ENDOGENOUS SPATIAL PRODUCTION NETWORKS: QUANTITATIVE IMPLICATIONS FOR TRADE & PRODUCTIVITY

By

Piyush Panigrahi

November 2021

COWLES FOUNDATION DISCUSSION PAPER NO. 2314

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

ENDOGENOUS SPATIAL PRODUCTION NETWORKS: QUANTITATIVE IMPLICATIONS FOR TRADE AND PRODUCTIVITY

PIYUSH PANIGRAHI
YALE UNIVERSITY

Abstract. Larger Indian firms selling inputs to other firms tend to have more customers, tend to be used more intensively by their customers, and tend to have larger customers. Motivated by these regularities, I propose a novel empirical model of trade featuring endogenous formation of input-output linkages between spatially distant firms. The empirical model consists of (a) a theoretical framework that accommodates first order features of firm-to-firm network data, (b) a maximum likelihood framework for structural estimation that is uninhibited by the scale of data, and (c) a procedure for counterfactual analysis that speaks to the effects of micro- and macro-shocks to the spatial network economy. In the model, firms with low production costs end up larger because they find more customers, are used more intensively by their customers and in turn their customers lower production costs and end up larger themselves. The model is estimated using novel micro-data on firm-to-firm sales between Indian firms. The estimated model implies that a 10% decline in inter-state border frictions in India leads to welfare gains ranging between 1% and 8% across districts. Moreover, over half of the variation in changes in firms' sales to other firms can be explained by endogenous changes in the network structure.

Date: October 31, 2021.
Key words and phrases. production networks, international trade, economic geography.
e-mail: piyush.panigrahi@yale.edu. I am immensely grateful to Andrés Rodríguez-Clare, Dave Donaldson, Thibault Fally, and Cécile Gaubert for enthusiastic support and generous guidance for this project. I thank Alan Auerbach, Matilde Bombardini, Lorenzo Caliendo, Ben Faber, Yuriy Gorodnichenko, Bryan Graham, Daniel Haanwinckel, Matthias Hoelzlein, Junyi Hou, Giammario Impullitti, Sam Kortum, Yuhei Miyauchi, Joseph Shapiro, Vitaliia Yaremko, seminar participants at Berkeley, Yale, Johns Hopkins SAIS, Penn State, Toronto, Singapore Management University, ECARES, Bocconi, and conference participants at CEPR-ERWIT, LACEA-LAMES, NAUEA, MVEA, MWIT, ETSG, EARIE, EEA-ESEM, ESAM, SED, CMES, AMES, NASMES, FREIT ITO, AFES, CESifo Area Conference on Global Economy, GEP/CEPR Annual Postgraduate Conference, EMUEA, RES, FIW for valuable comments. I thank the Department of Commercial Taxes in the Government of Gujarat, the Department of Goods and Services Tax in the Government of Maharashtra, the Commercial Taxes Department in the Government of Tamil Nadu, the Commissionerate of Commercial Taxes in the Government of Odisha, the Directorate of Commercial Taxes in the Government of West Bengal and the Ministry of Finance in the Government of India for sharing anonymized data. Financial support from the Clausen Center for International Business and Policy and Institute for South Asia Studies at UC Berkeley, the Weiss Family Fund for Research in Development Economics and, the CEPR and the UK Department for International Development (DFID) (under the jointly funded research initiative 'Private Enterprise Development in Low Income Countries' (PEDL), contract reference PEDL_LOA_7500_Panigrahi) is gratefully acknowledged. The views expressed are not necessarily those of CEPR or DFID. All errors are my own.

1. Introduction

Heterogeneity in production costs across firms is at the heart of modern general equilibrium models of firm heterogeneity and trade. Yet differences in firms' production costs are typically attributed to differences in productivity across firms. With firms operating in production networks, differences in production costs arise not just from differences in productivity but also from finding the most cost-effective suppliers of intermediate inputs. While trade does not directly affect the former, trade in intermediate inputs influences the latter. General equilibrium theories of trade with firms differing only in productivity do not grapple with microscopic heterogeneity in the extensive and intensive margins of firm-to-firm trade in intermediate inputs — who buys from whom and how much? How does endogenous formation of customer-supplier linkages between firms and the resultant network architecture drive differences in firms' overall sales, ability to sell across multiple destinations, and aggregate patterns of trade? How do we evaluate the impact of market integration, technology improvements, and improvements in allocative efficiency on aggregate outcomes when the production network of firms reorganizes in response to these shocks?

In this paper, I present a novel framework to evaluate the aggregate and firm-level consequences of micro- and macro- shocks to the spatial economy and answer these questions in four steps. First, I use novel micro-data to document empirical regularities arising from a new decomposition of firms' sales that underscores the salience of endogenous network formation between firms. Second, I develop a model of trade between multiple locations featuring endogenous formation of firm-to-firm production networks that not only rationalizes micro-data on firm-to-firm sales but is also consistent with structural gravity at the aggregate level. Third, I devise a procedure to structurally estimate the model that circumvents computational difficulties pervasive in estimation of network formation models with large numbers of firms. Fourth, I propose a procedure to evaluate counterfactual outcomes that accounts for randomness in network formation without requiring simulation of large networks which can be computationally burdensome due to interdependence in link formation.

Using data on 103 million firm-to-firm relationships assembled from administrative VAT records spanning across 5 years and pertaining to around 2.5 million Indian firms located across 141 districts, I find that firms with higher sales to other firms (a) tend to have more customers, (b) tend to be used more intensively by their customers and (c) tend to sell to larger customers. The first margin explains 35% of the variation in firms' sales, an additional 46% is explained by the second margin, leaving 19% for the third. On one hand, the third margin suggests that firms' heterogeneity in input sales is partially driven by demand from larger customers downstream in the supply chain. On the other hand, the first and the second margins suggest that firms' choice of suppliers and the intensity with which to use their goods potentially influences the attractiveness of the firm as a supplier to its own potential customers. While the former points to the role of network linkages in driving differences between firms, the latter highlights the role of endogenous formation of firm-to-firm linkages in it, both along the extensive and intensive margins. Theseregularities suggestthat endogenous network formation is pertinent to understanding the origins of firm heterogeneity.

The starting point of the theory is the Ricardian model of trade between multiple locations with geographic barriers and imperfect competition as in Bernard et al. (2003). I depart from their framework by accommodating heterogeneous consumer preferences, heterogeneous technological requirements by firms, and arbitrary production networks between firms.[1] Firms' production processes consist of multiple input requirements. Potential suppliers differ in the suitability of their goods for each of these requirements. Firms randomly encounter potential suppliers and select the most cost-effective suppliers for their production requirements. When selecting their suppliers, firms are more likely to select (and for a larger proportion of their requirements) a potential supplier that is able to sell at a lower price and produces a good that is more suitable for its production requirements.

The ability of a potential supplier to sell at a lower price than another is regulated by (a) its idiosyncratic productivity, (b) the efficiency with which its own suppliers were able to produce thus affording the firm a lower price for intermediate inputs, and (c) proximity to location of use thus having to incur lower geographic costs. Firms with lower production costs thus not only attract more customers but are also used more intensively in their customers' production processes. Since these customers use cheaper inputs, they end up with lower production costs themselves and become cost-effective suppliers to their customers. In the cross-section, firms with low production costs end up larger because they have more customers, are used more intensively by their customers and have larger customers.

Differences in the suitability of potential suppliers' goods for a firm's production requirements feature as match-specific productivities across firm pairs in a manner similar to the discrete choice framework. This leads to a multinomial logit model of

[1]While Caliendo and Parro (2014) allow for sectoral heterogeneity and intersectoral linkages in a Ricardian model of trade, they do not allow for arbitrary production networks between firms and are unable to accommodate the vast heterogeneity in input sourcing patterns at the firm-level observed in data.

supplier choice for each of the firm's production requirements. The estimation equation recognizes that while there is a positive probability of a firm sourcing inputs from every other firm, sourcing inputs for only a discrete number of requirements can give rise to sparsity in firm-to-firm connections. This sparsity can be extreme as is observed in the data where the number of firm-to-firm connections are many orders of magnitude lower than its potential given the number of firms in the economy. Predictions for inter-firm trade then allow estimation of the model utilizing the full volume of micro-data on firm-to-firm transactions via the method of maximum likelihood. Semi-parametric estimation of the model implies that firms' fixed effects serve as sufficient statistics for their implied marginal costs and bilateral inter-district fixed effects as a structural gravity specification for estimating trade frictions. Such estimation programs typically entail a high-dimensional non-linear optimization problem that quickly becomes cumbersome with large numbers of fixed effects. On the contrary, I show that these fixed effects can be computed in closed-form thus avoiding the problem altogether.

For counterfactual analysis, I propose a procedure that departs from the exact hat algebra approach commonly used in trade models (see Dekle et al. (2008) and Costinot and Rodríguez-Clare (2014)). In aggregate models of trade, the exact hat algebra approach evaluates the change in aggregate outcomes in response to shocks. In those models, aggregate data coincides with the expected value of aggregate outcomes in the initial state. In contrast, my model accommodates granularity and acknowledges that the observed data corresponds to only one of many possible realizations under the initial state. The data generating process implied by the model is therefore non-degenerate and hat algebra cannot be used as is. To evaluate counterfactual outcomes, I therefore use the model to obtain the expected value of the data generating process in the initial state and use hat algebra to evaluate the expected value in the counterfactual state. The model and the procedure are rich enough to not only speak about aggregate effects of aggregate shocks but also firm-level effects of aggregate shocks and aggregate and firm-level effects of micro-shocks.

Using the estimated model, I conduct three counterfactual experiments. First, I evaluate the consequences of reducing inter-state border frictions in the context of the recent Goods and Services Tax reform in India that aimed to mitigate such barriers to trade. I find that a 10% decline in border frictions leads to sizable welfare gains across districts ranging between 1% and 8%. Moreover, over half of the variation in changes in firms' sales to other firms can be explained by endogenous changes in the network structure.

Second, I examine firm-level implications of a uniform decline in trade frictions. Firms' sales to other firms adjust along two margins. On one hand, firms become either more or less successful in attracting more customers or increasing market share among existing ones. On the other hand, change in customers (or market shares within them) affects the average customer size for these firms. For firms in the top five percentiles, the negative adjustment along the second margin dominates positive adjustment along the first — their sales to other firms shrink. For the rest of the firms, sales to other firms expands. For firms in the next five percentiles, the positive adjustment along the first margin dominates the negative adjustment along the second. The remaining firms above the third quartile face positive adjustment along both margins. For firms below the third quartile, the positive adjustment along the second margin dominates the negative adjustment along the first.

Third, policy reforms can sometimes manifest as heterogeneous microeconomic shocks across firms. To illustrate the effects of micro-shocks on aggregate outcomes through the lens of the model, I evaluate the consequences of neutralizing firm-level distortions when they correlate positively versus negatively with size. I find that in either case endogenous changes in the network structure explain a dominant majority of changes in firms' sales to other firms. At the aggregate level, neutralizing positively size-dependent distortions has positive terms of trade effects for a majority of districts whereas neutralizing negatively size-dependent distortions has negative terms of trade effects for a majority of districts.

Related Literature. This paper contributes to four strands of literature. First, this paper is related to the nascent literature on endogenous production networks in general equilibrium which can be broadly classified into two categories. The first (Eaton et al. (2016); Oberfield (2018); Acemoglu and Azar (2020); Boehm and Oberfield (2020)) models formation of linkages as the outcome of selection from a discrete menu of choices whereas the second (Lim (2017); Taschereau-Dumouchel (2017); Huneeus (2018)) models formation of linkages between firms as the outcome of "love of variety" in input sourcing while being subject to relationship costs.[2] This paper is more closely related to the former to take advantage of extreme value functional forms that allow tractable empirical characterization for estimation. While this paper shares the mechanism for supplier selection with Oberfield (2018) and Boehm and Oberfield (2020) and formulation of technology and preferences with Eaton et al. (2016), none of these papers

[2]Other complementary approaches to endogenous production network formation include Carvalho and Voigtlander (2014), Chaney (2014), Tintelnot et al. (2018) and Antràs and de Gortari (2020).

provide a closed-form characterization of both the extensive and intensive margins of inter-firm trade that the model here delivers. While Oberfield (2018) and Boehm and Oberfield (2020) do not consider trade between locations, the model in Eaton et al. (2016) features trade. In contrast to Eaton et al. (2016), where there are no differences in suitability of goods across firms' requirements and estimation of the model requires use of simulation-based methods, the model here uniquely recognizes the fact that firms' input sourcing decisions comprise finding the supplier that not only offers the lowest price but is also the most suitable for production requirements. The extreme value formulation of this feature delivers a multinomial logit model of supplier choice and allows the model to be estimated directly using the full volume of data on firm-to-firm sales via maximum likelihood.

Second, this paper is related to a long literature on firm heterogeneity (for example, Jovanovic (1982); Hopenhayn (1992); Axtell (2001); Melitz (2003); Klette and Kortum (2004); Luttmer (2007); Arkolakis (2016)) and in particular the branch that studies the heterogeneity among firms arising from their engagement in input-output linkages — Oberfield (2018) and Bernard et al. (2019). The model here houses two sources of firm heterogeneity — from idiosyncratic productivities as in Kortum (1997) and from match-specific productivities and engagement in input-output linkages as in Oberfield (2018). Unlike Oberfield (2018), the model accommodates heterogeneity in the number of input suppliers across firms as well as in the intensity of use of suppliers across their customers. The model thus allows for variation in firms' average intensity of use by their customers. In the data, this margin explains 46% of the variation in firms' sales. The modeling approach here is distinct from Bernard et al. (2019) who use a fixed cost formulation that necessitates use of simulation-based estimation methods.

Third, the paper also relates to a growing literature on propagation of shocks and aggregation in distorted production networks including Jones (2011), Acemoglu et al. (2012), Swiecki (2017), Caliendo et al. (2017), Liu (2019), Baqaee and Farhi (2019a,b, 2020), and Bigio and LaO (2020). Some of these papers allow for non-Cobb–Douglas technologies and thus endogenize the intensity with which different inputs are used. However, they do not investigate which combinations of inputs will be used—that is, the extensive margin of firm-to-firm trade —which features prominently in this paper.

Finally, this paper is related to a rich literature in international trade. In the model, trade is driven by comparative advantage as in Ricardian trade models (Eaton and

Kortum (2002); Bernard et al. (2003)). However, since the model accommodates heterogeneity in consumer preferences and technological requirements across firms, comparative advantage is determined by each consumer and firm demanding inputs rather than at the level of each market. This allows the model to rationalize patterns of firm participation in international trade within the Ricardian framework which are typically relegated to new trade theory models such as Melitz (2003) and Eaton et al. (2011).[3] This paper is also related to the branch of the trade literature that develops firm-level models of importing that accommodate heterogeneity in input sourcing behavior between firms (for example, Antràs et al. (2017); Blaum et al. (2018)). While these papers consider models where firms choose the set of locations to source intermediate inputs or the share of intermediate inputs that are imported, here I develop a more disaggregated model where firms choose both the set of suppliers across multiple locations for intermediate inputs and the share purchased from each of them. The model also shares features with papers that emphasize the role of granularity in trade models such as Eaton et al. (2013), Armenter and Koren (2014), and Gaubert and Itskhoki (2021). The approach to counterfactual analysis parallels contemporaneous work by Dingel and Tintelnot (2020) who take a related approach in commuting choice models that feature granularity. [4]

Outline. Section 2 describes the data and the corresponding empirical regularities. Section 3 describes the model and lays out the probabilistic assumptions under which model predictions on inter-firm trade shares are derived. Section 4 begins with the estimation framework for firms' marginal costs, trade frictions and dispersion of firms' raw efficiencies. It then provides the procedure for conducting counterfactual analysis. Section 5 examines model implications for counterfactual scenarios, one that leads to improvements in allocative efficiency and another that causes market integration. Section 6 concludes. All proofs are relegated to the Online Appendix.

Notation. Throughout the paper, a firm is indexed by s when it is a *seller* of intermediate inputs or goods for final consumption and by b when it is a *buyer* of intermediate

[3]For example, Eaton et al. (2011) state that the Ricardian framework with a fixed range of commodities used in Bernard et al. (2003) does not deliver the feature that a larger market attracts more firms as observed in French data.

[4]In their case, non-degeneracy of counterfactual outcomes arises from a finite number of individuals making residential and workplace decisions. In this paper, non-degeneracy of counterfactual outcomes arises from interdependent decisions on input sourcing made by a finite number of firms. A similar problem of indeterminacy of the trade equilibrium in relative wages across locations arises in both cases. While they introduce the notion of continuum-case rational expectations to resolve this issue, I show that relative wages are deterministic under a large network approximation despite granularity at firm-level.

inputs. Households are indexed by i. A location is indexed by o when it is the *origin* of a trade flow and typically where firm s is located. Similarly, it is indexed by d when it is the *destination* of a trade flow and typically where firm b is located or household i resides. The set of all locations is denoted by $\mathcal{J}$. The set of all firms is denoted by $\mathcal{M}$ and the subset located at o is denoted by $\mathcal{M}_o$. The set of all households is denoted by $\mathcal{L}$ and the subset located at d is denoted by $\mathcal{L}_d$. The number of elements in these sets are denoted as $M = |\mathcal{M}|$, $L = |\mathcal{L}|$, $M_o = |\mathcal{M}_o|$, and $L_d = |\mathcal{L}_d|$.

2. Data & Empirical Regularities

2.1. **Sources of Data.** The primary dataset for this paper consists of the universe of *intra-state* firm-to-firm transactions assembled from commercial tax authorities of five Indian states (viz. Gujarat, Maharashtra, Tamil Nadu, Odisha, and West Bengal) between 2011-12 and 2015-16. Put together, these states had a nominal GDP of $738 billion in 2015-16, accounting for nearly 40% of national GDP. Among these states, the largest (Maharashtra) accounts for roughly 14% of national GDP while the smallest (Odisha) accounts for a little over 2%. I complement this dataset with data on the universe of *inter-state* firm-to-firm transactions obtained from the Ministry of Finance in the Government of India running for the same period. It includes transactions between all firms registered under the value-added tax system in their respective state. The combined dataset consists of transactions between goods-producing firms and does not include the services sector. It records 103 million inter-firm relationships between approximately 2.5 million firms across the years. Firms are located across 141 districts in these 5 states. [5]

2.2. **Network Margins of Firm Heterogeneity & Trade.** Indian firms are vastly heterogeneous in size, a pervasive finding in studies of firm-level data. Intuitively, firms' outcomes are shaped not only by their own intrinsic characteristics, like productivity, but also by the characteristics of the firms – suppliers and customers – that they connect with. In this paper, I am concerned with firm heterogeneity arising from their behavior in production networks along two margins — the upstream and the downstream margins. On one hand, firm behavior on the upstream margin — its decision of supplier choice on the extensive and intensive margins affects not only its own marginal cost but potentially that of customers that purchase goods from it. On the other hand, firm behavior on the downstream margin – its decision of quantity to produce and sell to customers affects its suppliers through demand for inputs from them. While the

[5] See Appendix D.1 for summary statistics.

downstream margin is operational in models with exogenous production networks, the upstream margin requires a model of endogenous network formation between firms — one where firms choose their suppliers and the intensity with which they use inputs from those suppliers.

To shed light on the economic importance of these margins and guide the main features of the model I will develop in Section 3, I leverage the rich network structure of the dataset to conduct a simple decomposition of firms' sales to other firms into three margins: number of customers, average intensity of use among those customers, and average customer size. Formally, input sales of firm s located at o can be decomposed into these three factors according to the following identity.

(2.1)

$$\text{input sales}_o(s) = N_o(s) \times \frac{\sum_d \sum_{b \in \mathcal{M}_d} \pi_{od}(s,b)}{N_o(s)} \times \frac{\sum_d \sum_{b \in \mathcal{M}_d} \pi_{od}(s,b) \times \text{input costs}_d(b)}{\sum_d \sum_{b \in \mathcal{M}_d} \pi_{od}(s,b)}.$$

In this expression, $N_o(s)$ is the number of customers and $\pi_{od}(s,b)$ is the intensity with which firm b located at d uses goods from seller s. Specifically, it is calculated as:

$$\pi_{od}(s,b) = \frac{\text{sales}_{od}(s,b)}{\text{input costs}_d(b)},$$

where $\text{sales}_{od}(s,b)$ denotes the value of goods sold by firm s to firm b and $\text{input costs}_d(b) = \sum_o \sum_{s \in \mathcal{M}_o} \text{sales}_{od}(s,b)$. Through variation in number of customers, the first factor captures the attractiveness of the firm to potential customers looking for input suppliers. Similarly through variation in intensity of use by customers, the second factor captures the attractiveness of the firm on the intensive margin of input choice by its customers. The third factor measures average size of customers as inferred from a weighted average of their input costs. The first two factors constitute the upstream margin and capture the direct importance of the firm in the production network since it captures how cost-effective the firm is irrespective of the characteristics of the customers it sells to. The third factor constitutes the downstream margin and captures the indirect importance of the firm in the production network through the importance of its customers, its customers' customers and so on. In addition, the upstream margin of firm's sales also captures the overall *intensity of use* of the firm — the sum of cost shares of all firms in the economy that can be attributed purchases from it.[6]

[6]The upstream margin is sometimes referred to as the firm's weighted out-degree. In recent work, Acemoglu et al. (2012) coin this term for similar statistics at the industry level.

I compute the share of variance of firms' sales that is explained by each of these factors.[7] Column (1) in Table 2.1 reports the results of the decomposition. Four-fifths of the variance in firms' sales can be attributed to the upstream margin leaving the rest for the downstream margin. It implies that larger firms are likely to have more customers (explains 35% of the variance), be used more intensively by those customers (46%), and have larger customers (19%). All three factors covary positively with sales and contribute a non-trivial share to the variance. The positive covariance of the downstream margin can be rationalized as follows. Firms with higher demand for their own goods produce larger quantities and to do so they purchase higher quantities of inputs from their suppliers. In turn, their suppliers end up with higher demand and they source larger quantities from their own suppliers and so on. Therefore, in the cross-section one observes that larger firms have larger customers on average. This points to the importance of supply chain linkages between firms even when the network structure is exogenously fixed.

However, it is the outsized contribution of the upstream margin that highlights the importance of endogenous network formation through two potential channels. First, when firms choose to source from more cost-effective suppliers, they are likely to inherit lower marginal costs from their suppliers. This makes them attractive to their own customers who become larger in turn. Therefore, in the cross-section one would observe a positive correlation between firms' sales and number of customers. This suggests that the endogeneity of production networks along the extensive margin of inter-firm trade is important. Second, when suppliers' goods are substitutable in a firms' input demand system, more cost-effective firms will account for a larger share of material costs of their customers. Since those customers source cheaper inputs intensively, they are likely to inherit lower marginal costs from their suppliers. This makes them attractive to their own customers and they become larger themselves. Therefore, in the cross-section one would observe a positive correlation between firms' sales and average intensity of use by customers. This suggests that the endogeneity of production networks along the intensive margin of inter-firm trade is important.

[7]In short, if a variable X can be decomposed into R factors, $\{X_r\}_{r=1}^{R}$ such that $X = X_1 \cdot X_2 \cdots X_R$, then the share of variance of X that can be attributed to any factor X_r is $\frac{Cov(\ln X, \ln X_r)}{Var[\ln X]}$. While these shares sum to unity by additivity of the covariance operator, they are not constrained to be positive individually. For example, see Klenow and Rodríguez-Clare (1997) for use in growth accounting and Eaton et al. (2011) for regression-based decomposition of margins of trade.

TABLE 2.1. **Network Margins of Firm Heterogeneity & Trade**

	(1)	(2)	(3)
# Customers	35%	37%	67%
Intensity per Customer	46%	57%	20%
Average Customer Size	19%	6%	13%
Fixed Effects:			
Seller×Year	—	✓	—
Origin×Year	—	—	✓
Data Level:			
Seller×Year	•	—	—
Seller×Destination×Year	—	•	—
Origin×Destination×Year	—	—	•
# observations	5.6×10^6	18.2×10^6	58,390

Note. Column (1) reports the contribution of factors: # customers, intensity per customer, and average customer size, to the variance of firms' sales (as per equation (2.1)). Column (2) reports the contribution of those factors to the variance of firms' destination-specific sales (as per equation (A.1)). Column (3) reports the same for trade flows between districts (as per equation (A.2)). See Appendix A for details and alternative specifications.

Furthermore, trade across space is costly and economic activity across space exhibits large dispersion. How does the relative position of firms across space affect their outcomes? How does geography affect the aforementioned margins of firm heterogeneity? To investigate this, I construct a similar decomposition at a more disaggregated level for firms' destination-specific sales and at a more aggregated level for trade flows between districts.[8] Column (2) in Table 2.1 reports results of variance decomposition of firm's destination-specific sales while controlling for firm-level fixed effects. This is done to capture the variation in individual firms' sales across multiple destinations. The upstream margin accounts for 94% of the variation leaving 6% for the downstream margin. Column (3) in Table 2.1 reports results of variance decomposition of aggregate trade flows between districts while controlling for origin fixed effects. The upstream margin accounts for 87% of the variation leaving 13% for the downstream margin. Since the upstream margin explains the lion's share of the variation in both cases, these results underscore the salience of geography in endogenous network formation between firms.

Taking stock, I find that firms that are larger also tend to have more customers, tend to be used more intensively by their customers, and tend to have larger customers.

[8]Further details are provided in Appendix A.

Of course, these decompositions capture equilibrium relationships and are not causal; nevertheless, they make clear that understanding the characteristics of firms' network is key to understanding origins of firm heterogeneity. While the economic intuition behind these results is straightforward, the decomposition results are, to the best of my knowledge, new to the literature.[9] With this in mind, I develop a model of endogenous production network formation in the next section that expressly takes these findings into account and leads to a multinomial logit model of supplier choice for estimation.

3. Theoretical Framework

In this section, I first describe a model of trade between multiple locations that accommodates heterogeneity in consumer preferences, heterogeneity in technological requirements of firms and arbitrary production networks. Subsequently, I employ functional form assumptions that make the model tractable and allow to derive aggregate implications before proceeding to describe the framework for estimation in the next section.

3.1. **Economic Environment.** The model economy $\mathcal{E} \equiv \{\mathcal{M},\mathcal{L},\mathcal{J}\}$ consists of many firms and households at many locations. Firms produce using local labor and intermediate inputs sourced from suppliers potentially spread across multiple locations. Each household supplies one unit of labor inelastically to local firms. Firms rebate any profits to local households. Trade between locations is subject to iceberg trade costs denoted by $\tau_{od} \geq 1$. That is, a firm producing at o needs to ship τ_{od} units of a good for one unit of good to arrive at d.

3.1.1. *Technology and Market Structure.* Firms' production processes involve combining labor and accomplishing a set of tasks. To accomplish tasks, firms source intermediate inputs from other firms. In particular, the production function for any firm b at location d is defined over labor and a discrete number of tasks (indexed by $k \in \mathcal{K}_d(b) \equiv \{1,\cdots,K_d(b)\}$) as:

$$y_d(b) = z_d(b)\left(\frac{l_d(b)}{1-\alpha_d}\right)^{1-\alpha_d}\left(\frac{\prod_{k\in\mathcal{K}_d(b)} m_d(b,k)^{1/K_d(b)}}{\alpha_d}\right)^{\alpha_d},$$

[9]In related work, Huneeus (2018) and Bernard et al. (2019) use Chilean and Belgian production network micro-data respectively to decompose firms' sales to other firms into # customers and sales per customer. At the aggregate level using trade flows, it is also related to the decomposition into extensive and intensive margins of trade such as in Eaton et al. (2011, 2016) and Fernandes et al. (2018). Here, I show that sales per customer in the former and the intensive margin in the latter can be further decomposed into two factors such that the decomposition delineates the role of endogenous network formation.

$$m_d(b,k) = \sum_{s \in \mathcal{S}_d(b)} m_{od}(s,b,k),$$

where $l_d(b)$ is the amount of labor input used by firm b, $m_d(b,k)$ is the quantity of materials utilized to accomplish task k, $z_d(b)$ is the idiosyncratic Hicks-neutral productivity with which firm b produces, and $K_d(b)$ is the number of tasks in firm b's production function.

Materials sourcing is subject to search frictions. Among all the firms in the economy, firm b encounters only a few and can source intermediate inputs to accomplish tasks only from those firms. This restricted set of potential suppliers is denoted by $\mathcal{S}_d(b)$. While outputs of potential suppliers are perfectly substitutable for accomplishing any task, they differ in their suitability for the task in question, captured by their respective match-specific productivities. For each of its tasks, firm b selects the supplier that offers the lowest effective price. Importantly, firm b may choose the same supplier for more than one tasks.

The market structure for intermediate inputs and final consumption is characterized by Bertrand competition. Firms face limit pricing behavior when sourcing intermediate inputs and engage in limit pricing themselves when supplying their goods for intermediate input use by other firms and for final consumption by households. This means that the lowest cost supplier for a firm or household sets a limit price to just undercut the next lowest cost supplier available to the firm for intermediate input use or to the household for final consumption.

I now turn to firms' cost minimization problem. For firm b, selecting the cost-minimizing input bundle consists of first selecting the most cost-effective supplier for each task among the set of potential suppliers, then choosing the quantity of inputs to purchase from those selected suppliers for each of the tasks and the amount of labor to hire. In other words, firm b first chooses who to source inputs from and then how much to buy from each of them.

