Econometrica Supplementary Material

SUPPLEMENT TO "MATCHING AND AGGLOMERATION: THEORY AND EVIDENCE FROM JAPANESE FIRM-TO-FIRM TRADE"

(*Econometrica*, Vol. 92, No. 6, November 2024, 1869–1905)

YUHEI MIYAUCHI
Department of Economics, Boston University

APPENDIX A: DATA APPENDIX

IN THIS SECTION of the Supplemental Appendix, I provide additional information about the firm-to-firm trade data from Tokyo Shoko Research (TSR). Table A.I and Figure A.1

TABLE A.I

SAMPLE SIZE AND COVERAGE OF TSR DATA SETS.

		TSR	Economic Census	TSR/Economic Census
2009	All	1,241,490	1,805,545	0.68
	Employment ≤ 4	587,611	1,067,825	0.55
	Employment ≥ 5	653,879	737,720	0.88
2016	All	1,473,991	1,877,438	0.78
	Employment ≤ 4	793,967	1,047,189	0.75
	Employment ≥ 5	680,024	830,249	0.81

Note: This table reports, for year (2009 or 2016) and firm size (over or under 5 employees), the sample size of the TSR data set (third column), the number of firms in Japan based on economic censuses (fourth column), and the ratio of the third and fourth columns (fifth column).

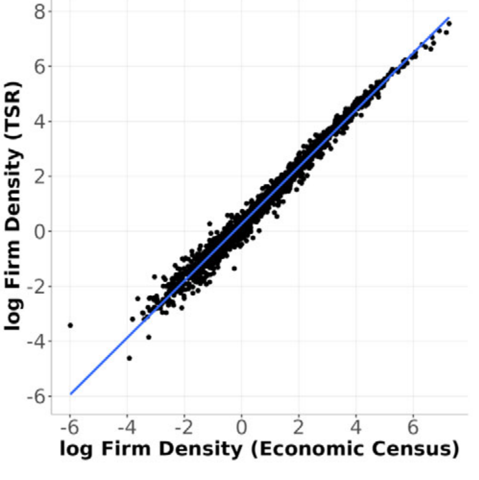

FIGURE A.1.—Coverage of TSR Data sets relative to Economic Census. *Note*: This figure plots the density of firms using two data sources: The economic census on the horizontal axis and TSR data on the vertical axis. Each dot represents a municipality in Japan. All data is from 2009. The straight line in the graph is the linear regression fit between the two variables. The slope of the regression line is 1.04 (with an intercept of 0.27) and the R-squared is 0.98.

Yuhei Miyauchi: miyauchi@bu.edu

 https://doi.org/10.3982/ECTA19697

TABLE A.II

LIST OF REASONS FOR BANKRUPTCIES.

Reason of Bankurptcy	Freq.	Freq. (At Least One Buyer)
Unanticipated Reasons	1588	269
Sales Decline	72,437	9460
Accumulation of Debt	10,700	2267
Spillovers from Other Bankruptcy	6222	938
Shortage of Capital	5577	1021
Management Failure	4843	601
Unknown	3585	458
Over-Investment in Capital	802	211
Deterioration of Credit Conditions	547	161
Difficulty in Collecting Account Receivables	454	126
Over-Accumulation of Inventory	73	29
Total	106,828	15,541

Note: This table reports the number of bankruptcies in each category of reported reasons. The second column ("Freq") reports the number of firms experiencing bankruptcies from 2008 to 2016 for each reason, and the third column ("Freq. (At Least One Buyer)") reports the number of bankrupt firms with at least one buyer (reported as a supplier by at least one firm). In an internal document by TSR, "unanticipated reasons" is described as "unanticipated accidental problems such as the death of representatives, flood disaster, fire, earthquake, traffic accident, fraud, theft, embezzlement, etc."

show the coverage rates of TSR data and show that the TSR data is broadly representative across geography. Table A.II presents the list of reported reasons for bankruptcies from which I identify unanticipated bankruptcies. Figure A.2 presents the frequency of unanticipated bankruptcies across space and time.

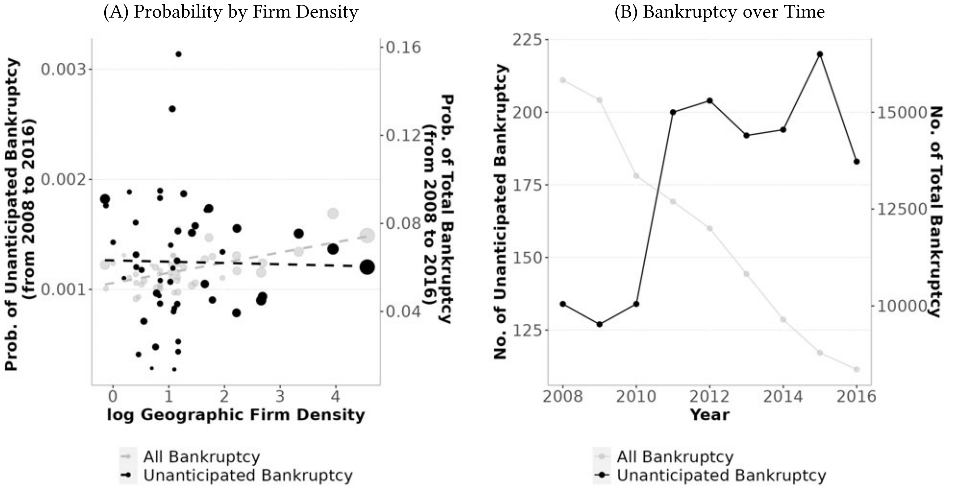

FIGURE A.2.—Spatial and Temporal Patterns of Unanticipated Bankruptcies. *Note*: Panel (A) plots the probability of unanticipated bankruptcies (colored in black; on the left vertical axis) and that of all bankruptcies (colored in gray; on the right vertical axis) against firm density (on the horizontal axis) during the sample period. Each dot represents a prefecture, and the area of the dot represents the log number of firms in the prefecture. Panel (B) plots the frequency of unanticipated bankruptcies (colored in black; on the left vertical axis) and that of all bankruptcies (colored in gray; on the right vertical axis) against year.

APPENDIX B: REDUCED-FORM RESULTS APPENDIX

In this section of the Supplemental Appendix, I present additional results and robustness of the impacts of unanticipated supplier bankruptcies reported in Section 3 of the main paper.

TABLE B.I

IMPACTS OF SUPPLIER BANKRUPTCY ON ADDITIONAL FIRM OUTCOMES.

	Employment Growth (Arc-Elasticity) (1)	Profit/ Sales (2)	Prob. Continuing Relationship with Existing Suppliers in Baseline Period (3)	Sales Growth (Arc-Elasticity) of Existing Suppliers in Baseline Period (4)	Number of New Suppliers of Firms in Same Industry and Prefecture (5)	Sales Growth (Arc-Elasticity) of Firms in Same Industry and Prefecture (6)
Trt × 1[t − BankruptYear = −2 or −3]	−0.006	−0.006	−0.030	−0.007	0.003	−0.002
	(0.010)	(0.008)	(0.048)	(0.005)	(0.016)	(0.005)
Trt × 1[t − BankruptYear = −1]						
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Trt × 1[t − BankruptYear = 0 or 1]	−0.017	−0.012	0.069	−0.004	0.003	0.002
	(0.014)	(0.008)	(0.034)	(0.008)	(0.017)	(0.007)
Trt × 1[t − BankruptYear = 2 or 3]	−0.025	0.006	0.080	0.002	0.001	0.002
	(0.022)	(0.013)	(0.044)	(0.011)	(0.027)	(0.009)
Observations	85,439	67,756	85,951	76,737	80,838	79,141

Note: Results of the stacked-by-event difference-in-difference regression specification (3) by omitting the terms corresponding to treatment heterogeneity ($\widetilde{\log S_{j,k}}$, $\widetilde{\log B_{j,m}}$, X_{fjkmg}). In columns (3) and (4), "existing suppliers in baseline period" indicates the set of suppliers (excluding bankrupt ones) that firms are connected to one year prior to the supplier bankruptcy. In columns (5) and (6), I take the number of new suppliers and sales growth of firms in the same two-digit industry and prefecture as treatment and control firms as outcome variables. The lack of statistically significant results in columns (5) and (6) provide additional robustness of column (4) of Table 2 to address the concern that treatment firms' supplier bankruptcy may have a direct effect on other firms in the same prefecture and industry. See footnote of Table 2 for further details about the specification.

TABLE B.II

HETEROGENEOUS IMPACTS OF SUPPLIER BANKRUPTCY BY REPORTED RANKING AND FIRM SIZE.

	Number of New Suppliers				Sales Growth (Arc-Elasticity)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Trt × 1[t − BankruptYear = −2 or −3]	−0.02	−0.01	−0.02	0.05	−0.01	−0.001	−0.02	−0.002
	(0.03)	(0.04)	(0.04)	(0.06)	(0.01)	(0.01)	(0.01)	(0.02)
Trt × 1[t − BankruptYear = −2 or −3] × X		−0.03	−0.05	0.03		−0.06	−0.01	0.001
		(0.08)	(0.04)	(0.03)		(0.04)	(0.01)	(0.01)
Trt × 1[t − BankruptYear = 0 or 1]	0.14	0.11	0.13	0.12	−0.03	−0.01	−0.03	−0.05
	(0.04)	(0.05)	(0.04)	(0.06)	(0.02)	(0.02)	(0.01)	(0.02)
Trt × 1[t − BankruptYear = 2 or 3]	0.19	0.17	0.19	0.29	−0.04	−0.02	−0.04	−0.05
	(0.06)	(0.07)	(0.06)	(0.09)	(0.02)	(0.03)	(0.02)	(0.03)
Trt × 1[t − BankruptYear = 0 or 1] × X		0.11	−0.11	−0.05		−0.11	0.01	−0.01
		(0.10)	(0.04)	(0.04)		(0.05)	(0.01)	(0.01)
Trt × 1[t − BankruptYear = 2 or 3] × X		0.07	−0.10	0.02		−0.12	−0.01	−0.02
		(0.15)	(0.05)	(0.05)		(0.06)	(0.02)	(0.02)
X		Top 1 Supplier	Firm Size (Buyer)	Firm Size (Supplier)		Top 1 Supplier	Firm Size (Buyer)	Firm Size (Supplier)
Observations	85,951	85,951	85,596	85,762	84,113	84,113	83,897	83,939

Note: Results of the stacked-by-event difference-in-difference regression specification (3) without supplier and buyer density ($\widetilde{\log S_{j,k}}$, $\widetilde{\log B_{j,m}}$) but with treatment heterogeneity (X_{fjkmg}) as specified in the bottom row. Columns (2) and (6) include the dummy variable that takes one if the bankrupt supplier is top 1 supplier. Columns (3) and (7) include the log of employment size of firm f (normalized to mean zero), and columns (4) and (8) include that of bankrupt suppliers. See footnote of Table 2 for further details about the specification.

TABLE B.III

HETEROGENEOUS IMPACTS OF UNANTICIPATED SUPPLIER BANKRUPTCY: ROBUSTNESS.

