Two-Sided Heterogeneity in Firm-to-Firm Production Networks*

Ming Li† CUHK Shenzhen

Xincheng Qiu‡ Peking University

Xinan Yan§ Peking University

March 26, 2026

(click here for latest version)

Abstract

Production networks make firm performance an aggregated outcome of bilateral relationships: who firms trade with and how much is transacted. This paper quantifies two-sided heterogeneity in firm-to-firm transactions using the universe of value-added tax invoices in China from 2014–2018. We adapt the two-way fixed-effects framework to production networks and decompose log transaction volume into supplier effects, buyer effects, and match-specific residuals. We find that supplier-side heterogeneity is the dominant source of transaction dispersion. Using the estimated components, we provide a novel characterization of firm-to-firm production networks. Finally, we document evidence of network misallocation among state-owned enterprises, who exhibit weaker supplier-quality sorting and preferential access to high-quality suppliers.

Keywords: Production networks, firm-to-firm transactions, two-way fixed effects, state-owned enterprises, misallocation

JEL Codes: D22, D57, L14, L23, P31

*We thank participants at Carnegie Mellon and Melbourne for helpful comments.

†Email: liming2020@cuhk.edu.cn. School of Management and Economics, Chinese University of Hong Kong (Shenzhen).

‡Email: xincheng.qiu@gsm.pku.edu.cn. Guanghua School of Management, Peking University.

§Email: yanxinan@stu.pku.edu.cn. Guanghua School of Management, Peking University.

1 Introduction

Modern economies are organized as production networks. Value added, by construction, is created through firm-to-firm input-output linkages: firms use the output of other firms as intermediate inputs, convert these intermediate inputs into their own product, and sell their own product that becomes inputs elsewhere. Empirical work often summarizes firm performance with a single number, such as productivity, value added, and sales, as if it were an intrinsic attribute of the firm. In a networked economy, such measures are aggregates of bilateral relationships: who a firm buys from, who it sells to, and what is transacted within each buyer–supplier match. This paper seeks to understand the importance of the two sides of these relationships in determining firm performance.

Using comprehensive administrative data of the universe of value-added tax (VAT) invoices in China from 2014–2018, we first document substantial heterogeneity in firm-to-firm transactions, even within buyer–supplier industry pairs. For instance, within industry pairs, the 90th-percentile firm-to-firm transaction is nearly 400 times larger than the 10th-percentile transaction. Why are some transactions larger than others? One possibility is that transaction volume reflects the supplier's ability to deliver value at scale: some firms operate with greater capacity and lower marginal costs, produce higher-quality or more customized inputs, and consistently meet delivery schedules and performance standards across partners. Another possibility is that transaction volume reflects the buyer's capacity to source and deploy inputs effectively: some firms are more productive and demand more intermediates, operate more efficient procurement and inventory systems, and build reputations for timely payment that make suppliers willing to allocate capacity. How much of a transaction is due to the supplier and how much the buyer? Distinguishing these two sides of forces is essential for understanding sources of firm size heterogeneity, how supply chains form and evolve, and frictions in buyer–supplier allocation that may generate misallocation and amplify shocks.

To answer this question, we adapt the canonical two-way fixed effects framework à la Abowd, Kramarz and Margolis (1999) to firm-to-firm production networks. Specifically, we model the log value of a transaction between a supplier and a buyer (demeaned within supplier-buyer industry-pairs) as the sum of a supplier fixed effect, a buyer fixed effect, and a residual. The supplier fixed effect summarizes how much a supplier sells to its average customer, capturing the supplier-side forces discussed above. The buyer fixed effect summarizes how much a buyer purchases from its average supplier, capturing the buyer-side forces. The residual represents match-specific components, such as relationship capital, specific customization, and other idiosyncratic factors. Because firms typically trade with many partners, the network is richly connected, which delivers the variation needed to separately identify supplier and buyer effects. This decomposition provides a parsimonious characterization of firm-to-firm production

networks that allows us to measure who are "good" suppliers and "good" buyers, quantify the extent to which good suppliers are matched with good buyers, and assess whether are good supplier also good buyers.

Our first set of results reveals a sharp asymmetry in the sources of transaction dispersion. The supplier component accounts for about 43–45% of the variation in firm-to-firm sales, whereas the buyer component accounts for only about 6–8%. This means implies a dominant role of suppliers in driving the heterogeneity in firm-to-firm transactions. That is, conditional on industry pairs, who you are as a supplier matters far more than who you are as a buyer. Moreover, the covariance term is economically negligible, and the correlation between supplier and buyer effects is approximately 0.01. We have conducted a several validation exercises to assess the validity of our specification. First, in a cross-validation test, out-of-sample predictions constructed from the estimated supplier and buyer effects in the training set closely track realized transaction values in the test set. The slope of realized on predicted values is almost one and the predictive R-squared is around 0.47. Second, we test for complementarity between supplier and buyer fixed effects and find little evidence for the quantitative importance of such complementarity. Together, these exercises support modeling log transaction values as an additive function of supplier and buyer components.

Next, we document several facts that further characterize firm-to-firm production networks and provide discipline for theories of network formation. First, we allow the supplier and buyer fixed effects to be year specific and find that the estimated fixed effects are highly persistent, with year-to-year correlations of about 0.9 for both supplier and buyer FEs. Second, the estimated fixed effects connect to other familiar measures of firm performance. Firms with higher FEs tend to be larger in terms of employment. Both supplier and buyer effects are strongly positively associated with firms' TFP, even after controlling for size. Third, good suppliers with high supplier effects also tend to be good customers with high buyer effects, pointing to some common underlying capability, such as managerial quality, technology, or reliability, that governs performance on both sides of the market. Finally, we find that good suppliers disproportionately source inputs from good suppliers.

We then document a set of descriptive patterns that highlight the distinction between the extensive margin of firm sales, i.e., the number of customers, and the intensive margin of firm sales captured by the estimated supplier fixed effects. On the buyer side, larger customers both have more suppliers and buy more from each supplier. On the supplier side, however, scale along the extensive margin is negatively related to the intensive margin. Suppliers with more customers tend to have lower supplier fixed effects and tend to match with worse customers. These facts point to a tension between breadth and depth in supplier relationships and caution that narratives based solely on the customer base can potentially miss important information

at the intensive margin.

Lastly, we use our estimated objects to uncover additional patterns by comparing state-owned enterprises (SOEs) with private firms. We find evidence consistent with market power and misallocation associated with SOEs. Conditional on customer effects, SOEs are matched with better suppliers than comparable private firms, as measured by supplier effects. Moreover, supplier sorting is substantially weaker among SOEs than among private firms, and it is also weaker in cities with greater SOE presence. These patterns suggest that SOE market power may reduce local market efficiency and create barriers to private firms.

This paper makes three contributions. First, it contributes to the analysis of firm heterogeneity. Following influential work by Hopenhayn (1992) and Melitz (2003), productivity-driven heterogeneity has become the workhorse in modern macroeconomics trade, and IO. Foster, Haltiwanger and Syverson (2008) emphasize that revenue-based outcomes also embed demand heterogeneity, and hence markups and price differences. A series work following Bloom and Van Reenen (2007) studies management practices and organizational capability that improve firm performance. A common feature in the approach in those studies is that outcomes such as sales or value added are treated as a firm's own attributes. We contribute to this agenda by showing that an important and separately measurable dimension of heterogeneity is transactional: firms differ systematically in how much they sell to an average customer and how much they buy from an average supplier, even after conditioning on industry pairs. The closest to ours is the recent work by Bernard, Dhyne, Magerman, Manova and Moxnes (2022) who also view firm outcomes as emerging from relationships in production networks.

Second, this paper provides new characterization of firm-to-firm production networks. A growing body of work documents the structure of buyer–supplier networks, such as degree distributions (Atalay et al., 2011), sparsity (Bernard and Zi, 2022), and geography (Bernard, Moxnes and Saito, 2019). We contribute by adding a new characterization based on who drives transaction dispersion. Using a two-way fixed-effect decomposition at the transaction level, we provide economy-wide evidence on the relative importance of supplier versus buyer components, the extent of assortative matching, and the persistence of supplier and buyer components over time. These moments are informative for quantitative theories of production networks.

Third, this identifies a new source of inefficiency of state-owned enterprises. A central theme in the Chinese economy literature is that state ownership and policy distortions affect productivity through factor misallocation (Hsieh and Klenow, 2009), preferential finance (Song, Storesletten and Zilibotti, 2011), and political pressure to hire excess labor (Berkowitz et al., 2017). Hsieh and Song (2015) find that SOE reforms contributed to a substantial portion of the aggregate productivity growth in China. We complement this perspective by identifying a distinct margin of inefficiency operating inside production networks: the allocation and quality of

buyer–supplier relationships. In particular, we show that the mapping between buyer-side capability and supplier quality is systematically weaker for SOEs, consistent with procurement and relationship formation being shaped by non-market objectives or institutional constraints. This "network misallocation" channel is conceptually different from capital and labor misallocation.

The remainder of the paper proceeds as follows. Section 2 describes the data sources and presents facts that motivate our analysis. Section 3 presents the estimation of the two-way fixed effects framework, reports the main variance decomposition results, and validates the log additively separable specification. Section 4 uses the estimated fixed effects to characterize the firm-to-firm production networks. Section 5 presents evidence of network misallocation among state-owned enterprises. Section 6 concludes.

2 Data and Motivating Facts

2.1 Data

We combine four data sets on Chinese firms for the empirical analysis.

VAT Invoices Central to this paper is the establishment-level dataset from the State Taxation Administration (STA) that records all value-added tax (VAT) invoices issued by firms in mainland China from 2014 through 2018. It encompasses roughly 16.1 billion transactions among 16 million entities, providing a granular view of commercial exchanges within the Chinese market. It provides information on each transaction, including the location, industry classification, and transaction values for all transactions subject to VAT reporting requirements in China. VAT filing in China is required for all transactions in order to claim input tax credits, and failures to comply are punishable by law. However, it is important to note that there are certain exceptions to transactions that fall outside the purview of the VAT invoice policy. In particular, invoices are not issued for direct sales to consumers and small firms are exempt from VAT payments: firms with average monthly sales below RMB 30,000 (approximately USD 4,300) in a given year. We use the 2018 cross section in the main analysis.

Firm Registration Data We combine the VAT dataset with the firm registration database released by the Chinese State Administration for Market Regulation. This data set covers the universe of all firms, more than 200 million in total, that were ever registered in China. It contains detailed information about a firm's location, the year of its establishment and exit (if any), the industry, the value of its registered capital, its investment history, its initial main shareholders, and the records of any subsequent changes in the main shareholders, etc. To

distinguish firms by its ownership type, we identify each firms' ultimate shareholder through the circulated tracing process (also known as the depth search algorithm). A firm is identified as a state-owned enterprise (SOE) if a government bureau is one of the firms ultimate owners, and a private-owned enterprise (POE) otherwise.