For any particular task k in firm b's production function, the cost-effectiveness of a supplier s from location o in $\mathcal{S}_d(b)$ depends on four factors: (a) the marginal cost of s, denoted $c_o(s)$; (b) the trade cost faced by s of shipping goods to d, τ_{od}; (c) the match-specific productivity when b utilizes the output of s to accomplish the task, denoted by $a_{od}(s,b,k)$, and (d) the markup charged by s when it sells its output to b for accomplishing the task, denoted $\bar{m}_{od}(s,b,k)$. For task k, firm b chooses the supplier that offers the cheapest price, that is,

$$s_d^*(b,k) = \arg\min_{s\in\mathcal{S}_d(b)} \left\{ \frac{\bar{m}_{od}(s,b,k)c_o(s)\tau_{od}}{a_{od}(s,b,k)} \right\}. \tag{3.1}$$

Firms spend equal shares of costs across tasks. Although the elasticity of substitution between tasks is equal to unity, this formulation captures richer patterns of substitution across outputs of other firms that are used to accomplish tasks. This is because a potential supplier charging a lower price is likely to be selected for a higher number of tasks by any firm and hence is likely to account for a higher cost share of the firm. The extensive margin of firms' input sourcing is determined by whether a potential supplier is chosen for at least one of the tasks whereas the intensive margin is determined by how many tasks the potential supplier gets selected for. Both these margins of inter-firm trade are determined endogenously in equilibrium. [10]

With limit pricing, the markup is determined by how much lower the effective cost faced by the best supplier is relative to the second best. Hence, the effective price faced by b for task k, denoted by $p_d(b,k)$, is given by

$$p_d(b,k) = \min_{s\in\mathcal{S}_d(b)\setminus\left\{s_d^*(b,k)\right\}} \left\{ \frac{c_o(s)\tau_{od}}{a_{od}(s,b,k)} \right\}. \tag{3.2}$$

Now, taking wage w_d and effective prices $\{p_d(b,k) : k \in \mathcal{K}_d(b)\}$ as given, the firm's unit cost function can be defined as:

$$c_d(b) = \min_{\{l_d(b),\{m_d(b,k):k\in\mathcal{K}_d(b)\}\}} w_d l_d(b) + \sum_{k\in\mathcal{K}_d(b)} p_d(b,k)m_d(b,k) \tag{3.3}$$

$$\text{subject to } z_d(b)\left(\frac{l_d(b)}{1-\alpha_d}\right)^{1-\alpha_d}\left(\frac{\prod_{k\in\mathcal{K}_d(b)} m_d(b,k)^{1/K_d(b)}}{\alpha_d}\right)^{\alpha_d} = 1$$

[10]In Eaton et al. (2016), the sets of potential suppliers for each task are exclusive (i.e., there are no suppliers in common across these sets). As a result, a supplier is selected for at most one task by the buyer and the number of tasks for which a supplier is selected is the same as the number of buyers. In contrast, here, the set of potential suppliers is common across all tasks. As a result, a supplier can be selected for more than one task by the buyer and the number of tasks for which a supplier is selected can potentially be higher than the number of buyers. While in Eaton et al. (2016), the number of suppliers is exogenously fixed (conditional in labor share) by the number of tasks, here the number of suppliers is determined endogenously.

With the cost function as defined above, the profit of a firm s located at o can be expressed as

$$\Pi_o(s) = \sum_d \sum_{b \in \mathcal{M}_d} \sum_{k \in \mathcal{K}_d(b)} (\bar{m}_{od}(s,b,k) - 1) \frac{c_o(s)\tau_{od}}{a_{od}(s,b,k)} m_{od}(s,b,k)$$
$$+ \sum_d \sum_{i \in \mathcal{L}_d} \sum_{n \in \mathcal{N}_d(i)} (\bar{m}_{od}(s,i,n) - 1) \frac{c_o(s)\tau_{od}}{a_{od}(s,i,n)} q_{od}(s,i,n),$$

where $m_{od}(s,b,k)$ denotes the quantity of goods sold by firm s to customer b for task k and $q_{od}(s,i,n)$ denotes the quantity of goods sold by firm s to households i for need n (described below). The quantity of goods sold $m_{od}(s,b,k)$ or $q_{od}(s,i,n)$ is positive if s is the most effective supplier for task k or need n respectively and zero otherwise.

3.1.2. *Household Preferences.* Households consume goods produced by firms to fulfill a set of needs. In particular, the utility function for any household i at location d is defined over a discrete number of needs (indexed by $n \in \mathcal{N}_d(i) \equiv \{1, \cdots, N_d(i)\}$) as:

$$u_d(i) = \prod_{n \in \mathcal{N}_d(i)} q_d(i,n)^{1/N_d(i)},$$
$$q_d(i,n) = \sum_{s \in \mathcal{S}_d(i)} q_{od}(s,i,n),$$

where $q_d(i,n)$ is the quantity of goods consumed to fulfill need n and $N_d(i)$ is the number of needs in the utility function.

Goods sourcing is subject to search frictions and is modeled similar to firms sourcing inputs. Outputs of potential suppliers are perfectly substitutable for fulfilling any need but differ in match-specific taste shocks. For each of its needs, household i selects the supplier that offers the lowest effective price and can sometimes select the same supplier for more than one needs. For household i, selecting the utility-maximizing consumption bundle comprises of first selecting the most cost-effective supplier for each need among the set of potential suppliers and then of choosing the quantity of goods to purchase from those selected suppliers for each of the needs. For any particular need n in i's utility function, the cost-effectiveness of a supplier s from location o in $\mathcal{S}_d(i)$ depends on four factors similar to those that affect firms sourcing inputs. In particular, for need n, household i chooses the supplier that offers the cheapest price, that is,

$$s_d^*(i,n) = \arg\min_{s \in \mathcal{S}_d(i)} \left\{ \frac{\bar{m}_{od}(s,i,n) c_o(s) \tau_{od}}{a_{od}(s,i,n)} \right\} \tag{3.4}$$

The markup is again determined by how much lower the effective cost faced by the best supplier is relative to the second best. The effective price faced by i for need n denoted by $p_d(i,n)$ is then given by

$$p_d(i,n) = \min_{s \in \mathcal{S}_d(i) \setminus \{s_d^*(i,n)\}} \left\{ \frac{c_o(s)\tau_{od}}{a_{od}(s,i,n)} \right\}. \tag{3.5}$$

Now, taking $\{p_d(i,n) : n \in \mathcal{N}_d(i)\}$ as given, the household's indirect utility function can be defined as:

$$V_d(i) = \max_{\{q_d(i,n): n \in \mathcal{N}_d(i)\}} \prod_{n \in \mathcal{N}_d(i)} q_d(i,n)^{1/N_d(i)} \tag{3.6}$$
$$\text{subject to} \sum_{n \in \mathcal{N}_d(i)} p_d(i,n) q_d(i,n) = w_d + \Pi_d$$

where $\Pi_d = \frac{\sum_{s \in \mathcal{M}_d} \Pi_d(b)}{L_d}$ is the per capita profit rebated to households residing at o.

3.1.3. *Equilibrium Definition and Characterization.* Let $\sigma \equiv \{\boldsymbol{z}, \boldsymbol{K}, \boldsymbol{N}, \boldsymbol{\tau}, \boldsymbol{\mathcal{S}}, \boldsymbol{a}\}$ denote the aggregate state of the economy. Here $\boldsymbol{z}$ denotes the vector of idiosyncratic productivities of firms, $\boldsymbol{K}$ denotes the numbers of tasks of all firms, $\boldsymbol{N}$ denotes the numbers of needs of all households, $\boldsymbol{\tau}$ denotes the vector of trade costs across all pairs of locations, $\boldsymbol{\mathcal{S}}$ denotes the sets of potential suppliers of all firms and households, and $\boldsymbol{a}$ denotes the vector of all match-specific productivities and match-specific taste shocks. All of these objects are exogenous.

An allocation in this economy is represented as $\xi \equiv \{\boldsymbol{l}(\sigma), \boldsymbol{m}(\sigma), \boldsymbol{q}(\sigma), \boldsymbol{y}(\sigma)\}$ and is defined as a set of functions,

$$\begin{aligned}
\boldsymbol{l}(\sigma) &\equiv \{l_d(b;\sigma) : b \in \mathcal{M}_d, d \in \mathcal{J}\}, \\
\boldsymbol{m}(\sigma) &\equiv \{m_{od}(s,b,k;\sigma) : k \in \mathcal{K}_d(b), (s,b) \in \mathcal{M}_o \times \mathcal{M}_d, (o,d) \in \mathcal{J} \times \mathcal{J}\}, \\
\boldsymbol{q}(\sigma) &\equiv \{q_{od}(s,i,n;\sigma) : n \in \mathcal{N}_d(i), (s,i) \in \mathcal{M}_o \times \mathcal{L}_d, (o,d) \in \mathcal{J} \times \mathcal{J}\}, \\
\boldsymbol{y}(\sigma) &\equiv \{y_o(s;\sigma) : s \in \mathcal{M}_s, o \in \mathcal{J}\},
\end{aligned}$$

that map the realization of the state to intermediate input and labor quantities, quantities consumed and quantities produced. A price system is represented as $\varrho \equiv \{\boldsymbol{c}(\sigma), \boldsymbol{p}(\sigma), \boldsymbol{w}(\sigma)\}$ and is defined as a set of functions,

$$\boldsymbol{c}(\sigma) \equiv \{c_o(s;\sigma) : s \in \mathcal{M}_o, o \in \mathcal{J}\},$$

$$\boldsymbol{p}(\sigma) \equiv \{p_d(i,n;\sigma) : n \in \mathcal{N}_d(i), i \in \mathcal{L}_d, d \in \mathcal{J}\} \cup \{p_d(b,k;\sigma) : k \in \mathcal{K}_d(b), b \in \mathcal{M}_d, d \in \mathcal{J}\},$$
$$\boldsymbol{w}(\sigma) \equiv \{w_d(\sigma) : d \in \mathcal{J}\},$$

that map the realization of the state to tasks' prices for firms, needs' prices for households, wage at each location and marginal costs of firms. This leads to the definition of equilibrium in this economy as follows.

Definition 1. For any given state σ, an equilibrium in this economy is defined as an allocation and price system, (ξ, ϱ) such that (a) households select suppliers for needs and firms select suppliers for tasks according to equations (3.1) and (3.4) respectively; (b) firms set prices for other firms and households according to equations (3.2) and (3.5) respectively; (c) households maximize utility according to equation (3.6); (d) firms minimize costs according to equation (3.3); and (e) market clears for each firm's goods and for labor at each location,

$$\sum_{d \in \mathcal{J}} \tau_{od} \left(\sum_{b \in \mathcal{M}_d} \sum_{k \in \mathcal{K}_d(b)} m_{od}(s,b,k) + \sum_{i \in \mathcal{L}_d} \sum_{n \in \mathcal{N}_d(i)} q_{od}(s,i,n) \right) = y_o(s),$$
$$\sum_{b \in \mathcal{M}_d} l_d(b) = L_d.$$

This completes description of the economic environment in the model. Moving ahead, the aggregate state can be divided into two parts. The first comprises of firms' productivities, firms' numbers of tasks, households' numbers of needs, and trade costs; this is denoted by $\sigma_0 \equiv \{\boldsymbol{z}, \boldsymbol{K}, \boldsymbol{N}, \boldsymbol{\tau}\}$. The second part comprises of sets of potential suppliers for firms and households and match-specific productivities and taste shocks; this is denoted by $\sigma_1 \equiv \{\boldsymbol{\mathcal{S}}, \boldsymbol{a}\}$. While σ_0 narrows down the set of networks that could be realized as an outcome of the network formation process, σ_1 pinpoints the exact network of firms that is realized. In the following subsections, I specify a probabilistic model so as to characterize the aggregate trade equilibrium between locations for any given σ_0.

3.2. **Probabilistic Model.** The probabilistic model is specified in three parts. First, I state distributional assumptions on firms' productivities, firms' numbers of tasks, and households' numbers of needs. This reduces the dimensionality of the firm-level state variables $\boldsymbol{z}$ and $\boldsymbol{K}$ and household-level state variables $\boldsymbol{N}$ so that they are characterized by parameters at the location level. Second, I describe the stochastic assumptions that govern random encounters with potential suppliers and the choice of suppliers thereof. This specifies the distribution of the numbers of potential suppliers available to each firm and each household ($\boldsymbol{\mathcal{S}}$) and that of the match-specific productivities and match-specific taste shocks associated with those suppliers ($\boldsymbol{a}$). Finally, I characterize the large economy limit of the model that enables aggregation and leads to the definition of the aggregate trade equilibrium.

Firms' productivities are drawn independently from Fréchet distributions parametrized such that the mean and dispersion across firms vary by location and are given by the following assumption.

Assumption 1. *Idiosyncratic ex ante productivities* $\{z_o(s):s\in\mathcal{M}_o\}$ *are drawn independently according to the following Fréchet distribution:*

$$\mathbb{P}(z_o(s)\leq z)=e^{-T_o z^{-\theta_o}}\mathbf{1}\{z\geq 0\}.$$

where T_o *and* θ_o *are respectively the scale and shape parameters of the productivity distribution at location* o*.*

For any location o, the average productivity of firms is determined by T_o and dispersion in productivities is determined by θ_o. A higher T_o implies higher average productivity and a higher θ_o implies lower dispersion in productivities. Firms' numbers of tasks and households' numbers of needs are drawn from zero-truncated Poisson distributions such that all firms have at least one task in their production function and all households have at least one need in their utility function.

Assumption 2. *The number of tasks* $\{K_d(b):b\in\mathcal{M}_d\}$ *are drawn independently according to the following zero-truncated Poisson distribution:*

$$\mathbb{P}(K_d(b)=K_d)=\frac{e^{-\kappa_d}\kappa_d^{K_d}}{(1-e^{-\kappa_d})K_d!}.$$

The number of needs $\{N_d(i):i\in\mathcal{L}_d\}$ *are drawn independently according to the following zero-truncated Poisson distribution:*

$$\mathbb{P}(N_d(i)=N_d)=\frac{e^{-\eta_d}\eta_d^{N_d}}{(1-e^{-\eta_d})N_d!}.$$

The distributions of the number of tasks across locations is parametrized such that the intensity κ_d varies by location. A higher κ_d implies that firms at d have a larger number of tasks on average and hence the potential to source inputs from a larger number of suppliers. A similar explanation holds for how households' number of needs depends on η_d.

Next, I turn to stochastic assumptions that govern random encounters with potential suppliers and the choice of suppliers thereof. Search frictions in the model are characterized by firms and households encountering potential suppliers via independent Bernoulli trials. The set of sets of potential suppliers $\boldsymbol{\mathcal{S}}$ is therefore completely determined as the outcome of these Bernoulli trials for meeting each firm. The success probabilities associated with these trials are given by the following assumption.

Assumption 3. *The probability with which firm b encounters firm s is given by*

$$\mathbb{P}(s \in \mathcal{S}_d(b)) = \frac{\lambda}{M},$$

where $\lambda > 0$. *Similarly, the probability with which household i encounters firm s is also given by* $\mathbb{P}(s \in \mathcal{S}_d(i)) = {}^{\lambda}/_{M}$.

These success probabilities are decreasing in the total number of firms in the economy. In economies with sufficiently large number of firms, these search frictions approximate Poisson processes where firms and households encounter potential suppliers with rate λ for their tasks and needs respectively.[11] Match-specific productivities and taste shocks are drawn independently for all potential suppliers for each of the tasks in firms' production functions and needs in households' utility functions from a Pareto distribution.

Assumption 4. *Match-specific productivities and taste shocks* $\boldsymbol{a}$ *are drawn independently according to the following Pareto distribution:*

$$F_a(a) = \left(1 - ({}^{a}/_{a_0})^{-\zeta}\right) \mathbf{1}\{a > a_0\}$$

with $\zeta < \theta_o \forall o \in \mathcal{J}$.

The shape parameter of this distribution ζ regulates the thickness of the right tails of the match-specific productivity and taste shock distributions. The lower ζ is, the higher is the likelihood of particularly high draws of match-specific productivities. With higher likelihood of high draws, the choice of supplier (according to equations 3.1 and 3.4) is less sensitive to marginal cost of the supplier or trade costs. The restriction that $\zeta < \theta_o$ for all locations implies that the likelihood of very high draws of idiosyncratic productivities is less than that of very high match-specific productivities. This ensures that the price index is well-defined in the limiting economy.

To enable the theoretical model to make aggregate predictions, I consider a limiting economy where firms and households are arranged on a continuum. In the limiting economy, the trade equilibrium conditional on σ_0 which is characterized by wages across locations $\{w_d : d \in \mathcal{J}\}$, is deterministic. Thus, there is no aggregate uncertainty at any

[11] For any firm b, the number of potential suppliers follows a binomial distribution, i.e., $\mathbb{P}(|\mathcal{S}_d(b)| = S_d) = \binom{M}{S_d}\left(\frac{\lambda}{M}\right)^{S_d}\left(1 - \frac{\lambda}{M}\right)^{M - S_d}$. For sufficiently large values of M, $|\mathcal{S}_d(b)| \sim \text{Poisson}(\lambda)$.

location in the limiting economy. In particular, I adopt the large economy model due to Al-Najjar (2004) which is characterized by a sequence of finite but increasingly large economies that progressively discretizes the unit continuum. The distribution of firms and households along the sequence is uniform. This allows use of the law of large numbers in the limiting continuum to derive cross-sectional distributions of effective prices and marginal costs for given wages. While effective prices of firms' tasks and marginal costs of firms might individually vary across realizations of σ_1, their cross-sectional distributions at each location are invariant across all such realizations in the limiting economy. The following definition formalizes the notion of the limiting economy in the context of this paper.

Definition 2. Consider a sequence of finite economies $\{\mathcal{E}_t : t \in \mathbb{N}\}$ where $\mathcal{E}_t \equiv \{\mathcal{M}_t, \mathcal{L}_t, \mathcal{J}_t\}$ is such that the t^{th} economy has the form $\mathcal{M}_t = \{m_1, \cdots, m_{M_t}\} \subset [0,1]$, $\mathcal{L}_t = \{\ell_1, \cdots, \ell_{L_t}\} \subset [0,1]$ and $\mathcal{J}_t = \mathcal{J}$. The uniform distribution on $\mathcal{M}_t$ is given by $\mathcal{U}_t^M(\mathcal{M}_t^0) = \frac{M_t^0}{M_t}$ for all $\mathcal{M}_t^0 \subset \mathcal{M}_t$. Similarly, the uniform distribution on $\mathcal{L}_t$ is given by $\mathcal{U}_t^L(\mathcal{L}_t^0) = \frac{L_t^0}{L_t}$ for all $\mathcal{L}_t^0 \subset \mathcal{L}_t$. Then, $\{\mathcal{E}_t : t \in \mathbb{N}\}$ is a discretizing sequence of economies if it satisfies:

(1) $\mathcal{M}_t \subset \mathcal{M}_{t+1}$ and $\mathcal{L}_t \subset \mathcal{L}_{t+1}$ for all t,
(2) $\lim_{t\to\infty} \mathcal{U}_t^M(\mathcal{M}_t \cap [a_l, a_h]) = \mathcal{U}([a_l, a_h])$,
(3) $\lim_{t\to\infty} \mathcal{U}_t^L(\mathcal{L}_t \cap [a_l, a_h]) = \mathcal{U}([a_l, a_h])$,

where $\mathcal{U}(\bullet)$ denotes the uniform distribution with support over $[0,1]$ and $[a_l, a_h] \subset [0,1]$.

Along the sequence $\{\mathcal{E}_t : t \in \mathbb{N}\}$ as the economy becomes more discretized, I make additional assumptions on σ_1 so that the model has a well-defined limit. The probability of meeting potential suppliers increases, i.e., $\lim_{t\to\infty} \lambda_t = \infty$, but at a rate slower than that at which the economy is discretized, i.e., $\lim_{t\to\infty} \frac{\lambda_t}{M_t} = 0$. At the same time, match-specific productivities are drawn from stochastically worse distributions as $\lim_{t\to\infty} a_{0,t} = 0$. While the number of potential suppliers grows arbitrarily large and the match-specific productivity associated with any single supplier is drawn from a stochastically worse distribution, the limit is well behaved because the probability of encountering a supplier with match-specific productivity greater than a does not change in the limiting economy, i.e., $\lim_{t\to\infty} \lambda_t a_{0,t}^{\zeta} = 1$.[12] Furthermore, the economy becomes discretized in a manner such that the proportion of firms and households at every location is non-zero and finite. The following assumption states this formally.

[12]This kind of assumption was shown to have a well-defined limit by Kortum (1997) and put to use for a similar purpose by Oberfield (2018).

Assumption 5. *The discretizing sequence of economies $\{\mathcal{E}_t : t \in \mathbb{N}\}$ satisfies the following conditions:*[13]

(1) $\{\lambda_t, a_{0,t} : t \in \mathbb{N}\}$ is such that $\lambda_t = o(M_t)$ and $\lambda_t a_{0,t}^{\zeta} = \Theta(1)$
(2) $\{M_{d,t}, L_{d,t} : d \in \mathcal{J}, t \in \mathbb{N}\}$ is such that $M_{d,t} = \Theta(M_t)$ and $L_{d,t} = \Theta(L_t)$ for all $d \in \mathcal{J}$

This completes the description of the probabilistic model. Let $\boldsymbol{T} \equiv \{T_d : d \in \mathcal{J}\}$, $\boldsymbol{\theta} \equiv \{\theta_d : d \in\}\mathcal{J}$, $\boldsymbol{\kappa} \equiv \{\kappa_d : d \in \mathcal{J}\}$, $\boldsymbol{\eta} \equiv \{\eta_d : d \in \mathcal{J}\}$, and $\boldsymbol{\alpha} \equiv \{\alpha_d : d \in \mathcal{J}\}$. Through Assumptions 1 and 2, the part of the aggregate state contained in σ_0 in the limiting economy can then be redefined as $\sigma_0 \equiv \{\boldsymbol{T}, \boldsymbol{\theta}, \boldsymbol{\kappa}, \boldsymbol{\eta}, \boldsymbol{\tau}\}$.

3.3. **Aggregate Implications.** I now proceed to characterize equilibrium prices $\varrho \equiv \{\boldsymbol{p}(\sigma), \boldsymbol{c}(\sigma), \boldsymbol{w}(\sigma)\}$ in the limiting economy, i.e., $\lim_{t\to\infty} \mathcal{E}_t$. In the limiting economy, for any given realization of σ_0, wages and cross-sectional distributions of effective prices and marginal costs at all locations are invariant across all realizations of σ_1. Therefore, equilibrium prices in the limiting economy can be expressed as $\varrho \equiv \{\boldsymbol{p}(\sigma_0), \boldsymbol{c}(\sigma_0), \boldsymbol{w}(\sigma_0)\}$. I begin with distributional properties of effective prices and marginal costs. Next, I provide model implications for firm-to-firm trade and trade between locations which lead to the characterization of wages in the trade equilibrium.

3.3.1. *Distributions of Effective Prices and Markups.* With limit pricing, the distribution of effective prices faced by a firm for any of its tasks or that faced by a household for any of its needs is characterized by the distribution of the offer with the second lowest effective cost to the supplier. The following proposition provides the distribution of effective prices in the limiting economy.

Proposition 1. *For any realization of σ_0, the effective prices of materials used by firm b to accomplish any task, $p_d(b,k)$, and that of goods consumed by household i to satisfy need n, $p_d(i,n)$, converge to the following distribution as $t \to \infty$:*

$$F_{p_d}(p) = \left(1 - e^{-A_d p^{\zeta}} - A_d p^{\zeta} e^{-A_d p^{\zeta}}\right) \mathbf{1}\{p > 0\},$$

where $\boldsymbol{A} \equiv \{A_d : d \in \mathcal{J}\}$ is the unique positive solution to the following fixed point problem:[14]

$$A_d = \sum_{o \in \mathcal{J}} \tau_{od}^{-\zeta} \mu_o \Gamma\left(1 - \frac{\zeta}{\theta_o}\right) T_o^{\frac{\zeta}{\theta_o}} w_o^{-\zeta(1-\alpha_o)} \mathbb{E}_{\{K_o\}}\left[\Gamma\left(2 - \frac{\alpha_o}{K_o}\right)^{K_o}\right] A_o^{\alpha_o}, \tag{3.7}$$

[13]For any two functions $f(n)$ and $g(n)$, $f(n) = o\,(g(n)) \implies \lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$ and $f(n) = \Theta(g(n) \implies \text{limsup}_{n\to\infty} \frac{|f(n)|}{g(n)} < \infty$ and $\text{limsup}_{n\to\infty} |\frac{f(n)}{g(n)}| > 0$.
[14]The gamma function $\Gamma(\cdot)$ is defined as $\Gamma(x) = \int_0^{\infty} e^{-x} m^{x-1} dm$.

where μ_o denotes the proportion of firms at o and $\mathbb{E}_{\{K_o\}}[\cdot]$ denotes the expectation over all realizations of numbers of tasks K_o across firms at o.

The distribution of effective prices conditional on σ_0 is obtained by appealing to a law of large numbers afforded by Definition 2. While the effective price faced by individual firms and households varies across realizations of σ_1, the cross-sectional distribution in the limit economy does not. These distributions are parametrized by a scale parameter A_d and a shape parameter ζ. Market access, given by A_d, is a key object of interest because it summarizes the probabilistic access of firms at d to inputs from all locations. The functional form suggests that firms at a location with higher market access face stochastically lower effective prices. Specifically, if $A_d > A_{d'}$, the distribution $F_{p_{d'}}(\cdot)$ first-order stochastically dominates $F_{p_d}(\cdot)$.

Focussing on equation (3.7), market access A_d is a trade friction ($\tau_{od}^{-\zeta}$) weighted sum of the attractiveness of all locations $o \in \mathcal{J}$, i.e., nearer locations receive higher weights because of lower trade costs τ_{od} and vice versa. The attractiveness of a location o for sourcing inputs is determined by four factors: (a) density of firms μ_o; (b) average productivity among firms T_o; (c) its own market access A_o; and (d) wages w_o. Locations with higher density, higher average productivity, higher market access or lower wages are more attractive. In addition, the attractiveness of a location o is more sensitive to its market access A_o and less so to wages w_o if materials share of costs α_o is higher at o and vice versa.

Although the effective price is characterized by the distribution of the offer with the second lowest effective cost to the supplier, it is still the supplier with the lowest effective cost that is selected. The distribution of markups faced by the firm or the household is characterized by that of the ratio of the second lowest to the lowest effective costs incurred by the second best and the best suppliers respectively. In addition, Assumption 5 implies that in the limiting economy, every firm or household encounters at least two potential suppliers with probability approaching one and this ensures that markups are well-behaved.[15] The following proposition provides the distribution of markups.

Proposition 2. *Markups over marginal cost of lowest cost supplier $\bar{m}_{od}(\cdot,\cdot,\cdot)$ are distributed according to the following Pareto distribution:*

$$F_{\bar{m}}(\bar{m}) = \left(1 - \bar{m}^{-\zeta}\right)\mathbf{1}\{\bar{m} > 1\}.$$

The shape parameter of the distribution of potential markups is ζ, the same parameter that governs dispersion in match-specific productivities. With lower ζ, higher

[15]To see this clearly, note that for any firm b, $\mathbb{P}(|\mathcal{S}_d(b)| < 2) = \left(1 - \frac{\lambda_t}{M_t}\right)^{M_t} + \binom{M_t}{1}\left(\frac{\lambda_t}{M_t}\right)\left(1 - \frac{\lambda_t}{M_t}\right)^{M_t - 1}$. It then follows that $\lim_{t\to\infty}\mathbb{P}(|\mathcal{S}_d(b)| < 2) = 0$.

markups are more likely since high match-specific productivities are more likely and hence are larger gaps between costs to the best and second best suppliers. Moreover, the distribution of markups is the same in any destination. An aggregate implication that follows from the distribution of markups is that the share of variable costs in gross output is given by $\frac{1}{1+1/\zeta}$ at all locations. This in turn implies that value-added share of gross output at location o is given by:

$$(3.8)\qquad \left({}^{VA}/_{GO}\right)_o = \frac{1-\alpha_o+1/\zeta}{1+1/\zeta}.$$

3.3.2. *Distributions of Marginal Costs.* The marginal cost of a firm determines (albeit, partially) whether it is selected by potential customers and if so, the intensity with which it is used. It is therefore a key variable governing network formation between firms. The marginal cost of the firm is itself determined by its own productivity, wage faced by it for hiring labor, and the effective price faced by it for its tasks. Since productivity, number of tasks and effective price faced for each task are randomly drawn for each firm, the marginal cost of any given firm is a random variable that is itself the product of a random number (number of tasks) of random variables (effective price for each task). In lieu of the distribution function which does not have a closed-form characterization, I provide closed-form expressions for moments of marginal costs distribution in the following proposition.

Proposition 3. *For any realization of σ_0, the distribution of marginal costs at any location o satisfies the following moment conditions:*

$$(3.9)\qquad \mathbb{E}[\log c_o(s)] = \frac{\alpha_o \psi^{(0)}(2)}{\zeta} + \frac{\psi^{(0)}(1)}{\theta_o} + (1-\alpha_o)\log w_o - \frac{\alpha_o}{\zeta}\log A_o - \frac{1}{\theta_o}\log T_o,$$

$$(3.10)\qquad Var[\log c_o(s)] = \frac{\psi^{(1)}(1)}{\theta_o^2} + \mathbb{E}_{\{K_o\}}[1/K_o]\frac{\alpha_o^2 \psi^{(1)}(2)}{\zeta^2},$$

where $\psi^{(n)}(\cdot)$ denotes polygamma functions.[16]

What factors affect average marginal costs of firms at a location? Intuitively, firms' marginal costs will be low on average if they are more productive, are able to source intermediate inputs at lower prices, and face lower cost of hiring labor. Equation (3.9) suggests that average marginal costs are lower at locations where firms (a) have higher average productivity (higher T_o); (b) face stochastically lower effective prices for their tasks thanks to better market access (higher A_o); and (c) face lower costs of hiring labor (lower w_o). Further, average marginal costs are more sensitive to market access and less

[16]Polygamma functions $\psi^{(n)}(\cdot)$ are defined as $\psi^{(n)}(x) = \frac{d^{n+1}\ln\Gamma(x)}{dx^{n+1}}$.

so to wages if materials share of costs α_o is higher and vice versa. Of these factors that influence average marginal costs, $\{\boldsymbol{T},\boldsymbol{\alpha}\}$ are exogenous location characteristics whereas $\{\boldsymbol{A},\boldsymbol{w}\}$ constitute endogenous price variables.