	Number of New Suppliers				
	(1)	(2)	(3)	(4)	(5)
Panel A: Average Effects					
Trt × 1[t − BankruptYear = −2 or −3]	−0.01	0.01	0.02	−0.01	0.03
	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)
Trt × 1[t − BankruptYear = 0 or 1]	0.15	0.15	0.16	0.14	0.18
	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
Trt × 1[t − BankruptYear = 2 or 3]	0.21	0.22	0.22	0.19	0.28
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Panel B: Heterogeneous Effects					
Trt × 1[t − BankruptYear = −2 or −3] × log Supplier Density	0.04	0.03	0.02	0.04	0.02
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Trt × 1[t − BankruptYear = −2 or −3] × log Buyer Density	0.03	0.06	0.07	0.04	0.02
	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)
Trt × 1[t − BankruptYear = 0 or 1] × log Supplier Density	0.14	0.13	0.11	0.15	0.14
	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)
Trt × 1[t − BankruptYear = 2 or 3] × log Supplier Density	0.17	0.14	0.09	0.15	0.16
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Trt × 1[t − BankruptYear = 0 or 1] × log Buyer Density	−0.03	−0.03	0.02	−0.01	0.001
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Trt × 1[t − BankruptYear = 2 or 3] × log Buyer Density	−0.09	−0.11	−0.05	−0.08	−0.06
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Trt FE × Year FE × Prefecture FE	X	X	X	X	X
Trt FE × Year FE × Buyer and Supplier Size	X	X	X	X	X
Trt FE × Year FE × Firm-Relationship Controls	X	X	X	X	X
Number of Treatment Firms	433	380	374	433	339
Number of Bankrupting Suppliers	181	159	155	181	146
Number of Control Firms	11,889	10,094	7358	11,889	8622
Observations	83,548	75,758	54,442	81,582	62,652

Note: Results of the stacked-by-event difference-in-difference regression specification (3). Panel A corresponds to the specification without treatment heterogeneity as in Table 2, and panel B corresponds to the specification (2) of Table 4. Column (1) excludes firms that drop out during the sample period. Column (2) excludes all supplier bankruptcies in 2009 (the year subsequent to the Great Financial Crisis). Column (3) excludes firms with headquarters in Tokyo prefecture. Column (4) excludes firms whose accounting information is not available after supplier bankruptcy event. Column (5) excludes firms that have establishments outside the headquarter prefecture. Standard errors are clustered at the firm level.

TABLE B.IV

HETEROGENEOUS IMPACTS: ALTERNATIVE SUPPLIER DENSITY.

	Dependent Variable: Number of New Suppliers					
	(1)	(2)	(3)	(4)	(5)	(6)
Trt × 1[t − BankruptYear = −2 or −3] × log Supplier Density	−0.02	0.05	0.05	0.03	0.03	0.05
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)
Trt × 1[t − BankruptYear = −2 or −3] × log Buyer Density	0.05	0.04	0.04	0.04	0.04	0.04
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Trt × 1[t − BankruptYear = 0 or 1] × log Supplier Density	0.11	0.12	0.13	0.14	0.14	0.07
	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)
Trt × 1[t − BankruptYear = 2 or 3] × log Supplier Density	0.15	0.13	0.14	0.15	0.16	0.08
	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.03)
Trt × 1[t − BankruptYear = 0 or 1] × log Buyer Density	−0.04	−0.02	−0.02	−0.03	−0.02	−0.02
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Trt × 1[t − BankruptYear = 2 or 3] × log Buyer Density	−0.10	−0.08	−0.08	−0.08	−0.08	−0.08
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Trt FE × Year FE × Prefecture FE	X	X	X	X	X	X
Trt FE × Year FE × Buyer and Supplier Size	X	X	X	X	X	X
Trt FE × Year FE × Firm-Relationship Controls	X	X	X	X	X	X
Definition of Seller Density	Count Local Suppliers	Sum of Suppliers × Travel Time $^{-0.5}$	Sum of Suppliers × Travel Time $^{-1}$	Sum of Suppliers × Travel Time $^{-2}$	Evaluated Right Before Bankrutpcy	Two-Digit Supplier Industry
Observations	85,407	85,407	85,407	85,407	85,407	85,407

Note: Results of the stacked-by-event difference-in-difference regression specification (3) with alternative definitions for supplier density. Column (1) defines the supplier density by the number of suppliers whose headquarters are established in the treatment firms' prefecture divided by the geographic area. Columns (2) to (4) define it as the sum of the number of suppliers in each prefecture times the power function of travel time between the prefecture and firm f's prefecture divided by the geographic area of firm f's prefecture. Column (5) evaluates the supplier density one year before each supplier bankruptcy, instead of the value in 2008 in my baseline specification. Column (6) defines the industry of suppliers at the two-digit level, instead of four-digit level in baseline. Standard errors are clustered at the firm level.

TABLE B.V

HETEROGENEOUS IMPACTS: ALTERNATIVE BUYER DENSITY.

	Dependent Variable: Number of New Suppliers		
	(1)	(2)	(3)
Trt × 1[t − BankruptYear = −2 or −3] × log Supplier Density	0.04	0.04	−0.04
	(0.03)	(0.03)	(0.04)
Trt × 1[t − BankruptYear = −2 or −3] × log Buyer Density	0.02	0.01	0.11
	(0.03)	(0.03)	(0.04)
Trt × 1[t − BankruptYear = 0 or 1] × log Supplier Density	0.13	0.13	0.14
	(0.04)	(0.04)	(0.04)
Trt × 1[t − BankruptYear = 2 or 3] × log Supplier Density	0.17	0.16	0.16
	(0.06)	(0.06)	(0.06)
Trt × 1[t − BankruptYear = 0 or 1] × log Buyer Density	−0.005	−0.02	−0.01
	(0.04)	(0.03)	(0.03)
Trt × 1[t − BankruptYear = 2 or 3] × log Buyer Density	−0.07	−0.02	0.01
	(0.06)	(0.04)	(0.04)
Definition of Buyers	Buyers in Same Two-Digit Industry	Buyers in Same Four-Digit Industry	Buyers facing Unanticipated Supplier Bankruptcies
Trt FE × Year FE × Prefecture FE	X	X	X
Trt FE × Year FE × Buyer and Supplier Size	X	X	X
Trt FE × Year FE × Firm-Relationship Controls	X	X	X
Observations	85,951	85,882	85,951

Note: Results of the stacked-by-event difference-in-difference regression specification (3) with alternative definitions for buyer density. Column (1) defines buyer density as the density of firms in the same 2-digit industry and prefecture. Column (2) follows the same definition except for using 4-digit industry. Column (3) defines it using the number of firms in the treatment firm's prefecture that faced an unanticipated supplier bankruptcy in the same two-digit industry up to 3 years prior to the event. Standard errors are clustered at the supplier level.

TABLE B.VI

HETEROGENEOUS IMPACTS: INVERSE HYPERBOLIC SINE (IHS) AND REVERSE REPORTING.

	Dependent variable:			
	Number of New Suppliers (IHS)			
	Baseline		Include Reverse Reporting	
	(1)	(2)	(3)	(4)
Trt × 1[t − BankruptYear = −2 or −3]	0.004		−0.04	
	(0.02)		(0.03)	
Trt × 1[t − BankruptYear = −2 or −3] × log Supplier Density		0.01		−0.02
		(0.02)		(0.02)
Trt × 1[t − BankruptYear = −2 or −3] × log Buyer Density		0.03		−0.01
		(0.02)		(0.03)
Trt × 1[t − BankruptYear = 0 or 1]	0.12		0.06	
	(0.02)		(0.03)	
Trt × 1[t − BankruptYear = 2 or 3]	0.15		0.09	
	(0.03)		(0.03)	
Trt × 1[t − BankruptYear = 0 or 1] × log Supplier Density		0.07		0.06
		(0.02)		(0.02)
Trt × 1[t − BankruptYear = 2 or 3] × log Supplier Density		0.05		0.02
		(0.03)		(0.03)
Trt × 1[t − BankruptYear = 0 or 1] × log Buyer Density		−0.01		0.02
		(0.02)		(0.03)
Trt × 1[t − BankruptYear = 2 or 3] × log Buyer Density		−0.05		0.01
		(0.03)		(0.03)
Trt FE × Year FE × Prefecture FE		X		X
Trt FE × Year FE × Buyer and Supplier Size		X		X
Trt FE × Year FE × Firm-Relationship Controls		X		X
Observations	85,951	85,407	85,965	85,421

Note: This table reports the robustness of the results in Table 4. Columns (1) and (2) apply the inverse hyperbolic sine (IHS) transformation to the number of new suppliers (outcome variable). Columns (3) and (4) define the number of new suppliers by including the supplier linkages reported by the supplier-side firms, in addition to the buyer-reported suppliers, as in the baseline specification. For the latter, I apply the inverse hyperbolic sine (IHS) transformation because of the fat-tailed distribution of the outcome variable, unlike the buyer-reported suppliers, which is bounded at 24 (Section 2). Standard errors are clustered at the firm level.

TABLE B.VII

AVERAGE IMPACTS OF SUPPLIER BANKRUPTCY DUE TO MANAGEMENT FAILURE.

	Dependent Variable: Number of New Suppliers								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Trt × 1[t − BankruptYear = −2 or −3]	0.05	0.06	0.01	0.05	0.13	0.05	0.05	0.05	−0.13
	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)
Trt × 1[t − BankruptYear = −1]									
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Trt × 1[t − BankruptYear = 0 or 1]	0.23	0.25	0.24	0.23	0.35	0.21	0.23	0.23	0.05
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.01)
Trt × 1[t − BankruptYear = 2 or 3]	0.35	0.38	0.34	0.34	0.60	0.32	0.38	0.35	0.16
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.04)	(0.02)
Specification	Baseline	Exclude Exit Firms	Exclude Tohoku & Hokkaido	Exclude Same-Industry Control Firms	Exclude Indirect Connection to Bankrupt Firms	Exclude Top 1 Supplier	Exclude Top 2 Suppliers	Exclude Ever-Treated Control Firms	Impute Pre-Period Fixed Effects
Number of Treatment Firms	1689	1689	1222	1689	1689	1317	1036	1689	1689
Number of Bankrupt Suppliers	337	337	258	337	337	301	275	337	337
Number of Control Firms	32,841	32,841	25,429	31,061	18,798	25,063	18,559	32,800	32,841
Observations	239,306	231,399	182,227	226,253	131,141	182,531	135,311	239,028	239,251

Note: A version of Table 2 where I instead use supplier bankruptcy due to “Management Failure” in Table A.II.

TABLE B.VIII

AVERAGE IMPACTS OF SUPPLIER BANKRUPTCY DUE TO MANAGEMENT FAILURE.

	Dependent Variable:		
	Sales Growth (Arc-Elasticity) (1)	Exit (2)	Sales Growth (Arc-Elasticity) (3)
Trt × 1[t − BankruptYear = −2 or −3]	−0.001		0.009
	(0.008)		(0.007)
Trt × 1[t − BankruptYear = −1]			
	(0.000)	(0.000)	(0.000)
Trt × 1[t − BankruptYear = 0 or 1]	−0.021	0.005	−0.004
	(0.009)	(0.004)	(0.006)
Trt × 1[t − BankruptYear = 2 or 3]	−0.057	0.018	−0.019
	(0.014)	(0.006)	(0.009)
Samples	All	All	Cond. on Survival
Control Mean	−0.185	0.052	−0.086
Number of Treatment Firms	1689	1689	1689
Number of Bankrupt Suppliers	337	337	337
Number of Control Firms	32,841	32,841	32,841
Observations	234,316	239,306	227,340

Note: A version of Table 3 where I instead use supplier bankruptcy due to "Management Failure" in Table A.II.

TABLE B.IX

HETEROGENEOUS IMPACTS OF SUPPLIER BANKRUPTCY DUE TO MANAGEMENT FAILURE.