Settlement Record of Enterprise Income Tax The third dataset is the annual settlement and payment record of enterprise income tax, which covers all firms that are subject to income tax filings.[1] From firms tax settlement records, we retain information on firm-level sales, employment, and detailed industry classification.

Administrative Tax Records Lastly, we supplement the data with the administrative enterprise income tax records from the Chinese State Administration of Tax (SAT) for the years 2008-2020, covering a representative sample of more than 1 million firms from stratified sampling.[2] For tax collection and audit purposes, SAT collects firm-level records of tax payments, as well as other financial statement information used in tax-related calculations. The advantages of tax data are two-fold. First, it is representative with wide coverage. As administrative data, it is not subject to the potential measurement error problem of self-reported data. Second, it contains detailed historical information on a wide array of tax-related information about the firms, allowing us to calculate the firms' TFP.

2.2 Sales Dispersion Across Firms and Transactions

We begin by documenting the highly dispersed distributions of both firm sales and firm-to-firm bilateral sales. Figure 1a shows the distribution of firms' log-transformed sales, demeaned within each industry, while Figure 1b shows the distribution of log firm-to-firm sales, demeaned within each supplier-customer industry pair. Even within industries and industry pairs, these variables exhibit significant heterogeneity. More importantly, firm-to-firm sales exhibit even greater variability than firm sales, providing evidence of the widespread heterogeneity in firm-to-firm sales.

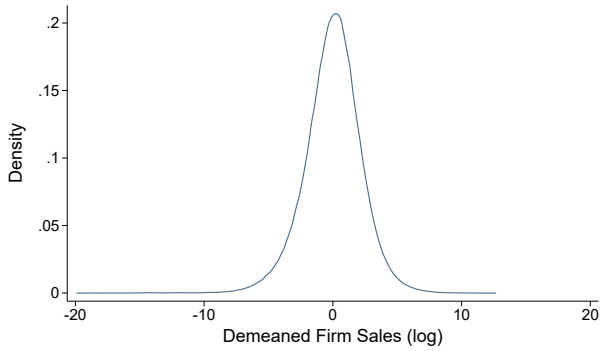
2.3 Firm Sales Decomposition

To understand why firm-to-firm sales heterogeneity matters and how it contributes to firm heterogeneity, we decompose firms' network sales into two components: the extensive margin (i.e., the number of customers) and the intensive margin (i.e., the average transaction size).

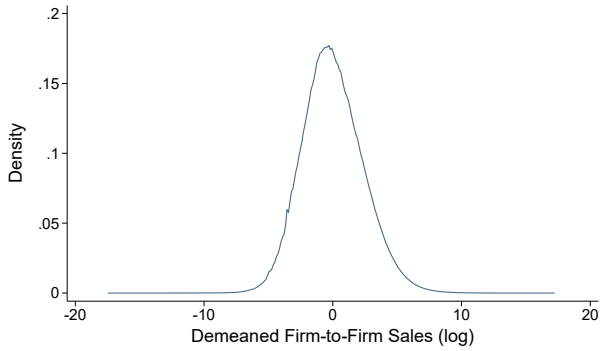
[1] Subsidiaries and firms with fixed-amount income tax are exempted from annual filing.
[2] China's SAT is the equivalent of the Internal Revenue Service (IRS) in the US.

Figure 1: Distribution of Firm Sales and Bilateral Sales

(a) Firm Sales Distribution



(b) Firm-to-Firm Sales Distribution



Notes: Sales are measured by aggregating the total annual sales for each firm across all transactions. Firm-to-firm sales are measured by aggregating the total annual transaction volume for each supplier-customer pair.

By definition, a firm's total sales are the sum of its bilateral sales with each customer in the network:

$$S_i = \sum_{j \in \mathcal{C}_i} m_{ij},$$

where $\mathcal{C}_i$ is the customer set of firm i, and m_{ij} is the transaction volume from firm i to its customers j. The expression can be rewritten as the product of the number of customers and the average transaction size with each customer. Taking the logarithm, a firm's sales can be decomposed as follows:

$$\ln S_i = \ln n_i + \ln \bar{s}_i, \tag{1}$$

where S_i represents the network sales of firm i, n_i^c denotes the number of customers of firm i, and $\bar{s}_i$ is the average sales of firm i with its customers. Regressing $\ln n_i$ and $\ln \bar{s}_i$ on $\ln S_i$ respectively yields the coefficients α^n and $\alpha^{\bar{s}}$, capturing how much of the variation in firm sales can be attributed to the extensive margin and the intensive margin. By the properties of OLS, this decomposition allocates the covariance term between the components of firm sales equally across components, implying that $\alpha^n + \alpha^{\bar{s}} = 1$. To control for industry heterogeneity, we demean all variables within each industry before the regression.

Table 1 presents the results of the variance decomposition. A striking finding is that 37% of the variation in firm sales is attributed to the variation in the number of customers (i.e., the extensive margin), while 63% is attributed to the variation in the average transaction size (i.e., the intensive margin). The coefficient for the intensive margin is nearly twice that of the extensive margin, highlighting the importance of understanding transaction size heterogeneity in explaining firm size heterogeneity. To be a large firm in terms of sales, it matters more whether a firm performs well in each transaction, rather than expanding its customer base.

Moreover, the two margins are negatively correlated, with a correlation coefficient of -0.13, suggesting a trade-off between the extensive and intensive margins. This pattern is consistent with mechanisms such as capacity constraints and relationship-specific investments (e.g., customized partnerships). Bernard et al. (2022) document a similar negative relationship using Belgian data and attribute it to the negative correlation between the two dimensions of firm heterogeneity, namely productivity and relationship capability. In this paper, we focus on the dominant role of average transaction size in shaping firm sales heterogeneity, rather than on the tension between the two margins. Because average transaction size is inherently a firm-to-firm object, this naturally directs our attention to firm-to-firm sales heterogeneity.

Table 1: Firm Sales Decomposition

N	$\ln n_i$	$\ln \bar{s_i}$	$\text{corr}(\ln n_i, \ln \bar{s_i})$
6,961,996	37%	63%	-0.13

Notes: The table presents the decomposition results of firm network sales. All variables are demeaned within industries.

We further conduct the decomposition for each industry to explore industry heterogeneity, as reported in Table 2. Consistent with the full sample decomposition, we find that the average transaction size is more important for most industries, with the coefficient $\alpha^{\bar{s}}$ greater than 0.5. This suggests that the pattern observed in Table 1 is more likely to reflect a common pattern across the economy, rather than being driven by a few dominant industries. The only exception is the Accommodation and Catering Industry, where the number of customers accounts for two-thirds of the variation in firm sales. Additionally, the negative correlation of the extensive and intensive margins also remains robust for most industries, with the Accommodation and Catering Industry exhibiting the strongest negative correlation (-0.41). Some manufacturing industries (e.g., Agricultural Food Manufacturing, Mining, Metal Smelting, Non-metallic Mineral Manufacturing) and the Construction industry show weakly positive or neutral correlation between the two margins. Overall, while industry heterogeneity exists, most industries share a common pattern: heterogeneity in the intensive margin is more pronounced than that in the extensive margin. This motivates us to further explore the origins of firm-to-firm transaction heterogeneity.

2.4 Roles of Suppliers and Customers

Having established that firm-to-firm transactions explain most of the variation in firm size, we further decompose it into supplier, customer, and match-specific margins. The supplier component reflects characteristics of a firm's suppliers that remain constant across its customers.

Table 2: Firm Sales Decomposition by Industry

			Decomposition		
Industry	Obs.	$\text{Var}(\ln S_i^{net})$	α^n	$\alpha^{\bar{s}}$	$\rho_{\ln n_i^c, \ln \bar{s}_i}$
All Industries	6,961,996	4.83	0.37	0.63	-0.13
Accommodation & Catering	44,167	5.59	0.66	0.34	-0.41
Agricultural Food Manufacturing	54,957	6.72	0.36	0.64	0.10
Agriculture, Forestry, Animal Husbandry & Fisheries	27,151	5.88	0.26	0.74	0.01
Chemical Manufacturing	213,761	4.80	0.45	0.55	-0.04
Computer Manufacturing	79,823	4.82	0.41	0.59	-0.06
Construction	551,839	4.55	0.30	0.70	0.05
Culture, Sports & Entertainment	43,027	5.08	0.29	0.71	-0.13
Education	2,104	5.78	0.46	0.54	-0.16
Electrical Manufacturing	120,157	4.79	0.43	0.57	-0.07
Electricity, Heat, Gas and Water Production & Supply	54,061	6.82	0.36	0.64	-0.16
Entertainment Equipment Manufacturing	38,208	4.13	0.36	0.64	-0.10
Finance	37,193	7.63	0.42	0.58	-0.14
General Equipment Manufacturing	227,226	3.67	0.44	0.56	-0.15
Health & Social Work	916	6.59	0.32	0.68	-0.08
Information Transmission, Software & IT Services	250,157	5.02	0.39	0.61	-0.13
Instrument Manufacturing	24,027	3.99	0.45	0.55	-0.24
Leasing & Business Services	487,056	4.97	0.31	0.69	-0.30
Metal Manufacturing	175,732	3.86	0.44	0.56	-0.10
Metal Smelting	30,328	6.39	0.40	0.60	0.04
Mining	20,863	6.18	0.28	0.72	0.12
Non-metallic Mineral Manufacturing	106,176	4.70	0.44	0.56	0.04
Other Manufacturing	62,472	4.25	0.36	0.64	-0.18
Others	50,274	5.99	0.42	0.58	-0.14
Printing Manufacturing	46,011	3.82	0.43	0.57	0.00
Real Estate	91,837	4.94	0.34	0.66	-0.22
Residential Services, Repairs & Other Services	78,918	4.24	0.34	0.66	-0.29
Scientific Research & Technical Services	627,087	4.75	0.40	0.60	-0.14
Specialized Equipment Manufacturing	130,302	4.14	0.42	0.58	-0.08
Textile Manufacturing	205,337	4.19	0.39	0.61	-0.11
Transportation Equipment Manufacturing	66,407	5.63	0.34	0.66	0.03
Transportation, Warehousing & Postal Services	251,455	4.39	0.36	0.64	-0.15
Water Conservancy, Environment & Public Facilities Mgmt.	29,633	4.51	0.31	0.69	-0.12
Wholesale & Retail Trade	2,599,362	5.07	0.37	0.63	-0.14
Wood & Fur Manufacturing	133,972	3.66	0.43	0.57	-0.10

Notes: The table presents the decomposition results of firm network sales by industry. The industry classification follows the Industrial Classification for National Economic Activities (2011), with manufacturing sector further subdivided into 16 more detailed subcategories. The category *Others* refers to the sectors of public administration, social security, social organizations, and international organizations.

