7-2024

Essays on estimating firm-level production function with spatial dependence

Bo Lin NG
Singapore Management University, bolin.ng.2018@phdecons.smu.edu.sg

Essays on Estimating Firm-Level Production Functions with Spatial Dependence

By

Ng Bo Lin

A DISSERTATION

In

ECONOMICS

Presented to the Singapore Management University in Partial Fulfilment

of the Requirements for the Degree of PhD in Economics

2024

Supervisor of Dissertation

PhD in Economics, Programme Director

Essays on Estimating Firm-Level Production Functions with Spatial Dependence

By

Ng Bo Lin

Submitted to the School of Economics in partial fulfilment of the requirements for the Degree of Philosophy in Economics

Dissertation Committee:

Chang Pao-li (Supervisor/Chair)

Associate Professor of Economics

Singapore Management University

Yang Zhenlin

Professor of Economics

Singapore Management University

Fan Haichao

Professor

School of Economics

Fudan University

Ma Lin

Assistant Professor of Economics

Singapore Management University

I hereby declare that this dissertation is my original work and it has been written by me in its entirety.

I have duly acknowledged all the sources of information which have been used in this dissertation.

This dissertation has also not been submitted for any degree in any university previously.

Ng Bo Lin

21 July 2024

Acknowledgements

To my supervisor, Professor Pao-Li Chang, I am eternally grateful for her unwavering support towards the completion of my dissertation. Without her guidance and patience, this would not have been possible.

To my co-authors Professors Pao-Li Chang, Zhenlin Yang and Ryo Makioka, I would like to thank them for their insightful comments and helpful feedback during the research process.

To my dissertation committee members Professors Pao-Li Chang, Zhenlin Yang, Lin Ma and Haichao Fan, I would like to thank them for taking time out of their extremely busy schedules to attend my dissertation defense, and providing valuable feedback which will undoubtedly improve the quality of the paper. Special thanks to Professor Haichao Fan for attending the defense at a short notice.

To Qiu Ling Thor and Amelia Tan at the School of Economics, I would like to thank them for providing excellent technical support.

To Paulo Ramos, Rohan Ray, Shenxi Song, Renjing Chen and Fan Zheng, sincerest gratitude for your patience and guidance throughout my stay. The journey was made much more bearable with your presence.

And finally, to my family and friends, for their unconditional support throughout the journey.

Abstract

This dissertation consists of two chapters on the estimation of firm-level production functions when spatial effects are present.

In joint work with Pao-Li Chang, chapter 1 focuses on the impacts of global value chains (GVC) on the firm-level outcomes of Singapore. First, we quantify Singapore's participation in global value chains (GVC) using the export decomposition framework of Borin and Mancini (2019), before using these indicators to analyse how GVC participation affects sectoral-level value-added and employment. We find that gross exports and foreign final demand have become more important for Singapore's value-added, largely driven by the Services sectors. We then use the GVC indicators to evaluate the impact of GVC participation on firm-level outcomes, including total factor productivity, labor productivity and employment. We find that firms tend to be more productive in sectors with stronger backward linkages (measured by the proportion of foreign content embedded in the production of a sector's GVC-related exports). On the other hand, firms tend to be less productive in sectors with stronger forward linkages (measured by the proportion of domestic content embedded in a sector's GVC-related exports). Our analyses provide policymakers with a better understanding of the impact of shifts in GVC on firm-level and sector-level performance measures.

In joint work with Pao-Li Chang, Ryo Makioka, and Zhenlin Yang, Chapter 2 proposes a three-stage efficient GMM estimation algorithm for estimating firm-level production functions given spatial dependence across firms due to supplier-customer relationships, sharing of input markets, or knowledge spillover. The procedure builds on Ackerberg, Caves and Frazer (2015) and Wooldridge (2009), but in addition, allows the productivity process to depend on the lagged output levels and lagged input usages of related firms, and spatially correlated productivity shocks across firms, where the set of related firms can differ across the three dimensions of spatial dependence. We establish the asymptotic properties of the proposed estimator, and conduct Monte Carlo simulations to evaluate the finite sample performance of the estimator. The proposed estimator is consistent under DGPs with or without spatial dependence, and with strong/weak or positive/negative spatial

dependence. In contrast, the conventional estimators lead to biased estimates of the production function parameters if the underlying DGPs have spatial dependence structure, and the magnitudes of the bias increase with the strength of spatial dependence in the underlying DGPs. We apply the proposed estimation algorithm to a Japanese firm-to-firm dataset during the period 2009-2018. We find significant and positive spatial coefficients in the Japanese firm-level productivity process via all three channels proposed above.

Contents

Chapter 1

Impacts of Global Value Chains on Singapore's Firm-level Employment and Productivity

1.1 Introduction

International trade plays a crucial role in driving economic growth for Singapore, as a small and open economy. Total merchandise trade and services trade increased to $1,206 billion and $836 billion, respectively, in 2023 from $884 billion and $389 billion, respectively, in 2015 (Ministry of Trade and Industry, 2015 and 2023). During this period, the Services sectors accounted for more than 72% of Singapore's GDP in 2023, up from 70.0% in 2015. The increasing importance of trade in services is taking place in an era when production processes become increasingly fragmented in stages yet integrated across countries, with falling transportation and communication costs.

In this paper, we quantity Singapore's participation in global value chains and analyse how GVC participation affects sectoral value-added and employment. In particular, we use the accounting framework of Borin and Mancini (2019) and adopt the formulas of Chang and Nguyen (2021, 2022a,b) to develop GVC indicators, which are then applied to the firm-level data of Singapore to infer the impact of GVC participation on firm-level outcomes.

1.2 GVC Participation of Singapore

To quantify Singapore's participation in global value chains (GVCs), we apply the gross-export decomposition framework of Borin and Mancini (2017, 2019) to the intercountry input-output (ICIO) tables compiled by the OECD TiVA Database.[1]

We measure Singapore's participation in the GVCs on an industry-year basis for both Manufacturing and Services sectors. In particular, we map the International Standard Industry Classification (ISIC) used by the TiVA ICIO tables to the Singapore Standard Industry Classification (SSIC), and further group the sectors into larger clusters at which the firm-level observations are associated with (regarding a firm's industry cluster affiliation). The concordance is provided in Table 1.1.

Table 1.1: SSIC versus TiVA ICIO sector concordance

Sector Code	Description	SSIC 2020	TiVA ICIO
1	Electronics	26	D26
2	Chemicals & Biomedical Manufacturing	19:23	D19:D23
3	Precision Engineering	24:25, 27:28	D24:D25, D27:D28
4	Transport Engineering	29:30	D29:D30
5	General Manufacturing Industries	10:18, 31:32, 33*	D10:D18, D31:D33
6	Construction	41:43	D41:D43
7	Wholesale Trade and Retail Trade	45*, 46:47	D45:D47
8	Transportation & Storage	49:53	D49:D53
9	Accommodation and Food & Beverage Services	55:56	D55:D56
10	Information & Communications	58:63	D58:D63
11	Finance & Insurance	64:66	D64:D66
12	Real Estate	68	D68
13	Professional Services	69:75	D69:D75
14	Administrative & Support Services	77:82	D77:D82
15	Others	01:03, 05:06*, 07*, 08:09, 35:38, 39*, 84*, 85:96, 97:98*	D01:D03, D05:D06, D07:D09, D35:D39, D84, D85:D88, D90:D96, D97:D98

Note: The column "Sector Code" lists the cluster of sectors at which the firm-level data are provided. The column "SSIC 2020" lists the corresponding SSIC sectors under each cluster. The column "TiVA ICIO" lists the corresponding TiVA ICIO sectors we map to the SSIC sectors under each cluster. Note that the firm-level dataset does not have observations in certain sectors, in which case, the sectors are marked with an asterisk * sign. This includes, e.g., D84: Public administration and defence; compulsory social security. For Sector 7, the original firm-level data can be further distinguished by the sub-sectors of Wholesale Trade versus Retail Trade. The two sub-sectors are combined, because the TiVA ICIO data are only available at the more aggregate level of D45:D47.

[1] In particular, we implement the decomposition using the "icio" STATA module developed by Belotti, Borin and Mancini (2021).

We measure a country-sector's involvement/position in the GVCs by the following four indices:

$$\text{BackwardLinkage}_{ci} = \frac{\sum_{p\neq c} FC_{ci,p}}{E_{ci,*}}, \tag{1.1}$$

$$\text{ForwardLinkage}_{ci} = \frac{\sum_{p\neq c}(DC_{ci,p} - TT_{ci,p})}{E_{ci,*}}, \tag{1.2}$$

$$\text{GVC}_{ci} = \text{BackwardLinkage}_{ci} + \text{ForwardLinkage}_{ci}, \tag{1.3}$$

$$\text{Downstreamness}_{ci} = \text{BackwardLinkage}_{ci}/\text{GVC}_{ci}. \tag{1.4}$$

where $E_{ci,*}, FC_{ci,p}, DC_{ci,p}$ and $TT_{ci,p}$ are respectively the gross exports of country c in sector i, the foreign and domestic contents embodied in the gross exports of country c to country p in sector i, and the "traditional trade" of country c to country p in sector i, respectively. In particular, traditional trade ($TT_{ci,p}$) corresponds to the domestic content of country c embedded in its exports of goods in sector i that is directly absorbed by importing country p, and hence regarded as trade contents not associated with GVC activities.

The backward linkage indicator measures the proportion of foreign content embedded in a country-sector's gross exports. The larger the index, the more intensive the sector is involved in the GVCs via the usage of foreign contents in its production for exports. On the other hand, the forward linkage indicator (domestic content net of traditional trade) measures the proportion of domestic content embedded in a country's gross exports that is not directly consumed (traditional trade) but further processed and exported by the direct trading partner p. Thus, a country-sector with a larger index of forward linkages participates more in the GVCs via downstream connections, in the sense that its domestic contents are more heavily used by the other economies as intermediate inputs for production of gross exports. The GVC indicator, by combining the backward and forward linkages, measures the total GVC-related trade as a proportion of a country-sector's gross exports. This is equivalent to the proportion of total export values that are not traditional trade, but contents that have travelled across country borders more than once and hence possibly associated with GVC activities. The downstreamness index in turn measures the relative importance of backward linkages in a country's GVC activities.

1.2.1 Manufacturing

In Tables 1.2–1.4 and Figure 1.1, we report the GVC indicators for the Manufacturing sectors during the period 1995–2020. As indicated by Table 1.2 and Figure 1.1, the backward linkage indicator for most sectors of Manufacturing fluctuated during the period of study, but remained in similar orders of magnitude. The Chemicals sector has experienced a larger extent of structural changes: its backward linkages grew by 45 percent (from 0.471 to 0.683) from 1995 to 2005, but fell by 23.6 percent (from 0.683 to 0.522) subsequently from 2005 to 2020. Although this phenomenon could partially be attributed to the supply chain disruptions during the Covid pandemic, the downward trend in the sector's backward linkages had started since 2015.

Table 1.2: GVC (Backward Linkage) – Manufacturing

Sub-sector	1995	2000	2005	2010	2015	2020
Electronics	0.535	0.547	0.425	0.548	0.562	0.510
Chemicals	0.471	0.516	0.683	0.642	0.607	0.522
Precision Eng.	0.498	0.516	0.549	0.580	0.511	0.509
Transport Eng.	0.454	0.445	0.482	0.471	0.485	0.499
General Manufacturing	0.366	0.406	0.431	0.414	0.458	0.423

Similarly, the forward linkages of most sectors remained stable during 1995–2020, with the exception of the Electronics sector and the Chemicals sector. The Electronics sector tended to increase in its forward linkages in the period of 2000–2015, while the Chemicals sector moved in the opposite direction. Interestingly, the Chemicals sector witnessed a slight increase in forward linkages in 2020, contrary to the other sectors. Overall, Singapore's manufacturing exports consist largely of foreign contents (in the range of 0.4–0.6 as a proportion of gross exports as summarized in Table 1.2), in contrast with a much smaller proportion attributable to domestic contents that are further processed and exported by other economies (less than 0.2 as a proportion of gross exports as indicated in Table 1.3).

As a result, the GVC participation of Singapore in manufacturing sectors are dominated by their backward linkages. We see a slow growth in the GVC participation across most manufacturing sectors, but some reversal in the trend for the Electronics and the Chemicals sectors, with the

Table 1.3: GVC (Forward Linkage) – Manufacturing

Sub-sector	1995	2000	2005	2010	2015	2020
Electronics	0.102	0.127	0.190	0.162	0.149	0.149
Chemicals	0.145	0.163	0.096	0.095	0.099	0.123
Precision Eng.	0.083	0.103	0.087	0.094	0.094	0.089
Transport Eng.	0.069	0.082	0.074	0.073	0.077	0.067
General Manufacturing	0.061	0.071	0.064	0.070	0.064	0.052

reversal taking place earlier around 2010 for the Chemicals sector and later around 2015 for the Electronics sector. Combining the relative trend of the backward and forward linkages, the Chemicals sector has become more downstream (its backward linkages strengthened relative to its forward linkages), while the Electronics sector has moved more upstream during the period.

Table 1.4: GVC (Backward and Forward Linkage) – Manufacturing

Sub-sector	1995	2000	2005	2010	2015	2020
Electronics	0.637	0.675	0.615	0.710	0.711	0.659
Chemicals	0.616	0.679	0.778	0.737	0.707	0.645
Precision Eng.	0.581	0.619	0.636	0.674	0.606	0.598
Transport Eng.	0.523	0.526	0.556	0.544	0.562	0.567
General Manufacturing	0.428	0.476	0.495	0.484	0.522	0.475

Table 1.5: Downstreamness – Manufacturing

Sub-sector	1995	2000	2005	2010	2015	2020
Electronics	0.840	0.811	0.691	0.772	0.790	0.774
Chemicals	0.764	0.760	0.877	0.871	0.859	0.809
Precision Eng.	0.857	0.833	0.863	0.861	0.844	0.850
Transport Eng.	0.867	0.845	0.866	0.866	0.862	0.881
General Manufacturing	0.857	0.852	0.871	0.856	0.877	0.891

1.2.2 Services

In Tables 1.6–1.8 and Figure 1.2, we tabulate the same indicators for the Services sectors. From 1995 to 2020, the backward linkages for most Services sectors generally increased. Construction was a clear exception, whose GVC participation by backward linkages decreased substantially.

Notably, supply chain disruptions in 2020 due to the global pandemic did not cause significant drags to the pre-existing trend in majority of the sectors. This could be attributed to the nature of these sectors, as trade in some services could be less sensitive to travel restrictions than trade in goods if they could be provided remotely/contactlessly. This pattern is particularly salient for the Information and Communication, and the Finance and Insurance sectors, whose backward linkages continued to strengthen from 2015 to 2020.

Table 1.6: GVC (Backward Linkage) – Services

Sub-sector	1995	2000	2005	2010	2015	2020
Construction	0.473	0.445	0.325	0.389	0.346	0.285
Wholesale and Retail Trade	0.245	0.306	0.332	0.327	0.403	0.374
Transportation and Storage	0.422	0.487	0.589	0.556	0.596	0.626
Accommodation and F&B	0.236	0.246	0.279	0.316	0.296	0.296
Information and Communication	0.283	0.305	0.341	0.481	0.473	0.519
Finance and Insurance	0.107	0.151	0.198	0.258	0.287	0.309
Real Estate	0.115	0.143	0.176	0.165	0.159	0.153
Professional Services	0.219	0.241	0.285	0.291	0.353	0.351
Administrative and Support Services	0.197	0.207	0.318	0.262	0.254	0.229

On the other hand, changes in the forward linkages over 1995 to 2020 is much more nuanced. The Wholesale and Retail Trade, and the Information and Communication sectors exhibited a decreasing trend in their forward linkages. On the other hand, the forward linkages of the Construction, the Finance and Insurance, and the Administrative and Support Services sectors tended to increase during the period (although with some fluctuations across years).

Table 1.7: GVC (Forward Linkage) – Services

Sub-sector	1995	2000	2005	2010	2015	2020
Construction	0.096	0.130	0.160	0.150	0.168	0.168
Wholesale and Retail Trade	0.153	0.179	0.177	0.177	0.151	0.139
Transportation and Storage	0.110	0.120	0.098	0.111	0.100	0.089
Accommodation and F&B	0.000	0.001	0.001	0.001	0.001	0.003
Information and Communication	0.092	0.106	0.104	0.080	0.069	0.058
Finance and Insurance	0.140	0.196	0.195	0.207	0.215	0.186
Real Estate	0.073	0.112	0.115	0.076	0.089	0.142
Professional Services	0.128	0.142	0.172	0.156	0.134	0.124
Administrative and Support Services	0.152	0.189	0.159	0.208	0.220	0.268

Overall, the GVC indicator for most Services sectors (apart from Construction) increased from 1995 to 2020, owing to the strong increase in backward linkages. Specifically, the Finance and Insurance, the Information and Communication, and the Transportation and Storage sectors saw an increase of 100, 54, and 34 percent, respectively, in their proportions of GVC-related trade (relative to sectoral exports) over this period. This suggests that these Services sectors of Singapore are now far more integrated in the GVCs than two decades ago.

Table 1.8: GVC (Backward and Forward Linkage) – Services

Sub-sector	1995	2000	2005	2010	2015	2020
Construction	0.568	0.574	0.485	0.539	0.514	0.453
Wholesale and Retail Trade	0.398	0.486	0.510	0.504	0.554	0.513
Transportation and Storage	0.533	0.607	0.687	0.667	0.696	0.714
Accommodation and F&B	0.236	0.246	0.280	0.318	0.297	0.299
Information and Communication	0.375	0.411	0.445	0.561	0.542	0.577
Finance and Insurance	0.247	0.348	0.394	0.465	0.503	0.495
Real Estate	0.188	0.255	0.291	0.241	0.247	0.295
Professional Services	0.347	0.384	0.457	0.447	0.487	0.474
Administrative and Support Services	0.349	0.396	0.477	0.470	0.474	0.496

Given the dominance of the backward linkages in the Services sectors and the growing importance of the backward linkages in most sectors, it also implies that the Services sectors in Singapore have become more downstream. This is especially prominent in the three sectors (Finance and Insurance, Information and Communication, and Transportation and Storage) highlighted above. Construction is again a clear exception, whose reduced reliance on foreign contents has moved the sector increasingly more upstream over time.

In summary, Singapore's participation in the GVCs strengthened during the period of study, and particularly so in the Services sectors. The trend was interrupted by the US-China trade war and the global pandemic during the period of 2017–2020 for the Manufacturing sectors. On the other hand, some Services sectors were better insulated from the disruptions, and continued their upward trajectory of participation in the GVCs during this period. The Services sectors started being less involved in GVC in 1995 (ranging from 0.188 to 0.568) than the Manufacturing sectors (ranging from 0.428 to 0.637). By 2020, however, some Services sectors have overtaken the Manufacturing

Table 1.9: Downstreamness – Services

Sub-sector	1995	2000	2005	2010	2015	2020
Construction	0.832	0.774	0.671	0.722	0.674	0.630
Wholesale and Retail Trade	0.616	0.631	0.652	0.649	0.728	0.729
Transportation and Storage	0.793	0.802	0.857	0.833	0.856	0.877
Accommodation and F&B	0.999	0.997	0.997	0.996	0.996	0.990
Information and Communication	0.755	0.741	0.765	0.857	0.872	0.899
Finance and Insurance	0.434	0.435	0.504	0.554	0.572	0.624
Real Estate	0.610	0.562	0.605	0.686	0.642	0.517
Professional Services	0.632	0.629	0.624	0.651	0.725	0.739
Administrative and Support Services	0.565	0.522	0.665	0.558	0.537	0.461

sectors. For example, the Transportation and Storage sector (with a GVC intensity of 0.714) has exceeded Electronics (0.659) and Chemicals (0.645), and become the most GVC-intensive sector of the Singapore economy in 2020.

1.3 Impacts of GVC on Sectoral Value-Added and Employment

In Section 1.2, we have presented the general pattern of Singapore's involvement in the GVCs at the sector level, where we have measured backward/forward linkages based on foreign/domestic contents, which in turn include foreign/domestic value-added and foreign/domestic double-counted. The double-counted content corresponds to the part of gross export values that cross country borders more than once and hence is considered double counted from the world GDP's perspective. Nonetheless, it is still a meaningful component from the GVC's perspective, as GVC activities in nature involve production stages across borders and hence the likelihood that the same content is repeatedly embedded in gross exports each time it leaves a country's border.

In this section, we focus on the impact of GVC participation in generating the domestic value-added (and in particular, the domestic *labor* value-added) and labor employment for the Singapore economy. For this set of analysis, we follow the methodologies developed by Horvat, Webb and Yamano (2020). In particular, we calculate indicators that measure the amount of domestic labor value-added and employment created by a country-sector and embodied in gross exports (regardless of sectors of exports) or in foreign final demand (regardless of sectors of foreign final demand). This also differs conceptually from the analysis in Section 1.2, where the foreign/domestic contents imputed are by the country-sector of exports (yet regardless of the sectors of origin of content).

Relative to a large economy, a small and open economy (such as Singapore) will in general have a disproportionately larger proportion of their value-added and employment embodied in gross exports (or in foreign final demand) than in local demand, since external demands are likely to outweigh the local demand of a small economy.

1.3.1 Domestic Value-Added / Labor Value-Added / Employment Embodied in Foreign Final Demand

We use the following formulas to measure the domestic value-added embodied in foreign final demand (DVAFFD), and similarly the domestic labor value-added and labor employment embodied in foreign final demand (DLVAFFD and DLFFD) respectively for each country-sector of the k countries and n sectors:

$$DVAFFD = \hat{v} \times B \times FD, \tag{1.5}$$

$$DLVAFFD = \widehat{e} \times B \times FD, \tag{1.6}$$

$$DLFFD = \widehat{l} \times B \times FD, \tag{1.7}$$

where $\widehat{v}$ is the $(kn \times kn)$ diagonalized value-added coefficient matrix measuring the amount of value-added per dollar of gross outputs, $\widehat{e}$ is the $(kn \times kn)$ diagonalized compensation of employees coefficient matrix measuring the amount of labor value-added per dollar of gross outputs, $\widehat{l}$ is the $(kn \times kn)$ diagonalized labor coefficient matrix measuring the number of workers employed per dollar of gross outputs, B is the $(kn \times kn)$ global Leontief inverse matrix measuring the total input required of each country-sector per dollar of gross outputs in each country-sector, and FD is the $(kn \times k)$ global final demand matrix measuring the dollar value of goods produced by each country-sector consumed by each destination country as final demand. The resulting matrix in (1.5) is a $(kn \times k)$ matrix, where the ci, p-th element of the matrix provides the domestic value-added of country c, industry i, in meeting the final demand of country p. Similarly, the ci, p-th element of the matrix in (1.6) and (1.7) provides the domestic labor value-added and labor employment, respectively, of country c, industry i, in meeting the final demand of country p. By focusing on the rows involving country c, and summing across columns $p \neq c$, we obtain the domestic value-added of country c embodied in foreign final demand, and similarly for domestic labor value-added and labor employment.

1.3.2 Domestic Value-Added / Labor Value-Added / Employment Embodied in Gross Exports

We impute the domestic value-added (or domestic labor value-added and labor employment, alternatively) embodied in the gross exports of country c as follows:

$$
\begin{aligned}
DVAGE_c &= \sum_p \widehat{v}_{c,c} \times B_{c,c} \times \widehat{GE}_{c,p} \times u, && (1.8) \\
DLVAGE_c &= \sum_p \widehat{e}_{c,c} \times B_{c,c} \times \widehat{GE}_{c,p} \times u, && (1.9) \\
DLGE_c &= \sum_p \widehat{l}_{c,c} \times B_{c,c} \times \widehat{GE}_{c,p} \times u, && (1.10)
\end{aligned}
$$

where $\widehat{v}_{c,c}$ is the cc-th block matrix of the value-added coefficient matrix $\widehat{v}$, $\widehat{e}_{c,c}$ is the cc-th block matrix of the labor value-added coefficient matrix $\widehat{e}$, $\widehat{l}_{c,c}$ is the cc-th block matrix of the labor coefficient matrix $\widehat{l}$, $B_{c,c}$ is the cc-th block matrix of the global Leontief matrix B, $\widehat{GE}_{c,p}$ is the $n \times n$ matrix of diagonalized gross exports vector from country c to its partner p across sectors, and u is a $n \times 1$ unit vector. The resulting matrix in (1.8) is a $(n \times 1)$ matrix, where the i-th element of the vector provides the domestic value-added of country c, industry i, embodied in country c's total gross exports across sectors. Similarly, the i-th element of the matrix in (1.9) and (1.10), respectively, provides the domestic labor value-added and labor employment of country c, industry

i, embodied in country c's total gross exports.

1.3.3 Data on Value-Added / Labor Value-Added / Employment

The value-added coefficient for each country-sector in $\widehat{v}$ is measured by the ratio of value added and gross output (corresponding to the 'VALU' and 'OUTPUT' rows in the TiVA ICIO tables). The labor value-added coefficient for each country-sector in $\widehat{e}$ is obtained by multiplying the "compensation of employees to value-added ratio" at the country-sector level (obtained from the TiM dataset, available up to 2018), with the "value-added to gross output ratio" of the same country-sector (contained in $\widehat{v}$) to obtain the labor value-added to gross output ratio. The labor employment data are not comprehensively available across countries and sectors. To obtain such data for Singapore across sectors, we take the following approximation approach. First, we obtain the labor employment data for the manufacturing cluster as a whole and for the services sub-sectors from the Singapore Department of Statistics (DOS) website. We then use the annual firm-level data as documented in Section 1.4 below to calculate the distribution of labor employment across sectors, based on which the aggregate labor employment in manufacturing from DOS is then apportioned across its sub-sectors. The annual labor employment in Others is imputed as the sum of labor employment in industries classified under Sector Code 15 in Table 1.1.[2] Note that the resulting data on labor employment for Singapore across sectors are available only for the period 2009–2018.

[2]To compute (1.7) for the Singapore block, we first calculate $B \times FD$ and focus on the rows involving Singapore. Denote it $(B \times FFD)_{sg}$, which has a dimension of n sectors by p trading partners. As the data on $\widehat{l}$ for Singapore (denote it $\widehat{l}_{sg,sg}$) are available only in 15 clusters, while the matrix B is available in more disaggregated TiVA sectors, we construct a weighting matrix W of dimension $15 \times n$ that measures the contribution of each of the n TiVA sectors to each of the 15 Singapore sector clusters in terms of gross outputs. In turn, the expression in (1.7) for the Singapore block is implemented as $\widehat{l}_{sg,sg} \times W \times (B \times FFD)_{sg}$. The implementation of (1.10) for Singapore uses a similar interface matrix W in between $\widehat{l}_{c,c}$ and $B_{c,c}$.

1.3.4 Results

Figure 1.3 illustrates the magnitude of the Singapore value-added / labor value-added / labor employment embodied in its gross exports and in foreign final demand, respectively, relative to the economy's total value-added / labor value-added / labor employment, for the broad clusters of the Manufacturing, Services, and Others sectors. Tables 1.10–1.11 provide the statistics in detail for the beginning year 1995 and the end year 2018 of the sample period. Note that gross exports and foreign final demand have become more important as the source of demand for Singapore value-added over the years, although the changes are very gradual. Specifically, the ratio of DVAGE to total VA increased from 62% to 65%, as indicated by Table 1.10. Second, the contribution of the Services sectors toward value-added via gross exports/foreign final demand has grown (from 38% to 46% in terms of DVAGE), but that of the Manufacturing sectors has decreased (from 21% to 17% in terms of DVAGE). The contribution of the Services sectors was more than two times that of the Manufacturing sectors by 2018 in terms of value-added. The gap was wider and close to four times in terms of labor value-added in 2018 (44% for Services versus 11% for Manufacturing in terms of DVAGE), suggesting the relatively larger role of labor in creating value-added in the Services sectors than the Manufacturing sectors. Third, the labor employment embodied in gross exports relative to the economy's total employment is systematically smaller than the labor value-added counterpart. This in part is due to the large presence of the Others sector in economy-wide labor employment (accounting for $> 20\%$ of labor employment), but near zero participation in GVC of the sector in terms of employment. Fourth, the patterns in terms of foreign final demand are very similar to those in terms of gross exports. This suggests that domestic value-added embodied in gross exports of Singapore are predominantly absorbed abroad, and very little of them returns home to be absorbed at home. Fifth, the dynamics of the value-added follows those of the labor value-added closely. This implies an economic structure where physical capital and the labor force (and human capital) move in tandem in creating value-added.

Figures 1.4–1.5 provide further breakdown and distribution of the Singapore value-added / labor value-added / labor employment embodied in its gross exports and in foreign final demand, respectively, across sub-sectors. First, we note that the Electronics and the Chemicals sectors are much more capital intensive than the other manufacturing sectors, as indicated by the much larger gaps between value-added and labor value-added embodied in external demands in these two sectors (13% and 11% in terms of DVAGE, and 9% and 7% in terms of DLVAGE, respectively for the two sectors in 1995). The reverse is true for the Wholesale and Retail Trade (17% in terms of DVAGE versus 22% in terms of DLVAGE) and Professional Services (7% versus 12%). Second, the Electronics sector has seen a decline in its contribution toward value-added embodied in gross exports / foreign final demand (from 13% to 7%), and in terms of labor value-added and labor employment as well. In contrast, the Wholesale and Retail Trade, and the Finance and Insurance sectors have grown in their importance in creating value-added via serving foreign markets. While the larger role of the Wholesale and Retail Trade sector reflects mainly that of capital value-added (from 17% to 22% in terms of DVAGE), the expansion of the Finance and Insurance sector has been driven predominantly by that of labor value-added (from 10% to 16% in terms of DLVAGE). Third, the Wholesale and Retail Trade sector's much larger presence in terms of labor employment than in terms of labor value-added embodied in gross exports / foreign final demand, suggests weaker wage rates for workers employed in this sector than in the other sectors. The same observation applies to the sector of Accommodation and Food & Beverage Services. The opposite is true for the sectors of Finance & Insurance and Information & Communication, where the sectors' presence in terms of labor value-added is much stronger than in terms of labor employment.

While the above discussions focus on the relative importance across sectors in generating value-added and labor value-added via external demands, the ranking might reflect the relative size of a sector or the intensity of a sector in its involvement in exports. Figures 1.6 and 1.7 provide further characterization on the second dimension. In particular, the figures illustrate Singapore's domestic value-added / labor value-added / labor employment embodied in gross exports / foreign final demand, respectively, at the sector level, relative to the sector's total value-added / labor value-added / labor employment. The larger the fraction, the more intensively the sector relies on external demands in generating value-added and labor value-added. We find that the ranking of the manufacturing sub-sectors in terms of the intensive margins (as shown in Figures 1.6 and 1.7) is generally in line with the overall ranking observed in Figures 1.4 and 1.5. For example, in terms of labor value-added, Electronics and Chemicals and Biomedical Engineering lead the manufacturing cluster in terms of the aggregate labor value-added embodied in external demands; they also lead the cluster in terms of the intensive margin, i.e., how much the sectors' labor value-added is generated via external demands. In contrast, Transport & Storage leads the services cluster in terms of the intensive margin, although it ranks behind Wholesale and Retail Trade, and Finance & Insurance in terms of the overall size of the aggregate value-added or labor value-added generated via external demands (and behind Professional Services in terms of DLVAGE or DLVAFFD). In addition, certain service sub-sectors are highly intensive in serving foreign demand, and yet are overall relatively small in absolute size. This includes, e.g., Administrative & Support Services, and to a lesser extent, Accommodation and Food & Beverage Services.

Table 1.10: Singapore domestic value-added / labor value-added embodied in gross exports

1995

		VA		DVAGE		$\frac{DVAGE}{Total(VA)}$	$\frac{DVAGE}{VA}$	LVA		DLVAGE		$\frac{DLVAGE}{Total(LVA)}$	$\frac{DLVAGE}{LVA}$
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
M	Electronics	6565.41	8%	6756.10	13%	8%	1.03	1771.35	5%	1710.11	9%	5%	0.97
M	Chemicals & Biomedical Manufacturing	6066.28	7%	5775.38	11%	7%	0.95	1836.14	5%	1461.44	7%	4%	0.80
M	Precision Engineering	2415.11	3%	2005.07	4%	2%	0.83	1288.29	4%	1018.96	5%	3%	0.79
M	Transport Engineering	1586.62	2%	808.10	2%	1%	0.51	887.25	3%	428.76	2%	1%	0.48
M	General Manufacturing Indices	2155.01	3%	1611.84	3%	2%	0.75	1000.73	3%	737.78	4%	2%	0.74
S	Construction	4027.74	5%	128.59	0%	0%	0.03	1804.83	5%	55.24	0%	0%	0.03
S	Wholesale & Retail Trade	12736.70	16%	8798.07	17%	11%	0.69	6384.91	19%	4330.50	22%	13%	0.68
S	Transportation & Storage	7005.03	9%	5610.98	11%	7%	0.80	2525.30	7%	1902.10	10%	6%	0.75
S	Accommodation and F&B Services	1714.58	2%	1170.22	2%	1%	0.68	971.65	3%	627.08	3%	2%	0.65
S	Information & Communications	3167.49	4%	2223.85	4%	3%	0.70	1452.33	4%	1017.24	5%	3%	0.70
S	Finance & Insurance	9425.16	12%	6072.51	12%	7%	0.64	2960.44	9%	1881.39	10%	6%	0.64
S	Real Estate	7102.65	9%	2032.77	4%	2%	0.29	719.50	2%	203.38	1%	1%	0.28
S	Professional Services	5043.24	6%	3347.30	7%	4%	0.66	3597.85	11%	2337.54	12%	7%	0.65
S	Administrative & Support Services	2765.10	3%	1845.70	4%	2%	0.67	1421.54	4%	941.24	5%	3%	0.66
O	Others	10113.86	12%	2297.64	5%	3%	0.23	5445.63	16%	941.32	5%	3%	0.17
	Total	81889.98	100%	50484.12	100%	62%		34067.74	100%	19594.08	100%	58%	
	Manufacturing		23%		34%	21%			20%		27%	16%	
	Services		65%		62%	38%			64%		68%	39%	

2018

		VA		DVAGE		$\frac{DVAGE}{Total(VA)}$	$\frac{DVAGE}{VA}$	LVA		DLVAGE		$\frac{DLVAGE}{Total(LVA)}$	$\frac{DLVAGE}{LVA}$
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
M	Electronics	16350.32	5%	15417.18	7%	4%	0.94	4842.96	3%	4335.75	5%	3%	0.90
M	Chemicals & Biomedical Manufacturing	25703.87	7%	25779.44	11%	7%	1.00	3955.84	3%	3538.56	4%	2%	0.89
M	Precision Engineering	8943.15	3%	7162.57	3%	2%	0.80	4096.33	3%	3127.82	4%	2%	0.76
M	Transport Engineering	6242.46	2%	5190.29	2%	1%	0.83	3766.20	2%	3115.07	4%	2%	0.83
M	General Manufacturing Indices	8736.92	2%	6030.29	3%	2%	0.69	3516.16	2%	2334.48	3%	2%	0.66
S	Construction	16773.73	5%	913.72	0%	0%	0.05	10549.00	7%	550.08	1%	0%	0.05
S	Wholesale & Retail Trade	64592.56	18%	50969.52	22%	14%	0.79	23111.22	15%	17985.12	21%	12%	0.78
S	Transportation & Storage	23626.67	7%	25221.37	11%	7%	1.07	9298.68	6%	8297.51	10%	5%	0.89
S	Accommodation and F&B Services	7814.40	2%	4542.96	2%	1%	0.58	4254.16	3%	2359.27	3%	2%	0.55
S	Information & Communications	14813.66	4%	11754.76	5%	3%	0.79	8425.60	6%	6903.02	8%	5%	0.82
S	Finance & Insurance	44109.73	12%	31909.65	14%	9%	0.72	19937.60	13%	14267.58	16%	9%	0.72
S	Real Estate	29711.46	8%	7480.20	3%	2%	0.25	2721.57	2%	677.86	1%	0%	0.25
S	Professional Services	21085.25	6%	15647.98	7%	4%	0.74	14808.17	10%	10822.18	12%	7%	0.73
S	Administrative & Support Services	20524.15	6%	15640.49	7%	4%	0.76	6327.59	4%	4791.49	5%	3%	0.76
O	Others	44527.49	13%	7488.50	3%	2%	0.17	32518.14	21%	4177.45	5%	3%	0.13
	Total	353555.82	100%	231148.92	100%	65%		152129.22	100%	87283.24	100%	57%	
	Manufacturing		19%		26%	17%			13%		19%	11%	
	Services		69%		71%	46%			65%		76%	44%	

Note: Column (1) reports the VA in millions USD. Column (2) reports the distribution of the VA across sectors in percentage terms. Column (3) reports the DVAGE in millions USD. Column (4) reports the distribution of the DVAGE across sectors in percentage terms. The statistics correspond to those illustrated in Figure 1.4. Column (5) reports the ratio of DVAGE in each sector relative to the economy's total VA. The statistics for Manufacturing/Services/Others correspond to those illustrated in Figure 1.3. Column (6) reports the ratio of DVAGE relative to VA for each sector. The statistics correspond to those illustrated in Figure 1.6. Column (7) reports the LVA in millions USD. Column (8) reports the distribution of the LVA across sectors in percentage terms. Column (9) reports the DLVAGE in millions USD. Column (10) reports the distribution of the DLVAGE across sectors in percentage terms. The statistics correspond to those illustrated in Figure 1.4. Column (11) reports the ratio of DLVAGE in each sector relative to the economy's total LVA. The statistics for Manufacturing/Services/Others correspond to those illustrated in Figure 1.3. Column (12) reports the ratio of DLVAGE relative to LVA for each sector. The statistics correspond to those illustrated in Figure 1.6. The ratios in Columns (6) and (12) could be greater than one, due to statistical discrepancy, where a country-sector's output exported to all destinations (i.e., the sum across its usages in columns in the IO tables) is not identical to the country-sector's gross output reported (i.e., the sum across the inputs used by the sector and value-added of the sector in rows in the IO tables).

