Export and Innovation:
The Role of Inter-Firm Knowledge Networks

Chong Liu*[1] and Zi Wang†[2]

[1]Peking University

[2]Shanghai University of Finance and Economics

June 14, 2020

Abstract

We propose a quantifiable general equilibrium model with trade and firm innovation in which knowledge diffuses across domestic firms via a highly selective network. Our theory suggests that in the presence of inter-firm knowledge networks, non-exporting firms can benefit indirectly from export liberalization by learning from exporting firms. We estimate the structure of inter-firm knowledge networks using the unique dataset on patent citations across Chinese manufacturing firms. Simulations of our model suggest that inter-firm knowledge networks account for about one third of the increase in the Chinese real income induced by trade liberalization over 2001-2006.

Keywords: *Trade Liberalization*; *Firm Innovation*; *Knowledge Diffusion.*
JEL classification: *O11*; *O33*; *O38*; *F62.*

*cliu.econ@pku.edu.cn
†wang.zi@mail.shufe.edu.cn

1 Introduction

Understanding the impacts of trade on firm innovation is central to evaluating long-term consequences of globalization. An extensive literature has documented heterogeneous innovation responses of firms to export liberalization: it tends to increase market size and innovation incentives for large exporters, but decrease the innovation of non-exporters.[1] Since most firms do not export,[2] their negative innovation responses could erode welfare gains from trade liberalization.

While recent studies highlight the firms' heterogeneous innovation responses to export liberalization, most of them assume that firms innovate in isolation. The literature on firm innovation suggests the opposite: firms learn with each other via a highly selective and unevenly spread network.[3] In the presence of inter-firm knowledge networks, a non-exporting firm could raise its productivity under export liberalization by learning from exporting firms. The quantitative importance of this indirect productivity effect of trade liberalization depends on the structure of inter-firm knowledge networks.

This paper aims to quantify the implications of inter-firm knowledge networks for the long-term consequences of trade liberalization. To achieve this, it first requires empirical measurements on the structure of inter-firm knowledge networks. In general, there is lack of data revealing such information. Few recent studies investigate *inter-sector* knowledge networks and document sectoral heterogeneity in the knowledge space.[4] However, within any sector there is only a very small fraction of firms engaged into international trade. So it is crucial to characterize technology diffusion across *firms*, in particular to what extent knowledge diffuses from exporters to non-exporters.

Following the literature on technology diffusion, we take patent citations across firms as *proxy* for inter-firm knowledge diffusion.[5] Specifically, we combine data on patent citations

---

[1] Aghion et al. (2018) propose a theory of firms' heterogeneous innovation responses to export shocks and find supporting evidence from the French firm data. These heterogeneous responses are also consistent with theories in Melitz (2003) and Akcigit et al. (2018), and empirical evidence in Lileeva and Trefler (2010) and Bustos (2011).

[2] Bernard and Jensen (1995), Mayer and Ottaviano (2007), and Bernard et al. (2012) have documented using firm-level data from a wide range of countries that firm participation in international trade is exceedingly rare.

[3] Jaffe et al. (1993), Jaffe et al. (2000), Hall et al. (2005), and Giuliani (2007) have documented the evenly spread networks through which firms or inventors transmit their tacit knowledge.

[4] Acemoglu et al. (2016) use the U.S. patent citation data to investigate knowledge networks across sectors and find substantial sectoral heterogeneity in knowledge diffusion. Cai and Li (2018) build a quantitative model with inter-sector knowledge networks.

[5] As argued by Jaffe et al. (1993), "Thus, in principle, a citation of Patent X by Patent Y means that X represents a piece of previously existing knowledge upon which Y builds".

across Chinese manufacturing firms with their performances in the Chinese Manufacturing Survey and their exports in the Chinese Customs Records. The combined dataset reveals substantial firm heterogeneity in both exports and knowledge diffusion. Figure 1 shows that (i) the firm's export intensity increases with its size, and (ii) the patents of larger exporters are cited by more firms. That is to say, larger firms export more and are more visible in the knowledge networks and therefore more likely to diffuse their knowledge to other firms. To our best knowledge, this is the first attempt to link firms' characteristics, especially their exports, to their positions in the patent citation networks.

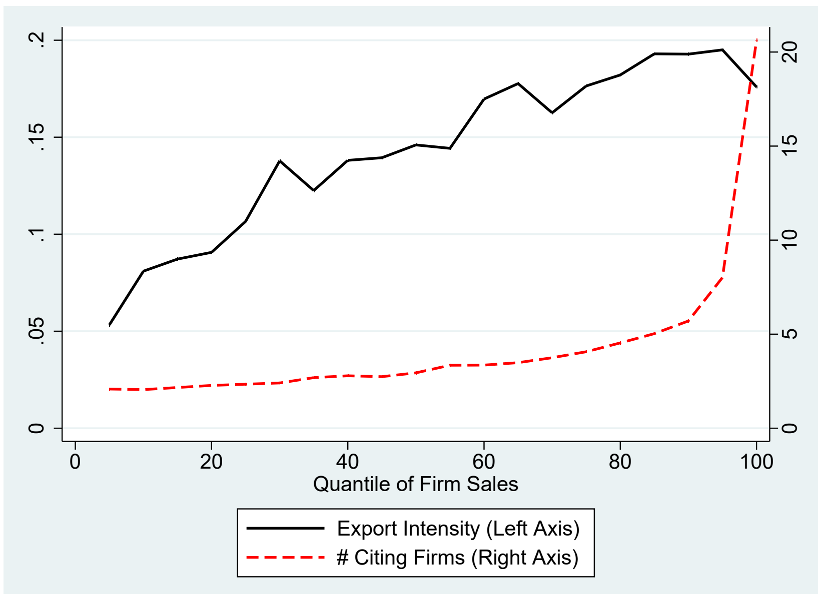


Figure 1: Firm Heterogeneity in Sales, Exports, and the Number of Citing Firms

(Note: The sample includes the Chinese manufacturing firms whose patents are cited by at least one other firms over 1998-2013. The firm sales and export intensities are averaged over 1998-2013. The sample firms are sorted by their sales and divided equally to 20 groups. For each group we compute the average export intensity and the number of citing firms (weighted by firm sales).)

To incorporate firm heterogeneity in exports and knowledge diffusion and draw aggregate implications, we propose a multi-country general equilibrium model with trade and firm innovation in which knowledge diffuses across domestic firms via a highly selective network. Our endogenous growth model features a continuum of firms that are heterogeneous in their current productivities and innovation capabilities. Innovation allows firms to raise their productivities by an amount that increases with innovation investment and innovation capabilities. Besides doing their own R&D, firms can also improve their technologies by learning from other firms. In each period, firms can randomly meet with other firms and absorb the knowledge of firms they meet. We allow the probability that two firms meet to be a flexible function of the characteristics of both firms. This flexible firm-to-firm matching function enables our model to capture firm heterogeneity in knowledge connections observed in the micro data. It is also convenient for aggregation: despite the multiple degrees of heterogeneity (in firms' current productivities, innovation capabilities, and meeting probabilities), the model is tractable and transparent for estimation.

Trade in our model is standard: firms incur both iceberg and fixed costs to export. Therefore, only the most productive firms would export. The decline in iceberg trade costs would increase market size and therefore innovation incentives for exporting firms, but decrease the innovation incentives for non-exporting firms by bidding up the wage. While these heterogeneous innovation responses have been well-discussed in the literature, we emphasize a new effect via knowledge diffusion: non-exporting firms can learn from exporting firms whose innovation incentives have been spurred by export liberalization. We show that this new effect could partially offset the negative productivity effect led by the decline in non-exporters' innovation and therefore increase the aggregate productivity and welfare gains from export liberalization.