What factors affect dispersion of marginal costs in equilibrium at a location? Marginal costs of firms differ from one another due to differences in productivity and due to differences in effective prices faced for their respective tasks. Equation (3.10) sheds light on the role of these two channels. The first term in equation (3.10) reveals the contribution of differences in productivities across firms. Locations where dispersion in productivities is higher (lower θ_o) will have a higher dispersion in marginal costs. More importantly, the second term in equation (3.10) reveals the contribution of differences in effective prices faced by firms. Focussing on equation (3.10), its contribution is governed by three factors. First and foremost, the contribution is decreasing in ζ. A lower ζ increases the likelihood of high draws of match-specific productivities and therefore generates higher dispersion in effective prices. This factor is common across all locations. Second, the contribution is higher at locations with higher materials share of costs (higher α_o). Naturally, if materials form a larger share of costs, dispersion in price of materials plays a larger role. Finally, the contribution is lower at locations with higher numbers of tasks (lower $\mathbb{E}[1/K_o]$ or higher κ_o). With higher numbers of tasks, for one firm's price of materials to be substantially higher than another, it requires a larger number of high draws of match-specific productivities. Since such an occurrence is unlikely, locations with higher κ_o have lower dispersion in materials prices across firms.

3.3.3. *Conditional Choice Probabilities & Firm-to-Firm Trade.* I turn to predictions for firm-to-firm trade. Since these are not aggregate implications but rather are at the firm-to-firm level, it is not meaningful to consider the limiting economy. Therefore, I consider a sufficiently large economy along the sequence in Definition 2 such that Assumption 5 holds, i.e., $\lambda/M \ll 1$, $|\lambda a_0^{\zeta} - 1| < \varepsilon_1$, and $|a_0| < \varepsilon_2$ for arbitrarily small values of ε_1 and ε_2. Recall from equation (3.1) that firms choose suppliers for tasks based on suppliers' marginal costs, trade costs faced by them, and match-specific productivities associated with the task under consideration. While trade costs $\boldsymbol{\tau}$ constitute σ_0, match-specific productivities are unknown and suppliers' marginal costs $c_o(s)$ are determined endogenously. I therefore characterize conditional choice probabilities for supplier choice, i.e., probabilities for choice of supplier conditional on its marginal cost but in expectation over match-specific productivities that are yet to be realized. Let $\pi^0_{od}(s,b)$ denote the probability with which firm b selects firm s for any one of its tasks. Prior to realizing the match-specific productivities for each task $\{a_{od}(s,b,k)\}_{k \in \mathcal{K}_d(b)}$, the probability of firm s

getting selected for any one of the tasks by firm b is common across all tasks. That is, $\pi^0_{od}(s,b) = \pi^0_{od}(s,b,k) = \mathbb{E}_{\{a\}}[\mathbf{1}\{s = s^*_d(b,k)\,|\,a\}]$ where the expectation operator is over all realizations of $a_{od}(s,b,k)$. The following proposition provides expressions for conditional choice probabilities $\pi^0_{od}(s,b)$ as well as for $\rho^0_{od}(s,b)$, the probability with which firm b selects firm s for at least one of its tasks – thereby determining the extensive margin of firm-to-firm trade. As it turns out, these probabilities are independent of the identity of the buyer at the destination and therefore can be written as $\pi^0_{od}(s,-)$ and $\rho^0_{od}(s,-)$.

Proposition 4. *For any realization of σ_0, conditional on firm s's marginal cost being $c_o(s)$, the probability with which any firm located in d selects firm s located in o for any given task is*

$$\pi^0_{od}(s,-) = \frac{c_o(s)^{-\zeta}\tau_{od}^{-\zeta}}{\sum_{s'\in\mathcal{M}} c_{o'}(s')^{-\zeta}\tau_{o'd}^{-\zeta}}. \tag{3.11}$$

Further, conditional on firm s's marginal cost being $c_o(s)$, the probability with which any firm located in d selects firm s located in o for at least one of its tasks is

$$\rho^0_{od}(s,-) = \frac{1 - e^{-\kappa_d \pi^0_{od}(s,-)}}{1 - e^{-\kappa_d}}. \tag{3.12}$$

The above proposition is key to understanding what drives network formation among firms in the model and how it enables the model to match empirical regularities described in Section 2. On one hand, equation (3.11) highlights the factors that influence the likelihood of a supplier s from o getting selected by a buyer at d for any one of its tasks. Firms with lower marginal costs, denoted by $c_o(s)$, are more likely to get selected for more tasks. Firms that are located nearer to the buyers and face lower trade costs, denoted by τ_{od}, are more likely to get selected for more tasks. Moreover, the elasticity of the likelihood of getting selected with respect to marginal costs or trade costs is decreasing in ζ. That is, $\frac{\partial \ln \pi^0_{od}(s,-)}{\partial \ln c_o(s)} = \frac{\partial \ln \pi^0_{od}(s,-)}{\partial \ln \tau_{od}} = -\zeta$. With lower ζ, Assumption 4 implies that high match-specific productivities are more likely and the choice of supplier is less sensitive to other factors, i.e., its marginal cost and the trade cost faced by it. On the other hand, equation (3.12) shows that the same factors also influence whether a supplier s from o gets selected by a buyer at d for at least one of its tasks or none at all, i.e., the extensive margin of firm-to-firm trade. Since $\rho^0_{od}(s,-)$ is increasing in $\pi^0_{od}(s,-)$, marginal costs $c_o(s)$, trade costs τ_{od} and dispersion in match-productivities governed by ζ affect the extensive margin of supplier choice in the same manner as above. In addition to these factors, equation (3.12) also suggests that firms are more likely to find customers at destinations where buyers have higher numbers of tasks (higher κ_d). Naturally, if buyers have larger numbers of tasks, the supplier draws a larger number of match-specific productivities, has a better chance of getting high draws and hence get selected by a buyer.

In summary, this proposition channels the role of the upstream margin — at any location d, firms with lower marginal costs are likely to find more customers and are also likely to be used intensively by them. The role of geography in the upstream margin comes from the dependence of these probabilities on trade costs — firms from o are less likely to be successful both at the extensive and intensive margins of firm-to-firm trade across potential customers at d if o is farther, i.e., τ_{od} is higher. These results then lead to predictions for trade between locations. Since those are aggregate predictions, they are derived for the limiting economy.

3.3.4. *Sourcing Probabilities & Trade between Locations.* Conditional choice probabilities of supplier choice naturally aggregate to sourcing probabilities. That is, the probability with which any buyer sources inputs from o for any one its tasks can be obtained as the sum of conditional choice probabilities associated with all the suppliers located at o. The limiting economy assumption comes in handy here as it allows aggregation across firms within a location. Conditional choice probabilities from Proposition 4 together with properties of the cross-sectional distributions of effective prices and marginal costs from Propositions 1 and 3 lead to the next proposition. This proposition characterizes sourcing probabilities across origins by firm b, denoted by $\pi^0_{od}(\bullet,b)$, as well as $\rho^0_{od}(\bullet,b)$, the probability with which firm b sources from o for at least one of its tasks. As in the previous proposition, these probabilities are independent of the identity of the buyer at the destination and therefore can be written as $\pi^0_{od}(\bullet,-)$ and $\rho^0_{od}(\bullet,-)$.

Proposition 5. *For any realization of σ_0, the probability with which any firm located in d selects a supplier from o for any given task is*

$$\pi^0_{od}(\bullet,-) = \frac{\mu_o \Gamma\left(1-\frac{\zeta}{\theta_o}\right) T_o^{\frac{\zeta}{\theta_o}} w_o^{-\zeta(1-\alpha_o)} \mathbb{E}_{\{K_o\}}\left[\Gamma\left(2-\frac{\alpha_o}{K_o}\right)^{K_o}\right] A_o^{\alpha_o} \tau_{od}^{-\zeta}}{A_d}. \tag{3.13}$$

Further, the probability with which any firm located in d selects a supplier from o for at least one of its tasks is

$$\rho^0_{od}(\bullet,-) = \frac{1-e^{-\kappa_d \pi^0_{od}(\bullet,-)}}{1-e^{-\kappa_d}}. \tag{3.14}$$

Sourcing probabilities in equation (3.13) hark back to market access defined in equation (3.7). Recall that market access is a weighted sum of attractiveness of all locations for a particular destination. Equation (3.13) suggests the probability with which a buyer from d sources intermediate inputs from o for any one of its tasks is given by the contribution of location o towards market access of firms at d. Firms at d are more likely to source inputs from o if there are a larger number of firms at o (higher μ_o), firms at o have higher productivities on average (higher T_o), wage w_o is lower, or firms at o have better market access (higher A_o). Other factors $\Gamma\left(1-\frac{\zeta}{\theta_o}\right)$ and $\mathbb{E}_{\{K_o\}}\left[\Gamma\left(2-\frac{\alpha_o}{K_o}\right)^{K_o}\right]$ capture that fact that when materials share α_o is higher or dispersion parameter θ_o is lower supply chains routed through firms at o are likely to be more efficient. The same factors also affect the likelihood of a buyer at d sourcing from o for at least one of its tasks or none at all. Since $\rho^0_{od}(\bullet,-)$ is increasing in $\pi^0_{od}(\bullet,-)$, a similar explanation holds for origin selection at the extensive margin. In addition to these factors, equation (3.14) also suggests that firms at d are more likely to source from o if they have higher numbers of tasks (higher κ_d). The explanation for this parallels that of how equation (3.12) affects the extensive margin of firm-to-firm trade.

Further, under the simplifying assumption that $\theta_o = \theta$, $\alpha_o = \alpha$, and $\kappa_d = \kappa$ at all locations o, the sourcing probabilities in equation (3.13) can be simplified as follows.

$$\pi_{od} = \frac{\mu_o\left(T_o w_o^{-\theta(1-\alpha)} A_o^{\alpha\cdot\frac{\theta}{\zeta}} \tau_{od}^{-\theta}\right)^{\frac{\zeta}{\theta}}}{A_d}, \tag{3.15}$$

where $A_d = \sum_o \left(\mu_o T_o w_o^{-\zeta(1-\alpha)} A_o^{\alpha}\right)\tau_{od}^{-\zeta}$ denotes the market access at location d. This bears resemblance to aggregate trade shares between locations obtained in Eaton and Kortum (2002) and Bernard et al. (2003). In their case, aggregate trade share is given by

$$\pi_{od} = \frac{T_o w_o^{-\theta(1-\alpha)} A_o^{\alpha}\tau_{od}^{-\theta}}{A_d},$$

where $A_d = \sum_o \left(T_o w_o^{-\theta(1-\alpha)} A_o^{\alpha}\right)\tau_{od}^{-\theta}$ denotes the market access at location d, α denotes the materials share of costs while T_o and θ are parameters of the Fréchet productivity distribution at location o given by $\mathbb{P}(z_o(s) \leq z) = e^{-T_o z^{-\theta}}\mathbf{1}\{z \geq 0\}$.

The sourcing probabilities also bear resemblance to aggregate trade shares between locations obtained in Melitz (2003) and Chaney (2008). In their case, aggregate trade share is given by

$$\pi_{od} = \frac{\mu_o\left(T_o w_o^{-\theta(1-\alpha)} A_o^{\alpha}\tau_{od}^{-\theta}\right) f_{od}^{-\left(\frac{\theta}{\sigma-1}-1\right)}}{A_d},$$

where $A_d = \sum_o \left(\mu_o T_o w_o^{-\theta(1-\alpha)} A_o^{\alpha}\right) \tau_{od}^{-\theta} f_{od}^{-\left(\frac{\theta}{\sigma-1}-1\right)}$ denotes the market access at location d, f_{od} denotes fixed costs of exporting from location o to location d, σ denotes the elasticity of substitution across differentiated goods while T_o and θ are parameters of the Pareto productivity distribution at location o given by $\mathbb{P}(z_o(s) \leq z) = \left(1 - T_o z^{-\theta}\right) \mathbf{1}\left\{z \geq T_o^{1/\theta}\right\}$.

In this context, two facts are worth noting about Equation (3.15): (a) the elasticity of trade shares with respect to trade costs comes from the shape parameter of match-specific productivities ζ and not the dispersion of productivities θ, and (b) trade shares are increasing in the density of firms at the origin μ_o. The former unlinks the dispersion in idiosyncratic productivities from the trade elasticity while the latter introduces a probabilistic notion of "love of variety" within the Ricardian framework.

3.3.5. *Trade Equilibrium.* Equation (3.13) suggests that the probability of sourcing from a particular origin o is common for all tasks across all firms at a destination d and that the choice is conditionally independent across firms at the destination. Therefore, the law of large numbers implies that in the limiting economy aggregate trade shares converge to the sourcing probabilities, i.e., $\lim_{t\to\infty} \pi_{od}^0(\mathcal{E}_t) = \pi_{od}^0(\bullet, -)$. This brings us to the proposition below which states that the trade equilibrium in the limiting economy is satisfied with trade shares given by $\pi_{od}^0(\bullet, -)$ for all networks that are realized for any given σ_0.

Proposition 6. *For any realization of σ_0, $\boldsymbol{w} \equiv \{w_d : d \in \mathcal{J}\}$ solves the following system of equations for all realizations of σ_1:*

$$\frac{w_o L_o}{1-\alpha_o} = \sum_{d\in\mathcal{J}} \pi_{od}^0(\bullet, -) \frac{w_d L_d}{1-\alpha_d} \tag{3.16}$$

This concludes the characterization of equilibrium prices and brings us to the definition of the trade equilibrium below.

Definition 3. For any given σ_0, the trade equilibrium in the limiting economy is defined as the vector of wages $\boldsymbol{w}$ such that (a) market access at each location satisfies equation (3.7); (b) trade shares coincide with sourcing probabilities in equation (3.13) and (c) the market clearing condition in equation (3.16) holds.

The trade equilibrium along with tractable expressions for firm-to-firm trade and aggregate trade in Propositions 4 and 5 give rise to transparent estimating equations for the model, to which I turn next.

4. Empirical Framework

This section lays out the procedure for estimation of the model and counterfactual analysis. The objective of estimation is to infer the aggregate state σ_0 that consists of trade costs, firms' idiosyncratic productivities, and firms' task intensities given the observed data. Estimation relies on Proposition 4. With the estimated model, counterfactual analysis for large economies is then conducted by relying on Proposition 6 to evaluate the change in aggregate outcomes that results in response to shocks deriving from a change in the aggregate state σ_0 to σ_0'. For clarity, state variables Δ, parameters Θ, and data $\mathbb{D}$ are grouped as follows:

$$\begin{aligned}
&\Delta \equiv \left\{ \{c_o(s)^{-\zeta} : s \in \mathcal{M}\}, \left\{\tau_{od}^{-\zeta} : (o,d) \in \mathcal{J} \times \mathcal{J}\right\} \right\} \\
&\Theta \equiv \{\zeta, \boldsymbol{\alpha}, \boldsymbol{\theta}, \boldsymbol{\kappa}\} \\
&\mathbb{D} \equiv \{\{\pi_{od}(s,b) : (s,b) \in \mathcal{M}_o \times \mathcal{M}_d\}, \boldsymbol{X}_{od} : (o,d) \in \mathcal{J} \times \mathcal{J}\}
\end{aligned}$$

where $\pi_{od}(s,b)$ denotes the share of firm s in firm b's material costs and $\boldsymbol{X}_{od}$ denotes the vector of bilateral origin-destination observables such as distance and borders etc. In what follows, terms with superscript $(\cdot)^0$ denote true values and those with superscript $(\cdot)^*$ denote corresponding estimates. Changes in quantities are denoted by $\widehat{(\cdot)}$.[17] For example, $\pi^0_{od}(s,b)$ denotes true values of conditional choice probabilities, $\pi^*_{od}(s,b)$ denotes estimates of conditional choice probabilities , and $\widehat{\pi^0_{od}(s,b)}$ denotes changes in conditional choice probabilities from the initial to the counterfactual state.

4.1. **Estimation of Marginal Costs and Trade Frictions.** I reformulate the economic model developed in the previous section as a multinomial logit model of supplier choice for tasks of each of the firms and estimate it semi-parametrically. Firm's marginal costs are estimated as firm fixed effects and bilateral origin-destination fixed effects correspond to a structural gravity specification for estimating trade frictions. Trade frictions are then estimated by projecting bilateral fixed effects on observables. Together, these provide estimates of conditional choice probabilities for firm-to-firm trade as well as sourcing probabilities for trade between locations.

4.1.1. *Marginal Costs & Structural Gravity.* The econometric model can be motivated using the balls and bins problem. Consider the multinomial random variable characterized by a firm b located at d throwing $K_d(b)$ balls (one for each of its tasks) into M bins.

[17]For any variable x that changes it value to x' in a counterfactual state, change in x is denoted as $\widehat{x} = {x'}/{x}$.

Each of these bins corresponds to a potential supplier, denoted by s. The probability with which any of these balls falls into the bin indexed s is given by the expression for $\pi^0_{od}(s,-)$ from Proposition 4. A realization of this random variable consists of the proportion of balls that landed in each of the bins. Since tasks are symmetric and the production function of firm b takes the Cobb-Douglas functional form, the model counterpart of this realization is the vector of cost shares of firm b across all suppliers in the economy. In other words, the cost share of firm b that can be attributed to firm s stands in for the relative frequency of firm s's successes in getting selected across firm b's tasks. Since there are a discrete number of tasks, $\pi^0_{od}(s,b)$ is only the expected share of tasks for which firm b uses the output of firm s. Any given realization may deviate from this expected value for particularly high or low realizations of match-specific productivities and from randomness in buyer-seller encounters between firms.[18] Therefore, making use of Proposition 4, the estimating equation can be expressed as a multinomial logit function:[19]

$$\mathbb{E}[\pi_{od}(s,b)] = \frac{c_o(s)^{-\zeta}\tau_{od}^{-\zeta}}{\sum_{s'\in\mathcal{M}} c_{o'}(s')^{-\zeta}\tau_{o'd}^{-\zeta}} \tag{4.1}$$

This multinomial logit specification is non-standard because of two reasons. On one hand, firms' marginal costs (included as firm fixed effects) are endogenously determined in the model through supplier choice decisions of all the firms in the economy. Since match-specific productivities are independent across firms and tasks in their production function, the supplier choice decision is however conditionally independent. Therefore, firm fixed effects estimated using this specification can be treated as the conditional distribution of marginal costs without resorting to full solution methods to estimate the model. This is analogous to the estimation of conditional choice probabilities in

[18]One could draw an analogy by reinterpreting the Eaton and Kortum (2002) model of trade between countries as the representative agent in the destination country throwing infinitely many balls (one for each commodity arranged on a continuum) into a finite number of bins (one for each origin country). Since the bins are finite in number while balls are infinitely many, sourcing probabilities coincide with aggregate trade shares deterministically. In contrast, the model here is of trade between firms where the customer firm throws a finite number of balls (one for each task) into potentially infinitely many bins (one for every firm in the economy). Since the bins are infinitely many in number while balls are finite in number, neither conditional choice probabilities determine firm-to-firm trade shares deterministically nor do sourcing probabilities determine aggregate trade shares deterministically.

[19]In related work, Eaton et al. (2013) also specify a multinomial likelihood function for international trade between countries derived from a different economic model and conduct estimation using pseudo-maximum likelihood estimation à la Gourieroux et al. (1984). The dimensionality of their estimation program is determined by the number of countries which is a much smaller number compared to the specification here where the dimensionality is determined by the number of firms that runs into millions.

dynamic discrete choice models following Hotz and Miller (1993) and its application to network formation with many agents in Menzel (2015).[20] On the other hand, since there are a large number of firms in the economy, estimation of the multinomial logit model would typically require high-dimensional non-linear optimization over a very large number of parameters to solve for the estimates. This can be computationally infeasible using standard Newton methods when the number of fixed effects runs into millions. However, this issue can be avoided by appealing to several special features of the multinomial likelihood function. First, estimates can be obtained using the Poisson likelihood function with additional fixed effects (see Baker (1994); Taddy (2015)). Second, Poisson likelihood estimation automatically satisfies adding up constraints implied by the model (see Fally (2015)). Third, Poisson likelihood specification allows solving for fixed effects in closed-form (for example, see Hausman et al. (1984)). Finally, subsequent estimation of trade frictions using bilateral fixed effects does not suffer from the incidental parameters problem and hence can be conducted through the conditional maximum likelihood approach. Formally, the estimation problem is as follows:

$$\Delta^* = \arg\max_{\Delta} \frac{1}{M} \sum_{b \in \mathcal{M}} \ln f_{\text{MNL}}(\mathbb{D} \mid \Delta), \tag{4.2}$$

$$f_{\text{MNL}}(\mathbb{D} \mid \Delta) \propto \prod_{s \in \mathcal{M}} \left(\frac{c_o(s)^{-\zeta} \tau_{od}^{-\zeta}}{\sum_{s' \in \mathcal{M}} c_{o'}(s')^{-\zeta} \tau_{o'd}^{-\zeta}} \right)^{\pi_{od}(s,b)}$$

The above specification with fixed effects however presents a problem of perfect multicollinearity in regressors. Note that dummy variables associated with $\left\{c_o(s)^{-\zeta} : s \in \mathcal{M}_o\right\}$ and $\left\{\tau_{od}^{-\zeta} : d \in \mathcal{J}\right\}$ are collinear for all such locations o. Hence, I make the following normalizations so that these fixed effects are identified up to scale.

Assumption 6. *For all $s \in \mathcal{M}_o, o \in \mathcal{J}$, let $c_o(s) = c_o \widetilde{c}_o(s)$ such that $\left(\sum_{s \in \mathcal{M}_o} \widetilde{c}_o(s)^{-\zeta}\right)^{-1/\zeta} = 1$.*

[20] To see this clearly, note that marginal costs of any firm b admits the following recursive representation.

$$\underbrace{c_d(b)}_{\text{value function}} = \frac{w_d^{1-\alpha_d}}{z_d(b)} \times \prod_{k=1}^{K_d(b)} \min_{s \in \mathcal{S}_d(b)} \left\{ \frac{\bar{m}_{od}(s,b,k) \tau_{od}}{a_{od}(s,b,k)} \times \underbrace{c_o(s)}_{\text{upstream value function}} \right\}^{\frac{\overbrace{\alpha_d}^{\text{discount factor}}}{K_d(b)}}$$

In this context, conditional choice probabilities from Proposition 4 are therefore the probabilities with which any given supplier s is chosen for any one of the buyer b's tasks.

The above assumption normalizes the power average $\left(\sum_{s\in\mathcal{M}_o}\widetilde{c}_o(s)^{-\zeta}\right)^{-1/\zeta}$ of firms' marginal costs relative to their location average to unity. It separates within and between location heterogeneity in firms' marginal costs. The within location component is captured by differences in $\widetilde{c}_o(s)$ while the between location component is captured by differences in c_o across locations.[21] Under this assumption, the first order conditions implied by the likelihood maximization problem in equation (4.2) can be solved to obtain closed-form estimators for fixed effects as described in the proposition below.

Proposition 7. *Under Assumption 6, estimates from equation (4.2) are given by:*

$$\left(\widetilde{c}_o(s)^{-\zeta}\right)^{*} = \frac{\sum_d \pi_{od}(s,\bullet)}{\sum_{s'\in\mathcal{M}_o}\sum_d \pi_{od}(s',\bullet)} \quad \forall s\in\mathcal{M}_o, o\in\mathcal{J}, \tag{4.3}$$

$$\left(\frac{c_o^{-\zeta}\tau_{od}^{-\zeta}}{\sum_{o'} c_{o'}^{-\zeta}\tau_{o'd}^{-\zeta}}\right)^{*} = \frac{1}{M_d}\sum_{b\in\mathcal{M}_d}\pi_{od}(\bullet,b) \quad \forall (o,d)\in\mathcal{J}\times\mathcal{J} \tag{4.4}$$

where $\pi_{od}(s,\bullet)\equiv\sum_{b\in\mathcal{M}_d}\pi_{od}(s,b)$ *and* $\pi_{od}(\bullet,b)\equiv\sum_{s\in\mathcal{M}_o}\pi_{od}(s,b)$.

The estimators for firm fixed effects in equation (4.3) neatly bridge theoretical predictions on firm-to-firm trade in equation (3.11) and empirical regularities arising from the decomposition in equation (2.1). The decomposition in equation (2.1) suggested that larger firms also tend to have higher intensity of use. Conditional choice probabilities in equation (3.11) predict that firms with low marginal costs are likely to have higher intensity of use. Equation (4.3) shows that firms' intensity of use is a sufficient statistic for its marginal costs, albeit scaled with an elasticity ζ. In addition, the theoretical expression for bilateral origin-destination fixed effects in equation (4.4) corresponds to a structural gravity specification. For any pair of locations (o,d), the estimator for this specification is the *simple* average of the cost share across firms at d that can be attributed to purchase of goods from firms in o. This is the empirical counterpart of sourcing probabilities in equation (3.13).

4.1.2. *Trade Frictions, Conditional Choice Probabilities, and Sourcing Probabilities.* With firm fixed effects out of the way, thanks to equation (4.3), trade frictions can now be estimated by projecting bilateral origin-destination fixed effects (from equation (4.4)) on bilateral observables such as distance, borders etc., similar to gravity

[21]The between location component captures both differences in average marginal cost between locations and also differences arising from having a higher number of firms at one location than another. To see this clearly, note that if marginal costs are identical across firms at location o, i.e., $c_o(s)=\bar{c}_o$. Then, $c_o = M_o^{-1/\zeta}\bar{c}_o$, which depends on both the number of firms and the average marginal cost.

regressions, with the following estimating equation:

$$\mathbb{E}\left[\left(\frac{c_o^{-\zeta}\tau_{od}^{-\zeta}}{\sum_{o'} c_{o'}^{-\zeta}\tau_{o'd}^{-\zeta}}\right)^*\right] = \frac{\exp\left(\ln\left(c_o^{-\zeta}\right) + \boldsymbol{X}'_{od}\boldsymbol{\beta}\right)}{\sum_{o'}\exp\left(\ln\left(c_{o'}^{-\zeta}\right) + \boldsymbol{X}'_{o'd}\boldsymbol{\beta}\right)}. \tag{4.5}$$

This delivers estimates of origin fixed effects $\left(c_o^{-\zeta}\right)^*$ and trade frictions $\left(\tau_{od}^{-\zeta}\right)^* = \exp(\boldsymbol{X}'_{od}\boldsymbol{\beta}^*)$. The manner in which trade frictions are estimated here differs from the standard approach of projecting aggregate trade flows on distance and border dummies (for example, see Agnosteva et al. (2019)). The dependent variable implied by the model is not aggregate trade flows (for example, Santos Silva and Tenreyro (2006)) or aggregate trade shares (as in Eaton et al. (2013)) but average trade share across buyers at the destination. More specifically, the dependent variable $\frac{1}{M_d}\sum_{b\in\mathcal{M}_d}\pi_{od}(\bullet,b)$ is an unweighted average of the sourcing share from o across all buyers at a destination. While this is not comparable to aggregate trade flows, it closely related to aggregate trade shares. In contrast to average trade shares which is a simple average of sourcing shares across firms, the aggregate trade share is a weighted average of individual sourcing probabilities where each individual buyer is weighted by its size.[22] To the extent that size of buyers is correlated with their sourcing probabilities from an origin, aggregate trade shares bias the estimates of the trade frictions faced by individual firms for the purposes of estimation here.

Fitted shares from the gravity regressions are the estimates of sourcing probabilities. Estimates of conditional choice probabilities are then obtained from firm fixed effects and estimates of sourcing probabilities. Formally, the estimates of conditional choice probabilities and sourcing probabilities are respectively given by

$$\pi_{od}^*(s,-) = \left(\widetilde{c}_o(s)^{-\zeta}\right)^* \cdot \pi_{od}^*(\bullet,-), \tag{4.6}$$

$$\pi_{od}^*(\bullet,-) = \frac{\left(c_o^{-\zeta}\right)^*\left(\tau_{od}^{-\zeta}\right)^*}{\sum_{o'\in\mathcal{J}}\left(c_{o'}^{-\zeta}\right)^*\left(\tau_{o'd}^{-\zeta}\right)^*}. \tag{4.7}$$

[22] To see this clearly, note that measured aggregate trade share can be expressed as

$$\begin{aligned}\pi_{od} &= \frac{\sum_{b\in\mathcal{M}_d}\text{purchases}_d(b)\times\pi_{od}(\bullet,b)}{\sum_{b'\in\mathcal{M}_d}\text{purchases}_d(b')}.\\ &= \frac{1}{M_d}\sum_{b\in\mathcal{M}_d}\pi_{od}(\bullet,b) + \frac{Cov(\pi_{od}(\bullet,b),\text{purchases}_d(b))}{\frac{1}{M_d}\sum_{b'\in\mathcal{M}_d}\text{purchases}_d(b')}\end{aligned}$$

The procedure for estimating structural elasticities Θ is relegated to Appendix C.2. Estimation results are presented in Appendix C.5.

4.2. **Counterfactual Analysis.** For counterfactual analysis, I consider the limiting economy as described in Definition 2. To operationalize Proposition 6 for counterfactual analysis, it is useful to express the trade equilibrium in changes. The following definition states that and motivates the algorithm for evaluating counterfactual outcomes in response to shocks that derive from a change in the aggregate state σ_0 to σ_0'.