Table 1.11: Singapore domestic value-added / labor value-added embodied in foreign final demand

		1995											
		VA		DVAFFD		$\frac{DVAFFD}{Total(VA)}$	$\frac{DVAFFD}{VA}$	LVA		DLVAFFD		$\frac{DLVAFFD}{Total(LVA)}$	$\frac{DLVAFFD}{LVA}$
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
M	Electronics	6565.41	8%	6672.31	13%	8%	1.02	1771.35	5%	1688.9	9%	5%	0.95
	Chemicals & Biomedical Manufacturing	6066.28	7%	5710.85	11%	7%	0.94	1836.14	5%	1444.91	7%	4%	0.79
	Precision Engineering	2415.11	3%	1990.83	4%	2%	0.82	1288.29	4%	1012.14	5%	3%	0.79
	Transport Engineering	1586.62	2%	802.85	2%	1%	0.51	887.25	3%	425.87	2%	1%	0.48
	General Manufacturing Indices	2155.01	3%	1604.54	3%	2%	0.74	1000.73	3%	734.54	4%	2%	0.73
S	Construction	4027.74	5%	127.52	0%	0%	0.03	1804.83	5%	54.78	0%	0%	0.03
	Wholesale & Retail Trade	12736.70	16%	8711.78	17%	11%	0.68	6384.91	19%	4288.03	22%	13%	0.67
	Transportation & Storage	7005.03	9%	5554.65	11%	7%	0.79	2525.30	7%	1882.53	10%	6%	0.75
	Accommodation and F&B Services	1714.58	2%	1169.23	2%	1%	0.68	971.65	3%	626.54	3%	2%	0.64
	Information & Communications	3167.49	4%	2207.07	4%	3%	0.70	1452.33	4%	1009.79	5%	3%	0.70
	Finance & Insurance	9425.16	12%	6030.64	12%	7%	0.64	2960.44	9%	1868.41	10%	5%	0.63
	Real Estate	7102.65	9%	2019.62	4%	2%	0.28	719.50	2%	202.06	1%	1%	0.28
	Professional Services	5043.24	6%	3318.41	7%	4%	0.66	3597.85	11%	2317.36	12%	7%	0.64
	Administrative & Support Services	2765.10	3%	1829.99	4%	2%	0.66	1421.54	4%	933.23	5%	3%	0.66
O	Others	10113.86	12%	2283.20	5%	3%	0.23	5445.63	16%	935.81	5%	3%	0.17
	Total	81889.98	100%	50033.49	100%	61%		34067.74	100%	19424.90	100%	57%	
	Manufacturing		23%		34%	20%			20%		27%	16%	
	Services		65%		62%	38%			64%		68%	39%	

		2018											
		VA		DVAFFD		$\frac{DVAFFD}{Total(VA)}$	$\frac{DVAFFD}{VA}$	LVA		DLVAFFD		$\frac{DLVAFFD}{Total(LVA)}$	$\frac{DLVAFFD}{LVA}$
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
M	Electronics	16350.32	5%	14525.68	7%	4%	0.89	4842.96	3%	4302.51	5%	3%	0.89
	Chemicals & Biomedical Manufacturing	25703.87	7%	23174.23	11%	7%	0.90	3955.84	3%	3514.26	4%	2%	0.89
	Precision Engineering	8943.15	3%	6853.31	3%	2%	0.77	4096.33	3%	3109.67	4%	2%	0.76
	Transport Engineering	6242.46	2%	5087.93	2%	1%	0.82	3766.20	2%	3094.58	4%	2%	0.82
	General Manufacturing Indices	8736.92	2%	5794.76	3%	2%	0.66	3516.16	2%	2324.26	3%	2%	0.66
S	Construction	16773.73	5%	868.45	0%	0%	0.05	10549.00	7%	546.17	1%	0%	0.05
	Wholesale & Retail Trade	64592.56	18%	49905.52	23%	14%	0.77	23111.22	15%	17856.19	21%	12%	0.77
	Transportation & Storage	23626.67	7%	20922.75	10%	6%	0.89	9298.68	6%	8190.12	9%	5%	0.88
	Accommodation and F&B Services	7814.40	2%	4323.16	2%	1%	0.55	4254.16	3%	2353.53	3%	2%	0.55
	Information & Communications	14813.66	4%	11404.56	5%	3%	0.77	8425.60	6%	6869.68	8%	5%	0.82
	Finance & Insurance	44109.73	12%	31359.84	14%	9%	0.71	19937.60	13%	14174.65	16%	9%	0.71
	Real Estate	29711.46	8%	7354.77	3%	2%	0.25	2721.57	2%	673.70	1%	0%	0.25
	Professional Services	21085.25	6%	15303.90	7%	4%	0.73	14808.17	10%	10747.93	12%	7%	0.73
	Administrative & Support Services	20524.15	6%	15418.64	7%	4%	0.75	6327.59	4%	4753.57	5%	3%	0.75
O	Others	44527.49	13%	7251.94	3%	2%	0.16	32518.14	21%	4161.29	5%	3%	0.13
	Total	353555.82	100%	219549.44	100%	62%		152129.22	100%	86672.11	100%	57%	
	Manufacturing		19%		25%	16%			13%		19%	11%	
	Services		69%		71%	44%			65%		76%	43%	

Note: Column (1) reports the VA in millions USD. Column (2) reports the distribution of the VA across sectors in percentage terms. Column (3) reports the DVAFFD in millions USD. Column (4) reports the distribution of the DVAFFD across sectors in percentage terms. The statistics correspond to those illustrated in Figure 1.5. Column (5) reports the ratio of DVAFFD in each sector relative to the economy's total VA. The statistics for Manufacturing/Services/Others correspond to those illustrated in Figure 1.3. Column (6) reports the ratio of DVAFFD relative to VA for each sector. The statistics correspond to those illustrated in Figure 1.7. Column (7) reports the LVA in millions USD. Column (8) reports the distribution of the LVA across sectors in percentage terms. Column (9) reports the DLVAFFD in millions USD. Column (10) reports the distribution of the DLVAFFD across sectors in percentage terms. The statistics correspond to those illustrated in Figure 1.5. Column (11) reports the ratio of DLVAFFD in each sector relative to the economy's total LVA. The statistics for Manufacturing/Services/Others correspond to those illustrated in Figure 1.3. Column (12) reports the ratio of DLVAFFD relative to LVA for each sector. The statistics correspond to those illustrated in Figure 1.7. The ratios in Columns (6) and (12) could be greater than one, due to statistical discrepancy, where a country-sector's output exported to all destinations (i.e., the sum across its usages in columns in the IO tables) is not identical to the country-sector's gross output reported (i.e., the sum across the inputs used by the sector and value-added of the sector in rows in the IO tables).

1.3.5 Cross-Country Comparison

In Tables 1.12–1.17, we tabulate the same indicators for the regional economies, in comparison with Singapore, during the period of 1995–2015. Singapore has had the largest share (more than 50 percent) of domestic value-added (and labor value-added) embodied in gross exports and in foreign final demand. This signifies the importance of external demands in supporting her domestic economy.

Table 1.12: Domestic value-added embodied in gross exports

Year	SGP	CHN	JPN	KOR	TWN	MYS	THA
1995	0.586	0.150	0.084	0.196	0.313	0.454	0.324
2000	0.569	0.179	0.100	0.251	0.329	0.530	0.427
2005	0.645	0.246	0.125	0.251	0.349	0.540	0.423
2010	0.612	0.198	0.130	0.304	0.380	0.474	0.403
2015	0.620	0.169	0.151	0.303	0.395	0.411	0.430

Note: The figures are in proportion to the total domestic value-added of the economy.

Table 1.13: Domestic labor value-added embodied in gross exports

Year	SGP	CHN	JPN	KOR	TWN	MYS	THA
1995	0.573	0.132	0.084	0.190	0.306	0.377	0.257
2000	0.550	0.143	0.099	0.227	0.313	0.437	0.333
2005	0.604	0.199	0.122	0.222	0.323	0.427	0.352
2010	0.564	0.170	0.133	0.256	0.347	0.372	0.330
2015	0.571	0.152	0.150	0.264	0.351	0.335	0.347

Note: The figures are in proportion to the total domestic labor value-added of the economy.

Table 1.14: Domestic employment embodied in gross exports

Year	SGP	CHN	JPN	KOR
2010	0.398	0.174	0.119	0.239
2013	0.410	0.173	0.122	0.259
2015	0.430	0.171	0.135	0.253
2018	0.430	0.142	0.138	0.233

Note: The figures are in proportion to the total employment of the economy.

During this period, most economies (apart from China and Malaysia) experienced an increase in the value-added generated from meeting external demands. For example, Japan's domestic value-added (and labor value-added) driven by gross exports and foreign final demand increased by nearly 50 percent. In contrast, Malaysia's domestic value-added (and labor value-added) due to external demands increased during the early 2000's but fell substantially subsequently. Meanwhile, Taiwan witnessed a steady growth in the contribution of gross exports and foreign final demand to its local (labor) value-added. This is in general the case for Korea as well.

Table 1.15: Domestic value-added embodied in foreign final demand

Year	SGP	CHN	JPN	KOR	TWN	MYS	THA
1995	0.584	0.149	0.083	0.195	0.311	0.452	0.323
2000	0.566	0.177	0.098	0.249	0.327	0.528	0.426
2005	0.644	0.242	0.123	0.250	0.348	0.538	0.422
2010	0.610	0.194	0.128	0.303	0.378	0.472	0.402
2015	0.618	0.164	0.149	0.301	0.393	0.410	0.429

Note: The figures are in proportion to the total domestic value-added of the economy.

Table 1.16: Domestic labor value-added embodied in foreign final demand

Year	SGP	CHN	JPN	KOR	TWN	MYS	THA
1995	0.570	0.131	0.083	0.190	0.305	0.375	0.257
2000	0.547	0.142	0.097	0.226	0.311	0.436	0.333
2005	0.602	0.196	0.120	0.220	0.322	0.426	0.351
2010	0.562	0.167	0.131	0.255	0.345	0.370	0.329
2015	0.570	0.148	0.148	0.266	0.349	0.334	0.346

Note: The figures are in proportion to the total domestic labor value-added of the economy.

Table 1.17: Domestic employment embodied in foreign final demand

Year	SGP	CHN	JPN	KOR
2010	0.393	0.174	0.117	0.236
2013	0.405	0.168	0.119	0.256
2015	0.426	0.166	0.133	0.250
2018	0.427	0.137	0.136	0.230

Note: The figures are in proportion to the total employment of the economy.

We also calculate for the regional economies their domestic employment embodied in gross exports and foreign final demand. The data on employment across sectors are obtained from the

TiM dataset. The data are however not available for Taiwan, Malaysia, and Thailand. Recall that the Singapore employment data are not available from the TiM dataset, but is constructed using information (among others) on the distribution of labor employment across manufacturing sub-sectors observed in the firm-level data (whose sampling period starts only in 2009). Thus, Tables 1.14 and 1.17 present the comparison for a smaller set of regional economies and for a shorter period (2010–2018). As observed in Figure 1.3, Singapore labor employment embodied in gross exports and foreign final demand is systematically lower than its labor value-added counterparts. The difference is as large as 17 percentage points. We observe similar patterns for Japan and Korea, but the gap is much smaller at 1-2 percentage points. In contrast, China has higher fractions of labor employment embodied in gross exports and in foreign final demand than its labor value-added counterparts (although the difference is minor). To some extents, this reflects potential measurement errors in the case of Singapore, because we have to impute the Singapore labor allocation across manufacturing sub-sectors. But the qualitative ranking observed above suggests that the workers of Singapore who are involved in GVC and serving external demands create higher labor value-added (and receive higher compensations) in general than workers not involved in GVC. The reverse tends to be true for the Chinese economy, where workers involved in GVC are paid lower on average, suggesting a high labor-intensity but low labor-compensation nature in China's export production activities.

1.4 Firm-level Productivity Estimation

In this section, we characterize the firm-level performance of the Singapore economy. The firm-level data for the period 2009–2020 are made available by the Singapore Department of Statistics (DOS). The dataset consists of an unbalanced panel of 2,843,000 firm-year observations. Table 1.18 provides a snapshot of the number of firm-level observations across sectors and years. Table 1.19 summarizes the list of variables contained in the dataset and their corresponding summary statistics. This includes the basic firm-level financial statement variables, firm's age (by bins), firm's exporting status, firm's foreign ownership status, and the sector of the firm. We measure a firm's usage of intermediate inputs in a year by the difference between its revenue and value-added. The nominal financial statement variables are then deflated using annual sector-level GDP deflators. Note that less than 10% of the firm-year observations report year-end capital stock (fixed assets). In particular, the observations on (beginning-of-period) capital stocks, which is required for the total factor productivity estimation, are entirely missing for the initial years of 2009–2011.

We measure a firm's productivity by: (i) labor productivity in log (calculated as the firm's value-added in a year per worker, in log), and (ii) total factor productivity in log (estimated based on the Wooldridge (2009) method using value added as the left-hand-side variable and intermediate input as the proxy variable). The resulting measures of firm-year productivities, based on respectively the value-added per worker and the residual from the value-added production function are summarized in Table 1.19. The large attrition in the number of observations for total factor productivity is mainly due to the large number of missing observations on capital stock, which is required in the production function estimation.

In particular, the production function parameters are estimated sector by sector, and reported in Table 1.20. Both labor and capital coefficient estimates are positive and statistically significant at 10% level (except for the Real Estate sector, whose capital coefficient estimate is marginally significant). The sectors of Precision Engineering, Information & Communication, Finance & Insurance, and Professional Services tended to be the most labor (human capital) intensive in value-added creation. In comparison, Electronics, Chemicals & Biomedical Manufacturing, General Manufacturing Industries, and Accommodation and Food & Beverage Services were more capital-intensive in value-added creation. The capital coefficient estimates, nonetheless, need to be taken with a grain of salt, given the potential large measurement errors and reporting omissions of firm-level fixed assets.

Table 1.18: Number of firm-level observations across sectors and years

	Reference Year (2009–2020)												
Sector	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2009–2020
1	325	324	317	347	347	334	354	390	409	594	610	623	4,974
2	843	825	828	839	802	780	772	835	846	1,013	1,052	1,074	10,509
3	3,521	3,474	3,414	3,494	3,427	3,299	3,225	3,338	3,409	3,941	3,997	4,033	42,572
4	1,142	1,155	1,173	1,329	1,312	1,306	1,278	1,361	1,321	1,377	1,409	1,468	15,631
5	3,675	3,578	3,545	3,780	3,679	3,611	3,520	3,770	3,852	4,018	4,158	4,276	45,462
6	13,812	14,413	15,139	15,852	16,728	18,291	18,848	18,878	18,894	18,851	18,909	18,623	207,238
7	66,822	68,349	69,230	69,830	71,256	77,755	77,206	75,097	74,473	76,839	78,940	82,266	888,063
8	9,790	10,141	10,463	10,646	10,893	11,727	12,006	14,043	17,456	13,464	13,810	14,184	148,623
9	7,678	8,166	8,367	8,693	9,000	9,623	9,928	10,273	11,046	11,413	11,591	11,983	117,761
10	7,906	8,397	9,000	9,590	10,788	13,645	15,021	15,525	15,985	17,231	18,334	19,689	161,111
11	11,072	12,323	13,718	14,623	15,554	16,210	16,734	16,768	18,025	19,083	20,484	21,588	196,182
12	7,135	7,420	7,585	7,750	7,979	8,226	8,287	8,319	7,853	7,757	7,896	8,019	94,226
13	22,149	23,755	25,268	26,931	28,758	32,596	33,215	33,624	34,063	36,770	38,793	40,239	376,161
14	9,448	9,903	10,320	10,655	11,100	12,720	12,612	12,319	12,770	13,491	14,203	14,035	143,576
15	25,869	27,027	27,685	28,493	29,928	33,501	34,036	34,008	34,867	37,393	38,756	39,348	390,911
1–15	191,187	199,250	206,052	212,852	221,551	243,624	247,042	248,548	255,269	263,235	272,942	281,448	2,843,000

Table 1.19: Summary statistics of firm-level variables

Variable	Definition	Mean	S.D.	25%-tile	75%-tile	No. of obs.
Y	Revenue	1.67e+07	6.18e+08	51976	915526	2667800
VA	Value-added	1470792	3.98e+07	6532	244358	2810221
L	Total number of workers in the firm	12.08116	118.9132	0	6	2843000
W	Total amount of wages paid by firm to workers	572968.7	8491577	0	189648.5	2843000
K	Fixed assets, end of year	1.09e+07	2.02e+08	2043.4	490675	252269
age	1. $Age <= 5$	2.711398	2.028502	1	4	2843000
	2. $5 < Age <= 10$					
	3. $10 < Age <= 15$					
	4. $15 < Age <= 20$					
	5. $20 < Age <= 25$					
	6. $25 < Age <= 30$					
	7. $30 < Age <= 35$					
	8. $Age > 35$					
exporter	0. Non-exporter	.1269578	.3329258	0	0	2842999
	1. Exporter					
foreign_owned	0. Local (with =30% local equity)	.1685614	.3743641	0	0	2843000
	1. Foreign					
y	ln Y, deflated	7.837737	2.404643	6.433623	9.149777	2561491
va	ln VA, deflated	6.74991	2.167953	5.483785	8.057137	2346408
l	ln L	1.580005	1.346234	.6931472	2.397895	1746007
w	ln W, deflated	7.123275	1.804398	6.046929	8.222347	1746007
m	ln (Y - VA), deflated	7.546274	2.421554	6.125949	8.878644	2425636
k	ln K, deflated, beginning of year	7.003602	3.12202	4.931695	8.954113	175200
Lprod_va	va - l	5.83589	1.268692	5.165665	6.564051	1597389
tfp_va	estimated total factor productivity	7.62756	1.367494	6.865269	8.399335	133955

Note: The variables of Y, VA, W, (Y-VA), and K are deflated using the GDP deflators downloaded from the Singapore DOS website. In particular, the series—Gross Domestic Product Deflators (2015 = 100), By Industry (SSIC 2020), Annual—for the period 2009–2020 is used. Note however that the deflators are not available for individual manufacturing sub-sectors at which the firm-level data are classified. Thus, we apply the common deflator for "Manufacturing" to the five manufacturing sub-sectors. Note also that the deflators are available for "Utilities", "Other Goods Industries", and "Other Services Industries" separately, while the sector Others defined in Table 1.1 includes them all. As an approximation, we apply the deflator for "Other Goods Industries" to the residual Others sector.

Table 1.20: Value-added production function parameter estimates

	(1) Electronics	(2) Chemicals & Biomedical Manufacturing	(3) Precision Engineering	(4) Transport Engineering	(5) General Manufacturing Industries	(6) Construction	(7) Wholesale Trade and Retail Trade
l	0.412***	0.482***	0.606***	0.436***	0.479***	0.472***	0.507***
	(0.0307)	(0.0298)	(0.0267)	(0.0259)	(0.0327)	(0.0176)	(0.00768)
k	0.0984**	0.0705**	0.0596***	0.0654**	0.0940***	0.0382**	0.0429***
	(0.0481)	(0.0340)	(0.0222)	(0.0285)	(0.0279)	(0.0174)	(0.00530)
Obs.	899	1620	2978	1088	1771	3504	26603

	(8) Transportation & Storage	(9) Accommodation and Food & Beverage Services	(10) Information & Communication	(11) Finance & Insurance	(12) Real Estate	(13) Professional Services	(14) Administrative & Support Services	(15) Others
l	0.522***	0.244***	0.637***	0.955***	0.459***	0.647***	0.593***	0.556***
	(0.0116)	(0.0230)	(0.0123)	(0.0164)	(0.0188)	(0.0110)	(0.0112)	(0.0140)
k	0.0451***	0.0773***	0.0372***	0.0155*	0.0422	0.0304***	0.0670***	0.0238**
	(0.0109)	(0.0151)	(0.00930)	(0.00802)	(0.0261)	(0.00702)	(0.0135)	(0.00997)
Obs.	6504	3931	7107	7736	1980	11203	4495	6462

Note: The production function is estimated based on the Wooldridge (2009) GMM method using value added as the left-hand-side variable, intermediate input as the proxy variable, and a 2nd-order polynomial function in k and m to approximate the productivity component ω. The production function parameters are estimated sector by sector.

1.5 Impacts of GVC on Firm-level Productivity and Employment

We then estimate the impact of GVC on firm-level productivity. In particular, we evaluate whether the GVC characteristics of a sector might have significant effects on firms operating in the sector. We consider all of the four GVC indicators discussed in Section 1.2, i.e., the intensity of backward and forward linkages, the depth of GVC participation, and the degree of downstreamness, as summarized by Equations (1.1)–(1.4). In particular, we estimate the following specification(s):

$$\begin{aligned}
\text{Productivity}_{f,s,t} = \quad & \beta_0 + \beta_1 \times \text{gvc}_{s,t} \\
+ \quad & \beta_2 \times \text{age}_{f,t} + \beta_3 \times \text{exporters}_{f,t} + \beta_4 \times \text{foreign-owned}_{f,t} \qquad (1.11) \\
+ \quad & FE_f + \beta_5 \times \text{Productivity}_{f,s,t-1} + \beta_6 \times \text{Productivity}_{f,s,t-2} + \varepsilon_{f,s,t},
\end{aligned}$$

where a firm's productivity (in log) is hypothesized to depend on the GVC characteristics of the sector that the firm operates in, the firm's age (in bins), the firm's exporting status, and the firm's ownership structure in the baseline specification. We gradually generalize the specification to control for firm fixed effects (FEs), and to allow for serially correlated productivity process with one lag and further with two lags.

1.5.1 Labor Productivity

The results based on labor productivity (value-added per worker) are summarized in Table 1.21. We note that the a firm's labor productivity is higher when the sector (where the firm operates in) experiences a stronger backward linkage. This is the case after the firm FEs are controlled for, as is done in the general specifications based on panel FE estimator or dynamic panel Arellano–Bond estimator. The effect of backward linkages on labor productivity is indeed stronger when we gradually use a more general specification.

In contrast, as a Singapore sector's intensity in forward linkages increases, the labor productivities of the firms operating in the sector tend to decrease. This pattern turns out to be quite robust, regardless of the productivity measure we use. Thus, this suggests that Singapore firms/workers tend to be more productive by participating in GVC via incorporating foreign contents in producing for gross exports, rather than by exporting its contents for further processing abroad down the value chain.

Given the dominance of backward linkages in Singapore's gross exports, the overall effect of GVC participation (backward and forward linkages combined) is positive and tends to align in sign with the effect of backward linkages. Similarly, the downstreamness of a sector tends to have positive effects on Singapore firms' labor productivity, as the downstreamness index in (1.4) measures the importance of backward linkages in total GVC-related trade.

It might also be of interest to note that although across firms, older established firms might have higher labor productivities, once firm FEs are controlled for, the effects of firm age tend to be negative, suggesting a decay of labor productivity within firms across time (all else being controlled). On the other hand, being/becoming an exporter and foreign owned tend to increase a firm's labor productivity, although the positive effect of foreign ownership tends to become insignificant once FEs are controlled for. This might be due to the fact that there are limited variations across time in a firm's foreign ownership status.

1.5.2 Total Factor Productivity

The results based on total factor productivity are summarized in Table 1.22. The findings are qualitatively very similar to those discussed above based on labor productivity, although the sample size is much smaller (limited to firms that report information on capital stock).

1.5.3 Employment

We then analyze how GVC participation/position influences firm-level employment. To do this, we replace firm-level productivity measures $\text{Productivity}_{f,s,t}$ in Equation (1.11) by firm-level labor employment $l_{f,s,t}$ (in log). The results are summarized in Table 1.23. Surprisingly, the patterns observed for the effects of GVC on labor productivity or total factor productivity apply to firm-level employment as well. That is, backward linkages have positive effects, while forward linkages have negative effects, on firm-level employment. In the case of Singapore, the presence of strong backward linkages dominates that of forward linkages, rendering a positive combined effects of GVC on firm-level employment. Firms operating in sectors that become more downstream also tend to increase in their labor employment. These patterns hold true after controlling for firm FEs and potential serial correlations in labor employment.

1.5.4 Robustness Checks

In the benchmark analysis, we omit the year 2020, because the Covid pandemic shock has had a far-reaching effect that could potentially alter both firm-level productivities and sector-level GVC participation/activities. Dropping the year 2020 reduces the concern of confounding factors. For robustness checks, we also consider a further restricted sample and omit both the years 2009 and 2020, in order to avoid the Great Financial Crisis period for similar reasons. Tables 1.24–1.25 report the alternative estimation results for labor productivity and employment. The findings remain similar qualitatively as those in Tables 1.21 and 1.23. In fact, the results for total factor productivity are the same as in Table 1.22, because there are no firm-level observations for (beginning-of-period) capital stocks in 2009–2011 and hence no firm-level total factor productivity estimates for these initial years.

In second set of robustness checks, we omit the sector 'Others' from the analysis, in view that the sector lumps together firms from many SSIC sub-sectors of different production natures (see Table 1.1). Both the production function estimated for the sector (from which the TFP of each firm in the sector is estimated) and the GVC indicators calculated for the sector (that pulls across sectors their gross exports, input usages, and output destinations) might be very crude approximations, and hence likely measured with large errors. As reported in Tables 1.26–1.28, the results turn out to be qualitatively quite similar to the baseline results, and the positive effects of the backward linkage and GVC intensity tend to be quantitatively stronger. The same conclusion applies to both firm-level productivities and labor employment.

1.6 Conclusion

Over the years, gross exports and foreign final demand have become more important as the source of demand for Singapore's value-added. This is a result of the increase in the domestic value-added embedded in gross exports of the Services sectors. Relative to the regional economies, Singapore has had the largest share of domestic value-added embodied in gross exports and foreign final demand, signifying the importance of external demand in supporting the Singapore economy. Analyzing the impact of GVC indicators on firm-level performance, we find that labor productivities and employment increase with stronger backward linkages, and decrease with stronger forward linkages, suggesting that a firm's employment and labor productivity are tied to the proportion of foreign content in the production of gross exports. The same pattern is observed for the firm-industry's downstreamness: a firm's employment and labor productivity increase with the downstreamness of the sector which the firm operates in. These results are consistent with the observations from Chang and Nguyen (2021), where they found Singapore to have a strong comparative advantage in the service sectors (such as finance/insurance and other business services), while most of these services are provided at the end of product value chains. This debunks the notion that it is always preferable for countries to be located in the upstream. The Singapore's case studied in this paper suggests that the firm-level labor employment and productivity in fact increase with a sector's downstreamness.

Bibliography

Belotti, Federico, Alessandro Borin, and Michele Mancini, "icio: Economic Analysis with Intercountry Input–Output Tables," *The Stata Journal*, 2021, *21* (3), 708–755.

Borin, Alessandro and Michele Mancini, "Follow the Value Added: Tracking Bilateral Relations in Global Value Chains," *MPRA Paper 82692*, 2017. University Library of Munich, Germany.

_ **and** _ , "Measuring What Matters in Global Value Chains and Value-Added Trade," *World Bank Policy Research Working Paper 8804*, 2019.

Chang, Pao-Li and Phuong T. B. Nguyen, "The Singapore Economy: Dynamism and Inclusion," Routledge 2021.

_ **and** _ , "Global Value Chains and the CPTPP," *The World Economy*, 2022, *45 (12)*, 3780–3832.

_ **and** _ , "Made in Singapore," *Singapore Economic Review*, 2022, *67 (06)*, 1835–1904.

Horvat, Peter, Colin Webb, and Norihiko Yamano, "Measuring Employment in Global Value Chains," *OECD Science, Technology and Industry Working Papers 2020/01*, 2020.

Wooldridge, Jeffrey M., "On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables," *Economics Letters*, 2009, *104*, 112–114.

Table 1.21: Impact of GVC participation/position on firm-level labor productivity, 2009–2019

	Backward Linkages					GVC Intensity			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2		(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_bwd	-0.0479***	0.461***	0.702***	1.176***	gvc_bwdfwd	0.222***	0.446***	0.536***	0.883***
	(0.0102)	(0.0312)	(0.0550)	(0.0603)		(0.00835)	(0.0353)	(0.0653)	(0.0707)
age	0.0467***	-0.0162***	-0.114***	-0.102***	age	0.0462***	-0.0164***	-0.114***	-0.105***
	(0.000460)	(0.00142)	(0.00201)	(0.00215)		(0.000461)	(0.00142)	(0.00203)	(0.00215)
exporter	0.431***	0.0789***	0.0164**	0.0252***	exporter	0.411***	0.0787***	0.0164**	0.0251***
	(0.00251)	(0.00410)	(0.00658)	(0.00649)		(0.00252)	(0.00410)	(0.00665)	(0.00653)
foreign_owned	1.028***	0.0361***	0.0197	0.0117	foreign_owned	1.021***	0.0358***	0.0198	0.0122
	(0.00290)	(0.00890)	(0.0175)	(0.0172)		(0.00290)	(0.00890)	(0.0177)	(0.0174)
L.Lprod_va			0.944***	0.831***	L.Lprod_va			0.967***	0.847***
			(0.0221)	(0.0196)				(0.0224)	(0.0201)
L2.Lprod_va				0.218***	L2.Lprod_va				0.222***
				(0.00679)					(0.00696)
Firm FE		Yes	Yes	Yes	Firm FE		Yes	Yes	Yes
N	1458792	1458792	882032	697495	N	1458792	1458792	882032	697495

	Forward Linkages					Downstreamness			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2		(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_fwd	1.004***	-0.340***	-1.318***	-2.170***	gvc_position	-0.773***	0.179***	0.786***	1.262***
	(0.0159)	(0.0705)	(0.134)	(0.144)		(0.00837)	(0.0282)	(0.0518)	(0.0570)
age	0.0456***	-0.0158***	-0.113***	-0.102***	age	0.0447***	-0.0153***	-0.111***	-0.0951***
	(0.000460)	(0.00142)	(0.00205)	(0.00225)		(0.000458)	(0.00142)	(0.00202)	(0.00223)
exporter	0.410***	0.0791***	0.0165**	0.0258***	exporter	0.422***	0.0791***	0.0164**	0.0255***
	(0.00244)	(0.00410)	(0.00674)	(0.00677)		(0.00243)	(0.00410)	(0.00658)	(0.00660)
foreign_owned	1.013***	0.0364***	0.0198	0.0134	foreign_owned	1.009***	0.0363***	0.0197	0.0125
	(0.00290)	(0.00889)	(0.0179)	(0.0180)		(0.00289)	(0.00889)	(0.0175)	(0.0175)
L.Lprod_va			0.993***	0.926***	L.Lprod_va			0.947***	0.870***
			(0.0218)	(0.0201)				(0.0216)	(0.0194)
L2.Lprod_va				0.248***	L2.Lprod_va				0.230***
				(0.00706)					(0.00680)
Firm FE		Yes	Yes	Yes	Firm FE		Yes	Yes	Yes
N	1458792	1458792	882032	697495	N	1458792	1458792	882032	697495

Table 1.22: Impact of GVC participation/position on firm-level total factor productivity, 2009–2019

	Backward Linkages			
	(1)	(2)	(3)	(4)
	OLS	FE	Dynamic Panel 1	Dynamic Panel 2
gvc_bwd	0.419***	1.369***	1.490***	1.257***
	(0.0390)	(0.132)	(0.179)	(0.257)
age	0.0811***	-0.00141	-0.0621***	-0.0425***
	(0.00186)	(0.00588)	(0.00809)	(0.0111)
exporter	0.308***	0.0564***	0.00987	-0.00822
	(0.00862)	(0.0108)	(0.0148)	(0.0181)
foreign_owned	0.586***	-0.0151	-0.0617	-0.0627
	(0.00784)	(0.0225)	(0.0388)	(0.0422)
L.tfp_va			0.399***	0.502***
			(0.0607)	(0.0812)
L2.tfp_va				0.133***
				(0.0297)
Firm FE		Yes	Yes	Yes
N	113797	113797	54978	37453

	GVC Intensity			
	(1)	(2)	(3)	(4)
	OLS	FE	Dynamic Panel 1	Dynamic Panel 2
gvc_bwdfwd	-0.129***	1.217***	1.295***	0.866***
	(0.0390)	(0.166)	(0.223)	(0.324)
age	0.0816***	-0.00493	-0.0677***	-0.0440***
	(0.00187)	(0.00594)	(0.00810)	(0.0110)
exporter	0.342***	0.0564***	0.00952	-0.00828
	(0.00866)	(0.0108)	(0.0147)	(0.0181)
foreign_owned	0.598***	-0.0183	-0.0636*	-0.0641
	(0.00788)	(0.0225)	(0.0386)	(0.0422)
L.tfp_va			0.387***	0.497***
			(0.0605)	(0.0816)
L2.tfp_va				0.131***
				(0.0298)
Firm FE		Yes	Yes	Yes
N	113797	113797	54978	37453

	Forward Linkages			
	(1)	(2)	(3)	(4)
	OLS	FE	Dynamic Panel 1	Dynamic Panel 2
gvc_fwd	-1.970***	-1.742***	-2.522***	-2.640***
	(0.0676)	(0.304)	(0.376)	(0.529)
age	0.0859***	-0.0141**	-0.0776***	-0.0468***
	(0.00186)	(0.00580)	(0.00792)	(0.0113)
exporter	0.343***	0.0576***	0.0117	-0.00751
	(0.00818)	(0.0108)	(0.0152)	(0.0187)
foreign_owned	0.631***	-0.0158	-0.0661*	-0.0639
	(0.00785)	(0.0225)	(0.0397)	(0.0435)
L.tfp_va			0.445***	0.565***
			(0.0624)	(0.0873)
L2.tfp_va				0.155***
				(0.0319)
Firm FE		Yes	Yes	Yes
N	113797	113797	54978	37453