To quantify macro-level implications, we conduct a two-tiered empirical analysis. In the first tier, we parameterize our firm-to-firm matching function and estimate it using inter-firm knowledge networks observed in the Chinese patent citation data. Using the simulated method of moments with equilibrium conditions as constraints, we find that the probability of two firms meeting increases with the size of both parties and is log-supermodular with respect to firm size. Our parameterized matching function is able to reproduce key targeted and untargeted moments of inter-firm knowledge networks in the data.

In the second tier of our empirical inquiry, we focus on general equilibrium welfare analysis in a two-economy world. We insert the estimated inter-firm knowledge networks into the model and calibrate changes in trade costs to the observed trade shares between China and the rest of the world over 2001-2006. The calibration results suggest considerable decline in the Chinese import and export costs during this period.

We then quantify the impacts of trade liberalization on the Chinese real income by fixing trade costs between China and the rest of the world in their levels of 2001. Our simulation shows that the decline in trade costs over 2001-2006 increases the time-discounted Chinese real income by 1.8%. Then we eliminate inter-firm knowledge networks and re-conduct this counterfactual exercise. Without inter-firm knowledge networks, trade liberalization only increases the time-discounted Chinese real income by 1.2%. As a result, inter-firm knowledge diffusion accounts for about one third of the Chinese welfare gains from trade liberalization over 2001-2006. Without inter-firm knowledge networks, trade liberalization would dramatically reduce the innovation of non-exporting firms in China, which erodes the Chinese productivity and welfare gains from trade liberalization.

Related Literature: Relative to the literature on firm innovation and trade liberalization, our model adds inter-firm knowledge networks, and it extends previous models to a quantitative setting. Aghion et al. (2018) and Akcigit et al. (2018) develop trade models with endogenous firm innovation and endogenous competition, but without inter-firm knowledge diffusion. Also related are quantitative models of innovation and inter-sector knowledge diffusion, such as Klette and Kortum (2004) and Cai and Li (2018). As discussed above, to explore the innovation effects of trade liberalization, we must focus on inter-firm rather than inter-sector knowledge diffusion since in any sector firm participation in international trade is very rare.

This paper also relates to empirical studies that document firms' heterogeneous innovation responses to trade liberalization. Aghion et al. (2018), Bustos (2011), Lileeva and Trefler (2010), and Coelli et al. (2016) document firm differential responses to export shocks. Bloom et al. (2016), Iacovone et al. (2011), Autor et al. (2016), and Bombardini et al. (2017) document heterogeneous innovation responses to import competition. Consistent with these empirical explorations, our model suggests that export liberalization promotes the innovation of larger and more productivity firms but reduces the innovation of smaller firms, in particular non-exporters.

Our model builds on the recent quantitative models of firm-to-firm networks. There is an extensive literature that quantifies firm-to-firm production linkages (Lim, 2019; Huneeus, 2019; Acemoglu and Azar, 2017; Acemgolu et al., 2012; Atalay et al., 2011; Bernard and Moxnes, 2018; Tintelnot et al., 2018). Efforts for quantifying inter-firm knowledge networks are rare.[6] These tractable characterizations of firm-to-firm linkages, including the matching function in our model, have integrated micro structure of production and innovation into macro production function and drawn aggregate implications.

The rest of this paper is arranged as follows. Section 2 documents key features of patent citation networks across Chinese manufacturing firms in order to motivate our structural model. Section 3 builds the model and characterizes the model's implications for trade and innovation based on a special case with two symmetric countries. Section 4 estimates inter-firm knowledge networks. Section 5 conducts counterfactual experiments to quantify

[6] Bloom et al. (2013) investigate inter-firm knowledge diffusion in a stylized empirical framework. Akcigit and Kerr (2016) and Akcigit et al. (2017) incorporate inter-firm knowledge diffusion into quantitative frameworks, but do not capture firm heterogeneity in the knowledge network.

the importance of inter-firm knowledge networks to welfare gains from trade liberalization. Section 6 concludes.

2 Motivational Facts

This section documents firm heterogeneity in exports and knowledge connections. Following the literature of innovation and technology diffusion, we use patent citation as proxy of inter-firm knowledge diffusion. We then link a firm's position in patent citation networks to its characteristics. These micro data patterns motivate the specification of knowledge networks in our model and provides empirical moments for the structural estimation of our model's key parameters.

2.1 Data Sources

Our data consists of three parts. First, we observe the performance of Chinese manufacturing firms whose annual sales exceed 5 million RMB (about $650,000) over 1998-2013 from the Annual Survey of Industrial Firms (ASIF). The main performance variables include sales, employment, capital stock, materials, and export values. Second, we observe all patents assigned by the Chinese intellectual property offices, with the patent number, the contact information of patent owners, the inventors, and the description of the patents. Third, we obtain the citation linkages across these patents from the Google patent. We merge the firms in ASIF and patent records by their names and contact information. The details of data sources and data merge are presented in Appendix A.

We focus on the time-invariant deep structure of inter-firm knowledge networks. Therefore, we construct our database used in motivational facts and quantification as follows. First, we define that firm A learns from firm B if and only if A has cited at least one patent owned by B over 2000-2013. Second, we exclude the patents that are cited over 2008-2013 because there is no sufficient time for these patents to be widely visible. Third, we take averages on firms' performances and exports over 2000-2013. As a result, we construct a cross-sectional dataset of Chinese manufacturing firms with their long-term average performances, exports, and inter-firm patent citation linkages. The detailed summary statistics of our database are presented in Appendix A.

2.2 Regularities of Inter-firm Patent Citation Networks

Armed with the database constructed as described in Section 2.1, we present several regularities about firm heterogeneity in patent citations and exports. We first investigate firm heterogeneity in the patent citation network. In particular, we find that:

Fact 1: *A considerable fraction of patent citations occur within industries. Even within a narrowly-defined industry, firms still occupy heterogeneous positions in patent citation networks.*

Notably, a large fraction of patent citations occur across firms within narrowly-defined industries. Table 1 suggests that about 40% of firm pairs with patent citations are within 2-digit CIC industry. Even if we look at 3-digit CIC industries there are still about one quarter of firm pairs within industry. This implies that a large share of knowledge diffusion occurs within industry across firms. As a result, understanding inter-firm, not only inter-industry, knowledge networks is crucially to understand the structure of knowledge diffusion.

Table 1: Patent Citation Linkages across Chinese Manufacturing Firms

# Firm pairs with patent citations	85471
In which:	
Cited and citing firms in the same 2-digit CIC industry	36939
Cited and citing firms in the same 3-digit CIC industry	22012

Inter-firm knowledge diffusion is not pervasive. Figure 2 shows that firms occupy very heterogeneous positions in the knowledge network: in a small network connecting 130 household appliance producers, few large firms lie at the center with a large number of connections, whereas most small firms lie at the peripheral with few connections. We will show later that this heterogeneity is not specific to the household appliance industry but common in the whole manufacturing sector.

We proceed by linking firm size and export status to firms' positions in patent citation networks. In particular, we find that:

Fact 2: *Large exporters lie at the center of inter-firm patent citation networks.*

Figure 3 plots the number of firms from which a firm cites (in-degree) against the fraction of firms who cite that many firms. We also plot the number of firms citing a firm (out-degree) against the fraction of firms who are cited by that many firms. The distributions appear to be very close to a Pareto distribution as the cdfs are close to linear. In other words, the distributions of in-degree and out-degree are characterized by many firms with few connections and few firms with many connections.