Definition 4. For any change in aggregate state σ_0 to σ_0', equilibrium change in wages $\widehat{\boldsymbol{w}} \equiv \{\widehat{w}_d : d \in \mathcal{J}\}$ and welfare $\widehat{\boldsymbol{V}} \equiv \left\{\widehat{V}_d : d \in \mathcal{J}\right\}$ are characterized the following system of equations for all realizations of σ_1 or σ_1':[23]

$$
\begin{aligned}
\widehat{A}_d &= \sum_o \pi^0_{od}(\bullet,-)\widehat{\delta}_{od}\widehat{w}_o^{-\zeta(1-\alpha_o)}\widehat{A}_o^{\alpha_o} \\
\widehat{\pi^0_{od}(\bullet,-)} &= \frac{\widehat{\delta}_{od}\widehat{w}_o^{-\zeta(1-\alpha_o)}\widehat{A}_o^{\alpha_o}}{\widehat{A}_d} \\
\frac{\widehat{w}_o w_o L_o}{1-\alpha_o} &= \sum_d \widehat{\pi^0_{od}(\bullet,-)}\pi^0_{od}(\bullet,-)\frac{\widehat{w}_d w_d L_d}{1-\alpha_d} \\
\widehat{V}_d &= \widehat{w}_d \widehat{A}_d^{1/\zeta}
\end{aligned}
$$

where $\widehat{\boldsymbol{\delta}} \equiv \left\{\widehat{\delta}_{od} : (o,d) \in \mathcal{J} \times \mathcal{J}\right\}$ is function of shocks that capture the resultant effect of change from σ_0 to σ_0'.

With this definition of the equilibrium in changes in the limiting economy, the procedure for computing counterfactual outcomes consists of three steps. First, I evaluate the expected value of aggregate and firm-level outcomes such as intensity of use and sales in the initial state. Second, I evaluate changes in aggregate outcomes when going from the initial state to the counterfactual state. This is done using a tâtonnement algorithm similar to Alvarez and Lucas (2007) and Dekle et al. (2008). Finally, I evaluate the expected value of aggregate and firm-level outcomes in the counterfactual state. Details of the procedure are stated in Appendix C.4.

5. Counterfactual Analysis

This section illustrates how the model can be used to assess the consequences of micro- and macro- shocks to the spatial economy. The procedure for counterfactual analysis proposed in Section 4.2 allows evaluation of welfare gains at the district level

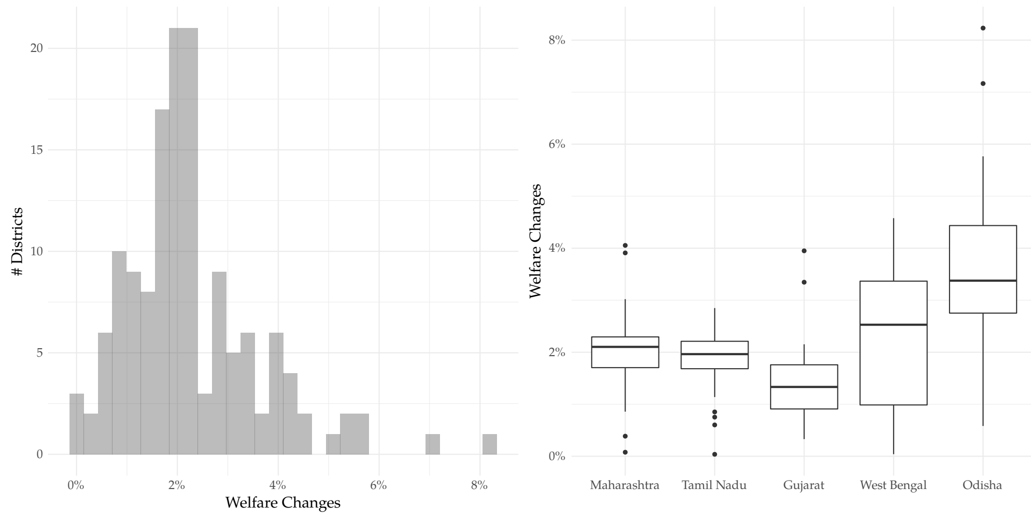
[23]The expression for welfare changes is derived in Appendix C.3.

as well as the impact on firms' sales and intensity of use of these shocks. First, I discuss a counterfactual experiment that reduces trade frictions across state borders. Second, I discuss how the production network of firms changes in response to an aggregate shock that uniformly reduces external trade frictions. Finally, I examine the implications of neutralizing firm-level distortions when they are either positively or negatively correlated with firm size on aggregate and firm-level outcomes.

5.1. **Decline in Border Frictions.** When India adopted the VAT in the early 2000s, its implementation was uneven. India has a federal system of government — one that divides the powers of government between the national and the state governments. Commercial taxation being overseen by the state government, individual states implemented their own respective VAT systems. This resulted in over 30 such systems coming into place across India. While this increased formality and tax compliance, it had the unintended consequence of regional segregation in organization of production, for three reasons. First, VAT increases formality because firms prefer to source inputs from other firms within the system to be able to collect tax credits on input purchases. Consequently, individual firms preferred to source inputs from firms within their own state's VAT system as opposed to one in a different state or VAT system. Second, the national government levied a sales tax on firm-to-firm transactions across state borders which made more efficient suppliers of intermediate inputs relatively more expensive if they were in a different state. Third, there were cumbersome inspections, especially at state borders that caused logistical delays. In July 2017, the federal government in India abolished all state VAT systems and introduced the Goods and Services Tax to serve as a single national VAT system. This eliminated sales taxes on inter-state movement of goods and harmonized the VAT structure across states in an attempt to reduce such barriers to intra-national trade.

In this context, I consider the aggregate and firm-level impact of a 10% decline in trade costs between district pairs crossing state borders to understand the potential impact of the GST reform on production networks in intra-national trade. Figure 5.1 suggests that this leads to sizable welfare gains of 1% in some districts to as large as 8% in others. Across states, the median district in larger states Gujarat, Maharashtra, and Tamil Nadu gains less than those in smaller states West Bengal and Odisha. Changes in firms' sales to other firms can be decomposed into changes in its intensity of use and changes in its average customer size as follows:

Figure 5.1. **Gains from Decline in Border Frictions**



Note. The left panel is a stacked histogram of welfare changes across districts. The right panel is a box and whiskers plot of welfare gains across districts within each state. States are arranged by economic size in descending order. The data used in this figure pertains to 2015-16.

$$\frac{\Delta\text{Sales}}{\text{Sales}} = \overbrace{\frac{\Delta\text{Intensity of Use}}{\text{Intensity of Use}}}^{\Delta\%\text{upstream margin}} + \underbrace{\frac{\Delta\text{Average Customer Size}}{\text{Average Customer Size}}}_{\Delta\%\text{downstream margin}}$$
$$+ \underbrace{\frac{\Delta\text{Intensity of Use}}{\text{Intensity of Use}} \times \frac{\Delta\text{Average Customer Size}}{\text{Average Customer Size}}}_{\text{second order term}}$$

To determine the relative contribution of the upstream and downstream margins to the dispersion in changes in firms' sales, I apply a Shapley decomposition (see Shorrocks (2013)). The Shapley decomposition determines the expected marginal contribution of each of these margins and the interaction term to the total variation in changes in firms' input sales; intuitively, it assigns the fraction of the R^2 of a regression that is due to each set of explanatory variables. Table 5.1 reports the results of this decomposition. Column (6) suggests that over half of the variation in changes in firms' sales can be attributed to endogenous changes in the network or the upstream margin while a third can be attributed to the downstream margin. In columns (1)-(5), when considering variation among firms within each state, the upstream margin accounts for only over third of the variation. This is because the incidence of the shock is at the state borders,

Table 5.1. **Decline in Border Frictions: Margins of Changes in Firms' Sales**

State	MH (1)	TN (2)	GJ (3)	WB (4)	OD (5)	All (6)
Δ% upstream margin	40.76%	40.81%	36.49%	39.44%	38.06%	55.69%
Δ% downstream margin	29.37%	34.14%	45.74%	31.44%	43.02%	33.45%
second order term	29.86%	25.04%	17.76%	29.14%	18.91%	10.85%

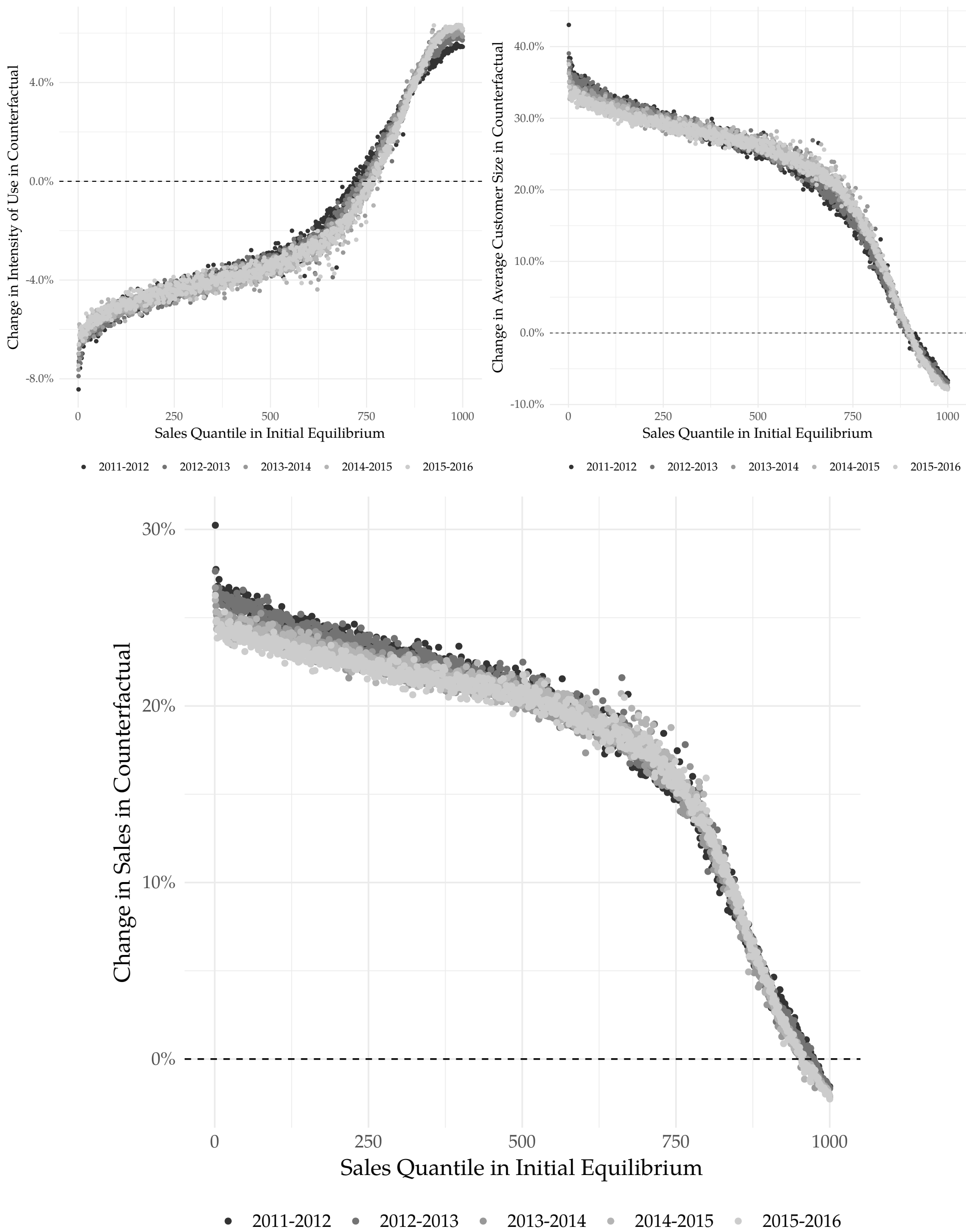
Note. This table reports the contribution of changes in firm's margins to the variation in changes in firms' sales calculated using a Shapley decomposition when firm-year observations are split by state. Here, MH stands for Maharashtra, GJ for Gujarat, TN for Tamil Nadu, WB for West Bengal, and OD for Odisha.

so the contribution of the upstream margin is not as high as that seen in the cross-state comparison in column (6).

A few points are in order. First, this decomposition is of sales to other firms and so would not exist in models without input-output linkages. Second, in models with exogenous production networks, i.e., with Cobb-Douglas technologies between firms, intensity of use does not respond to shocks. The large variation in the upstream margin would therefore be missing. Finally, in models with non-Cobb–Douglas technologies that endogenize the intensity with which existing suppliers are used but where the extensive margin of firm-to-firm trade does not respond to shocks, the explanatory power of the upstream margin would be understated. This is because changes in intensity of use accrue not only from changes in intensity of use by existing customers but also from changes in the number of customers. By allowing for substitution across both existing suppliers and new potential suppliers, the model is not only more general but also more tractable since it does not require calibrating the extensive margin of firm-to-firm trade to observed data.

5.2. **Market Integration.** A large body of recent literature studies barriers that impede trade between regions within a country and the gains that accrue from a reduction in those barriers (for a review, see Donaldson (2015)). I study the firm-level implications of a decline in relative costs of trading with firms in other districts. This experiment conceptually captures improvements in transportation infrastructure as well as any other policy changes that affect trade outside an agent's own location relative to within its own location. I consider the counterfactual scenario where external trade frictions

FIGURE 5.2. **Decline in Trade Frictions: Change in Firms' Sales and its Margins**



Note. For each year, firms are grouped into 1000 bins according to their sales in the initial equilibrium. Each bin consists of around 1000 firms. For firms in each of these bins, the top left panel plots the average percent change in intensity of use when trade frictions decline, the top right panel does the same for average customer size, and the bottom panel for sales to other firms.

decline by 10%.[24] With a decline in external trade costs, a large majority of firms are subject to opposing forces along the upstream and downstream margins.

Figure 5.2 depicts the effect of these margins of firms' sales to other firms. To understand this, it is useful to look at firms in four groups: (a) those in the top 5% in terms of sales; (b) those in the top 10% but not in the top 5%; (c) those in the top 25% but not in the top 10%; and (d) those in the bottom 75%. First, consider firms in group (a). Starting with the top left panel, these firms gain the most in intensity of use. At the same time, they are more likely to have had customers who are large, i.e., in the top 5% and whose sales declined. This implies that the average customer size of these firms declines as shown in the top right panel. These firms are subject to opposing forces on the upstream and downstream margins. While they gain in intensity of use, the lose sufficiently in average customer size that their sales decline. Second, consider firms in group (b). These firms still gain above 4% in intensity of use but are also likely to have had customers in the top 5% (whose sales declined). These firms are also subject to opposing forces on the upstream and downstream margins such that their sales increase. Third, consider firms in group (c). These firms gain less than 4% in intensity of use, are less likely to have had customers in the top 5% and so their average customer size increases. These firms are also subject to reinforcing forces on the upstream and downstream margins such that their sales increase. Finally, consider the large majority of firms in group (d). These firms lose in intensity of use, but are also much less likely to have had customers in the top 5%, so their average customer size increases. These firms are subject to opposing forces on the upstream and downstream margins. While they lose in intensity of use, the gain sufficiently in average customer size that their sales increase.

Taking stock, as trade frictions decline, firms with low production costs become more successful at farther or less remote destinations in getting selected for customers' tasks. This comes at the expense of firms with higher production costs who are now less successful in getting selected for tasks both locally and elsewhere. While intensity of use of firms in the bottom three quartiles decreases by as much as 8%, intensity of use for firms in the top quartile increases by as much as 4%. At the same time, firms in the top decile are more likely to have customers in the top 5% those for whom sales has

[24] Counterfactual outcomes are evaluated using the procedure described in Appendix C.4 with aggregate shocks given by:

$$\widehat{\delta}_{od} = \begin{cases} \frac{1}{1.1^{-\varsigma}} & o \neq d \\ 1 & o = d \end{cases}$$

There is no heterogeneity in shocks at the firm-level in this counterfactual experiment.

declined. Those customers produce less and source fewer inputs from firms in the top decile. Average customer size for firms in the top decile and quantity demanded from them declines. On the contrary, firms in the bottom nine deciles are less likely to have customers in the top 5% for whom sales has declined. For these firms, average customer size has increased. The net outcome of these margins acting on firms at all quantiles is that large firms' sales to other firms shrink where as those of a large majority of firms in the lower quantiles expands.

5.3. **Size-Dependent Distortions & Improvements in Allocative Efficiency.** A substantial literature has documented the presence of firm-level distortions in developing economies (for a review, see Atkin and Khandelwal (2020)). In this counterfactual experiment, I study the implications of neutralizing positively versus negatively size-dependent distortions affecting firms' labor input choice. The notion for such gains is similar in spirit to that in the closed economy model with labor wedges as in Hsieh and Klenow (2009), multiplier effects from inter-sectoral linkages as in Jones (2013), and trade as in Swiecki (2017). Unlike these papers, I consider the effect of removing firm-level distortions through the lens of a model of trade where production networks between firms respond endogenously. The experiment I consider homogenizes labor market distortions. That is, it eliminates dispersion in those firm-specific labor market "taxes" and hence consists of shocks at the firm level. In conducting this analysis, I assume that all tax revenue is rebated equally to local households both in the initial state and the counterfactual state and hence the level of the homogeneous tax rate in the counterfactual scenario does not affect welfare calculations.[25] Figure 5.3 shows that terms of trade effects are negative in a large number of districts when removing

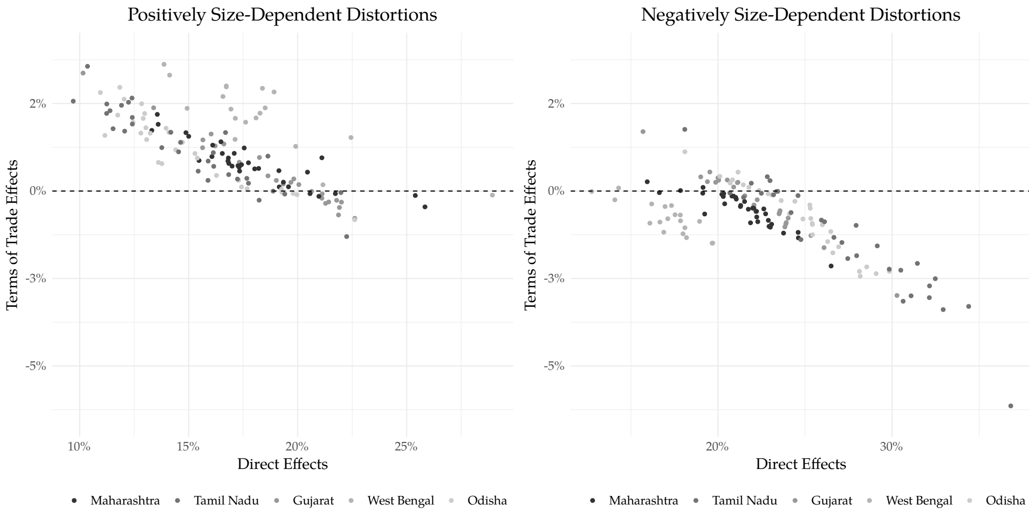
[25]Size-dependent distortions are generated as:

$$1+t_o(s)=\begin{cases}(1-q)^{-\frac{1}{\eta}} & \text{if distortions are positively size-dependent}\\ q^{-\frac{1}{\eta}} & \text{if distortions are negatively size-dependent}\end{cases},$$

where q denotes the quantile of the firm for sales to other firms and η denotes the shape parameter of Pareto distributed distortions drawn from the following cumulative distribution function:$\mathbb{P}\left(1+t_o(s)\leq 1+t\right) = \left(1-(1+t)^{-\eta}\right)\mathbf{1}\left\{t\geq 0\right\}$. For generating distortions, η was calibrated to 5. Counterfactual outcomes are evaluated using the procedure described in Appendix C.4 with firm-level and aggregate shocks respectively given by:

$$\widehat{\delta}_{od}(s)={}^{1}/_{(1+t_o(s))^{-\zeta(1-\alpha_o)}},$$
$$\widehat{\delta}_{od}={}^{1}/_{\mathbb{E}_{\{t_o\}}\left[(1+t_o)^{-\zeta(1-\alpha_o)}\right]}.$$

FIGURE 5.3. **Elimination of Size-Dependent Distortions: Direct & Indirect Effects**



Note. The left panel plots direct and terms of trade effects when distortions are positively size-dependent and the right panel when distortions are negatively size-dependent. Points are shaded by state in both panels, darker shades indicate richer states. For each district, direct effects are calculated as the increase the total factor productivity if each district were a closed economy. Terms-of-trade effects are calculated as the difference between the welfare change from the experiment and the direct effects.

negatively size-dependent distortions while they are largely positive when removing positively size-dependent distortions.

The result of removing distortions at the firm-level is that firms that faced higher tax rates and were too small, now expand, with labor being reallocated to them as in models of misallocation such as Restuccia and Rogerson (2008) and Hsieh and Klenow (2009). While this captures direct effects, the analysis here also takes into account indirect effects through input-output linkages between firms and the endogenous response of the network structure to these shocks. To examine how this experiment affects the production network between firms, I consider the decomposition of changes in firms' sales to other firms into changes in its intensity of use and changes in its average customer size. Table 5.2 reports the results of a Shapley decomposition of margins of sales. I find that changes in intensity of use explain majority of variation in changes in firms' sales — around 80% with positively size-dependent distortions and 75% with negatively size-dependent distortions. The downstream margin is however less important in the case of negatively size-dependent distortions than in the case of positively size-dependent distortions. This is because firms with lower sales and facing larger distortions are

TABLE 5.2. **Elimination of Size-Dependent Distortions: Margins of Changes in Firms' Sales**

State	MH (1)	TN (2)	GJ (3)	WB (4)	OD (5)	All (6)
Positively Size-Dependent Distortions						
Δ% upstream margin	73.87%	82.02%	82.52%	80.47%	74.82%	81.16%
Δ% downstream margin	11.81%	8.23%	6.75%	9.89%	12.00%	8.34%
second order term	14.31%	9.74%	10.71%	9.63%	13.17%	10.48%
Negatively Size-Dependent Distortions						
Δ% upstream margin	66.57%	73.23%	80.73%	78.01%	71.25%	75.08%
Δ% downstream margin	1.34%	1.40%	1.57%	3.11%	1.11%	1.58%
second order term	32.07%	25.35%	17.69%	18.80%	27.57%	23.32%

Note. This table reports the contribution of changes in firm's margins to the variation in changes in firms' sales calculated using a Shapley decomposition when firm-year observations are split by state. Here, MH stands for Maharashtra, GJ for Gujarat, TN for Tamil Nadu, WB for West Bengal, and OD for Odisha.

likely to have had higher production costs. Since their customers sourced inputs from relatively expensive suppliers, they likely had higher production costs themselves and therefore change relatively less in size when such distortions are neutralized.

6. Conclusion

This paper developed a new framework for analyzing aggregate and firm-level consequences of shocks to the spatial economy when customer-supplier linkages between firms evolve endogenously. I documented that Indian firms with higher sales to other firms tend to have more customers, tend to be used more intensively by those customers, and tend to have larger customers. Firms' intensity of use explains a vast majority of variation in their sales to other firms. The model explains this through a single dimension of firm heterogeneity: production costs. Firms with low production costs find more customers, are used more intensively by them and since their customers use cheaper inputs intensively, they lower production costs and become larger themselves. Furthermore, firms differ not only in their relative position in the production network, but also across space thereby facing different wages when hiring labor as well as different trade costs when sourcing inputs from potentially multiple locations.

Interdependence of link formation between firms in general equilibrium models of network formation typically restrains the use of simulation-based estimation to arbitrary scale, i.e., with very large numbers of firms. On the contrary, the procedure developed here makes estimation and counterfactual analysis both scalable and tractable. Firms' intensity of use was shown to be a sufficient statistic for their production costs — a key endogenous object of interest. As a result, estimation did not necessitate full solution of the model to obtain the distribution of production costs. Besides, counterfactual analysis did not require large-scale simulation either and was done under a large economy approximation to resolve aggregate uncertainty. In an empirical application, I show that a 10% decline in inter-state border frictions has sizable welfare gains ranging from 1% in some districts to as high as 8% in others. Moreover, over half of the variation in changes in firms' sales to other firms can be explained by endogenous changes in the network structure.

The framework developed here can be directly applied to answer questions that could be broadly classified as market integration, technology improvements, and improvement in allocative efficiency; nevertheless, it can serve as a fertile baseline model to answer a wider variety of questions where changes in the production network across firms can lead to aggregate consequences. In pursuit of parsimonious parametrization, the model abstracts from several realistic features of the network economy such as sectoral heterogeneity in technological requirements, supply chain dynamics, industry dynamics of entry and exit, heterogeneous search frictions, and richer bargaining environment between buyers and suppliers all of which are potential avenues for future research.

References

Acemoglu, D. and P. D. Azar (2020): "Endogenous Production Networks," *Econometrica*, 88, 33–82.

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012): "The Network Origins of Aggregate Fluctuations," *Econometrica*, 80, 1977–2016.

Agnosteva, D. E., J. E. Anderson, and Y. V. Yotov (2019): "Intra-national Trade Costs: Assaying Regional Frictions," *European Economic Review*, 112, 32 – 50.

Al-Najjar, N. I. (2004): "Aggregation and the Law of Large Numbers in Large Economies," *Games and Economic Behavior*, 47, 1 – 35.

Allen, T., C. Arkolakis, and X. Li (2020): "On the Equilibrium Properties of Network Models with Heterogeneous Agents," Working Paper 27837, National Bureau of Economic Research.

Alvarez, F. and R. E. Lucas (2007): "General Equilibrium Analysis of the Eaton-Kortum model of International Trade," *Journal of Monetary Economics*, 54, 1726 – 1768.

Antràs, P. and A. de Gortari (2020): "On the Geography of Global Value Chains," *Econometrica*, 88, 1553–1598.

Antràs, P., T. C. Fort, and F. Tintelnot (2017): "The Margins of Global Sourcing: Theory and Evidence from U.S. Firms," *American Economic Review*, 107, 2514–64.

Arkolakis, C. (2016): " A Unified Theory of Firm Selection and Growth," *The Quarterly Journal of Economics*, 131, 89–155.

Armenter, R. and M. Koren (2014): "A Balls-and-Bins Model of Trade," *American Economic Review*, 104, 2127–51.

Atkin, D. and A. K. Khandelwal (2020): "How Distortions Alter the Impacts of International Trade in Developing Countries," *Annual Review of Economics*, 12, 213–238.

Axtell, R. L. (2001): "Zipf Distribution of U.S. Firm Sizes," *Science*, 293, 1818–1820.

Baker, S. G. (1994): "The Multinomial-Poisson Transformation," *Journal of the Royal Statistical Society. Series D (The Statistician)*, 43, 495–504.

Baqaee, D. R. and E. Farhi (2019a): "The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem," *Econometrica*, 87, 1155–1203.

——— (2019b): "Networks, Barriers, and Trade," Working Paper 26108, National Bureau of Economic Research.

——— (2020): "Productivity and Misallocation in General Equilibrium," *The Quarterly Journal of Economics*, 135, 105–163.

Bernard, A., E. Dhyne, G. Magerman, K. Manova, and A. Moxnes (2019): "The Origins of Firm Heterogeneity: A Production Network Approach," Working Paper, Universite Libre de Bruxelles.

Bernard, A. B., J. Eaton, J. B. Jensen, and S. Kortum (2003): "Plants and Productivity in International Trade," *American Economic Review*, 93, 1268–1290.

Bigio, S. and J. LaO (2020): "Distortions in Production Networks," *The Quarterly Journal of Economics*, 135, 2187–2253.

Blaum, J., C. Lelarge, and M. Peters (2018): "The Gains from Input Trade with Heterogeneous Importers," *American Economic Journal: Macroeconomics*, 10, 77–127.

Boehm, J. and E. Oberfield (2020): "Misallocation in the Market for Inputs: Enforcement and the Organization of Production," *The Quarterly Journal of Economics*, 135, 2007–2058.

Breunig, R. (2001): "An Almost Unbiased Estimator of the Coefficient of Variation," *Economics Letters*, 70, 15 – 19.

Caliendo, L. and F. Parro (2014): "Estimates of the Trade and Welfare Effects of NAFTA," *The Review of Economic Studies*, 82, 1–44.

Caliendo, L., F. Parro, and A. Tsyvinski (2017): "Distortions and the Structure of the World Economy," Working Paper 23332, National Bureau of Economic Research.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011): "Robust Inference With Multiway Clustering," *Journal of Business and Economic Statistics*, 29, 238–249.

Carvalho, V. M. and N. Voigtlander (2014): "Input Diffusion and the Evolution of Production Networks," Working Paper 20025, National Bureau of Economic Research.

Chaney, T. (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 98, 1707–21.

——— (2014): "The Network Structure of International Trade," *American Economic Review*, 104, 3600–3634.

Costinot, A. and A. Rodríguez-Clare (2014): *Trade Theory with Numbers: Quantifying the Consequences of Globalization*, Elsevier, vol. 4 of *Handbook of International Economics*, 197–261.

Dekle, R., J. Eaton, and S. Kortum (2008): "Global Rebalancing with Gravity: Measuring the Burden of Adjustment," *IMF Staff Papers*, 55, 511–540.

Dingel, J. I. and F. Tintelnot (2020): "Spatial Economics for Granular Settings," Working Paper 27287, National Bureau of Economic Research.

Donaldson, D. (2015): "The Gains from Market Integration," *Annual Review of Economics*, 7, 619–647.