A good supplier may be characterized by strong appeal, scalable production, reliable delivery, and so on. Similarly, the customer component reflects characteristics of a firm's customers that remain constant across suppliers. A good buyer may be characterized by large demand, efficient procurement, prompt payment, and more. Firm-to-firm sales are jointly determined by supplier and customer capabilities, raising a natural question about their relative importance. Before

the formal analysis, we present a preliminary visualization to provide some initial insights.

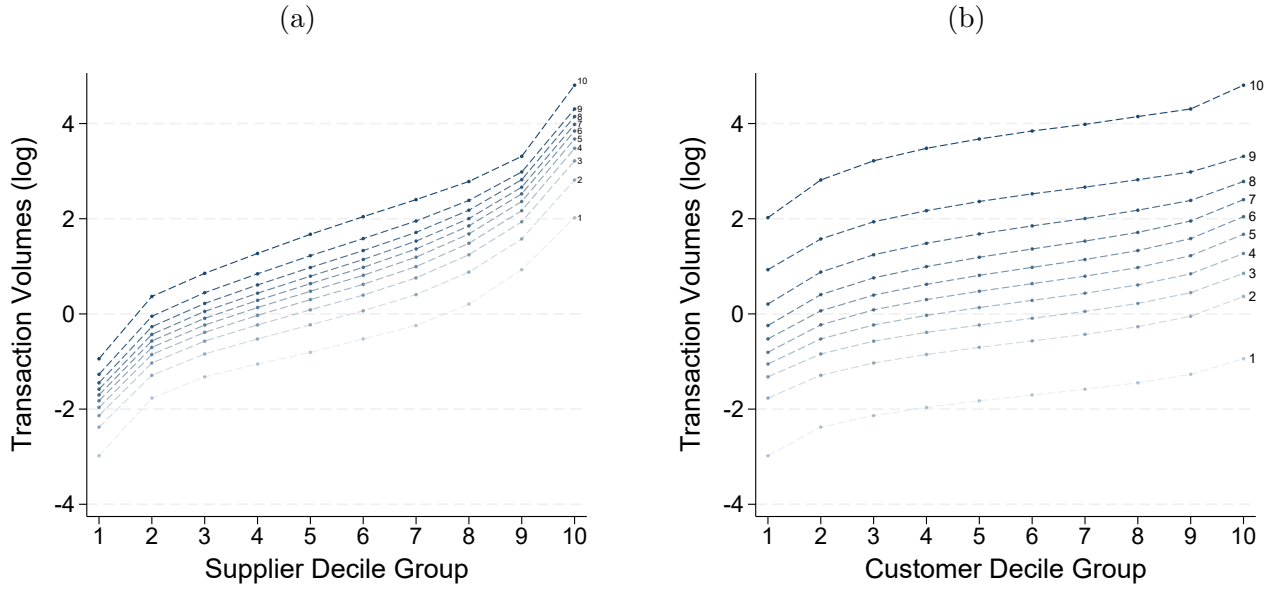
Leveraging the rich connectivity of the firm-to-firm production network, we examine the role of two-sided heterogeneity by studying how average firm-to-firm sales vary across supplier and customer decile groups. We first demean the log of firm-to-firm sales within each supplier-customer industry pair to control for industry-pair complementarity, ensuring that any variation in sales is not driven by systematic differences in industry-specific demand conditions. For example, industry pairs such as construction and real estate, or mining and metal smelting, may exhibit strong complementarity, so that transactions in these pairs are systematically larger than those in other industry pairs. We then compute, for each supplier, the mean log sales across its customers, denoted by $\overline{\ln s_i}$, and, for each customer, the mean log purchases across its suppliers, denoted by $\overline{\ln m_j}$. Firms with only one trading partner are excluded, since they provide no within-firm variation across trading partners. Next, we sort suppliers by $\overline{\ln s_i}$ and customers by $\overline{\ln m_j}$ into deciles, assigning them to groups $q_s = 1, \ldots, 10$ and $q_m = 1, \ldots, 10$, respectively. Higher deciles correspond to larger suppliers or larger customers. Finally, for each supplier-customer decile pair, we compute the mean of $\ln m_{ij}$, yielding a 10×10 matrix in which each element represents the average log transaction value between suppliers in the i-th supplier decile group and customers in the j-th customer decile group.

Figure 2 provides a structured visualization of the matrix. In Figure 2a, each line represents a customer group, and the slope of the line illustrates how the average transaction volume changes across different supplier groups. Similarly, Figure 2b shows the average transaction volume across customer groups, with each line representing a supplier group and the slope indicating how the average transaction volume changes across different customer groups.

Figure 2 highlights two important patterns: *parallelism* and *asymmetry*. *Parallelism* refers to the near-parallel shifts in transaction volumes across groups, suggesting that firm-to-firm transaction volumes can be decomposed additively. *Asymmetry*, on the other hand, refers to the steeper increase in transaction volumes across supplier groups than across customer groups. This suggests that supplier heterogeneity plays a more important role in shaping firm-to-firm transactions than customer heterogeneity.

Collectively, these results point us toward a more quantitative understanding of two-sided heterogeneity in the firm-to-firm production network, which is the central focus of this paper.

Figure 2: Average Transaction Volumes across Decile Groups



Notes: The figures display y_{ij} across supplier and customer decile groups. Each line represents a supplier or customer decile group, with darker colors and larger values indicating larger groups. The firm-to-firm sales are demeaned within supplier-customer industry-pairs.

3 Estimation

In this section, we aim to investigate the two-sided heterogeneity in the firm-to-firm production network using a formalized and structured empirical framework. Disentangling the components of firm-to-firm sales variation attributable to supplier-specific and customer-specific heterogeneity allows us to gain a clearer understanding of the sources of variation in firm-to-firm transaction volumes.

3.1 Firm-to-Firm Sales Decomposition

Do firm-to-firm transactions depend more on suppliers or on buyers? To study two-sided heterogeneity in a disciplined way, it is natural to draw on the canonical wage decomposition framework of Abowd, Kramarz and Margolis (1999) (hereafter, AKM), which is widely used in labor economics. Our outcome of interest, firm-to-firm sales, is analogous to wages in that it is generated by a two-sided match between a supplier and a customer. Moreover, the near-additive separability documented in Figure 2 provides direct empirical support for a two-way fixed effects specification in the spirit of AKM. Taken together, these conceptual and structural parallels motivate us to adapt the AKM approach to quantify two-sided heterogeneity in firm-to-firm production networks.

Specifically, we model firm-to-firm sales from supplier i to customer j as follows:

$$y_{ij} = G + \psi_i + \theta_j + \epsilon_{ij}, \tag{2}$$

where y_{ij} denotes log sales from supplier i to customer j, demeaned at the supplier-customer industry-pair level. Here, ψ_i denotes the supplier fixed effect, θ_j denotes the customer fixed effect, and ϵ_{ij} is the residual. The two sets of fixed effects, $\{\psi_i, \theta_j\}$, capture two-sided heterogeneity in firm-to-firm production networks.

We demean bilateral sales at the supplier-customer industry-pair level for two main reasons. First, our primary focus is on heterogeneity in firm-to-firm sales within industry pairs. By doing so, we eliminate variation driven by differences in input-output structures across industries. This demeaning process helps remove demand shifts arising from industry-specific heterogeneity. Second, the specification in equation (2) assumes additive separability in firm-to-firm sales. To ensure consistent estimation, it is crucial to eliminate any systematic complementarities in bilateral sales across firm-type matches. In practice, demeaning at the industry-pair level effectively mitigates these potential complementarities.

Table 3: Degree Distribution

		P10	P25	P50	P75	P90
# Suppliers per Firm	Full	1	1	6	21	55
	LOO	2	4	11	30	71
# Customers per Firm	Full	1	2	7	20	57
	LOO	2	4	9	26	68

Notes: *Full* refers to the full 2018 cross section, and *LOO* refers to the leave-one-out connected set of firms used in the decomposition. P10 denotes the 10-th percentile of the degree distribution; other percentile labels are defined analogously.

Compared to the standard wage decomposition, we expect our specification to be more relevant and reasonable. A key distinguishing feature of firm-to-firm production networks is the presence of many-to-many matches, as shown in Table 3. The median numbers of suppliers and customers in the full sample are 6 and 7, respectively. First, these many-to-many matches significantly enhance the connectedness of the data, thereby reducing the severity of the limited mobility problem. Second, they allow us to identify both supplier and customer fixed effects using cross-sectional variation, relaxing the standard assumption that fixed effects remain constant over time. Furthermore, we can assess this assumption by examining the persistence of the fixed effects.

Identification Estimating equation (2) using OLS raises identification concerns, primarily due to potential complementarity in bilateral sales and the connectedness requirement for separately identifying supplier and customer fixed effects. To obtain unbiased estimates, we impose two assumptions for identification. First, a conditional exogeneity assumption rules out selection on idiosyncratic match-specific shocks. Conditional on supplier and customer fixed effects, the assignment of suppliers to customers must be exogenous with respect to the match shock:

$$\mathbb{E}[\varepsilon_{ij} \mid i, j] = 0.$$

We conduct various tests of exogenous mobility and find strong support for the conditional exogenous mobility assumption.

Second, the identification of fixed effects requires matches with multiple trading partners. Firms must be part of a connected set, meaning that they are directly or indirectly linked through supplier-buyer relationships. Without this, if all matches are one-to-one, supplier and customer fixed effects cannot be separately identified. Following Kline, Saggio and Sølvsten (2020) (hereafter, KSS), we restrict the sample to the largest leave-one-out (LOO) connected set, which is the largest group of firms that remains connected after removing any single firm. The KSS procedure does not alter the fixed-effect point estimates but corrects for finite-sample bias in second-order moments due to estimation error. Therefore, we estimate the model on the LOO connected set and apply the KSS correction to obtain bias-corrected covariances of supplier and customer fixed effects, which are then used to evaluate two-sided heterogeneity. As shown in Table 4, the firm-to-firm network is well connected, with the LOO set covering over 98% of supplier-customer pairs in the full set.

Table 4: Sample Coverage

	# Suppliers	# Customers	# Pairs
Full Set	6,961,996	10,053,187	240,767,236
LOO Set	5,756,331	7,362,667	236,881,882
Coverage	82.68%	73.24%	98.39%

Notes: *Full* refers to the full 2018 cross section, and *LOO* refers to the leave-one-out connected set of firms used in the decomposition.