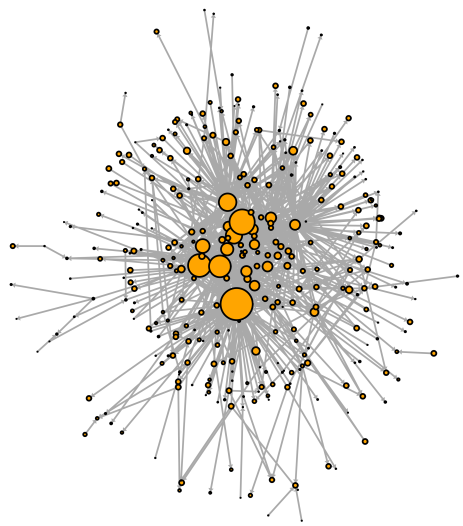
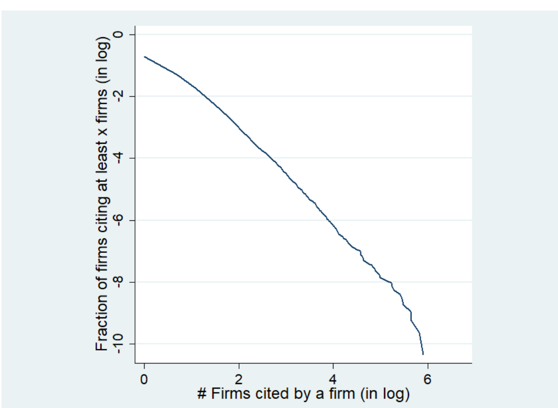
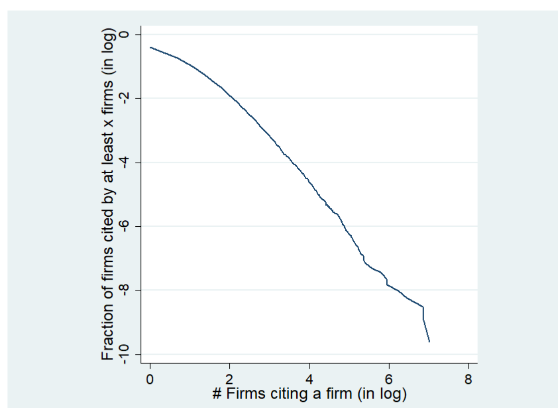


Figure 2: The Inter-firm Patent Citation Network in Household Appliances (CIC 395)



(a) In-Degree

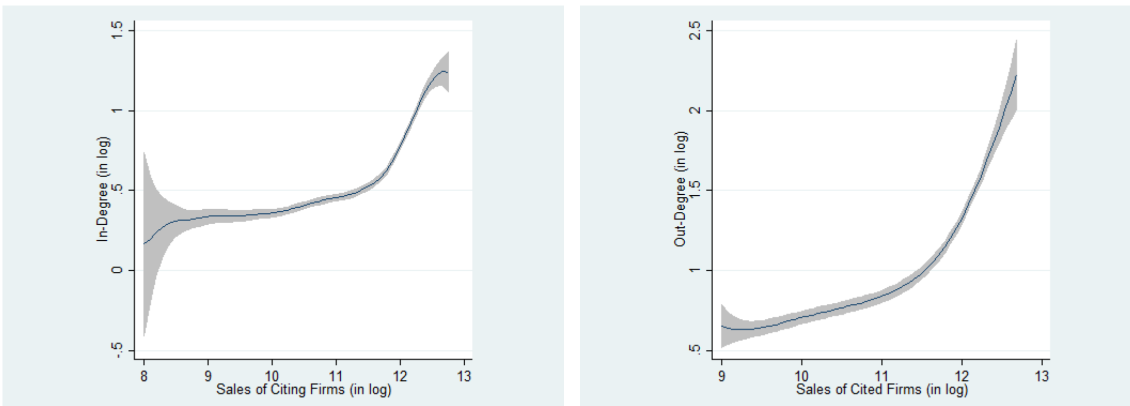


(b) Out-Degree

Figure 3: Distributions of In- and Out-Degree

(Notes: In-degree is defined as the number of firms from which a firm cites. Out-degree is defined as the number of firms citing a firm.)

The firms with many connections in the patent citation network are not randomly assigned. Figure 4 shows that larger firms cite and are cited by more firms. Figure 5 shows that exporters have larger in-degree and out-degree than non-exporters. Combining the regularities documented in Figure 3, 4, and 5, we find that few large exporters lie at the center of the inter-firm patent citation network. As a result, shocks in export markets directly affect these larger exporters, and then propagate via inter-firm knowledge networks and translate into aggregate effects on all firms.

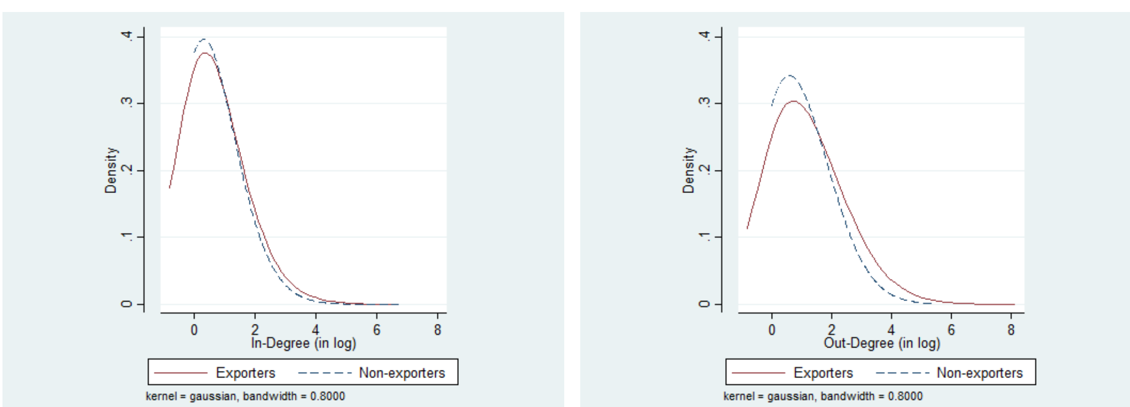


(a) In-Degree

(b) Out-Degree

Figure 4: Firms' Sales and In-/Out-Degree

(Notes: In-degree is defined as the number of firms from which a firm cites. Out-degree is defined as the number of firms citing a firm.)



(a) In-Degree

(b) Out-Degree

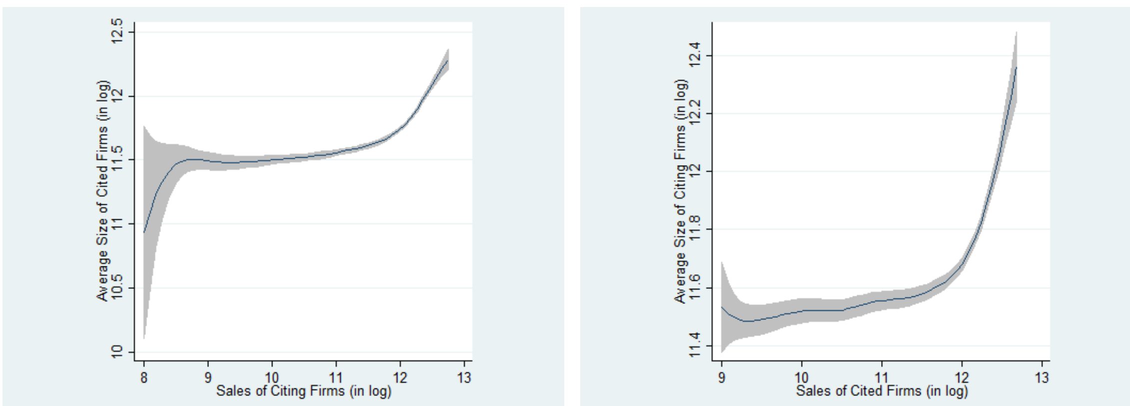
Figure 5: Firms' Export Status and In-/Out-Degree

(Notes: In-degree is defined as the number of firms from which a firm cites. Out-degree is defined as the number of firms citing a firm.)