	Downstreamness			
	(1)	(2)	(3)	(4)
	OLS	FE	Dynamic Panel 1	Dynamic Panel 2
gvc_position	0.848***	1.029***	1.447***	1.431***
	(0.0337)	(0.130)	(0.169)	(0.242)
age	0.0856***	-0.00631	-0.0659***	-0.0422***
	(0.00186)	(0.00586)	(0.00804)	(0.0112)
exporter	0.317***	0.0572***	0.0111	-0.00737
	(0.00823)	(0.0108)	(0.0150)	(0.0184)
foreign_owned	0.616***	-0.0142	-0.0619	-0.0611
	(0.00786)	(0.0225)	(0.0392)	(0.0427)
L.tfp_va			0.421***	0.529***
			(0.0613)	(0.0830)
L2.tfp_va				0.143***
				(0.0304)
Firm FE		Yes	Yes	Yes
N	113797	113797	54978	37453

Table 1.23: Impact of GVC participation/position on firm-level employment, 2009–2019

	Backward Linkages			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_bwd	0.0492***	0.323***	0.860***	0.211***
	(0.0106)	(0.0286)	(0.0237)	(0.0395)
age	0.105***	-0.00135	-0.193***	-0.0616***
	(0.000534)	(0.00134)	(0.00146)	(0.00189)
exporter	0.782***	0.235***	0.0344***	-0.00122
	(0.00308)	(0.00334)	(0.00204)	(0.00365)
foreign_owned	0.0294***	-0.0246***	-0.0314***	-0.00740
	(0.00296)	(0.00757)	(0.00629)	(0.0113)
L.l			0.493***	1.816***
			(0.0100)	(0.0185)
L2.l				-0.105***
				(0.00298)
Firm FE		Yes	Yes	Yes
N	1589339	1589339	1016336	815190

	GVC Intensity			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_bwdfwd	-0.597***	0.232***	0.733***	0.0965**
	(0.00857)	(0.0306)	(0.0271)	(0.0454)
age	0.107***	-0.00136	-0.194***	-0.0617***
	(0.000535)	(0.00134)	(0.00146)	(0.00190)
exporter	0.830***	0.235***	0.0340***	-0.00163
	(0.00311)	(0.00334)	(0.00205)	(0.00367)
foreign_owned	0.0488***	-0.0248***	-0.0315***	-0.00718
	(0.00297)	(0.00757)	(0.00631)	(0.0113)
L.l			0.501***	1.830***
			(0.00999)	(0.0186)
L2.l				-0.106***
				(0.00300)
Firm FE		Yes	Yes	Yes
N	1589339	1589339	1016336	815190

	Forward Linkages			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_fwd	-2.482***	-0.529***	-1.271***	-0.572***
	(0.0175)	(0.0649)	(0.0551)	(0.0925)
age	0.109***	-0.000933	-0.194***	-0.0612***
	(0.000532)	(0.00133)	(0.00145)	(0.00189)
exporter	0.830***	0.235***	0.0339***	-0.00138
	(0.00302)	(0.00334)	(0.00206)	(0.00366)
foreign_owned	0.0666***	-0.0242***	-0.0307***	-0.00696
	(0.00295)	(0.00757)	(0.00633)	(0.0113)
L.l			0.510***	1.825***
			(0.00994)	(0.0182)
L2.l				-0.106***
				(0.00299)
Firm FE		Yes	Yes	Yes
N	1589339	1589339	1016336	815190

	Downstreamness			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_position	1.389***	0.215***	0.906***	0.340***
	(0.00952)	(0.0265)	(0.0233)	(0.0386)
age	0.109***	-0.000447	-0.190***	-0.0607***
	(0.000531)	(0.00133)	(0.00145)	(0.00188)
exporter	0.797***	0.235***	0.0345***	-0.000710
	(0.00299)	(0.00334)	(0.00204)	(0.00363)
foreign_owned	0.0626***	-0.0243***	-0.0308***	-0.00728
	(0.00295)	(0.00757)	(0.00629)	(0.0112)
L.l			0.495***	1.801***
			(0.00997)	(0.0183)
L2.l				-0.105***
				(0.00296)
Firm FE		Yes	Yes	Yes
N	1589339	1589339	1016336	815190

Table 1.24: Impact of GVC participation/position on firm-level labor productivity, 2010–2019

	Backward Linkages			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_bwd	-0.0398***	0.458***	0.702***	1.176***
	(0.0106)	(0.0331)	(0.0550)	(0.0603)
age	0.0470***	-0.0314***	-0.114***	-0.102***
	(0.000479)	(0.00151)	(0.00201)	(0.00215)
exporter	0.434***	0.0752***	0.0164**	0.0252***
	(0.00263)	(0.00425)	(0.00658)	(0.00649)
foreign_owned	1.040***	0.0334***	0.0197	0.0117
	(0.00302)	(0.00938)	(0.0175)	(0.0172)
L.Lprod_va			0.944***	0.831***
			(0.0221)	(0.0196)
L2.Lprod_va				0.218***
				(0.00679)
Firm FE		Yes	Yes	Yes
N	1347861	1347861	882032	697495

	GVC Intensity			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_bwdfwd	0.227***	0.458***	0.536***	0.883***
	(0.00868)	(0.0377)	(0.0653)	(0.0707)
age	0.0466***	-0.0319***	-0.114***	-0.105***
	(0.000479)	(0.00151)	(0.00203)	(0.00215)
exporter	0.414***	0.0749***	0.0164**	0.0251***
	(0.00264)	(0.00425)	(0.00665)	(0.00653)
foreign_owned	1.032***	0.0330***	0.0198	0.0122
	(0.00302)	(0.00938)	(0.0177)	(0.0174)
L.Lprod_va			0.967***	0.847***
			(0.0224)	(0.0201)
L2.Lprod_va				0.222***
				(0.00696)
Firm FE		Yes	Yes	Yes
N	1347861	1347861	882032	697495

	Forward Linkages			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_fwd	1.005***	-0.262***	-1.318***	-2.170***
	(0.0166)	(0.0740)	(0.134)	(0.144)
age	0.0459***	-0.0327***	-0.113***	-0.102***
	(0.000478)	(0.00151)	(0.00205)	(0.00225)
exporter	0.414***	0.0752***	0.0165**	0.0258***
	(0.00256)	(0.00426)	(0.00674)	(0.00677)
foreign_owned	1.024***	0.0337***	0.0198	0.0134
	(0.00301)	(0.00938)	(0.0179)	(0.0180)
L.Lprod_va			0.993***	0.926***
			(0.0218)	(0.0201)
L2.Lprod_va				0.248***
				(0.00706)
Firm FE		Yes	Yes	Yes
N	1347861	1347861	882032	697495

	Downstreamness			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_position	-0.769***	0.133***	0.786***	1.262***
	(0.00874)	(0.0298)	(0.0518)	(0.0570)
age	0.0450***	-0.0320***	-0.111***	-0.0951***
	(0.000477)	(0.00153)	(0.00202)	(0.00223)
exporter	0.427***	0.0752***	0.0164**	0.0255***
	(0.00254)	(0.00425)	(0.00658)	(0.00660)
foreign_owned	1.021***	0.0336***	0.0197	0.0125
	(0.00301)	(0.00938)	(0.0175)	(0.0175)
L.Lprod_va			0.947***	0.870***
			(0.0216)	(0.0194)
L2.Lprod_va				0.230***
				(0.00680)
Firm FE		Yes	Yes	Yes
N	1347861	1347861	882032	697495

Table 1.25: Impact of GVC participation/position on firm-level employment, 2010–2019

	Backward Linkages			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_bwd	-0.0436***	0.308***	0.860***	0.211***
	(0.0109)	(0.0295)	(0.0237)	(0.0395)
age	0.107***	-0.00690***	-0.193***	-0.0616***
	(0.000554)	(0.00136)	(0.00146)	(0.00189)
exporter	0.794***	0.225***	0.0344***	-0.00122
	(0.00323)	(0.00338)	(0.00204)	(0.00365)
foreign_owned	0.0337***	-0.0238***	-0.0314***	-0.00740
	(0.00308)	(0.00777)	(0.00629)	(0.0113)
L.l			0.493***	1.816***
			(0.0100)	(0.0185)
L2.l				-0.105***
				(0.00298)
Firm FE		Yes	Yes	Yes
N	1469987	1469987	1016336	815190

	GVC Intensity			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_bwdfwd	-0.642***	0.227***	0.733***	0.0965**
	(0.00886)	(0.0321)	(0.0271)	(0.0454)
age	0.108***	-0.00740***	-0.194***	-0.0617***
	(0.000554)	(0.00137)	(0.00146)	(0.00190)
exporter	0.840***	0.225***	0.0340***	-0.00163
	(0.00325)	(0.00338)	(0.00205)	(0.00367)
foreign_owned	0.0535***	-0.0240***	-0.0315***	-0.00718
	(0.00308)	(0.00777)	(0.00631)	(0.0113)
L.l			0.501***	1.830***
			(0.00999)	(0.0186)
L2.l				-0.106***
				(0.00300)
Firm FE		Yes	Yes	Yes
N	1469987	1469987	1016336	815190

	Forward Linkages			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_fwd	-2.425***	-0.462***	-1.271***	-0.572***
	(0.0182)	(0.0658)	(0.0551)	(0.0925)
age	0.110***	-0.00751***	-0.194***	-0.0612***
	(0.000552)	(0.00137)	(0.00145)	(0.00189)
exporter	0.833***	0.225***	0.0339***	-0.00138
	(0.00316)	(0.00338)	(0.00206)	(0.00366)
foreign_owned	0.0682***	-0.0234***	-0.0307***	-0.00696
	(0.00307)	(0.00777)	(0.00633)	(0.0113)
L.l			0.510***	1.825***
			(0.00994)	(0.0182)
L2.l				-0.106***
				(0.00299)
Firm FE		Yes	Yes	Yes
N	1469987	1469987	1016336	815190

	Downstreamness			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_position	1.335***	0.178***	0.906***	0.340***
	(0.00994)	(0.0271)	(0.0233)	(0.0386)
age	0.111***	-0.00672***	-0.190***	-0.0607***
	(0.000551)	(0.00136)	(0.00145)	(0.00188)
exporter	0.800***	0.225***	0.0345***	-0.000710
	(0.00313)	(0.00338)	(0.00204)	(0.00363)
foreign_owned	0.0633***	-0.0236***	-0.0308***	-0.00728
	(0.00306)	(0.00777)	(0.00629)	(0.0112)
L.l			0.495***	1.801***
			(0.00997)	(0.0183)
L2.l				-0.105***
				(0.00296)
Firm FE		Yes	Yes	Yes
N	1469987	1469987	1016336	815190

Table 1.26: Impact of GVC participation/position on firm-level labor productivity, 2009–2019, excluding the ‘Others’ sector

	Backward Linkages			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_bwd	-0.153***	0.531***	0.877***	1.412***
	(0.0126)	(0.0327)	(0.0552)	(0.0624)
age	0.0474***	-0.0238***	-0.114***	-0.0980***
	(0.000508)	(0.00157)	(0.00222)	(0.00239)
exporter	0.427***	0.0801***	0.0166**	0.0257***
	(0.00257)	(0.00424)	(0.00665)	(0.00665)
foreign_owned	1.050***	0.0305***	0.00834	0.00127
	(0.00301)	(0.00950)	(0.0184)	(0.0184)
L.Lprod_va			0.893***	0.801***
			(0.0218)	(0.0200)
L2.Lprod_va				0.212***
				(0.00718)
Firm FE		Yes	Yes	Yes
N	1238644	1238644	743385	586448

	GVC Intensity			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_bwdfwd	0.357***	0.625***	1.049***	1.534***
	(0.0134)	(0.0414)	(0.0696)	(0.0792)
age	0.0468***	-0.0231***	-0.112***	-0.0981***
	(0.000511)	(0.00157)	(0.00222)	(0.00238)
exporter	0.405***	0.0798***	0.0165**	0.0253***
	(0.00258)	(0.00424)	(0.00663)	(0.00660)
foreign_owned	1.048***	0.0301***	0.00845	0.00108
	(0.00301)	(0.00950)	(0.0183)	(0.0183)
L.Lprod_va			0.889***	0.784***
			(0.0219)	(0.0205)
L2.Lprod_va				0.206***
				(0.00729)
Firm FE		Yes	Yes	Yes
N	1238644	1238644	743385	586448

	Forward Linkages			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_fwd	1.355***	-0.556***	-0.912***	-1.992***
	(0.0193)	(0.0818)	(0.147)	(0.162)
age	0.0460***	-0.0243***	-0.115***	-0.104***
	(0.000508)	(0.00156)	(0.00230)	(0.00250)
exporter	0.414***	0.0802***	0.0167**	0.0263***
	(0.00252)	(0.00424)	(0.00689)	(0.00701)
foreign_owned	1.042***	0.0309***	0.00818	0.00349
	(0.00301)	(0.00950)	(0.0190)	(0.0194)
L.Lprod_va			0.964***	0.917***
			(0.0219)	(0.0208)
L2.Lprod_va				0.248***
				(0.00759)
Firm FE		Yes	Yes	Yes
N	1238644	1238644	743385	586448

	Downstreamness			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_position	-0.851***	0.322***	0.594***	1.156***
	(0.00894)	(0.0317)	(0.0559)	(0.0629)
age	0.0452***	-0.0242***	-0.115***	-0.101***
	(0.000506)	(0.00156)	(0.00226)	(0.00244)
exporter	0.428***	0.0802***	0.0167**	0.0263***
	(0.00252)	(0.00424)	(0.00677)	(0.00683)
foreign_owned	1.038***	0.0309***	0.00835	0.00279
	(0.00301)	(0.00950)	(0.0187)	(0.0189)
L.Lprod_va			0.930***	0.860***
			(0.0220)	(0.0202)
L2.Lprod_va				0.230***
				(0.00731)
Firm FE		Yes	Yes	Yes
N	1238644	1238644	743385	586448

Table 1.27: Impact of GVC participation/position on firm-level total factor productivity, 2009–2019, excluding the ‘Others’ sector

	Backward Linkages					GVC Intensity			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2		(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_bwd	0.496***	1.475***	1.653***	1.435***	gvc_bwdfwd	-0.208***	1.539***	1.746***	1.322***
	(0.0432)	(0.137)	(0.182)	(0.262)		(0.0517)	(0.181)	(0.231)	(0.339)
age	0.0773***	-0.00355	-0.0586***	-0.0343***	age	0.0768***	-0.00406	-0.0602***	-0.0345***
	(0.00192)	(0.00615)	(0.00858)	(0.0120)		(0.00192)	(0.00628)	(0.00862)	(0.0119)
exporter	0.293***	0.0590***	0.0146	-0.00564	exporter	0.330***	0.0590***	0.0139	-0.00580
	(0.00878)	(0.0112)	(0.0152)	(0.0188)		(0.00885)	(0.0112)	(0.0151)	(0.0187)
foreign_owned	0.606***	-0.0270	-0.0699*	-0.0646	foreign_owned	0.613***	-0.0294	-0.0704*	-0.0650
	(0.00823)	(0.0242)	(0.0414)	(0.0472)		(0.00822)	(0.0242)	(0.0409)	(0.0469)
L.tfp_va			0.390***	0.514***	L.tfp_va			0.369***	0.496***
			(0.0607)	(0.0814)				(0.0602)	(0.0812)
L2.tfp_va				0.138***	L2.tfp_va				0.131***
				(0.0303)					(0.0301)
Firm FE		Yes	Yes	Yes	Firm FE		Yes	Yes	Yes
N	105538	105538	51226	34972	N	105538	105538	51226	34972

	Forward Linkages					Downstreamness			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2		(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_fwd	-2.390***	-1.798***	-2.401***	-2.515***	gvc_position	0.897***	1.111***	1.458***	1.424***
	(0.0750)	(0.318)	(0.385)	(0.545)		(0.0346)	(0.136)	(0.174)	(0.250)
age	0.0806***	-0.0197***	-0.0796***	-0.0402***	age	0.0807***	-0.0121**	-0.0678***	-0.0360***
	(0.00191)	(0.00604)	(0.00836)	(0.0124)		(0.00192)	(0.00607)	(0.00847)	(0.0122)
exporter	0.317***	0.0602***	0.0165	-0.00531	exporter	0.296***	0.0598***	0.0159	-0.00519
	(0.00843)	(0.0112)	(0.0156)	(0.0196)		(0.00852)	(0.0112)	(0.0155)	(0.0193)
foreign_owned	0.641***	-0.0301	-0.0755*	-0.0673	foreign_owned	0.629***	-0.0287	-0.0720*	-0.0649
	(0.00820)	(0.0242)	(0.0426)	(0.0494)		(0.00824)	(0.0242)	(0.0421)	(0.0484)
L.tfp_va			0.439***	0.598***	L.tfp_va			0.419***	0.560***
			(0.0626)	(0.0895)				(0.0617)	(0.0848)
L2.tfp_va				0.166***	L2.tfp_va				0.154***
				(0.0332)					(0.0316)
Firm FE		Yes	Yes	Yes	Firm FE		Yes	Yes	Yes
N	105538	105538	51226	34972	N	105538	105538	51226	34972

Table 1.28: Impact of GVC participation/position on firm-level employment, 2009–2019, excluding the 'Others' sector

	Backward Linkages			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_bwd	0.135***	0.337***	0.959***	0.195***
	(0.0133)	(0.0303)	(0.0250)	(0.0405)
age	0.112***	-0.00536***	-0.196***	-0.0583***
	(0.000584)	(0.00147)	(0.00159)	(0.00203)
exporter	0.759***	0.237***	0.0357***	0.000273
	(0.00310)	(0.00343)	(0.00211)	(0.00374)
foreign_owned	0.0364***	-0.0257***	-0.0309***	-0.00539
	(0.00307)	(0.00811)	(0.00663)	(0.0118)
L.l			0.488***	1.789***
			(0.0103)	(0.0187)
L2.l				-0.125***
				(0.00320)
Firm FE		Yes	Yes	Yes
N	1355087	1355087	864414	692915

	GVC Intensity			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_bwdfwd	-1.252***	0.213***	1.136***	0.161***
	(0.0142)	(0.0363)	(0.0315)	(0.0501)
age	0.114***	-0.00501***	-0.195***	-0.0584***
	(0.000586)	(0.00146)	(0.00159)	(0.00204)
exporter	0.818***	0.237***	0.0354***	0.0000734
	(0.00314)	(0.00344)	(0.00211)	(0.00375)
foreign_owned	0.0461***	-0.0259***	-0.0311***	-0.00529
	(0.00307)	(0.00811)	(0.00664)	(0.0119)
L.l			0.490***	1.795***
			(0.0103)	(0.0189)
L2.l				-0.125***
				(0.00321)
Firm FE		Yes	Yes	Yes
N	1355087	1355087	864414	692915

	Forward Linkages			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_fwd	-3.638***	-0.991***	-0.976***	-0.406***
	(0.0218)	(0.0769)	(0.0611)	(0.0992)
age	0.117***	-0.00649***	-0.201***	-0.0586***
	(0.000580)	(0.00148)	(0.00160)	(0.00205)
exporter	0.783***	0.238***	0.0348***	-0.0000793
	(0.00305)	(0.00343)	(0.00213)	(0.00376)
foreign_owned	0.0572***	-0.0249***	-0.0302***	-0.00499
	(0.00305)	(0.00810)	(0.00671)	(0.0119)
L.l			0.512***	1.804***
			(0.0102)	(0.0183)
L2.l				-0.126***
				(0.00322)
Firm FE		Yes	Yes	Yes
N	1355087	1355087	864414	692915

	Downstreamness			
	(1) OLS	(2) FE	(3) Dynamic Panel 1	(4) Dynamic Panel 2
gvc_position	1.596***	0.413***	0.727***	0.187***
	(0.0103)	(0.0303)	(0.0250)	(0.0411)
age	0.117***	-0.00603***	-0.199***	-0.0585***
	(0.000580)	(0.00147)	(0.00160)	(0.00204)
exporter	0.753***	0.238***	0.0355***	0.000180
	(0.00304)	(0.00343)	(0.00212)	(0.00375)
foreign_owned	0.0580***	-0.0252***	-0.0305***	-0.00518
	(0.00305)	(0.00811)	(0.00667)	(0.0119)
L.l			0.498***	1.794***
			(0.0103)	(0.0185)
L2.l				-0.125***
				(0.00321)
Firm FE		Yes	Yes	Yes
N	1355087	1355087	864414	692915

Figure 1.1: Singapore's GVC participation/position in Manufacturing

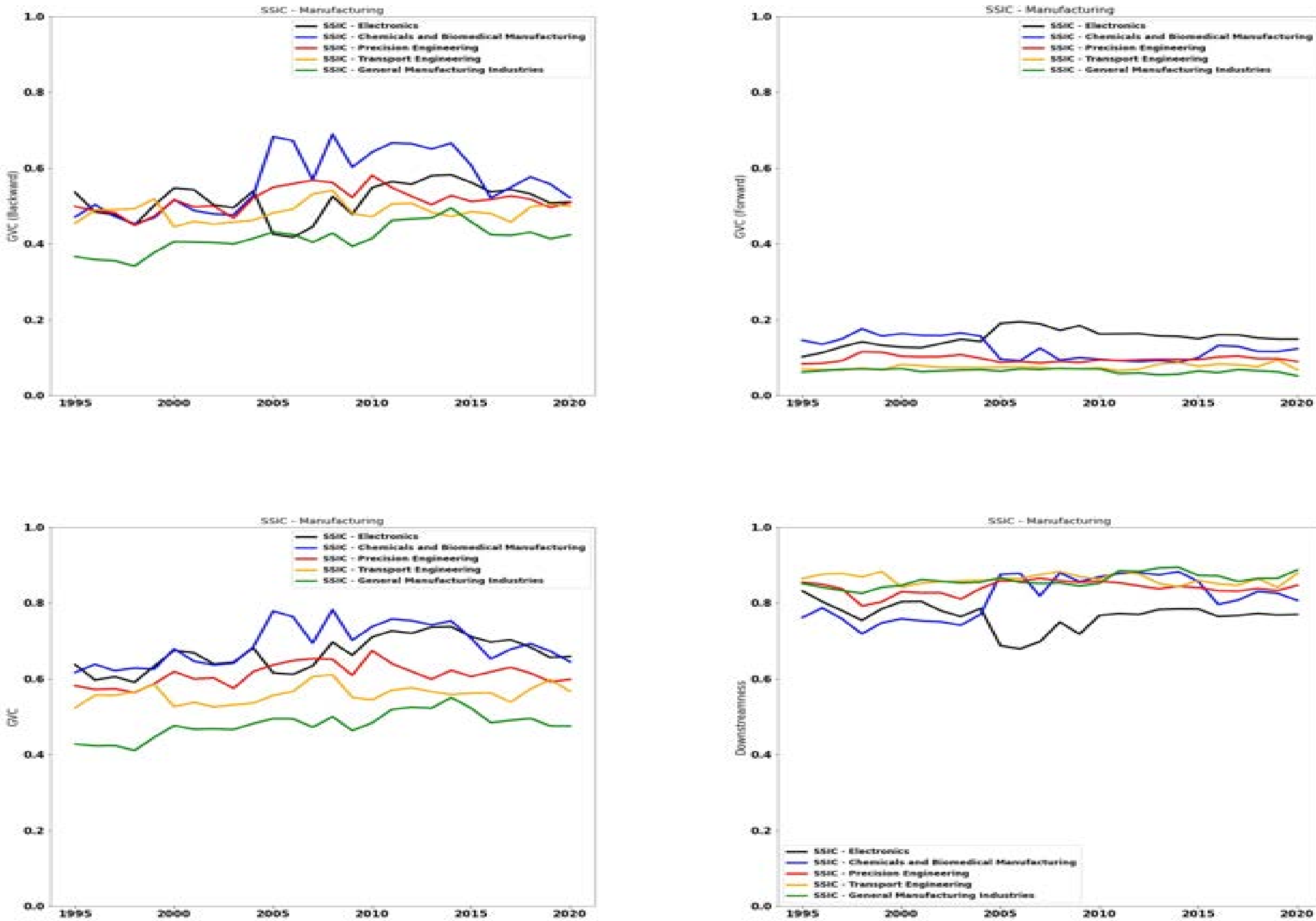


Figure 1.2: Singapore's GVC participation/position in Services

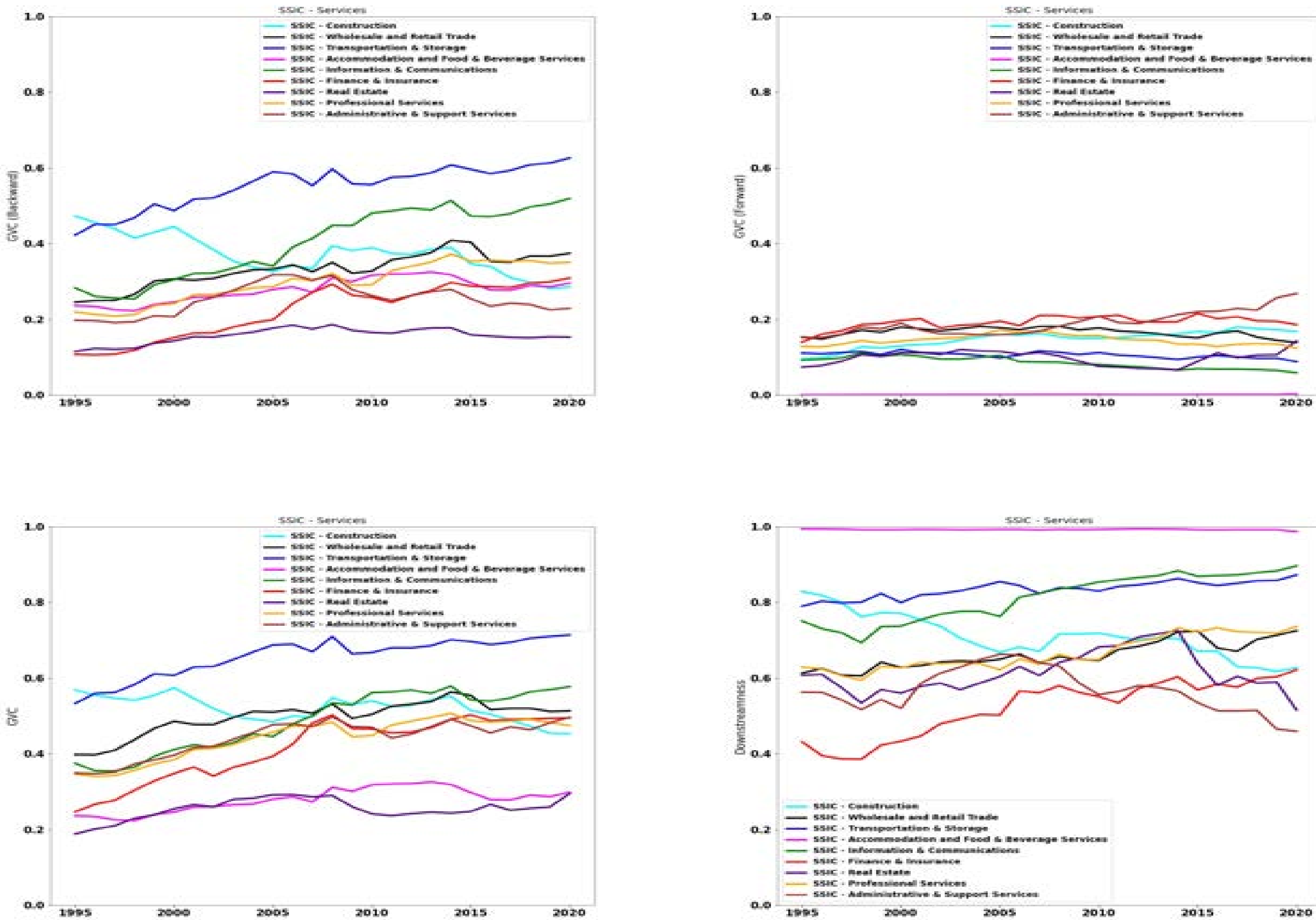


Figure 1.3: Singapore domestic value-added / labor value-added / labor employment embodied in gross exports / foreign final demand, as share of the economy's total value-added / labor value-added / labor employment

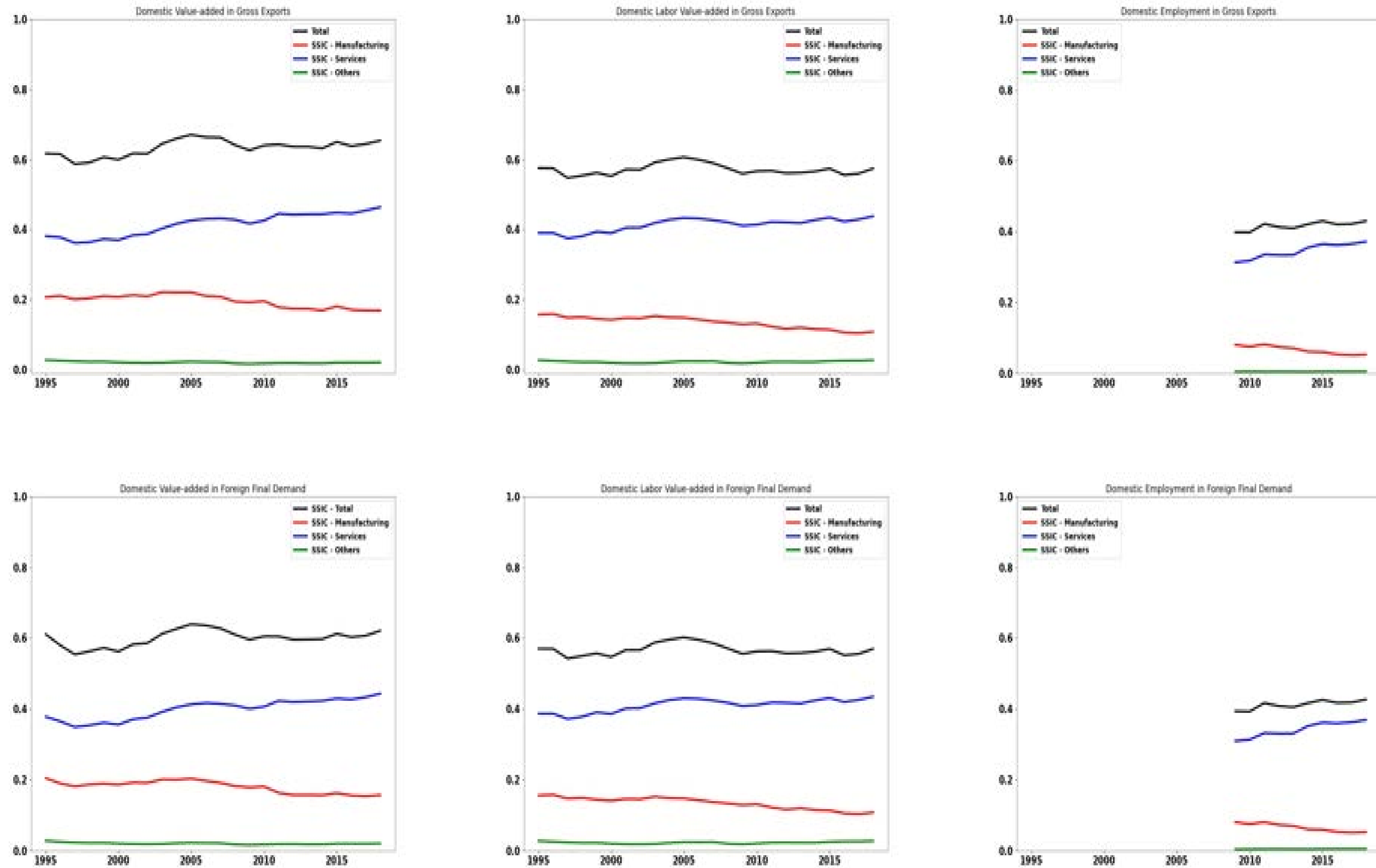


Figure 1.4: Singapore domestic value-added / labor value-added / labor employment embodied in gross exports (distribution across sectors)

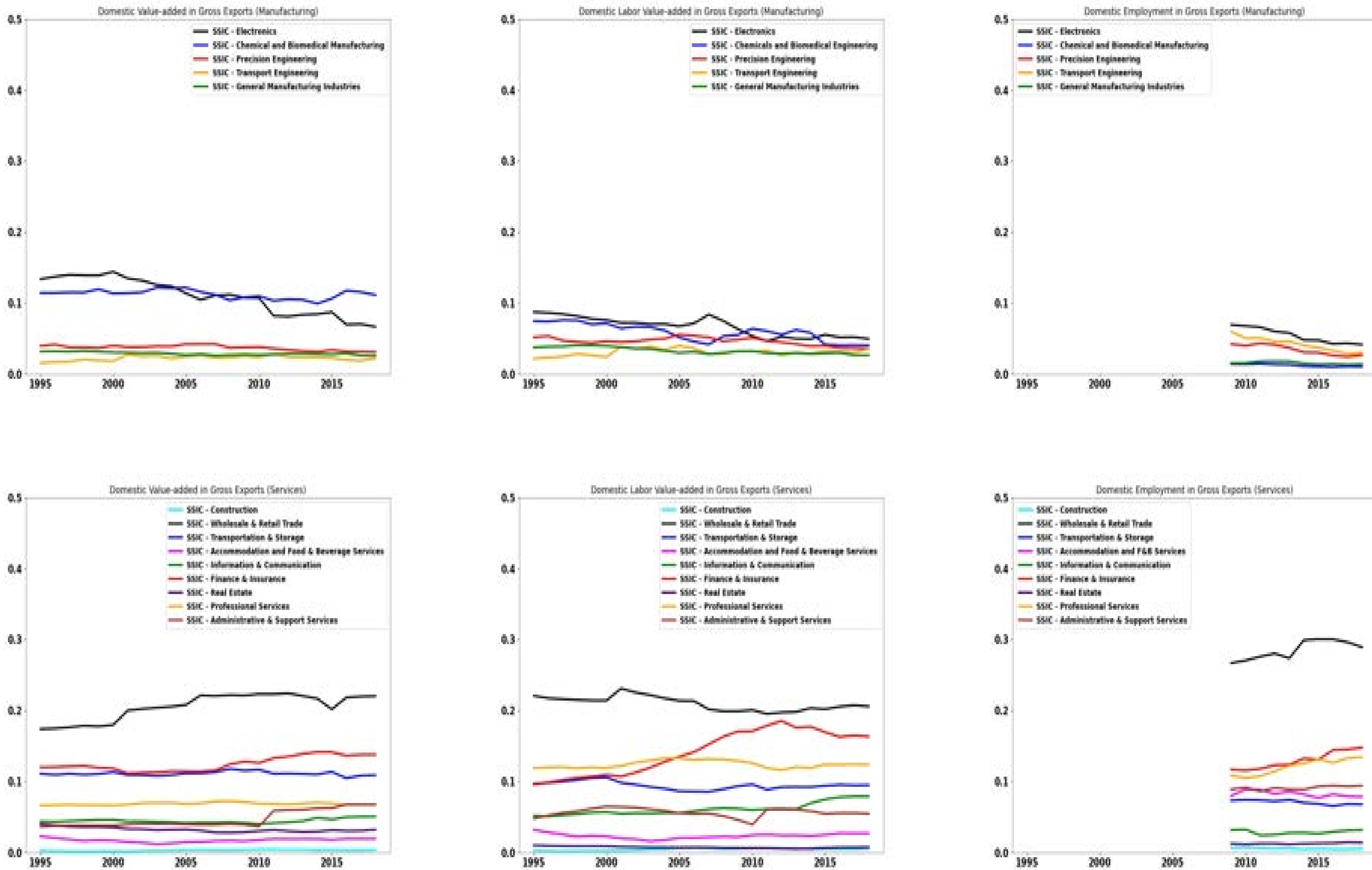


Figure 1.5: Singapore domestic value-added / labor value-added / labor employment embodied in foreign final demand (distribution across sectors)