	Dependent Variable: Number of New Suppliers							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Pre-period							
Trt × 1[t − BankruptYear = −3 or −2] × log Supplier Density	0.03	0.002	−0.001	−0.02	0.01	−0.01	0.01	0.02
	(0.02)	(0.02)	(0.04)	(0.03)	(0.04)	(0.03)	(0.02)	(0.03)
Trt × 1[t − BankruptYear = −3 or −2] × log Buyer Density	0.05	0.03	0.05	0.02	0.03	0.02	0.03	0.03
	(0.02)	(0.02)	(0.04)	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)
	Panel B: Post-period							
Trt × 1[t − BankruptYear = 0 or 1] × log Supplier Density	0.06	0.05	0.04	0.06	0.09	0.08	0.05	0.07
	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
Trt × 1[t − BankruptYear = 2 or 3] × log Supplier Density	0.10	0.09	0.12	0.11	0.12	0.13	0.09	0.11
	(0.03)	(0.03)	(0.05)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)
Trt × 1[t − BankruptYear = 0 or 1] × log Buyer Density	0.03	−0.01	−0.04	−0.02	−0.02	−0.01	−0.02	0.01
	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)
Trt × 1[t − BankruptYear = 2 or 3] × log Buyer Density	0.07	0.01	0.02	−0.02	0.03	0.01	−0.01	0.05
	(0.04)	(0.04)	(0.06)	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)
Specification			IV: Firm Density in CEO's Birth Prefecture	Exclude Top 1 Supplier	Exclude Top 2 Suppliers	Exclude Tohoku & Hokkaido	Exclude Same-Industry Control Firms	Exclude Indirect Connection to Bankrupt Firms
Trt FE × Year FE × Prefecture FE	X	X	X	X	X	X	X	X
Trt FE × Year FE × Buyer and Supplier Size		X	X	X	X	X	X	X
Trt FE × Year FE × Firm-Relationship Controls		X	X	X	X	X	X	X
Trt FE × Year FE × CEO Birth Prefecture FE			X					
and CEO Characteristics								
Number of Treatment Firms	1689	1689	1566	1317	1036	1222	1689	1689
Number of Bankrupt Suppliers	337	337	325	301	275	258	337	337
Number of Control Firms	32,841	32,841	30,007	25,063	18,559	25,429	31,061	18,798
Observations	239,289	237,113	168,070	180,807	133,963	180,354	224,166	134,734

Note: A version of Table 4 where I instead use supplier bankruptcy arising due to "Management Failure" in Table A.II.

TABLE B.X
HETEROGENEOUS IMPACTS: FIRM SALES.

	Dependent variable:	
	Sales Growth (Arc Elasticity)	
	OLS	OLS
	(1)	(2)
Trt × 1[t − BankruptYear = −2 or −3]	−0.01	
	(0.01)	
Trt × 1[t − BankruptYear = −2 or −3] × log Supplier Density		−0.01
		(0.01)
Trt × 1[t − BankruptYear = −2 or −3] × log Buyer Density		0.01
		(0.01)
Trt × 1[t − BankruptYear = 0 or 1]	−0.03	
	(0.02)	
Trt × 1[t − BankruptYear = 2 or 3]	−0.04	
	(0.02)	
Trt × 1[t − BankruptYear = 0 or 1] × log Supplier Density		0.01
		(0.01)
Trt × 1[t − BankruptYear = 2 or 3] × log Supplier Density		0.01
		(0.02)
Trt × 1[t − BankruptYear = 0 or 1] × log Buyer Density		0.003
		(0.02)
Trt × 1[t − BankruptYear = 2 or 3] × log Buyer Density		−0.01
		(0.02)
Trt FE x Year FE x Prefecture FE		X
Trt FE x Year FE x Buyer and Supplier Size		X
Trt FE x Year FE x Firm-Relationship Controls		X
Observations	84,113	83,723

Note: This table reports the results of the stacked-by-event difference-in-difference regression specification (3) with an outcome variable of firm sales (defined by arc elasticity). Standard errors are clustered at the firm level.

APPENDIX C: MODEL DERIVATIONS

This appendix discusses additional details of the model.

C.1. *First-Order Approximation of Matching Rates*

This appendix derives equation (2) of the main paper. Using the properties of a Poisson process, $\text{NewSuppliers}_{fjkm\Delta} = 1 - \exp(-\eta S_{j,k}^{*\lambda^S} B_{j,m}^{*\lambda^B - 1}\Delta)$. First-order approximation of this equation around $\Delta \approx 0$ yields

$$\text{NewSuppliers}_{fjkm\Delta} \approx \eta \left(S_{j,k}^*\right)^{\lambda^S} \left(B_{j,m}^*\right)^{\lambda^B - 1} \Delta. \tag{C.1}$$

Further, taking the first-order approximation of this equation for $\log S_{j,k}$ and $\log B_{j,m}$ around their means $\log \overline{S}$ and $\log \overline{B}$, yields

$$\text{NewSuppliers}_{fjkm\Delta} \approx \eta \overline{S}^{\lambda^S} \overline{B}^{\lambda^B-1} \Delta \times \left(1 + \lambda^S \widetilde{\log S^*_{j,k}} + \left(\lambda^B - 1\right) \widetilde{\log B^*_{j,m}}\right), \tag{C.2}$$

where $\widetilde{\log S^*_{j,k}} \equiv \log S^*_{j,k} - \log \overline{S}$ and $\widetilde{\log B^*_{j,m}} \equiv \log B^*_{j,m} - \log \overline{B}$. This equation coincides with equation (2) in the main paper.

C.2. *Cutoff for Entry for Sales*

Let denote the final good sales of firms in location j with unit cost c (net of trade cost) when they enter location j as $\varrho_{j,k} c^{-\sigma+1}$, where $\varrho_{j,k}$ is a demand shifter that depends on aggregate equilibrium conditions. Denoting the unit cost threshold of entry as $\overline{c}_{j,k}$, the goods market clearing condition is given by

$$Y^F_{j,k} = S_{j,k} \int_0^{\overline{c}_{j,k}} \varrho_{j,k} c^{-\sigma+1} \theta c^{\theta-1} \overline{c}_{j,k}^{-\theta} \, dc = \frac{\theta}{\theta - \sigma + 1} S_{j,k} \varrho_{j,k} (\overline{c}_{j,k})^{-\sigma+1}.$$

Now, combining with the zero-profit condition for a marginal seller $f_{j,k} w_j = \frac{1}{\sigma} \times \varrho_{j,k} (\overline{c}_{j,k})^{-\sigma+1}$, the measure of suppliers $S_{j,k}$ is solved as

$$S_{j,k} = \frac{\theta - \sigma + 1}{\theta \sigma} \frac{Y^F_{j,k}}{w_j f_{j,k}}. \tag{C.3}$$

Furthermore, the the entry cutoff $\overline{c}_{j,k}$ is given by $\overline{c}_{j,k} = (S_{j,k}/\Omega_{j,k})^{1/\theta}$.

Furthermore, the aggregate sales fixed-cost payment by firms that sell in location j, $\mathcal{F}_{j,k}$, is given by

$$\mathcal{F}_{j,k} = f_{j,k} w_j S_{j,k} = \frac{\theta - \sigma + 1}{\sigma \theta} Y^F_{j,k}. \tag{C.4}$$

Therefore, $\frac{\theta - \sigma + 1}{\sigma \theta}$ fraction of aggregate final goods sales are required as sales fixed-cost payment. Therefore, the share of profit (net of sales fixed cost) to aggregate final sales is given by $\frac{1}{\sigma} - \frac{\theta - \sigma + 1}{\sigma \theta} = \frac{\sigma - 1}{\sigma \theta}$.

C.3. *Inverse Cost Shifter* $\Gamma_{i,m}$

From the assumption of the measure of firms such that $\mu_{i,m}(\varphi) = N_{i,m} \varphi^{-\theta}$, the measure of firms below unit cost c is given by

$$\begin{aligned} H_{i,m}(c) = \Gamma_{i,m} c^\theta &= \int_{p_1,\dots,p_K} \mu_{i,m}\left(\frac{w_i^{\gamma_{L,m}} \prod_{k \in K} p_k{}^{\gamma_{km}}}{c A_{i,m}}\right) \prod_{k \in K} dG^I_{i,k}(p_k) \\ &= \left(N_{i,m} A^\theta_{i,m} w_i^{-\theta \gamma_{L,m}} \prod_{k \in K} \int_{p_k} p_k{}^{-\theta \gamma_{km}} \, dG^I_{i,k}(p_k)\right) c^\theta, \end{aligned} \tag{C.5}$$

where $G^I_{i,k}(\cdot)$ is the steady-state distribution of potential suppliers in location i and sector k, which follows the inverse of the Pareto distribution with an upper bound $\overline{c}_{i,k}$. Given the steady-state match probability $\Lambda_{i,km}$, I have

$$\begin{aligned}\int p_k{}^{-\theta\gamma_{km}}\,dG^I_{i,k}(p_k) &= \Lambda_{i,km}\int_0^{\overline{c}_{i,k}} c^{-\gamma_{km}\theta}\,dG_{i,k}(c) + (1-\Lambda_{i,km})\int_0^{\overline{c}_{i,k}} (\chi c)^{-\gamma_{km}\theta}\,dG_{i,k}(c)\\ &= \Lambda_{i,km}\frac{1}{1-\gamma_{km}}(\overline{c}_{i,k})^{-\theta\gamma_{km}} + (1-\Lambda_{i,km})\frac{\chi^{-\theta\gamma_{km}}}{1-\gamma_{km}}(\overline{c}_{i,k})^{-\theta\gamma_{km}}\\ &= \frac{\chi^{-\theta\gamma_{km}}}{1-\gamma_{km}}(\overline{c}_{i,k})^{-\theta\gamma_{km}}\bigl\{1+\Lambda_{i,km}\bigl(\chi^{\theta\gamma_{km}}-1\bigr)\bigr\}, \qquad (C.6)\end{aligned}$$

where $G_{i,k}(\cdot)$ the CDF of inverse Pareto distribution with upper bound $\overline{c}_{i,k}$, that is, $G_{i,k}(c) = c^\theta/\overline{c}^\theta_{i,k}$. (Note that $\{\Gamma_{n,m}\}$ affect the supplier cost distribution only through $\overline{c}_{i,k}$ because production costs follow a power law distribution.) Combining equations (C.5) and (C.6) lead to equation (8) of the main paper with $\varrho_m \equiv \prod_k \frac{\chi^{-\theta\gamma_{km}}}{1-\gamma_{km}}$.

C.4. *Derivation for Equation (17)*

$$d\log\bigl(1+\Lambda_{j,km}\bigl(\chi^{\theta\gamma_{km}}-1\bigr)\bigr) = \frac{d\log\bigl(1+\Lambda_{j,km}\bigl(\chi^{\theta\gamma_{km}}-1\bigr)\bigr)}{d\log\Lambda_{j,km}}\frac{d\log\Lambda_{j,km}}{d\log\upsilon_{j,km}}\,d\log\upsilon_{j,km},$$

where

$$\frac{d\log\bigl(1+\Lambda_{j,km}\bigl(\chi^{\theta\gamma_{km}}-1\bigr)\bigr)}{d\log\Lambda_{j,km}} = \frac{\Lambda_{j,km}\bigl(\chi^{\theta\gamma_{km}}-1\bigr)}{1+\Lambda_{j,km}\bigl(\chi^{\theta\gamma_{km}}-1\bigr)},$$

and

$$\frac{d\log\Lambda_{j,km}}{d\log\upsilon_{j,km}} = \frac{\rho_{j,km}}{(\upsilon_{j,km}+\rho_{j,km})^2}\frac{\upsilon_{j,km}}{\Lambda_{j,km}} = \frac{\rho_{j,km}}{\upsilon_{j,km}+\rho_{j,km}} = 1-\Lambda_{j,km},$$

and

$$d\log\upsilon_{j,km} = \lambda^S\,d\log S_{j,k} + \bigl(\lambda^B-1\bigr)\,d\log B_{j,m}.$$

From equations (6) and (12), $d\log S_{j,k} = d\log Y^F_{j,k} = d\log L_j$. Furthermore, from the discussion in Section 4.1.3, $d\log B_{j,m} = d\log N_{j,m}$.