We turn to investigate what kinds of firms are more likely to be connected. In particular, we find that:

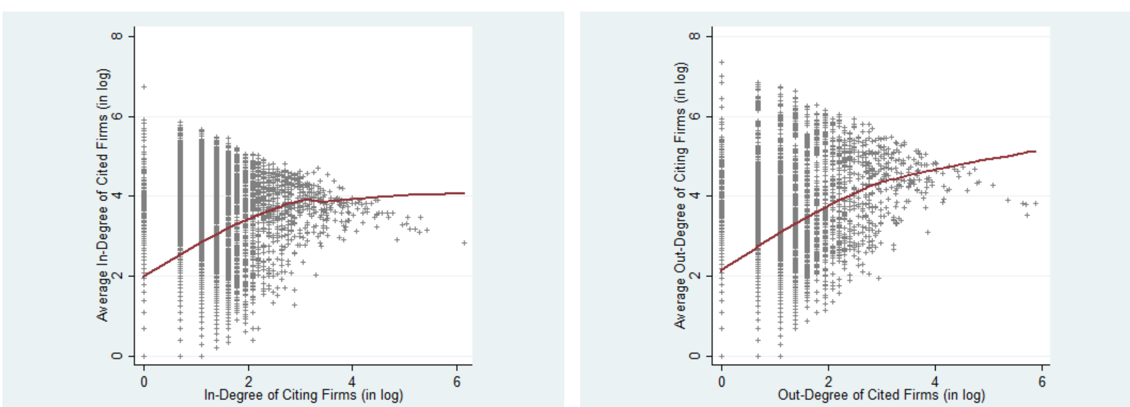
Fact 3: *The inter-firm knowledge network exhibits strong positive assortivity: larger and more connected firms are connected with larger and more connected firms.*

Figure 6 shows that larger firms are more likely to cite and be cited by larger firms. Figure 7 shows that more connected firms are more likely to connect with more connected firms. These results suggest that inter-firm patent citation networks are highly selective: a small number of large firms lie at the center of the network and are well-connected with each other, whereas the majority of small firms lie at the peripheral of the network, with very few connections. In the next section, we will show how our model captures these characteristics of inter-firm knowledge networks.



(a) citing firms (b) cited firms

Figure 6: Matching Assortivity: Sales



(a) In-Degree (b) Out-Degree

Figure 7: Matching Assortativity: Degree

(Notes: In-degree is defined as the number of firms from which a firm cites. Out-degree is defined as the number of firms citing a firm.)

3 The Model

Consider a world made up of two countries, $\ell = 1, 2$. Time is discrete and goes to infinity. The preferences and production possibilities of each country ℓ are as follows.

3.1 Preferences

Country ℓ is endowed with labor L_ℓ who are infinitely lived, immobile across countries, and inelastic in supply. The representative consumer in country ℓ has constant-elasticity-of-substitution (CES) preferences over a unit mass of of varieties and seeks to maximize:

$$U_\ell = \sum_{t=0}^{\infty} \iota^t \left[\int_0^1 c_{\ell,t}(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{\sigma}{\sigma-1}}, \tag{1}$$

where $c_{\ell,t}(\omega)$ denotes consumption of good ω, $\sigma > 1$ is the elasticity of substitution, and $\iota \in [0, 1]$ is the time discounting factor.

3.2 Technologies and Trade

Each variety is produced by a firm using labor. The market structure for all firm sales is assumed to be monopolistic competition. Each firm is owned by a family of entrepreneurs. Each entrepreneur lives for one period. At the beginning of period t, the entrepreneur of firm ω born at period t in country ℓ inherits labor productivity $\phi_{\ell,t}(\omega)$ from her ancestors and makes the innovation decision $\kappa_{\ell,t}(\omega) \geq 0$. Firm ω with inherited productivity $\phi_{\ell,t}(\omega)$ and innovation $\kappa_{\ell,t}(\omega)$ has effective productivity $\kappa_{\ell,t}(\omega)\phi_{\ell,t}(\omega)$ and thus marginal cost of production $\frac{w_{\ell,t}}{\kappa_{\ell,t}(\omega)\phi_{\ell,t}(\omega)}$ where $w_{\ell,t}$ is the wage of country ℓ at period t. The initial productivity distribution $\{\phi_{\ell,0}(\omega)\}$ is exogenous. We assume that the mass of firms in each country is fixed and there is no entry and exit of firms.

Firms can sell their products both at home and abroad. However, as in Melitz (2003) firms that select into exporting face both fixed and variable costs of trade. Firms in country ℓ that export to country n at period t incur a fixed cost $f^X_{\ell n,t} \geq 0$ in terms of country ℓ's labor, with $f^X_{\ell\ell,t} = 0$, while variable trade costs take the iceberg form. To deliver one unit of output from country ℓ to country n at period t a firm must ship $\tau_{\ell n,t} \geq 1$ units, with $\tau_{\ell\ell,t} = 1$.

Conditional on the distribution of effective productivity $\kappa_{\ell,t}(\omega)\phi_{\ell,t}(\omega)$, the structure of production and demand in this economy is equivalent to that in Melitz (2003) and solving firms' static profit maximization problem is straightforward. Firms face isoelastic demand

and set factory gate prices as a constant mark-up $\frac{\sigma}{\sigma-1}$ over marginal costs. Since $f^X_{\ell\ell,t} = 0$, all firms will produce and make sales in the domestic market. Firms only export to a given market if their variable profits in that market are sufficient to cover the fixed export cost. In particular, firms choose to export from country ℓ to country n if and only if their effective productivity is at least $\tau_{\ell n,t}\left(\frac{\sigma^{\frac{\sigma}{\sigma-1}}}{\sigma-1}\right)\left(\frac{w^{\sigma}_{\ell,t} f^X_{\ell n,t}}{D_{n,t}}\right)^{\frac{1}{\sigma-1}}$ where $D_{n,t} := P^{\sigma-1}_{n,t} X_{n,t}$ is the aggregate demand shifter in country n at period t and where $P_{n,t}$ is the price index and $X_{n,t}$ is the total expenditure.

3.3 Firm Innovation

Firms are heterogeneous in innovation efficiency. To achieve innovation $\kappa_{\ell,t}(\omega)$, firm ω has to employ $\frac{\kappa_{\ell,t}(\omega)^{\alpha}}{z(\omega)}$ additional units of labor where $z(\omega) > 0$ denotes the innovation efficiency of firm ω. We regard $z(\omega)$ as the *fundamental characteristics* of firm ω so that $z(\omega)$ is exogenous and time-invariant. Therefore, we index firm ω by its innovation efficiency z and denote the cumulative distribution function of z in country ℓ as $G_{\ell}(z)$ with the support S_z.