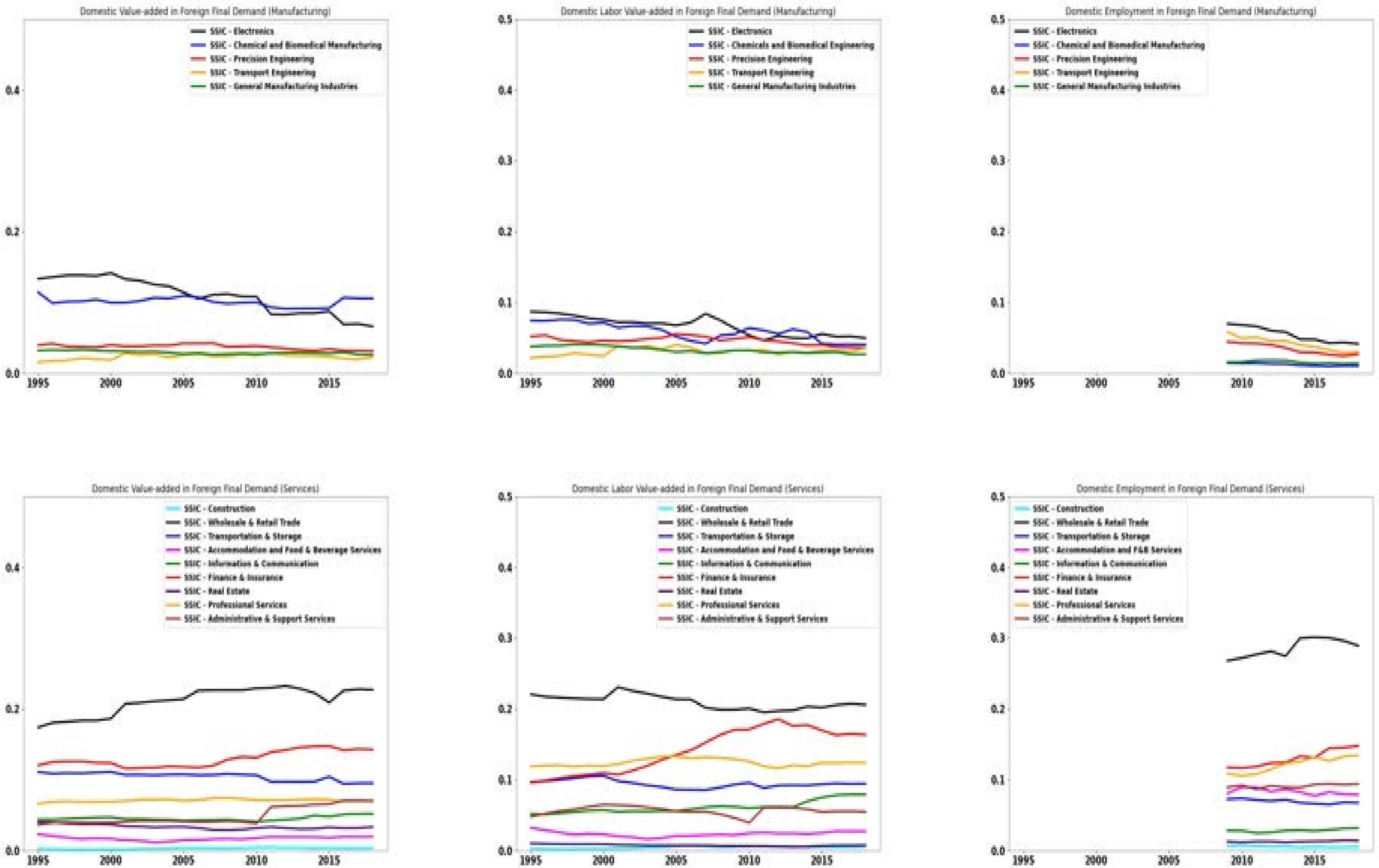


Figure 1.6: Singapore domestic value-added / labor value-added / labor employment embodied in gross exports, relative to the value-added / labor value-added / labor employment of the sector

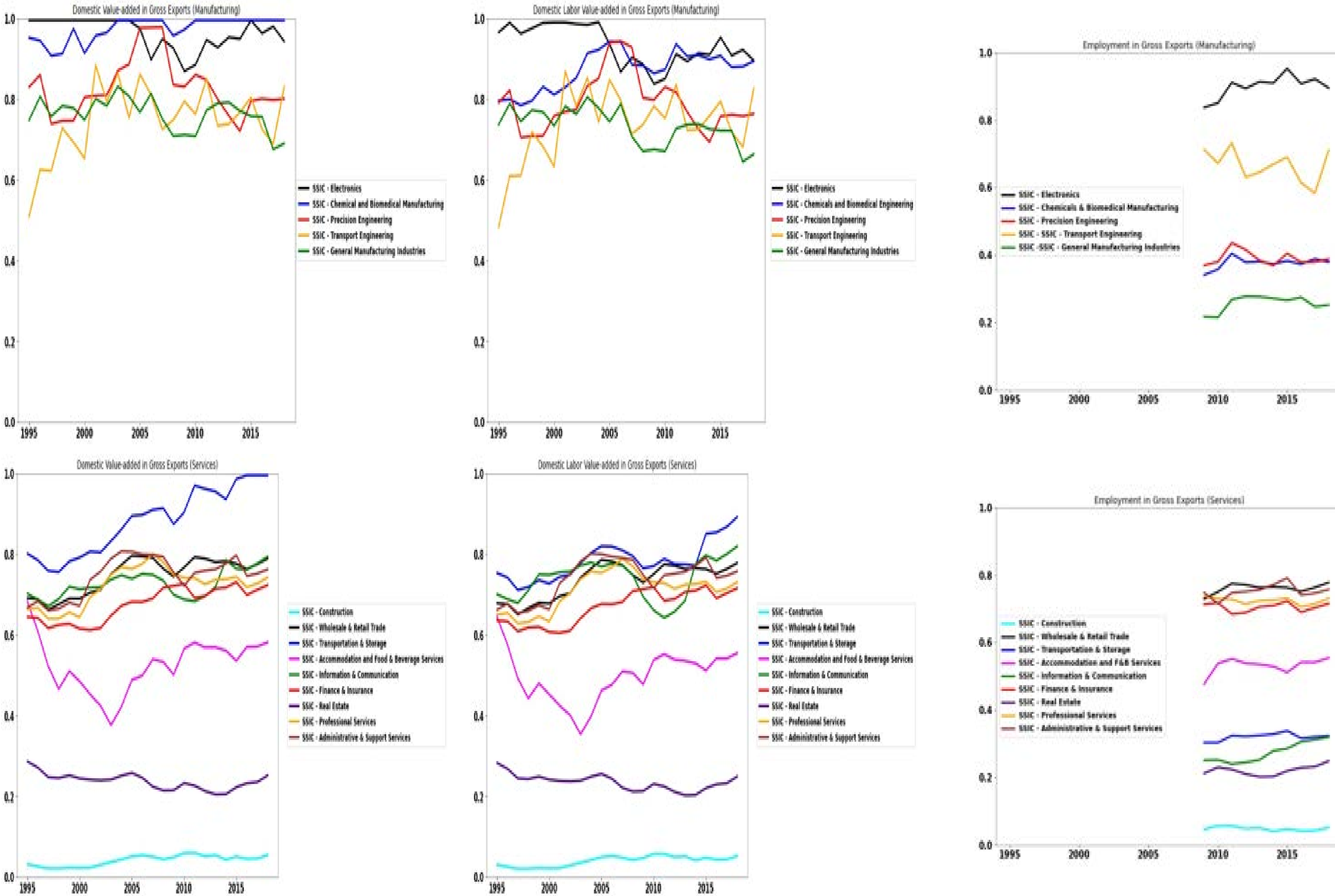
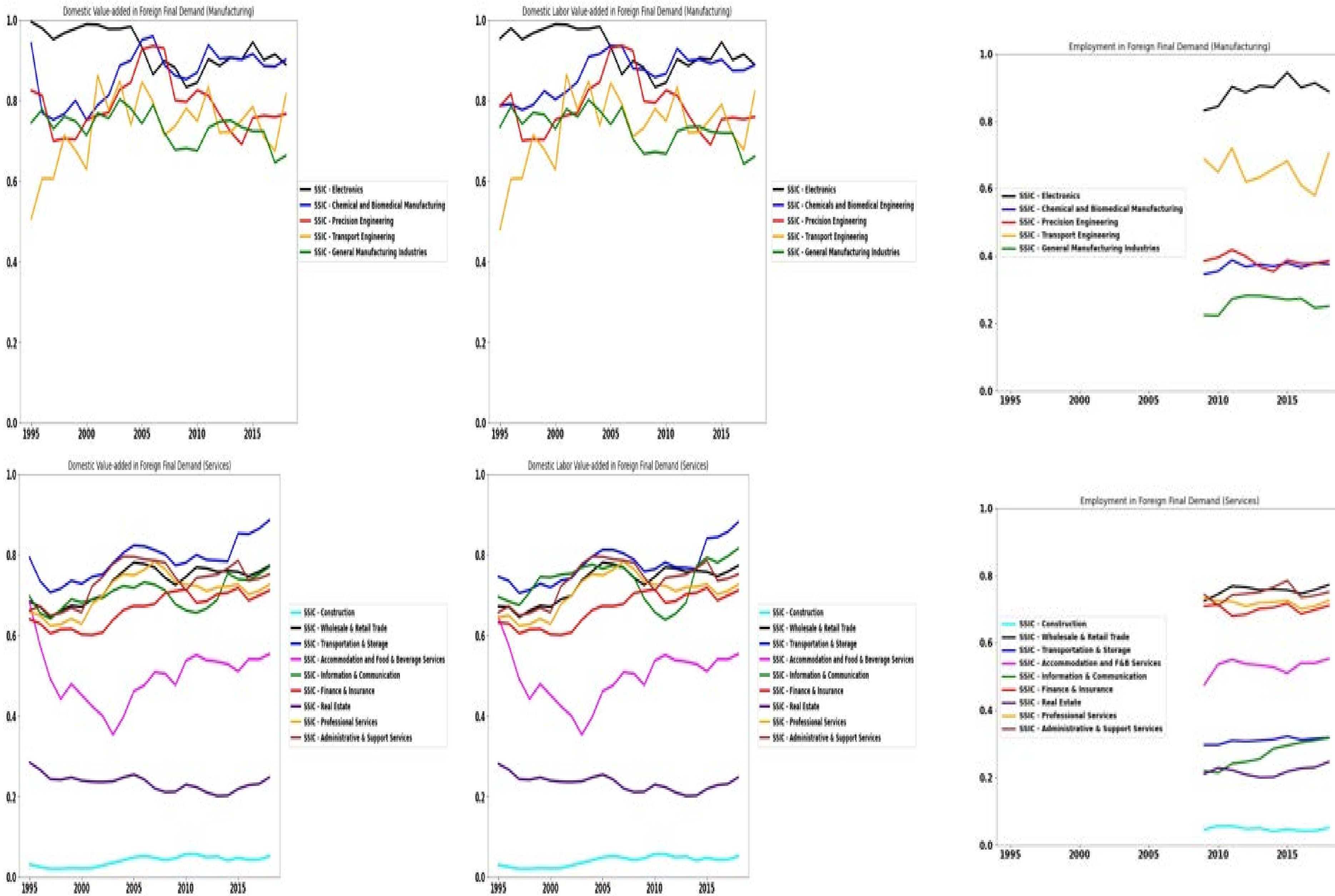


Figure 1.7: Singapore domestic value-added / labor value-added / labor employment embodied in foreign final demand, relative to the value-added / labor value-added / labor employment of the sector



Chapter 2

Estimating Firm-Level Production Functions with Spatial Dependence

2.1 Introduction

Firm-level productivity (production function) estimation is critical to both positive and normative research, in inferring the characteristics of firm-level production activities and identifying the effect of policy/exogenous shocks (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg, Caves and Frazer, 2015; Wooldridge, 2009). Equally important, a large literature has documented/analyzed how firms interact with each other via the input-output linkages, factor markets, and knowledge spillovers (e.g., Demir, Fieler, Xu and Yang, 2024; Miyauchi, 2023; Alfaro-Ureña, Manelici and Vasquez, 2022; Ellison, Glaeser and Kerr, 2010; Helsley and Strange, 1990; Diamond and Simon, 1990; Helsley and Strange, 2002; Jaffe, Trajtenberg and Henderson, 1993; Audretsch and Feldman, 1996; Matray, 2021). This paper contributes to the literature by proposing methodologies for estimating firm-level productivity (production function), simultaneously taking into account potential spatial interactions across firms. In particular, a firm's productivity is allowed to depend on related firms' lagged outputs (e.g., via local input-output linkages), on related firms' lagged labor inputs (e.g., via sharing local labor pools), and on related firms' current productivity shocks (e.g., via knowledge spillovers with the boundary defined by the geographical area and/or by the network of suppliers-customers).

We develop a three-stage efficient GMM estimation algorithm, and show by theory the asymptotic properties of the proposed estimator and by Monte Carlo simulations the finite sample performance of the estimator. The procedure provides the estimates of the production function parameters (the labor and capital elasticities in value-added), the degree of autoregressive correlation in the productivity process, and the spatial parameters (characterizing the dependence of productivity on related firms' lagged outputs and lagged inputs respectively, and the strength of spatial error correlation of the productivity shocks). The Monte Carlo simulations demonstrate that the proposed estimator yields point estimates that are consistent for the true parameters both in the absence and in the presence of spatial effects. In other words, it returns statistically insignificant coefficient estimates of the spatial dependence parameters, when the underlying DGPs are free of such structures, and consistent estimates of the spatial dependence parameters when the underlying DGPs are characterized with such structures (via the lagged output, the lagged labor input, or the productivity shock channel). This finding holds for DGPs with strong or weak spatial dependence, and DGPs with positive or negative spatial dependence. The proposed estimation algorithm also generates standard error estimates of the parameters that are consistent with the Monte Carlo simulated standard deviations, and with a convergence rate (when the sample size changes) in line with the theory. In contrast, the conventional productivity estimators (which ignore potential spatial interactions across firms) lead to biased estimates of the production function parameters when the underlying DGPs exhibit spatial dependence structure (and thus, for which the conventional estimator is misspecified). The conventional estimates are upward (downward) biased when the underlying DGPs have positive (negative) spatial dependence structure, and the extents of the bias worsen when the underlying DGPs' spatial dependence strengthens.

We apply the developed methodology and estimation algorithm to the Japanese BSJBSA-TSR linked dataset for the period 2009–2018. The dataset combines the firm-level financial statement information from the Basic Survey of Japanese Business Structure and Activities (BSJBSA), and the firm-to-firm supplier-customer relationship from Tokyo Shoko Research (TSR). The estimation sample covers 13,001 firms per year (both publicly listed and unlisted firms in Japan of medium/large sizes, across 194 commuting zones, and 14 industries), and in particular, provides information on each firm's most important domestic suppliers and customers (up to 24 connections, respectively). We find significant and positive spatial coefficients in the Japanese firm-level productivity process via all three proposed channels. In particular, a 1% increase in the average sales of a firm's suppliers/customers in the previous period in the same commuting zone helps improve the firm's current productivity by 0.005%. A larger local labor market also enhances a firm's productivity: specifically, a 1% increase in the average labor inputs in the previous period by firms located in the same commuting zone raises a firm's current productivity by 0.05%. There is also evidence of contemporary knowledge spillovers among firms located in the same commuting zone (with a positive and significant spatial error correlation coefficient of 0.41) and/or with supplier-customer relationships. In sum, the proposed estimator suggests that spatial interactions across firms play a significant role in determining the Japanese firm-level productivity both statistically and economically.

The prior literatures on estimating firm-level production functions typically ignore potential spatial dependence across firms. The firm-level production functions are often taken to be independent and estimated, before the estimated productivities are used to analyze potential determinants of firm productivity (such as exposure to foreign direct investment or imports at the sector level or supplying to multinational corporations at the firm level). The recent work by Iyoha (2022) highlights the need to estimate firm productivities in a modified framework taking into account the presence of productivity spillovers. Her work, however, models the interdependence across firms "in reduced form" in terms of their productivities, and not directly in terms of outputs, inputs, or the productivity shocks of related firms. This leads to a rather difficult setup for estimations, and for establishing the asymptotic properties (and variance-covariance) of the estimator.

In this paper, we propose methodologies that model productivity dependence across firms structurally where the spatial effects operate via potentially lagged outputs, lagged inputs, and current productivity shocks of related firms, motivated by the mechanisms highlighted by the literatures. We draw on the approaches proposed in the productivity estimation literature (e.g., Wooldridge, 2009; Ackerberg, Caves and Frazer, 2015), and the spatial econometrics literature (e.g., Kelejian and Prucha, 1998, 1999; Kapoor, Kelejian and Prucha, 2007; Lee and Yu, 2014; Elhorst, 2014). The resulting three-stage efficient GMM estimator has standard asymptotic properties, with variance-covariance estimators that take into account the spatial interactions across firms in each of the three dimensions proposed. The sets of instruments suitable for each stage are also straightforward extensions of those suggested by each of these two individual literatures. As discussed above, our proposed estimator is shown to be consistent under DGPs with or without spatial dependence. In contrast, the conventional estimators are biased when the true DGPs are indeed characterized by spatial dependence. These findings imply that analyzing spatial interactions across firms based on the productivities estimated by the conventional estimators will lead to biased inferences. Instead, the proposed estimator in this paper offers a framework to simultaneously estimate firm production functions and spatial interactions across firms in one unified setup.

The rest of the paper is organized as follows. We set up the model in Section 2.2. In Section 2.3, we develop the estimation algorithms and establish the asymptotic properties of the proposed estimator. Section 2.4 introduces the Japanese firm-level and firm-to-firm datasets. Section 2.5 conducts Monte Carlo simulations to evaluate the performance of the proposed estimator in comparison with the conventional estimator. In Section 2.6, we apply the proposed methodology empirically to the Japanese dataset, and Section 2.7 concludes.

2.2 Model

Consider the following production function, where a firm's value-added depends on its primary factor inputs and productivity:

$$va_{it} = \alpha_0 + \alpha_l l_{it} + \alpha_k k_{it} + \omega_{it} + \xi_{it}, \tag{2.1}$$

where va_{it}, l_{it}, k_{it} and ω_{it} denote the log of: value-added, labor input, capital stock at the beginning of period, and productivity, respectively, of firm i in period t, with ξ_{it} denoting the value-added shock to firm i in period t. In the conventional setup, a firm's productivity ω_{it} is assumed to be dependent on its lagged productivity $\omega_{i,t-1}$ via an unknown function $f(\cdot)$ as in Olley and Pakes (1996), Levinsohn and Petrin (2003), Ackerberg, Caves and Frazer (2015), and Wooldridge (2009). One might also consider a firm's current productivity to depend on the lagged characteristics $\mathbf{x}_{i,t-1}$ of the firm (such as its lagged exporting status and R&D expenditure) à la De Loecker (2013), Doraszelski and Jaumandreu (2013), and Braguinsky, Ohyama, Okazaki and Syverson (2015).

We generalize the conventional setup and further allow spatial dependence across firms. In particular, a firm i's current productivity is allowed to depend on the lagged output $y_{j,t-1}$ of its related firms j in the set $\mathcal{N}^{y}_{i,t-1}$, and the lagged inputs $\boldsymbol{\Omega}_{j,t-1} \equiv \{l_{j,t-1}, k_{j,t-1}, \mathbf{m}_{j,t-1}\}$ of a possibly different set $\mathcal{N}^{\Omega}_{i,t-1}$ of related firms, where $\mathbf{m}_{j,t-1}$ denotes the log of: $1 \times M$ vector of intermediate inputs of firm j in period $t-1$:

$$\omega_{it} = f(\omega_{i,t-1}) + \lambda \sum_{j \in \mathcal{N}^{y}_{i,t-1}} \mathtt{w}^{y}_{ij,t-1} y_{j,t-1} + \sum_{j \in \mathcal{N}^{\Omega}_{i,t-1}} \mathtt{w}^{\Omega}_{ij,t-1} \boldsymbol{\Omega}_{j,t-1} \boldsymbol{\beta}_{\bar{\Omega}} + \mathbf{x}_{i,t-1} \boldsymbol{\beta}_{x} + u_{it}, \tag{2.2}$$

where $y_{j,t-1}$ denotes the log of gross output of firm j in period $t-1$. Furthermore, the innovation u_{it} to the productivity of firm i in period t is allowed to be spatially correlated with those of related firms in the set $\mathcal{N}^{u}_{it}$ contemporarily:

$$u_{it} = \mu \sum_{j \in \mathcal{N}^{u}_{it}} \mathtt{w}^{u}_{ij,t} u_{jt} + v_{it}, \qquad i, j = 1, 2, 3, ..., N \text{ and } t = 2, 3, ..., T. \tag{2.3}$$

The weight assigned to each of the related firms in the set $\mathcal{N}^{y}_{i,t-1}$ is specified by $\mathtt{w}^{y}_{ij,t-1}$, and correspondingly those for firms in $\mathcal{N}^{\Omega}_{i,t-1}$ and $\mathcal{N}^{u}_{it}$ are specified by $\mathtt{w}^{\Omega}_{ij,t-1}$ and $\mathtt{w}^{u}_{ij,t}$, respectively. Note that the set of related firms that a firm's productivity depends upon can be defined by supplier-customer relationship, by ownership structure, by physical location, by industry of sales, or by combinations of them, and can differ across the three channels of spatial dependence, as the context of the study may deem appropriate.

As in the literature, it is assumed that firms observe their productivities before making production decisions, but the productivity is unknown to the econometricians and is the target of estimation (along with the production function parameters).

2.2.1 Assumptions

We adopt standard assumptions similar to those in the productivity estimation literature, but extend them to accommodate the setup with spatial dependence across firms as introduced in Equations (2.1)–(2.3).

Assumption 1. $E(\xi_{it}|l_{it}, k_{it}, \boldsymbol{m}_{it}) = 0$.

Assumption 2. $E(\xi_{it}|l_{jt}, k_{jt}, \boldsymbol{m}_{jt}, l_{j,t-1}, k_{j,t-1}, \boldsymbol{m}_{j,t-1}, \ldots, l_{j1}, k_{j1}, \boldsymbol{m}_{j1}) = 0$.

Assumption 3. $E(u_{it}|k_{it}, l_{j,t-1}, k_{j,t-1}, \boldsymbol{m}_{j,t-1}, \boldsymbol{x}_{j,t-1}, \ldots, l_{j1}, k_{j1}, \boldsymbol{m}_{j1}, \boldsymbol{x}_{j1}) = 0$.

Assumption 4. *The residuals, ξ_{it}, are assumed to be i.i.d. across both i and t, and have finite fourth moments: $\xi_{it} \overset{iid}{\sim} (0, \sigma^2_\xi),\ E|\xi_{it}^{4+\eta}| < \infty$, for some $\eta > 0$.*

Assumption 5. *The productivity innovations, u_{it}, are spatially correlated as specified in Equation (2.3), with residual v_{it} assumed to be i.i.d. across both i and t, and have finite fourth moments: $v_{it} \overset{iid}{\sim} (0, \sigma^2_v),\ E|v_{it}^{4+\eta}| < \infty$, for some $\eta > 0$.*

Assumption 6. *The residuals, ξ_{it} and v_{it}, are uncorrelated.*

Assumption 7. *$(\boldsymbol{I}_N - \mu \boldsymbol{W}_t^u)$ are non-singular for $t = 1, 2, \ldots, T$, with $\mu \in (\max_t \frac{1}{\lambda_{\min,t}}, \min_t \frac{1}{\lambda_{\max,t}})$, where $\boldsymbol{I}_N$ is the identity matrix of size N, $\boldsymbol{W}_t^u \equiv \{w_{ij,t}^u\}$, and $\lambda_{\min,t}$ and $\lambda_{\max,t}$ are the smallest and largest eigenvalues of $\boldsymbol{W}_t^u$.*

Assumption 8. *The row and column sums of the matrices, $\boldsymbol{W}_{t-1}^y, \boldsymbol{W}_{t-1}^\Omega, \boldsymbol{W}_t^u$ and $(\boldsymbol{I}_N - \mu \boldsymbol{W}_t^u)$ are uniformly bounded in absolute value for $t = 2, 3, \ldots, T$, where $\boldsymbol{W}_{t-1}^y \equiv \{w_{ij,t-1}^y\}$ and $\boldsymbol{W}_{t-1}^\Omega \equiv \{w_{ij,t-1}^\Omega\}$. The elements of the three spatial weight matrices are at most of order $\hbar_n^{-1}$ such that $\hbar_n/N \to 0$ as $N \to \infty$.*

Assumption 9. *The regressor matrix* $\{\mathbf{\Omega}_t, y_{t-1}, \mathbf{\Omega}_{t-1}, \boldsymbol{x}_{t-1}\}$ *has a full column rank, and the elements are uniformly bounded for* $t = 2, 3, \ldots, T$.

Assumption 1 is the standard assumption made in the literature for firm-level productivity estimations. Assumption 2 requires that the residuals ξ_{it} in the value-added equation (2.1) are conditionally mean independent of the current and past input usages of the firm itself, *and* also those of the other firms. This is not as stringent an assumption as it might appear, because the productivity term ω_{it} in Equation (2.1) has absorbed potential spatial dependence across firms to the extent modelled by Equation (2.2). Note that $\mathbf{m}_{it}$ and $\mathbf{m}_{jt}$ appear in the conditioning set of the first two assumptions, although only $\mathbf{m}_{j,t-1}$ appears in the model, because the inference of ω_{it} will utilize the information of $\mathbf{m}_{it}$ (as will become clear in Section 2.3.2). Assumption 3 basically states that the innovation u_{it} to productivity is conditionally mean independent of the state variable (capital), as well as the past input choices and characteristics of the firm itself and the other firms. Together, the first three assumptions will help identify the set of moment conditions and instruments for estimating the parameters in Equations (2.1) and (2.2). Assumptions 4–6 are made to develop the variance-covariance estimator of the parameters. In particular, the finite fourth moment condition for v_{it} is required for the estimation of the spatial parameter μ in Equation (2.3). Assumptions 7–9 are adopted from Kelejian and Prucha (1999), Kapoor, Kelejian and Pucha (2007), and Elhorst (2014) to ensure that the spatial parameter estimates exist. Note that we will construct the connectivity matrices such that they are row-normalized (with zeros in the diagonal by construction) and satisfy Assumption 8.

2.3 Estimation Algorithms

In this section, we propose a three-stage estimation procedure based on Generalized Method of Moments (GMM) to obtain consistent estimates of the parameters in Equations (2.1)–(2.3).

2.3.1 Moment Conditions

Given Assumptions 1–5, the following conditions hold with respect to the error terms in Equation (2.1) and Equation (2.2):

$$E(\xi_t | l_t, k_t, \mathbf{m}_t, \boldsymbol{\Omega}_{t-1}) = \mathbf{0}, \tag{2.4}$$

$$E(\xi_t + u_t | k_t, \boldsymbol{\Omega}_{t-1}, \mathbf{W}^y_{t-1} y_{t-1}, \mathbf{W}^\Omega_{t-1} \boldsymbol{\Omega}_{t-1}, \mathbf{x}_{t-1}) = \mathbf{0}, \tag{2.5}$$

where $\xi_t \equiv (\xi_{1t}, ..., \xi_{Nt})'$ and $u_t \equiv (u_{1t}, ..., u_{Nt})'$ denote the $N \times 1$ vector of the residual terms from Equation (2.1) and Equation (2.2), respectively, across firms in period t; $l_t \equiv (l_{1t}, ..., l_{Nt})'$ denotes the $N \times 1$ vector of labor inputs across firms in period t; k_t and $\mathbf{m}_t$ are similarly defined; $\boldsymbol{\Omega}_{t-1} \equiv [l_{t-1}\, k_{t-1}\, \mathbf{m}_{t-1}]$; $y_{t-1} \equiv (y_{1,t-1}, ..., y_{N,t-1})'$ denotes the $N \times 1$ vector of gross outputs across firms in period $t-1$; $\mathbf{x}_{t-1}$ is similarly defined. The matrices $\mathbf{W}^y_{t-1} \equiv \{\mathrm{w}^y_{ij,t-1}\}$ and $\mathbf{W}^\Omega_{t-1} \equiv \{\mathrm{w}^\Omega_{ij,t-1}\}$ are $N \times N$ connectivity matrices in period $t-1$ that specify the dependence of firm i's productivity in period t on related firms j's lagged outputs and lagged inputs, respectively. Note that the conditional mean is defined element (firm) wise in each period t. This set of conditionally mean independent conditions will lead to the moment conditions specified below in Equation (2.16).

Furthermore, by Kelejian and Prucha (1999) and Kapoor, Kelejian and Prucha (2007), the following three moment conditions hold with respect to the error term in Equation (2.3):

$$E\begin{bmatrix} \frac{1}{N} v_t' v_t \\ \frac{1}{N} v_t' \mathbf{W}^{u\prime}_t \mathbf{W}^u_t v_t \\ \frac{1}{N} v_t' \mathbf{W}^u_t v_t \end{bmatrix} = \begin{bmatrix} \sigma^2_v \\ \frac{\sigma^2_v}{N} \mathrm{tr}(\mathbf{W}^{u\prime}_t \mathbf{W}^u_t) \\ 0 \end{bmatrix}. \tag{2.6}$$

where $v_t \equiv (v_{1t}, ..., v_{Nt})'$ denotes the $N \times 1$ vector of the residual term from Equation (2.3) across firms in period $t = 2, 3, \ldots, T$.

2.3.2 Estimation Strategy

Stage 1

This stage basically follows the productivity estimation literature (e.g., Levinsohn and Petrin, 2003; Wooldridge, 2009). The productivity ω_{it} is assumed to be observable to the firm (or managers of the firm), but not to the econometrician. However, since a firm would in theory choose the optimal level of intermediate input $\mathbf{m}_{it}$ to maximize profits, given its initial capital stock k_{it}, labor force l_{it}, and realized productivity level ω_{it}, the econometrician could invert the relationship to infer a firm's productivity given its initial capital stock and observed input choices:

$$\omega_{it} = g(l_{it}, k_{it}, \mathbf{m}_{it}), \tag{2.7}$$

where $g(\cdot,\cdot,\cdot)$ is some unknown general function in the observed input levels. Equation (2.7), together with Equation (2.1), imply the following reduced-form value-added function:

$$\begin{aligned} va_t &= \alpha_0 \iota_N + \alpha_l l_t + \alpha_k k_t + \omega_t + \xi_t \\ &= \alpha_0 \iota_N + \alpha_l l_t + \alpha_k k_t + \mathbf{g}(l_t, k_t, \mathbf{m}_t) + \xi_t \\ &\equiv \mathbf{h}(l_t, k_t, \mathbf{m}_t) + \xi_t, \end{aligned} \tag{2.8}$$

where $va_t \equiv (va_{1t}, ..., va_{Nt})'$ denotes the $N \times 1$ vector of value-added across firms in period t; and ι_N is a $N \times 1$ vector of one's. The shorthand $\mathbf{g}(l_t, k_t, \mathbf{m}_t)$ is a $N \times 1$ column vector with $g(l_{it}, k_{it}, \mathbf{m}_{it})$ as its i-th entry; similarly, $\mathbf{h}(l_t, k_t, \mathbf{m}_t)$ is a $N \times 1$ column vector with $h(l_{it}, k_{it}, \mathbf{m}_{it})$ as its i-th entry, where $h(l_{it}, k_{it}, \mathbf{m}_{it}) \equiv \alpha_0 + \alpha_l l_{it} + \alpha_k k_{it} + g(l_{it}, k_{it}, \mathbf{m}_{it})$. As in Ackerberg, Caves and Frazer (2015) and Wooldridge (2009), one could approximate $h(\cdot,\cdot,\cdot)$ in Equation (2.8) by a n-degree polynomial that contain at least l_{it}, k_{it} and $\mathbf{m}_{it}$. For example, in the case where $\mathbf{m}_{it}$ contains only one type of intermediate input and is hence a scalar, $h(l_{it}, k_{it}, m_{it})$ can be approximated by

$\sum_{p,q,r} \left(l_{it}^p k_{it}^q m_{it}^r\right) \delta_{p,q,r}$, with nonnegative integers p, q and r such that $p + q + r \leq n$. That is:

$$h(l_{it}, k_{it}, \mathbf{m}_{it}) = \alpha_0 + c(l_{it}, k_{it}, \mathbf{m}_{it})\boldsymbol{\delta}, \tag{2.9}$$

where $c(l_{it}, k_{it}, \mathbf{m}_{it})$ is a $1 \times Q$ vector of functions in $(l_{it}, k_{it}, \mathbf{m}_{it})$ and $\boldsymbol{\delta}$ a $Q \times 1$ vector of parameters. For example, for a 2nd-order polynomial h function ($n = 2$), $c(l_{it}, k_{it}, \mathbf{m}_{it}) = [l_{it}, k_{it}, m_{it}, l_{it}^2, l_{it}k_{it}, l_{it}m_{it}, k_{it}^2, k_{it}m_{it}, m_{it}^2]$.

Given the condition (2.4), Equation (2.8) given Equation (2.9) can be estimated using the following set of instrument variables (IVs) for period t:

$$\mathcal{Z}_{t,I} = (\iota_N, \mathbf{c}_t, \mathbf{c}_{t-1}), \tag{2.10}$$

where the shorthand $\mathbf{c}_t$ is a $N \times Q$ matrix with $c(l_{it}, k_{it}, \mathbf{m}_{it})$ as its i-th row entry. Note that since $g(\cdot, \cdot, \cdot)$ is allowed to be a general function (including linearity in the arguments as a special case), the slope coefficients (α_l, α_k) on the inputs are not identified from Equation (2.8), as highlighted by Ackerberg, Caves and Frazer (2015). However, it enables an estimate $\hat{h}(l_{it}, k_{it}, \mathbf{m}_{it})$ of $h(l_{it}, k_{it}, \mathbf{m}_{it})$. In turn, the slope coefficients of the production function can be identified in a later stage, along with the other parameters, as laid out in the next section.

The set of IVs listed in (2.10)—and in the condition (2.4)—includes the input variables only up to one lag, and hence corresponds to weaker conditions than stated in Assumption 2. One could potentially enlarge the set and include longer lags of the input variables in the conditioning set, given Assumption 2.