APPENDIX D: ADDITIONAL THEORETICAL RESULTS

This appendix provides additional theoretical results. Section D.1 derives the equilibrium in a special case of my model with a single sector. Section D.2 provides sufficient conditions for equilibrium existence and uniqueness. Section D.3 analyzes the planner's problem and highlights the sources of misallocation in the equilibrium.

D.1. *Single Sector Model*

In this section, I derive the system of equations for a special case of my model with a single sector (i.e., $|K| = 1$). Note that in the single sector model final consumption is proportional to intermediate goods absorption ($Y^F_i \propto Y^I_i$) and entry is proportional to population size ($N_i \propto L_i$).[S1] Moreover, from equation (6), $\bar{c}_i \propto (L_i/\Omega_i)^{1/\theta}$. The equilibrium is then summarized by two sets of equilibrium conditions, "buyer access" and "supplier access" equations, analogous to Arkolakis, Huneeus, and Miyauchi (2023).

First, it can be shown that the labor market clearing condition (equation (14)) becomes

$$w_i L_i \psi_i = \sum_{j \in \mathcal{N}} w_j L_j \psi_j \pi_{ij}.$$

This equation corresponds to "buyer-access" equation. Together with the gravity equation of $\pi_{ij} = \Gamma_i \tau_{ij}/\Omega_j$ (equation (9)) and the expression for Γ_i (equation (8)), I have

$$\frac{w_i^{1+\theta\gamma} L_i^{-\theta\epsilon+(1-\gamma)} \Omega_i^{-(1-\gamma)}}{1+\Lambda_i\left(\chi^{\theta(1-\gamma)}-1\right)} = \sum_{j \in \mathcal{N}} K_{ij}^B w_j L_j \Omega_j^{-1}, \tag{D.1}$$

where K_{ij}^B is a constant that only depends on the exogenous variables and parameters. Using the expression of the steady-state match probability (equation (7)),

$$\Lambda_i = \left[1 + K_i^\Lambda L_i^{1-\lambda^S-\lambda^B}\right]^{-1}, \tag{D.2}$$

where K_i^Λ is a constant, and I used the fact that $S_i \propto L_i$ and $B_i \propto N_i \propto L_i$ for single-sector model.

Second, from the definition of $\Omega_j = \sum_i \Gamma_i \tau_{ij}^\theta$ (equation (6)), I have

$$\Omega_j = \sum_{i \in \mathcal{N}} K_{ij}^S w_i^{-\theta\gamma} L_i^{1+\theta\varepsilon-(1-\gamma)} \Omega_i^{1-\gamma}\left(1+\Lambda_i\left(\chi^{\theta(1-\gamma)}-1\right)\right), \tag{D.3}$$

where K_{ij}^S is a constant. This equation corresponds to "supplier-access" equation.

Lastly, population mobility equation (16) is given by

$$w_i^\upsilon L_i^{-1+\upsilon\frac{\theta-\sigma+1}{\theta(\sigma-1)}} \Omega_i^{\frac{\upsilon}{\theta}} = \sum_{j \in \mathcal{N}} K_j^L w_j^\upsilon L_j^{\upsilon\frac{\theta-\sigma+1}{\theta(\sigma-1)}} \Omega_j^{\frac{\upsilon}{\theta}}, \tag{D.4}$$

where K_j^L is a constant.

Together, the single-sector model with exogenous population is characterized by $\{w_i, \Omega_i\}$ that satisfy equations (D.1), (D.2), (D.3). The equilibrium with endogenous population mobility is characterized by $\{w_i, \Omega_i, L_i\}$ that satisfy equations (D.1), (D.2), (D.3), and (D.4).

D.2. *Equilibrium Existence and Uniqueness*

In this appendix, I discuss conditions for equilibrium existence and uniqueness.

Equilibrium Existence. The equilibrium existence of my multisector and location model is immediate from Brower's fixed-point theorem. To see this, equilibrium variables $\{w_i, L_i, \Lambda_{i,km}, \pi_{ij,k}\}$ are bounded under normalization $\sum_i w_i = 1$, and all the mappings are continuous and differentiable.

[S1] See Arkolakis, Huneeus, and Miyauchi (2023) for a related derivation.

Equilibrium Uniqueness Without Population Mobility. Deriving clear analytical results for the equilibrium uniqueness is challenging using my model with multiple sectors. However, one can show that equilibrium is unique (up to scale) without population mobility with a single sector (Appendix D.1). To see this, from Allen, Arkolakis, and Li (2022), equilibrium is unique if the matrix $|B\Gamma^{-1}|$ has a spectral radius equal to or less than one, where

$$\Gamma = \begin{bmatrix} 1+\theta\gamma & -(1-\gamma) \\ 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & -1 \\ -\theta\gamma & 1-\gamma \end{bmatrix}.$$

By invoking the Collatz–Wielandt formula (see Remark 5 in Allen, Arkolakis, and Li (2022)), one can show that the largest eigenvalue of $|B\Gamma^{-1}|$ is one regardless of the parameter values.

Sufficient Conditions for Equilibrium Uniqueness With Population Mobility. I next provide sufficient conditions for uniqueness with population mobility in a single-sector model. Notice that the equilibrium system is not constant elasticity so I apply the results of Allen, Arkolakis, and Li (2022) that allow for the variable elasticity system. It can be show that equilibrium is unique if the spectral radius of $\max_{\overline{\Lambda}_i\in[0,1]}|B\Gamma^{-1}|$ is equal to or less than one, where $\max_{\overline{\Lambda}_i\in[0,1]}$ is the element-by-element maximum and

$$\Gamma = \begin{bmatrix} 0 & 0 & 1 \\ \upsilon & -1+\tilde{\sigma}\theta & -\dfrac{\upsilon}{1-\sigma}-\tilde{\sigma}\theta \\ 1+\theta\gamma & -\theta\epsilon+(1-\gamma)-\overline{\Lambda}_i\big(\lambda^S+\lambda^B-1\big) & -(1-\gamma) \end{bmatrix},$$

$$B = \begin{bmatrix} -\theta\gamma & 1+\theta\epsilon-(1-\gamma)+\overline{\Lambda}_j\big(\lambda^S+\lambda^B-1\big) & 1-\gamma \\ \upsilon & \tilde{\sigma}\theta & -\dfrac{\upsilon}{1-\sigma}-\tilde{\sigma}\theta \\ 1 & 1 & -1 \end{bmatrix},$$

where $\tilde{\sigma} = \frac{\theta-\sigma+1}{\theta(\sigma-1)}$ and $\overline{\Lambda}_j = \frac{\Lambda_j(\chi^{\theta(1-\gamma)}-1)(1-\Lambda_j)}{1+\Lambda_j(\chi^{\theta(1-\gamma)}-1)}$.

Note that while this condition is sufficient, it is not necessary. As discussed in Allen, Arkolakis, and Li (2022), in the context of a variable elasticity system (as in this case), the sufficient condition may be significantly more conservative compared to the necessary and sufficient condition. In fact, although my baseline parametrization in Section 5 does not satisfy the aforementioned condition, I have confirmed that the choice of initial values does not affect my counterfactual simulation results. This indicates that the existence of multiple equilibria is unlikely to pose an issue under my baseline calibration.

D.3. *Planning Problem and Sources of Misallocation*

In this appendix, I discuss how the thick market and congestion externalities lead to inefficiency in equilibrium entry in a special case of a single location and sector by analyzing an optimal planning problem. I focus on the case with single location and sector to provide a clear and straightforward explanation.

Laissez-Faire Equilibrium With a Single Sector and Location. Given that there is single sector, I normalize the wage $w = 1$. The inverse cost shifter, Γ, is simplified from equation (8) as

$$\Gamma = N\overline{c}^{-\theta\gamma}\big(1+\Lambda\big(\chi^{\theta\gamma}-1\big)\big), \tag{D.5}$$

where I normalized $\varrho A^{\theta} = 1$.[S2] The cut-off of marginal cost below which firms enter a location as a seller, $\overline{c}$, and the measure of sellers, S, is simplified from equation (6) as

$$\overline{c} = \Gamma^{-1/\theta}, \qquad S = \Gamma \overline{c}^{\theta}, \tag{D.6}$$

where I normalized $\frac{\theta-\sigma+1}{\theta\sigma}\frac{L}{f} = 1$.[S3] The measure of entrepreneurs, N, is simplified from equation (10) as

$$N = 1, \tag{D.7}$$

where I normalized $\frac{\sigma-1}{\sigma\theta}\frac{L}{F} = 1$. Together, the laissez-faire equilibrium is characterized by $\{\Gamma, \overline{c}, N\}$ that satisfy equations (D.5), (D.6), (D.7) above.

Planning Problem. I consider an optimal taxation problem where the planner has access to taxes for entrepreneurs' entry, sellers' entry, and income tax. First, I assume that the planner imposes labor income tax, τ^W, as a fraction of labor income. Keeping the same normalization that wage $w = 1$, individual post-tax income is given by

$$w^* = 1 - \tau^W. \tag{D.8}$$

Second, I assume that the planner imposes taxes for sales entry, τ^S, as a fraction of fixed-cost payment f. Noting that the aggregate sales are also affected by the income tax above, the seller entry cutoff is modified from equation (D.6) as

$$\overline{c} = \left(\Gamma^{-1}\left(1-\tau^W\right)\left(1-\tau^S\right)\right)^{1/\theta}, \qquad S = \Gamma \overline{c}^{\theta}. \tag{D.9}$$

Third, I assume that the planner imposes taxes for entrepreneurs' entry, τ^N, as a fraction of fixed-cost payment F. The entrepreneur entry condition is modified from equation (D.7) as

$$N = \left(1-\tau^W\right)\left(1-\tau^N\right). \tag{D.10}$$

The planner chooses τ^W, τ^S, τ^N subject to the government's budget constraint

$$\tau^W L + \frac{\tau^S}{1-\tau^S} Sf + \frac{\tau^N}{1-\tau^N} NF = 0,$$

where $\frac{1}{1-\tau^S} Sf$ and $\frac{1}{1-\tau^N} NF$ are the aggregate fixed-cost payment for entrepreneurs' entry inclusive of taxes and that for sellers' entry inclusive of taxes, respectively. From the discussions in Appendix C.2, these two objects are $\frac{\theta-\sigma+1}{\theta\sigma}$ and $\frac{\sigma-1}{\sigma\theta}$ fraction of aggregate final goods sales $(1-\tau^W)L$, respectively. Together, the government budget constraint is rewritten as

$$\frac{\tau^W}{1-\tau^W} = \tau^N \frac{\sigma-1}{\theta\sigma} + \tau^S \frac{\theta-\sigma+1}{\theta\sigma}. \tag{D.11}$$

The welfare in this economy is given by

$$\log U = \log \frac{w^*}{P} = \log\left(1-\tau^W\right) - \frac{1}{1-\sigma}\log\Gamma - \frac{\theta-\sigma+1}{1-\sigma}\log\overline{c}, \tag{D.12}$$

[S2]This normalization, and subsequent normalization in this section, has no effect on any of the results presented in this section. Note also that L, and hence A, is exogenous in the single-location model.

