuts in China. *American Economic Review*, 111(7):2065–2100.

Collard-Wexler, A. and De Loecker, J. (2015). Reallocation and technology: Evidence from the US steel industry. *American Economic Review*, 105(1):131–171.

De Loecker, J. (2011). Product differentiation, multi-product firms, and estimating the impact of trade liberalization on productivity. *Econometrica*, 79(5):1407–1451.

De Loecker, J., Eeckhout, J., and Unger, G. (2020). The rise of market power and the macroeconomic implications. *Quarterly Journal of Economics*, 135(2):561–644.

De Loecker, J., Goldberg, P. K., Khandelwal, A. K., and Pavcnik, N. (2016). Prices, markups, and trade reform. *Econometrica*, 84(2):445–510.

De Loecker, J. and Warzynski, F. (2012). Markups and firm-level export status. *American Economic Review*, 102(6):2437–2471.

Doraszelski, U. and Jaumandreu, J. (2013). R&D and productivity: Estimating endogenous productivity. *Review of Economic Studies*, 80(4):1338–1383.

Fort, T. C. (2017). Technology and production fragmentation: Domestic versus foreign sourcing. *The Review of Economic Studies*, 84(2):650–687.

Foster, L., Haltiwanger, J., and Syverson, C. (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review*, 98(1):394–425.

Gandhi, A., Navarro, S., and Rivers, D. A. (2020). On the identification of gross output production functions. *Journal of Political Economy*, 128(8):2973–3016.

Goldschmidt, D. and Schmieder, J. F. (2017). The rise of domestic outsourcing and the evolution of the german wage structure. *Quarterly Journal of Economics*, 132(3):1165–1217.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Grossman, G. M. and Helpman, E. (2002). Integration versus outsourcing in industry equilibrium. *Quarterly Journal of Economics*, 117(1):85–120.

Grossman, G. M. and Rossi-Hansberg, E. (2008). Trading tasks: A simple theory of offshoring. *American Economic Review*, 98(5):1978–1997.

Grossman, S. J. and Hart, O. D. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy*, 94(4):691–719.

Guo, Y. M. and Li, X. (2023). The impact of greater VAT tax neutrality on total factor productivity: Evidence from china's VAT credit refund reform in 2018. *Economic Analysis and Policy*, 78:922–936.

Halpern, L., Koren, M., and Szeidl, A. (2015). Imported inputs and productivity. *American Economic Review*, 105(12):3660–3703.

Hart, O. and Moore, J. (1990). Property rights and the nature of the firm. *Journal of Political Economy*, 98(6):1119–1158.

Huang, Y., Sun, Z., and Huang, Y. (2025). Exploring the effect of VAT carryforward and refund policy on enhancing innovation efficiency among china's A-share listed firms. *Finance Research Letters*, 72:106568.

Keen, M. and Lockwood, B. (2010). The value added tax: Its causes and consequences. *Journal of Development Economics*, 92(2):138–151.

Klette, T. J. and Griliches, Z. (1996). The inconsistency of common scale estimators when output prices are unobserved and endogenous. *Journal of Applied Econometrics*, 11(4):343–361.

Levinsohn, J. and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, 70(2):317–341.

Li, M., Liu, C., and Zhou, S. (2026a). Tax incentives and corporate intelligent transformation: Evidence from china's VAT credit refund policy. *Journal of Asian Economics*, 102:102100.

Li, M., Qiu, X., and Yan, X. (2026b). Two-sided heterogeneity in firm-to-firm production networks. Working paper.

Liao, C. and Li, H. (2026). Value-added tax neutrality and enterprise green strategy: Evidence from china's credit refund reform. *Finance Research Letters*, 89:109169.

Liu, L., Cao, L., Cao, Y., Lu, M., and Shan, Y. (2024). VAT credit refunds and firm productivity: Evidence from china's VAT reform. *International Review of Financial Analysis*, 93:103171.

Liu, Y. and Mao, J. (2019). How do tax incentives affect investment and productivity? firm-level evidence from China. *American Economic Journal: Economic Policy*, 11(3):261–291.

Ministry of Finance and State Administration of Taxation (2016). Notice on comprehensively launching the pilot program of replacing business tax with value-added tax. https://www.gov.cn/gongbao/content/2016/content_5088782.htm. Caishui [2016] No. 36.