Stage 2

Next, given the productivity process's dynamic and spatial dependence structure specified in Equation (2.2), and Equation (2.7), we can also write the value-added function in the following alternative reduced form:

$$\begin{aligned} va_t &= \alpha_0\iota_N + \alpha_l l_t + \alpha_k k_t + \omega_t + \xi_t \\ &= \alpha_0\iota_N + \alpha_l l_t + \alpha_k k_t + \mathbf{f}[\mathbf{g}(l_{t-1}, k_{t-1}, \mathbf{m}_{t-1})] \\ &\quad + \lambda \mathbf{W}^y_{t-1} y_{t-1} + \mathbf{W}^\Omega_{t-1}\boldsymbol{\Omega}_{t-1}\boldsymbol{\beta}_{\bar{\Omega}} + \mathbf{x}_{t-1}\boldsymbol{\beta}_x + u_t + \xi_t, \end{aligned} \tag{2.11}$$

where the shorthand $\mathbf{f}[\mathbf{g}(l_{t-1}, k_{t-1}, \mathbf{m}_{t-1})]$ is a $N \times 1$ column vector with $f[g(l_{i,t-1}, k_{i,t-1}, \mathbf{m}_{i,t-1})]$ as its i-th entry. Recall that the matrices $\mathbf{W}^y_{t-1} \equiv \{\mathtt{w}^y_{ij,t-1}\}$ and $\mathbf{W}^\Omega_{t-1} \equiv \{\mathtt{w}^\Omega_{ij,t-1}\}$ are $N \times N$ connectivity matrices in period $t-1$ that specify the dependence of firm i's productivity in period t on related firms j's lagged outputs and lagged inputs, respectively. In deriving Equation (2.11), we have used Equation (2.2) to replace ω_t and Equation (2.7) to replace $\omega_{i,t-1}$ in the $f(\cdot)$ function such that $f(\omega_{i,t-1}) = f[g(l_{i,t-1}, k_{i,t-1}, \mathbf{m}_{i,t-1})]$. As suggested by Wooldridge (2009), one could use a G-th degree polynomial to approximate $f(\cdot)$ such that:

$$f(\nu) = \rho_1\nu + \rho_2\nu^2 + \ldots + \rho_G\nu^G. \tag{2.12}$$

Given the condition in (2.5), Equation (2.11) can be estimated using the following set of IVs for period t:

$$\mathcal{Z}_{t,II} = (\iota_N, k_t, \mathbf{c}_{t-1}, \mathbf{W}^y_{t-1} y_{t-1}, \mathbf{W}^\Omega_{t-1}\boldsymbol{\Omega}_{t-1}, \mathbf{x}_{t-1}), \tag{2.13}$$

with one lag (or a longer past history) of the variables. Additional spatio-temporal lags of explanatory variables, such as $(\mathbf{W}^y_{t-1})^2 y_{t-1}$, $(\mathbf{W}^y_{t-1})^3 y_{t-1}$, $(\mathbf{W}^\Omega_{t-1})^2\boldsymbol{\Omega}_{t-1}$ and $(\mathbf{W}^\Omega_{t-1})^3\boldsymbol{\Omega}_{t-1}$, may also be added to the set of IVs to help identify the spatial coefficients.

While Ackerberg, Caves and Frazer (2015) propose to estimate Equations (2.8) and (2.11)—without the spatial structure—sequentially, by plugging in estimates from Equation (2.8) into Equation (2.11), we adopt the approach proposed by Wooldridge (2009) and estimate them jointly, as it leads to more efficient estimators. In particular, denote the parameters of the system by $\boldsymbol{\theta} = (\alpha_0, \boldsymbol{\delta}', \alpha_l, \alpha_k, \lambda, \boldsymbol{\beta}'_{\bar{\Omega}}, \boldsymbol{\beta}'_x, \rho_1, ..., \rho_G)'$. The residuals from Equations (2.8) and (2.11) given the parameters are, respectively:

$$
\begin{aligned}
r_{t,I}(\boldsymbol{\theta}) &= va_t - \alpha_0 \iota_N - \mathbf{c}_t \boldsymbol{\delta}, && (2.14)\\
r_{t,II}(\boldsymbol{\theta}) &= va_t - \alpha_0 \iota_N - \alpha_l l_t - \alpha_k k_t - \mathbf{f}[\mathbf{c}_{t-1}\boldsymbol{\delta} - \alpha_l l_{t-1} - \alpha_k k_{t-1}] \\
&\quad -\lambda \mathbf{W}^y_{t-1} y_{t-1} - \mathbf{W}^{\Omega}_{t-1} \boldsymbol{\Omega}_{t-1} \boldsymbol{\beta}_{\bar{\Omega}} - \mathbf{x}_{t-1} \boldsymbol{\beta}_x, && (2.15)
\end{aligned}
$$

where recall that $g(l_{i,t-1}, k_{i,t-1}, \mathbf{m}_{i,t-1}) = h(l_{i,t-1}, k_{i,t-1}, \mathbf{m}_{i,t-1}) - \alpha_0 - \alpha_l l_{i,t-1} - \alpha_k k_{i,t-1} = c(l_{i,t-1},\, k_{i,t-1}, \mathbf{m}_{i,t-1})\boldsymbol{\delta} - \alpha_l l_{i,t-1} - \alpha_k k_{i,t-1}$, given Equations (2.8) and (2.9). The conditions in (2.4) and (2.5) imply that:

$$
E[\mathcal{Z}'_{it}\, r_{it}(\boldsymbol{\theta})] \equiv E\left[\begin{pmatrix} \mathcal{Z}'_{it,I} & \mathbf{0} \\ \mathbf{0} & \mathcal{Z}'_{it,II} \end{pmatrix} \begin{pmatrix} r_{it,I}(\boldsymbol{\theta}) \\ r_{it,II}(\boldsymbol{\theta}) \end{pmatrix} \right] = \mathbf{0}, \tag{2.16}
$$

where $\mathcal{Z}_{it,I}$, $\mathcal{Z}_{it,II}$, $r_{it,I}(\boldsymbol{\theta})$, and $r_{it,II}(\boldsymbol{\theta})$ are the i-th row entry of $\mathcal{Z}_{t,I}$, $\mathcal{Z}_{t,II}$, $r_{t,I}(\boldsymbol{\theta})$, and $r_{t,II}(\boldsymbol{\theta})$, respectively. Given Equation (2.16), we proceed with GMM estimation of $\boldsymbol{\theta}$.

Stage 3

We estimate the spatial error structure in Equation (2.3) based on the GMM approach of Kelejian and Prucha (1999) and Kapoor, Kelejian and Prucha (2007). Specifically, given the parameter estimates $\hat{\boldsymbol{\theta}}$ from the previous stages, we impute estimates of the productivity innovation term, $\hat{u}_t$, by taking the difference between (2.15) and (2.14), since the residuals from the second stage is $\widehat{\xi_t + u_t}$ and the residuals from the first stage is $\widehat{\xi_t}$:

$$
\begin{aligned}
\hat{u}_t \equiv\ & \mathbf{c}_t \hat{\boldsymbol{\delta}} - \hat{\alpha}_l l_t - \hat{\alpha}_k k_t \\
& -\mathbf{f}[\mathbf{c}_{t-1} \hat{\boldsymbol{\delta}} - \hat{\alpha}_l l_{t-1} - \hat{\alpha}_k k_{t-1}] \\
& -\hat{\lambda} \mathbf{W}^y_{t-1} y_{t-1} - \mathbf{W}^{\Omega}_{t-1} \boldsymbol{\Omega}_{t-1} \hat{\boldsymbol{\beta}}_{\bar{\Omega}} - \mathbf{x}_{t-1} \hat{\boldsymbol{\beta}}_x,
\end{aligned} \tag{2.17}
$$

and use the moment conditions implied by Equation (2.6) to estimate μ and σ_v^2 jointly by GMM. Note that if we define $\overline{u}_t \equiv \mathbf{W}_t^u u_t$, $\overline{v}_t \equiv \mathbf{W}_t^u v_t$, and $\overline{\overline{u}}_t = (\mathbf{W}_t^u)^2 u_t$, it follows that $v_t = u_t - \mu\overline{u}_t$ and $\overline{v}_t = \overline{u}_t - \mu\overline{\overline{u}}_t$. By replacing v_t in the moment condition (2.6) with $u_t - \mu\overline{u}_t$ rewrites the three moment conditions in terms of u_t and $\mu\overline{u}_t$. We can follow similar steps as in Kelejian and Prucha (1999) and Kapoor, Kelejian and Prucha (2007) to derive Equation (2.18) below:

$$\frac{1}{T-1}\sum_{t=2}^{T}\left(\boldsymbol{\gamma}_t - \boldsymbol{\Gamma}_t\begin{pmatrix}\mu\\ \mu^2\\ \sigma_v^2\end{pmatrix}\right) = \mathbf{0}, \tag{2.18}$$

where

$$\boldsymbol{\gamma}_t = \frac{1}{N}\begin{pmatrix}E(u_t'u_t)\\ E(\overline{u}_t'\overline{u}_t)\\ E(u_t'\overline{u}_t)\end{pmatrix}, \tag{2.19}$$

$$\boldsymbol{\Gamma}_t = \frac{1}{N}\begin{pmatrix}2E(u_t'\overline{u}_t) & -E(\overline{u}_t'\overline{u}_t) & N\\ 2E(\overline{\overline{u}}_t'\overline{u}_t) & -E(\overline{\overline{u}}_t'\overline{\overline{u}}_t) & \mathrm{tr}(\mathbf{W}_t^{u\prime}\mathbf{W}_t^u)\\ E(u_t'\overline{\overline{u}}_t + \overline{u}_t'\overline{u}_t) & -E(\overline{u}_t'\overline{\overline{u}}_t) & 0\end{pmatrix}. \tag{2.20}$$

Use the estimates of the productivity innovation term from Equation (2.17), $\hat{u}_t$, to construct the sample counterparts of the $\boldsymbol{\gamma}_t$ vector and the $\boldsymbol{\Gamma}_t$ matrix:[1]

$$\varsigma_t \equiv \frac{1}{N}\begin{pmatrix} \hat{u}_t{}'\hat{u}_t \\ \hat{u}_t{}'\mathbf{W}_t^{u\prime}\mathbf{W}_t^u\hat{u}_t \\ \hat{u}_t{}'\mathbf{W}_t^u\hat{u}_t \end{pmatrix}, \tag{2.21}$$

$$\boldsymbol{F}_t \equiv \frac{1}{N}\begin{pmatrix} 2\hat{u}_t{}'\mathbf{W}_t^u\hat{u}_t & -\hat{u}_t{}'\mathbf{W}_t^{u\prime}\mathbf{W}_t^u\hat{u}_t & N \\ 2\hat{u}_t{}'\mathbf{W}_t^{u\prime}\mathbf{W}_t^{u\prime}\mathbf{W}_t^u\hat{u}_t & -\hat{u}_t{}'\mathbf{W}_t^{u\prime}\mathbf{W}_t^{u\prime}\mathbf{W}_t^u\mathbf{W}_t^u\hat{u}_t & \mathrm{tr}(\mathbf{W}_t^{u\prime}\mathbf{W}_t^u) \\ \hat{u}_t'\mathbf{W}_t^u\mathbf{W}_t^u\hat{u}_t + \hat{u}_t'\mathbf{W}_t^{u\prime}\mathbf{W}_t^u\hat{u}_t & -\hat{u}_t'\mathbf{W}_t^{u\prime}\mathbf{W}_t^u\mathbf{W}_t^u\hat{u}_t & 0 \end{pmatrix}, \tag{2.22}$$

and form the sample counterpart of the condition in Equation (2.18):

$$\frac{1}{T-1}\sum_{t=2}^{T}\left(\varsigma_t - \boldsymbol{F}_t\begin{pmatrix} \mu \\ \mu^2 \\ \sigma_v^2 \end{pmatrix}\right) \equiv \frac{1}{T-1}\sum_{t=2}^{T}\epsilon_t, \tag{2.23}$$

where ϵ_t is a 3×1 vector of residuals. We can then estimate μ and σ_v^2 by the transformed moment condition $E(\epsilon_t) = \mathbf{0}$. Specifically,

$$\epsilon_t(\mu,\sigma_v^2) = \frac{1}{N}\begin{pmatrix} \hat{u}_t{}'(\mathbf{I}_N - 2\mu\mathbf{W}_t^u + \mu^2\mathbf{W}_t^{u\prime}\mathbf{W}_t^u)\hat{u}_t - \sigma_v^2 N \\ \hat{u}_t{}'\mathbf{W}_t^{u\prime}(\mathbf{I}_N - 2\mu\mathbf{W}_t^{u\prime} + \mu^2\mathbf{W}_t^{u\prime}\mathbf{W}_t^u)\mathbf{W}_t^u\hat{u}_t - \sigma_v^2\mathrm{tr}(\mathbf{W}_t^{u\prime}\mathbf{W}_t^u) \\ \hat{u}_t{}'\left(\mathbf{I}_N - \mu(\mathbf{W}_t^u + \mathbf{W}_t^{u\prime}) + \mu^2\mathbf{W}_t^{u\prime}\mathbf{W}_t^u\right)\mathbf{W}_t^u\hat{u}_t \end{pmatrix}. \tag{2.24}$$

The algorithm above provides a set of estimates consistent for $\boldsymbol{\theta}$, μ and σ_v^2. We can improve the efficiency of the estimators by deriving the weighting matrix for the GMM estimator, and repeat the procedure until the parameter estimates converge. Section 2.3.3 characterizes the algorithm to obtain the efficient GMM estimator.

[1] The derivations are provided in Section 2.8.1 of the Theoretical Appendix.

2.3.3 Efficient GMM Estimator

This section itemizes the steps to implement the proposed estimation strategy and obtain the efficient GMM estimator of $\boldsymbol{\theta}$ and $\psi \equiv \{\mu, \sigma_v^2\}$.

1. Minimize the objective function: $\left[\frac{1}{N(T-1)} \sum_{i=1}^{N} \sum_{t=2}^{T} \mathcal{Z}_{it}' r_{it}(\boldsymbol{\theta})\right]' \boldsymbol{\mathcal{W}}_{\boldsymbol{\theta}} \left[\frac{1}{N(T-1)} \sum_{i=1}^{N} \sum_{t=2}^{T} \mathcal{Z}_{it}' r_{it}(\boldsymbol{\theta})\right]$ with respect to $\boldsymbol{\theta}$ by setting $\boldsymbol{\mathcal{W}}_{\boldsymbol{\theta}} = \mathbf{I}_{\mathcal{M}}$ to obtain the one-step estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$, where $\boldsymbol{\mathcal{W}}_{\boldsymbol{\theta}}$ of dimension $\mathcal{M} \times \mathcal{M}$ refers to the weighting matrix for the moment conditions used in the estimation of $\boldsymbol{\theta}$, and $\mathcal{M}$ is the combined number of moment conditions (instruments) from Stage 1 and Stage 2.

2. Given the one-step estimate of $\boldsymbol{\theta}$, obtain the residuals $\{\hat{u}_t\}_{t=2}^{T}$ by Equation (2.17). This in turn can be used to obtain an estimator of μ and σ_v^2 based on Equations (2.23) and (2.24):

$$\arg\min_{\mu,\sigma_v^2} \frac{1}{T-1} \sum_{t=2}^{T} \left(\varsigma_t - \boldsymbol{F}_t \begin{pmatrix} \mu \\ \mu^2 \\ \sigma^2 \end{pmatrix} \right)' \boldsymbol{\mathcal{W}}_{\psi} \frac{1}{T-1} \sum_{t=2}^{T} \left(\varsigma_t - \boldsymbol{F}_t \begin{pmatrix} \mu \\ \mu^2 \\ \sigma^2 \end{pmatrix} \right), \tag{2.25}$$

   by setting $\boldsymbol{\mathcal{W}}_{\psi} = \mathbf{I}_3$ to obtain the one-step estimator of μ and σ_v^2, where $\boldsymbol{\mathcal{W}}_{\psi}$ of dimension 3×3 refers to the weighting matrix for the moment conditions used in the estimation of ψ, and there are three moment conditions in this case.

3. Derive a variance-covariance estimator $\widehat{\mathbf{V}}_{\boldsymbol{\theta}}$ of the moment conditions used in the estimation

of $\boldsymbol{\theta}$, $\mathbf{V}_{\boldsymbol{\theta}} = \text{Var}\left(\frac{1}{\sqrt{N(T-1)}}\sum_{i=1}^{N}\sum_{t=2}^{T} Z'_{it}r_{it}\right)$, noting that:[2]

$$
\begin{aligned}
\mathbf{V}_{\boldsymbol{\theta}} &= \frac{1}{N(T-1)}\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{j=1}^{N}\sum_{s=2}^{T} E[(Z'_{it}r_{it})(Z'_{js}r_{js})'] \\
&= \frac{1}{N(T-1)}\sum_{i=1}^{N}\sum_{t=2}^{T}\begin{pmatrix} \mathcal{Z}'_{it,I}\mathcal{Z}_{it,I}E(\xi_{it}^2) & \mathcal{Z}'_{it,I}\mathcal{Z}_{it,II}E(\xi_{it}^2) \\ \mathcal{Z}'_{it,II}\mathcal{Z}_{it,I}E(\xi_{it}^2) & \mathcal{Z}'_{it,II}\mathcal{Z}_{it,II}E(\xi_{it}^2) \end{pmatrix} \\
&\quad + \frac{1}{N(T-1)}\sum_{t=2}^{T}\boldsymbol{\mathcal{Z}}'_t\begin{pmatrix} \mathbf{0}_{N\times N} & \mathbf{0}_{N\times N} \\ \mathbf{0}_{N\times N} & \sigma_v^2[(\mathbf{I}_N - \mu\mathbf{W}_t^u)^{-1}(\mathbf{I}_N - \mu\mathbf{W}_t^u)^{-1\prime}] \end{pmatrix}\boldsymbol{\mathcal{Z}}_t
\end{aligned} \tag{2.26}
$$

by replacing the residuals ξ_t with the sample counterpart $\widehat{\xi}_t$, and the parameters (μ and σ_v^2) with their estimates obtained from Steps 1–2 above.

4. Repeat Step 1, but update the weighting matrix by $\boldsymbol{\mathcal{W}}_{\boldsymbol{\theta}} = \widehat{\mathbf{V}}_{\boldsymbol{\theta}}^{-1}$.

5. To obtain an estimate of the variance-covariance matrix for the moment conditions used in the estimation of $\boldsymbol{\psi}$, we extend the framework of Kapoor, Kelejian and Prucha (2007) to allow for non-normal errors and time-varying connectivity matrices.[3] The variance-covariance matrix for the sample counterpart of the left side of the moment conditions in Equation (2.6) is given by:

$$
\mathbf{V}_{\psi} = \begin{pmatrix} V_{\psi,11} & V_{\psi,12} & 0 \\ V_{\psi,21} & V_{\psi,22} & V_{\psi,23} \\ 0 & V_{\psi,32} & V_{\psi,33} \end{pmatrix}, \tag{2.27}
$$

where $V_{\psi,11} = \sigma_v^4(\kappa_v + 2)$; $V_{\psi,12} = \frac{\sigma_v^4}{N(T-1)}(\kappa_v + 2)\text{tr}(\mathbf{W}^{u\prime}\mathbf{W}^u)$; $V_{\psi,21} = V_{\psi,12}$; $V_{\psi,22} = \frac{\sigma_v^4}{N(T-1)}[\kappa_v\ \text{diagv}(\mathbf{W}^{u\prime}\mathbf{W}^u)'\text{diagv}(\mathbf{W}^{u\prime}\mathbf{W}^u) + \text{tr}(\mathbf{W}^{u\prime}\mathbf{W}^u(\mathbf{W}^{u\prime}\mathbf{W}^u + \mathbf{W}^u\mathbf{W}^{u\prime}))]$; $V_{\psi,23} = \frac{\sigma_v^4}{N(T-1)}\text{tr}((\mathbf{W}^{u\prime}\mathbf{W}^u)(\mathbf{W}$ $\mathbf{W}^{u\prime}))$; $V_{\psi,32} = V_{\psi,23}$; and $V_{\psi,33} = \frac{\sigma_v^4}{N(T-1)}\text{tr}(\mathbf{W}^u(\mathbf{W}^u + \mathbf{W}^{u\prime}))$. κ_v is the excess kurtosis of v_{it}, and $\mathbf{W}^u$ is a $N(T-1)\times N(T-1)$ block-diagonal matrix with $\mathbf{W}_2^u, \mathbf{W}_3^u, \ldots, \mathbf{W}_T^u$ on

[2]The derivations of $\mathbf{V}_{\boldsymbol{\theta}}$ are provided in Section 2.8.2 of the Theoretical Appendix.
[3]The derivations of $\mathbf{V}_{\psi}$ are provided in Section 2.8.3 of the Theoretical Appendix.

the diagonal; the operator 'diagv' takes the diagonal elements of a matrix and converts them to a column vector.

The excess kurtosis κ_v can be estimated using the following formula given the estimates of μ and σ_v^2:

$$\hat{\kappa}_v = \frac{\sum_{i=1}^N \sum_{t=2}^T (\hat{v}_{it} - \frac{1}{N(T-1)} \sum_{i=1}^N \sum_{t=2}^T \hat{v}_{it})^4}{N(T-1)\hat{\sigma}_v^4} - 3, \tag{2.28}$$

where $\hat{v}_{it}$ is the i-th element of $\hat{v}_t = (\mathbf{I}_N - \hat{\mu}\mathbf{W}_t^u)\hat{u}_t$ and $\hat{\sigma}_v^4 = (\hat{\sigma}_v^2)^2$.

6. Repeat Step 2, this time setting $\boldsymbol{\mathcal{W}}_\psi = \widehat{\mathbf{V}}_\psi^{-1}$.

7. Repeat Steps 3–6 until convergence in the estimates: $\widehat{\boldsymbol{\theta}}$, $\hat{\mu}$, $\hat{\sigma}_v^2$, $\widehat{\mathbf{V}}_{\boldsymbol{\theta}}$, and $\widehat{\mathbf{V}}_\psi$.

8. Obtain a variance-covariance matrix estimator of the parameters $\boldsymbol{\theta}$ based on the asymptotic property established for the efficient GMM estimator (Lee and Yu, 2014):

$$\sqrt{N(T-1)}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \sim \boldsymbol{\mathcal{N}}(0, \text{plim}_{N,T\to\infty}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}), \tag{2.29}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = (\boldsymbol{\mathcal{Z}}_\Delta' \mathbf{V}_{\boldsymbol{\theta}}^{-1} \boldsymbol{\mathcal{Z}}_\Delta)^{-1}$ and $\boldsymbol{\mathcal{Z}}_\Delta = E\left[\frac{d}{d\boldsymbol{\theta}'} Z_{it}' r_{it}(\boldsymbol{\theta})\right]$. The sample counterpart is correspondingly:

$$\widehat{\boldsymbol{\mathcal{Z}}}_\Delta = \frac{1}{N(T-1)} \sum_{i=1}^N \sum_{t=2}^T \left[\frac{d}{d\boldsymbol{\theta}'} Z_{it}' r_{it}(\boldsymbol{\theta})\right].$$

By the Slutsky theorem, we have $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}} \equiv (\widehat{\boldsymbol{\mathcal{Z}}}_\Delta' \widehat{\mathbf{V}}_{\boldsymbol{\theta}}^{-1} \widehat{\boldsymbol{\mathcal{Z}}}_\Delta)^{-1} \xrightarrow{p} (\boldsymbol{\mathcal{Z}}_\Delta' \mathbf{V}_{\boldsymbol{\theta}}^{-1} \boldsymbol{\mathcal{Z}}_\Delta)^{-1} \equiv \boldsymbol{\Sigma}_{\boldsymbol{\theta}}$.

9. Similarly, the variance-covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\psi}}$ (*in the original scale*) for the parameters $\boldsymbol{\psi}$ can be estimated by:

$$\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\psi}} = \frac{1}{N(T-1)} (\widehat{\mathcal{G}}_\Delta' \widehat{\mathbf{V}}_\psi^{-1} \widehat{\mathcal{G}}_\Delta)^{-1}, \tag{2.30}$$

where

$$\tag{2.31}$$

$$\widehat{\mathcal{G}}_{\Delta} = \frac{1}{(T-1)}\sum_{t=2}^{T}\frac{d\epsilon_t(\widehat{\boldsymbol{\psi}})}{d\boldsymbol{\psi}'} \tag{2.32}$$

$$= \frac{1}{N(T-1)}\sum_{t=2}^{T}\begin{pmatrix} 2\hat{u}_t{}'(\widehat{\mu}\mathbf{W}_t^{u\prime}-\mathbf{I}_N)\mathbf{W}_t^u\hat{u}_t & -N \\ 2\hat{u}_t{}'\mathbf{W}_t^{u\prime}(\widehat{\mu}\mathbf{W}_t^{u\prime}\mathbf{W}_t^u-\mathbf{W}_t^{u\prime})\mathbf{W}_t^u\hat{u}_t & -\mathrm{tr}(\mathbf{W}_t^{u\prime}\mathbf{W}_t^u) \\ \hat{u}_t{}'(2\widehat{\mu}\mathbf{W}_t^{u\prime}\mathbf{W}_t^u-(\mathbf{W}_t^u+\mathbf{W}_t^{u\prime}))\mathbf{W}_t^u\hat{u}_t & 0 \end{pmatrix}.$$

2.4 Data

Our dataset is constructed by combining two Japanese datasets. The first dataset is the Basic Survey of Japanese Business Structure and Activities (BSJBSA), provided by the Ministry of Economy, Trade and Industry (METI), Japan. The data include a firm-level annual survey of detailed business information, such as sales, employment, capital stock, industry classification (Japan Standard Industry Classification, JSIC) and intermediate purchases. The data cover both manufacturing and non-manufacturing firms that have: (1) more than 50 employees, and (2) capital stocks of more than 30 million yens (approximately 250 thousand USD in 2015).

The second dataset contains information on firm-to-firm relationship provided by Tokyo Shoko Research (TSR), a major credit reporting company in Japan. It provides a firm's most important domestic suppliers and customers (up to 24 connections in each direction) and covers both publicly listed and unlisted firms in Japan of all sizes and industries. Because these two datasets do not use the same firm identification codes, we match them on the basis of firm name, address, phone number, and postal code. Using the BSJBSA as the denominator (since it provides the required firm-level variables for productivity estimations), the percentage of firms in BSJBSA that are matched with its counterpart in TSR is very high, typically at 93%–94%, across years during the sample period 2009–2018. Table 2.1 provides the detailed firm counts.

We supplement the BSJBSA-TSR linked dataset with the JIP database 2021 provided by the

Research Institute of Economy, Trade and Industry (RIETI), and with the information on commuting zones (CZs) constructed by Adachi, Fukai, Kawaguchi and Saito (2021). We impute the industry-level price deflators based on the JIP database. In particular, it contains the nominal and real values of outputs, intermediate inputs, investment, and value added for the 100 industries classified by the JIP database. We construct the deflators by the ratios of the nominal and real values for each of these variables, and merge them with the BSJBSA data (based on concordance between the BSJBSA JSIC industries and the JIP industries, provided in the JIP database). There are in total 433 JSIC 3-digit industries. A JIP industry is matched on average with 4.8 JSIC 3-digit industries. These deflators are then used to convert the BSJBSA corresponding variables into real terms. We also impute the average work hours per person in a year in an industry based on the JIP database and merge the variable with the BSJBSA data (using again the concordance between the JIP and JSIC industries).

The information on CZs is used as one criterion below in defining connectivity matrices across firms. Adachi, Fukai, Kawaguchi and Saito (2021) construct the CZ information for Japan using the hierarchical agglomerative clustering (HAC) method of Tolbert and Killian (1987) and Tolbert and Sizer (1996) for constructing the US CZs. By Adachi et al. (2021), there are 267 CZs in 2010 and 265 in 2015 in Japan. We use the 2010 CZ definition, and merge the CZ information with the BSJBSA observations based on the prefecture and city names of a firm's address, and examine/adjust manually if necessary.

The variables at the firm-year level used for the production function estimation is constructed in the following manners. The number of workers is constructed as the sum of regular workers and part-time workers (excluding temporary workers) in headquarter, head office, branch office, and assignee company (available from the BSJBSA). The labor hour is constructed as the number of workers (from the BSJBSA) of a firm in a year, times the average work hours per person in the corresponding year and industry (from the JIP database). The real physical capital stock is constructed by the perpetual inventory method with 2007 as the base year, using the real physical capital value in 2007 (year end) and the real investment in physical capital in each year from the BSJBSA, together with the depreciation rates at the industry level from the JIP database. The real intermediate inputs are constructed by the sum of the cost of goods sold, and general and administrative expenses, minus wages, rental costs, depreciation, and taxes reported in the BSJBSA, deflated by the input deflator (constructed using the JIP database as documented above). The real revenue is measured by sales, deflated by the output deflator imputed from the JIP database. The real value added is constructed by nominal value-added (i.e., nominal sales minus nominal intermediate inputs) deflated by the value-added deflator from the JIP database.

2.4.1 Summary Statistics

Table 2.2(a) provides the summary statistics of the key variables for the BSJBSA-TSR linked sample in year 2015 (based on the denominator of BSJBSA firms, not all of which have corresponding entries in TSR). The effective number of observations differs from Table 2.1 due to potentially missing observations on the variable of interest.

A few remarks are in order. First, the average firms tend to be large (e.g., having 490 workers, and 7.2 billion JPY physical capital, roughly equivalent to 60 million USD). This is due to the fact that the BSJBSA only covers medium and large firms. Second, the average firms report 6.8 customers and 6.7 suppliers, suggesting that the TSR's limit of reporting the top suppliers and customers up to 24 connections in each direction is not practically binding for most of the firms.

Figures 2.1(a)–2.3(a) illustrate the number of firms, their average size in terms of employment, and their average number of customers and suppliers for each 1-digit JSIC rev12 industry. Figure 2.1(a) indicates that most of the firms in the sample are in the manufacturing, wholesale & retail trade, and information & communications industries. Among them, those in the service industries tend to be large in terms of employment (e.g., accommodation and food & beverage services, electricity, gas, heat supply & water, and finance & insurance; see Figure 2.2(a)). In terms of connectedness with other firms, those in the construction, manufacturing, and mining & quarrying industries tend to have a larger number of customers and suppliers, while those in the service industries tend to have a smaller number of business customers but just as many suppliers as in other industries

(Figure 2.3(a)).

Figures 2.4(a)–2.6(a) show a large heterogeneity across prefectures in terms of the number of firms, their average size, and their average number of customers and suppliers. Most of the firms in the sample are located in economically large prefectures, such as Tokyo, Osaka, Aichi, Kanagawa, and Hyogo (Figure 2.4(a)). Firms in these large prefectures tend also to be large in terms of employment (Figure 2.5(a)) and connected with a larger number of both customers and suppliers (Figure 2.6(a)).

The potential spatial dependence across firms through the supply chain and the input markets can be more local than the prefecture level. Bernard, Moxnes and Saito (2019) show that the median distance of any customer-supplier pair in the TSR data is 30 km and thus smaller than the typical size of prefectures. Thus, the following figures further provide the characterization at the commuting zone level. Figure 2.7(a) shows the number of CZs within each prefecture in 2015. Prefectures with large areas (e.g., Hokkaido and Nagano) tend to have many CZs, while those with small areas and economic sizes (e.g., Kagawa and Fukui) tend to have few CZs. Figures 2.8(a)–2.10(a) show the counterparts of Figures 2.4(a)–2.6(a) at the commuting-zone level. The commuting zone with the largest number of firms (8993) is CZ89 that covers the busiest areas around Tokyo (parts of Tokyo, Kanagawa, Chiba, and Saitama). Economically large CZs also tend to have larger firms and more connected firms. For example, the same CZ89 ranks top 5th in terms of average firm's employment size (643.99), and top 24th in terms of average firm's customer connections (7.44). Note also that the average firm size is much more dispersed at the right tail when we zoom in at the commuting-zone level (Figure 2.9(a)) compared to that at the prefecture level (Figure 2.5(a)). Similarly, the distributions of supplier/customer connections are much more dispersed at the commuting-zone level (Figure 2.10(a)) than at the prefecture level (Figure 2.6(a)), suggesting a large degree of heterogeneity across CZs within prefectures.

2.4.2 Definition of Connectivity Matrices

To model the spatial-temporal lag dependence in outputs, we define the output connectivity matrix $\mathbf{W}^y_{t-1}$ based on the set of a firm's customers and/or suppliers located in the same commuting zone. The ij-th element of $\mathbf{W}^y_{t-1}$ takes on the value one if both firms i and j are located in the same commuting zone, and in addition, firm j is a customer or supplier of firm i, in period $t-1$. This is in line with the research conducted by Demir, Fieler, Xu and Yang (2024), Alfaro-Ureña, Manelici and Vasquez (2022) and Ellison, Glaeser and Kerr (2010). They find that the input-output linkages across firms and the geographical proximity of firms play an important role in productivity spillovers.

Next, for spatial dependence across firms through the input markets, we restrict our focus to the labor market channel, and define input connectivity matrix $\mathbf{W}^\Omega_{t-1}$ such that the ij-th element of $\mathbf{W}^\Omega_{t-1}$ takes on the value one if firms i and j are located in the same commuting zone in period $t-1$. Firms located in the same commuting zone are more likely to tap into the same labor pool, considering the potential labor mobility frictions across zones. Multiple theories have been proposed about the benefits associated with a large labor pool. When firms in the same location employ more workers, the potential pool of labor in the location increases. This facilitates better worker-firm matches (e.g., Helsley and Strange, 1990); allows risk sharing and worker turnover across firms (e.g., Diamond and Simon, 1990; Krugman, 1991); and induces stronger incentives for workers to invest in human capital knowing that they do not face ex post appropriation (Rotemberg and Saloner, 2000). As a result, conditional on the amount of labor input hired by a firm, the quality of labor input (and hence firm productivity) is likely higher when the total labor employed in the same location in the past period is larger. Relatedly, Greenstone, Hornbeck and Moretti (2010) find that estimated spillover effects resulting from the opening of Million Dollar Plants are larger for other plants that share labor pools and similar technologies with the new plant.

To model the spatial diffusion of the productivity shock u_t, we consider three variants of the connectivity matrix $\mathbf{W}^u_t$, depending on whether two firms are located in the same commuting zone, whether they have supplier/customer relationships, or both. In particular, the ij-th element of $\mathbf{W}_t^u$ takes on the value one: (i) if both firms i and j are located in the same commuting zone in period t, (ii) if both firms i and j are located in the same commuting zone, and firm j is a customer or supplier of firm i, in period t, and (iii) if firm j is a customer or supplier of firm i in period t, respectively. Supporting evidence of the first criterion used includes the work by Jaffe et al. (1993), Audretsch and Feldman (1996), and Matray (2021). These studies suggest that knowledge/innovation spillovers tend to be geographically localized. In the second variant, the spillover is further restricted specifically to firms in supplier-customer relationships. An unexpected productivity shock experienced by a firm may trickle down to its buyers via the provision of higher quality inputs, allowing its buyers to scale up their productivities. Alternatively, the technology innovation or discovery may occur simultaneously to the network of firms that belong to the same supply or value chain. The third variant instead focuses on the supply chain as the conduit of innovation spillovers, but disregards the potential distance between the customers/suppliers. Note that the second variant is a relatively sparse matrix compared to the other two variants.

The connectivity matrices defined above are then row-normalized, such that each row has a row sum equal to one (and zero if all elements in a row are zeros).

2.5 Monte Carlo Simulations

In this section, we conduct Monte Carlo simulations to assess the consistency and efficiency of the estimator we proposed in Section 2.3 (which allows for spatial dependence across firms), and compare it with the conventional estimators (that assume no such spatial dependence). We consider five data generating processes (DGPs). The first DGP (DGP1) is favorable to the conventional estimator and assumes that the productivity ω_t follows an AR(1) process. The remaining DGPs consider spatial dependence of different structures and strengths across firms. The second DGP (DGP2) assumes the productivity ω_t to depend on own lagged productivity and the lagged outputs/inputs of connected firms as specified in Equation (2.2). The third DGP (DGP3) further allows the productivity shock u_t to be spatially correlated as specified in Equation (2.3). The fourth DGP

(DGP4) is the same as DGP3 but assumes stronger spatial dependence in the lagged output/inputs of connected firms. The fifth DGP (DGP5) is the same as DGP4 but considers instead negative spatial dependence in the lagged output/inputs of connected firms.

We generate the simulation data based on the empirical sample statistics of the Japanese BSJBSA-TSR linked dataset. Appendix 2.9 provides detailed documentations of the simulation setups, which we summarize below. We follow Ackerberg, Caves and Frazer (2015) and adopt a Leontief production function such that:

$$VA_{it} = \min\{e^{\alpha_0} L_{it}^{\alpha_l} K_{it}^{\alpha_k} e^{\omega_{it}}, e^{\alpha_m} M_{it}\} e^{\xi_{it}},$$

which implies Equation (2.1). In turn, gross output is linear in value-added. In particular, we set $e^{\alpha_m} = 1$ in simulating the gross output. The firm-level productivity is simulated based on Equations (2.2)–(2.3), with variations in the parameter values across the DGPs studied.