[S3]To derive this equation, note that $Y^F = wL = L$ and $\Omega = \Gamma$ in a single-location case.

where the expression for the price index P followed from equation (15). Together, the planner's problem is given by maximizing equation (D.12) with respect to τ^W, τ^S, τ^N subject to equations (D.8), (D.9), (D.10), and (D.11).

Reformulating the first-order conditions, I can show that the optimal tax system satisfies the following two sets of conditions:

$$(1-\tau^S)(1-\tau^W)-1 = \underbrace{-\frac{(\sigma-1)\gamma}{(\theta-\sigma+1)(1-\gamma)}}_{\text{pooling externality},\leq 0} + \underbrace{\frac{1}{1-\gamma}\frac{\Lambda(\chi^{\theta\gamma}-1)(1-\Lambda)}{1+\Lambda(\chi^{\theta\gamma}-1)}\lambda^S}_{\text{thick-market externality},\geq 0}, \tag{D.13}$$

$$(1-\tau^N)(1-\tau^W)-1 = \underbrace{\frac{\gamma}{1-\gamma}}_{\text{input–output externality},\geq 0} + \underbrace{\frac{1}{1-\gamma}\frac{\Lambda(\chi^{\theta\gamma}-1)(1-\Lambda)}{1+\Lambda(\chi^{\theta\gamma}-1)}(\lambda^B-1)}_{\text{congestion externality},\leq 0}. \tag{D.14}$$

These two equations succinctly summarize the sources of misallocation present in the laissez-faire equilibrium, which the planner aims to correct. Equation (D.13) summarizes the inefficiency in sellers' entry. The first term, labeled "pooling externality," arises from the fact that increased seller entry raises the average cost of suppliers in the matching market. The second term, labeled "thick market externality," is positive if $\gamma > 0$ and $\lambda^S > 0$. Both of these effects disappear when $\gamma = 0$ (i.e., no intermediate inputs), which is consistent with the assumption that there are no matching frictions in the final goods market.

Equation (D.14) summarizes the inefficiency in entrepreneur's entry. The first term, labeled "input–output externality," arises from firms not internalizing the impact of their entry on reducing intermediate goods costs and generating social surplus. This effect is commonly observed in models with firm entry and input–output linkages (e.g., Krugman and Venables (1995) and Antràs, Fort, Gutiérrez, and Tintelnot (2024)). The second term, labeled "congestion externality," is positive if $\gamma > 0$ and $\lambda^B < 1$.

APPENDIX E: MODEL EXTENSIONS

This appendix provides several extensions of my theoretical framework. I explore these alternative specifications as a part of sensitivity analysis in Appendix H.

E.1. *Markups in Intermediate Input Sales*

In this appendix, I consider an alternative setting of the model where firms charge markups for intermediate input sales instead of marginal cost pricing in the main paper. More specifically, I assume that firms apply the same markup ratio $\sigma/(\sigma-1)$ as final goods consumers. This modification changes the condition for the threshold of marginal cost to enter location j and sector k, $\overline{c}_{j,k}$, and the measure of suppliers selling in location j, $S_{j,k}$, as

$$S_{j,k} = \frac{\theta-\sigma+1}{\theta\sigma}\frac{Y^F_{j,k}+\sum_m Y^I_{j,km}}{w_j f_{j,k}\Omega_{j,k}}, \tag{E.1}$$

where the difference from equation (6) is the addition of input demand $\sum_m Y^I_{j,km}$. Similarly, the condition for the free entry of entrepreneurs is replaced from equation (10) as

$$N_{j,m} = \frac{\sigma - 1}{\sigma\theta}\frac{X^F_{j,m} + X^I_{j,m}}{w_j F_{j,m}}, \tag{E.2}$$

where $X^I_{j,m} \equiv \sum_n \pi_{jn,m} \sum_{m'} Y^I_{n,mm'}$. Market clearing conditions are also replaced from equations (13) and (14) as

$$\tilde{X}_{i,m} = \sum_{j\in\mathcal{N}}\left(\sum_{m'\in K}\frac{\sigma-1}{\sigma}Y^I_{j,mm'} + \frac{\sigma-1}{\sigma}Y^F_{j,m'}\right)\pi_{ij,m'}, \tag{E.3}$$

$$\begin{aligned} w_i L_i \psi_i = \sum_k \Bigg(&\gamma_{L,k}\tilde{X}_{i,k} + \frac{\sigma-1}{\sigma\theta}\sum_{j\in\mathcal{N}}\left(Y^F_{j,k} + \sum_m Y^I_{j,km}\right)\pi_{ij,k} \\ &+ \frac{\theta-\sigma+1}{\sigma\theta}\left(Y^F_{i,k} + \sum_m Y^I_{j,km}\right)\Bigg). \end{aligned} \tag{E.4}$$

All other equilibrium conditions remain the same.

E.2. *Forward-Looking Acceptance Decision*

In this appendix, I extend the model by incorporating a forward-looking acceptance decision regarding matching with a supplier. In particular, I assume that buyer-side firms now have the choice to accept or reject a match, leading to the formation of a long-term relationship. This is in contrast to the baseline model discussed in the main text, where firms always form a relationship conditional on a match. I maintain the assumption that suppliers set prices at their marginal cost, meaning that buyers have all the bargaining power. Consequently, suppliers make no profit from the relationships, and the suppliers' forward-looking decision becomes irrelevant in this context.

Let $V^B_{\omega t,k}(c)$ denote the continuation value of a buyer ω engaged in an ongoing relationship with a supplier in sector k with a unit cost of c (net of iceberg cost). Additionally, let $U^B_{\omega t,k}$ represent the value of a firm without an ongoing supplier relationship in sector k. The Bellman equation for the matched buyer can be expressed as follows:

$$\xi V^B_{\omega t,k}(c) = \Pi^F_{\omega t,k}(c) - \rho_{km}\left(V^B_{\omega t,k}(c) - U^B_{\omega t,k}\right) + \dot{V}^B_{\omega t,k}(c), \tag{E.5}$$

where ξ is the discount rate of the firm; $\Pi^F_{\omega t,k}(c)$ is ω's final goods profit when the unit cost of intermediate goods in sector k is c; ρ_{km} is the Poisson rate at which the relationship is destroyed; and $\dot{V}^B_{\omega t,k}(c)$ indicates the time derivative of the value function $V^B_{\omega t,k}(c)$.

The Bellman equation for the unmatched buyer is given by

$$\xi U^B_{\omega t,k} = \Pi^{F,U}_{\omega t,k} + v_{i,km}a_{\omega,k}\int_0^{c^*_{\omega,k}}\left(V^B_{\omega t,k}(c) - U^B_{\omega t,k}\right)dG_{\omega,k}(c) + \dot{U}^B_{\omega t,k}, \tag{E.6}$$

where $\Pi^{F,U}_{\omega t,k}$ is the profit from final goods when the firm does not have a directly matched supplier in sector k; $v_{i,km}$ is the Poisson rate of matching with a supplier; $a_{\omega,k}$ is the unconditional probability that the buyer accepts a match; $c^*_{\omega,k}$ is the threshold of the supplier's unit cost below, which the buyer decides to form a relationship; $G_{\omega,k}(\cdot)$ is the cumulative distribution function of the suppliers' unit cost conditional on match acceptance (such that $G_{\omega,k}(0) = 0$ and $G_{\omega,k}(c^*_{\omega,k}) = 1$); and $\dot{U}^B_{\omega t,k}$ indicates time derivatives of $U^B_{\omega t,k}$.

To solve these Bellman equations analytically, I take a limit of sales fixed cost to zero ($f_{j,k} \to 0$). This assumption implies that the instantaneous profit $\Pi^F_{\omega t,k}(c)$ and $\Pi^{F,U}_{\omega t,k}$ are isoelastic in costs and allows me to derive an analytical solution. Under this assumption, I have

$$\Pi^F_{\omega t,k}(c) = K_{i,m}\left(p_{\omega t,-k} c^{\gamma_{km}}\right)^{1-\sigma}, \tag{E.7}$$

$$\begin{aligned} \Pi^{F,U}_{\omega t,k} &= \int_0^{\overline{c}_{i,k}} K_{i,m}\left(p_{\omega t,-k}(\chi c)^{\gamma_{km}}\right)^{1-\sigma} dG^I_{i,k}(c) \\ &= K_{i,m}\left(p_{\omega t,-k}\chi^{\gamma_{km}}\right)^{1-\sigma} \frac{\theta}{\theta + \gamma_{km}(1-\sigma)} (\overline{c}_{i,k})^{\gamma_{km}(1-\sigma)}, \end{aligned} \tag{E.8}$$

where $\overline{c}_{i,k}$ is the entry cutoff of suppliers in market i as defined by equation (6), $G^I_{i,k}(\cdot)$ denotes the cost distribution of suppliers in market i (which follows the inverse of Pareto distribution with upper bound $\overline{c}_{i,k}$, and $p_{\omega t,-k}$ indicates the component of marginal cost of firm ω other than the component from input sector k.

I assume that buyers set $c^*_{\omega,k}$ to maximize the expected value of the unmatched state.[S4] This implies that $c^*_{\omega,k}$ is determined so that firms are in expectation indifferent between accepting or rejecting a match:

$$E\left[V^B_{\omega t,k}\left(c^*_{\omega,k}\right)\right] = E\left[U^B_{\omega t,k}\right], \tag{E.9}$$

where the expectation is taken with respect to intermediate input cost other than input sector k. Evaluating equation (E.5) at $c = c^*_{\omega,k}$ and taking expectation in the steady state (i.e., $E[\dot{V}^B_{\omega t,k}] = 0$) yields

$$\xi E\left[V^B_{\omega t,k}\left(c^*_{\omega,k}\right)\right] = E\left[\Pi^F_{\omega t,k}\left(c^*_{\omega,k}\right)\right] = K_{i,m} {c^*_{\omega,k}}^{(1-\sigma)\gamma_{km}} E\left[(p_{\omega t,-k})^{1-\sigma}\right], \tag{E.10}$$

Furthermore, by defining $J_{\omega t,k}(c) = V^B_{\omega t,k}(c) - U^B_{\omega t,k}$, equations (E.5) and (E.6) yield

$$(\xi + \rho_{km}) J_{\omega t,k}(c) = \left(\Pi^F_{\omega t,k}(c) - \Pi^{F,U}_{\omega t,k}\right) - v_{i,km} a_{\omega,k} \int_0^{c^*_{\omega,k}} J_{\omega t,k}(c)\, dG_{\omega,k}(c) + \dot{J}_{\omega t,k}(c).$$

Taking the derivative of this equation with respect to c yields

$$(\xi + \rho_{km}) \frac{\partial}{\partial c} J_{\omega t,k}(c) = \gamma_{km}(\sigma - 1) K_{i,m}(p_{\omega t,-k})^{1-\sigma} c^{(1-\sigma)\gamma_{km}-1} + \frac{\partial}{\partial c}\dot{J}_{\omega t,k}(c)$$

By integrating this expression from c to $c^*_{\omega,k}$, I have

$$\begin{aligned} J_{\omega t,k}(c) - J_{\omega t,k}\left(c^*_{\omega,k}\right) = {} & \frac{K_{i,m}(p_{\omega t,-k})^{1-\sigma}}{\xi + \rho_{km}} \left[\left(c^*_{\omega,k}\right)^{(1-\sigma)\gamma_{km}} - c^{(1-\sigma)\gamma_{km}}\right] \\ & + \left[\dot{J}_{\omega t,k}(c) - \dot{J}_{\omega t,k}\left(c^*_{\omega,k}\right)\right]. \end{aligned}$$

[S4] I assume that the cutoff value $c^*_{\omega,k}$ is determined ex ante. If the cutoff were dependent on input prices in other sectors at each time point, obtaining a closed-form solution would be infeasible.