Eaton, J. and S. Kortum (2002): "Technology, Geography, and Trade," *Econometrica*, 70, 1741–1779.

Eaton, J., S. Kortum, and F. Kramarz (2011): "An Anatomy of International Trade: Evidence From French Firms," *Econometrica*, 79, 1453–1498.

——— (2016): "Firm-to-Firm Trade: Imports, Exports, and the Labor Market," Discussion Papers 16048, Research Institute of Economy, Trade and Industry (RIETI).

Eaton, J., S. Kortum, and S. Sotelo (2013): *International Trade: Linking Micro and Macro*, Cambridge University Press, vol. 2 of *Econometric Society Monographs*, 329–370.

Fally, T. (2015): "Structural Gravity and Fixed Effects," *Journal of International Economics*, 97, 76 – 85.

Fernandes, A. M., P. J. Klenow, S. Meleshchuk, D. Pierola, and A. Rodríguez-Clare (2018): "The Intensive Margin in Trade," Working Paper 25195, National Bureau of Economic Research.

Gaubert, C. and O. Itskhoki (2021): "Granular Comparative Advantage," *Journal of Political Economy*, 129, 871–939.

Gourieroux, C., A. Monfort, and A. Trognon (1984): "Pseudo Maximum Likelihood Methods: Theory," *Econometrica*, 52, 681–700.

Hausman, J., B. H. Hall, and Z. Griliches (1984): "Econometric Models for Count Data with an Application to the Patents-R & D Relationship," *Econometrica*, 52, 909–938.

Head, K. and T. Mayer (2014): "Gravity Equations: Workhorse,Toolkit, and Cookbook," in *Handbook of International Economics*, ed. by G. Gopinath, E. Helpman, and K. Rogoff, Elsevier, vol. 4 of *Handbook of International Economics*, 131 – 195.

Head, K. and J. Ries (2001): "Increasing Returns versus National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade," *The American Economic Review*, 91, 858–876.

Hopenhayn, H. A. (1992): "Entry, Exit, and firm Dynamics in Long Run Equilibrium," *Econometrica*, 60, 1127–1150.

Hotz, V. J. and R. A. Miller (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *The Review of Economic Studies*, 60, 497–529.

Hsieh, C.-T. and P. J. Klenow (2009): "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 124, 1403–1448.

Huneeus, F. (2018): "Production Network Dynamics and the Propagation of Shocks," Working papers, Yale University.

Jones, C. I. (2011): "Intermediate Goods and Weak Links in the Theory of Economic Development," *American Economic Journal: Macroeconomics*, 3, 1–28.

——— (2013): *Misallocation, Economic Growth, and Input–Output Economics*, Cambridge University Press, vol. 2 of *Econometric Society Monographs*, 419–456.

Jovanovic, B. (1982): "Selection and the Evolution of Industry," *Econometrica*, 50, 649–670.

Klenow, P. and A. Rodríguez-Clare (1997): "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?" in *NBER Macroeconomics Annual 1997, Volume 12*, National Bureau of Economic Research, Inc, NBER Chapters, 73–114.

Klette, T. J. and S. Kortum (2004): "Innovating Firms and Aggregate Innovation," *Journal of Political Economy*, 112, 986–1018.

Kortum, S. (1997): "Research, Patenting, and Technological Change," *Econometrica*, 65, 1389–1419.

Lim, K. (2017): "Firm-to-Firm Trade in Sticky Production Networks," Meeting Papers 280, Society for Economic Dynamics.

Liu, E. (2019): "Industrial Policies in Production Networks," *The Quarterly Journal of Economics*, 134, 1883–1948.

Luttmer, E. G. J. (2007): "Selection, Growth, and the Size Distribution of Firms," *The Quarterly Journal of Economics*, 122, 1103–1144.

McFadden, D. (1974): "Conditional Logit Analysis of Qualitative Choice Behavior," *Frontiers in Econometrics*, 105–142.

Melitz, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71, 1695–1725.

Menzel, K. (2015): "Strategic Network Formation with Many Agents," Working Papers, New York University.

Oberfield, E. (2018): "A Theory of Input-Output Architecture," *Econometrica*, 86, 559–589.

Restuccia, D. and R. Rogerson (2008): "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments," *Review of Economic Dynamics*, 11, 707 – 720.

Santos Silva, J. and S. Tenreyro (2006): "The Log of Gravity," *The Review of Economics and statistics*, 88, 641–658.

Shorrocks, A. F. (2013): "Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value," *The Journal of Economic Inequality*, 11, 99–126.

Swiecki, T. (2017): "Intersectoral Distortions and the Welfare Gains from Trade," *Journal of International Economics*, 104, 138 – 156.

Taddy, M. (2015): "Distributed Multinomial Regression," *The Annals of Applied Statistics*, 9, 1394–1414.

Taschereau-Dumouchel, M. (2017): "Cascades and Fluctuations in an Economy with an Endogenous Production Network," Meeting Papers 700, Society for Economic Dynamics.

Tintelnot, F., A. K. Kikkawa, M. Mogstad, and E. Dhyne (2018): "Trade and Domestic Production Networks," Working Paper 25120, National Bureau of Economic Research.

Online Appendix

Appendix A. Data & Empirical Regularities: Appendix

A.1. **Margins of Firms' Sales.** Similar to Equation 2.1, I construct a decomposition of firms' destination-specific sales as:

$$\text{input sales}_{od}(s) = \overbrace{N_{od}(s) \times \frac{\sum_{b\in\mathcal{M}_d}\pi_{od}(s,b)}{N_{od}(s)}}^{\text{upstream margin}} \times \underbrace{\frac{\sum_{b\in\mathcal{M}_d}\pi_{od}(s,b)\times\text{input costs}_d(b)}{\sum_{b\in\mathcal{M}_d}\pi_{od}(s,b)}}_{\text{downstream margin}}, \tag{A.1}$$

where input $\text{sales}_{od}(s)$ denotes input sales of firm s to customers at d and $N_{od}(s)$ denotes the number of customers of s who are located at d. Table A.1 provides results of this decomposition under different specifications.

A.2. **Margins of Intranational Trade.** Trade flows between Indian districts aggregated from firm-to-firm sales show that districts within the same state are more likely to trade than those across states. Among all possible pairs of districts, around 40% do not trade at all. For district pairs that trade with each other, I construct the following decomposition of trade flows into four factors:

$$\begin{aligned}\text{sales}_{od} = {} & \overbrace{N_{od} \times \frac{\sum_{s\in\mathcal{M}_o}N_{od}(s)}{N_{od}} \times \frac{\sum_{s\in\mathcal{M}_o}\sum_{b\in\mathcal{M}_d}\pi_{od}(s,b)}{\sum_{s\in\mathcal{M}_o}N_{od}(s)}}^{\text{upstream margin}} \\ & \times \underbrace{\frac{\sum_{s\in\mathcal{M}_o}\sum_{b\in\mathcal{M}_d}\pi_{od}(s,b)\times\text{input costs}_d(b)}{\sum_{s\in\mathcal{M}_o}\sum_{b\in\mathcal{M}_d}\pi_{od}(s,b)}}_{\text{downstream margin}},\end{aligned} \tag{A.2}$$

where $\text{sales}_{od} = \sum_{s\in\mathcal{M}_o}\sum_{b\in\mathcal{M}_d}\text{sales}_{od}(s,b)$, N_{od} denotes the sellers from o that sell at d. In this decomposition, the first three margins capture the role of the upstream margin whereas the third margin captures the role of the downstream margin in driving differences in aggregate trade flows. In considering this decomposition, I depart from the trade literature where these margins are regrouped such that the first margin is called the extensive margin of trade defined as the number of firms from o that sell at d and the remaining three margins are together called the intensive margin of trade

Table A.1. **Margins of Firms' Sales: Contribution to Total Variance**

	Sales		Destination-Specific Sales		
	(1)	(2)	(3)	(4)	(5)
# Customers	35%	36%	37%	23%	22%
Intensity per Customer	46%	46%	57%	57%	59%
Average Customer Size	19%	18%	6%	20%	19%
Fixed Effects:					
Seller×Year	—	—	✓	—	—
Origin×Year	—	✓	—	—	✓
Destination×Year	—	—	—	—	✓
Data Level:					
Seller×Year	•	•	—	—	—
Seller×Destination×Year	—	—	•	•	•
# observations	$5.6{\times}10^6$	$5.6{\times}10^6$	$18.2{\times}10^6$	$18.2{\times}10^6$	$18.2{\times}10^6$

Note. Columns (1) and (2) report the contribution of factors: # customers, intensity per customer, and average customer size, to the variance of firms' sales as per equation (2.1). Column (3), (4), and (5) report the contribution of those factors to the variance of firms' destination-specific sales as per equation (A.1).

average sales across the firms from o that enter d. [26] This is so as to emphasize the role of endogenous network formation and cross-border supply chains in determining aggregate trade flows. Table A.2 reports the results from this decomposition.

Appendix B. Theoretical Framework: Appendix

B.1. **Proof of Propositions 1 and 2.**

B.1.1. *Joint Distribution of the Lowest and the Second Lowest Effective Costs.* We begin by characterizing the joint distribution of the lowest and second lowest effective cost available to buyer b located at d, $\widetilde{F}_{p_d}\left(p^{(1)},p^{(2)}\right)=\mathbb{P}\left(p_d^*(b,k)\leq p^{(1)},p_d(b,k)\geq p^{(2)}\right)$. To do so, we evaluate the probability with which b receives exactly one offer with an effective cost no greater than $p^{(1)}$ and no other offers less than $p^{(2)}(>p^{(1)})$. The lowest cost offer $p^{(1)}$ can be from any one of the locations in $\mathcal{J}$. We evaluate the probability with which this offer is from any given location o and sum it across all locations. The probability with which b receives one offer with an effective cost no greater than $p^{(1)}$ from o and no other offers less than $p^{(2)}$ across all locations is given by:

[26] For example, see Eaton et al. (2011) and Fernandes et al. (2018) for such decomposition of the margins of international trade between countries where it is documented that the extensive margin accounts for over half the variation in trade flows between countries.

Table A.2. **Margins of Intranational Trade: Contribution to Total Variance**

	(1)	(2)	(3)	(4)
# Sellers	59%	57%	61%	58%
# Customers per Seller	8%	9%	8%	10%
Intensity per Customer	20%	20%	24%	26%
Average Customer Size	13%	13%	7%	6%
Fixed Effects:				
Origin×Year	—	✓	—	✓
Destination×Year	—	—	✓	✓
Data Level:				
Origin×Destination×Year	•	•	•	•
# observations	58,390	58,390	58,390	58,390
# dropped observations (zeros)	41,015	41,015	41,015	41,015
# district pairs	$141^2 \times 5$	$141^2 \times 5$	$141^2 \times 5$	$141^2 \times 5$

Note. This table reports the contribution of factors: # sellers, # customers per seller, intensity per customer, and average customer size, to the variance of trade flows between districts, as per equation (A.2).

$$\begin{cases} \binom{M_o}{1}\frac{\lambda}{M}\mathbb{P}\left(\frac{c_o(s)\tau_{od}}{a_{od}(s,b,k)} \leq p^{(1)}\right) & \text{if } o \neq d \\ \times\left(1-\frac{\lambda}{M}\mathbb{P}\left(\frac{c_o(s)\tau_{od}}{a_{od}(s,b,k)} \leq p^{(2)}\right)\right)^{M_o-1} & \\ \times\left(1-\frac{\lambda}{M}\mathbb{P}\left(\frac{c_d(s)\tau_{dd}}{a_{dd}(s,b,k)} \leq p^{(2)}\right)\right)^{M_d-1} & \\ \times\prod_{o' \notin \{o,d\}}\left(1-\frac{\lambda}{M}\mathbb{P}\left(\frac{c_{o'}(s)\tau_{o'd}}{a_{o'd}(s,b,k)} \leq p^{(2)}\right)\right)^{M_{o'}} & \\ & \\ \binom{M_o-1}{1}\frac{\lambda}{M}\mathbb{P}\left(\frac{c_o(s)\tau_{od}}{a_{od}(s,b,k)} \leq p^{(1)}\right) & \text{if } o = d \\ \times\left(1-\frac{\lambda}{M}\mathbb{P}\left(\frac{c_o(s)\tau_{od}}{a_{od}(s,b,k)} \leq p^{(2)}\right)\right)^{M_o-2} & \\ \times\prod_{o' \neq o}\left(1-\frac{\lambda}{M}\mathbb{P}\left(\frac{c_{o'}(s)\tau_{o'd}}{a_{o'd}(s,b,k)} \leq p^{(2)}\right)\right)^{M_{o'}} & \end{cases}$$

Under Assumption 5, the probability with which b encounters exactly one supplier who can deliver at a cost no greater than $p^{(1)}$ and encounters no other suppliers with offers less than $p^{(2)}$ across all locations is given by:

$$\widetilde{F}_{p_d}(p^{(1)}, p^{(2)}) = \sum_o \lambda\mu_o \mathbb{P}\left(\frac{c_o(s)\tau_{od}}{a_{od}(s,b,k)} \leq p^{(1)}\right)\exp\left(-\sum_{o'}\lambda\mu_{o'}\mathbb{P}\left(\frac{c_{o'}(s)\tau_{o'd}}{a_{o'd}(s,b,k)} \leq p^{(2)}\right)\right)$$

Using the limit $\lim_{t\to\infty}\lambda_t a_{0,t}^{\zeta}\to 1$, this can be further simplified as $A_d\left(p^{(1)}\right)^{\zeta}\exp\left(-A_d\left(p^{(2)}\right)^{\zeta}\right)$ where $A_d=\sum_o \mu_o\tau_{od}^{-\zeta}\mathbb{E}\left[c_o(\cdot)^{-\zeta}\right]$ is obtained as follows:

$$
\begin{aligned}
A_d p^{\zeta} &= \sum_o \lambda\mu_o \mathbb{P}\left(\frac{c_o(s)\tau_{od}}{a_{od}(s,b,k)}\le p\right)\\
&= \sum_o \lambda\mu_o \mathbb{E}_{\{c_o\}}\left[1-F_a\left(\frac{c_o(s)\tau_{od}}{p}\right)\right]\\
&= \left(\sum_o \mu_o\tau_{od}^{-\zeta}\mathbb{E}\left[c_o(\cdot)^{-\zeta}\right]\right)p^{\zeta}\\
\implies A_d &= \sum_o \mu_o\tau_{od}^{-\zeta}\mathbb{E}\left[c_o(\cdot)^{-\zeta}\right]
\end{aligned}
$$

The density function is then obtained by the negative cross-derivative of $\widetilde{F}_{p_d}\left(p^{(1)},p^{(2)}\right)$ as follows:

$$
\begin{aligned}
\widetilde{F}'_{p_d}\left(p^{(1)},p^{(2)}\right) &= -\frac{\partial^2 F_{p_d}\left(p^{(1)},p^{(2)}\right)}{\partial p^{(1)}\partial p^{(2)}}\\
&= -\frac{\partial\left(A_d\left(p^{(1)}\right)^{\zeta}\right)}{\partial p^{(1)}}\frac{\partial\left(\exp\left(-A_d\left(p^{(2)}\right)^{\zeta}\right)\right)}{\partial p^{(2)}}\\
&= \zeta^2 A_d^2\left(p^{(1)}p^{(2)}\right)^{\zeta-1}e^{-A_d\left(p^{(2)}\right)^{\zeta}}
\end{aligned}
$$

B.1.2. *Distribution of Effective Prices.* We derive an expression for $F_{p_d}(p)$, that is, the probability with which any firm b located in d faces an effective price no greater than p for one of its tasks k. Firm b faces an effective price no greater than p if the second-lowest cost available to it is no less than p. This is obtained as:

$$
\begin{aligned}
F_{p_d}(p) &= \int_0^p\left(\int_0^{p^{(2)}} F'_{p_d}\left(p^{(1)},p^{(2)}\right)dp^{(1)}\right)dp^{(2)}\\
&= 1-A_d p^{\zeta}\exp\left(-A_d p^{\zeta}\right)-\exp\left(-A_d p^{\zeta}\right)
\end{aligned}
$$

B.1.3. *Derivation of Market Access.*

$$
c_o(\cdot)=w_o^{1-\alpha}\left(\prod_{k=1}^{K_o(\cdot)} p_o(\cdot,k)^{1/K_o}\right)^{\alpha}
$$

$$\implies \mathbb{E}\left[c_o(\cdot)^{-\zeta}\right] = \mathbb{E}\left[\left(\frac{w_o^{1-\alpha}\left(\prod_{k=1}^{K_o(\cdot)} p_o(\cdot,k)^{1/K_o}\right)^{\alpha}}{z_o(\cdot)}\right)^{-\zeta}\right]$$

$$= w_o^{-\zeta(1-\alpha_o)}\mathbb{E}\left[\prod_{k=1}^{K_o(\cdot)} p_o(\cdot,k)^{-\alpha_o\zeta/K_o(s)}\right]\mathbb{E}\left[z_o(\cdot)^{\zeta}\right]$$

$$= w_o^{-\zeta(1-\alpha_o)}\left(\mathbb{E}\left[\mathbb{E}\left[\prod_{k=1}^{K_o(\cdot)} p_o(\cdot,k)^{-\alpha_o\zeta/K_o(\cdot)} \mid K_o\right]\right]\right)\mathbb{E}\left[z_o(\cdot)^{\zeta}\right]$$

$$= w_o^{-\zeta(1-\alpha_o)}\left(\mathbb{E}\left[\prod_{k=1}^{K_o(\cdot)} \Gamma\left(2-\frac{\alpha_o}{K_o(\cdot)}\right)A_o^{\frac{\alpha_o}{K_o(\cdot)}}\right]\right)\Gamma\left(1-\frac{\zeta}{\theta_o}\right)T_o^{\frac{\zeta}{\theta_o}}$$

$$= \mathbb{E}\left[\Gamma\left(2-\frac{\alpha_o}{K_o(\cdot)}\right)^{K_o(\cdot)}\right]\Gamma\left(1-\frac{\zeta}{\theta_o}\right)T_o^{\frac{\zeta}{\theta_o}}w_o^{-\zeta(1-\alpha_o)}A_o^{\alpha_o}$$

This implies that $\{A_d\}_{d\in\mathcal{J}}$ solves the following fixed point problem:

$$A_d = \sum_o \mu_o \tau_{od}^{-\zeta}\Gamma\left(1-\frac{\zeta}{\theta_o}\right)\mathbb{E}\left[\Gamma\left(2-\frac{\alpha_o}{K_o(\cdot)}\right)^{K_o(\cdot)}\right]T_o^{\frac{\zeta}{\theta_o}}w_o^{-\zeta(1-\alpha_o)}A_o^{\alpha_o}$$

It can be similarly shown that effective prices for needs faced by households is also given by $F_{p_d}(\cdot)$ The following lemma states that the above fixed point problem that solves for market access is well-defined in the sense that it admits a unique positive solution. The proof strategy follows from Allen et al. (2020).

Lemma. *The following system of equations*

$$A_d = \sum_o R_{od} A_o^{\alpha_o},$$

$$R_{od} = \mu_o \tau_{od}^{-\zeta}\Gamma\left(1-\frac{\zeta}{\theta_o}\right)\mathbb{E}_{\{K_o\}}\left[\Gamma\left(2-\frac{\alpha}{K_o}\right)^{K_o}\right]T_o^{\frac{\zeta}{\theta_o}}w_o^{-\zeta(1-\alpha_o)}.$$

(1) *has at least one positive solution*
(2) *has at most one positive solution (up to scale)*
(3) *the unique solution can be computed as the limit of a simple iterative procedure.*

Proof. First, I establish existence of positive solution to the system of equations. Define operator $T:\mathbb{R}_{++}^J \to \mathbb{R}_{++}^J$ where $T(\boldsymbol{A}) = (\sum_o R_{o1}A_o^{\alpha_o}, \cdots, \sum_o R_{oJ}A_o^{\alpha_o})'$. Note that all components of R_{od} are positive and finite. Then, by construction, for any d, not all R_{od}s are zero. Therefore, for any $\boldsymbol{A} \gg 0$, $\sum_o R_{o1}A_o^{\alpha_o} \geq \underline{A} > 0$. Further, there exists $\bar{A} < \infty$ such that $\sum_o R_{od} A_o^{\alpha_o} \leq \bar{A}$. Now consider the operator $T:\mathcal{A}\rightarrow\mathcal{A}$ defined by $T(A_1,\cdots,A_J)=(\sum_o R_{o1} A_o^{\alpha_o},\cdots,\sum_o R_{oJ} A_o^{\alpha_o})'$. Suppose $\mathcal{A}=\left\{\boldsymbol{A}\in\mathbb{R}_{++}^J \,|\, \underline{A}\leq A_d\leq\bar{A}\forall d\right\}$.Then, if $\boldsymbol{A}\gg 0$, it follows that $T(\boldsymbol{A})\gg 0$. Note that $\mathcal{A}$ is closed and bounded. Since $\mathcal{A}\subset\mathbb{R}_{++}^J$, this implies that $\mathcal{A}$ is compact. Further, $\mathcal{A}$ is non-empty and convex, and T is continuous. Then, by Brouwer's fixed point theorem, $T(\bullet)$ has a fixed point. This establishes existence of a solution the system of equations.

To establish uniqueness, let's suppose by way of contradiction that the system of equations has two different solutions $\boldsymbol{A}^{(0)},\boldsymbol{A}^{(1)}$ that are not linear transformations of each other. Denote $\bar{a}=\max_d \frac{A_d^{(1)}}{A_d^{(0)}}$ and $\underline{a}=\min_d \frac{A_d^{(1)}}{A_d^{(0)}}$. Notice that $\frac{\bar{a}}{\underline{a}}\geq 1$. Thus the system of equations can be expressed as:

$$\frac{A_d^{(1)}}{A_d^{(0)}}=\frac{\sum_o R_{od}\left(\frac{A_d^{(1)}}{A_d^{(0)}}\right)^{1-\alpha_o}\left(A_d^{(0)}\right)^{1-\alpha_o}}{A_d^{(0)}}$$

Suppose $\bar{d}=\text{argmax}_d\left(\frac{A_d^{(1)}}{A_d^{(0)}}\right)$ and $\underline{\alpha}=\min\alpha_o$, then we have:

$$\begin{aligned}
\frac{A_{\bar{d}}^{(1)}}{A_{\bar{d}}^{(0)}} &= \bar{a} \\
\implies \frac{\sum_o R_{o\bar{d}}\left(\frac{A_o^{(1)}}{A_o^{(0)}}\right)^{1-\alpha_o}\left(A_o^{(0)}\right)^{1-\alpha_o}}{A_{\bar{d}}^{(0)}} &= \bar{a} \\
\implies \frac{\sum_o R_{o\bar{d}}\bar{a}^{1-\underline{\alpha}}\left(A_o^{(0)}\right)^{1-\alpha_o}}{A_{\bar{d}}^{(0)}} &\geq M \\
\implies \frac{\sum_o R_{o\bar{d}}\left(A_o^{(0)}\right)^{1-\alpha_o}}{A_{\bar{d}}^{(0)}}\bar{a}^{1-\underline{\alpha}} &\geq \bar{a} \\
\implies \bar{a}^{\underline{\alpha}} &\leq 1 \\
\implies \bar{a} &\leq 1
\end{aligned}$$

Similarly, we can show that $\underline{a}\geq 1$. This implies that $\frac{\bar{a}}{\underline{a}}\leq 1$. But by construction $\frac{\bar{a}}{\underline{a}}\geq 1$. Therefore, it must be the case that $\frac{\bar{a}}{\underline{a}}=1$ or $\boldsymbol{A}^{(0)}=\boldsymbol{A}^{(1)}$. This establishes uniqueness.

Next, I show that the solution to the system of equations can be obtained via a simple iterative procedure. Starting from any strictly positive $\boldsymbol{A}^{(0)}$, we construct a sequence

$\boldsymbol{A}^{(t)}$ successively in the following way,

$$A_d^{(t)} = \sum_o R_{od} \left(A_o^{(t-1)} \right)^{\alpha_o}$$

Denote $\bar{a}^{(t)} = \max_d \frac{A_d^{(t)}}{A_d^{(t-1)}}$ and $\underline{a}^{(t)} = \min_d \frac{A_d^{(t)}}{A_d^{(t-1)}}$. Notice that $\frac{\bar{a}^{(t)}}{\underline{a}^{(t)}} \geq 1$.

Suppose $\bar{d} = \operatorname{argmax}_d \left(\frac{A_d^{(t)}}{A_d^{(t-1)}} \right)$ and $\underline{\alpha} = \min \alpha_o$, then we have:

$$\begin{aligned}
\frac{A_{\bar{d}}^{(t)}}{A_{\bar{d}}^{(t-1)}} &= \bar{a}^{(t)} \\
\implies \frac{\sum_o R_{o\bar{d}} \left(\frac{A_o^{(t-1)}}{A_o^{(t-2)}} \right)^{1-\alpha_o} \left(A_o^{(t-2)} \right)^{1-\alpha_o}}{A_{\bar{d}}^{(t-1)}} &= \bar{a}^{(t)} \\
\implies \frac{\sum_o R_{o\bar{d}} \left(A_o^{(0)} \right)^{1-\alpha_o}}{A_{\bar{d}}^{(0)}} \left(\bar{a}^{(t-1)} \right)^{1-\underline{\alpha}} &\geq \bar{a}^{(t)} \\
\implies \frac{\bar{a}^{(t)}}{\left(\bar{a}^{(t-1)} \right)^{1-\underline{\alpha}}} &\leq 1
\end{aligned}$$

Similarly, we can show that $\frac{\underline{a}^{(t)}}{\left(\underline{a}^{(t-1)} \right)^{1-\bar{\alpha}}} \geq 1$. This implies the following

$$\begin{aligned}
\frac{\bar{a}^{(t)}}{\left(\bar{a}^{(t-1)} \right)^{1-\underline{\alpha}}} &\leq \frac{\underline{a}^{(t)}}{\left(\underline{a}^{(t-1)} \right)^{1-\bar{\alpha}}} \\
\implies \frac{\bar{a}^{(t)}}{\underline{a}^{(t)}} &\leq \frac{\left(\bar{a}^{(t-1)} \right)^{1-\underline{\alpha}}}{\left(\underline{a}^{(t-1)} \right)^{1-\bar{\alpha}}} \\
&\leq \frac{\left(\bar{a}^{(t-1)} \right)^{1-\underline{\alpha}}}{\left(\underline{a}^{(t-1)} \right)^{1-\underline{\alpha}}} \\
\implies \frac{\bar{a}^{(t)}}{\underline{a}^{(t)}} &\leq \frac{\bar{a}^{(t-1)}}{\underline{a}^{(t-1)}}
\end{aligned}$$

Since $\frac{\bar{a}^{(t)}}{\underline{a}^{(t)}} \geq 1 \forall t$, this implies that $\lim_{t \to \infty} \frac{\bar{a}^{(t)}}{\underline{a}^{(t)}} = 1$. That is, the solution can be computed as the limit of a simple iterative procedure. $\square$

B.2. **Proof of Proposition 2.**

$$\mathbb{P}\left(\frac{p_d(b,k)}{p_d^*(b,k)} \leq \bar{m} \,|\, p_d(b,k) = p^{(2)} \right) = \mathbb{P}\left(p_d^*(b,k) \geq \frac{p_d(b,k)}{\bar{m}} \,|\, p_d(b,k) = p^{(2)} \right)$$

$$=1-\int_0^{\frac{p_d(b,k)}{\bar{m}}}\frac{\widetilde{F}'_{p_d}\left(p^{(1)},\frac{p_d(b,k)}{\bar{m}}\right)}{F'_{p_d}\left(\frac{p_d(b,k)}{\bar{m}}\right)}dp^{(1)}$$
$$=1-\bar{m}^{-\zeta}$$

B.3. **Proof of Proposition 3.**

Corollary 1. *The distribution of idiosyncratic productivities of firms at location o satisfies the following moment conditions:*

$$\mathbb{E}[\log z_o(\cdot)]=-\frac{\psi^{(0)}(1)}{\theta}+\frac{1}{\theta}\log T_o,$$
$$\mathbb{E}\left[(\log z_o(\cdot)-\mathbb{E}[\log z_o(\cdot)])^2\right]=\frac{\psi^{(1)}(1)}{\theta^2}.$$

Proof. From Assumption 1, notice that:

$$\mathbb{P}(z_o\leq z)=e^{-T_o z^{-\theta}}\mathbf{1}\{z>0\}$$
$$=\exp\left(-\left(\frac{z}{T_o^{1/\theta}}\right)^{-\theta}\right)\mathbf{1}\{z>0\}$$

The results then follow from Lemma 1. □

Corollary 2. *The distribution of effective price faced by firms at location d satisfies the following moment conditions:*

$$\mathbb{E}[\log p_d(\cdot,\cdot)\,|\,A_o]=\frac{\psi^{(0)}(2)}{\zeta}-\frac{1}{\zeta}\log A_d,$$
$$\mathbb{E}\left[(\log p_d(\cdot,\cdot)-\mathbb{E}[\log p_d(\cdot,\cdot)])^2\right]=\frac{\psi^{(1)}(2)}{\zeta^2}.$$

Proof. From Proposition 1, notice that:

$$\mathbb{P}(p_d(\cdot,\cdot)\leq p)=\left(1-e^{-A_d p^\zeta}-A_d p^\zeta e^{-A_d p^\zeta}\right)\mathbf{1}\{p\geq 0\}$$
$$=\left(1-\exp\left(-\left(\frac{p}{A_d^{-1/\zeta}}\right)^\zeta\right)-\left(\frac{p}{A_d^{-1/\zeta}}\right)^\zeta\exp\left(-\left(\frac{p}{A_d^{-1/\zeta}}\right)^\zeta\right)\right)\mathbf{1}\{p>0\}$$