The firm-level input variables (labor and capital inputs) and the firm-to-firm connectivity matrices are simulated based on the firm-level statistics and the supplier-customer network statistics of the BSJBSA-TSR linked dataset. For example, based on the BSJBSA-TSR linked dataset, we tabulate the distribution of firms that supply to one, two, three, $\ldots$, and up to 24 other firms; and respectively, the distribution of firms that purchase from one, two, three, $\ldots$, and up to 24 other firms. We use these distribution statistics across years to simulate time-varying supplier-customer networks, which takes into account network addition, attrition, and persistency observed in the data.

Given the model structure, we assume that the error terms (ξ_{it}, v_{it}) are normally distributed with mean zeros and standard deviations of $\sigma_\xi = 0.3$ and $\sigma_v = 0.7$. We simulate a balanced panel of 500, 750 or 1000 firms for 10 or 19 time periods. For each DGP, 1000 simulated samples are drawn and estimated. We report the mean (Mean) and the standard deviation (SD) of the parameter point estimates across the 1000 Monte Carlo simulations, together with the estimated standard errors (SE) derived from the variance-covariance matrices of the estimators. The exact parameter values used in the DGPs are listed in the first row of Tables 2.3–2.7. The parameter values that are common across DGPs are: $\alpha_0 = 0$, $\alpha_l = 0.6$, $\alpha_k = 0.4$; $\rho_1 = 0.5$ and $\rho_2 = \ldots = \rho_G = 0$. To simplify the Monte Carlo exercises, we drop $\mathbf{x}_{i,t-1}$ (a firm's lagged exporting status and/or R&D expenditure) from consideration in the simulation. In DGP2, the strength of spillovers in terms of lagged outputs and lagged labor inputs of related firms is set at: $\lambda = \beta_l = 0.01$. DGP4 considers stronger spillovers such that $\lambda = \beta_l = 0.1$, while DGP5 considers negative spillovers such that $\lambda = \beta_l = -0.1$. In DGP3–DGP5, with spatial error dependence, we set $\mu = 0.25$.

Given the simulated sample, we use the Wooldridge (2009) GMM estimator in combination with the ACF estimator to represent the conventional estimator (henceforth WGMM), which assumes no spatial dependence across firms.[4] For our proposed estimator (SGMM), we use the instruments indicated in Equations (2.10) and (2.13) in estimations. In particular, the current and first lag of labor, capital and material inputs are used as the instruments for the first-stage equation (2.8), with a degree-1 h function in labor, capital and material inputs (à la Ackerberg, Caves and Frazer, 2015). For the second-stage equation (2.11), the current capital along with the first lag of labor, capital and material inputs, and the lagged-one-period outputs and labor inputs of related firms ($\mathbf{W}^y_{t-1}y_{t-1}$, $(\mathbf{W}^y_{t-1})^2y_{t-1}$, $\mathbf{W}^l_{t-1}l_{t-1}$, $(\mathbf{W}^l_{t-1})^2l_{t-1}$) are used as instruments. The connectivity matrices are as defined in Section 2.4.2. In particular, the connectivity matrix $\mathbf{W}^u_t$ specifying the spatial correlation of productivity shocks is defined based on the customer-supplier relationships across firms. The same set of instruments are used for the WGMM estimations, but excluding the related firms' lagged outputs and lagged labor inputs ($\mathbf{W}^y_{t-1}y_{t-1}$, $(\mathbf{W}^y_{t-1})^2y_{t-1}$, $\mathbf{W}^l_{t-1}l_{t-1}$, $(\mathbf{W}^l_{t-1})^2l_{t-1}$).

2.5.1 Simulation Results

Table 2.3 reports the results for DGP1. The conventional estimator (WGMM) performs well as it should, when the DGP has no spatial dependence across firms. Importantly, our proposed

[4] As noted above, the Wooldridge (2009) procedure estimates Equations (2.8) and (2.11) jointly (instead of sequentially as in ACF), while the ACF estimator allows $g(\cdot,\cdot,\cdot)$ to be a general function in all the input variables (including linearity as a special case) and hence does not aim to identify the slope coefficients (α_l, α_k) in Stage 1 from Equation (2.8) but in Stage 2 from Equation (2.11).

estimator (SGMM) performs just as well. The point estimates of both estimators are close to the true parameter values, and the 95% confidence intervals (CIs) include the true parameter values for the input coefficients of the production function (α_l, α_k). While our estimator has wider confidence intervals than the conventional estimator for the input coefficients, it returns mean estimates of the spatial coefficients (λ, β_l, μ) nearly identical to zeros, consistent with the true parameter values of the underlying DGP. Both estimators obtain estimates for the autoregressive parameter (ρ_1) that are close to the true parameter value, even when the duration of the panel is relatively short. Both the conventional estimator and our proposed estimator yield standard error estimates (SE) that are close to their Monte Carlo standard deviations (SD). The SEs also reduce as the sample size increases at a rate consistent with the asymptotic properties laid out in Section 2.3.3.

Table 2.4 reports the findings for the second set of simulations based on DGP2. When spatial dependence across firms via lagged outputs and lagged labor inputs are indeed present, the conventional estimator leads to biased estimates of the input coefficients. In particular, its mean estimates for α_l across variations in N and T are higher than the true parameter value. The bias does not shrink with a larger sample size, suggesting the inconsistency of the conventional estimator when spatial dependence is present in the underlying DGP. In contrast, our proposed SGMM estimator yields estimates that are close to the true parameter values for both input elasticities (α_l, α_k) and the spatial coefficients (λ, β_l), with 95% CIs that well cover the true parameter values. Finally, our proposed SGMM estimator reports statistically insignificant estimates of μ, consistent with the underlying DGP where no spatial correlation in the error terms (i.e., the productivity shocks u_t) is present.

In DGP3, the data generating process for the productivity term further allows for spatial error correlation across related firms. Table 2.5 shows that the conventional estimator of the labor coefficient of the production function remains to be upward biased, while our proposed estimator yields consistent estimates that are close to the true parameter values for all the coefficients of interest. In particular, we note that the SGMM estimator returns estimates of the spatial error coefficient (μ) that are close to its true parameter value when it is indeed non-zero.

Table 2.6 reports the simulation results for DGP4. With larger spatial coefficients ($\lambda = \beta_l = 0.1$, instead of 0.01), the conventional estimator of all coefficients (α_l, α_k, ρ_1) are upward biased, and the extents of bias are substantial (by around 23 percentage points for α_l, 4-8 percentage points for α_k, and 5-13 percentage points for ρ_1). Furthermore, the standard errors (SE) obtained by the conventional estimator deviate significantly from the Monte Carlo standard deviations (SD). In contrast, our proposed SGMM estimator continues to yield consistent estimates for the true parameters, with estimates of the standard errors (SE) that are very close to the Monte Carlo standard deviations (SD).

Table 2.7 indicates that if the underlying DGP is characterized with large negative spatial coefficients ($\lambda = \beta_l = -0.1$, instead of 0.01), the conventional estimator of the input coefficients (α_l, α_k) are instead downward biased, and the extents of bias continues to be substantial (by around 21 percentage points for α_l, 3-4 percentage points for α_k). The standard errors (SE) obtained by the conventional estimator also deviate significantly from the Monte Carlo standard deviations (SD) for these two input coefficients. In contrast, our proposed SGMM estimator yields consistent estimates for the true parameters, with estimates of the standard errors (SE) being very close to the Monte Carlo standard deviations (SD). In particular, it is able to capture the negative signs of the two spatial coefficients ($\lambda = \beta_l = -0.1$) and their magnitudes.

In sum, across all the DGPs, we find that the proposed SGMM estimator yields point estimates that are consistent for the true parameter values both in the absence and in the presence of spatial effects. By the SGMM estimator, the standard error estimates (SE) of the parameters are also very close to the Monte Carlo standard deviations (SD). As the sample size $N(T-1)$ doubles (either due to doubling of N or $T-1$), both the standard error estimates (SE) and the Monte Carlo standard deviations (SD) shrink at a rate close to $1/\sqrt{2}$, consistent with a convergence rate of $1/\sqrt{N(T-1)}$.

2.6 Empirical Analysis

2.6.1 Estimation Sample

We apply the methodology and estimation algorithms proposed in Section 2.3 to the Japanese dataset introduced in Section 2.4. Given the BSJBSA-TSR linked data, we further restrict the sample to a balanced panel of firms with observations on the set of variables required for productivity estimations. In particular, the sample is based on firms that were surveyed for 10 consecutive years from 2009 to 2018.[5] Second, the sample of firms used for analysis also need to have non-missing values for log of labor hours (used to measure l_{it}), log of real capital stock (k_{it}), log of real intermediate inputs (m_{it}), log of real revenues (y_{it}), and log of real value added (va_{it}), during the entire sample period 2009–2018. Recall that the real capital stock is calculated based on the perpetual inventory method with the real capital stock in 2007 (year end) as the initial value. Observations on real capital stock for a firm could be missing, for example, if the firm was not observed in 2007.

The resulting sample is a balanced panel of 13,001 firms for the period 2009–2018. Given the balanced panel of firms, the set of a firm's customers/suppliers identified via the TSR entries is effectively restricted to those whose firm-level data also exist in BSJBSA. In particular, firm j is regarded effectively as a customer/supplier of firm i in the estimation if firm i reports firm j as a customer or supplier, *and* if firm j exists in the BSJBSA dataset (with consecutive observations on output and inputs, as required for the estimation of Equation (2.2)).

Table 2.2(b) provides the summary statistics for the estimation sample. Relative to the raw sample reported in Table 2.2(a), the firms in the estimation sample tends to be larger in terms of both inputs and output, and have more customers and suppliers. This is expected, as larger firms are more likely to be surveyed consecutively throughout the years and have positive inputs/output. Although larger firms tend to have more customers/suppliers, the orders of magnitude in the number of connections on average do not differ substantially between the raw and estimation

[5] This excludes, for example, firms whose number of employees fell under 50 at some point during the sample period.

samples. Despite the much smaller set of firms covered, the estimation sample accounts for 60.97% of aggregate real value added and 62.79% of real gross output of the raw sample in 2015 (and a majority of the other economic activities in terms of employment, labor hours, real capital stock, and real spending on intermediate inputs).

Figures 2.1(b)–2.10(b) repeat the characterization as in Figures 2.1(a)–2.10(a), and show that the estimation sample has similar patterns as documented for the raw sample. Notably, four industries are not present in the estimation sample: construction, mining & quarrying, agriculture & forestry, and fisheries (so firms in these industries tend not to be large enough to be consecutively surveyed by BSJBSA). Focusing on the remaining industries, the rank across industries is almost identical in terms of the number of firms: manufacturing, wholesale & retail trade, and information & communications remain to be the top three industries with the largest numbers of firms (Figure 2.1). The set of prefectures with the largest numbers of firms is also similar to that previously documented (Figure 2.4). Basically, firms in the estimation sample tend to be larger in terms of employment size (Figures 2.2, 2.5, and 2.9), and are slightly more connected in terms of customers/suppliers (Figures 2.3, 2.6, and 2.10), in comparison with the raw sample.

2.6.2 Estimation Results

We estimate the model proposed in Equations (2.1)–(2.3) based on the estimation methodology laid out in Section 2.3 and the connectivity matrices defined in Section 2.4.2. In short, we define the output connectivity matrix $\mathbf{W}^y_{t-1}$ based on the set of a firm's customers/suppliers located in the same commuting zone. The ij-th element of $\mathbf{W}^y_{t-1}$ takes on the value one if both firms i and j are located in the same commuting zone, and in addition, firm j is a customer or supplier of firm i, in period $t-1$. Second, we define input connectivity matrix $\mathbf{W}^\Omega_{t-1}$ such that the ij-th element of $\mathbf{W}^\Omega_{t-1}$ takes on the value one if both firms i and j are located in the same commuting zone in period $t-1$. The input variable being analyzed corresponds to the lagged labor inputs of the connected firms defined by $\mathbf{W}^\Omega_{t-1}$. Third, we consider three variants of the spatial error connectivity matrix $\mathbf{W}^u_t$ and define it such that the ij-th element of $\mathbf{W}^u_t$ takes on the value one: (i) if both firms i and

j are located in the same commuting zone in period t, (ii) if both firms i and j are located in the same commuting zone, and firm j is a customer or supplier of firm i, in period t, and (iii) if firm j is a customer or supplier of firm i in period t, respectively. The connectivity matrices are then row-normalized, such that each row has a row sum equal to one (and zero if all elements in a row are zeros).

Table 2.8 reports the estimation results. Column 1 (based on the first definition of $\mathbf{W}_t^u$) indicates that all three spatial coefficients are significant and positive. A 1% increase in the sales of customers/suppliers in the same commuting zone in the previous period helps improve a firm's current productivity by 0.005%. A larger local labor market also enhances a firm's productivity: specifically, a 1% increase in the employment of firms located in the same commuting zone in the previous period raises a firm's current productivity by 0.05%. Finally, there is evidence of contemporary knowledge spillovers across firms located in the same commuting zone: the productivity innovations u_{it} are spatially correlated with a positive and significant slope coefficient of 0.41. The other production function parameters are also precisely estimated. Column 1 reports a labor value-added share of 0.84, a capital value-added share of 0.04, and a partial AR(1) coefficient of 0.961 for the productivity process.

The estimates for key parameters of interest remain similar if we adopt alternative definitions of spatial error connectivity matrices $\mathbf{W}_t^u$, as reported in Column 2 and Column 3. The key difference is the strength of contemporary spatial correlation in the productivity shocks u_{it}. They tend to be quantitatively smaller by an order of magnitude if $\mathbf{W}_t^u$ is defined in a more restricted manner, relative to the finding in Column 1 based on common commuting zone alone.

To compare the spatial GMM findings with those if one ignores the spatial dependence across firms, Column 4 reports the estimates based on conventional estimators (restricting the spatial coefficients to be zeros but otherwise adopting the same GMM approach). We find that the labor value-added share tends to be downward biased (0.44), the capital value-added share upward biased (0.19), and the AR(1) coefficient upward biased (0.964), in comparison with the spatial GMM estimates reported above.

2.7 Conclusion

In this paper, we develop a framework to simultaneously estimate firm production functions and spatial interactions across firms in one unified setup. We propose a three-stage efficient GMM estimation algorithm, and show by theory the asymptotic properties of the proposed estimator and by Monte Carlo simulations the finite sample performance of the estimator. The Monte Carlo simulations demonstrate that the proposed estimator is consistent under DGPs with or without spatial dependence across firms. In contrast, the conventional estimators are biased when the true DGPs are indeed characterized with spatial dependence. By applying the developed methodology and estimation algorithm to the Japanese BSJBSA-TSR linked dataset for the period 2009–2018, we find that spatial interactions across firms play a significant role in determining the Japanese firm-level productivity both statistically and economically.

The paper can be extended in several directions in future research. First, the connectivity matrices in our setup are allowed to differ across different mechanisms of spatial interactions. One can potentially hypothesize alternative candidates for the connectivity matrices and conduct specification tests that select the specification that best fits the model. This will also provide insights into the nature of spatial interactions across firms and allow tests for competing hypotheses. Second, the current framework allows for time-varying connectivity matrices. This is useful, as we can use the framework to analyze how shocks (such as high-speed rails and earthquakes) affect the connectivity matrices across firms, and in turn, the firm-level performance measures (such as productivity, and production technology). Third, the current framework could also be used to analyze the centrality of firms in the sense that a firm is more central if by increasing connectivity (links) to this firm, the average productivity of all firms is improved by more than if by increasing connectivity (links) to another firm. This is useful for policy design that aims to target subsidies at the critical links of a firm network structure for the greater benefits of the economy.

Bibliography

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer, "Identification Properties of Recent Production Function Estimators," *Econometrica*, 2015, *83*, 2411–2451.

Adachi, Daisuke, Taiyo Fukai, Daiji Kawaguchi, and Yukiko Umeno Saito, "Commuting Zones in Japan," RIETI Discussion Paper 20-E-021 2021.

Alfaro-Ureña, Alonso, Isabela Manelici, and Jose P. Vasquez, "The Effects of Joining Multinational Supply Chains: New Evidence from Firm-to-Firm Linkages," *Quarterly Journal of Economics*, 2022, *137* (3), 1495–1552.

Audretsch, David B. and Maryann P. Feldman, "R&D Spillovers and the Geography of Innovation and Production," *American Economic Review*, 1996, *86* (3), 630–640.

Bernard, Andrew B., Andreas Moxnes, and Yukiko U. Saito, "Production Networks, Geography and Firm Performance," *Journal of Political Economy*, 2019, *127*, 639–688.

Braguinsky, Serguey, Atsushi Ohyama, Tetsuji Okazaki, and Chad Syverson, "Acquisitions, Productivity, and Profitability: Evidence from the Japanese Cotton Spinning Industry," *American Economic Review*, 2015, *105* (7), 2086–2119.

De Loecker, Jan, "Detecting Learning by Exporting," *American Economic Journal: Microeconomics*, 2013, *5* (3), 1–21.

Demir, Banu, Ana Cecilia Fieler, Daniel Yi Xu, and Kelly Kaili Yang, "O-Ring Production Networks," *Journal of Political Economy*, 2024, *132* (1), forthcoming.

Diamond, Charles A. and Curtis J. Simon, "Industrial Specialization and the Returns to Labor," *Journal of Labor Economics*, 1990, *8* (2), 175–201.

Doraszelski, Ulrich and Jordi Jaumandreu, "R&D and Productivity: Estimating Endogenous Productivity," *Review of Economic Studies*, 2013, *80*, 1338–1383.

Elhorst, Paul, *From Cross-Sectional Data to Spatial Panels*, Heidelberg: Springer, 2014.

Ellison, Glenn, Edward L. Glaeser, and William R. Kerr, "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns," *American Economic Review*, 2010, *100* (3), 1195–1213.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti, "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings," *Journal of Political Economy*, 2010, *118* (3), 536–598.

Helsley, Robert W. and William C. Strange, "Matching and Agglomeration Economies in a System of Cities," *Regional Science and Urban Economics*, 1990, *20* (2), 189–212.

_ and _ , "Innovation and input sharing," *Journal of Urban Economics*, 2002, *51* (1), 25–45.

Iyoha, Ebehi, "Estimating Productivity in the Presence of Spillovers: Firm-level Evidence from the US Production Network," manuscript 2022.

Jaffe, A. B., M. Trajtenberg, and R. Henderson, "Geographic localization of knowledge spillovers as evidenced by patent citations," *Quarterly Journal of Economics*, 1993, *108* (3), 577–598.

Kapoor, Mudit, Harry Kelejian, and Ingmar Prucha, "Panel Data Models with Spatially Correlated Error Components," *Journal of Econometrics*, 2007, *140*, 97–130.

Kelejian, Harry and Ingmar R Prucha, "A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances," *The Journal of Real Estate Finance and Economics*, 1998, *17*, 99–121.

Kelejian, Harry H and Ingmar R Prucha, "A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model," *International Economic Review*, 1999, *40* (2), 509–533.

Krugman, Paul, *Geography and Trade*, Cambridge, MA: MIT Press, 1991.

Lee, Lung-Fei and Jihai Yu, "Efficient GMM Estimation of Spatial Dynamic Panel Data Models with Fixed Effects," *Journal of Econometrics*, 2014, *180*, 174–197.

Levinsohn, James and Amil Petrin, "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 2003, *70*, 317–341.

Matray, Adrien, "The Local Innovation Spillovers of Listed Firms," *Journal of Financial Economics*, 2021, *141*, 395–412.

Miyauchi, Yuhei, "Matching and Agglomeration: Theory and Evidence from Japanese Firm-to-Firm Trade," *Econometrica*, 2023, p. conditionally accepted.

Olley, G. Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1996, *64*, 1263–1297.

Rotemberg, Julio J. and Garth Saloner, "Competition and Human Capital Accumulation: A Theory of Interregional Specialization and Trade," *Regional Science and Urban Economics*, 2000, *30* (4), 373–404.

Tolbert, Charles M. and Molly Sizer, "U.S. Commuting Zones and Labor Market Areas: A 1990 Update," Staff Reports 278812, USDA 1996.

_ **and Molly Sizer Killian**, "Labor Market Areas for the United States," ERS staff report, USDA 1987.

Wooldridge, Jeffrey M., "On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables," *Economics Letters*, 2009, *104*, 112–114.

2.8 Theoretical Appendix

2.8.1 Deriving the Moment Condition in Stage 3

Substituting $v_t = u_t - \mu \mathbf{W}_t^u u_t$ into the left side of Equation (2.6) yields:

$$\frac{1}{N}E\begin{bmatrix} (u_t - \mu\mathbf{W}_t^u u_t)'(u_t - \mu\mathbf{W}_t^u u_t) \\ (u_t - \mu\mathbf{W}_t^u u_t)'\mathbf{W}_t^{u\prime}\mathbf{W}_t^u(u_t - \mu\mathbf{W}_t^u u_t) \\ (u_t - \mu\mathbf{W}_t^u u_t)'\mathbf{W}_t^u(u_t - \mu\mathbf{W}_t^u u_t) \end{bmatrix} = \frac{1}{N}E\begin{bmatrix} (u_t - \mu\overline{u}_t)'(u_t - \mu\overline{u}_t) \\ (u_t - \mu\overline{u}_t)'\mathbf{W}_t^{u\prime}\mathbf{W}_t^u(u_t - \mu\overline{u}_t) \\ (u_t - \mu\overline{u}_t)'\mathbf{W}_t^u(u_t - \mu\overline{u}_t) \end{bmatrix}$$

$$= \frac{1}{N}E\begin{bmatrix} (u_t'u_t - 2\mu\overline{u}_t{}'u_t + \mu^2\overline{u}_t'\overline{u}_t) \\ (\overline{u}_t'\overline{u}_t - 2\mu\overline{\overline{u}}_t'\overline{u}_t + \mu^2\overline{\overline{u}}_t'\overline{\overline{u}}_t) \\ \overline{u}_t'u_t - \mu(\overline{u}_t'\overline{u}_t + \overline{\overline{u}}_t'u_t) + \mu^2\overline{\overline{u}}_t'\overline{u}_t \end{bmatrix}.$$

Therefore, Equation (2.6) becomes:

$$\frac{1}{N}E\begin{bmatrix} (u_t'u_t - 2\mu\overline{u}_t{}'u_t + \mu^2\overline{u}_t'\overline{u}_t) \\ (\overline{u}_t'\overline{u}_t - 2\mu\overline{\overline{u}}_t'\overline{u}_t + \mu^2\overline{\overline{u}}_t'\overline{\overline{u}}_t) \\ \overline{u}_t'u_t - \mu(\overline{u}_t'\overline{u}_t + \overline{\overline{u}}_t'u_t) + \mu^2\overline{\overline{u}}_t'\overline{u}_t \end{bmatrix} = \begin{bmatrix} \sigma_v^2 \\ \frac{\sigma_v^2}{N}\mathrm{tr}(\mathbf{W}_t^{u\prime}\mathbf{W}_t^u) \\ 0 \end{bmatrix}. \tag{2.33}$$

Rearranging terms, we have:

$$\frac{1}{N}E\begin{bmatrix} u_t'u_t \\ \overline{u}_t'\overline{u}_t \\ \overline{u}_t'u_t \end{bmatrix} = \frac{1}{N}\begin{bmatrix} 2\mu\overline{u}_t{}'u_t - \mu^2\overline{u}_t'\overline{u}_t + \sigma_v^2 \\ 2\mu\overline{\overline{u}}_t'\overline{u}_t - \mu^2\overline{\overline{u}}_t'\overline{\overline{u}}_t + \frac{\sigma_v^2}{N}\mathrm{tr}(\mathbf{W}_t^{u\prime}\mathbf{W}_t^u) \\ \mu(\overline{u}_t'\overline{u}_t - \overline{\overline{u}}_t'u_t) - \mu^2\overline{\overline{u}}_t'\overline{u}_t \end{bmatrix}$$

$$= \frac{1}{N}\begin{bmatrix} 2\overline{u}_t{}'u_t & -\overline{u}_t'\overline{u}_t & 1 \\ 2\overline{\overline{u}}_t'\overline{u}_t & -\overline{\overline{u}}_t'\overline{\overline{u}}_t & \frac{1}{N}\mathrm{tr}(\mathbf{W}_t^{u\prime}\mathbf{W}_t^u) \\ (\overline{u}_t'\overline{u}_t - \overline{\overline{u}}_t'u_t) & -\overline{\overline{u}}_t'\overline{u}_t & 0 \end{bmatrix}\begin{bmatrix} \mu \\ \mu^2 \\ \sigma_v^2 \end{bmatrix}, \tag{2.34}$$

which yields the following relationship:

$$\gamma_t = \mathbf{\Gamma}_t \begin{bmatrix} \mu \\ \mu^2 \\ \sigma_v^2 \end{bmatrix}.$$

2.8.2 Deriving the Variance-Covariance Matrix for the Moment Conditions in Stages 1 and 2

Using the definition of the variance-covariance matrix of the moment conditions in the first and second stages, we have:

$$\begin{aligned}
\mathbf{V}_{\boldsymbol{\theta}} &= \mathrm{Var}\left(\frac{1}{\sqrt{N(T-1)}}\sum_{i=1}^{N}\sum_{t=2}^{T} Z'_{it} r_{it}\right) \\
&= \frac{1}{N(T-1)}\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{j=1}^{N}\sum_{s=2}^{T} E[(Z'_{it} r_{it})(Z'_{js} r_{js})'] \\
&= \frac{1}{N(T-1)}\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{j=1}^{N}\sum_{s=2}^{T} Z'_{it} E[r_{it} r'_{js}] Z_{js} \\
&= \frac{1}{N(T-1)}\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{j=1}^{N}\sum_{s=2}^{T} Z'_{it} E\left[\begin{pmatrix} \xi_{it} \\ \xi_{it} + u_{it} \end{pmatrix} (\xi_{js} \quad \xi_{js} + u_{js})'\right] Z_{js} \\
&= \frac{1}{N(T-1)}\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{j=1}^{N}\sum_{s=2}^{T} Z'_{it} E\begin{pmatrix} \xi_{it}\xi_{js} & \xi_{it}(\xi_{js} + u_{js}) \\ (\xi_{it} + u_{it})\xi_{js} & (\xi_{it} + u_{it})(\xi_{js} + u_{js}) \end{pmatrix} Z_{js}. \qquad (2.35)
\end{aligned}$$

To derive $E\begin{pmatrix} \xi_{it}\xi_{js} & \xi_{it}(\xi_{js} + u_{js}) \\ (\xi_{it} + u_{it})\xi_{js} & (\xi_{it} + u_{it})(\xi_{js} + u_{js}) \end{pmatrix}$, consider the following 4 cases:

1. $i = j, t = s$

$$E\begin{pmatrix} \xi_{it}\xi_{js} & \xi_{it}(\xi_{js}+u_{js}) \\ (\xi_{it}+u_{it})\xi_{js} & (\xi_{it}+u_{it})(\xi_{js}+u_{js}) \end{pmatrix} = E\begin{pmatrix} \xi_{it}\xi_{it} & \xi_{it}(\xi_{it}+u_{it}) \\ (\xi_{it}+u_{it})\xi_{it} & (\xi_{it}+u_{it})(\xi_{it}+u_{it}) \end{pmatrix}$$
$$= \begin{pmatrix} \sigma_\xi^2 & \sigma_\xi^2 \\ \sigma_\xi^2 & \sigma_\xi^2 + E(u_{it}^2) \end{pmatrix}.$$

2. $i = j, t \neq s$

$$E\begin{pmatrix} \xi_{it}\xi_{js} & \xi_{it}(\xi_{js}+u_{js}) \\ (\xi_{it}+u_{it})\xi_{js} & (\xi_{it}+u_{it})(\xi_{js}+u_{js}) \end{pmatrix} = E\begin{pmatrix} \xi_{it}\xi_{is} & \xi_{it}(\xi_{is}+u_{is}) \\ (\xi_{it}+u_{it})\xi_{is} & (\xi_{it}+u_{it})(\xi_{is}+u_{is}) \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

3. $i \neq j, t = s$

$$E\begin{pmatrix} \xi_{it}\xi_{js} & \xi_{it}(\xi_{js}+u_{js}) \\ (\xi_{it}+u_{it})\xi_{js} & (\xi_{it}+u_{it})(\xi_{js}+u_{js}) \end{pmatrix} = E\begin{pmatrix} \xi_{it}\xi_{jt} & \xi_{it}(\xi_{jt}+u_{jt}) \\ (\xi_{it}+u_{it})\xi_{jt} & (\xi_{it}+u_{it})(\xi_{jt}+u_{jt}) \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 0 \\ 0 & E(u_{it}u_{jt}) \end{pmatrix}.$$

4. $i \neq j, t \neq s$

$$E\begin{pmatrix} \xi_{it}\xi_{js} & \xi_{it}(\xi_{js}+u_{js}) \\ (\xi_{it}+u_{it})\xi_{js} & (\xi_{it}+u_{it})(\xi_{js}+u_{js}) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

To estimate $E(u_{it}^2)$, note that u_{it} can be obtained by taking the product of the i-th row of

$(I_N - \mu\mathbf{W}_t^u)^{-1}$ and v_t, since $u_t = (\mathbf{I}_N - \mu\mathbf{W}_t^u)^{-1}v_t$. Then, we have:

$$\begin{aligned} E(u_{it}^2) &= E([(I_N - \mu\mathbf{W}_t^u)^{-1}]_i v_t v_t'([(I_N - \mu\mathbf{W}_t^u)^{-1}]_i)']) \\ &= [(I_N - \mu\mathbf{W}_t^u)^{-1}]_i E[v_t v_t']([(I_N - \mu\mathbf{W}_t^u)^{-1}]_i)' \\ &= \sigma_v^2[(I_N - \mu\mathbf{W}_t^u)^{-1}]_i([(I_N - \mu\mathbf{W}_t^u)^{-1}]_i)', \end{aligned}$$

where $[X]_i$ refers to the i-th row of matrix X and $E(v_t v_t') = \sigma_v^2\mathbf{I}_N$. Similarly, $E(u_{it}u_{jt})$ is given by:

$$\begin{aligned} E(u_{it}u_{jt}) &= E([(I_N - \mu\mathbf{W}_t^u)^{-1}]_i v_t v_t'([(I_N - \mu\mathbf{W}_t^u)^{-1}]_j)']) \\ &= [(I_N - \mu\mathbf{W}_t^u)^{-1}]_i E[v_t v_t']([(I_N - \mu\mathbf{W}_t^u)^{-1}]_j)' \\ &= \sigma_v^2[(I_N - \mu\mathbf{W}_t^u)^{-1}]_i([(I_N - \mu\mathbf{W}_t^u)^{-1}]_j)'. \end{aligned}$$

Substituting all the terms back into $\mathbf{V}_{\boldsymbol{\theta}}$, we obtain:

$$\begin{aligned}
\mathbf{V}_{\boldsymbol{\theta}} =& \frac{1}{N(T-1)} \sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{j=1}^{N}\sum_{s=2}^{T} \mathcal{Z}_{it}' E \begin{pmatrix} \xi_{it}\xi_{js} & \xi_{it}(\xi_{js}+u_{js}) \\ (\xi_{it}+u_{it})\xi_{js} & (\xi_{it}+u_{it})(\xi_{js}+u_{js}) \end{pmatrix} \mathcal{Z}_{js} \\
=& \frac{1}{N(T-1)} \sum_{i=1}^{N}\sum_{t=2}^{T} \begin{pmatrix} \mathcal{Z}_{it,I}' & 0 \\ 0 & \mathcal{Z}_{it,II}' \end{pmatrix} E \begin{pmatrix} \xi_{it}^2 & \xi_{it}^2 \\ \xi_{it}^2 & \xi_{it}^2 + u_{it}^2 \end{pmatrix} \begin{pmatrix} \mathcal{Z}_{it,I} & 0 \\ 0 & \mathcal{Z}_{it,II} \end{pmatrix} \\
&+ \frac{1}{N(T-1)} \sum_{i=1}^{N}\sum_{j\neq i}^{N}\sum_{t=2}^{T} \begin{pmatrix} \mathcal{Z}_{it,I}' & 0 \\ 0 & \mathcal{Z}_{it,II}' \end{pmatrix} E \begin{pmatrix} 0 & 0 \\ 0 & u_{it}u_{jt} \end{pmatrix} \begin{pmatrix} \mathcal{Z}_{jt,I} & 0 \\ 0 & \mathcal{Z}_{jt,II} \end{pmatrix} \\
=& \frac{1}{N(T-1)} \sum_{i=1}^{N}\sum_{t=2}^{T} \begin{pmatrix} \mathcal{Z}_{it,I}'\mathcal{Z}_{it,I}E(\xi_{it}^2) & \mathcal{Z}_{it,I}'\mathcal{Z}_{it,II}E(\xi_{it}^2) \\ \mathcal{Z}_{it,II}'\mathcal{Z}_{it,I}E(\xi_{it}^2) & \mathcal{Z}_{it,II}'\mathcal{Z}_{it,II}E(\xi_{it}^2) \end{pmatrix} \\
&+ \frac{1}{N(T-1)} \sum_{t=2}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N} \begin{pmatrix} \mathcal{Z}_{it,I}' & 0 \\ 0 & \mathcal{Z}_{it,II}' \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & \sigma_v^2[(I_N-\mu\mathbf{W}_t^u)^{-1}]_i([(I_N-\mu\mathbf{W}_t^u)^{-1}]_j)' \end{pmatrix} \begin{pmatrix} \mathcal{Z}_{jt,I} & 0 \\ 0 & \mathcal{Z}_{jt,II} \end{pmatrix} \\
=& \frac{1}{N(T-1)} \sum_{i=1}^{N}\sum_{t=2}^{T} \begin{pmatrix} \mathcal{Z}_{it,I}'\mathcal{Z}_{it,I}E(\xi_{it}^2) & \mathcal{Z}_{it,I}'\mathcal{Z}_{it,II}E(\xi_{it}^2) \\ \mathcal{Z}_{it,II}'\mathcal{Z}_{it,I}E(\xi_{it}^2) & \mathcal{Z}_{it,II}'\mathcal{Z}_{it,II}E(\xi_{it}^2) \end{pmatrix} \\
&+ \frac{1}{N(T-1)} \sum_{t=2}^{T} \begin{pmatrix} \mathbf{0}_{\mathcal{M}_1\times\mathcal{M}_1} & \mathbf{0}_{\mathcal{M}_1\times\mathcal{M}_2} \\ \mathbf{0}_{\mathcal{M}_2\times\mathcal{M}_1} & \sigma_v^2\mathcal{Z}_{t,II}'[(\mathbf{I}_N-\mu\mathbf{W}_t^u)^{-1}(\mathbf{I}_N-\mu\mathbf{W}_t^u)^{-1\prime}]\mathcal{Z}_{t,II} \end{pmatrix},
\end{aligned}$$

where $\mathcal{M}_1$ and $\mathcal{M}_2$ are the number of moment conditions (instruments) used in the first and second stages respectively (such that $\mathcal{M}_1 + \mathcal{M}_2 = \mathcal{M}$), and $\mathcal{Z}_{t,II} = [Z_{1t,II}', Z_{2t,II}', \ldots, Z_{Nt,II}']'$ is a $N \times \mathcal{M}_2$ matrix.