Using this equation, equation (E.6) is rewritten as

$$\begin{aligned}
\xi U^B_{\omega t,k} &= \Pi^{F,U}_{\omega t,k} + v_{i,km} a_{\omega,k} \int_0^{c^*_{\omega,k}} J_{\omega t,k}(c)\, dG_{\omega,k}(c) + \dot{U}_{\omega t,k} \\
&= \Pi^{F,U}_{\omega t,k} - v_{i,km} a_{\omega,k} \frac{K_{i,m}(p_{\omega t,-k})^{1-\sigma}}{\xi + \rho_{km}} \left(c^*_{\omega,k}\right)^{(1-\sigma)\gamma_{km}} \frac{(1-\sigma)\gamma_{km}}{(1-\sigma)\gamma_{km} + \theta} \\
&\quad + v_{i,km} a_{\omega,k} J_{\omega t,k}\left(c^*_{\omega,k}\right) + \int_0^{c^*_{\omega,k}} \left[\dot{J}_{\omega t,k}(c) - \dot{J}_{\omega t,k}\left(c^*_{\omega,k}\right)\right] dG_{\omega,k}(c) + \dot{U}^B_{\omega t,k},
\end{aligned} \tag{E.11}$$

where the last transformation uses the fact that $G_{\omega,k}(\cdot)$ is the inverse of the Pareto distribution with dispersion parameter θ and upper bound $c^*_{\omega,k}$. By taking the expectation of this equation,

$$\xi E\left[U^B_{\omega t,k}\right] = E\left[\Pi^{F,U}_{\omega t,k}\right] - v_{i,km} a_{\omega,k} \frac{K_{i,m} E\left[(p_{\omega t,-k})^{1-\sigma}\right]}{\xi + \rho_{km}} \left(c^*_{\omega,k}\right)^{(1-\sigma)\gamma_{km}} \frac{(1-\sigma)\gamma_{km}}{(1-\sigma)\gamma_{km} + \theta}. \tag{E.12}$$

Together with equations (E.9) and (E.10), and by solving for $E[\Pi^{F,U}_{\omega t,k}]$ using the Pareto distribution of input cost,

$$\begin{aligned}
& K_{i,m} {c^*_{\omega,k}}^{(1-\sigma)\gamma_{km}} E\left[(p_{\omega t,-k})^{1-\sigma}\right] \\
&\quad = K_{i,m} E\left[(p_{\omega t,-k})^{1-\sigma}\right] \left(\chi^{\gamma_{km}}\right)^{1-\sigma} \frac{\theta}{\theta + \gamma_{km}(1-\sigma)} (\overline{c}_{i,k})^{\gamma_{km}(1-\sigma)} \\
&\qquad - v_{i,km} a_{\omega,k} \frac{K_{i,m} E\left[(p_{\omega t,-k})^{1-\sigma}\right]}{\xi + \rho_{km}} \left(c^*_{\omega,k}\right)^{(1-\sigma)\gamma_{km}} \frac{(1-\sigma)\gamma_{km}}{(1-\sigma)\gamma_{km} + \theta} \\
&\iff \quad {c^*_{\omega,k}}^{(1-\sigma)\gamma_{km}} \left[1 - \frac{v_{i,km} a_{\omega,k}}{\xi + \rho_{km}} \frac{(\sigma-1)\gamma_{km}}{\theta - (\sigma-1)\gamma_{km}}\right] \\
&\qquad = \chi^{\gamma_{km}(1-\sigma)} (\overline{c}_{i,k})^{(1-\sigma)\gamma_{km}} \frac{\theta}{\theta + \gamma_{km}(1-\sigma)} \\
&\iff \quad \left(\frac{c^*_{\omega,k}}{\overline{c}_{i,k}}\right)^{\theta} = \chi^{\theta} \left(\frac{\theta + \gamma_{km}(1-\sigma)}{\theta}\right)^{\frac{\theta}{\gamma_{km}(\sigma-1)}} \\
&\qquad \times \left[1 - \frac{v_{i,km} a_{\omega,k}}{\xi + \rho_{km}} \frac{(\sigma-1)\gamma_{km}}{\theta - (\sigma-1)\gamma_{km}}\right]^{\frac{\theta}{(\sigma-1)\gamma_{km}}}.
\end{aligned} \tag{E.13}$$

Now, noting that $a_{\omega,k} = (\frac{c^*_{\omega,k}}{\overline{c}_{i,k}})^{\theta}$ and $a_{\omega,k} \le 1$,

$$a_{\omega,k} = \min\left\{1, \chi^{\theta} \left(\frac{\theta + \gamma_{km}(1-\sigma)}{\theta}\right)^{\frac{\theta}{\gamma_{km}(\sigma-1)}} \left[1 - \frac{v_{i,km} a_{\omega,k}}{\xi + \rho_{km}} \frac{(\sigma-1)\gamma_{km}}{\theta - (\sigma-1)\gamma_{km}}\right]^{\frac{\theta}{(\sigma-1)\gamma_{km}}}\right\}, \tag{E.14}$$

and

$$c^*_{\omega,k} = \overline{c}_{i,k} a^{\frac{1}{\theta}}_{\omega,k}. \tag{E.15}$$

Furthermore, these expressions imply that $c^*_{\omega,k}$ and $a_{\omega,k}$ depend only on firms' location i and the supplier sector k such that $c^*_{\omega,k} = c^*_{i,km}$ and $a_{\omega,k} = a_{i,km}$.

E.3. *Sourcing From Intermediaries While Supplier Relationship Exists*

In the main text, I assume that, whenever a firm already has an existing supplier relationship in input sector k, they are in a binding contract so that they cannot source from intermediaries. In this appendix, I consider an alternative setup where firms can source from intermediaries even when they have an ongoing relationship.

Consider a firm ω in j to source from input sector k. If firm ω has an existing relationship with a supplier, the input cost from the matched suppliers follows the inverse Pareto distribution with upper bound $\overline{c}$, where I dropped the subscripts for j, k for ease of exposition. If firm ω sources from intermediaries, the input cost follows the inverse Pareto distribution with an upper bound $\chi\overline{c}$. Assuming that these two events are random, the minimum of these two costs, c^*, is always below $\overline{c}$ and distributed according to

$$\Pr\left[c^* < \tilde{c}^*\right] = 1 - \Pr\left[c > \tilde{c}^*\right] \Pr\left[\chi c > \tilde{c}^*\right] = 1 - \left(1 - \left(\frac{\tilde{c}^*}{\overline{c}}\right)^{\theta}\right)\left(1 - \left(\frac{\tilde{c}^*}{\overline{c}\chi}\right)^{\theta}\right),$$

and hence,

$$\frac{\partial}{\partial \tilde{c}^*} \Pr\left[c^* < \tilde{c}^*\right] = \theta \frac{\tilde{c}^{*\theta-1}}{\overline{c}^{\theta}}\left[\left(1 + \chi^{-\theta}\right) - 2\chi^{-\theta}\left(\frac{\tilde{c}^*}{\overline{c}}\right)^{\theta}\right].$$

Hence, the expected value of c^* is given by

$$\mathbb{E}\left[c^*\right] = \int_0^{\overline{c}} c^* \frac{\partial}{\partial \tilde{c}^*} \Pr\left[c^* < \tilde{c}^*\right] dc^* = \frac{\theta}{\theta+1}\left(1 + \chi^{-\theta}\right)\overline{c} - \frac{\theta}{2\theta+1} 2\chi^{-\theta}\overline{c}.$$

Note that the expected cost if unmatched with a supplier is given by

$$\mathbb{E}[\chi c] = \int_0^{\chi\overline{c}} \theta c \frac{c^{\theta-1}}{\overline{c}^{\theta}}\, dc = \frac{\theta}{\theta+1}\chi\overline{c}.$$

Therefore, the ratio of the expected cost if matched and unmatched is given by

$$\frac{\mathbb{E}\left[c^*\right]}{\mathbb{E}[\chi c]} = \frac{1}{\chi}\left(1 - \frac{1}{2\theta+1}\chi^{-\theta}\right).$$

Therefore, compared to the main text (where this ratio is $1/\chi$), the expected benefit of direct sourcing is higher than that in my baseline specification.

E.4. *Firm Heterogeneity for Supplier Demand*

In this appendix, I extend the model to incorporate the feature that not all firms possess demand to match with external suppliers. In particular, I assume that only a fraction $\delta_{j,km}$ of firms in location j and sector m have the demand to match with a supplier in sector k. These exogenous parameters, $\delta_{j,km}$, can vary based on j, k, and m. By introducing these additional parameters, I can rationalize the differential patterns of supplier matching between conditional and unconditional on supplier bankruptcy as observed and documented in Appendix G and Table G.I.

The model remains largely unchanged, with the only modification being the replacement of the steady-state probability that a firm in location j and sector m has a direct relationship with a supplier in sector k, denoted as $\Lambda_{j,km}$ in equation (7), with the following expression:

$$\Lambda_{j,km} = \delta_{j,km} \frac{v_{j,km}}{v_{j,km} + \rho_{j,km}}. \tag{E.16}$$

Similarly, the counterfactual equilibrium remains the same as in Appendix F except that equation (F.5) is replaced with

$$\hat{\Lambda}_{j,km} = \frac{(\Lambda_{j,km}/\delta_{j,km})(\hat{S}_{i,k})^{\lambda^S}(\hat{B}_{i,m})^{\lambda^B-1}}{(\Lambda_{j,km}/\delta_{j,km})(\hat{S}_{i,k})^{\lambda^S}(\hat{B}_{i,m})^{\lambda^B-1} + (1 - \Lambda_{j,km}/\delta_{j,km})}. \tag{E.17}$$

To undertake counterfactual simulation using this extended model in my sensitivity analysis (Appendix H), I additionally need to know the values of $\delta_{j,km}$. I calibrate $\delta_{j,km}$ using equation (E.16) with the observed steady-state match probability ($\Lambda_{j,km}$), model-predicted matching rates ($v_{j,km}$), and the observed link separation rates ($\rho_{j,km}$).