The results then follow from Lemma 2. □

B.3.1. *Proof of Equation (3.9).*

$$\log c_o(s)=(1-\alpha_o)\log w_o+\frac{\alpha_o}{K_o(s)}\sum_{k=1}^{K_o(s)}\log p_o(s,k)-\log z_o(s)$$

$$\implies \mathbb{E}[\log c_o(\cdot)] = (1-\alpha_o)\log w_o + \mathbb{E}\left[\mathbb{E}\left[\frac{\alpha_o}{K_o(\cdot)}\sum_{k=1}^{K_o(\cdot)}\log p_o(\cdot,k)\,|\,K_o\right]\right] - \mathbb{E}[\log z_o(\cdot)]$$

$$= (1-\alpha_o)\log w_o + \mathbb{E}\left[\frac{\alpha_o}{K_o(\cdot)}\sum_{k=1}^{K_o(\cdot)}\mathbb{E}[\log p_o(\cdot,k)\,|\,K_o]\right] - \mathbb{E}[\log z_o(\cdot)]$$

$$= (1-\alpha_o)\log w_o + \alpha_o\mathbb{E}[\log p_o(\cdot,\cdot)] - \mathbb{E}[\log z_o(\cdot)]$$

$$= \frac{\alpha_o\psi^{(0)}(2)}{\zeta} + \frac{\psi^{(0)}(1)}{\theta_o} - \frac{\alpha_o}{\zeta}\log A_o - \frac{1}{\theta_o}\log\left(T_o w_o^{-\theta(1-\alpha)}\right)$$

B.3.2. *Proof of Equation (3.10).*

$$Var[\log c_o(\cdot)] = \mathbb{E}\left[(\log c_o(\cdot) - \mathbb{E}[\log c_o(\cdot)])^2\right]$$

$$= \mathbb{E}\left[\mathbb{E}\left[\left(\frac{\alpha_o}{K_o(\cdot)}\sum_{k=1}^{K_o(\cdot)}(\log p_o(\cdot,k) - \mathbb{E}[\log p_o(\cdot,\cdot)])\right)^2 |\,K_o\right]\right]$$

$$+\mathbb{E}\left[(\log z_o(\cdot) - \mathbb{E}[\log z_o(\cdot)])^2\right]$$

$$= \mathbb{E}\left[\frac{\alpha_o^2}{K_o(\cdot)^2}\sum_{k=1}^{K_o(\cdot)}\mathbb{E}\left[(\log p_o(\cdot,k) - \mathbb{E}[\log p_o(\cdot,\cdot)])^2\,|\,K_o\right]\right]$$

$$+Var[\log z_o(\cdot)]$$

$$= \mathbb{E}\left[\frac{1}{K_o(\cdot)}\right]\alpha_o^2 Var[\log p_o(\cdot,\cdot)] + Var[\log z_o(\cdot)]$$

$$= \mathbb{E}\left[\frac{1}{K_o}\right]\frac{\alpha_o^2\psi^{(1)}(2)}{\zeta^2} + \frac{\psi^{(1)}(1)}{\theta_o^2}$$

B.4. **Proof of Proposition 4.**

B.4.1. *Proof of Equation (3.11).* Consider a pair of firms s located in o and b located in d. Now, suppose the marginal cost of firm s from o and it's cost of shipping goods to d are $c_o(s)$ and τ_{od} respectively. For any task k and match-specific productivity $a_{od}(s,b,k) = a$, the effective cost incurred by s of delivering its goods for task k by b is $\frac{c_o(s)\tau_{od}}{a}$. Supplier s is selected by b for task k if b encounters s with match-specific productivity a and b does not encounter any other supplier for whom it is effectively less costly to deliver the good (including the event that b meets s and the match-specific productivity realized is higher than a). The probability with which b selects s for any of its tasks with match-specific productivity a is given by:

$$\pi^0_{od}(s,b,k\,|\,a)=\frac{\lambda}{M}\times\prod_{s'\in\mathcal{M}}\left(1-\frac{\lambda}{M}\mathbb{P}\left(\frac{c_{o'}(s')\tau_{o'd}}{a_{o'd}(s',b,k)}\leq\frac{c_o(s)\tau_{od}}{a}\right)\right)$$
$$=\frac{\lambda}{M}\times\exp\left(\sum_{s'\in\mathcal{M}}\ln\left(1-\frac{\lambda}{M}\mathbb{P}\left(\frac{c_{o'}(s')\tau_{o'd}}{a_{o'd}(s',b,k)}\leq\frac{c_o(s)\tau_{od}}{a}\right)\right)\right)$$

Since $\lambda=o(M)$, considering $\frac{\lambda}{M}\ll 1$ and using the approximation $\ln(1+x)\approx x$ for $|x|\ll 1$, the above expression simplifies as:

$$\pi^0_{od}(s,b,k\,|\,a)=\frac{\lambda}{M}\exp\left(-\frac{\lambda}{M}\sum_{s'\in\mathcal{M}}\mathbb{P}\left(\frac{c_{o'}(s')\tau_{o'd}}{a_{o'd}(s',b,k)}\leq\frac{c_o(s)\tau_{od}}{a}\right)\right)$$

Taking expectation over all possible realizations of $a_{od}(s,b,k)$, we obtain:

$$\begin{aligned}
\pi^0_{od}(s,b,k)&=\mathbb{E}_{\{a\}}\left[\pi^0_{od}(s,b,k\,|\,a)\right]\\
&=\frac{\lambda}{M}\int_0^\infty\exp\left(-\frac{\lambda}{M}\sum_{s'\in\mathcal{M}}\mathbb{P}\left(\frac{c_{o'}(s')\tau_{o'd}}{a_{o'd}(s',b,k)}\leq\frac{c_o(s)\tau_{od}}{a}\right)\right)dF_a(a)\\
&=\frac{\lambda}{M}\int_{a_0}^\infty\exp\left(-\frac{\lambda}{M}\sum_{s'\in\mathcal{M}}\mathbb{P}\left(a_{o'd}(s',b,k)\geq\frac{c_{o'}(s')\tau_{o'd}}{c_o(s)\tau_{od}}a\right)\right)d\left(1-(a/a_0)^{-\zeta}\right)\\
&=\frac{\lambda a_0^\zeta}{M}\int_{a_0}^\infty\exp\left(-\frac{\lambda a_0^\zeta}{M}\sum_{s'\in\mathcal{M}}\mathbf{1}\left(\frac{c_{o'}(s')\tau_{o'd}}{c_o(s)\tau_{od}}a\geq a_0\right)\left(\frac{c_{o'}(s')\tau_{o'd}}{c_o(s)\tau_{od}}a\right)^{-\zeta}\right.\\
&\left.-\frac{\lambda}{M}\sum_{s'\in\mathcal{M}}\mathbf{1}\left(\frac{c_{o'}(s')\tau_{o'd}}{c_o(s)\tau_{od}}a\leq a_0\right)\right)\zeta a^{-\zeta-1}da\\
&=\frac{1}{M}\int_0^\infty\exp\left(-\frac{1}{M}\sum_{s'\in\mathcal{M}}\left(\frac{c_{o'}(s')\tau_{o'd}}{c_o(s)\tau_{od}}\right)^{-\zeta}a^{-\zeta}\right)d\left(-a^{-\zeta}\right)\\
&=\frac{c_o(s)^{-\zeta}\tau_{od}^{-\zeta}}{\sum_{s'\in\mathcal{M}}c_{o'}(s')^{-\zeta}\tau_{o'd}^{-\zeta}}\Gamma(1)\\
&=\frac{c_o(s)^{-\zeta}\tau_{od}^{-\zeta}}{\sum_{s'\in\mathcal{M}}c_{o'}(s')^{-\zeta}\tau_{o'd}^{-\zeta}}
\end{aligned}$$

Here, in the fifth line we utilize Assumption 5 which implies that in sufficiently large economies $\lim_{t\to\infty}\lambda_t a_{0,t}^\zeta\to 1$ and $\lim_{t\to\infty}a_{0,t}\to 0$ such that $\frac{\lambda}{M}\sum_{s'\in\mathcal{M}}\mathbf{1}\left(\frac{c_{o'}(s')\tau_{o'd}}{c_o(s)\tau_{od}}a\leq a_0\right)\to 0$ for all firms s'. Since $\pi_{od}(s,b,k)$ is independent of the identity of the task k, we write

$\pi^0_{od}(s,b) = \pi^0_{od}(s,b,k)$. Further, since $\pi^0_{od}(s,b)$ is independent of the identity of the buyer at any location d, we write $\pi^0_{od}(s,-) = \pi^0_{od}(s,b)$.

B.4.2. *Proof of Equation (3.12).* The probability with which a firm s located in o is selected by any given firm at d for at least one of its tasks is given by:

$$\begin{aligned}
\mathbb{P}(s \text{ gets selected for at least one task at } d) &= \sum_{K=1}^{\infty} \mathbb{P}(K_d(b)=K)\left(1-\left(1-\pi^0_{od}(s,-)\right)^K\right) \\
&= \left(1-\sum_{K=1}^{\infty} \mathbb{P}(K_d(b)=K)\left(1-\pi^0_{od}(s,-)\right)^K\right) \\
&= \left(1-\frac{e^{-\kappa_d}}{(1-e^{-\kappa_d})}\sum_{K=1}^{\infty}\frac{\left(\kappa_d(1-\pi^0_{od}(s,-))\right)^K}{K!}\right) \\
&= \frac{1-e^{-\kappa_d \pi^0_{od}(s,-)}}{1-e^{-\kappa_d}}
\end{aligned}$$

B.5. **Proof of Proposition 5.**

B.5.1. *Proof of Equation (3.13).* The probability with which any firm at d sources from firms at o for any of its tasks is given by

$$\begin{aligned}
\pi^0_{od}(\bullet,-) &= \left(\lim_{t\to\infty}\frac{M_o}{M}\right)\left(\lim_{t\to\infty}\frac{1}{M_o}\sum_{s\in\mathcal{M}_o}\pi^0_{od}(s,-)\right) \\
&= \left(\lim_{t\to\infty}\frac{M_o}{M}\right)\left(\lim_{t\to\infty}\frac{1}{M_o}\sum_{s\in\mathcal{M}_o}\frac{c_o(s)^{-\zeta}\tau_{od}^{-\zeta}}{A_d}\right) \\
&= \frac{\mu_o \mathbb{E}\left[c_o(\cdot)^{-\zeta}\right]\tau_{od}^{-\zeta}}{A_d} \\
&= \frac{\mu_o \Gamma\left(1-\frac{\zeta}{\theta_o}\right) T_o^{\frac{\zeta}{\theta_o}} w_o^{-\zeta(1-\alpha_o)} \mathbb{E}\left[\Gamma\left(2-\frac{\alpha_o}{K_o(\cdot)}\right)^{K_o(\cdot)}\right] A_o^{\alpha_o}\tau_{od}^{-\zeta}}{A_d}
\end{aligned}$$

B.5.2. *Proof of Equation (3.14).* The probability with which any firm at d sources at least one task from o is given by:

$$\begin{aligned}
\mathbb{P}(d \text{ firm sources at least one task from } o) &= \sum_{K=1}^{\infty}\mathbb{P}(K_d(\cdot)=K)\left(1-\left(1-\pi^0_{od}(\bullet,-)\right)^K\right) \\
&= \left(1-\sum_{K=1}^{\infty}\mathbb{P}(K_d(\cdot)=K)\left(1-\pi^0_{od}(\bullet,-)\right)^K\right)
\end{aligned}$$

$$
\begin{aligned}
&= \left(1 - \frac{e^{-\kappa_d}}{(1-e^{-\kappa_d})} \sum_{K=1}^{\infty} \frac{\left(\kappa_d(1-\pi_{od}^0(\bullet,-))\right)^K}{K!}\right) \\
&= \left(1 - \frac{e^{-\kappa_d}}{(1-e^{-\kappa_d})} \left(e^{\kappa_d\left(1-\pi_{od}^0(\bullet,-)\right)} - 1\right)\right) \\
&= \frac{1-e^{-\kappa_d \pi_{od}^0(\bullet,-))}}{1-e^{-\kappa_d}}
\end{aligned}
$$

B.6. **Proof of Proposition 6.** For any realization of σ, labor demand by firm b at d can be expressed as:

$$
l_d(b,\sigma) = \frac{1}{w_d(\sigma)}(1-\alpha_d)c_d(b,\sigma)y_d(b,\sigma)
$$

Substituting the above expression in the labor market clearing for location d, we obtain:

$$
\begin{aligned}
L_d &= \sum_{b\in\mathcal{M}_d} l_d(b,\sigma) \\
&= \sum_{b\in\mathcal{M}_d} \frac{1}{w_d(\sigma)}(1-\alpha_d)c_d(b,\sigma)y_d(b,\sigma) \\
\implies \sum_{b\in\mathcal{M}_d} c_d(b,\sigma)y_d(b,\sigma) &= \frac{w_d(\sigma)L_d}{1-\alpha_d}
\end{aligned}
$$

Goods market clearing condition for firm s located at o can be simplified as:

$$
\begin{aligned}
y_o(s,\sigma) &= \sum_d \sum_{b\in\mathcal{M}_d} \sum_{k\in\mathcal{K}_d(b)} \frac{\tau_{od}(s,\sigma)m_{od}(s,b,k,\sigma)}{a_{od}(s,b,k,\sigma)} \\
&+ \sum_d \sum_{i\in\mathcal{L}_d} \sum_{n\in\mathcal{N}_d(i)} \frac{\tau_{od}(s,\sigma)q_{od}(s,i,n,\sigma)}{g_{od}(s,i,n,\sigma)} \\
\implies c_o(s,\sigma)y_o(s,\sigma) &= \sum_d \alpha_d \sum_{b\in\mathcal{M}_d} \left(\frac{1}{K_d(b)} \sum_{k\in\mathcal{K}_d(b)} \frac{\mathbf{1}\{s=s_d^*(b,k,\sigma)\}}{\bar{m}_d(b,k,\sigma)}\right) c_d(b,\sigma)y_d(b,\sigma) \\
&+ \sum_d \sum_{i\in\mathcal{L}_d} \left(\frac{1}{N_d(i)} \sum_{n\in\mathcal{N}_d(i)} \frac{\mathbf{1}\{s=s_d^*(i,n,\sigma)\}}{\bar{m}_d(i,n,\sigma)}\right) (w_d(\sigma)+\Pi_d(\sigma))
\end{aligned}
$$

$$\Longrightarrow \underbrace{\sum_{s\in\mathcal{M}_o} c_o(s,\sigma)y_o(s,\sigma)}_{\text{(1) Supply}} = \underbrace{\sum_d \alpha_d \sum_{b\in\mathcal{M}_d}\left(\frac{1}{K_d(b)}\sum_{k\in\mathcal{K}_d(b)}\frac{\mathbf{1}\{s_d^*(b,k,\sigma)\in\mathcal{M}_o\}}{\bar{m}_d(b,k,\sigma)}\right)c_d(b,\sigma)y_d(b,\sigma)}_{\text{(2) Intermediate Input Demand}}$$

$$+\underbrace{\sum_d\sum_{i\in\mathcal{L}_d}\left(\frac{1}{N_d(i)}\sum_{n\in\mathcal{N}_d(i)}\frac{\mathbf{1}\{s_d^*(i,n,\sigma)\in\mathcal{M}_o\}}{\bar{m}_d(i,n,\sigma)}\right)(w_d(\sigma)+\Pi_d(\sigma))}_{\text{(3) Final Consumption Demand}}$$

We can simplify term (1) by making use of the labor market clearing condition as:

$$\begin{aligned}\text{Supply} &= \sum_{s\in\mathcal{M}_o} c_o(s,\sigma)y_o(s,\sigma)\\ &= \frac{w_o(\sigma)L_o}{1-\alpha_o}\end{aligned}$$

We can simplify term (2) as follows:

$$\begin{aligned}&\text{Intermediate Input Demand}\\ &=\sum_d \alpha_d \sum_{b\in\mathcal{M}_d}\left(\frac{1}{K_d(b)}\sum_{k\in\mathcal{K}_d(b)}\frac{\mathbf{1}\{s_d^*(b,k,\sigma)\in\mathcal{M}_o\}}{\bar{m}_d(b,k,\sigma)}\right)c_d(b,\sigma)y_d(b,\sigma)\\ &=\sum_d \alpha_d \frac{\overbrace{\frac{1}{M_d}\sum_{b\in\mathcal{M}_d}\left(\frac{1}{K_d(b)}\sum_{k\in\mathcal{K}_d(b)}\frac{\mathbf{1}\{s_d^*(b,k,\sigma)\in\mathcal{M}_o\}}{\bar{m}_d(b,k,\sigma)}\right)c_d(b,\sigma)y_d(b,\sigma)}^{(A)}}{\underbrace{\frac{1}{M_d}\sum_{b\in\mathcal{M}_d}c_d(b,\sigma)y_d(b,\sigma)}_{(B)}}\\ &\times \underbrace{\sum_{b\in\mathcal{M}_d}c_d(b,\sigma)y_d(b,\sigma)}_{=\frac{w_d(\sigma)L_d}{1-\alpha_d}}\end{aligned}$$

Term (A) can be simplified as follows:

$$
\begin{aligned}
(A) &= \frac{1}{M_d}\sum_{b\in\mathcal{M}_d}\left(\frac{1}{K_d(b)}\sum_{k\in\mathcal{K}_d(b)}\frac{\mathbf{1}\{s_d^*(b,k,\sigma)\in\mathcal{M}_o\}}{\bar{m}_d(b,k,\sigma)}\right)c_d(b,\sigma)y_d(b,\sigma)\\
&\xrightarrow{t\to\infty}\mathbb{E}\left[\left(\frac{1}{K_d(\cdot)}\sum_{k\in\mathcal{K}_d(\cdot)}\frac{\mathbf{1}\{s_d^*(\cdot,k,\sigma)\in\mathcal{M}_o\}}{\bar{m}_d(\cdot,k,\sigma)}\right)c_d(\cdot,\sigma)y_d(\cdot,\sigma)\right]\\
&=\mathbb{E}\left[\left(\frac{1}{K_d(\cdot)}\sum_{k\in\mathcal{K}_d(\cdot)}\frac{\mathbf{1}\{s_d^*(\cdot,k,\sigma)\in\mathcal{M}_o\}}{\bar{m}_d(\cdot,k,\sigma)}\right)\right]\mathbb{E}[c_d(\cdot,\sigma)y_d(\cdot,\sigma)]\\
&=\mathbb{E}\left[\mathbb{E}\left[\left(\frac{1}{K_d(\cdot)}\sum_{k\in\mathcal{K}_d(\cdot)}\frac{\mathbf{1}\{s_d^*(\cdot,k,\sigma)\in\mathcal{M}_o\}}{\bar{m}_d(\cdot,k,\sigma)}\right)\mid K_d\right]\right]\mathbb{E}[c_d(\cdot,\sigma)y_d(\cdot,\sigma)]\\
&=\mathbb{E}\left[\frac{1}{K_d(\cdot)}\sum_{k\in\mathcal{K}_d(\cdot)}\mathbb{E}\left[\frac{\mathbf{1}\{s_d^*(\cdot,k,\sigma)\in\mathcal{M}_o\}}{\bar{m}_d(\cdot,k,\sigma)}\mid K_d\right]\right]\mathbb{E}[c_d(\cdot,\sigma)y_d(\cdot,\sigma)]\\
&=\mathbb{E}\left[\frac{1}{K_d(\cdot)}\sum_{k\in\mathcal{K}_d(\cdot)}\mathbb{E}\left[\frac{\mathbf{1}\{s_d^*(\cdot,\cdot,\sigma)\in\mathcal{M}_o\}}{\bar{m}_d(\cdot,\cdot,\sigma)}\right]\right]\mathbb{E}[c_d(\cdot,\sigma)y_d(\cdot,\sigma)]\\
&=\mathbb{E}\left[\frac{\mathbf{1}\{s_d^*(\cdot,\cdot,\sigma)\in\mathcal{M}_o\}}{\bar{m}_d(\cdot,\cdot,\sigma)}\right]\mathbb{E}[c_d(\cdot,\sigma)y_d(\cdot,\sigma)]\\
&=\mathbb{E}\left[\frac{1}{\bar{m}_d(\cdot,\cdot,\sigma)}\right]\mathbb{E}[\mathbf{1}\{s_d^*(\cdot,\cdot,\sigma)\in\mathcal{M}_o\}]\mathbb{E}[c_d(\cdot,\sigma)y_d(\cdot,\sigma)]\\
&=\frac{\zeta}{\zeta+1}\pi_{od}(\bullet,-,\sigma_0)\mathbb{E}[c_d(\cdot,\sigma)y_d(\cdot,\sigma)]
\end{aligned}
$$

Term (B) can be simplified as follows:

$$
\begin{aligned}
(B) &= \frac{1}{M_d}\sum_{b\in\mathcal{M}_d}c_d(b,\sigma)y_d(b,\sigma)\\
&\xrightarrow{t\to\infty}\mathbb{E}[c_d(\cdot,\sigma)y_d(\cdot,\sigma)]
\end{aligned}
$$

Substituting (A) and (B) back in the Intermediate Input Demand,we obtain:

$$
\text{Intermediate Input Demand} = \sum_d \alpha_d \frac{\zeta}{\zeta+1}\pi_{od}(\bullet,-,\sigma_0)\frac{w_d(\sigma)L_d}{1-\alpha_d}
$$

We can simplify term (3) as follows:

Final Consumption Demand

$$
\begin{aligned}
&= \sum_d \sum_{i \in \mathcal{L}_d} \left(\frac{1}{N_d(i)} \sum_{n \in \mathcal{N}_d(i)} \frac{\mathbf{1}\{s_d^*(i,n,\sigma) \in \mathcal{M}_o\}}{\bar{m}_d(i,n,\sigma_1)} \right) (w_d(\sigma) + \Pi_d(\sigma)) \\
&= \sum_d \left(\frac{1}{L_d} \sum_{i \in \mathcal{L}_d} \left(\frac{1}{N_d(i)} \sum_{n \in \mathcal{N}_d(i)} \frac{\mathbf{1}\{s_d^*(i,n,\sigma) \in \mathcal{M}_o\}}{\bar{m}_d(i,n,\sigma_1)} \right) \right) (w_d(\sigma) + \Pi_d(\sigma)) L_d \\
&\xrightarrow{t \to \infty} \sum_d \mathbb{E}\left[\frac{1}{N_d(\cdot)} \sum_{n \in \mathcal{N}_d(\cdot)} \frac{\mathbf{1}\{s_d^*(\cdot,n,\sigma) \in \mathcal{M}_o\}}{\bar{m}_d(\cdot,n,\sigma)} \right] (w_d(\sigma) + \Pi_d(\sigma)) L_d \\
&= \sum_d \mathbb{E}\left[\mathbb{E}\left[\frac{1}{N_d(\cdot)} \sum_{n \in \mathcal{N}_d(\cdot)} \frac{\mathbf{1}\{s_d^*(\cdot,n,\sigma) \in \mathcal{M}_o\}}{\bar{m}_d(\cdot,n,\sigma)} \,|\, N_d \right] \right] (w_d(\sigma) + \Pi_d(\sigma)) L_d \\
&= \sum_d \mathbb{E}\left[\frac{1}{N_d(\cdot)} \sum_{n \in \mathcal{N}_d(\cdot)} \mathbb{E}\left[\frac{\mathbf{1}\{s_d^*(\cdot,n,\sigma) \in \mathcal{M}_o\}}{\bar{m}_d(\cdot,n,\sigma)} \,|\, N_d \right] \right] (w_d(\sigma) + \Pi_d(\sigma)) L_d \\
&= \sum_d \mathbb{E}\left[\frac{1}{N_d(\cdot)} \sum_{n \in \mathcal{N}_d(\cdot)} \mathbb{E}\left[\frac{\mathbf{1}\{s_d^*(\cdot,\cdot,\sigma) \in \mathcal{M}_o\}}{\bar{m}_d(\cdot,\cdot,\sigma)} \right] \right] (w_d(\sigma) + \Pi_d(\sigma)) L_d \\
&= \sum_d \mathbb{E}\left[\frac{\mathbf{1}\{s_d^*(\cdot,\cdot,\sigma) \in \mathcal{M}_o\}}{\bar{m}_d(\cdot,\cdot,\sigma)} \right] (w_d(\sigma) + \Pi_d(\sigma)) L_d \\
&= \sum_d \mathbb{E}\left[\frac{1}{\bar{m}_d(\cdot,\cdot,\sigma)} \right] \mathbb{E}[\mathbf{1}\{s_d^*(\cdot,\cdot,\sigma) \in \mathcal{M}_o\}] (w_d(\sigma) + \Pi_d(\sigma)) L_d \\
&= \sum_d \frac{\zeta}{\zeta+1} \pi_{od}(\bullet,-,\sigma_0)(w_d(\sigma) + \Pi_d(\sigma)) L_d
\end{aligned}
$$

Also, note that $\Pi_d(\sigma) L_d = \left(\frac{\zeta+1}{\zeta} - 1\right) \sum_{b \in \mathcal{M}_d} c_d(b,\sigma) y_d(b,\sigma) = \frac{1}{\zeta} \frac{w_d(\sigma) L_d}{1-\alpha_d}$. Putting these together we can further simplify the goods market clearing condition to obtain the desired result as follows:

$$
\begin{aligned}
\frac{w_o(\sigma) L_o}{1-\alpha_o} &= \frac{\zeta}{\zeta+1} \sum_d \pi_{od}(\bullet,-,\sigma_0) \left(\frac{\alpha_d}{1-\alpha_d} + 1 + \frac{1}{\zeta(1-\alpha_d)} \right) w_d(\sigma) L_d \\
&= \sum_d \pi_{od}(\bullet,-,\sigma_0) \frac{w_d(\sigma) L_d}{1-\alpha_d}
\end{aligned}
$$

$$\implies \frac{w_o(\sigma)L_o}{1-\alpha_o} = \sum_d \pi_{od}(\bullet,-,\sigma_0)\frac{w_d(\sigma)L_d}{1-\alpha_d}$$

Since $\{w_d(\sigma)\}_d$ solves the above system of equations for a given realization of σ_0, irrespective of the realization of σ_1, we conclude that $w_d(\sigma)=w_d(\sigma_0)$. That is, $\{w_d : d\in\mathcal{J}\}$ solves the following system of equations for given realization of σ_0, irrespective to realization of σ_1.

$$\frac{w_o L_o}{1-\alpha_o} = \sum_d \pi_{od}(\bullet,-)\frac{w_d L_d}{1-\alpha_d}$$

Appendix C. Empirical Framework: Appendix

C.1. **Proof of Proposition 7.** In our context, the multinomial random variable counts the number of successes in each of the M categories (one for each other supplier s), after $K_d(b)$ independent trials (one for each task associated with b). Let $\pi^0_{od}(s,b)$ denote the probability of success and $K_{od}(s,b)$ denote the number of successes in category s, the probability of observing $\{K_{od}(s,b) : s\in\mathcal{M}_o, o\in\mathcal{J}\}$ conditional on the number of tasks $K_d(b)$ is:

$$\mathbb{P}(\{K_{od}(s,b) : s\in\mathcal{M}_o, o\in\mathcal{J}\}\,|\,K_d(b)) = K_d(b)!\prod_{o\in\mathcal{J}}\prod_{s\in\mathcal{M}_o}\frac{(\pi^0_{od}(s,b))^{K_{od}(s,b)}}{K_{od}(s,b)!}$$

where $\sum_{o\in\mathcal{J}}\sum_{s\in\mathcal{M}_o}\pi^0_{od}(s,b)=1$ and $\sum_{o\in\mathcal{J}}\sum_{s\in\mathcal{M}_o}K_{od}(s,b)=K_d(b)$. From assumption 2, the unconditional probability is given by:

$$\begin{aligned}\mathbb{P}(\{K_{od}(s,b) : s\in\mathcal{M}_o, o\in\mathcal{J}\}) &= \left(K_d(b)!\prod_{o\in\mathcal{J}}\prod_{s\in\mathcal{M}_o}\frac{(\pi^0_{od}(s,b))^{K_{od}(s,b)}}{K_{od}(s,b)!}\right)\times\frac{e^{-\kappa_d}\kappa_d^{K_d(b)}}{(1-e^{-\kappa_d})K_d(b)!}\\ &= \frac{e^{-\kappa_d}}{1-e^{-\kappa_d}}\left(\prod_{o\in\mathcal{J}}\prod_{s\in\mathcal{M}_o}\frac{(\kappa_d\pi^0_{od}(s,b))^{K_{od}(s,b)}}{K_{od}(s,b)!}\right)\end{aligned}$$

The likelihood for the complete sample, $\mathbb{K}\equiv\{K_{od}(s,b) : (s,b)\in\mathcal{M}_o\times\mathcal{M}_d, (o,d)\in\mathcal{J}\times\mathcal{J}\}$ with probabilities $\mathbf{\Pi}^0\equiv\{\pi^0_{od}(s,b) : (s,b)\in\mathcal{M}_o\times\mathcal{M}_d, (o,d)\in\mathcal{J}\times\mathcal{J}\}$is:

$$\ell(\mathbb{K}\,|\,\mathbf{\Pi}^0) = \prod_{d\in\mathcal{J}}\prod_{b\in\mathcal{M}_d}\left(\frac{1}{1-e^{-\kappa_d}}\left(\prod_{o\in\mathcal{J}}\prod_{s\in\mathcal{M}_o}\frac{e^{-\kappa_d\pi^0_{od}(s,b)}(\kappa_d\pi^0_{od}(s,b))^{K_{od}(s,b)}}{K_{od}(s,b)!}\right)\right)^{\frac{1}{K_d(b)}}$$