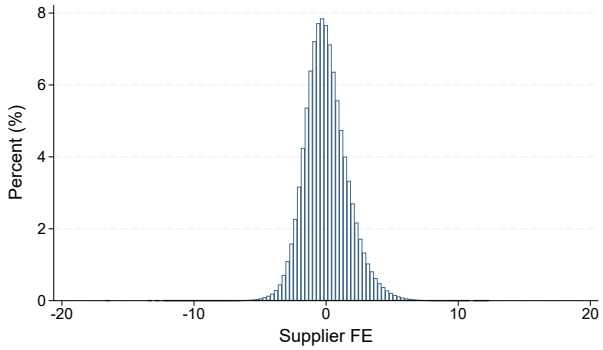
Results Figure 3 shows the distribution of the estimated supplier and customer fixed effects. A striking result is that the distribution of supplier fixed effects is much more dispersed than that of customer fixed effects. Moreover, equation (2) can be used for variance decomposition

as follows:

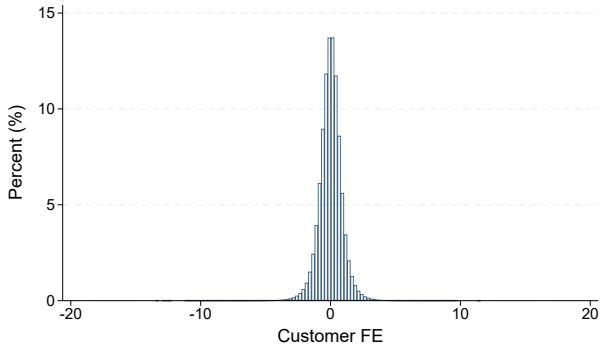
$$\text{Var}\left(y_{ij}\right) = \underbrace{\text{Var}\left(\psi_i\right)}_{\text{Supplier effects}} + \underbrace{\text{Var}\left(\theta_j\right)}_{\text{Customer effects}} + \underbrace{2\,\text{Cov}\left(\psi_i, \theta_j\right)}_{\text{Sorting}} + \underbrace{\text{Var}\left(\varepsilon_{ij}\right)}_{\text{Residual}}. \tag{3}$$

Figure 3: Distribution of Supplier and Customer FEs

(a) Supplier FE Distribution



(b) Customer FE Distribution



Notes: Figure 3a plots the distribution of the supplier fixed effects. Figure 3b plots the distribution of the customer fixed effects. Both sets of fixed effects are demeaned.

Table 5 reports the decomposition results, including both the plug-in estimators and the bias-corrected estimators proposed by Andrews et al. (2008) and Kline et al. (2020), respectively. As shown in Table 5, the variance of supplier fixed effects is more than six times larger than that of customer fixed effects, accounting for over 40% of the variation in firm-to-firm sales. The substantial difference in the variances of supplier and customer fixed effects further supports the conjecture that supplier heterogeneity is much larger, consistent with the preliminary visualization presented in Section 2.4. Together, the two sets of fixed effects explain over one-half of the variation in firm-to-firm sales, with an R^2 of 0.53 in the AKM baseline. Results from the KSS correction closely align with the AKM plug-in estimators, suggesting that finite-sample bias and heteroskedasticity have little impact on the fixed-effects estimates in our sample. This similarity also indicates that our estimation sample is well connected and that the fixed-effect point estimates are relatively reliable, supporting the validity of our subsequent analysis. Additionally, there is little sorting in the transaction-level data, with the correlation close to zero.

3.2 Validation

To assess the reliability of our estimated components, we conduct a battery of validation exercises that probe the conditional exogeneity assumption and the credibility of our empirical design.

Table 5: Firm-to-Firm Sales Decomposition

	AKM	AGSU	KSS
$\text{Var}(\psi_i)$	44.75%	43.50%	43.34%
$\text{Var}(\theta_j)$	7.82%	6.23%	6.47%
$\text{Cov}(\psi_i, \theta_j)$	0.18%	0.22%	0.22%
$\text{Var}(\psi_i)/\text{Var}(\psi_i + \theta_j)$	85%	87%	86%
$\text{Corr}(\psi_i, \theta_j)$	0.01	0.01	0.01
R^2	0.53	0.50	0.50

Notes: This table reports the decomposition of firm-to-firm sales for the 2018 cross section. The firm-to-firm sales are demeaned at the supplier-customer industry-pair level. AKM refers to unadjusted fixed effects of Abowd, Kramarz and Margolis (1999). AGSU refers to the homoskedastic fixed-effects bias-correction method of Andrews, Gill, Schank and Upward (2008). KSS refers to the heteroskedastic fixed-effects bias-correction method of Kline, Saggio and Sølvsten (2020).

Exogenous matches In equation (2), to identify the supplier and customer fixed effects, we impose the assumption of exogenous matches. Specifically, we assume that matches between suppliers and customers are uncorrelated with unobserved factors that could affect transaction volumes. In wage decomposition using panel data, a common approach to testing this assumption is the event-study analysis proposed by Card et al. (2013). The idea is that, under exogenous mobility, wage gains and losses should be symmetric for movers between higher- and lower-wage firms. In cross-sectional data, matches are observed only as static many-to-many relationships; thus, we do not observe the dynamics and cannot conduct an event study. However, we borrow this insight and conduct similar tests, as shown in Figure 2 in Section 2.4. Specifically, we group suppliers and customers into decile groups based on their average transaction size, which serves as a proxy for supplier and customer type. If matches are exogenous, we would expect broadly parallel changes in y_{ij} across supplier (customer) groups for different trading-partner groups. In Figure 2, we do not see systematic departures from parallelism, which is consistent with this assumption.

Complementarity Test By decomposing firm-to-firm sales into the sum of supplier and customer fixed effects, we implicitly impose the assumption of no complementarity between supplier-customer type pairs. Figure 2 visually supports this assumption, and we further examine it through regression analysis. Specifically, we regress y_{ij} on the estimated fixed effects $\{\hat{\psi}_i, \hat{\theta}_j\}$ and an interaction term capturing potential complementarity. The results are reported in Table 6. Three key points should be highlighted. First, in columns (1) and (2), the coefficients on supplier and customer fixed effects are close to, but slightly larger than, one, implying weak positive assortativity in matching, which is consistent with the correlation of 0.01 reported in

Table 5. Second, the R^2 from the supplier fixed effects (0.45) is more than five times larger than that from the customer fixed effects (0.08), underscoring the importance of supplier heterogeneity. Third, there is little evidence of complementarity. Comparing columns (3) and (4), the interaction of the two fixed effects explains very little of the variation in firm-to-firm sales, with the R^2 increasing by only 0.0006. Moreover, even with the interaction term, the coefficients on both supplier and customer fixed effects remain close to 1.

Table 6: Complementarity Test

	log (Firm-to-Firm Sales)			
	(1)	(2)	(3)	(4)
Supplier FE	1.004***		1.000***	1.070***
	(.000)		(.000)	(.000)
	[13944.845]		[15046.107]	[6429.724]
Customer FE		1.022***	1.000***	0.923***
		(.000)	(.000)	(.000)
		[4411.835]	[5846.884]	[3956.736]
Supplier × Customer FE				0.053***
				(.000)
				[484.855]
Constant	-1.353***	1.386***	0.000	-0.103***
	(.000)	(.000)	(.000)	(.000)
	[-9524.607]	[3844.999]	[0.000]	[-285.590]
R^2	0.4510	0.0818	0.5292	0.5298
Observations	236,881,882	236,881,882	236,881,882	236,881,882

Standard errors in parentheses, and t statistics in square brackets.

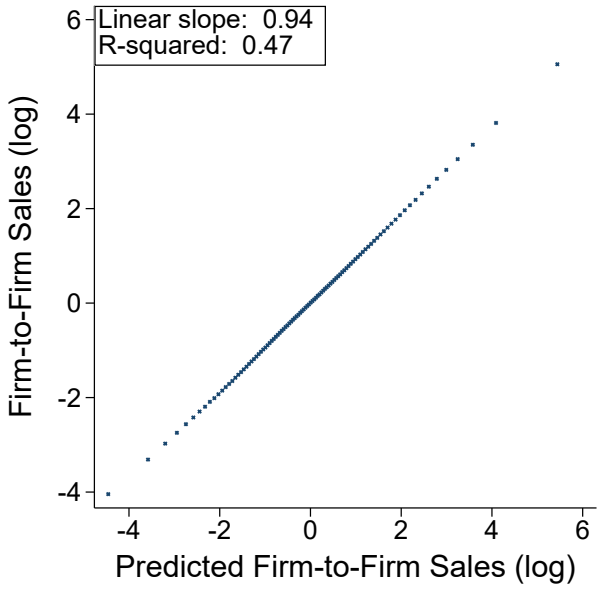
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table reports the regression results of firm-to-firm sales on the estimated supplier and customer fixed effects. Firm-to-firm sales are demeaned at the supplier-customer industry-pair level.

Cross-validation Since fixed effects are estimated with some noise, we pursue a split-sample approach to examine the relevance of our estimated parameters and the validity of the decomposition framework. We randomly split the full sample into a training set and a test set, with the training set used to estimate the supplier and customer effects and the test set used to assess the predictive power of these estimates. With the supplier and customer effects estimated from the training set, we can examine how well they predict actual firm-to-firm transactions in the test

set. We use three different test sizes (20%, 10%, and 2%) to progressively mitigate the impact of sample size reduction on the estimates. Figure 4 shows that the predicted firm-to-firm sales align closely with the actual sales. Table A-1 reports the results from regressing firm-to-firm sales in the test set on the fixed effects estimated from the training set. Both supplier and customer effects are significantly correlated with firm-to-firm sales in the test set and perform better when both effects are included in the specification.

Figure 4: Actual vs. Predicted Firm-to-Firm Sales



Notes: This figure shows the relationship between predicted firm-to-firm sales, obtained from the fixed effects estimated using the training set, and actual firm-to-firm sales. The predicted firm-to-firm sales are calculated as the sum of the supplier and customer effects estimated from the training-set decomposition. The training set constitutes 80% of the sample, and the test set constitutes 20%.

4 Characterization of Supplier and Customer Effects

The firm-to-firm sales decomposition in Section 3 provides a comprehensive view of two-sided heterogeneity in firm-to-firm production networks, emphasizing the asymmetric roles of suppliers and customers and the critical importance of suppliers in these transactions. In this paper, we take the analysis one step further by going beyond the variance components. How do these components correlate with other dimensions of firms? In this section, we examine the point estimates of the supplier and customer fixed effects by linking them to other firm-level characteristics to better understand their implications.