We can estimate $\mathbf{V}_{\boldsymbol{\theta}}$ by its sample counterpart:

$$\begin{aligned}
\widehat{\mathbf{V}}_{\boldsymbol{\theta}} =& \frac{1}{N(T-1)} \sum_{i=1}^{N}\sum_{t=2}^{T} \begin{pmatrix} \mathcal{Z}_{it,I}'\mathcal{Z}_{it,I}\widehat{\xi}_{it}^2 & \mathcal{Z}_{it,I}'\mathcal{Z}_{it,II}\widehat{\xi}_{it}^2 \\ \mathcal{Z}_{it,II}'\mathcal{Z}_{it,I}\widehat{\xi}_{it}^2 & \mathcal{Z}_{it.II}'\mathcal{Z}_{it,II}\widehat{\xi}_{it}^2 \end{pmatrix} \\
&+ \frac{1}{N(T-1)} \sum_{t=2}^{T} \boldsymbol{\mathcal{Z}}_t' \begin{pmatrix} \mathbf{0}_{N\times N} & \mathbf{0}_{N\times N} \\ \mathbf{0}_{N\times N} & \widehat{\sigma_v^2}[(\mathbf{I}_N-\widehat{\mu}\mathbf{W}_t^u)^{-1}(\mathbf{I}_N-\widehat{\mu}\mathbf{W}_t^u)^{-1\prime}]) \end{pmatrix} \boldsymbol{\mathcal{Z}}_t,
\end{aligned}$$

where $\widehat{\xi}_{it}$, $\widehat{\mu}$ and $\widehat{\sigma_v^2}$ are obtained from Steps 1–2 in Section 2.3.3.

2.8.3 Deriving the Variance-Covariance Matrix for the Moment Conditions in Stage 3

In this section, we derive the variance-covariance matrix for the moment conditions used in Stage 3. Let $\mathbf{A}$ and $\mathbf{B}$ be $n \times n$ non-stochastic matrices. For a $n \times 1$ random vector $\mathbf{e}$ with mean 0, variance σ_e^2 and finite excess kurtosis κ_e:

$$\text{cov}(\mathbf{e}'\mathbf{A}\mathbf{e}, \mathbf{e}'\mathbf{B}\mathbf{e}) = \sigma_e^4 \kappa_e a'b + \sigma_e^4 \text{tr}(\mathbf{A}(\mathbf{B}' + \mathbf{B})), \tag{2.36}$$

where a = diagv($\mathbf{A}$) and b = diagv($\mathbf{B}$). The operator 'diagv' takes the diagonal elements of a matrix and converts them to a column vector.

Define $\mathbf{v} = [v_2', v_3', \ldots, v_T']$ to be a $N(T-1) \times 1$ vector, and $\overline{\mathbf{v}} = \mathbf{W}^u\mathbf{v}$. Let κ_v be the excess kurtosis of $\mathbf{v}$. The sample counterpart of the moment conditions in Equation (2.6) is:

$$\frac{1}{N(T-1)} \begin{pmatrix} \mathbf{v}'\mathbf{v} \\ \mathbf{v}'\mathbf{W}^{u\prime}\mathbf{W}^u\mathbf{v} \\ \mathbf{v}'\mathbf{W}^u\mathbf{v} \end{pmatrix} = \begin{pmatrix} \hat{\sigma}_v^2 \\ \frac{\hat{\sigma}_v^2}{N(T-1)} \text{tr}(\mathbf{W}^{u\prime}\mathbf{W}^u) \\ 0 \end{pmatrix}. \tag{2.37}$$

The following computes each cell of the variance-covariance matrix of the vector on the left side of Equation (2.37):

$$\begin{aligned}
V_{\psi,11} &= N(T-1)\text{cov}\left(\frac{1}{N(T-1)}\mathbf{v}'\mathbf{v}, \frac{1}{N(T-1)}\mathbf{v}'\mathbf{v}\right) \\
&= N(T-1)\frac{\sigma_v^4}{(N(T-1))^2}[\kappa_v \iota_{N(T-1)}' \iota_{N(T-1)} + \text{tr}(\mathbf{I}_{N(T-1)}(\mathbf{I}_{N(T-1)}' + \mathbf{I}_{N(T-1)}))] \\
&= \sigma_v^4(\kappa_v + 2); \\
V_{\psi,12} &= N(T-1)\text{cov}\left(\frac{1}{N(T-1)}\mathbf{v}'\mathbf{v}, \frac{1}{N(T-1)}\overline{\mathbf{v}}'\overline{\mathbf{v}}\right) \\
&= N(T-1)\text{cov}\left(\frac{1}{N(T-1)}\mathbf{v}'\mathbf{v}, \frac{1}{N(T-1)}\mathbf{v}'\mathbf{W}^{u\prime}\mathbf{W}^u\mathbf{v}\right) \\
&= \frac{\sigma_v^4}{N(T-1)}[\kappa_v \text{tr}(\mathbf{W}^{u\prime}\mathbf{W}^u) + \text{tr}(\mathbf{I}_{N(T-1)}(\mathbf{W}^{u\prime}\mathbf{W}^u + \mathbf{W}^u\mathbf{W}^{u\prime}))] \\
&= \frac{\sigma_v^4}{N(T-1)}(\kappa_v + 2)\text{tr}(\mathbf{W}^{u\prime}\mathbf{W}^u);
\end{aligned}$$

$$
\begin{aligned}
V_{\psi,13} &= N(T-1)\text{cov}\left(\frac{1}{N(T-1)}\mathbf{v}'\mathbf{v}, \frac{1}{N(T-1)}\mathbf{v}'\overline{\mathbf{v}}\right) \\
&= N(T-1)\text{cov}\left(\frac{1}{N(T-1)}\mathbf{v}'\mathbf{v}, \frac{1}{N(T-1)}\mathbf{v}'\mathbf{W}^{u\prime}\mathbf{v}\right) \\
&= \frac{\sigma_v^4}{N(T-1)}[\kappa_v \text{tr}(\mathbf{W}^u) + \text{tr}(\mathbf{W}^u)] \\
&= 0; \\
V_{\psi,22} &= N(T-1)\text{cov}\left(\frac{1}{N(T-1)}\overline{\mathbf{v}}'\overline{\mathbf{v}}, \frac{1}{N(T-1)}\overline{\mathbf{v}}'\overline{\mathbf{v}}\right) \\
&= N(T-1)\text{cov}\left(\frac{1}{N(T-1)}\mathbf{v}'\mathbf{W}^{u\prime}\mathbf{W}^u\mathbf{v}, \frac{1}{N(T-1)}\mathbf{v}'\mathbf{W}^{u\prime}\mathbf{W}^u\mathbf{v}\right) \\
&= \frac{\sigma_v^4}{N(T-1)}[\kappa_v \text{ diagv}(\mathbf{W}^{u\prime}\mathbf{W}^u)' \text{ diagv}(\mathbf{W}^{u\prime}\mathbf{W}^u)) + \text{tr}((\mathbf{W}^{u\prime}\mathbf{W}^u)(\mathbf{W}^{u\prime}\mathbf{W}^u + \mathbf{W}^u\mathbf{W}^{u\prime}))]; \\
V_{\psi,23} &= N(T-1)\text{cov}\left(\frac{1}{N(T-1)}\overline{\mathbf{v}}'\overline{\mathbf{v}}, \frac{1}{N(T-1)}\mathbf{v}'\overline{\mathbf{v}}\right) \\
&= N(T-1)\text{cov}\left(\frac{1}{N(T-1)}\mathbf{v}'\mathbf{W}^{u\prime}\mathbf{W}^u\mathbf{v}, \frac{1}{N(T-1)}\mathbf{v}'\mathbf{W}^u\mathbf{v}\right) \\
&= \frac{\sigma_v^4}{N(T-1)}[\kappa_v \text{ diagv}(\mathbf{W}^{u\prime}\mathbf{W}^u)' \text{ diagv}(\mathbf{W}^u) + \text{tr}((\mathbf{W}^{u\prime}\mathbf{W}^u)(\mathbf{W}^u + \mathbf{W}^{u\prime}))] \\
&= \frac{\sigma_v^4}{N(T-1)}\text{tr}((\mathbf{W}^{u\prime}\mathbf{W}^u)(\mathbf{W}^u + \mathbf{W}^{u\prime})); \\
V_{\psi,32} &= N(T-1)\text{cov}\left(\frac{1}{N(T-1)}\mathbf{v}'\overline{\mathbf{v}}, \frac{1}{N(T-1)}\overline{\mathbf{v}}'\overline{\mathbf{v}}\right) \\
&= N(T-1)\text{cov}\left(\frac{1}{N(T-1)}\mathbf{v}'\mathbf{W}^u\mathbf{v}, \frac{1}{N(T-1)}\mathbf{v}'\mathbf{W}^{u\prime}\mathbf{W}^u\mathbf{v}\right) \\
&= \frac{\sigma_v^4}{N(T-1)}[\kappa_v \text{ diagv}(\mathbf{W}^u)' \text{ diagv}(\mathbf{W}^{u\prime}\mathbf{W}^u) + \text{tr}(\mathbf{W}^u(\mathbf{W}^{u\prime}\mathbf{W}^u + \mathbf{W}^{u\prime}\mathbf{W}^u))] \\
&= 2\frac{\sigma_v^4}{N(T-1)}\text{tr}(\mathbf{W}^u\mathbf{W}^{u\prime}\mathbf{W}^u); \\
V_{\psi,33} &= N(T-1)\text{cov}\left(\frac{1}{N(T-1)}\mathbf{v}'\overline{\mathbf{v}}, \frac{1}{N(T-1)}\mathbf{v}'\overline{\mathbf{v}}\right) \\
&= N(T-1)\text{cov}\left(\frac{1}{N(T-1)}\mathbf{v}'\mathbf{W}^u\mathbf{v}, \frac{1}{N(T-1)}\mathbf{v}'\mathbf{W}^u\mathbf{v}\right) \\
&= \frac{\sigma_v^4}{N(T-1)}[\kappa_v \text{ diagv}(\mathbf{W}^u)' \text{ diagv}(\mathbf{W}^u) + \text{tr}(\mathbf{W}^u(\mathbf{W}^u + \mathbf{W}^{u\prime}))] \\
&= \frac{\sigma_v^4}{N(T-1)}\text{tr}(\mathbf{W}^u(\mathbf{W}^u + \mathbf{W}^{u\prime})).
\end{aligned}
$$

2.9 Simulation Appendix

2.9.1 Simulation of Connectivity Matrices

The BSJBSA-TSR linked dataset provides the distribution of the number of customers (and respectively suppliers) that a firm has, up to 24 customers (and suppliers). We assign a time-invariant random number for each firm, r_i, which is uniformly distributed in $[0, 1]$, for $i \in \{1, 2, \ldots, N\}$. For the initial period, we use a weakly monotonic mapping function, $q_t^{num}(\cdot)$, to map the firm random number $r_i \in [0, 1]$ to the number of customers, given the empirical distribution. In other words, $q_t^{num}(r_i) = num_{it}$, where $q_t^{num}(\cdot)$ is the inverse of the empirical distribution function of the number of customers in period t. Given the number of customers assigned to each firm in the initial period, we randomly draw its customers from the pool of firms. Subsequently, given the mapping from the firm random number to the number of customers that firm i has at time t, $q_t^{num}(r_i) = num_{it}$, we randomly drop firms from the set of customers that a firm initially has in the previous period if $num_{it} < num_{i,t-1} * persistency_{t-1}$, where $persistency_{t-1}$ is the fraction of firm-to-firm relationships in period $t - 1$ that survive in period t as observed in the data. Alternatively, we add firms (randomly drawn from the pool of unrelated firms) to the set of customers that a firm has in the previous period after attrition (the identity of the connections dropped being randomly drawn from the pool of existing customers of a firm) if $num_{it} > num_{i,t-1} * persistency_{t-1}$. The number of suppliers that a firm has across time is simulated in similar manner.

Given data on the distribution of firms across commuting zones, we use the inverse of the empirical distribution function of commuting zones, $q_t^{cz}(\cdot)$, to map each firm $r_i \in [0, 1]$ to commuting zone in each period, such that $q_t^{cz}(r_i) = cz_{it}$. We then generate the connectivity matrix based on common commuting zone. The ij-th element of the matrix is set equal to 1, if firms i and j are located in the same commuting zone in period t and 0 otherwise. As we vary the number of firms N above 500, we limit the number of connections per firm. Specifically, we generate random numbers $r_{ij} \in [0, 1]$ from a uniform distribution for each firm-pair ij. For each ij-th element of the connectivity matrix that equals 1 in period t, it remains to be 1 if $r_{ij} < 500/N$ and reduces to

0 otherwise.

2.9.2 Simulation of Input and Output Variables

Based on the BSJBSA-TSR linked dataset, we obtain the mean and standard deviation of labor input (and respectively, capital) across firms in each year from 2009 to 2018. We then simulate the usage of labor input (and respectively, capital) for each firm, by drawing randomly from Normal distributions that have the same mean and standard deviation as empirically observed specific to each input variable and year.

For the Monte Carlo simulations, we adopt a Leontief production function as in Ackerberg, Caves and Frazer (2015) such that:

$$VA_{it} = \min \{e^{\alpha_0} L_{it}^{\alpha_l} K_{it}^{\alpha_k} e^{\omega_{it}}, e^{\alpha_m} M_{it}\} e^{\xi_{it}}, \tag{2.38}$$

which gives rise to the following relationship between material inputs and productivity after taking logs:

$$\alpha_m + m_{it} = \alpha_0 + \alpha_l l_{it} + \alpha_k k_{it} + \omega_{it}. \tag{2.39}$$

Setting $e^{\alpha_m} = 1$ as in ACF, we have: $m_{it} = \alpha_0 + \alpha_l l_{it} + \alpha_k k_{it} + \omega_{it}$. The logged output, $y_{it} = \ln Y_{it}$, is then derived using the sum of value-added and material inputs: $y_{it} = \ln(VA_{it}+M_{it})$.

2.9.3 Simulation Procedure

Given simulated data on $\{l_{it}\}_{i=1,t=1}^{i=N,t=T}$, $\{k_{it}\}_{i=1,t=1}^{i=N,t=T}$, $\{\mathbf{W}_t^y\}_{t=1}^{T-1}$, $\{\mathbf{W}_t^l\}_{t=1}^{T-1}$ and $\{\mathbf{W}_t^u\}_{t=1}^{T}$ and the parameter values for $\{\alpha_0, \alpha_l, \alpha_k, \lambda, \beta_l, \rho_1, \mu, \sigma_\xi, \sigma_v\}$, the data used for the simulations are generated as follows:

1. Set $\omega_{i,t-1}$ = 0, for $t = 1$.

2. Generate $va_{i,t-1}$ based on the simulated $l_{i,t-1}$ and $k_{i,t-1}$, the parameter values for $\{\alpha_0, \alpha_l, \alpha_k\}$, the productivity $\omega_{i,t-1}$, and the random draw of $\xi_{i,t-1}$ from a Normal distribution with mean

0 and variance σ_ξ^2.

3. Set $m_{i,t-1}$ according to Equation (2.39) and derive $y_{i,t-1} = \ln(VA_{i,t-1} + M_{i,t-1})$.

4. Generate ω_{it} based on Equations (2.2)–(2.3), given $y_{i,t-1}$, simulated data on $\{\mathbf{W}_t^y\}_{t=1}^{T-1}$, $\{\mathbf{W}_t^l\}_{t=1}^{T-1}$ and $\{\mathbf{W}_t^u\}_{t=1}^{T}$, the parameter values for $\{\lambda, \beta_l, \rho_1, \mu\}$, and the random draw of v_{it} from a Normal distribution with mean 0 and variance σ_v^2.

5. Iterate Steps 2–4 to generate simulated data on $\{va_{it}\}_{i=1,t=1}^{i=N,t=T}$, $\{m_{it}\}_{i=1,t=1}^{i=N,t=T}$, $\{y_{it}\}_{i=1,t=1}^{i=N,t=T}$, and $\{\omega_{it}\}_{i=1,t=1}^{i=N,t=T}$.

Table 2.1: BSJBSA and TSR Matching Percentage

Sample	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
# firms in BSJBSA	29096	29570	30647	30584	30217	30180	30231	30151	29530	29780
# TSR firms matched	26947	27559	28486	28557	28237	28263	28196	28448	27715	27978
Percentage	92.61	93.20	92.95	93.37	93.45	93.65	93.27	94.35	93.85	93.95

Notes: This table reports the percentage of firms in BSJBSA that are matched with its counterpart in TSR. The BSJBSA set of firms is used as the denominator, since it provides the required firm-level variables for productivity estimations.

Table 2.2: Summary Statistics (for 2015 cross section)

(a) BSJBSA-TSR Linked Sample

	Observations	Mean	Std	Min	Max
Labor headcounts	29044	490.48	1940.36	50	130725
Labor hours	29044	794921.01	3111792	0	206818495
Real capital	18580	7205.55	75883.27	0	5394504
Real spending on intermediate inputs	29044	18826.32	127064.81	-4802.70	9208360
Real revenue	29044	21954.38	137121.55	0.93	1.06+07
Real value added	29044	3109.33	15810.43	-51864.47	1417696
Number of customers	27788	6.77	5,54	0	24
Number of suppliers	27788	6.73	4.75	0	24
Number of customers existing in BSJBSA	27788	4.09	3.95	0	23
Number of suppliers existing in BSJBSA	27788	3.73	3.17	0	22

(b) Estimation Sample

	Observations	Mean	Std	Min	Max
Labor headcounts	13001	581.90	2219.38	50	81740
Labor hours	13001	972987.7	3712966	5621.48	162990234
Real capital	13001	8096.92	83753.2	0.17	5394504
Real spending on intermediate inputs	13001	25639.27	161801.6	12.36	9208361
Real revenue	13001	29641.12	176162.1	123.84	10620083
Real value added	13001	4045.19	20239.19	9.67	1417696
Number of customers	12608	7.26	5.74	0	24
Number of suppliers	12608	7.38	4.93	0	24
Number of customers existing in BSJBSA	12608	4.53	4.16	0	23
Number of suppliers existing in BSJBSA	12608	4.26	3.37	0	22

Notes: Refer to Section 2.4 for the measurement of the variables. The revenue, value added, capital, and intermediate inputs are in million Japanese yens. The number of customers / suppliers is identified by the BSJBSA firm's corresponding entries in TSR, which provides the list of a firm's top 24 customers / suppliers. The number of customers / suppliers existing in BSJBSA refers to the subset of a firm's top 24 customers / suppliers listed in TSR that also have firm-level information in BSJBSA.

Table 2.3: DGP1 – No Spatial Dependence in Productivity

Estimator	N	T-1	Stat.	α_l 0.6	α_k 0.4	λ 0	β_l 0	ρ_1 0.5	μ 0	σ_v^2 0.49
WGMM	500	9	Mean	0.6004	0.3997	-	-	0.4993	-	-
			SD	(0.0070)	(0.0102)	-	-	(0.0139)	-	-
			SE	(0.0071)	(0.0105)	-	-	(0.0140)	-	-
	500	18	Mean	0.6001	0.4000	-	-	0.4995	-	-
			SD	(0.0050)	(0.0073)	-	-	(0.0095)	-	-
			SE	(0.0050)	(0.0075)	-	-	(0.0095)	-	-
	750	9	Mean	0.6004	0.3992	-	-	0.4996	-	-
			SD	(0.0059)	(0.0085)	-	-	(0.0119)	-	-
			SE	(0.0058)	(0.0087)	-	-	(0.0114)	-	-
	750	18	Mean	0.6001	0.3998	-	-	0.4998	-	-
			SD	(0.0041)	(0.0059)	-	-	(0.0077)	-	-
			SE	(0.0041)	(0.0061)	-	-	(0.0077)	-	-
	1000	9	Mean	0.6001	0.4001	-	-	0.4996	-	-
			SD	(0.0050)	(0.0073)	-	-	(0.0097)	-	-
			SE	(0.0049)	(0.0073)	-	-	(0.0099)	-	-
	1000	18	Mean	0.5999	0.4001	-	-	0.4998	-	-
			SD	(0.0035)	(0.0051)	-	-	(0.0069)	-	-
			SE	(0.0036)	(0.0053)	-	-	(0.0067)	-	-
SGMM	500	9	Mean	0.6001	0.3997	0.0000	0.0001	0.4991	-0.0010	0.4885
			SD	(0.0211)	(0.0105)	(0.0022)	(0.0109)	(0.0139)	(0.0360)	(0.0118)
			SE	(0.0213)	(0.0110)	(0.0022)	(0.0110)	(0.0140)	(0.0353)	(0.0103)
	500	18	Mean	0.6013	0.4001	0.0001	-0.0001	0.4994	-0.0008	0.4895
			SD	(0.0157)	(0.0075)	(0.0016)	(0.0079)	(0.0095)	(0.0243)	(0.0083)
			SE	(0.0152)	(0.0077)	(0.0016)	(0.0078)	(0.0095)	(0.0254)	(0.0072)
	750	9	Mean	0.6003	0.3993	0.0001	0.0000	0.4995	-0.0011	0.4899
			SD	(0.0169)	(0.0090)	(0.0020)	(0.0088)	(0.0119)	(0.0283)	(0.0097)
			SE	(0.0172)	(0.0090)	(0.0019)	(0.0087)	(0.0115)	(0.0284)	(0.0084)
	750	18	Mean	0.6000	0.3997	0.0001	0.0000	0.4998	-0.0005	0.4898
			SD	(0.0118)	(0.0061)	(0.0014)	(0.0060)	(0.0077)	(0.0213)	(0.0067)
			SE	(0.0122)	(0.0063)	(0.0013)	(0.0062)	(0.0089)	(0.0207)	(0.0059)
	1000	9	Mean	0.6002	0.4001	-0.0001	0.0000	0.4995	0.0005	0.4894
			SD	(0.0143)	(0.0076)	(0.0017)	(0.0073)	(0.0097)	(0.0240)	(0.0084)
			SE	(0.0142)	(0.0077)	(0.0018)	(0.0073)	(0.0099)	(0.0242)	(0.0073)
	1000	18	Mean	0.5997	0.4000	0.0000	0.0001	0.4997	-0.0002	0.4899
			SD	(0.0097)	(0.0053)	(0.0012)	(0.0050)	(0.0069)	(0.0173)	(0.0059)
			SE	(0.0099)	(0.0055)	(0.0012)	(0.0051)	(0.0067)	(0.0177)	(0.0051)

Notes: For each DGP, 1000 simulated samples are drawn and estimated. We report the mean (Mean) and the standard deviation (SD) of the parameter point estimates across the 1000 Monte Carlo simulations, together with the estimated standard errors (SE) derived from the variance-covariance matrices of the estimators. The exact parameter values used in the DGPs are listed in the first row of the table.

Table 2.4: DGP2 – Spatial Dependence in Productivity via Lagged Outputs and Lagged Labor Inputs of Related Firms

				α_l	α_k	λ	β_l	ρ_1	μ	σ_v^2
Estimator	N	T-1	Stat.	0.6	0.4	0.01	0.01	0.5	0	0.49
WGMM	500	9	Mean	0.6233	0.4045	-	-	0.5014	-	-
			SD	(0.0070)	(0.0103)	-	-	(0.0137)	-	-
			SE	(0.0071)	(0.0106)	-	-	(0.0138)	-	-
	500	18	Mean	0.6232	0.4047	-	-	0.5022	-	-
			SD	(0.0050)	(0.0073)	-	-	(0.0095)	-	-
			SE	(0.0050)	(0.0075)	-	-	(0.0094)	-	-
	750	9	Mean	0.6230	0.4026	-	-	0.5011	-	-
			SD	(0.0059)	(0.0085)	-	-	(0.0118)	-	-
			SE	(0.0059)	(0.0088)	-	-	(0.0113)	-	-
	750	18	Mean	0.6231	0.4027	-	-	0.5016	-	-
			SD	(0.0041)	(0.0059)	-	-	(0.0078)	-	-
			SE	(0.0041)	(0.0061)	-	-	(0.0077)	-	-
	1000	9	Mean	0.6219	0.4033	-	-	0.5003	-	-
			SD	(0.0050)	(0.0074)	-	-	(0.0098)	-	-
			SE	(0.0049)	(0.0073)	-	-	(0.0098)	-	-
	1000	18	Mean	0.6221	0.4029	-	-	0.5011	-	-
			SD	(0.0035)	(0.0051)	-	-	(0.0069)	-	-
			SE	(0.0035)	(0.0053)	-	-	(0.0067)	-	-
SGMM	500	9	Mean	0.6003	0.3996	0.0100	0.0101	0.4990	-0.0012	0.4886
			SD	(0.0209)	(0.0106)	(0.0021)	(0.0108)	(0.0137)	(0.0359)	(0.0117)
			SE	(0.0213)	(0.0110)	(0.0022)	(0.0110)	(0.0138)	(0.0353)	(0.0103)
	500	18	Mean	0.6015	0.4001	0.0100	0.0092	0.4994	-0.0008	0.4896
			SD	(0.0156)	(0.0075)	(0.0015)	(0.0079)	(0.0093)	(0.0242)	(0.0083)
			SE	(0.0152)	(0.0077)	(0.0015)	(0.0078)	(0.0094)	(0.0254)	(0.0072)
	750	9	Mean	0.6004	0.3992	0.0101	0.0100	0.4995	-0.0009	0.4899
			SD	(0.0169)	(0.0089)	(0.0019)	(0.0088)	(0.0117)	(0.0282)	(0.0098)
			SE	(0.0172)	(0.0090)	(0.0019)	(0.0087)	(0.0113)	(0.0284)	(0.0084)
	750	18	Mean	0.6001	0.3998	0.0101	0.0100	0.4998	-0.0005	0.4898
			SD	(0.0118)	(0.0061)	(0.0014)	(0.0060)	(0.0077)	(0.0213)	(0.0067)
			SE	(0.0122)	(0.0063)	(0.0013)	(0.0062)	(0.0077)	(0.0207)	(0.0059)
	1000	9	Mean	0.6002	0.4001	0.0099	0.0100	0.4995	0.0004	0.4893
			SD	(0.0143)	(0.0077)	(0.0017)	(0.0073)	(0.0097)	(0.0240)	(0.0084)
			SE	(0.0142)	(0.0077)	(0.0018)	(0.0073)	(0.0098)	(0.0242)	(0.0073)
	1000	18	Mean	0.5997	0.4000	0.0100	0.0101	0.4998	-0.0002	0.4898
			SD	(0.0097)	(0.0054)	(0.0012)	(0.0050)	(0.0068)	(0.0173)	(0.0059)
			SE	(0.0100)	(0.0055)	(0.0012)	(0.0051)	(0.0067)	(0.0177)	(0.0051)

Notes: For each DGP, 1000 simulated samples are drawn and estimated. We report the mean (Mean) and the standard deviation (SD) of the parameter point estimates across the 1000 Monte Carlo simulations, together with the estimated standard errors (SE) derived from the variance-covariance matrices of the estimators. The exact parameter values used in the DGPs are listed in the first row of the table.

Table 2.5: DGP3 – Spatial Dependence in Productivity via Lagged Outputs and Lagged Labor Inputs of Related Firms, and via Productivity Shocks

Estimator	N	T-1	Stat.	α_l 0.6	α_k 0.4	λ 0.01	β_l 0.01	ρ_1 0.5	μ 0.25	σ_v^2 0.49
WGMM	500	9	Mean	0.6233	0.4045	-	-	0.5012	-	-
			SD	(0.0071)	(0.0104)	-	-	(0.0139)	-	-
			SE	(0.0071)	(0.0106)	-	-	(0.0138)	-	-
	500	18	Mean	0.6232	0.4047	-	-	0.5022	-	-
			SD	(0.0050)	(0.0074)	-	-	(0.0095)	-	-
			SE	(0.0050)	(0.0076)	-	-	(0.0094)	-	-
	750	9	Mean	0.6230	0.4026	-	-	0.5011	-	-
			SD	(0.0060)	(0.0086)	-	-	(0.0120)	-	-
			SE	(0.0059)	(0.0088)	-	-	(0.0113)	-	-
	750	18	Mean	0.6185	0.4013	-	-	0.5054	-	-
			SD	(0.0054)	(0.0035)	-	-	(0.0099)	-	-
			SE	(0.0057)	(0.0037)	-	-	(0.0094)	-	-
	1000	19	Mean	0.6219	0.4033	-	-	0.5003	-	-
			SD	(0.0051)	(0.0074)	-	-	(0.0099)	-	-
			SE	(0.0050)	(0.0074)	-	-	(0.0098)	-	-
	1000	18	Mean	0.6221	0.4030	-	-	0.5011	-	-
			SD	(0.0036)	(0.0052)	-	-	(0.0070)	-	-
			SE	(0.0036)	(0.0053)	-	-	(0.0067)	-	-
SGMM	500	9	Mean	0.6004	0.3996	0.0100	0.0100	0.4988	0.2483	0.4885
			SD	(0.0213)	(0.0107)	(0.0021)	(0.0110)	(0.0139)	(0.0359)	(0.0117)
			SE	(0.0215)	(0.0110)	(0.0022)	(0.0111)	(0.0140)	(0.0353)	(0.0103)
	500	18	Mean	0.6015	0.4001	0.0101	0.0093	0.4993	0.2489	0.4896
			SD	(0.0158)	(0.0075)	(0.0015)	(0.0079)	(0.0094)	(0.0240)	(0.0082)
			SE	(0.0153)	(0.0077)	(0.0015)	(0.0078)	(0.0095)	(0.0254)	(0.0072)
	750	9	Mean	0.6003	0.3992	0.0101	0.0100	0.4994	0.2486	0.4899
			SD	(0.0170)	(0.0090)	(0.0019)	(0.0088)	(0.0119)	(0.0282)	(0.0098)
			SE	(0.0172)	(0.0090)	(0.0019)	(0.0088)	(0.0115)	(0.0284)	(0.0084)
	750	18	Mean	0.6000	0.3997	0.0101	0.0100	0.4997	0.2493	0.4898
			SD	(0.0118)	(0.0062)	(0.0014)	(0.0060)	(0.0077)	(0.0213)	(0.0066)
			SE	(0.0122)	(0.0063)	(0.0013)	(0.0062)	(0.0078)	(0.0207)	(0.0059)
	1000	9	Mean	0.6002	0.4001	0.0099	0.0101	0.4994	0.2501	0.4894
			SD	(0.0143)	(0.0077)	(0.0017)	(0.0074)	(0.0099)	(0.0240)	(0.0083)
			SE	(0.0143)	(0.0077)	(0.0018)	(0.0074)	(0.0099)	(0.0242)	(0.0074)
	1000	18	Mean	0.5997	0.4000	0.0100	0.0102	0.4998	0.2498	0.4899
			SD	(0.0098)	(0.0054)	(0.0012)	(0.0050)	(0.0070)	(0.0173)	(0.0058)
			SE	(0.0100)	(0.0055)	(0.0012)	(0.0051)	(0.0067)	(0.0177)	(0.0051)

Notes: For each DGP, 1000 simulated samples are drawn and estimated. We report the mean (Mean) and the standard deviation (SD) of the parameter point estimates across the 1000 Monte Carlo simulations, together with the estimated standard errors (SE) derived from the variance-covariance matrices of the estimators. The exact parameter values used in the DGPs are listed in the first row of the table.

Table 2.6: DGP4 – Stronger Spatial Dependence in Productivity via Lagged Outputs and Lagged Labor Inputs of Related Firms, and via Productivity Shocks

Estimator	N	T	Var	α_l 0.6	α_k 0.4	λ 0.1	β_l 0.1	ρ_1 0.5	μ 0.25	σ_v^2 0.49
WGMM	500	9	Mean	0.8358	0.4801	-	-	0.5938	-	-
			SD	(0.0077)	(0.0124)	-	-	(0.0079)	-	-
			SE	(0.0101)	(0.0172)	-	-	(0.0093)	-	-
	500	18	Mean	0.8277	0.4949	-	-	0.6280	-	-
			SD	(0.0055)	(0.0089)	-	-	(0.0062)	-	-
			SE	(0.0076)	(0.0130)	-	-	(0.0069)	-	-
	750	9	Mean	0.8337	0.4492	-	-	0.5688	-	-
			SD	(0.0063)	(0.0096)	-	-	(0.0072)	-	-
			SE	(0.0080)	(0.0132)	-	-	(0.0080)	-	-
	750	18	Mean	0.8329	0.4513	-	-	0.5945	-	-
			SD	(0.0045)	(0.0070)	-	-	(0.0054)	-	-
			SE	(0.0058)	(0.0098)	-	-	(0.0059)	-	-
	1000	9	Mean	0.8247	0.4408	-	-	0.5457	-	-
			SD	(0.0053)	(0.0083)	-	-	(0.0066)	-	-
			SE	(0.0064)	(0.0105)	-	-	(0.0071)	-	-
	1000	18	Mean	0.8232	0.4478	-	-	0.5772	-	-
			SD	(0.0038)	(0.0061)	-	-	(0.0051)	-	-
			SE	(0.0048)	(0.0080)	-	-	(0.0053)	-	-
SGMM	500	9	Mean	0.6002	0.3996	0.1000	0.1002	0.4993	0.2480	0.4885
			SD	(0.0212)	(0.0108)	(0.0018)	(0.0110)	(0.0074)	(0.0360)	(0.0108)
			SE	(0.0214)	(0.0110)	(0.0018)	(0.0111)	(0.0078)	(0.0352)	(0.0103)
	500	18	Mean	0.6015	0.4001	0.1000	0.0993	0.4998	0.2488	0.4896
			SD	(0.0157)	(0.0075)	(0.0013)	(0.0080)	(0.0058)	(0.0242)	(0.0077)
			SE	(0.0153)	(0.0077)	(0.0013)	(0.0079)	(0.0057)	(0.0254)	(0.0072)
	750	9	Mean	0.6004	0.3992	0.1001	0.1000	0.5001	0.2484	0.4899
			SD	(0.0170)	(0.0090)	(0.0016)	(0.0088)	(0.0067)	(0.0282)	(0.0091)
			SE	(0.0172)	(0.0090)	(0.0016)	(0.0088)	(0.0067)	(0.0283)	(0.0084)
	750	18	Mean	0.6000	0.3998	0.1001	0.1001	0.4998	0.2492	0.4898
			SD	(0.0118)	(0.0061)	(0.0011)	(0.0061)	(0.0048)	(0.0213)	(0.0062)
			SE	(0.0122)	(0.0063)	(0.0011)	(0.0063)	(0.0048)	(0.0207)	(0.0059)
	1000	9	Mean	0.6002	0.4001	0.1000	0.1001	0.4996	0.2500	0.4896
			SD	(0.0144)	(0.0077)	(0.0014)	(0.0074)	(0.0060)	(0.0240)	(0.0079)
			SE	(0.0143)	(0.0077)	(0.0015)	(0.0074)	(0.0060)	(0.0242)	(0.0073)
	1000	18	Mean	0.5997	0.4000	0.1000	0.1001	0.5001	0.2496	0.4899
			SD	(0.0097)	(0.0054)	(0.0010)	(0.0051)	(0.0043)	(0.0173)	(0.0056)
			SE	(0.0100)	(0.0055)	(0.0010)	(0.0052)	(0.0043)	(0.0177)	(0.0051)

Notes: For each DGP, 1000 simulated samples are drawn and estimated. We report the mean (Mean) and the standard deviation (SD) of the parameter point estimates across the 1000 Monte Carlo simulations, together with the estimated standard errors (SE) derived from the variance-covariance matrices of the estimators. The exact parameter values used in the DGPs are listed in the first row of the table.