Since entrepreneurs are assumed to be one-period-lived, they decide their innovation by solving the following one-period profit-maximization problem:

$$\max_{\kappa_{\ell,t}(z)} \left\{ \tilde{\sigma} w^{1-\sigma}_{\ell,t} \left[\kappa_{\ell,t}(z)\phi_{\ell,t}(z)\right]^{\sigma-1} D_{\ell,t}(z, \phi_{\ell,t}(z)) - w_{\ell,t}\frac{\kappa_{\ell,t}(z)^{\alpha}}{z} \right\}, \tag{2}$$

where $D_{\ell,t}(z, \phi_{\ell,t}(z))$ is the demand shifter faced by firm z of country ℓ, which will be specified below. The constant $\tilde{\sigma} := \frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}$. To guarantee an interior solution to Problem (2), we assume that $\alpha > \sigma - 1$.

Firm innovation in our model is isomorphic to the model of firm innovation developed by Desmet, Nagy, and Rossi-Hansberg (2018) in which firms innovate in order to maximize their bid for land and obtain zero profits after covering their innovation costs. Following this idea, the complex dynamic problem of firm innovation is simplified into a repeated static problem.

The optimal innovation for firm z in country ℓ at period t is determined by its demand, its innovation cost, and its inherited productivity:

$$\kappa^*_{\ell,t}(z) = \left[\frac{(\sigma-1)\tilde{\sigma}}{\alpha} D_{\ell,t}(z, \phi_{\ell,t}(z))\right]^{\frac{1}{\alpha-(\sigma-1)}} \left[z\phi_{\ell,t}(z)^{\sigma-1}\right]^{\frac{1}{\alpha-(\sigma-1)}} w_{\ell,t}^{-\frac{\sigma}{\alpha-(\sigma-1)}}. \tag{3}$$

The demand shifter faced by firm z in country ℓ, $D_{\ell,t}(z, \phi_{\ell,t}(z))$, depends on whether it is an exporter at period t. We have shown that a firm will export if and only if its effective productivity is above a cutoff productivity. Armed with the optimal innovation in Equation (3), the demand shifter faced by firm z in country ℓ is

$$D_{\ell,t}\left(z,\phi_{\ell,t}(z)\right)=\begin{cases}D_{\ell,t}+\tau_{\ell n,t}^{1-\sigma}D_{n,t}, & \text{if } \left(z,\phi_{\ell,t}(z)\right)\in\mathcal{E}_{\ell,t}\\ D_{\ell,t}, & \text{if } \left(z,\phi_{\ell,t}(z)\right)\notin\mathcal{E}_{\ell,t}\end{cases}\tag{4}$$

for $\ell \neq n$, where the set of exporters, $\mathcal{E}_{\ell,t}$, is defined by

$$\mathcal{E}_{\ell,t}:=\left\{\left(z,\phi_{\ell,t}(z)\right)|\tilde{\Lambda}_1 z^{\frac{\sigma-1}{\alpha-(\sigma-1)}}\phi_{\ell,t}(z)^{\frac{\alpha(\sigma-1)}{\alpha-(\sigma-1)}}w_{\ell,t}^{-\frac{(\sigma-1)(1+\alpha)}{\alpha-(\sigma-1)}}\left[\left(D_{\ell,t}+\tau_{\ell n,t}^{1-\sigma}D_{n,t}\right)^{\frac{\alpha}{\alpha-(\sigma-1)}}-D_{\ell,t}^{\frac{\alpha}{\alpha-(\sigma-1)}}\right]\geq w_{\ell,t}f_{\ell n,t}^{X}\right\},\tag{5}$$

and where the constant $\tilde{\Lambda}_1:=\left(1-\frac{\sigma-1}{\alpha}\right)\tilde{\sigma}\left[\frac{\sigma-1}{\alpha}\tilde{\sigma}\right]^{\frac{\sigma-1}{\alpha-(\sigma-1)}}$,

3.4 Inter-Firm Knowledge Networks

We assume that the productivity at the beginning of period $t+1$, $\phi_{\ell,t+1}(z)$, does not only depend on the firm's own innovation outcomes at period t, but also on other firms' innovation outcomes. Specifically, we assume that productivity $\{\phi_{\ell,t}(z)\}$ is evolved as follows:

$$\phi_{\ell,t+1}(z)=\left[\kappa_{\ell,t}^{*}(z)\phi_{\ell,t}(z)+\delta\int_{S_z}m(z',z)\kappa_{\ell,t}^{*}(z')\phi_{\ell,t}\left(z'\right)dG_{\ell}(z')\right]^{\beta},\quad \beta\leq 1,\quad \delta>0.\tag{6}$$

Several issues are worth further discussing. First, firm z is only able to receive knowledge from firm z' with probability $m(z',z)\in[0,1]$. Given that there exists a continuum of firms of every state z, $m(z',z)$ is also equal to the fraction of z'-firms that diffuse their knowledge to a given z-firm, as well as the fraction of z-firms that receive knowledge from a given z'-firm. As a result, the structure of inter-firm knowledge networks can be fully characterized by this matching function $m(z',z)$. Notably, Lim (2018) has utilized a similar matching function to characterize production networks across firms.

Second, in our baseline specification, we assume that the matching function $m(z',z)$ is time-invariant and depends only on firms' *fundamental characteristics* z. In this paper, we focus on the implications of stable and time-invariant characteristics of inter-firm knowledge networks. Moreover, as shown below, this specification leads to simple conditions that ensure the uniqueness of steady-state equilibrium and transparent estimates of the model's key parameters.

Third, we do not specify a search and matching process to rationalize our matching function $m(z',z)$. The general form of $m(z',z)$ makes our model sufficiently flexible to capture rich patterns of inter-firm knowledge spillovers observed in the patent citation data. We leave the micro-foundation of matching function to future work.

Finally, it is straightforward to incorporate other firm characteristics than innovation efficiency z into our matching function. In this paper, we concentrate on the single-dimensional innovation efficiency because most of the firm characteristics are strongly correlated with firm size. In some context, firm characteristics other than size may be in special interest. For example, if we are interested in how foreign ownership affects inter-firm knowledge spillovers, we have to incorporate ownership status into our matching function. Moreover, variables that characterize physical and technological distances across firms can also be included. We leave these concerns to future work.

3.5 Equilibrium

We close the model by clearing the markets. Labor market clearing implies that

$$w_{\ell,t}L_\ell = \underbrace{\left[1-\frac{1}{\sigma}\left(1-\frac{\sigma-1}{\alpha}\right)\right]\int_{S_z} X_{\ell,t}(z)dG_\ell(z)}_{\text{Production}} + \underbrace{\int_{\mathcal{E}_{\ell,t}} w_{\ell,t}f^X_{\ell n,t}dG_\ell(z)}_{\text{Fixed Export Costs}}, \tag{7}$$

where the firm-level sales is given by:

$$X_{\ell,t}(z) = \sigma\tilde{\sigma}w_{\ell,t}^{1-\sigma}\left[\kappa^*_{\ell,t}(z)\phi_{\ell,t}(z)\right]^{\sigma-1}D_{\ell,t}\left(z,\phi_{\ell,t}(z)\right). \tag{8}$$

The total expenditure of country ℓ is given by

$$X_{\ell,t} = w_{\ell,t}L_\ell + \underbrace{\frac{1}{\sigma}\left(1-\frac{\sigma-1}{\alpha}\right)\int_{S_z} X_{\ell,t}(z)dG_\ell(z) - \int_{\mathcal{E}_{\ell,t}} w_{\ell,t}f^X_{\ell n,t}dG_\ell(z)}_{\text{Net Profits}}. \tag{9}$$

Finally, the aggregate price index is determined by the prices of domestic and imported varieties. For $n \neq \ell$, we have

$$P_{\ell,t} = \frac{\sigma}{\sigma-1}\left[\int_{S_z} w_{\ell,t}^{1-\sigma}\left[\kappa^*_{\ell,t}(z)\phi_{\ell,t}(z)\right]^{\sigma-1}dG_\ell(z) + \int_{\mathcal{E}_{n,t}}(w_{n,t}\tau_{n\ell,t})^{1-\sigma}\left[\kappa^*_{n,t}(z)\phi_{n,t}(z)\right]^{\sigma-1}dG_n(z)\right]^{\frac{1}{1-\sigma}}. \tag{10}$$

The instantaneous welfare in country ℓ can be measured by real income $W_{\ell,t} = \frac{X_{\ell,t}}{P_{\ell,t}}$.