4.1 Persistence of Supplier and Customer Effects

How persistent are supplier and customer effects? Do they capture only temporary shocks, or do they also reflect more stable, long-term characteristics of firms? To address these questions, we examine the extent to which these effects remain stable over time. If the estimated effects exhibit significant persistence, this would suggest that they are driven by long-term, structural characteristics of firms, such as product appeal, market position, production capabilities, or enduring relationships with suppliers and customers. These persistent effects are likely influenced by factors that are less sensitive to short-term shocks, implying that supplier and customer heterogeneity, as captured by the fixed effects, plays a key role in shaping firm-to-firm interactions and performance in the long run. Conversely, if supplier and customer effects show low persistence, this may indicate that they are primarily driven by short-term shocks, such as temporary shifts in firm behavior or market demand. In this case, the effects would be less stable and more responsive to immediate environmental changes, suggesting that firm-to-firm sales patterns are primarily influenced by transient factors rather than long-term strategic decisions or structural differences among suppliers and customers.

Table 7: Correlation of Fixed Effects over Time

(a) Supplier Fixed Effects

Year	2014	2015	2016	2017	2018
2014	1.00				
2015	0.91	1.00			
2016	0.82	0.86	1.00		
2017	0.64	0.65	0.76	1.00	
2018	0.59	0.59	0.68	0.84	1.00

(b) Customer Fixed Effects

Year	2014	2015	2016	2017	2018
2014	1.00				
2015	0.95	1.00			
2016	0.87	0.87	1.00		
2017	0.82	0.82	0.87	1.00	
2018	0.59	0.58	0.60	0.61	1.00

Notes: Panel a reports the correlation matrix of supplier fixed effects across years, and Panel b reports the correlation matrix of customer fixed effects across years. The sample period spans 2014–2018. All correlations are statistically significant at the 1% level.

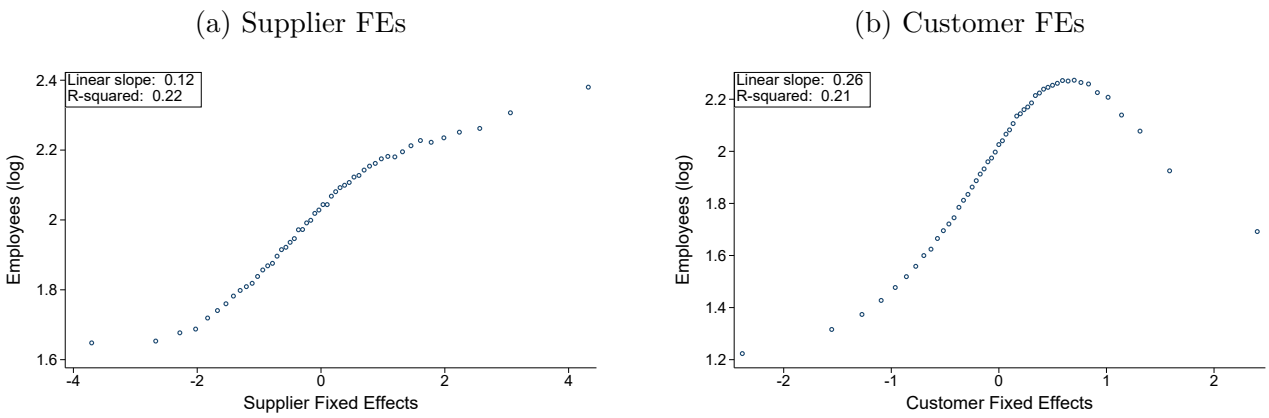
We re-estimate equation (2) using cross-sectional data from 2014 to 2018 and obtain $\{\psi_{it}, \theta_{jt}\}$ for $t \in \{2014, 2015, 2016, 2017, 2018\}$. Table 7a and Table 7b present the pairwise correlations of supplier and customer effects for the same firms across years, respectively. For adjacent years, the correlation is consistently high, suggesting that these effects reflect firms' intrinsic characteristics rather than being driven by temporary shocks or year-specific noise. As the time gap between years increases, the correlations gradually decline, indicating that the persistence of supplier and customer effects diminishes over time. This decline in correlation is also consistent with the dynamic nature of firms, reflecting long-term adjustments and life-cycle growth. The dynamics of the fixed effects suggest that the assumption of their constancy over time may be reasonable to some extent. However, the panel should not span too long a period, as this

assumption may become less valid over extended time horizons.

4.2 Relationship with Firm Size and Productivity

Firm size Supplier fixed effects, in essence, capture the average sales of suppliers after controlling for the composition of their customer sets. Similarly, customer fixed effects capture the average purchases of customers after controlling for the composition of their supplier sets. These effects are naturally linked to firm size. Figure 5 illustrates the relationship between firm size and these fixed effects, using the number of employees as a proxy for firm size. Firms with higher supplier fixed effects tend to be larger, and this relationship is fairly monotonic. However, customer fixed effects do not exhibit a monotonically increasing relationship with firm size. Instead, they show a concave relationship characterized by an inverse U-shape. A similar concave pattern is documented in the labor market by Bloom et al. (2018), where the largest firms have wage effects comparable to or lower than those of their smaller counterparts. Overall, these patterns suggest that although firm size plays a significant role, it is not the sole determinant of two-sided heterogeneity within the network. The estimated supplier and customer components capture much more than firm size alone.

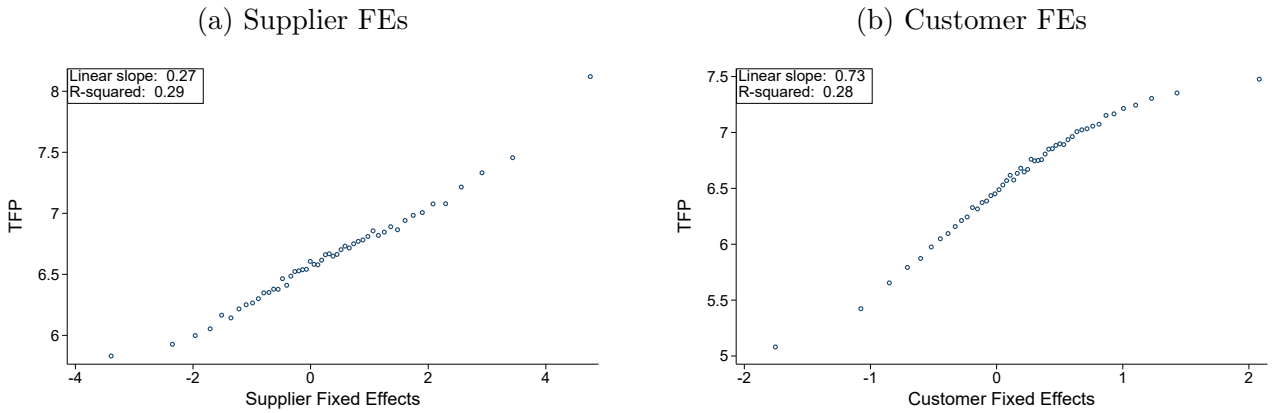
Figure 5: Fixed Effects and Firm Size



Notes: The figures plot the relationship between firm size and supplier (customer) fixed effects. Supplier and customer fixed effects are demeaned, and we control for industry fixed effects in the binned scatter plots.

Productivity Another variable of broad interest is firm productivity, whose estimation is challenging. Do our estimated supplier and customer fixed effects capture information about firm productivity? The answer is yes. Firms with higher supplier or customer effects tend to exhibit higher productivity in the data. Figure 6 shows that both supplier and customer fixed effects are positively correlated with firms' TFP. This relationship remains robust even after controlling

Figure 6: Fixed Effects and TFP



Notes: The figures plot the relationship between firms' TFP and supplier (customer) fixed effects. Supplier and customer fixed effects are demeaned, and we control for industry fixed effects in the binned scatter plots.

for firm size or the number of suppliers and customers, further supporting the notion that these fixed effects provide insights beyond firm size. Additionally, the consistent positive correlation suggests that we can use the estimated fixed effects as proxies for TFP, with a much simpler estimation procedure.

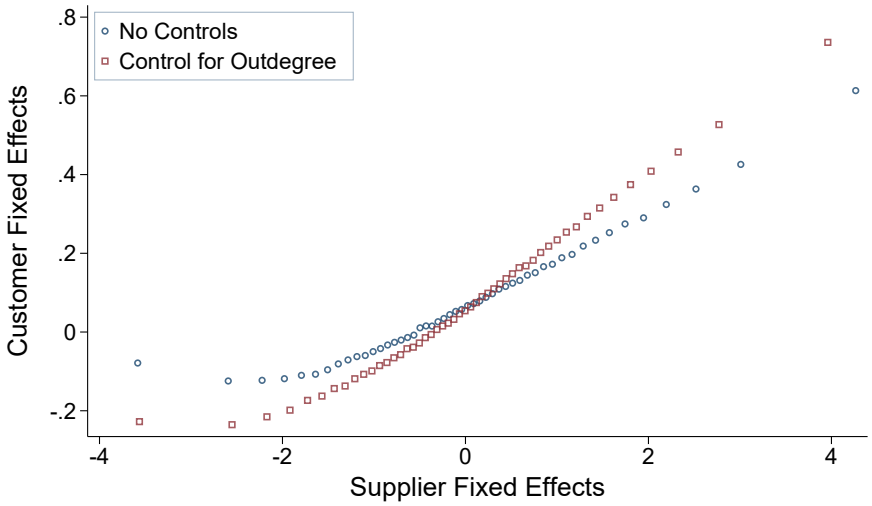
The above analysis establishes that the estimated supplier and customer effects are informative and empirically relevant, as they are positively correlated with both firm size and productivity. In the rest of the paper, for simplicity, we refer to firms with higher supplier (customer) fixed effects as *good* suppliers (customers).

4.3 Good Suppliers Are Also Good Customers

A distinct feature of firm-to-firm production networks is that firms act as both suppliers and customers, and we estimate their supplier and customer effects separately. Are good suppliers also good customers? Figure 7 visualizes the relationship between firms' customer fixed effects and supplier fixed effects, revealing that firms with higher supplier fixed effects tend to have higher customer fixed effects. The positive relationship becomes even stronger after controlling for the number of customers. This suggests that supplier and customer effects reflect firms' intrinsic characteristics, influencing their dual roles as both suppliers and customers.

It is precisely because firms serve as both suppliers and customers that the much more pronounced supplier heterogeneity is such a striking finding. In firm-to-firm production networks, suppliers and customers are essentially the same set of firms, yet the heterogeneity they exhibit as suppliers and as customers differs markedly. To the best of our knowledge, few theories in the literature are able to rationalize this finding.

Figure 7: Customer Fixed Effects vs. Supplier Fixed Effects



Notes: The figure plots the relationship between firms' customer fixed effects and supplier fixed effects. Supplier and customer fixed effects are demeaned, and we control for industry fixed effects in the binned scatter plots.