Ministry of Finance and State Administration of Taxation (2018a). Notice on adjusting value-added tax rates. https://www.gov.cn/zhengce/zhengceku/2018-12/31/content_5439278.htm. Caishui [2018] No. 32.

Ministry of Finance and State Administration of Taxation (2018b). Notice on tax policies for refunding excess input value-added tax credits in selected industries in 2018. https://www.chinatax.gov.cn/chinatax/n810341/n810765/n3359382/201807/c3730844/content.html. Caishui [2018] No. 70.

Naritomi, J. (2019). Consumers as tax auditors. *American Economic Review*, 109(9):3031–3072.

Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):1263–1297.

Ono, Y. (2007). Market thickness and outsourcing services. *Regional Science and Urban Economics*, 37(2):220–238.

Pomeranz, D. (2015). No taxation without information: Deterrence and self-enforcement in the value added tax. *American Economic Review*, 105(8):2539–2569.

State Council of the People's Republic of China (2017). Provisional regulations of the people's republic of china on value-added tax (2017 revision). https://www.gov.cn/gongbao/content/2017/content_5248215.htm. State Council Order No. 691.

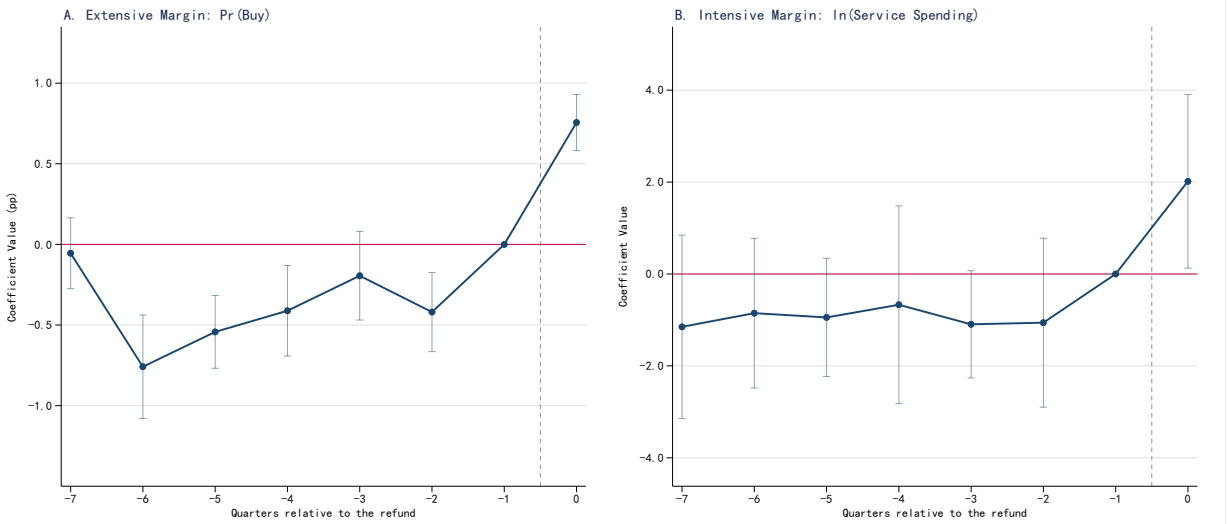
Syverson, C. (2004). Market structure and productivity: A concrete example. *Journal of Political Economy*, 112(6):1181–1222.

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, 104(3):112–114.

Yang, G.-Z. and Si, D.-K. (2025). The impact of VAT input tax refund policy on firms' labor income share: Evidence from china. *Economic Analysis and Policy*, 86:2233–2246.

A Appendix A

Figure 8: Event Study by Margin: Extensive and Intensive



Note: The figure decomposes Figure 7 into the two margins reported in columns (3) and (2) of Table 5. Panel A takes as dependent variable an indicator for positive spending, estimated by OLS, and is measured in percentage points. Panel B takes log spending among cells with positive spending and is measured in log points $\times$ 100; it therefore conditions on a decision the reform itself moves, and the entrants it excludes are on average smaller buyers, so Panel B is best read as a net figure. Quarter -1 is 2018Q3, normalised to zero; quarter 0 is 2018Q4. Fixed effects, clustering, and the detrending are as in Figure 7. Bars are 95% confidence intervals.