Therefore, the log-likelihood is proportional to:[27]

[27] Note that:

$$\frac{K_{od}(s,b)}{K_d(b)} = \frac{\sum_{k\in\mathcal{K}_d(b)}\mathbf{1}\{s=s^*_d(b,k)\}}{K_d(b)}$$

$$\mathcal{L}\left(\mathbb{K}\,|\,\boldsymbol{\Pi}^0\right) \propto \sum_{o\in\mathcal{J}}\sum_{s\in\mathcal{M}_o}\left(\sum_{d\in\mathcal{J}}\sum_{b\in\mathcal{M}_d}\pi_{od}(s,b)\right)\ln\left(c_o(s)^{-\zeta}\tau_{od}^{-\zeta}\right)$$
$$-\sum_{d\in\mathcal{J}}M_d\ln\left(\sum_{s'\in\mathcal{M}}c_{o'}(s')^{-\zeta}\tau_{o'd}^{-\zeta}\right)$$
$$+\sum_{d}\left(\sum_{b\in\mathcal{M}_d}\frac{1}{K_d(b)}\right)\ln\left(\frac{e^{-\kappa_d}}{1-e^{-\kappa_d}}\right)+\sum_{d}M_d\ln\kappa_d$$

Under Assumption 6, note that $c_o(s)=\widetilde{c}_o(s)c_o$ and $\sum_{s'\in\mathcal{M}}c_{o'}(s')^{-\zeta}\tau_{o'd}^{-\zeta}=\sum_{o'}c_{o'}^{-\zeta}\tau_{o'd}^{-\zeta}$, therefore the likelihood equations for $\widetilde{c}_o(s)$ are given by:

$$\frac{\sum_d\sum_{b\in\mathcal{M}_d}\pi_{od}(s,b)}{\widetilde{c}_o(s)^{-\zeta}c_o^{-\zeta}}=\sum_d\frac{M_d}{\sum_{o'}c_{o'}^{-\zeta}\tau_{o'd}^{-\zeta}}\tau_{od}^{-\zeta}$$

The likelihood equations for $\tau_{od}^{-\zeta}$ are given by:

$$\begin{aligned}
\frac{\left(\sum_{b\in\mathcal{M}_d}\sum_{s\in\mathcal{M}_o}\pi_{od}(s,b)\right)}{\tau_{od}^{-\zeta}} &= \frac{M_d}{\sum_{o'}c_{o'}^{-\zeta}\tau_{o'd}^{-\zeta}}\left(\sum_{s\in o}c_o(s)^{-\zeta}\right)\\
&= \frac{M_d}{\sum_{o'}c_{o'}^{-\zeta}\tau_{o'd}^{-\zeta}}c_o^{-\zeta}\\
\implies \tau_{od}^{-\zeta} &= \frac{\left(\sum_{b\in\mathcal{M}_d}\sum_{s\in\mathcal{M}_o}\pi_{od}(s,b)\right)}{\frac{M_d}{\sum_{o'}c_{o'}^{-\zeta}\tau_{o'd}^{-\zeta}}c_o^{-\zeta}}\\
&= \frac{\left(\sum_{b\in\mathcal{M}_d}\pi_{od}(\bullet,b)\right)}{\frac{M_d}{\sum_{o'}c_{o'}^{-\zeta}\tau_{o'd}^{-\zeta}}c_o^{-\zeta}}
\end{aligned}$$

$$\begin{aligned}
&=\sum_{k\in\mathcal{K}_d(b)}\mathbf{1}\{s=s_d^*(b,k)\}\left(\frac{1}{K_d(b)}\right)\\
&=\sum_{k\in\mathcal{K}_d(b)}\mathbf{1}\{s=s_d^*(b,k)\}\left(\frac{X_d(b,k)}{X_d(b)}\right)\\
&=\left(\frac{\sum_{k\in\mathcal{K}_d(b)}\mathbf{1}\{s=s_d^*(b,k)\}X_d(b,k)}{X_d(b)}\right)\\
&=\frac{X_{od}(s,b)}{X_d(b)}\\
&=\pi_{od}(s,b)
\end{aligned}$$

Substituting the expression for $\tau_{od}^{-\zeta}$, we obtain an estimator for $\widetilde{c}_o(s)^{-\zeta}$ as:

$$\begin{aligned}\widetilde{c}_o(s)^{-\zeta} &= \frac{\sum_{b\in\mathcal{M}}\pi_{od}(s,b)}{\sum_{b\in\mathcal{M}}\pi_{od}(\bullet,b)} \\ &= \frac{\sum_d \pi_{od}(s,\bullet)}{\sum_{s'\in\mathcal{M}_o}\sum_d \pi_{od}(s',\bullet)}\end{aligned}$$

This then provides us with an estimator for $\frac{c_o^{-\zeta}\tau_{od}^{-\zeta}}{\sum_{o'} c_{o'}^{-\zeta}\tau_{o'd}^{-\zeta}}$ as follows:

$$\begin{aligned}\tau_{od}^{-\zeta} &= \frac{\left(\sum_{b\in\mathcal{M}_d}\sum_{s\in\mathcal{M}_o}\pi_{od}(s,b)\right)}{\frac{M_d}{\sum_{o'} c_{o'}^{-\zeta}\tau_{o'd}^{-\zeta}} c_o^{-\zeta}} \\ \implies \frac{c_o^{-\zeta}\tau_{od}^{-\zeta}}{\sum_{o'} c_{o'}^{-\zeta}\tau_{o'd}^{-\zeta}} &= \frac{\sum_{b\in\mathcal{M}_d}\sum_{s\in\mathcal{M}_o}\pi_{od}(s,b)}{M_d} \\ &= \frac{1}{M_d}\sum_{b\in\mathcal{M}_d}\pi_{od}(\bullet,b)\end{aligned}$$

C.2. **Estimation of Structural Elasticities, Θ.**

C.2.1. *Trade Elasticity* ζ. Since the model satisfies structural gravity at the aggregate level (see Equation (3.13)) and the dispersion of match-specific productivities ζ coincides with the elasticity of trade with respect to trade costs, I calibrate the value of this parameter to 5 from median of the estimates of price elasticity in structural gravity equations (see Head and Mayer (2014)).

C.2.2. *Materials Share* $\boldsymbol{\alpha}$. The distribution of markups from Proposition 2 provides expressions for value-added share of gross output $(VA/GO)_o$. Using equation (3.8), materials share α_o is calibrated as $\alpha_o = (1+1/\zeta)(1-(VA/GO)_o)$, where $(VA/GO)_o$ across districts are constructed using aggregate production statistics as follows.

I obtain district-level sectoral GDP $\{VA_o^j\}$ from Nielsen Analytics, a private data firm and industry-level data on value-added share of gross output at the national level, $\left\{(VA/GO)^j : j\in\mathcal{I}\right\}$ from the World Input-Output Database. Using these, I construct a measure of value-added share of gross output at the district level as

$$(VA/GO)_o = \frac{\sum_{j\in\mathcal{I}} VA_o^j}{\sum_{j\in\mathcal{I}} \frac{VA_o^j}{(VA/GO)^j}}. \tag{C.1}$$

I use data pertaining to six industry groups for this calculation. They are (a) Mining and Quarrying; (b) Construction; (c) Manufacturing; (d) Electricity, Gas and Water

Supply; (e) Transport, Storage and Communication; and (f) Trade, Hotels and Restaurants.

C.2.3. *Task Intensities* $\boldsymbol{\kappa}$. Estimation of task intensities $\boldsymbol{\kappa}$ relies on equation (3.12). The expression for $\rho^0_{od}(s,-)$ shows how κ_d regulates the number of customers a firm s finds at location d given its conditional choice probabilities $\pi^0_{od}(s,-)$. Since both $\rho^0_{od}(s,-)$ and $\pi^0_{od}(s,-)$ are independent of the identity of buyer at d, the empirical counterpart of $\rho^0_{od}(s,-)$ is the fraction of firms at d that buy goods from s. The theoretical value of $\rho^0_{od}(s,-)$ in terms of κ_d is computed using estimated conditional choice probabilities $\pi^*_{od}(s,-)$. Estimates of task intensities $\boldsymbol{\kappa}$ are obtained by minimizing the distance between theoretical and empirical values of $\rho^0_{od}(s,-)$.

C.2.4. *Productivity Dispersion* $\boldsymbol{\theta}$. The theoretical expression for the squared coefficient of variation of $c_o(s)^{-\zeta}$ which are estimated as firm fixed effects in equation (4.2) can be derived to be:

$$CV_o\left(c_o(s)^{-\zeta/2}\right)^2 = \frac{\Gamma\left(1-\frac{\zeta}{\theta_o}\right)}{\Gamma\left(1-\frac{\zeta}{2\theta_o}\right)^2} \cdot \frac{\mathbb{E}_{\{K_o\}}\left[\Gamma\left(2-\frac{\alpha_o}{K_o}\right)^{K_o}\right]}{\mathbb{E}_{\{K_o\}}\left[\Gamma\left(2-\frac{\alpha_o}{2K_o}\right)^{K_o}\right]^2} - 1$$

For any value of $\boldsymbol{\theta}$, the theoretical value is evaluated using estimates of task intensities $\boldsymbol{\kappa}$ and materials shares $\boldsymbol{\alpha}$, as per the above expression. For each district, the empirical value of the squared coefficient of variation is obtained using the estimator proposed in Breunig (2001) from estimates $\left\{\left(c_o(s)^{-\zeta}\right)^* : s \in \mathcal{M}\right\}$. Estimates of $\boldsymbol{\theta}$ are obtained by minimizing the distance between the theoretical and the empirical values of squared coefficient of variation.

C.3. **Expected Utility & Welfare Changes.** Households residing at location d are heterogeneous both in their numbers of needs and match-specific taste shocks of using different suppliers' goods to fulfill their needs. Welfare at any location is then calculated in expectation. That is, $V_d = \mathbb{E}[V_d(\cdot)]$. With Cobb-Douglas utilities across needs, indirect utility of household i residing at d is given by:

$$V_d(i) = \frac{w_d\left(1 + 1/\zeta(1-\alpha_o)\right)}{\prod_{n=1}^{N_d(i)} p_d(i,n)^{1/N_d(i)}}$$

Expected indirect utility of households at location d can then be derived as:

$$V_d = \mathbb{E}[V_d(\cdot)]$$

$$
\begin{aligned}
&= \mathbb{E}\left[w_d (1 + {}^1\!/\zeta(1-\alpha_o)) \prod_{n=1}^{N_d(\cdot)} p_d(\cdot, n)^{-1/N_d(\cdot)} \right] \\
&= w_d (1 + {}^1\!/\zeta(1-\alpha_o)) \mathbb{E}\left[\mathbb{E}\left[\prod_{n=1}^{N_d(\cdot)} p_d(\cdot, n)^{-1/N_d(\cdot)} \mid N_d \right] \right] \\
&= w_d (1 + {}^1\!/\zeta(1-\alpha_o)) \mathbb{E}\left[\prod_{n=1}^{N_d(\cdot)} \mathbb{E}\left[p_d(\cdot, \cdot)^{-1/N_d(\cdot)} \mid N_d \right] \right] \\
&= w_d (1 + {}^1\!/\zeta(1-\alpha_o)) \mathbb{E}\left[\prod_{n=1}^{N_d(\cdot)} \Gamma\left(2 - \frac{1}{\zeta N_d(\cdot)} \right) A_d^{\frac{1}{\zeta N_d(\cdot)}} \right] \\
&= (1 + {}^1\!/\zeta(1-\alpha_o)) \mathbb{E}_{\{N_d\}}\left[\Gamma\left(2 - \frac{1}{\zeta N_d} \right)^{N_d} \right] w_d A_d^{\frac{1}{\zeta}}
\end{aligned}
$$

Welfare changes, i.e., changes in expected indirect utility at location d in response to shocks can be calculated as:

$$
\widehat{V}_d = \widehat{w}_d \widehat{A}_d^{1/\zeta},
$$

where $\widehat{w}_d$ denotes the change in wage and $\widehat{A}_d$ denotes change in market access at d.

C.4. **Procedure for Computing Counterfactual Outcomes.** Counterfactual analysis is conducted in three steps. First, I evaluate the expected value of aggregate and firm-level outcomes in the initial state. Second, I compute changes in aggregate outcomes that result from the counterfactual shock. Finally, I evaluate the expected value of aggregate and firm-level outcomes in the counterfactual state

Step 1: Compute expected value of aggregate and firm-level outcomes in initial state. In the initial state, $\boldsymbol{wL} \equiv \{w_d L_d : d \in \mathcal{J}\}$ is obtained as the solution to the following system of equations:

$$
\frac{w_d L_d}{1 - \alpha_d} = \sum_d \pi^*_{od}(\bullet, -) \frac{w_o L_o}{1 - \alpha_o},
$$

where $\pi^*_{od}(\bullet, -)$ is calculated as in equation (4.7). Using the solution to these equations, value-added and gross output for each district are respectively calculated as:

$$
VA_d = w_d L_d \left(\frac{({}^{VA}\!/_{GO})_d}{({}^{VA}\!/_{GO})_d - {}^1\!/\zeta + 1} \right),
$$

$$GO_d = w_d L_d \left(\frac{1}{\left({}^{VA}\!/\!{}_{GO}\right)_d - {}^{1}\!/\!{}_{\zeta+1}} \right),$$

where $\left({}^{VA}\!/\!{}_{GO}\right)_d$ for district d is calculated in equation (C.1). Total value-added across all districts is chosen as the numeraire, i.e., $\sum_d VA_d = 1$. At the firm-level, input sales, total sales, intensity of use, and average customer size are respectively calculated as:

$$\text{input sales}_o(s) = \sum_d \pi^*_{od}(s,-)(GO_d - VA_d),$$

$$\text{total sales}_o(s) = \sum_d \pi^*_{od}(s,-)GO_d,$$

$$\text{intensity of use}_o(s) = \sum_d \pi^*_{od}(s,-)M_d,$$

$$\text{average customer size}_o(s) = \frac{\text{input sales}_o(s)}{\text{intensity of use}_o(s)},$$

where $\pi^*_{od}(s,-)$ is calculated as in equation (4.6).

Step 2: Evaluate change in aggregate outcomes from initial to counterfactual state. For any change in σ_0, $\widehat{\boldsymbol{\delta}} \equiv \left\{ \widehat{\delta}_{od} : (o,d) \in \mathcal{J} \times \mathcal{J} \right\}$, one can solve for change in wages $\widehat{\boldsymbol{w}} \equiv \{\widehat{w}_d : d \in \mathcal{J}\}$ with the following tâtonnement algorithm for some positive constant μ and tolerance value tol:

(1) Start with a guess for the vector of change in wages, $\widehat{\boldsymbol{w}}^{(0)}$

(2) For the vector of wage changes, in the t^{th} iteration $\widehat{\boldsymbol{w}}^{(t)}$, compute change in market access as the solution to the following system of equations:

$$\widehat{A}_d^{(t)} = \sum_o \pi_{od}(\bullet,-) \widehat{\delta}_{od} \left(\widehat{w}_o^{(t)}\right)^{-\zeta(1-\alpha_o)} \left(\widehat{A}_o^{(t)}\right)^{\alpha_o}$$

(3) Compute counterfactual sourcing probabilities as:

$$\left(\pi_{od}^{(t)}(\bullet,-)\right)' = \pi_{od}(\bullet,-) \frac{\widehat{\delta}_{od} \left(\widehat{w}_o^{(t)}\right)^{-\zeta(1-\alpha_o)} \left(\widehat{A}_o^{(t)}\right)^{\alpha_o}}{\widehat{A}_d^{(t)}}$$

(4) Compute excess demand for labor $\boldsymbol{Z}\left(\widehat{\boldsymbol{w}}^{(t)}\right) \equiv \left\{ Z_o\left(\widehat{\boldsymbol{w}}^{(t)}\right) : o \in \mathcal{J} \right\}$ as:

$$Z_o\left(\widehat{\boldsymbol{w}}^{(t)}\right) = \frac{1-\alpha_o}{w_o L_o} \sum_d \left(\pi_{od}^{(t)}(\bullet,-)\right)' \widehat{w}_d^{(t)} \frac{w_d L_d}{1-\alpha_d} - \widehat{w}_o$$

(5) Update the vector of change in wages as $\widehat{\boldsymbol{w}}^{(t+1)} \leftarrow \widehat{\boldsymbol{w}}^{(t)} + \mu \boldsymbol{Z}\left(\widehat{\boldsymbol{w}}^{(t)}\right)$.

(6) If $\|\widehat{\boldsymbol{w}}^{(t+1)} - \widehat{\boldsymbol{w}}^{(t)}\| > tol$, go back to (2), else end.

Welfare changes can then be computed as $\widehat{V}_d = \widehat{w}_d^{(\infty)} \left(\widehat{A}_d^{(\infty)}\right)^{\frac{1}{\zeta}}$.
Step 3: Compute expected value of aggregate and firm-level outcomes in counterfactual state. As in the initial state, here again VA'_d and GO'_d are computed for each district using $(\boldsymbol{wL})'$ instead of $\boldsymbol{wL}$.

$$VA'_d = \widehat{w}_d^{(\infty)} w_d L_d \left(\frac{(VA/GO)_d}{(VA/GO)_d - 1/\zeta+1} \right),$$

$$GO'_d = \widehat{w}_d^{(\infty)} w_d L_d \left(\frac{1}{(VA/GO)_d - 1/\zeta+1} \right).$$

Firm-level outcomes are then calculated by using $\pi_{od}^{(\infty)}(\bullet,-)$ instead of $\pi_{od}^{*}(\bullet,-)$ as follows:

$$(\text{input sales}_o(s))' = \sum_d \pi_{od}^{(\infty)}(s,-)(GO'_d - VA'_d),$$

$$(\text{total sales}_o(s))' = \sum_d \pi_{od}^{(\infty)}(s,-)GO'_d,$$

$$(\text{intensity of use}_o(s))' = \left(\sum_d \pi_{od}^{(\infty)}(s,-)M_d \right),$$

$$(\text{average customer size}_o(s))' = \frac{(\text{input sales}_o(s))'}{(\text{intensity of use}_o(s))'},$$

where $\pi_{od}^{(\infty)}(s,-) = \frac{\left(\tilde{c}_o(s)^{-\zeta}\right)^* \widehat{\delta}_{od}(s)}{\sum_{s' \in \mathcal{M}_o} \left(\tilde{c}_o(s')^{-\zeta}\right)^* \widehat{\delta}_{od}(s')} \pi_{od}^{(\infty)}(\bullet,-)$ and $\widehat{\delta}_{od}(s)$ is the firm-level shock from the change in σ_0.

C.5. **Estimation Results.** This section first goes over estimates of trade frictions and conditional choice probabilities, and then model predictions for firms' sales and intensity of use.[28] Then, I evaluate the model by seeing how well it replicates empirical regularities documented in Section 2.

C.5.1. *Estimates.* Trade frictions are estimated using gravity regressions. Table C.1 reports estimated coefficients for distance and border dummies in column (4) and compares them to common methods in the trade literature in columns (1)-(3). Column (1) is an atheoretical regression specification that is not appealing when there are zeros in trade data and hence not comparable to other columns. Column (2) is still an atheoretical specification but is consistent with handling zeros in the data. Column (3) is a model-based specification and accommodates zeros in the data. Column (4)

[28]Estimates of elasticities contained in Θ are relegated to Appendix C.5.

is the specification that is implied by the model here. Comparing (2) or (3) to (4) shows that using aggregate trade flows or shares underestimates trade frictions for estimation of the model here. With estimated trade frictions in hand, estimates of sourcing probabilities, denoted by $\pi^*_{od}(\bullet,-)$, and firms' conditional choice probabilities across destinations, denoted by $\pi^*_{od}(s,-)$, are obtained using equations (4.7) and (4.6). I solve for wages that satisfy the trade equilibrium in the limiting economy using equation (3.16) with the estimated sourcing probabilities. The expected value of firms' destination-specific intensity of use and sales are respectively calculated, using wages and estimated conditional choice probabilities, as:

$$\text{intensity of use}_{od}(s) = \pi^*_{od}(s,-) M_d, \tag{C.2}$$

$$\text{input sales}_{od}(s) = \pi^*_{od}(s,-)\left(\frac{\alpha_d}{1-\alpha_d+{}^1\!/\!_\zeta}\right) w_d L_d. \tag{C.3}$$

Intensity of use and sales of any given firm are then computed by summing over the above values across all destinations. For aggregate trade flows between an origin-destination pair, corresponding values are obtained by summing over all firms at the origin.

C.5.2. *Model Fit.* A key finding in Proposition 7 is that the fixed effect estimate for a firm s with the multinomial likelihood specification is in fact its measured intensity of use, $\sum_d \pi_{od}(s,\bullet)$.[29] According to the model (in equation (3.11)), this fixed effect is related marginal costs as $c_o(s)^{-\zeta}$. This directly features in equation (4.6) and plays a vital role in enabling the model to reproduce the empirical regularities. Apart from this, goodness of fit is governed by four factors. First, imperfect correlation between data and fitted values in Table C.1, Column (4) causes differences in $\pi_{od}(\bullet,-)$ and $\pi^*_{od}(\bullet,-)$. Second, estimating equation (4.1) is parsimoniously specified as it does not allow heterogeneity in trade frictions faced by firms. While the data is at the firm-to-firm level, fixed effects are only at the firm and origin–destination level. Third, equilibrium wages computed for the limiting economy differ from data. These differences capture the granularity of data which are assumed away in the limiting economy. Finally, estimates of material share of costs $\boldsymbol{\alpha}$ and dispersion in match-specific productivities ζ also affect predicted values calculated via equation (C.3).

[29]Fixed effect for firm s is the product of the within location component $\widetilde{c}_o(s)^{-\zeta}$ and the between location component $c_o^{-\zeta}$. Equation (4.3) provides a estimator for the former. The latter is estimated in column (4) in Table C.1 using a multinomial likelihood specification. By properties of the multinomial likelihood, this estimate is given by $\sum_d \pi_{od}(\bullet,\bullet)$. Together, they imply that the fixed effect estimate for firm s can be expressed as $\left(c_o(s)^{-\zeta}\right)^* = \sum_{b\in\mathcal{M}} \pi_{od}(s,b)$.

Estimates of intensity of use are only affected by the first two factors whereas those of sales are affected by all of them. Columns (1) and (3) in Table C.2 report the coefficient of determination of log-log regressions where observed values are projected on predicted values of intensity of use and sales. Average customer size is omitted from this table because it is obtained as the ratio of sales and intensity of use and so it is not meaningful to measure its goodness of fit. Columns (2) and (4) in Table C.2 report similar results but using average trade shares observed in data $\pi_{od}(\bullet,-)$ instead of the corresponding fitted values $\pi^*_{od}(\bullet,-)$ for sourcing probabilities. These columns help assess the loss of fit arising from gravity regressions. These results suggest that (a) fits for sales are worse than intensity of use due to the third and fourth factors, (b) fits for firms' destination-specific sales are the worse than firms' overall sales due to the second factor, (c) fit of gravity regressions causes substantial loss of fit only for aggregate trade flows due to the first factor.

Table C.3 reports how the estimated model performs in comparison to the empirical regularities documented in Section 2. I focus only on the upstream (intensity of use) and downstream (average customer size) margins and not the three-way decomposition in Section 2. This is because the model does not meaningfully differentiate between the first and second factors in expectation and further, it is the joint contribution of both these factors that plays a role in endogenous network formation. Table C.3 shows that the intensity of use margin explains a vast majority of the variation in firms' sales in the estimated model as is the case in the data. This is true across all columns in the data qualitatively. Quantitatively, all columns except (3) provide a reasonably good fit. In column (3), the loss of fit can be attributed to the second factor.

Table C.1. **Gravity Regressions**

	$\sinh^{-1}(\text{sales}_{od})$ (1): OLS	sales_{od} (2): PPML	$\frac{\text{sales}_{od}}{\sum_{o'}\text{sales}_{o'd}}$ (3): MPML	$\left(\frac{c_o^{-\zeta}\tau_{od}^{-\zeta}}{\sum_{o'}c_{o'}^{-\zeta}\tau_{o'd}^{-\zeta}}\right)^*$ (4): MPML
log(distance)	-2.947***	-0.219***	-0.712***	-0.990***
	(0.039)	(0.042)	(0.045)	(0.044)
$\mathbf{1}$\{inter-state\}	-5.032***	-1.971***	-2.125***	-2.579***
	(0.069)	(0.104)	(0.090)	(0.089)
$\mathbf{1}$\{inter-district\}	0.086	-1.484***	-1.852***	-2.262***
	(0.215)	(0.117)	(0.078)	(0.067)
$\mathbf{1}$\{neighbor\}	-1.121***	0.562***	0.251***	0.516***
	(0.113)	(0.053)	(0.052)	(0.047)
Fixed Effects:				
Origin × Year	✓	✓	✓	✓
Destination × Year	✓	✓	✓	✓
Adjusted R^2	0.669	—	—	—
Pseudo R^2	—	0.945	0.435	0.488
Squared Correlation	0.674	0.953	0.793	0.898
# observations	$141^2 \times 5$	$141^2 \times 5$	$141^2 \times 5$	$141^2 \times 5$

Note. Standard errors in parentheses, two-way clustered by origin–year and destination–year. $^{*}p < 0.05, ^{**}p < 0.01, ^{***}p < 0.001$. Observations pertain to all bilateral pairs between 141 districts for 5 years. The distance between district pairs is calculated as the distance between their centroids. A district's distance to itself is calculated as the radius of the circle with the same area as the district. Column (1) is estimated using an OLS specification with the inverse hyperbolic sine of trade flows as dependent variable. Column (2) is estimated using a Poisson PML specification with aggregate trade flows as the dependent variable as in Santos Silva and Tenreyro (2006). Column (3) is estimated using a multinomial PML specification with aggregate trade shares as the dependent variable as in Eaton et al. (2013). Column (4) is estimated using a multinomial PML specification from equation (4.5). Two-way clustering is done as in Cameron et al. (2011). Pseudo R^2 is calculated as in McFadden (1974).

Table C.2. **Goodness of Fit: Firms' Intensity of Use and Sales**

	Intensity of Use		Overall	
	$\pi^*_{od}(\bullet,-)$	$\pi_{od}(\bullet,-)$	$\pi^*_{od}(\bullet,-)$	$\pi_{od}(\bullet,-)$
	(1)	(2)	(3)	(4)
Sales	0.996	1.000	0.428	0.467
Destination-Specific Sales	0.324	0.381	0.176	0.186
Trade Flows	0.512	0.999	0.503	0.709

Note. This table reports R^2 of log-log regressions when predicted values for intensity of use and sales are projected on observed data at three levels of aggregation: firms' destination-specific sales, firms' sales, and aggregate trade flows. Columns (1) and (3) use estimated average trade shares from equation (4.7) while (2) and (4) use exact average trade share from equation (4.4) for the calculations.

Table C.3. **Model Fit: Margins of Firms' Sales**

	Sales		Destination-Specific Sales		Trade Flows	
	(1)	(2)	(3)	(4)	(5)	(6)
Data:						
Intensity of Use	81%	82%	94%	80%	87%	94%
Average Customer Size	19%	18%	6%	20%	13%	6%
Model:						
Intensity of Use	73%	100%	61%	76%	71%	100%
Average Customer Size	27%	0%	39%	24%	29%	0%
Fixed Effects:						
Seller×Year	—	—	✓	—	—	—
Origin×Year	—	✓	—	—	—	✓
Destination×Year	—	—	—	—	—	✓
Data Level:						
Seller×Year	•	•	—	—	—	—
Seller×Destination×Year	—	—	•	•	—	—
Origin×Destination×Year	—	—	—	—	•	•
# observations	5.6×10^6	5.6×10^6	18.2×10^6	18.2×10^6	58,390	58,390

Note. Columns (1) and (2) report the contribution of factors: intensity of use and average customer size, to the variance of firms' sales (as per equation (2.1)) in the data (top panel) and in the model (bottom panel). Columns (3) and (4) report the contribution of those factors to the variance of firms' destination-specific sales (as per equation (A.1)). Columns (5) and (6) report the same for trade flows between districts (as per equation (A.2)).