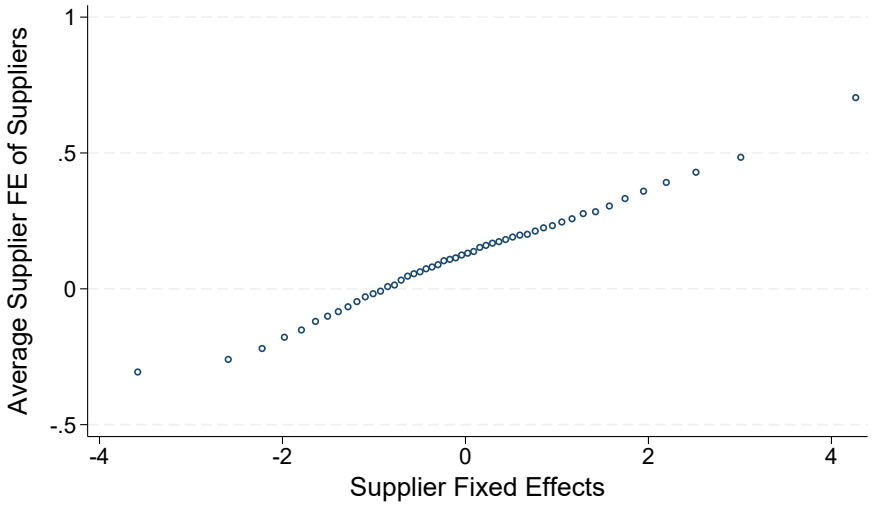
4.4 Good Suppliers Source from Good Suppliers

Good suppliers tend to source from other good suppliers, reflecting a positive sorting pattern in terms of supplier quality within firm-to-firm production networks. Figure 8 illustrates the positive correlation between firms' supplier quality, measured by the average supplier fixed effects of their suppliers, and their own supplier fixed effects. Note that the sorting pattern identified in this analysis is not contradicted by the nearly zero correlation observed in the variance decomposition, as the concept of sorting differs fundamentally between the two analyses. The correlation in the variance decomposition examines whether good suppliers are matched with good customers in firm-to-firm production networks, while the positive correlation shown in Figure 8 focuses on firms' roles as suppliers, indicating that good suppliers are more likely to source from other good suppliers. Furthermore, while the variance decomposition is based on network data, the measures in this section are obtained from firm-level data, with firms being mechanically reweighted after the data are reshaped

4.5 Fixed Effects and Degree

Fixed effects and degree capture different dimensions of a firm's importance within the network. Fixed effects, representing the intensive margin, reflect a firm's trading capacity in each transaction with its trading partners, while degree captures the extensive margin by measuring the number of connections each firm has. We examine how the two margins are correlated for

Figure 8: Supplier FE and Average Supplier FE of Suppliers



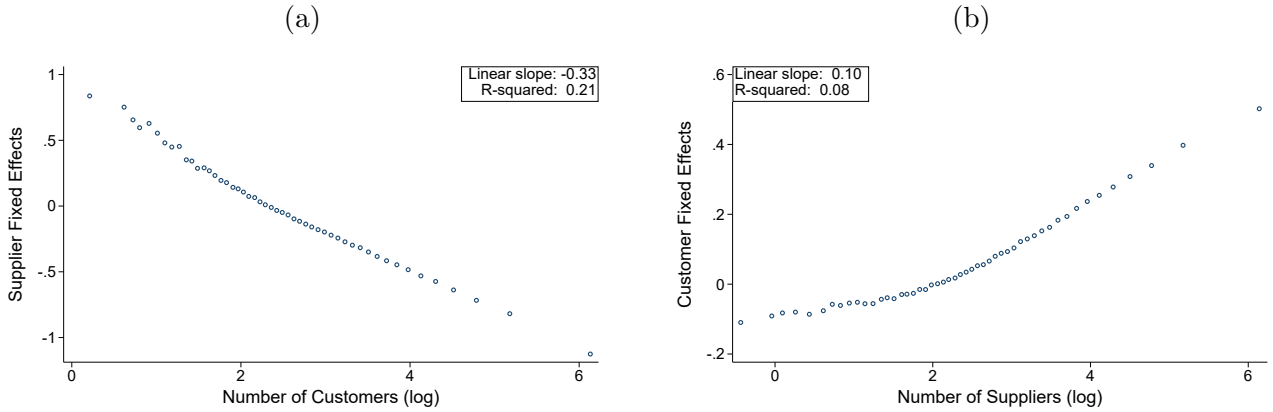
Notes: The figure plots the relationship between firms' supplier fixed effects and the average supplier fixed effects of their suppliers. Supplier fixed effects are demeaned, and we control for industry fixed effects in the binned scatter plots.

firms within the network.

Figure 9a documents a negative correlation between outdegree and supplier fixed effects. Outdegree refers to the number of distinct customer links a firm maintains in the firm-to-firm network, while supplier fixed effects capture heterogeneity on the seller side in terms of average transaction volume. This negative correlation suggests a tension between the extensive and intensive margins when evaluating a supplier's importance in the network. Several underlying mechanisms may drive this negative correlation, such as firms' limited capacity, the nature of customized or exclusive relationships, or other structural constraints within the network that limit a firm's ability to expand its customer base while maintaining high transaction volume. However, this tension between the extensive and intensive margins does not arise in the case of customers. As shown in Figure 9b, firms sourcing from more suppliers tend to exhibit higher customer fixed effects. This further highlights the fundamental distinction in firms' behavior when acting as suppliers versus customers within the network. Understanding the dual role of firms and the heterogeneity between their supplier and customer roles warrants further investigation in future research.

Good suppliers vs. large suppliers We refer to firms with higher supplier fixed effects as good suppliers and those with more customers as large suppliers. Figure 9a suggests that a supplier is less likely to be both good and large at the same time, reflecting a trade-off that many firms face. We investigate whether systematic differences exist between good suppliers

Figure 9: Fixed Effects and Degrees

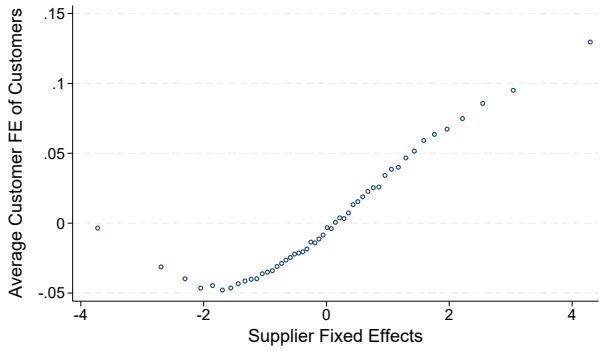


Notes: Figure 9a plots the relationship between firms' number of customers and their supplier fixed effects. Figure 9b plots the relationship between firms' number of suppliers and their customer fixed effects. Supplier and customer fixed effects are demeaned, and industry fixed effects are controlled for in the binned scatter plots.

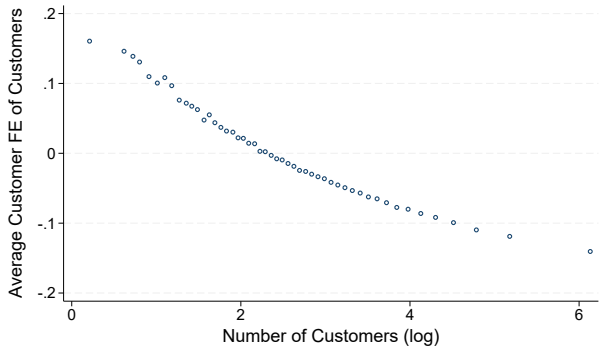
and large suppliers, particularly in terms of the characteristics of their customer base. Figure 10 highlights a clear difference in quality between the customer sets of good and large suppliers. Using the average customer fixed effect as a proxy for customer quality, Figure 10a shows that good suppliers tend to match with higher-quality customers, whereas Figure 10b shows that large suppliers tend to match with lower-quality customers, on average. More importantly, this pattern is not driven by composition effects among customers. Figure 10 shows that customer effects at different percentiles increase with a firm's own supplier fixed effects and decrease with the firm's outdegree. Therefore, the observed differences in customer quality between good and large suppliers are not driven only by differences in their top or bottom customers, but instead reflect a broader difference in the overall quality of their customer base.

Figure 10

(a) Supplier FE and Average Customer



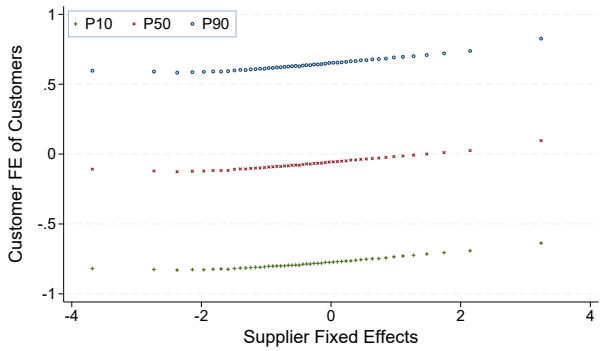
(b) Outdegree and Average Customer



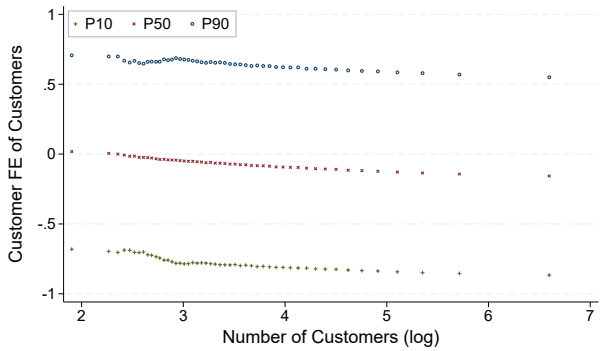
Notes: Figure 10a plots the relationship between firms' supplier fixed effects and the average customer effects of their customers. Figure 10b plots the relationship between firms' number of customers and the average customer effects of their customers. Supplier and customer fixed effects are demeaned, and both binned-scatter plots control for industry fixed effects.

Figure 11

(a) Supplier FE and Customer Quality



(b) Outdegree and Customer Quality



Notes: Figure 11a plots the relationship between firms' supplier fixed effects and their customers' customer fixed effects at the 10th, 50th, and 90th percentiles (P10, P50, and P90). Figure 11b plots the relationship between firms' number of customers and their customers' customer fixed effects at the 10th, 50th, and 90th percentiles (P10, P50, and P90). Supplier and customer fixed effects are demeaned, and both binned-scatter plots control for industry fixed effects.