APPENDIX F: SYSTEM OF EQUATIONS FOR COUNTERFACTUAL EQUILIBRIUM

To conduct these counterfactual simulations, I follow the exact-hat algebra approach of Dekle, Eaton, and Kortum (2008) and rewrite the counterfactual equilibrium conditions in terms of the unobserved changes in the endogenous variables between the counterfactual and initial equilibria. I denote the value of a variable in the initial equilibrium by x, the value of this variable in the counterfactual equilibrium by x' (with a prime), and the relative change in this variable by $\hat{x} = x'/x$ (with a hat). Consider a counterfactual to change exogenous productivity $\hat{A}_{j,m}$ and iceberg trade costs $\hat{\tau}_{ij,m}$.[S5] Given the values of the production and preference parameters $\{\alpha_k, \gamma_{L,m}, \gamma_{km}, \theta, \sigma, \varepsilon\}$, baseline population size and trade flows $\{L_i, \pi_{ij,k}, Y^F_{i,k}\}$, parameters and baseline variables for firm-to-firm matching $\{\lambda^S, \lambda^B, \chi\}$ and $\{\Lambda_{i,km}\}$, counterfactual equilibrium is computed in terms of the changes in the endogenous variables $\{\hat{\Lambda}_{i,km}, \hat{\Gamma}_{i,k}, \hat{\bar{c}}_{i,k}, \hat{w}_j, \hat{N}_{j,m}, \hat{L}_j, \hat{A}_{j,m}\}$:

(i) Production and Trade Linkages:.

$$\hat{\pi}_{ij,m} = \frac{\hat{\Gamma}_{i,m}(\hat{\tau}_{ij,m})^{\theta}}{\sum_{\ell \in \mathcal{N}} \hat{\Gamma}_{\ell,m}(\hat{\tau}_{\ell j,m})^{\theta} \pi_{\ell j,m}}, \tag{F.1}$$

$$\hat{\Gamma}_{i,m} = \hat{N}_{i,m} \hat{A}^{\theta}_{i,m} \hat{w}_i^{-\theta\gamma_{L,m}} \prod_{k \in K} \hat{\bar{c}}_{i,k}^{-\theta\gamma_{km}} \frac{1 + \Lambda'_{i,km}(\chi^{\theta\gamma_{km}} - 1)}{1 + \Lambda_{i,km}(\chi^{\theta\gamma_{km}} - 1)}, \tag{F.2}$$

$$\hat{\Omega}_{j,m} = \sum_{\ell \in \mathcal{N}} \hat{\Gamma}_{\ell,m}(\hat{\tau}_{\ell j,m})^{\theta} \pi_{\ell j,m}, \qquad \hat{\bar{c}}_{j,k} = \left[\frac{\hat{L}_j}{\hat{\Omega}_{j,k}}\right]^{1/\theta}. \tag{F.3}$$

[S5] The counterfactual to change population size $\hat{L}_j$ follows the same procedure, except that equation (F.11) is replaced by the assumed exogenous values.

(ii) Matching:.

$$\hat{S}_{j,k} = \hat{\Omega}_{j,k}(\hat{\bar{c}}_{j,k})^{\theta}, \qquad \hat{B}_{j,m} = \hat{N}_{j,m}, \tag{F.4}$$

$$\hat{\Lambda}_{j,km} = \Lambda_{j,km}(\hat{S}_{j,k})^{\lambda^S}(\hat{B}_{j,m})^{\lambda^B - 1} + (1 - \Lambda_{j,km}). \tag{F.5}$$

(iii) General Equilibrium:.

$$\hat{N}_{j,m} = \frac{1}{\hat{w}_j} \frac{\sum_i \pi'_{ji,m} Y^{F'}_{i,m}}{\sum_i \pi_{ji,m} Y^{F}_{i,m}}, \tag{F.6}$$

$$\hat{A}_{i,m} = \hat{\bar{A}}_{i,m} \hat{L}_i^{\varepsilon}, \tag{F.7}$$

$$\hat{Y}^F_{i,k} = \hat{w}_i \hat{L}_i, \qquad Y^{I'}_{i,km} = \gamma_{km} \tilde{X}'_{i,m}, \tag{F.8}$$

$$\tilde{X}'_{i,m} = \sum_{j \in \mathcal{N}} \left(\sum_{m' \in K} Y^{I'}_{j,mm'} + \frac{\sigma - 1}{\sigma} Y^{F'}_{j,m'} \right) \pi'_{ij,m'}, \tag{F.9}$$

$$\hat{w}_i = \frac{1}{\hat{L}_i} \frac{\sum_k \left(\gamma_{L,k} \tilde{X}'_{i,k} + \frac{\sigma - 1}{\sigma\theta} \sum_{j \in \mathcal{N}} Y^{F'}_{j,k} \pi'_{ij,k} + \frac{\theta - \sigma + 1}{\sigma\theta} Y^{F'}_{i,k} \right)}{\sum_k \left(\gamma_{L,k} \tilde{X}_{i,k} + \frac{\sigma - 1}{\sigma\theta} \sum_{j \in \mathcal{N}} Y^{F}_{j,k} \pi_{ij,k} + \frac{\theta - \sigma + 1}{\sigma\theta} Y^{F}_{i,k} \right)}, \tag{F.10}$$

$$\hat{L}_j = \frac{(\hat{w}_j / \hat{P}_j)^{\nu}}{\sum_{\ell} (\hat{w}_{\ell} / \hat{P}_{\ell})^{\nu} L_{\ell}}. \tag{F.11}$$

APPENDIX G: MODEL FIT TO UNTARGETED MOMENTS

In this appendix, I discuss additional evidence of model fit.

New Supplier Link Creation Rates Unconditional on Supplier Bankruptcy. In Section 5, I calibrate the matching function elasticities targeting the spatial heterogeneity of new supplier matching rates *conditional on* supplier bankruptcy. In this calibration process, I do not specifically target the spatial heterogeneity of supplier matching rates *unconditional on* supplier bankruptcy. To assess how these untargeted statistics align between the data and model predictions, in Table G.I, I report the regression coefficients of the log of new supplier link creation rates *unconditional on supplier loss* on the log of supplier density ($S^*_{j,k}$, as defined in Section 3) for each sector pairs and (buyer) location. For the model predictions (columns (1) and (3)), I calculate the independent variable using the expression $(1 - \exp(\eta S^{*\lambda^S}_{j,k} B^{\lambda^B}_{j,m})) \times (1 - \Lambda_{j,km})$, where the multiplication by $1 - \Lambda_{j,km}$ reflects the fact that firms with ongoing supplier relationships do not match with new suppliers. For the data (columns (2) and (4)), I calculate the independent variable as the average number of new linkages generated per year and buyer.

Both the model prediction and the data reveal a significant positive relationship between new supplier link creation rates and supplier density. However, it is worth noting

TABLE G.I

SUPPLIER LINK CREATION RATES: MODEL VS. DATA.

	log New Supplier Link Creation Rates (Unconditional on Supplier Loss)			
	Model (1)	Data (2)	Model (3)	Data (4)
log Supplier Density	0.51 (0.06)	0.22 (0.03)	0.66 (0.06)	0.35 (0.03)
Constant	−3.05 (0.13)	−4.33 (0.14)		
Prefecture FE			X	X
Supplier Sector FE			X	X
Observations	1,561,864	1,561,864	1,561,864	1,561,864
Adjusted R^2	0.81	0.17	0.86	0.43

Note: Standard errors are clustered at the supplier sector and prefecture level.

that the coefficient on the log supplier density is larger in the model prediction (0.51, column (1)) compared to the data (0.22, column (2)). Additionally, the intercepts are larger in the model (−3.05) than in the data (−4.33). This evidence indicates that the model tends to overpredict the new supplier creation rates, particularly for locations and sectors with high supplier density.

One potential explanation for this discrepancy lies in the existence of unmodeled heterogeneity in the demand for matching with suppliers. To grasp the intuition, consider a scenario where a subset of firms in the economy has no demand for external suppliers. As a result, these firms never encounter supplier bankruptcies, and thus they are excluded from the samples in Section 3 and do not influence the calibrated matching function elasticities (λ^S and λ^B). However, their presence does impact the heterogeneity of supplier matching rates *unconditional on* supplier bankruptcy.

To address this potential model misspecification, I introduce an extension to the model in Appendix E.4, which incorporates firm heterogeneity regarding the demand for external suppliers across different sectors and locations. In my sensitivity analysis for counterfactual simulations (Appendix H), I demonstrate the robustness of my findings when considering this alternative specification.

Aggregate Sales by Location and Sector. Another untargeted moment in my calibration procedure is the aggregate firm sales by sector and location. Unlike the approach taken by Caliendo and Parro (2014), where they precisely match the world input–output tables at a detailed sector and location level, my calibration only targets the aggregate input–output table (aggregated across locations) and the cross-regional trade patterns.

Table G.II presents the regression coefficients for the log aggregate sales, both predicted by the model and observed in the data, against a measure of supplier access ($\prod_{k\in K}(1+\Lambda_{i,km}(\chi^{\theta\gamma_{km}}-1))$ that appears in equation (8). Both the model prediction and the data exhibit a positive relationship between aggregate sales and the supplier access measure. The regression coefficients are similar, with the model prediction at 2.35 (column (1)) and the data at 2.79 (column (2)). These findings hold even when controlling for prefecture and sector fixed effects (columns (3) and (4)). These results provide further support for the adequacy of my model, particularly the market clearing assumptions used in calibrating the input and final goods demand, $Y^I_{j,km}$ and $Y^F_{j,k}$.

TABLE G.II

AGGREGATE SALES: MODEL VS. DATA.

	log Sales (Normalized)			
	Model (1)	Data (2)	Model (3)	Data (4)
log Supplier Access	2.35	2.79	1.28	1.70
	(1.04)	(1.18)	(0.46)	(0.65)
Prefecture FE			X	X
Sector FE			X	X
Observations	3632	3632	3632	3632
Adjusted R^2	0.45	0.63	0.79	0.89

Note: Standard errors are clustered at the prefecture level. log supplier access is defined by $\prod_{k\in K}(1+\Lambda_{i,km}(\chi^{\theta\gamma_{km}}-1))$.

APPENDIX H: ADDITIONAL RESULTS FOR COUNTERFACTUAL SIMULATIONS

In this appendix, I provide additional results and sensitivity analysis of the counterfactual simulation in Section 5.2.

Increasing Population Size in Tokyo. Figure H.1 presents the results of a counterfactual simulation of increasing Tokyo's population size on average real wages in Japan and in various prefectures within Japan.

Table H.I presents a sensitivity analysis of the same counterfactual simulation regarding the increase in Tokyo's population size. For each model specification, column (a) presents the elasticity of Tokyo's real wage with population size in baseline specification, column

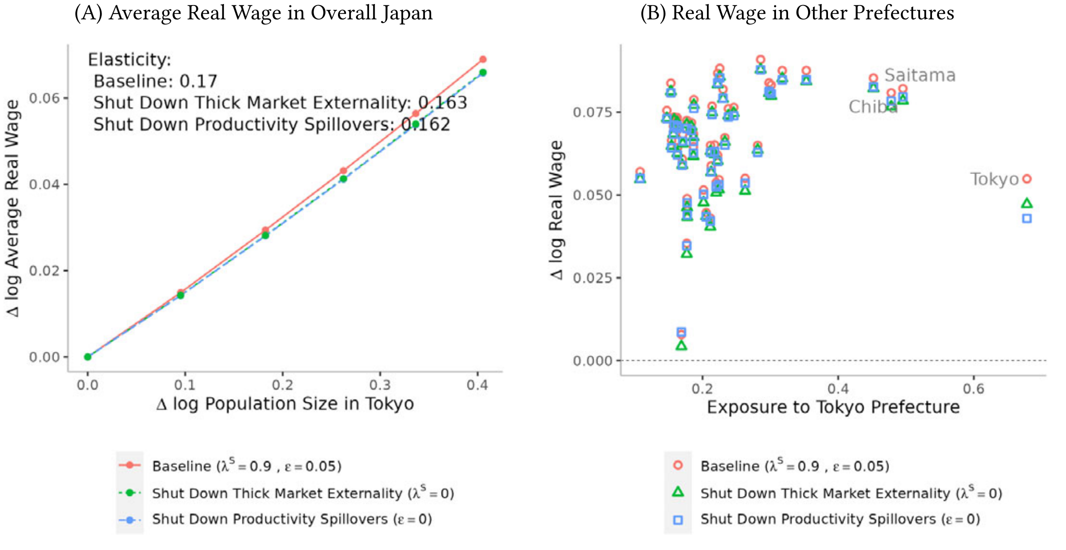

FIGURE H.1.—Effects of Increasing Population Size in Tokyo. *Note*: Results of the counterfactual simulation to increase population size in Tokyo prefecture. Panel (A) plots the log changes in average real wages across Japan against the log change in Tokyo prefecture's population size. Panel (B) shows the log changes in real wages in each prefecture for a 50% increase in Tokyo's population size against the prefecture's exposure to Tokyo prefecture, defined as the share of supplier linkages from Tokyo prefecture.