Table 2.7: DGP5 – Negative Spatial Dependence in Productivity via Lagged Outputs and Lagged Labor Inputs of Related Firms, and via Productivity Shocks

Estimator	N	T	Var	α_l 0.6	α_k 0.4	λ -0.1	β_l -0.1	ρ_1 0.5	μ 0.25	σ_v^2 0.49
WGMM	500	9	Mean	0.3802	0.3566	-	-	0.5133	-	-
			SD	(0.0072)	(0.0107)	-	-	(0.0096)	-	-
			SE	(0.0083)	(0.0131)	-	-	(0.0097)	-	-
	500	18	Mean	0.3819	0.3553	-	-	0.5324	-	-
			SD	(0.0051)	(0.0076)	-	-	(0.0074)	-	-
			SE	(0.0059)	(0.0095)	-	-	(0.0077)	-	-
	750	9	Mean	0.3820	0.3710	-	-	0.5032	-	-
			SD	(0.0061)	(0.0091)	-	-	(0.0084)	-	-
			SE	(0.0067)	(0.0105)	-	-	(0.0081)	-	-
	750	18	Mean	0.3816	0.3713	-	-	0.5204	-	-
			SD	(0.0043)	(0.0064)	-	-	(0.0061)	-	-
			SE	(0.0048)	(0.0075)	-	-	(0.0062)	-	-
	1000	9	Mean	0.3877	0.3738	-	-	0.4956	-	-
			SD	(0.0052)	(0.0079)	-	-	(0.0070)	-	-
			SE	(0.0055)	(0.0086)	-	-	(0.0071)	-	-
	1000	18	Mean	0.3872	0.3727	-	-	0.5124	-	-
			SD	(0.0037)	(0.0057)	-	-	(0.0057)	-	-
			SE	(0.0040)	(0.0063)	-	-	(0.0054)	-	-
SGMM	500	9	Mean	0.6003	0.3996	-0.1000	-0.0999	0.5001	0.2479	0.4886
			SD	(0.0212)	(0.0108)	(0.0027)	(0.0110)	(0.0087)	(0.0360)	(0.0110)
			SE	(0.0214)	(0.0110)	(0.0028)	(0.0111)	(0.0089)	(0.0352)	(0.0103)
	500	18	Mean	0.6014	0.4001	-0.0999	-0.1008	0.4996	0.2488	0.4894
			SD	(0.0157)	(0.0075)	(0.0021)	(0.0080)	(0.0069)	(0.0242)	(0.0079)
			SE	(0.0153)	(0.0077)	(0.0021)	(0.0079)	(0.0068)	(0.0254)	(0.0072)
	750	9	Mean	0.6004	0.3992	-0.0999	-0.1001	0.4994	0.2485	0.4898
			SD	(0.0170)	(0.0089)	(0.0025)	(0.0088)	(0.0077)	(0.0282)	(0.0092)
			SE	(0.0172)	(0.0090)	(0.0024)	(0.0088)	(0.0075)	(0.0283)	(0.0084)
	750	18	Mean	0.6000	0.3998	-0.0999	-0.0999	0.5000	0.2492	0.4898
			SD	(0.0118)	(0.0062)	(0.0018)	(0.0060)	(0.0055)	(0.0213)	(0.0064)
			SE	(0.0122)	(0.0063)	(0.0018)	(0.0063)	(0.0056)	(0.0207)	(0.0059)
	1000	9	Mean	0.6002	0.4001	-0.1001	-0.1000	0.5001	0.2500	0.4895
			SD	(0.0143)	(0.0077)	(0.0022)	(0.0074)	(0.0064)	(0.0240)	(0.0081)
			SE	(0.0142)	(0.0077)	(0.0022)	(0.0074)	(0.0066)	(0.0241)	(0.0073)
	1000	18	Mean	0.5997	0.4000	-0.1000	-0.0999	0.4997	0.2497	0.4899
			SD	(0.0097)	(0.0054)	(0.0016)	(0.0051)	(0.0051)	(0.0173)	(0.0056)
			SE	(0.0100)	(0.0055)	(0.0016)	(0.0052)	(0.0050)	(0.0177)	(0.0051)

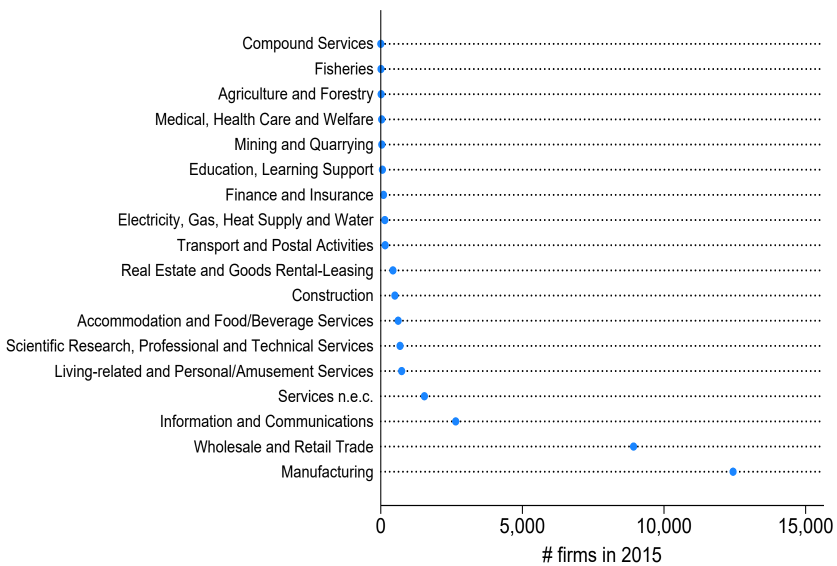
Notes: For each DGP, 1000 simulated samples are drawn and estimated. We report the mean (Mean) and the standard deviation (SD) of the parameter point estimates across the 1000 Monte Carlo simulations, together with the estimated standard errors (SE) derived from the variance-covariance matrices of the estimators. The exact parameter values used in the DGPs are listed in the first row of the table.

Table 2.8: Production Function Estimations (Japanese Firms 2009–2018)

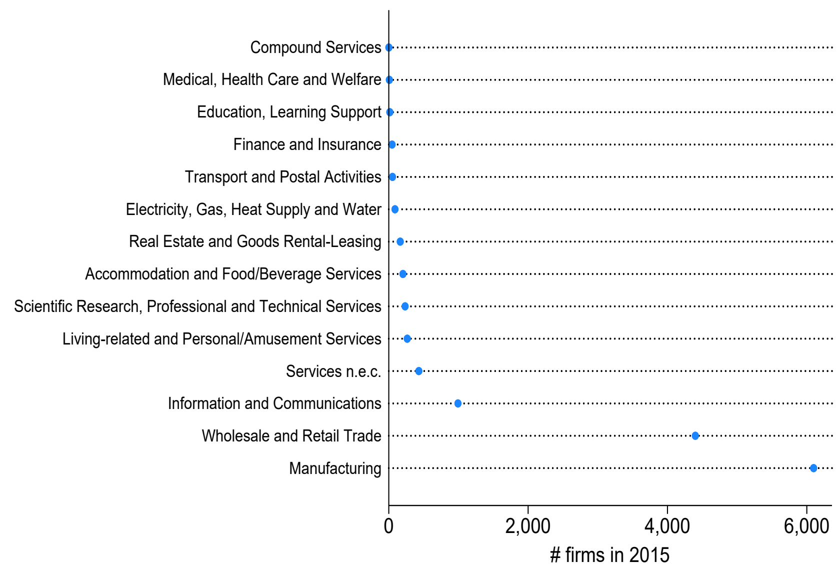
	SGMM			WGMM
	(1)	(2)	(3)	(4)
$\mathbf{W}_{t-1}^{y}$	buyer-seller*CZ	buyer-seller*CZ	buyer-seller*CZ	
$\mathbf{W}_{t-1}^{\Omega}$	CZ	CZ	CZ	
$\mathbf{W}_{t}^{u}$	CZ	buyer-seller*CZ	buyer-seller	
α_0	-22.3467	-20.2377	-22.6221	0.5059
	(1.1313)	(0.9708)	(1.1482)	(1.4718)
α_l	0.8397	0.8045	0.8575	0.4394
	(0.0298)	(0.0264)	(0.0303)	(0.1222)
α_k	0.0368	0.0546	0.0326	0.1854
	(0.0127)	(0.0119)	(0.0130)	(0.0375)
ρ_1	0.9610	0.9569	0.9600	0.9638
	(0.0028)	(0.0030)	(0.0029)	(0.0090)
λ	**0.0048**	**0.0051**	**0.0049**	
	(0.0005)	(0.0005)	(0.0005)	
β	**0.0541**	**0.0540**	**0.0556**	
	(0.0034)	(0.0034)	(0.0035)	
μ	**0.4123**	**0.0227**	**0.0285**	
	(0.0171)	(0.0030)	(0.0030)	
σ_v^2	0.0302	0.0279	0.0321	
	(0.0019)	(0.0018)	(0.0020)	
no. of observations	130,010	130,010	130,010	130,010
no. of firms	13,001	13,001	13,001	13,001

Notes: This table reports the estimations of Equations (2.1)–(2.3) based on the estimation methodology laid out in Section 2.3 and the connectivity matrices defined in Section 2.4.2. The function $h(l_{it}, k_{it}, m_{it})$ in Equation (2.9) is approximated by a second-order polynomial function: $l_{it}^{p}k_{it}^{q}m_{it}^{r}$ for $p+q+r \leq 2$, with nonnegative integers p, q and r. The slope coefficient estimates $\boldsymbol{\delta}$ are omitted from the table above. The function $f(\nu)$ in Equation (2.12) is assumed to be of first order as in the conventional estimator. The list of instruments used for SGMM is: $\mathcal{Z}_{t,I} = (\iota_N, \mathbf{c}_t, \mathbf{c}_{t-1})$ and $\mathcal{Z}_{t,II} = (\iota_N, k_t, \mathbf{c}_{t-1}, \mathbf{W}_{t-1}^{y}y_{t-1}, \mathbf{W}_{t-1}^{l}l_{t-1}, (\mathbf{W}_{t-1}^{y})^2 y_{t-1}, (\mathbf{W}_{t-1}^{l})^2 l_{t-1})$. The list of instruments used for WGMM is the same as those for SGMM, but excluding the related firms' outputs and labor inputs ($\mathbf{W}_{t-1}^{y}y_{t-1}$, $(\mathbf{W}_{t-1}^{y})^2 y_{t-1}$, $\mathbf{W}_{t-1}^{l}l_{t-1}$, $(\mathbf{W}_{t-1}^{l})^2 l_{t-1}$). We iterate the efficient GMM estimation procedure until the set of parameter estimates converges.

Figure 2.1: Number of firms in each industry in 2015

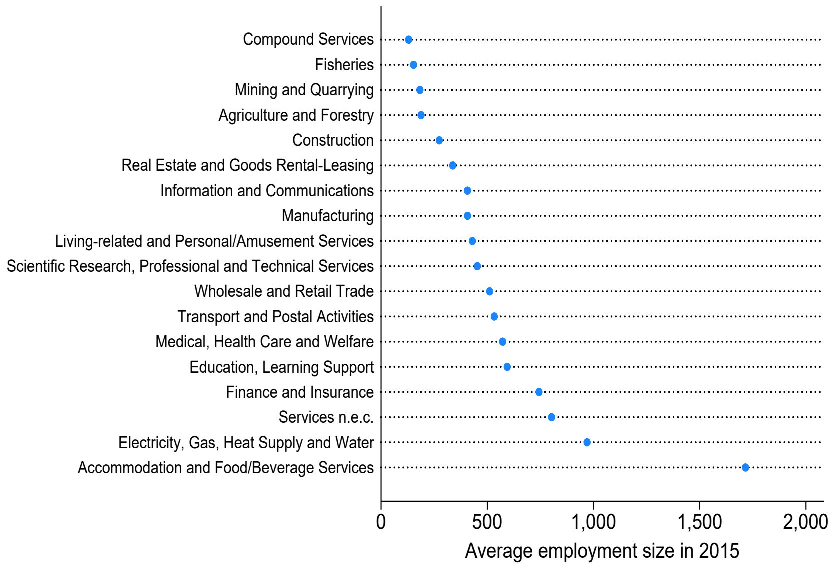


(a) BSJBSA-TSR Linked Sample

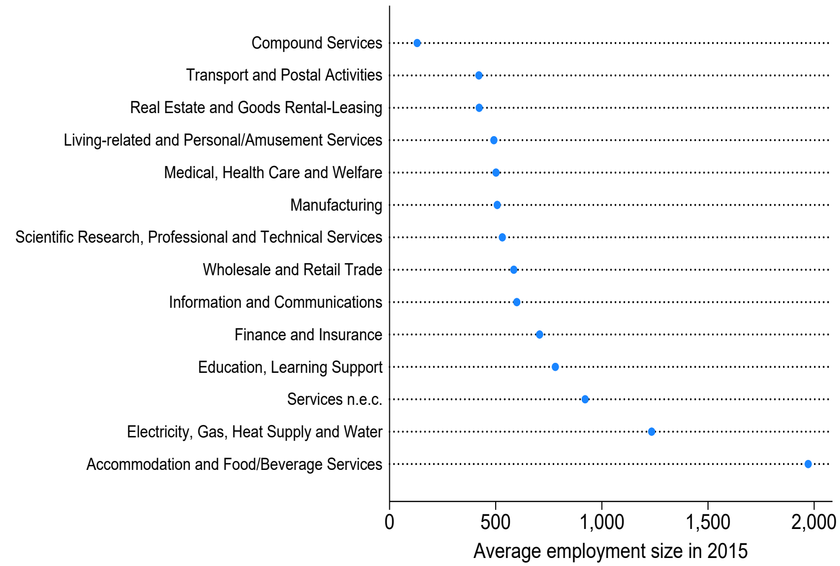


(b) Estimation Sample

Figure 2.2: Average firm's employment in each industry in 2015

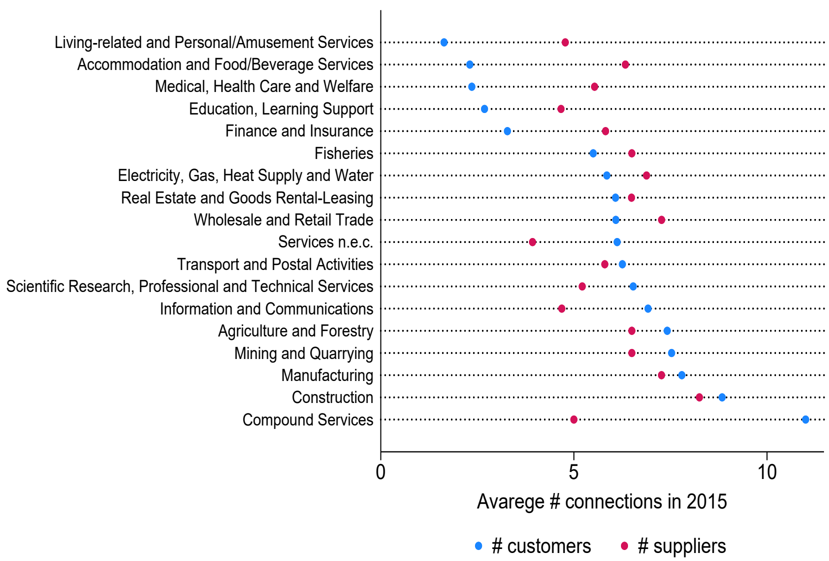


(a) BSJBSA-TSR Linked Sample

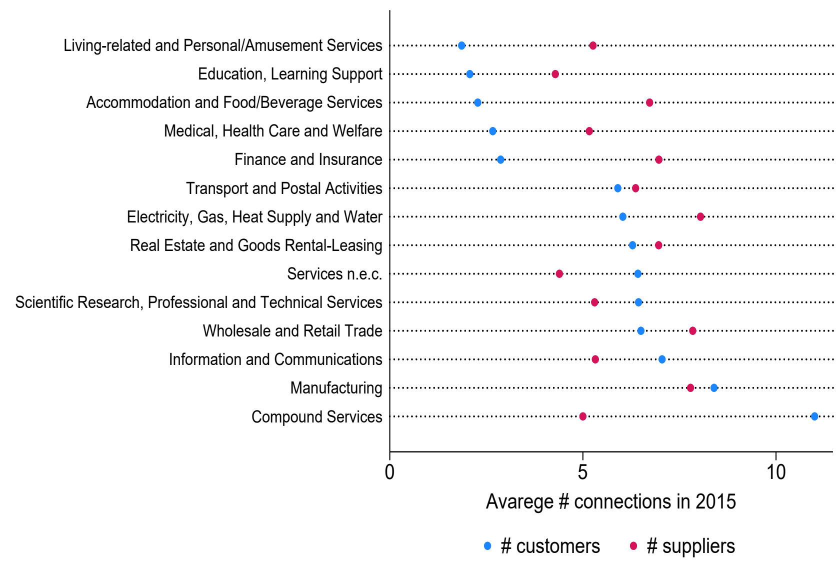


(b) Estimation Sample

Figure 2.3: Average number of connections in each industry in 2015

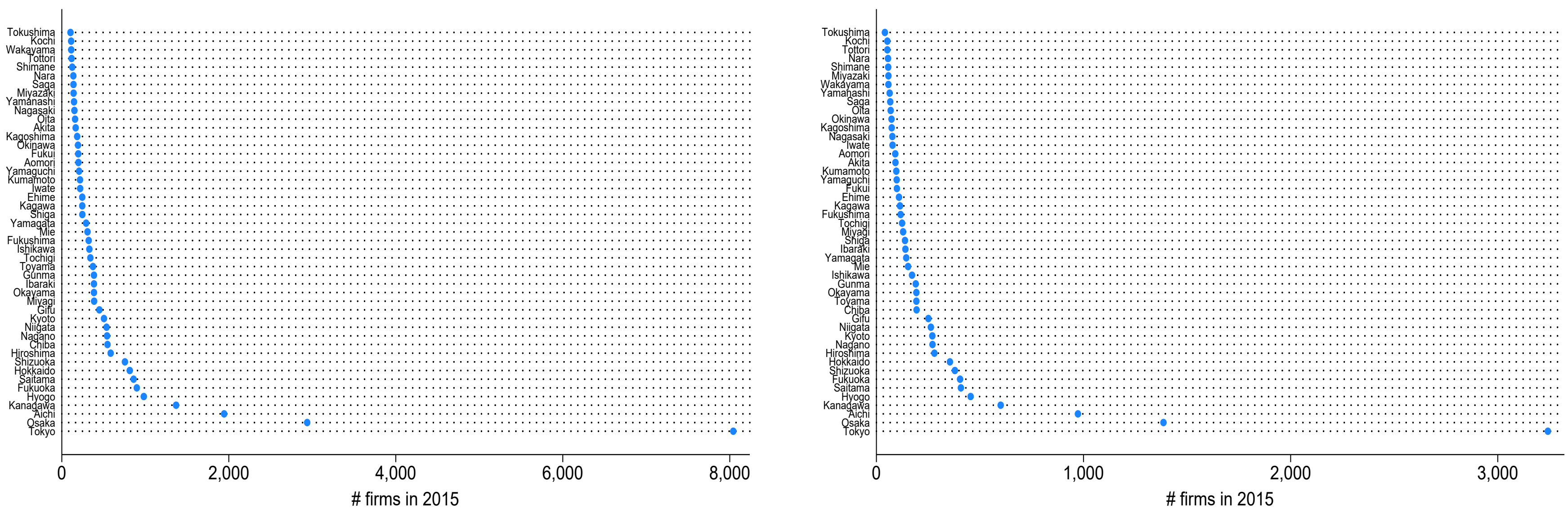


(a) BSJBSA-TSR Linked Sample



(b) Estimation Sample

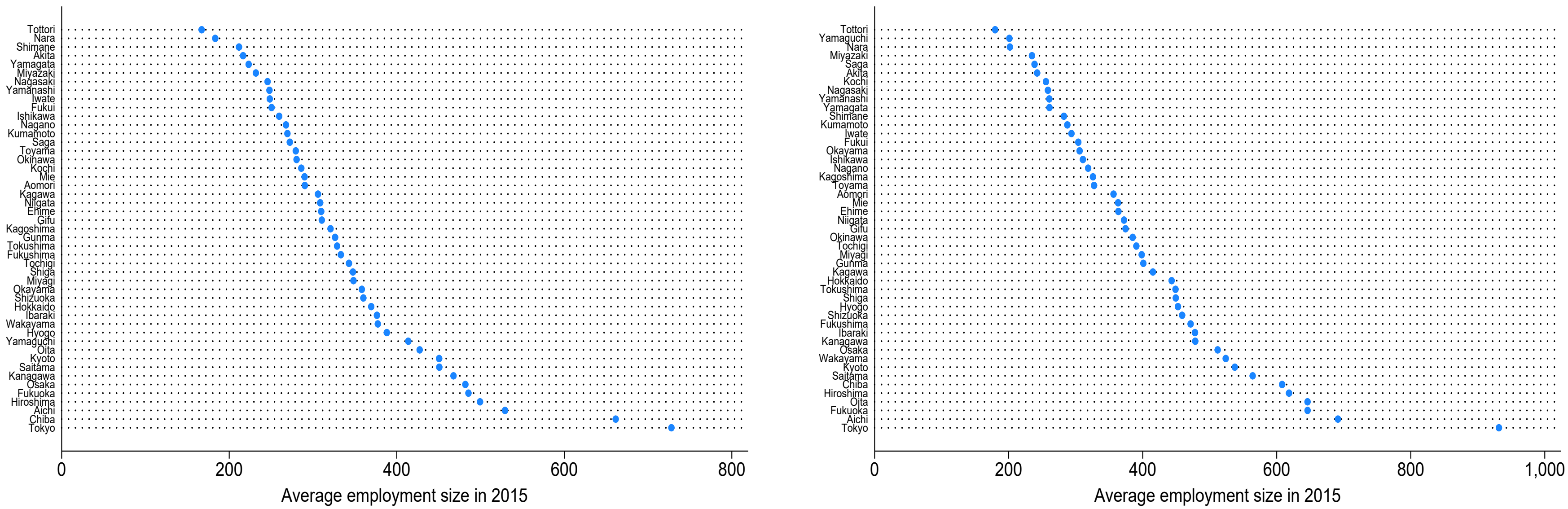
Figure 2.4: Number of firms in each prefecture in 2015



(a) BSJBSA-TSR Linked Sample

(b) Estimation Sample

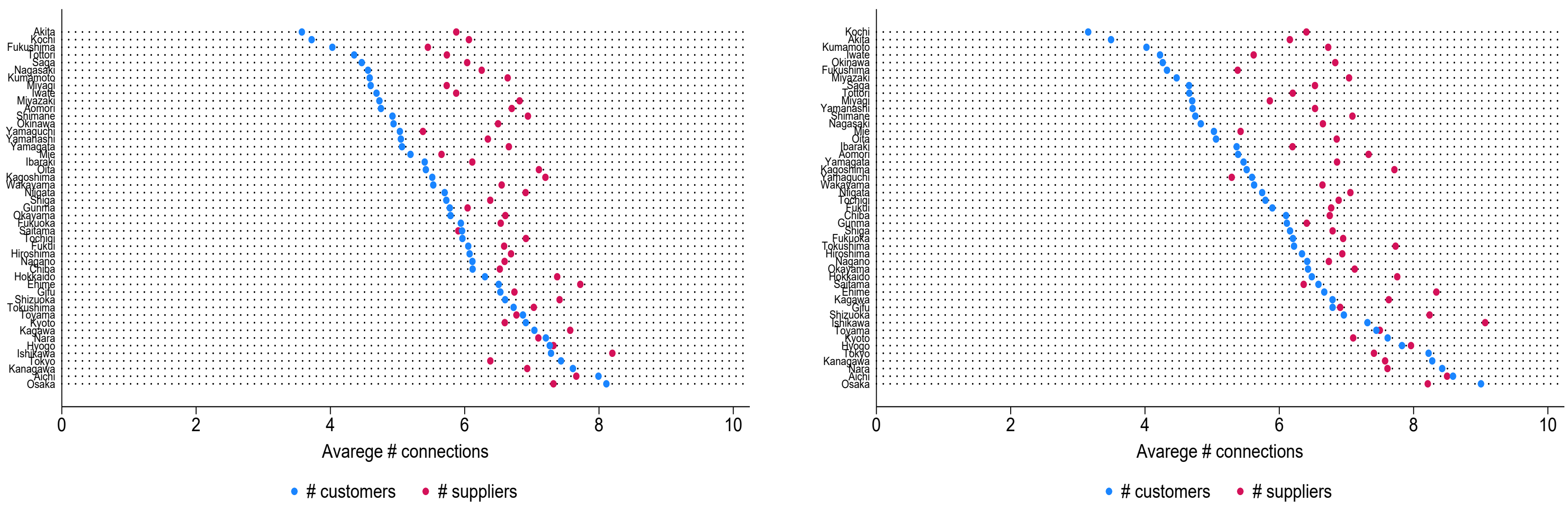
Figure 2.5: Average firm's employment in each prefecture in 2015



(a) BSJBSA-TSR Linked Sample

(b) Estimation Sample

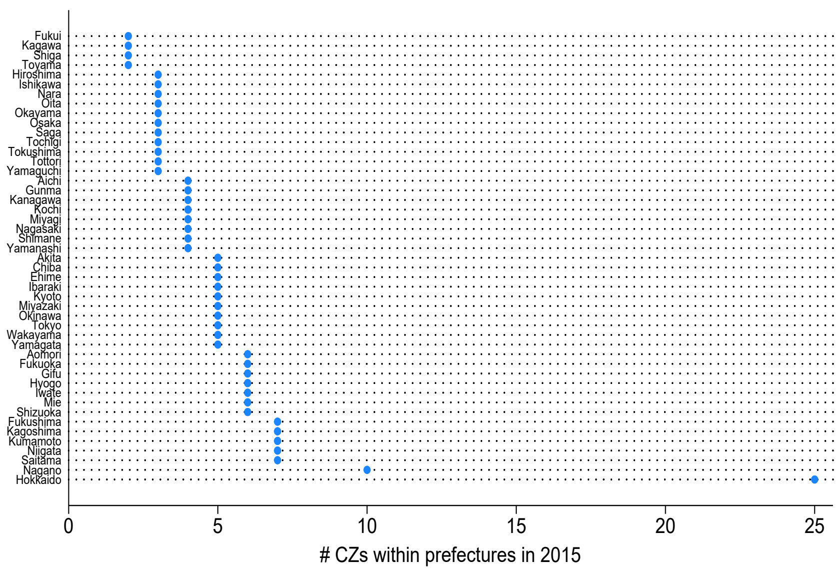
Figure 2.6: Average number of connections in each prefecture in 2015



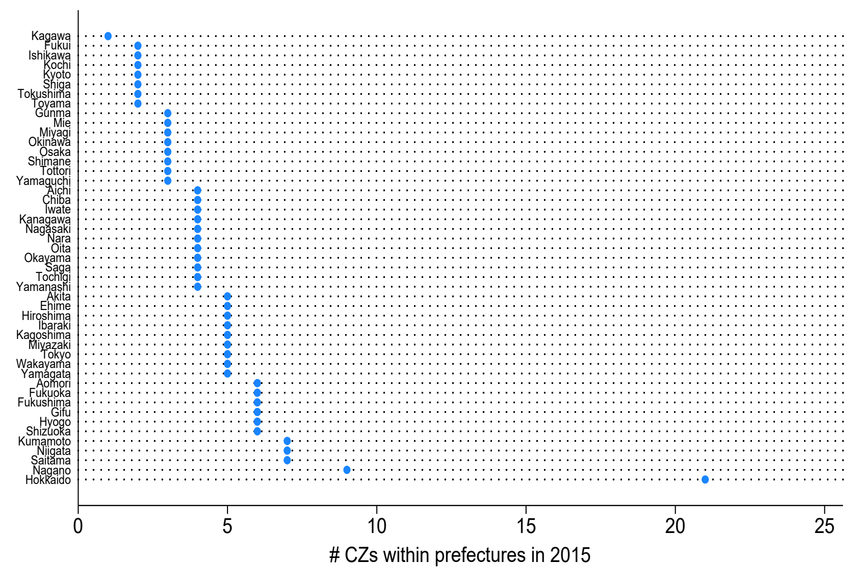
(a) BSJBSA-TSR Linked Sample

(b) Estimation Sample

Figure 2.7: Number of commuting zones in each prefecture in 2015

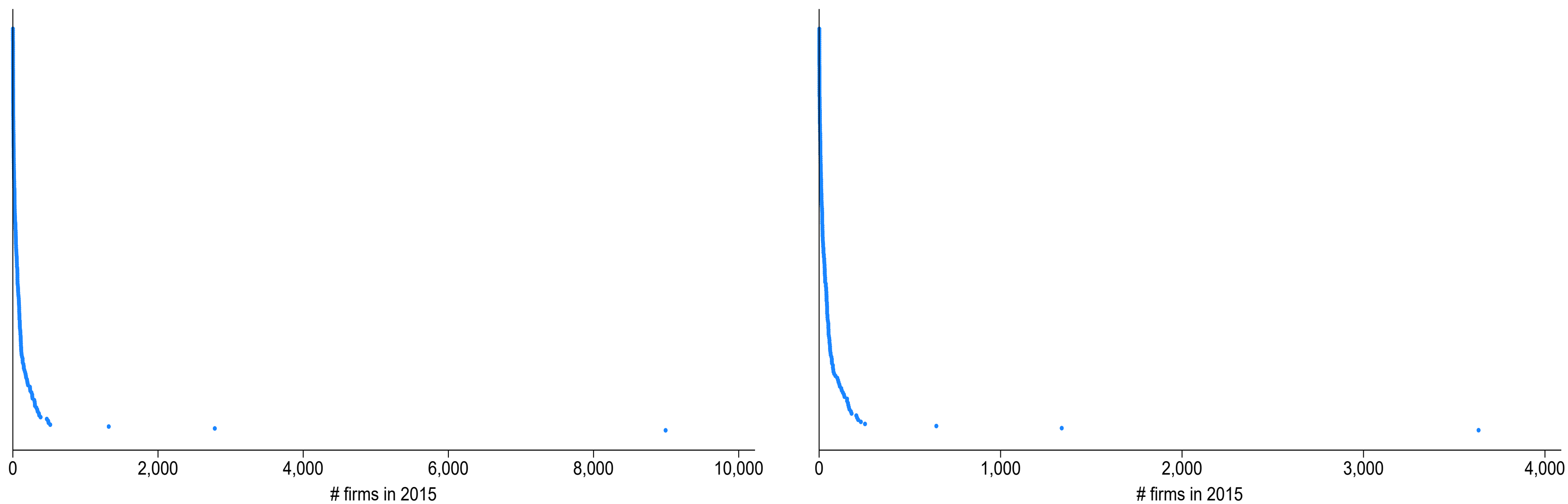


(a) BSJBSA-TSR Linked Sample



(b) Estimation Sample

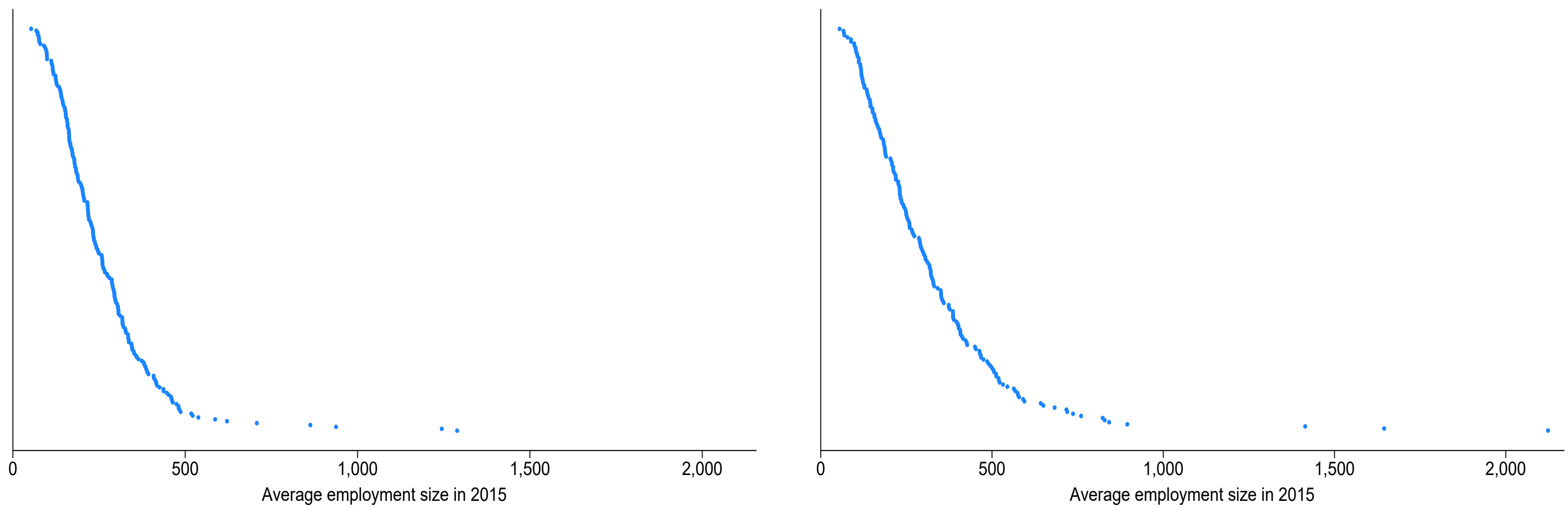
Figure 2.8: Number of firms in each commuting zone in 2015



(a) BSJBSA-TSR Linked Sample

(b) Estimation Sample

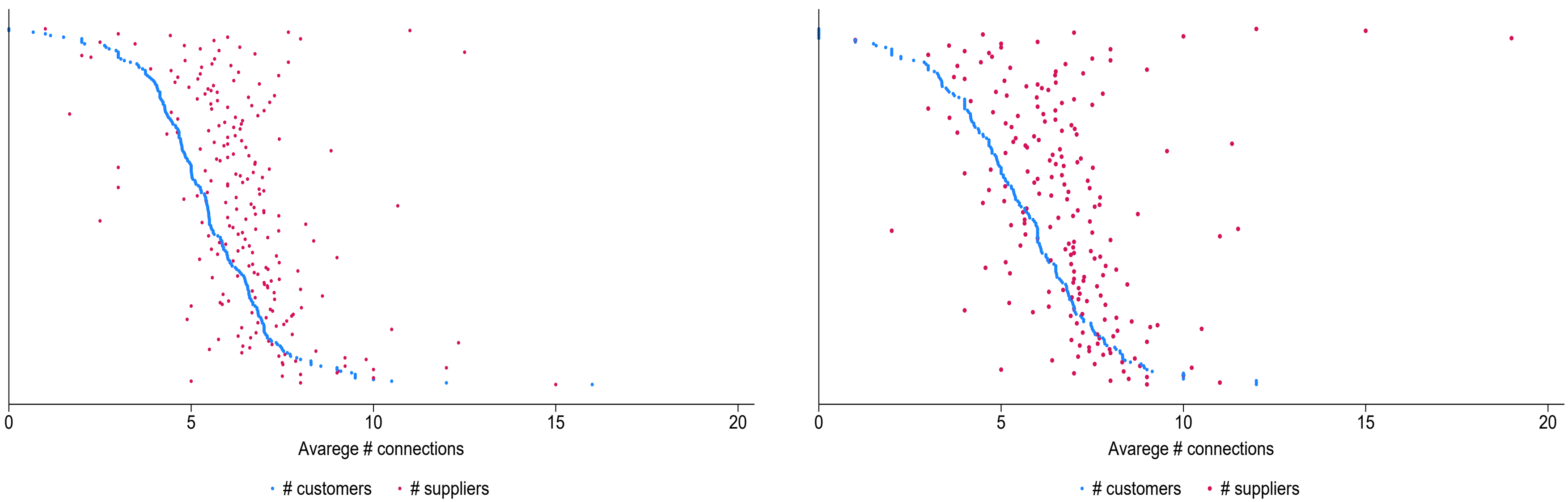
Figure 2.9: Average firm's employment in each commuting zone in 2015



(a) BSJBSA-TSR Linked Sample

(b) Estimation Sample

Figure 2.10: Average number of connections in each commuting zone in 2015



(a) BSJBSA-TSR Linked Sample

(b) Estimation Sample