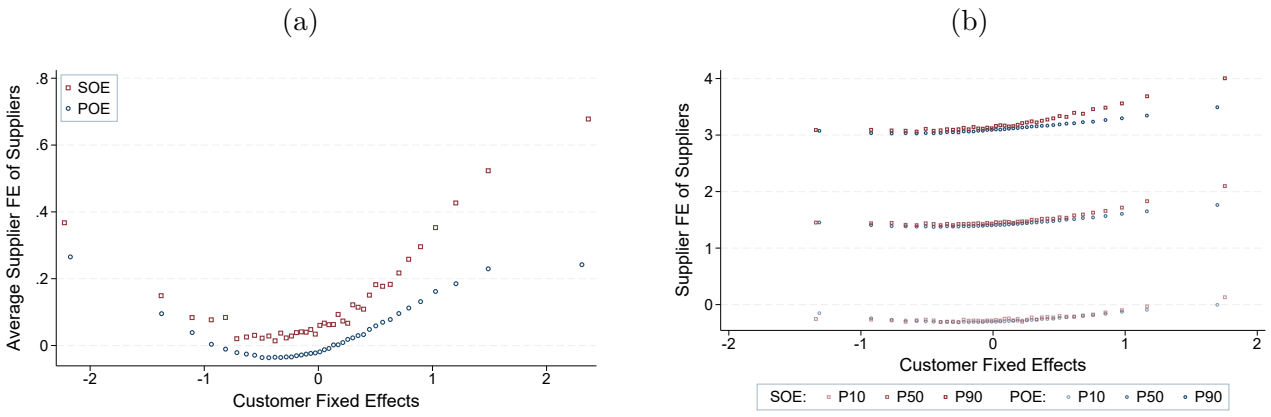
5 State-Owned Enterprises

In this section, we demonstrate, as an application, how the estimated supplier and customer components can be used to identify potential market power and misallocation associated with SOEs by comparing their matching patterns with those of privately owned enterprises (POEs, hereafter). In our sample, approximately 2.5% of firms are state-owned.

As shown in Figure 12, conditional on estimated customer effects, SOEs tend to match with higher-quality suppliers, suggesting the presence of market power associated with SOEs

or a supplier-side preference for SOEs. Panel b further breaks down supplier fixed effects by percentile and shows that this pattern is primarily driven by firms' top suppliers. In other words, when attention is restricted to top suppliers, SOEs are more likely than comparable POEs to match with exceptionally good suppliers.

Figure 12

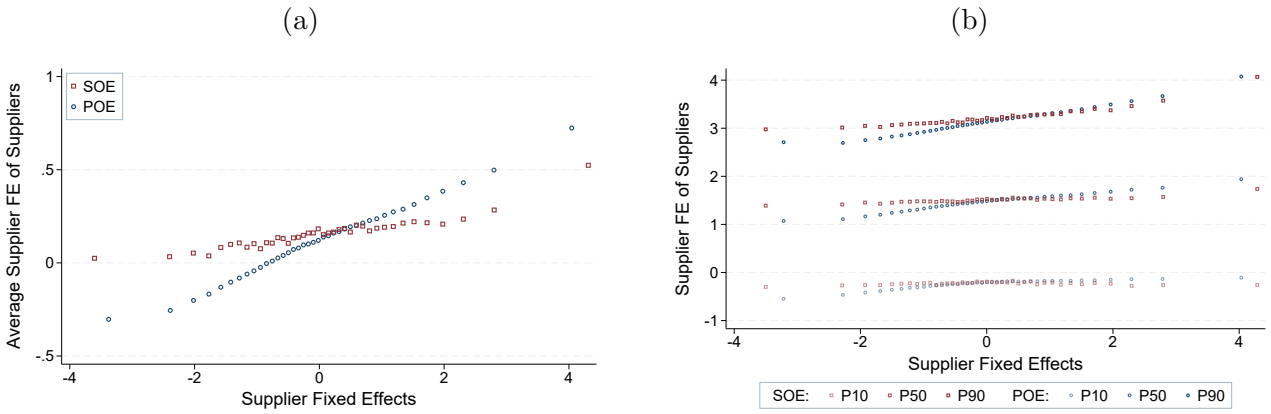


Notes: Figure 12a shows the relationship between firms' customer fixed effects and the average supplier fixed effect of their suppliers. Figure 12b plots the relationship between firms' customer fixed effects and the supplier fixed effects of their suppliers at the 10th, 50th, and 90th percentiles (P10, P50, and P90), where we restrict the sample to customers with at least 10 suppliers. We first demean both supplier and customer fixed effects. Both binned-scatter plots control for firm size (log employment) and industry fixed effects.

Beyond access to better suppliers, SOEs also differ from POEs in terms of supplier sorting. Figure 13 replots the pattern discussed in Section 4.4, namely that good suppliers tend to source from other good suppliers, by grouping firms according to ownership type. A notable finding is that supplier sorting is particularly pronounced among POEs, whereas among SOEs, the quality of matched suppliers remains relatively stable regardless of the firms own supplier effects. This suggests that procurement by SOEs is less sensitive to supplier strength than procurement by comparable private firms, possibly reflecting non-market objectives or constraints such as directed procurement, relationship-based sourcing, regional employment goals, soft budget constraints that mute cost and quality pressure, or bureaucratic procurement rules. This pattern holds throughout the supplier distribution.

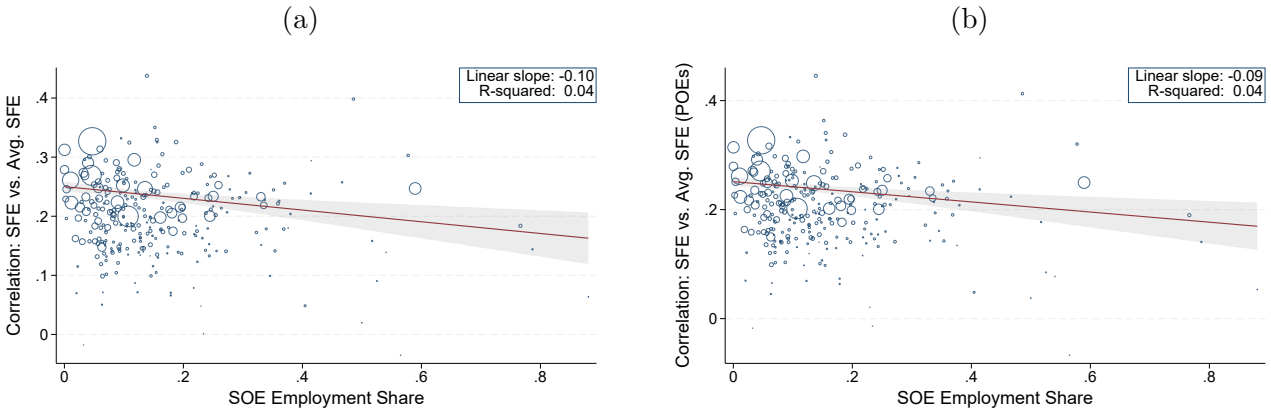
SOEs not only exhibit systematic differences in supplier access and sorting, but also generate spillover effects on local sorting patterns. Figure 14 shows a negative association between city-level SOE presence and market-based inter-firm sorting. This pattern is evident both for sorting among all firms and for sorting restricted to private firms (POEs). This evidence is consistent with the mechanism in Brandt et al. (2025), whereby local SOE dominance raises barriers faced by POEs. Our results suggest that these distortions reduce aggregate efficiency by weakening the sorting channel.

Figure 13



Notes: Figure 13a shows the relationship between firms' supplier fixed effects and the average supplier fixed effect of their suppliers. Figure 13b plots the relationship between firms' supplier fixed effects and the supplier fixed effects of their suppliers at the 10th, 50th, and 90th percentiles (P10, P50, and P90), where we restrict the sample to firms with at least 10 suppliers. We first demean both supplier and customer fixed effects. Both binned-scatter plots control for firm size (log employment) and industry fixed effects.

Figure 14: Sorting and State Ownership across Cities



Notes: Each dot represents a city and is weighted by the number of firms. Figure 14a plots inter-firm sorting against the city-level SOE employment share. Figure 14b plots the same relationship, with sorting computed using private firms (POEs) only. Sorting is measured as the correlation between a firms supplier fixed effect and the average supplier fixed effect of its suppliers.

Taken together, these patterns indicate ownership-linked wedges in the intermediate input market and constitute evidence of misallocation. In an efficient supply chain with complementarities, we expect positive assortative matching between firm capability and partner capability. We find that SOEs upstream procurement is less assortative, consistent with muted quality discipline and the persistence of low-productivity suppliers. At the same time, SOEs as customers systematically obtain higher-capability suppliers, consistent with preferential access that reallocates high-quality intermediate inputs toward SOEs and away from alternative (potentially

higher marginal product) users.

6 Conclusion

This paper treats firm size as an equilibrium object emerging from bilateral relationships in a production network rather than a purely intrinsic attribute. By adapting an AKM-style two-way fixed-effects framework to Chinese VAT invoice records, we decompose industry-demeaned bilateral sales into supplier components, buyer components, and match-specific residuals. The many-to-many nature of these trading relationships ensures strong network connectedness, enabling the separate identification of supplier and buyer effects within a single cross-section and permitting bias-corrected second-moment decompositions.

The core empirical message is one of sharp asymmetry. Dispersion in firm-to-firm sales is driven primarily by supplier-side heterogeneity, while buyer-side effects play a more limited role, and the implied sorting component is small. This supports an interpretation where systematic variation operates through stable, one-dimensional firm "types" rather than strong match-specific complementarities. These estimated components are persistent and relate closely to firm size and productivity, offering a characterization of network structure that moves beyond degree-based notions of importance by distinguishing the intensive margin of transaction scale from the extensive margin of link formation.

Applying these estimates to firm ownership reveals a distinct wedge in the Chinese production network. SOEs exhibit weaker supplier-quality sorting in their procurement processes, meaning their input quality is less responsive to their own capabilities. However, conditional on their own capability, SOEs systematically secure higher-quality suppliers. This pattern suggests a form of network misallocation distinct from traditional factor misallocation. Ownership-linked constraints and advantages distort the matching process, muting market discipline for SOEs while reallocating high-quality supplier capacity toward them and away from potentially more efficient private-sector users.

These results provide a scalable framework for disciplining models of production networks and firm distortions. For policymakers, the findings suggest that productivity gains may depend not only on the reallocation of capital and labor but also on removing the institutional wedges that govern how firms match and scale within input markets.

References

Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 1999, *67* (2), 251–333.

Andrews, M. J., L. Gill, T. Schank, and R. Upward, "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias?," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2008, *171* (3), 673–697.

Atalay, Enghin, Ali Hortacsu, James Roberts, and Chad Syverson, "Network structure of production," *Proceedings of the National Academy of Sciences*, 2011, *108* (13), 5199–5202.

Berkowitz, Daniel, Hong Ma, and Shuichiro Nishioka, "Recasting the iron rice bowl: The reform of China's state-owned enterprises," *Review of Economics and Statistics*, 2017, *99* (4), 735–747.