Definition 1 (Dynamic Equilibrium) *Given $\{L_\ell, G_\ell(z)\}$, $m(z', z)$, and $\{\phi_{\ell,0}(z)\}$, the dynamic equilibrium of our model consists of $\{\kappa^*_{\ell,t}(z), \phi_{\ell,t}(z), P_{\ell,t}, X_{\ell,t}, w_{\ell,t}\}$ such that*

1. *Consumers maximize their utility.*

2. *Given $\{\phi_{\ell,t}(z), P_{\ell,t}, X_{\ell,t}, w_{\ell,t}\}$, each firm decides $\kappa_t^*(z)$ as in Equation (3) where its individual demand is characterized by Equation (4) and (5).*

3. *Given $\{\phi_{\ell,0}(z)\}$, productivities evolve as in Equation (6).*

4. *Wage is determined by Equation (7).*

5. *Aggregate price index $P_{\ell,t}$ is given by Equation (10).*

6. *Total expenditure $X_{\ell,t}$ is given by Equation (9).*

3.6 Steady-State

We proceed by defining and characterizing the steady-state of the equilibrium. We are particularly interested in the uniqueness of the steady-state since we will estimate $m(z', z)$ by matching the equilibrium outcomes in the steady-state to the data. In the steady-state, firms' productivities are time-invariant, i.e. $\phi_{\ell,t}(z) = \phi_\ell(z)$ for all t. As a result, $w_{\ell,t}$, $X_{\ell,t}$, $P_{\ell,t}$ and $\kappa_{\ell,t}^*(z)$ are all time-invariant in the steady-state. The steady-state $(\phi_\ell(z))_{z \in S_z}$ is determined by Equation (6). In general, the uniqueness of steady-state relies on the structure of inter-firm knowledge network, $m(z', z)$. This is one special case in which we can establish the uniqueness of steady-state for any exogenous $m(z', z) \geq 0$. The following result is an application of Theorem 1 of Allen, Arkolakis, and Li (2017):

Proposition 2 (Steady-State in the Close Economy) *Consider a close economy. Suppose that $1 - \frac{\alpha\beta}{\alpha-(\sigma-1)} > 0$. Then there exists a unique distribution of steady-state productivity for any matching function $m(z', z)$. Moreover, given $\tilde{D} := \left[\frac{(\sigma-1)\tilde{\sigma}}{\alpha} D\right]^{\frac{1}{\alpha-(\sigma-1)}}$, the steady-state $\{\phi(z)\}$ can be computed by iterating the following system of equations:*

$$\phi(z) = \tilde{D}^\beta \left[z^{\frac{1}{\alpha-(\sigma-1)}} \phi(z)^{\frac{\alpha}{\alpha-(\sigma-1)}} + \delta \int_{S_z} m(z', z) \, (z')^{\frac{1}{\alpha-(\sigma-1)}} \, \phi(z')^{\frac{\alpha}{\alpha-(\sigma-1)}} dG(z') \right]^\beta, \quad \forall z \in S_z. \tag{11}$$

Proposition 2 suggests that in the close economy, as long as $m(z', z)$ is exogenous, the sufficient conditions for uniqueness include only three parameters, (α, β, σ). This result greatly simplifies structural estimation.

Notably, there is no growth in the steady-state since we assume that L_ℓ is constant. This resembles the semi-endogenous growth model whose growth rate in the balanced growth path is equal to the exogenous labor growth rate.

3.7 Trade and Innovation in Two Symmetric Countries

In this subsection, we analytically characterize transitional dynamics in a special case of two symmetric economies. The purpose is to understand the mechanisms through which trade liberalization affects innovation and productivity growth of heterogeneous firms. We assume that two economies are symmetric in a steady-state before period t, with $\tau_{12,k} = \tau_{21,k} = \tau$ for all $k < t$. At the beginning of period t, there is a permanent symmetric change in trade costs: $d\tau_{12,t} = d\tau_{21,t} = d\tau$. Due to the symmetric of two economies, we take wage as a numeraire and drop the country subscript in equilibrium conditions.

We first characterize the equilibrium at period t. In the steady-state at the beginning of period t, $\phi_t(z)$ is non-decreasing with z. Therefore, firms choose to export if and only if $z \geq \tilde{z}_t$ where the cutoff $\tilde{z}_t$ satisfies:

$$\tilde{\Lambda}_1 \left[\tilde{z}_t^{\sigma-1} \phi_t(\tilde{z}_t)^{\alpha(\sigma-1)}\right]^{\frac{1}{\alpha-(\sigma-1)}} \left[\left(1+\tau_t^{1-\sigma}\right)^{\frac{\alpha}{\alpha-(\sigma-1)}} - 1\right] D_t^{\frac{\alpha}{\alpha-(\sigma-1)}} = f_t^X. \tag{12}$$

Equation (12) is the zero-profit (ZP) condition which implies a negative relationship between the cutoff $\tilde{z}_t$ and the aggregate demand shifter D_t. Intuitively, when the aggregate demand is larger, more firms are profitable to export, which reduces the cutoff $\tilde{z}_t$.

Under two symmetric countries, the goods market clearing condition can be simplified into:

$$X_t = L + \frac{1}{\sigma}\left(1 - \frac{\sigma-1}{\alpha}\right) X_t - \left[1 - G(\tilde{z}_t)\right] f_t^X. \tag{13}$$

Moreover, the aggregate price index can be expressed as

$$P_t = \frac{\sigma}{\sigma-1}\left[\int_{S_z} \left[\kappa_t^*(z)\phi_t(z)\right]^{\sigma-1} dG(z) + \int_{\tilde{z}_t}^{\infty} \tau_t^{1-\sigma}\left[\kappa_t^*(z)\phi_t(z)\right]^{\sigma-1} dG(z)\right]^{\frac{1}{1-\sigma}}. \tag{14}$$

Combining Equation (13) and (14), the aggregate demand shifter can be expressed as

$$D_t = \left\{\frac{L - \left[1 - G\left(\tilde{z}_t\right)\right] f_t^X}{1 - \frac{1}{\sigma}\left(1 - \frac{\sigma-1}{\alpha}\right)}\right\}^{1-\frac{\sigma-1}{\alpha}}$$
$$\left\{\frac{1}{\Lambda_2}\left[\int_0^{\tilde{z}_t} \left[z\phi_t(z)^{\alpha}\right]^{\frac{\sigma-1}{\alpha-(\sigma-1)}} dG(z) + \left(1+\tau_t^{1-\sigma}\right)^{\frac{\alpha}{\alpha-(\sigma-1)}} \int_{\tilde{z}_t}^{\infty} \left[z\phi_t(z)^{\alpha}\right]^{\frac{\sigma-1}{\alpha-(\sigma-1)}} dG(z)\right]\right\}^{-\frac{\alpha-(\sigma-1)}{\alpha}}, \tag{15}$$

where the constant $\Lambda_2 = \left(\frac{\sigma}{\sigma-1}\right)^{\sigma-1}\left[\frac{(\sigma-1)\tilde{\sigma}}{\alpha}\right]^{\frac{1-\sigma}{\alpha-(\sigma-1)}}$.