Supplementary Material

Appendix D. Data & Empirical Regularities: Supplementary Material

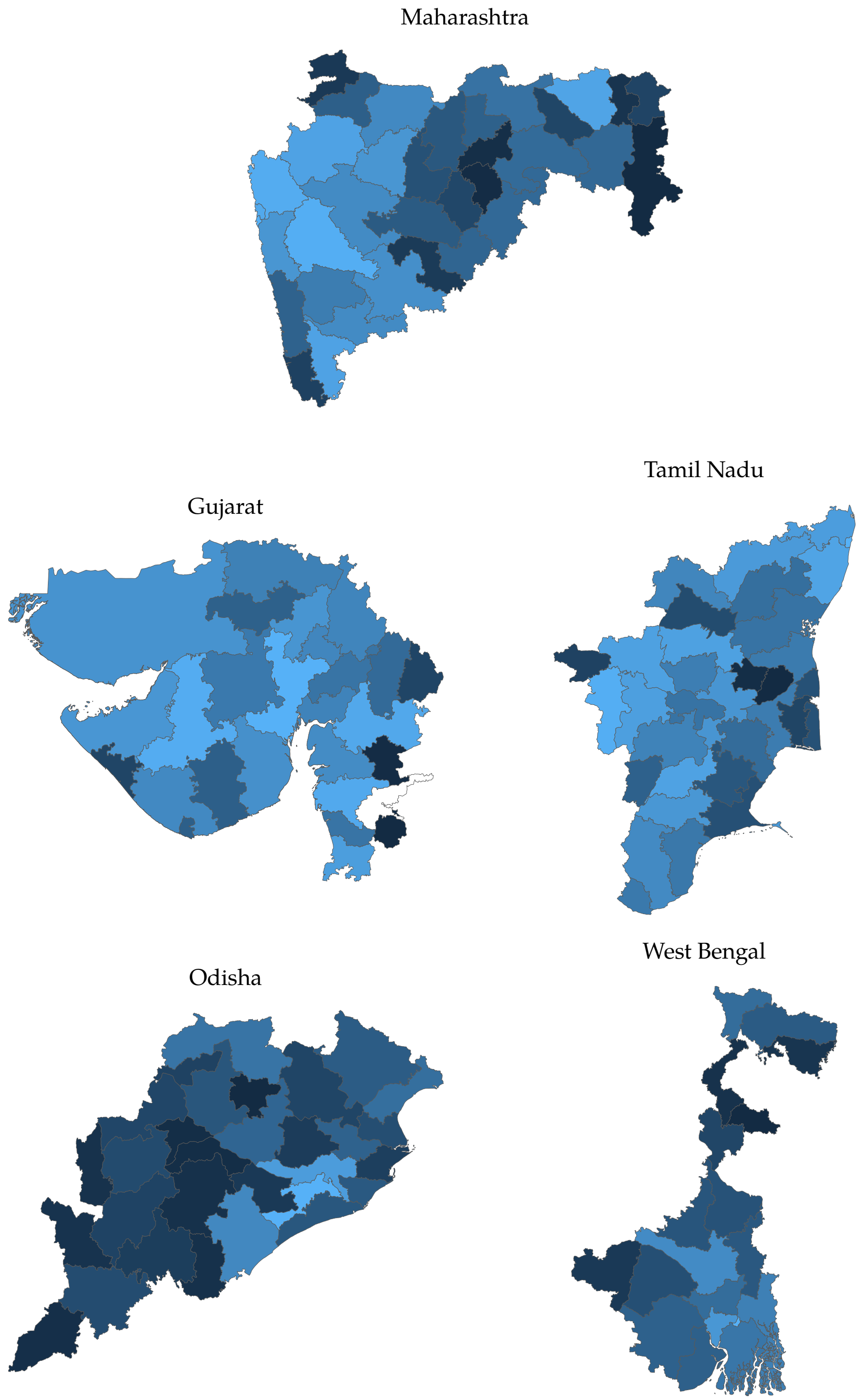
D.1. **Summary Statistics.** Table D.1 reports count statistics of firms and relationships between them each year accompanied by their breakdown into different categories. Figure D.1 plots the spatial distribution of firm–year pairs across districts on their respective state maps. Figure D.2 plots the distribution of firm-to-firm relationships across district pairs. Table D.2 reports distributions of firms' sales to other firms, # customers, and sales per customer. Table D.3 reports distributions of firms' purchases from other firms, # suppliers, and purchases per supplier.

D.2. **Margins of Firms' Sales.** The joint distribution of firms' sales with intensity of use and average customer size is depicted in Figure D.3. Figure D.4 provides the results of decomposition of firms' sales by district. Figure D.5 provides the results of decomposition of firms' sales by percentile bins.

D.3. **Margins of Intranational Trade.** The overall level of trade integration between districts as measured by the Head and Ries (2001) index is depicted in Figure D.6.[30] Figure D.7 depicts the breakdown of trade flows across district pairs into the upstream and downstream margins.

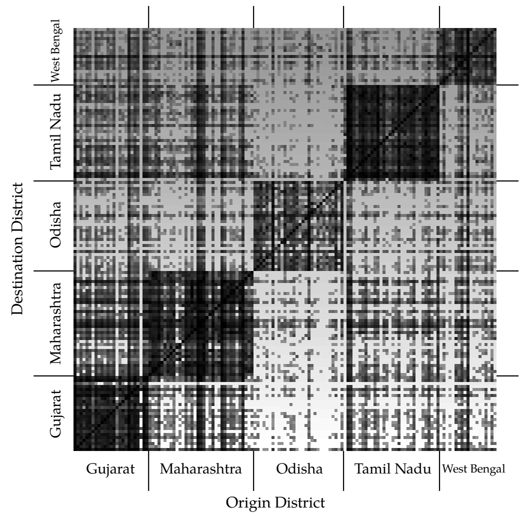
[30] For any pair of districts (o,d), the Head and Ries (2001) index is computed as the ratio $\sqrt{\frac{\text{sales}_{od} \times \text{sales}_{do}}{\text{sales}_{oo} \times \text{sales}_{dd}}}$.

Figure D.1. **Distribution of Firms across Districts**



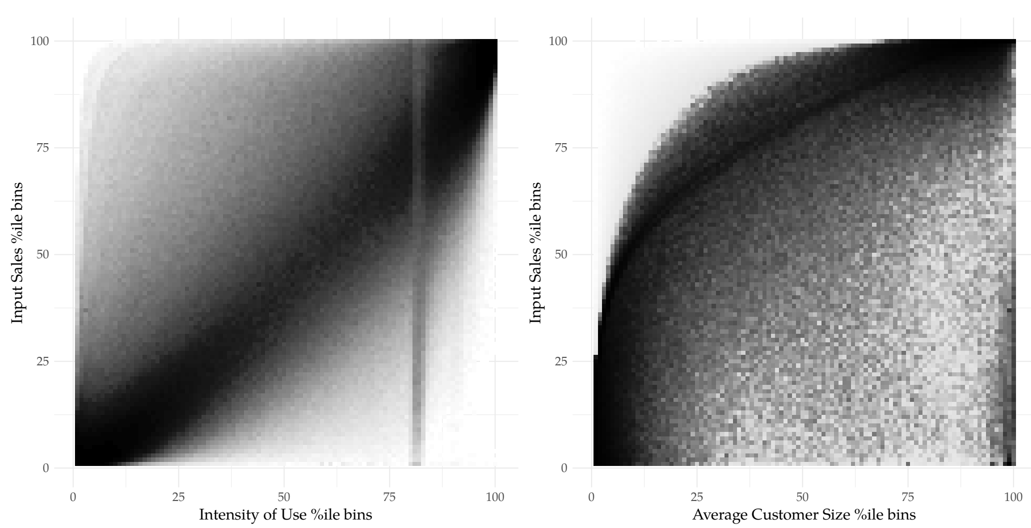
Note. Districts are shaded by # firm–year observations (from columns (1–5) in Table D.1, middle panel). Darker shades reflect lower values. Relative areal extent of states is not up to scale.

FIGURE D.2. **Distribution of Firm-to-Firm Relationships across District Pairs**



Note. This figure depicts the 141×141 matrix of # relationships between district pairs in 2015-2016 (from column (5) in Table D.1, bottom panel). Darker cells reflect higher values. Districts are arranged first by state and then alphabetically within states on both axes.

FIGURE D.3. **Margins of Firms' Sales: Joint Distribution with Sales**



Note. In this figure, firms are classified into 100×100 bins based on their total sales in input markets and intensity of use (left panel) or average customer size (right panel). This is a two-dimensional histogram where each cell in this 100×100 matrix is shaded as per the quantile of the count of firms in the bin such that darker shades correspond to higher quantiles.

Table D.1. **Summary Statistics: Firms and their Relationships**

	2011-2012 (1)	2012-2013 (2)	2013-2014 (3)	2014-2015 (4)	2015-2016 (5)	All (6)
# **Firms**	1,616	1,743	1,899	2,040	2,107	2,572
Neither	23%	25%	24%	24%	24%	18%
Buy	16%	16%	16%	18%	18%	15%
Sell	19%	17%	17%	15%	14%	17%
Both	42%	42%	43%	43%	43%	50%
Gujarat	21%	21%	22%	24%	24%	25%
Maharashtra	42%	41%	40%	38%	36%	34%
Odisha	5%	5%	5%	6%	6%	6%
Tamil Nadu	24%	25%	25%	25%	25%	27%
West Bengal	8%	8%	8%	8%	9%	8%
# **Relationships**	17,681	18,547	21,031	22,600	23,786	103,646
Intra-District	58%	57%	57%	58%	57%	58%
Inter-District	37%	38%	38%	38%	39%	38%
Inter-State	5%	5%	5%	4%	4%	4%

Note. Figures for # firms and # relationships are in units of thousands. Columns (1–5) report values for each year and column (6) the total across all years. The top panel breaks down # firms by their participation in the network, i.e., whether they buy from and sell to other firms. The middle panel breaks down firms by the state in which they are located. The bottom panel reports the breakdown of # relationships based on whether the customer and supplier are located in the same district, different districts or different states altogether. For the top and middle panels, column (6) reports statistics pertaining to unique firms across all years. For the bottom panel, column (6) reports the sum of columns (1–5).

Table D.2. **Summary Statistics: Firms' Customers and Sales**

Percentile:	90 (1)	75 (2)	50 (3)	25 (4)	10 (5)
# **Customers**:					
2011-2012	39	14	4	2	1
2012-2013	39	14	4	2	1
2013-2014	41	15	5	2	1
2014-2015	42	15	5	2	1
2015-2016	44	15	5	2	1
Sales per Customer:					
2011-2012	1,850	536	153	44	13
2012-2013	1,993	570	162	47	14
2013-2014	2,027	577	165	49	14
2014-2015	2,160	609	171	50	14
2015-2016	2,121	610	174	51	14
Sales to other Firms:					
2011-2012	18,620	4,606	987	158	25
2012-2013	19,702	4,882	1,037	167	26
2013-2014	21,225	5,199	1,095	175	27
2014-2015	22,531	5,517	1,143	177	27
2015-2016	22,936	5,690	1,184	183	27

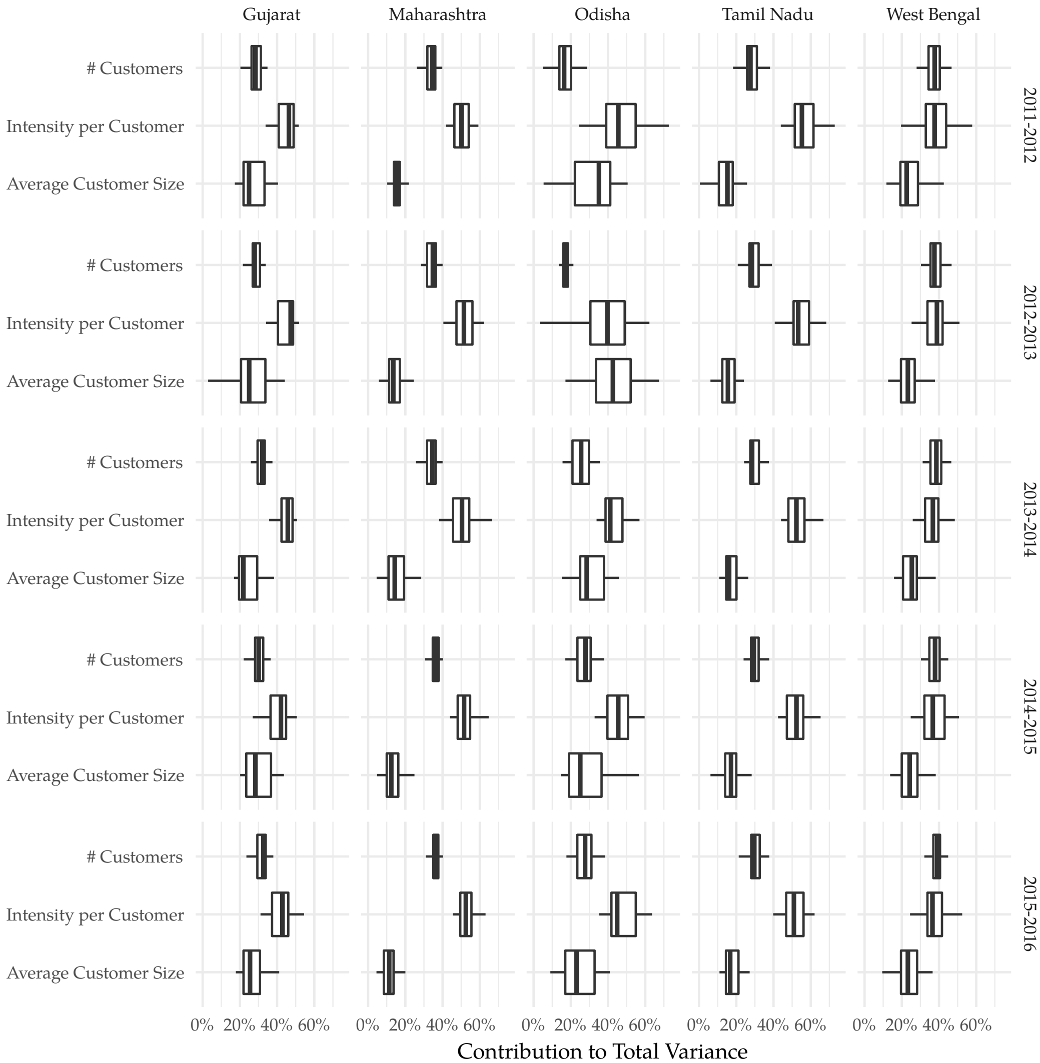
Note. For each year and for firms that sell to other firms, the top panel reports the median, top and bottom deciles, and upper and lower quartiles of # customers. The middle panel reports the same for sales per customer and the bottom panel for sales to other firms. All sales figures are reported in units of 1000 INR.

TABLE D.3. **Summary Statistics: Firms' Suppliers and Purchases**

Percentile:	90 (1)	75 (2)	50 (3)	25 (4)	10 (5)
# **Suppliers**:					
2011-2012	46	21	8	3	1
2012-2013	44	20	8	3	1
2013-2014	45	21	8	3	1
2014-2015	43	20	8	3	1
2015-2016	44	21	8	3	1
Purchases per Supplier:					
2011-2012	2,628	752	228	77	30
2012-2013	2,871	811	243	81	32
2013-2014	2,978	841	251	84	33
2014-2015	2,983	845	249	83	33
2015-2016	2,941	835	249	84	33
Purchases from other Firms:					
2011-2012	28,005	7,822	2,182	575	150
2012-2013	29,137	8,067	2,252	589	152
2013-2014	31,825	8,826	2,410	630	163
2014-2015	30,445	8,311	2,318	614	159
2015-2016	31,136	8,506	2,403	633	164

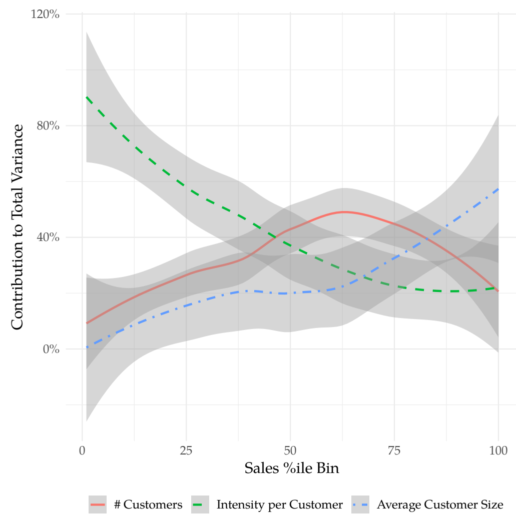
Note. For each year and for firms that purchase from other firms, the top panel reports the median, top and bottom deciles, and upper and lower quartiles of # suppliers. The middle panel reports the same for purchases per supplier and the bottom panel for purchases from other firms. All purchases figures are reported in units of 1000 INR.

FIGURE D.4. **Margins of Firms' Sales: Contribution to Variance, by District**



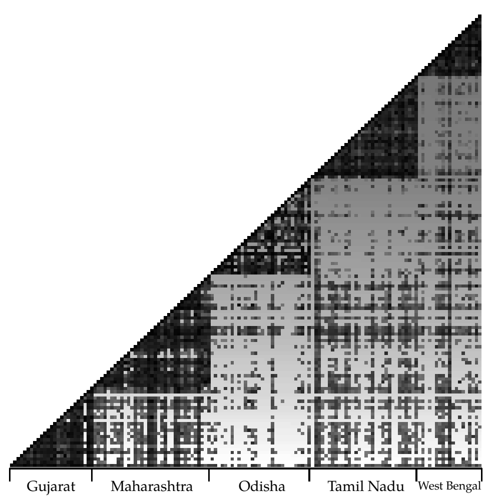
Note. For firms grouped by district–year, the contribution of factors: # customers, intensity per customer, and average customer size, to the variance of firms' sales was calculated as per equation (2.1). This figure is a box and whiskers plot of the contribution of these factors across districts arranged in a state×year grid.

FIGURE D.5. **Margins of Firms' Sales: Contribution to Variance, by Sales Quantile**



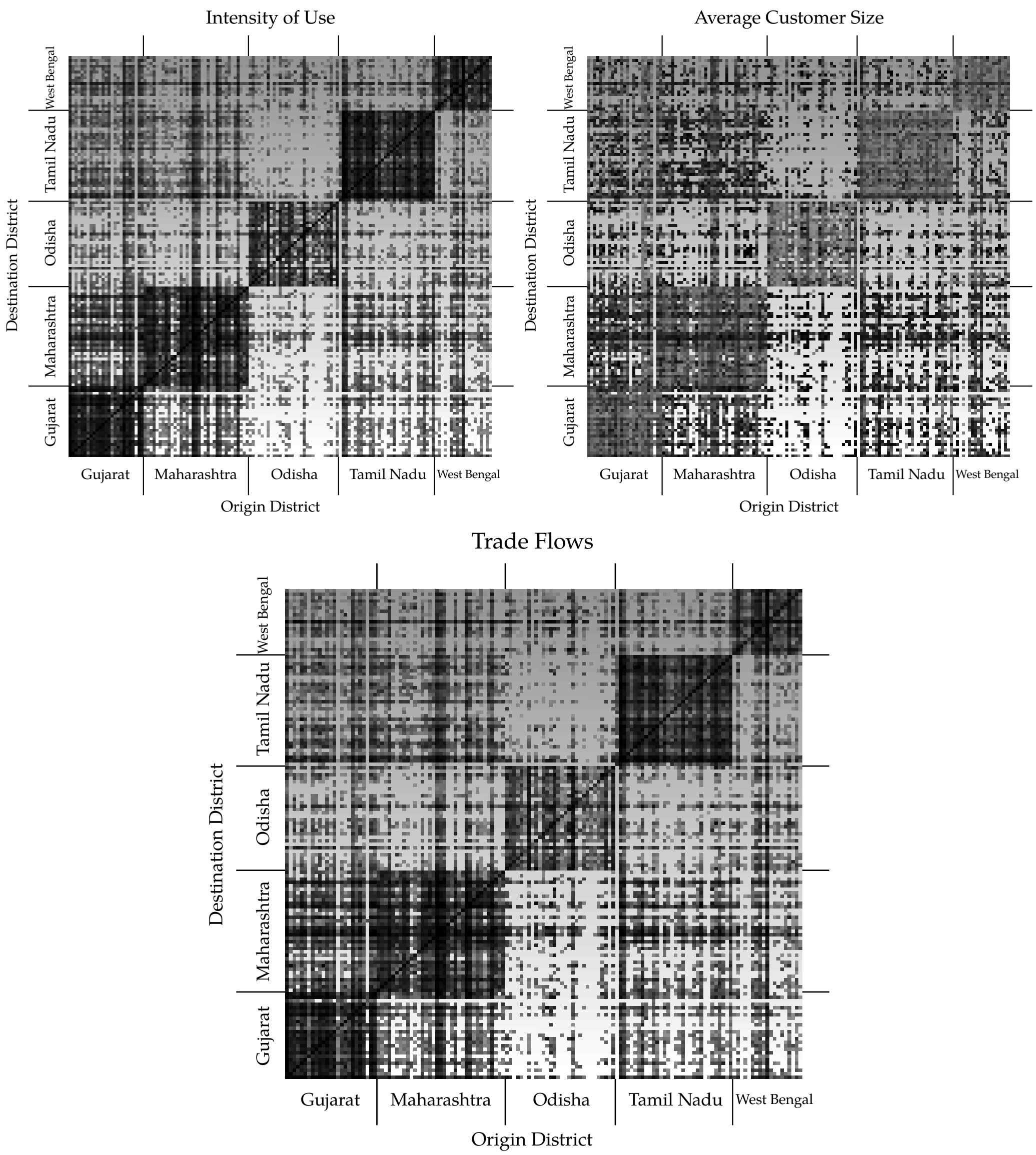
Note. For firms grouped into 100 equal-sized bins, the contribution of factors: # customers, intensity per customer, and average customer size, to the variance of firms' sales was calculated as per equation (2.1). This figure is a smoothed regression plot of the contribution of these factors across those bins.

FIGURE D.6. **Margins of Intranational Trade: Trade Integration between Districts**



Note. This figure depicts the (symmetric) 141×141 matrix of Head and Ries (2001) indexes of district pairs where cells with higher values are shaded darker. Districts are first ordered by state and then alphabetically within each state. Blocks along the diagonal depict values for intra-state district pairs while other areas depict inter-state district pairs.

FIGURE D.7. **Margins of Intranational Trade: Upstream & Downstream Margins**



Note. Across district pairs, the top left panel depicts the 141×141 matrix of intensity of use (i.e., the upstream margin), the top right depicts average customer size (i.e., the downstream margin), and the bottom depicts trade flows or sales from origin to destination. Darker cells reflect higher values. Districts are arranged first by state and then alphabetically within states on both axes. All values pertain to 2015-2016.

Appendix E. Theoretical Framework: Supplementary Material

E.1. Some Properties of Extreme Value Distributions.

Note. Some relevant values of polygamma functions can be calculated as follows:

$$\psi^{(0)}(1) = -\gamma,$$
$$\psi^{(1)}(1) = \frac{\pi^2}{6},$$
$$\psi^{(2)}(1) = 2\zeta(3),$$
$$\psi^{(3)}(1) = \frac{\pi^4}{15},$$
$$\psi^{(n)}(2) = \psi^{(n)}(1) + (-1)^n n!,$$

where γ is the Euler-Mascheroni constant, π is Archimedes' constant, and $\zeta(3)$ is the Apéry constant.

Lemma 1. *If a random variable X with support over $\mathbb{R}_{>0}$ is such that $\mathbb{P}(X \leq x) = e^{-\left(\frac{x}{s}\right)^{-\alpha}}$, then the following moment conditions hold:*

$$\mathbb{E}\left[X^j\right] = \Gamma\left(1 - \frac{j}{\alpha}\right) s^j \quad \forall j < \alpha,$$
$$\mathbb{E}[\log X] = -\frac{\psi^{(0)}(1)}{\alpha} + \log s,$$
$$\mathbb{E}\left[(\log X - \mathbb{E}[\log X])^2\right] = \frac{\psi^{(1)}(1)}{\alpha^2}.$$

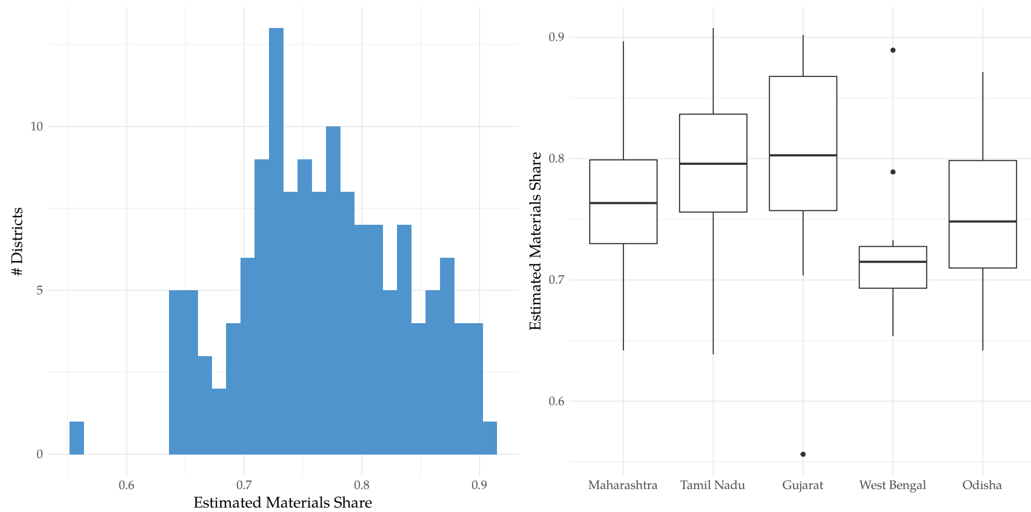
Lemma 2. *If a random variable X with support over $\mathbb{R}_{>0}$ is such that $\mathbb{P}(X \leq x) = 1 - \left(\frac{x}{s}\right)^{\alpha} e^{-\left(\frac{x}{s}\right)^{\alpha}} - e^{-\left(\frac{x}{s}\right)^{\alpha}}$, then the following moment conditions hold:*

$$\mathbb{E}\left[X^j\right] = \Gamma\left(2 + \frac{j}{\alpha}\right) s^j \quad \forall j > -2\alpha,$$
$$\mathbb{E}[\log X] = \frac{\psi^{(0)}(2)}{\alpha} + \log s,$$
$$\mathbb{E}\left[(\log X - \mathbb{E}[\log X])^2\right] = \frac{\psi^{(1)}(2)}{\alpha^2}.$$

Appendix F. Empirical Framework: Supplementary Material

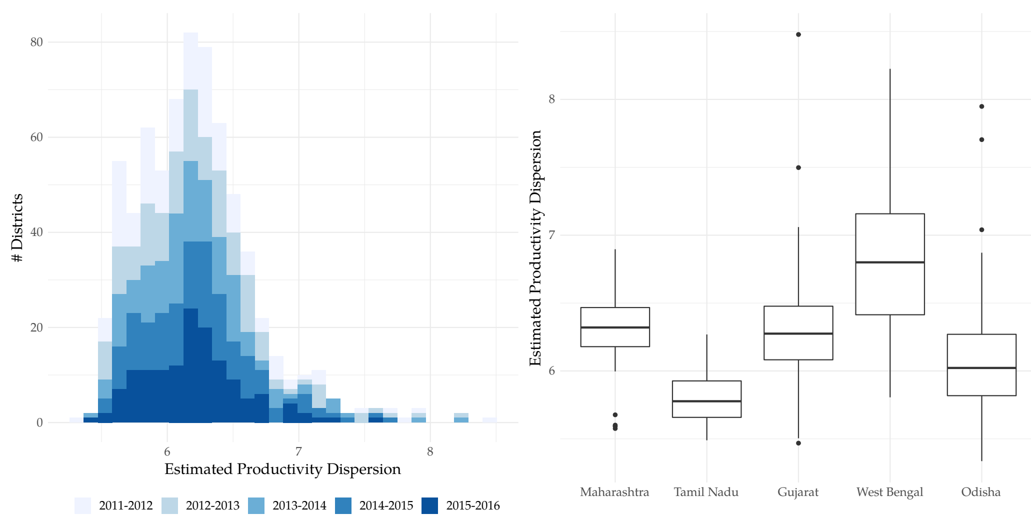
For each district, Figure F.1 presents estimates of material share of costs across districts. For each district–year pair, Figure F.2 presents estimates of shape parameters of Fréchet distribution of productivities and Figure F.3 presents estimates of task intensities of zero-truncated Poisson distribution of numbers of tasks.

Figure F.1. **Estimates of Material Shares α**



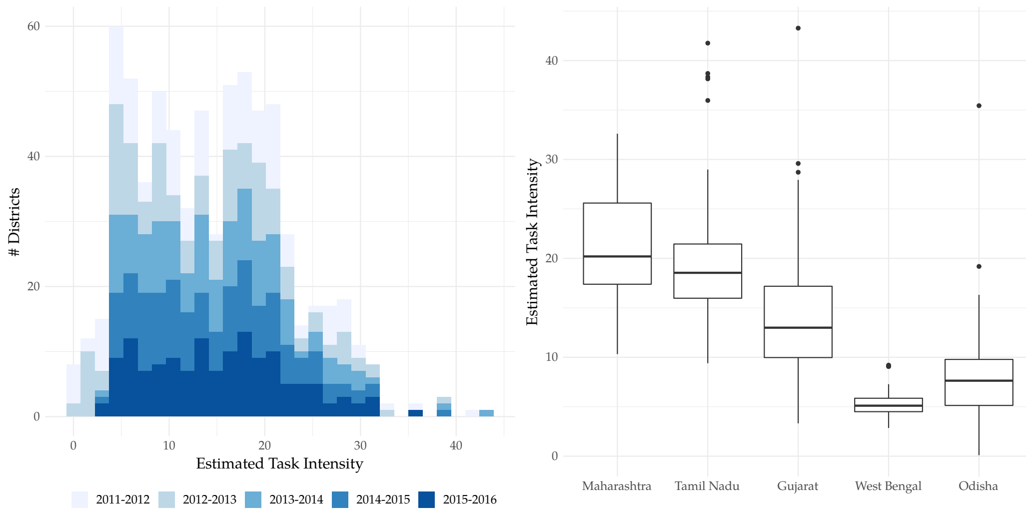
Note. The left panel is a histogram of estimated material share of costs across districts. The right panel is a box and whiskers plot of estimated material share of costs across districts within each state. States are arranged by economic size in descending order.

Figure F.2. **Estimates of Productivity Dispersion θ**



Note. The left panel is a stacked histogram of estimated shape parameter of Fréchet distributions of firms' productivities, across districts (from Assumption (1)). The right panel is a box and whiskers plot of the estimated shape parameters across district–year pairs within each state. States are arranged by economic size in descending order.

Figure F.3. **Estimates of Task Intensities $\boldsymbol{\kappa}$**



Note. The left panel is a stacked histogram of estimated task intensities of zero-truncated Poisson distributions of firms' numbers of tasks, across districts (from Assumption (2)). The right panel is a box and whiskers plot of the estimated task intensities across district–year pairs within each state. States are arranged by economic size in descending order.