Bernard, Andrew B and Yuan Zi, "Sparse production networks," Technical Report, National Bureau of Economic Research 2022.

_ , Andreas Moxnes, and Yukiko U Saito, "Production networks, geography, and firm performance," *Journal of Political Economy*, 2019, *127* (2), 639–688.

Bernard, Andrew B., Emmanuel Dhyne, Glenn Magerman, Kalina Manova, and Andreas Moxnes, "The Origins of Firm Heterogeneity: A Production Network Approach," *Journal of Political Economy*, 2022, *130* (7), 1765–1804.

Bloom, Nicholas and John Van Reenen, "Measuring and explaining management practices across firms and countries," *The Quarterly Journal of Economics*, 2007, *122* (4), 1351–1408.

_ , Fatih Guvenen, Benjamin S. Smith, Jae Song, and Till von Wachter, "The Disappearing Large-Firm Wage Premium," *AEA Papers and Proceedings*, 2018, *108*, 317–322.

Brandt, Loren, Gueorgui Kambourov, and Kjetil Storesletten, "Barriers to Entry and Regional Economic Growth in China," *The Review of Economic Studies*, 2025, *93* (1), 286–326.

Card, David, Jörg Heining, and Patrick Kline, "Workplace heterogeneity and the rise of West German wage inequality," *The Quarterly Journal of Economics*, 2013, *128* (3), 967–1015.

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

Hopenhayn, Hugo A, "Entry, exit, and firm dynamics in long run equilibrium," *Econometrica*, 1992, pp. 1127–1150.

Hsieh, Chang-Tai and Peter J Klenow, "Misallocation and manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 2009, *124* (4), 1403–1448.

_ **and Zheng Michael Song**, "Grasp the Large, Let Go of the Small: The Transformation of the State Sector in China," *Brookings Papers on Economic Activity*, 2015.

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten, "Leave-Out Estimation of Variance Components," *Econometrica*, 2020, *88* (5), 1859–1898.

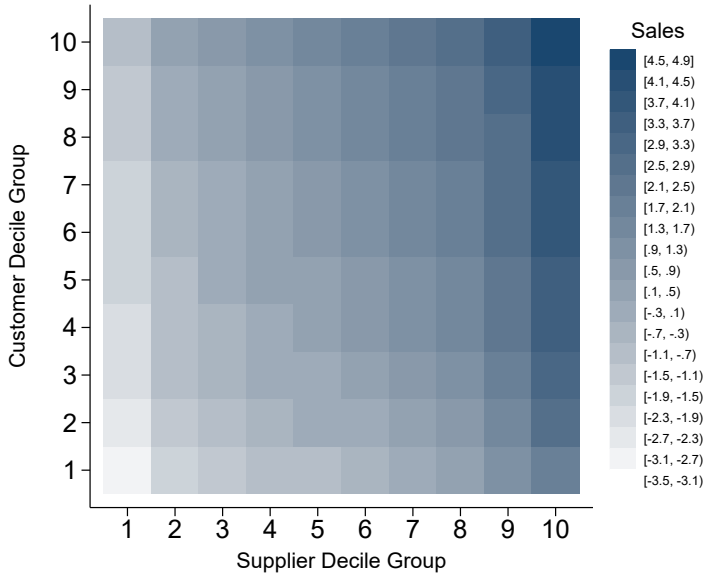
Melitz, Marc J, "The impact of trade on intra-industry reallocations and aggregate industry productivity," *Econometrica*, 2003, *71* (6), 1695–1725.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti, "Growing like china," *American economic review*, 2011, *101* (1), 196–233.

APPENDICES FOR ONLINE PUBLICATION

I Details of Empirical Analysis

Figure A-1: Average Log Sales by Supplier-Customer Decile Pairs



Notes: This figure illustrates the average log firm-to-firm sales across the supplier and customer decile groups. The firm-to-firm sales are demeaned within supplier-customer industry-pairs.

Figure A-1 presents a heatmap of $\ln \bar{m}_{ij}$, showing that the average log of firm-to-firm sales increases with both a supplier's average sales to its customers and a customer's average purchases from its suppliers.

Table A-1: Test Set Regressions

	log (Firm-to-Firm Sales)					
	(1)	(2)	(3)	(4)	(5)	(6)
Supplier FE	0.963***	1.012***	0.966***	1.019***	0.968***	1.021***
	(.000)	(.000)	(.000)	(.000)	(.001)	(.001)
	[5855.547]	[2448.877]	[4174.024]	[1700.164]	[1875.213]	[781.724]
Customer FE	0.792***	0.739***	0.808***	0.751***	0.816***	0.759***
	(.000)	(.001)	(.001)	(.001)	(.001)	(.002)
	[1921.759]	[1331.562]	[1379.095]	[939.514]	[620.680]	[429.583]
Supplier × Customer FE		0.037***		0.039***		0.040***
		(.000)		(.000)		(.001)
		[138.288]		[101.571]		[46.463]
Constant	-0.232***	-0.304***	-0.220***	-0.299***	-0.205***	-0.281***
	(.001)	(.001)	(.001)	(.000)	(.002)	(.003)
	[-346.240]	[-356.082]	[-225.895]	[-237.954]	[-96.634]	[-104.169]
R^2	0.4685	0.4688	0.4713	0.4716	0.4733	0.4736
Observations	46,376,854	46,376,854	23,282,060	23,282,060	4,668,648	4,668,648
Test Share	20%	20%	10%	10%	2%	2%

Standard errors in parentheses, and t statistics in square brackets.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table reports the regression results by regressing firm-to-firm sales from the test set on the fixed effects estimated from the training set. The firm-to-firm sales are demeaned at the supplier-customer industry-pair level.

Table A-2: Firm-to-Firm Sales Decomposition by Industry

		Variance Share					Sample Size		
Industry	$\text{Var}(\ln m_{ij})$	ψ_i	θ_j	ω_{ij}	$\frac{\text{Var}(\psi_i)}{\text{Var}(\psi_i+\theta_j)}$	ρ_{ψ_i,θ_j}	Observations	# Suppliers	# Customers
Accommodation & Catering	1.85	19.18%	16.92%	80.32%	0.51	0.04	18,711,076	42,365	2,592,556
Education	3.41	43.06%	10.71%	55.86%	0.82	-0.03	114,817	1,874	95,346
Printing Manufacturing	4.41	25.08%	9.17%	65.44%	0.77	-0.05	1,463,917	45,082	740,809
Wood & Fur Manufacturing	4.41	35.24%	8.97%	53.75%	0.81	-0.02	3,271,300	128,053	1,256,145
Finance	4.42	29.98%	10.32%	55.74%	0.67	0.13	5,240,027	35,611	2,565,404
Non-metallic Mineral Manufacturing	4.47	28.31%	7.97%	58.33%	0.74	0.06	2,783,111	100,683	947,663
Textile Manufacturing	4.56	30.43%	9.85%	56.89%	0.75	0.00	4,716,501	197,075	929,464
Health & Social Work	4.76	47.42%	9.81%	49.54%	0.92	-0.13	21,929	872	21,172
Residential Services, Repairs & Other Services	4.82	43.67%	9.04%	49.83%	0.83	-0.00	2,706,616	73,002	1,409,455
General Equipment Manufacturing	4.85	28.01%	7.82%	59.59%	0.72	0.10	6,630,631	220,587	1,651,996
Instrument Manufacturing	4.93	28.07%	7.61%	59.20%	0.69	0.16	1,254,300	23,461	561,497
Metal Manufacturing	5.06	36.07%	7.38%	52.97%	0.80	0.04	4,889,968	169,371	1,660,028
Chemical Manufacturing	5.08	30.31%	8.27%	56.62%	0.75	0.05	9,162,568	207,909	2,216,603
Specialized Equipment Manufacturing	5.08	27.02%	7.81%	60.50%	0.71	0.11	3,751,273	126,294	1,382,955
Other Manufacturing	5.12	37.16%	7.72%	51.96%	0.81	0.03	1,605,819	59,389	964,061
Smelting Manufacturing	5.49	34.13%	7.14%	52.25%	0.76	0.11	1,482,273	29,274	699,219
Water Conservancy, Environment & Public Facilities Mgmt.	5.52	46.35%	7.23%	43.83%	0.84	0.05	450,505	24,902	322,845
Scientific Research & Technical Services	5.58	43.00%	7.19%	47.91%	0.82	0.06	19,029,068	584,746	3,772,287
Electrical Manufacturing	5.64	29.99%	7.26%	56.06%	0.74	0.11	4,774,555	116,671	1,446,267
Agricultural Food Manufacturing	5.75	30.80%	9.45%	56.63%	0.77	-0.01	1,523,651	52,150	556,867
Others	5.91	51.64%	7.12%	43.61%	0.91	-0.04	3,500,117	46,953	1,756,482
Computer Manufacturing	5.93	29.05%	7.28%	58.58%	0.75	0.09	2,880,229	77,385	992,953
Leasing & Business Services	5.94	65.49%	6.08%	32.13%	0.93	-0.03	22,083,611	417,000	3,724,195
Real Estate	5.97	55.54%	6.81%	41.13%	0.92	-0.05	2,752,217	82,122	1,258,964
Electricity, Heat, Gas and Water Production & Supply	5.98	31.00%	9.11%	57.30%	0.82	-0.07	2,377,553	42,205	1,422,759
Construction	6.01	43.24%	6.09%	45.96%	0.82	0.11	4,264,048	440,191	1,226,919
Culture, Sports & Entertainment	6.03	43.16%	7.56%	46.74%	0.81	0.08	845,093	38,592	392,805
Information Transmission, Software & IT Services	6.10	46.95%	7.11%	43.45%	0.81	0.11	9,882,554	223,757	3,209,863
Wholesale & Retail Trade	6.11	48.76%	7.30%	42.74%	0.88	-0.01	84,287,071	2,444,826	6,490,566
Entertainment Equipment Manufacturing	6.27	44.11%	7.02%	45.80%	0.84	0.04	711,261	36,150	412,755
Agriculture, Forestry, Animal Husbandry & Fisheries	6.79	50.75%	7.36%	41.71%	0.91	-0.05	274,270	22,038	201,250
Transportation Equipment Manufacturing	6.95	26.06%	7.32%	59.61%	0.72	0.10	1,441,197	63,978	542,629
Transportation, Warehousing & Postal Services	7.19	39.96%	6.43%	49.28%	0.81	0.08	8,743,079	240,879	2,466,728
Mining	7.60	41.73%	6.48%	47.95%	0.87	-0.00	227,157	18,830	139,095

Notes: The table presents the decomposition results of firm-to-firm sales by industries. Each transaction is classified according to the supplier's industry, and the variance of both supplier and customer effects within industries is computed, based on the decomposition results reported in Table 5.