Equation (15) is the aggregate-demand (AD) condition which implies a positive relationship between $\tilde{z}_t$ and D_t. Intuitively, the increase in $\tilde{z}_t$ drives firms out of export markets, which raises P_t and therefore D_t.

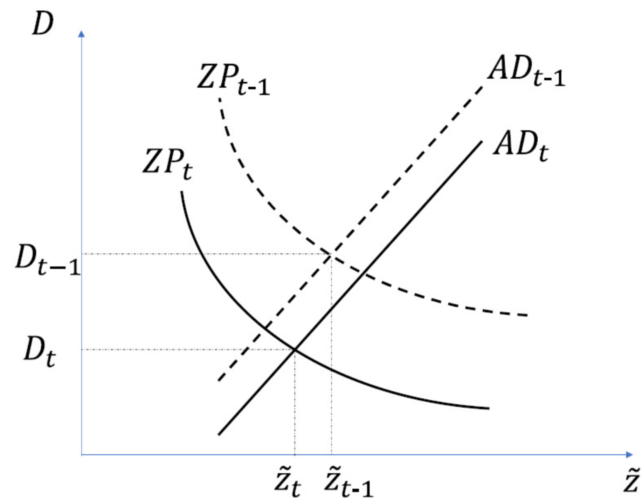


Figure 8: The Changes in Equilibrium Outcomes under $d\tau < 0$

The intersection of ZP and AD curves determines instantaneous equilibrium. Therefore, we have the following result:

Lemma 3 (Transitional Dynamics under Symmetric Trade Liberalization) *Consider two symmetric economies at a steady state before period t. Under symmetric trade liberalization at period t, there is a unique transitional path converging to the new steady state.*

We then investigate the impacts of trade liberalization on firm innovation and growth. The innovation of exporting firms can be given by

$$\kappa_t^*(z) = \left[\frac{(\sigma-1)\tilde{\sigma}}{\alpha}\left(1+\tau_t^{1-\sigma}\right)D_t\right]^{\frac{1}{\alpha-(\sigma-1)}}\left[z\phi_t(z)^{\sigma-1}\right]^{\frac{1}{\alpha-(\sigma-1)}}. \tag{16}$$

It is straightforward that holding the aggregate demand D_t constant, the innovation of exporting firms increases as τ_t falls. The decline in trade costs increases the market size of exporting firms and therefore their marginal benefits of innovation.

We are particularly interested in non-exporting firms due to their lack of direct access to exporting markets:

Proposition 4 (Static Innovation Effects of Trade Liberalization on Non-Exporters) *Consider two symmetric economies described above. If $z \leq \min\{\tilde{z}_t, \tilde{z}_{t-1}\}$, then $-\frac{d\kappa_t^*(z)}{d\tau} < 0$.*

Proposition 4 suggests that trade liberalization would instantaneously reduce the innovation of non-exporting firms. This is mainly because exporting firms bid up the wage. Notably, most of the firms do not export and they provide a large fraction of products consumed by domestic consumers. As a result, the negative innovation response of non-exporting firms could considerably increase the price index faced by domestic consumers and erode welfare gains from trade liberalization.

However, a firm's productivity growth depends not only on its own innovation, but also its learning from other firms. As a result, the dynamic productivity effect of trade liberalization relies crucially on the structure of inter-firm knowledge networks. This new innovation effect of trade liberalization via knowledge spillovers has not been discussed in the literature.

The following result demonstrates the importance of inter-firm knowledge networks in shaping the firms' productivity growth. Suppose that $z \leq \min\{\tilde{z}_t, \tilde{z}_{t-1}\}$. Then

$$\begin{aligned}
-\frac{\phi_{t+1}(z)^{\frac{1}{\beta}-1}}{\beta}\frac{d\phi_{t+1}(z)}{d\tau} = \quad & \underbrace{-\frac{1}{\alpha-(\sigma-1)}\kappa_t^*(z)\frac{d\log D_t}{d\tau}}_{\text{(1) The Impact of Competition on Own Innovation}} \\
& \underbrace{-\frac{\delta}{\alpha-(\sigma-1)}\left[\int_{S_z} m(z',z)\kappa_t^*(z')\phi_t(z')dG(z')\right]\frac{d\log D_t}{d\tau}}_{\text{(2) The Impact of Competition on Other firms' Innovation}} \\
& \underbrace{+\delta m(\tilde{z}_t,z)\phi_t(\tilde{z}_t)\left[\frac{(\sigma-1)\tilde{\sigma}}{\alpha}D_t\right]^{\frac{1}{\alpha-(\sigma-1)}}\left[\tilde{z}_t\phi_t\left(\tilde{z}_t\right)^{\sigma-1}\right]^{\frac{1}{\alpha-(\sigma-1)}}\left[\left(1+\tau^{1-\sigma}\right)^{\frac{1}{\alpha-(\sigma-1)}}-1\right]\frac{d\tilde{z}_t}{d\tau}}_{\text{(3) Learning from New Exporters}} \\
& \underbrace{+\frac{\delta}{\alpha-(\sigma-1)}\frac{(\sigma-1)\tau_t^{-\sigma}}{1+\tau_t^{1-\sigma}}\int_{\tilde{z}_t}^{\infty} m(z',z)\kappa_t^*(z')\phi_t(z')dG(z')}_{\text{(4) Learning from Incumbent Exporters}}.
\end{aligned} \tag{17}$$

In Equation (17), we decompose the impacts of trade liberalization on the productivity growth of a non-exporting firm into four terms. The first term summarizes the competition effect of trade liberalization on the non-exporter's own innovation, which has been characterized by Proposition 4. The second term reflects the competition effect of trade liberalization via inter-firm knowledge networks. The third and fourth terms reflects the diffusion of knowledge gains from trade liberalization. The third term in Equation (17) suggests that a non-exporter could learn from new exporters whose innovation incentives have been spurred by trade liberalization. Likewise, the fourth term in Equation (17) indicates the knowledge diffusion from incumbent exporters to non-exporters.

In sum, while trade liberalization only directly boosts the innovation of a small fraction of firms, their technology progress could benefit other firms via inter-firm knowledge networks. As a result, investigating the structure of inter-firm knowledge networks, in particular "who learns from whom", is key to understanding the dynamic effects of trade liberalization.

4 Estimating Inter-Firm Knowledge Networks

To quantify the implications of knowledge diffusion for welfare gains from trade liberalization, we estimate the structure of inter-firm knowledge networks using data on patent citations across Chinese manufacturing firms. Since our data on patent citations is only for Chinese firms, we base our estimation on the close economy equilibrium.

This section proceeds as follows. Section 4.1 proposes parametric assumptions for the matching function $m(z', z)$. Section 4.2 describes the estimation procedures, which are conducted sequentially in Section 4.3 and 4.4. Section 4.6 discusses the estimation results and examines the model fit.