TABLE H.I

SENSITIVITY ANALYSIS: EFFECTS OF INCREASING POPULATION SIZE IN TOKYO.

Specification	(a) Baseline	(b) $\lambda^S = 0$	Diff. from Baseline	(c) $\varepsilon = 0$	Diff. from Baseline
(1) baseline	0.135	0.116	−0.019	0.106	−0.029
(2) $\lambda^S = 1.39$	0.156	0.116	−0.040	0.126	−0.030
(3) $\lambda^S = 0.46$	0.123	0.116	−0.007	0.093	−0.029
(4) $\lambda^B = 1.29$	0.145	0.120	−0.025	0.115	−0.029
(5) $\lambda^B = 0.60$	0.126	0.114	−0.012	0.096	−0.030
(6) $\chi = 2.48$	0.178	0.116	−0.061	0.148	−0.030
(7) $\chi = 1.15$	0.124	0.116	−0.008	0.095	−0.029
(8) $\sigma = 5$ and $\theta = 5$	−0.094	−0.113	−0.019	−0.124	−0.029
(9) $\sigma = 5$ and $\theta = 10$	0.043	0.019	−0.024	0.008	−0.035
(10) $\varepsilon = 0.1$	0.165	0.146	−0.019	0.106	−0.059
(11) $\varepsilon = 0.03$	0.124	0.105	−0.019	0.106	−0.018
(12) Incorporate Profit for Intermediate Goods Sales	0.136	0.120	−0.016	0.107	−0.030
(13) Introduce Forward-Looking Match Acceptance	0.134	0.116	−0.018	0.105	−0.029
(14) Introduce Firms without Demand for Suppliers	0.150	0.116	−0.034	0.121	−0.029

Note: Results of the counterfactual simulation to increase population size in Tokyo prefecture under alternative specifications indicated by the first column. Column (a) shows the elasticity of Tokyo's real wages with Tokyo's population (following the same definition as Table 7) under baseline specification; Column (b) shows the results under no thick market externality ($\lambda^S = 0$); Column (c) shows the results under no agglomeration productivity spillover ($\varepsilon = 0$). See the main text for further details about each specification.

(b) presents the same elasticity by shutting down the thick market externality, and column (c) presents the same elasticity by shutting down productivity spillovers. Row (1) represents the baseline specification from Table H.1, which serves as the reference point for further analysis. In the subsequent analysis, I specifically focus on the contribution of the thick market externality to this elasticity (column (b); "Diff. from Baseline").

Rows (2) to (7) provide a sensitivity analysis of the matching function elasticities (λ^S and λ^B) and the iceberg cost of indirect sourcing (χ). These alternative parameter values are calibrated based on the same methodology as outlined in Section 5.1.2, targeting plus and minus 1.64 times the standard errors of the point estimates from the reduced-form regression coefficients in columns (3) and (5) of Table 6. As anticipated, the contribution of the thick market externality varies significantly across different parameter values, although it consistently has a negative sign, indicating a positive influence of the thick market externality on agglomeration benefits.

Rows (8) and (9) consider alternative values for σ and θ, which impact the overall agglomeration benefit (column (a)) through the love-of-variety externality from firm entry and the pooling externality discussed in Section 4.3. Notably, with $\sigma = 5$ and $\theta = 5$, the agglomeration benefit becomes negative due to the dominance of the pooling externality over other positive agglomeration externalities. However, the contribution of the thick market externality remains stable in this scenario. Rows (10) and (11) consider different values of agglomeration productivity spillovers ε. The contribution of the thick market externality remains stable.

In row (12), I consider an alternative model specification to incorporate profits from intermediate input sales as discussed in Appendix E.1. In row (13), I consider an alternative model specification to accommodate forward-looking match acceptance decisions

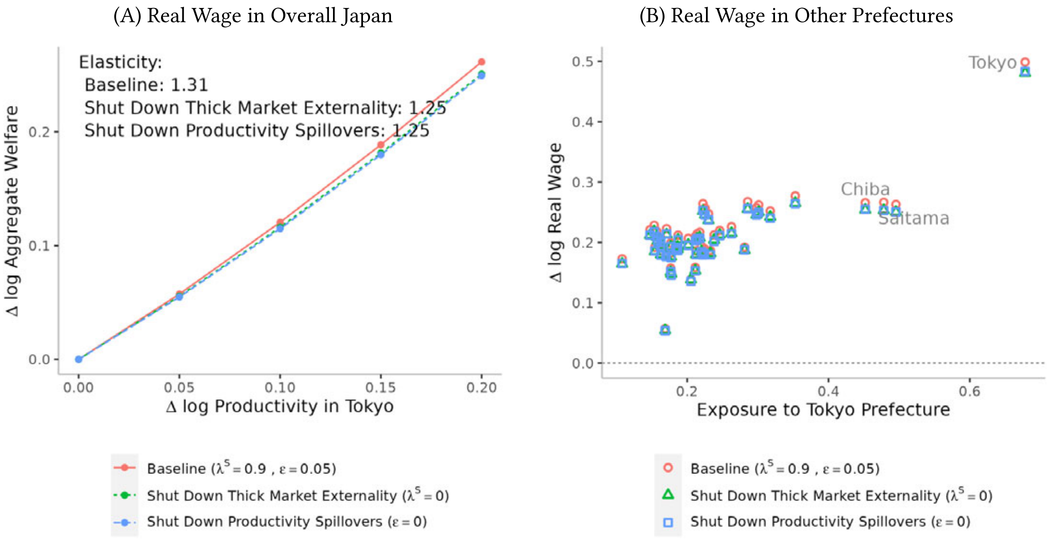

FIGURE H.2.—Effects of Increasing Productivity in Tokyo. *Note*: Results of the counterfactual simulation to increase productivity in Tokyo prefecture. Panel (A) plots the log changes in aggregate welfare in Japan ($\mathcal{U}$) against the log change in Tokyo prefecture's population size. Panel (B) shows the log changes in real wages in each prefecture for a 20% increase in Tokyo's productivity against the prefecture's exposure to Tokyo prefecture, defined as the share of supplier linkages from Tokyo prefecture.

by input buyers as discussed in Appendix E.2.[S6] In both specifications, the contribution of the thick market externality remains similar.

In row (14), I consider an alternative model specification discussed in Appendix E.4, where only a fraction $\delta_{j,km}$ of firms in location j and sector m have demand to match with a supplier in sector k. Interestingly, I observe a somewhat larger contribution of the thick market externality in this specification. This can be attributed to the fact that the changes in matching rates ($v_{j,km}$) triggered by an increase in population size have a more pronounced impact on the steady-state match probability ($\Lambda_{j,km}$) in this model specification (as seen in equation (E.17)). Consequently, the contribution of the thick market externality becomes more significant in this alternative model specification.

Increasing Productivity in Tokyo. Figure H.2 presents the results of a counterfactual simulation to increase Tokyo's productivity on aggregate welfare in Japan and real wages in different prefectures in Japan.

In Table H.II, I provide a sensitivity analysis for the same counterfactual simulation. The patterns observed in the sensitivity analysis closely mirror those observed in Table H.I. Rows (15) introduces an alternative calibration for a smaller value of migration elasticity υ (1 instead of 2). As anticipated, a smaller value of υ leads to a smaller contribution of the thick market externality. This occurs because a smaller migration elasticity induces less population responses to the productivity shock, thereby shrinking the effects of the thick market externality.

[S6]For this simulation, I set the discount rate ζ to 0.03.

TABLE H.II

SENSITIVITY ANALYSIS: EFFECTS OF INCREASING PRODUCTIVITY IN TOKYO.

Specification	(a) Baseline	(b) $\lambda^S = 0$	Diff. from Baseline	(c) $\varepsilon = 0$	Diff. from Baseline
(1) baseline	2.49	2.40	−0.09	2.37	−0.12
(2) $\lambda^S = 1.39$	2.59	2.40	−0.19	2.46	−0.13
(3) $\lambda^S = 0.46$	2.44	2.40	−0.03	2.32	−0.12
(4) $\lambda^B = 1.29$	2.53	2.41	−0.12	2.41	−0.13
(5) $\lambda^B = 0.60$	2.46	2.40	−0.06	2.34	−0.12
(6) $\chi = 2.48$	2.72	2.40	−0.32	2.57	−0.15
(7) $\chi = 1.15$	2.44	2.40	−0.04	2.32	−0.12
(8) $\sigma = 5$ and $\theta = 5$	1.86	1.81	−0.04	1.80	−0.06
(9) $\sigma = 5$ and $\theta = 10$	2.71	2.57	−0.14	2.54	−0.17
(10) $\varepsilon = 0.1$	2.63	2.53	−0.10	2.37	−0.26
(11) $\varepsilon = 0.03$	2.44	2.36	−0.09	2.37	−0.07
(12) Incorporate Profit for Intermediate Goods Sales	2.51	2.44	−0.07	2.39	−0.12
(13) Introduce Forward-Looking Match Acceptance	2.49	2.40	−0.08	2.37	−0.12
(14) Introduce Firms without Demand for Suppliers	2.55	2.40	−0.14	2.42	−0.13
(15) $\upsilon = 1$	2.30	2.25	−0.05	2.26	−0.05

Note: Results of the counterfactual simulation to increase productivity in Tokyo prefecture under alternative specifications indicated by the first column. Column (a) shows the elasticity of Tokyo's real wages with Tokyo's productivity (following the same definition as Table 8) under baseline specification; Column (b) shows the results under no thick market externality ($\lambda^S = 0$); Column (c) shows the results under no agglomeration productivity spillover ($\varepsilon = 0$). See the main text for further details about each specification.

REFERENCES

ALLEN, TREB, COSTAS ARKOLAKIS, AND XIANGLIANG LI (2022): "On the Equilibrium Properties of Network Models With Heterogeneous Agents," NBER Working Paper. [0016]

ANTRÀS, POL, TERESA C. FORT, AGUSTÍN GUTIÉRREZ, AND FELIX TINTELNOT (2024): "Trade Policy and Global Sourcing: An Efficiency Rationale for Tariff Escalation," *Journal of Political Economy Macroeconomics*, 2, 1–44. [0018]

ARKOLAKIS, COSTAS, FEDERICO HUNEEUS, AND YUHEI MIYAUCHI (2023): "Spatial Production Networks," Working Paper. [0015]

CALIENDO, LORENZO, AND FERNANDO PARRO (2014): "Estimates of the Trade and Welfare Effects of NAFTA," *Review of Economic Studiesic*, 82, 1–44. [0025]

DEKLE, ROBERT, JONATHAN EATON, AND SAMUEL KORTUM (2008): "Global Rebalancing With Gravity: Measuring the Burden of Adjustment," *IMF Staff Papers*, 55, 511–540. [0023]

KRUGMAN, PAUL, AND ANTHONY J. VENABLES (1995): "Globalization and the Inequality of Nations," *The Quarterly Journal of Economics*, 110, 857–880. [0018]

Co-editor Dave Donaldson handled this manuscript.

Manuscript received 26 April, 2021; final version accepted 20 August, 2024; available online 10 September, 2024.