4.1 Parametric Assumptions

The data on patent citations suggests that firms' positions in the knowledge network depend closely on their sizes. Guided by these patterns, we parameterize the distribution of innovation efficiency, $G(z)$, and the matching function $m(z', z)$ as follows. We first assume that the innovation efficiency z are log-normally distributed, with the mean μ_z and variance σ_z^2. In our structural estimation, we normalize $\mu_z = 0$. This log-normality assumption is consistent with the literature suggesting that the major part of firm size distribution, except for the tail, can be well-approximated by the log-normal distribution. Since our empirical regularities in Section 2 have shown that most of the firms have very few knowledge connections whereas few giant firms have many connections, our model will yield a log-normal firm size distribution, except for the tail.

Then we parameterize $m(z', z)$ as

$$m(z', z) = \frac{\gamma}{1 + \exp\{-\left[\xi_1 \log z' + \xi_2 \log z + \rho \log z' \log z\right]\}}, \quad \gamma \in [0, 1]. \tag{18}$$

There are four parameters in this matching function. First, $\gamma \in [0, 1]$ characterizes the average matching rate across firms. Second, ξ_1 and ξ_2 characterize how matching rates vary, respectively, with respect to innovation efficiency of the cited firm z' and the citing firm z. Finally, ρ characterizes the assortativity of inter-firm knowledge networks, i.e. whether larger and more connected firms are connected with firms that are also larger and more connected. Notably, we have

$$\frac{\partial^2 \log m(z',z)}{\partial \log z \partial \log z'} = \left[\rho + \frac{(\xi_1 + \rho \log z)(\xi_2 + \rho \log z')}{m(z',z)/\gamma}\right] \frac{\exp\{-[\xi_1 \log z' + \xi_2 \log z + \rho \log z' \log z]\}}{m(z',z)/\gamma}. \tag{19}$$

Therefore, $m(z', z)$ is log-supermodular if $\rho > 0$ and $\xi_1, \xi_2 > 0$.

4.2 Estimation Procedure

We estimate parameters for $G(z)$ and $m(z', z)$ using data on sales and patent citations of Chinese manufacturers. We start by discussing the parameters that are not estimated from data. First, we set the value of the elasticity of substitution σ to 4, which is close to the estimates in the literature. Second, we normalize L so that $\tilde{D} = 1$. Third, we set $\alpha = 15$ so that the net profit share is equal to $\frac{1}{\sigma}\left(1 - \frac{\sigma-1}{\alpha}\right) = 0.2$. Finally, we set the curvature of productivity evolution $\beta = 0.72$. Without knowledge diffusion, this leads to $\frac{\partial \log(\phi_{t+1}(z))}{\partial \log(\phi_t(z))} = \frac{\alpha\beta}{\alpha-(\sigma-1)} = 0.9$, which is close to the estimates in the literature.[7]

The remaining parameters, $(\gamma, \xi_1, \xi_2, \rho)$ and (δ, σ_z^2), are then estimated using data on inter-firm citation linkages and firms' performances. We first estimate (δ, σ_z^2) from the observed firm sales distribution and citation linkages by a maximum likelihood estimator. Then given the estimates on (δ, σ_z^2), we estimate parameters of matching function, $(\gamma, \xi_1, \xi_2, \rho)$, from citation linkages by the simulated method of moments with equilibrium conditions as constraints.

4.3 Maximum Likelihood Estimator on (δ, σ_z^2)

In this subsection, we estimate (δ, σ_z^2) using data on firms' long-term average sales and citation linkages $\{\mathbf{1}\,[\text{firm j cites from firm i}]\}$. Through the lens of our model, the firm's long-term average sales can be expressed as $x_i = \left(z_i^{\frac{1}{\alpha-(\sigma-1)}} \phi_i^{\frac{\alpha}{\alpha-(\sigma-1)}}\right)^{\sigma-1}$. Then we can recover the innovation efficiency of firm j directly by combining the sales equation with Equation (11):

$$z_j(\delta) = \frac{\left(x_j^{\frac{1}{\sigma-1}}\right)^{\alpha-(\sigma-1)}}{\left[x_j^{\frac{1}{\sigma-1}} + \delta \sum_{i=1}^{R} m_{ij} x_i^{\frac{1}{\sigma-1}}\right]^{\alpha\beta}}, \tag{20}$$

[7] See, for example, Roberts, et al. (2011).

where the empirical matching rate is constructed by $m_{ij} := \frac{\mathbf{1}[\text{firm j cites from firm i}]}{R}$.

Equation (20) provides identification for δ. If $\delta = 0$, then $\{x_j\}$ should be log-normally distributed, as we have assumed for $\{z_j\}$. The extent to which the observed $\{x_j\}$ deviate from log-normal distribution identifies the magnitude of inter-firm knowledge spillover, δ. More specifically, let $K_j(\delta) = \log z_j(\delta) - \frac{1}{S}\sum_{i=1}^{S} \log z_i(\delta)$. Under the assumption that z is log-normally distributed, $K_j(\delta) \sim N(0, \sigma_z^2)$. Therefore, (δ, σ_z^2) can be estimated by the maximum likelihood estimator (MLE):

$$\max_{(\delta,\sigma_z^2)} \ell(\delta, \sigma_z^2; \{x_i, m_{ij}\}) = -\frac{S}{2}\log\left(\sigma_z^2\right) - \frac{1}{2\sigma_z^2}\sum_{j=1}^{S} K_j(\delta)^2. \tag{21}$$

Table 2: Estimates on (δ, σ_z^2)

Parameter		Value	Standard Error
Magnitude of spillover	δ	3.118	.204
Variance of z	σ_z^2	.130	.001

(Notes: the standard errors are estimated based on the asymptotics of extreme estimator.)

The MLE estimates on (δ, σ_z^2) are shown in Table 2. The estimate on δ is sizable and significantly positive. This result indicates that the observed distribution of long-term sales $\{x_j\}$ substantially deviate from log-normal distribution. In other words, although the innovation efficiency z_j is assumed to be log-normal, the resulting productivity and sales deviate from log-normality because of the heterogeneous inter-firm knowledge linkages.

4.4 Constrained Simulated Method of Moments on $(\gamma, \xi_1, \xi_2, \rho)$

In this subsection, we estimate parameters of matching function, $(\gamma, \xi_1, \xi_2, \rho)$, by the following constrained simulated method of moments. We draw N observations $\{U_i\}_{i=1}^{N}$ independently from $N(0, \sigma_z^2)$ and compute $z_i = \exp\{U_i\}$. Then the simulated sales $\{x_i\}_{i=1}^{N}$ can be computed by

$$x_j^{\frac{1}{\sigma-1}} = z_j^{\frac{1}{\alpha-(\sigma-1)}}\left[x_j^{\frac{1}{\sigma-1}} + \delta\sum_{i=1}^{N} m(z_i, z_j; \gamma, \xi_1, \xi_2, \rho)x_i^{\frac{1}{\sigma-1}}\right]^{\frac{\alpha\beta}{\alpha-(\sigma-1)}}. \tag{22}$$

As described above, we observe firm j citing patents from firm i. We thereby can compute firms' in-degree and out-degree. Our simulation can generate the corresponding statistics as:

$$tm_j = \sum_{i=1}^{N} m(z_i, z_j), \quad tm_i = \sum_{j=1}^{N} m(z_i, z_j). \tag{23}$$

Moreover, we can compute the average sales of firms from which firm j cites and the average sales of firms citing firm i. The corresponding simulated statistics can be computed by

$$am_j = \frac{\sum_{i=1}^{N} m(z_i, z_j) x_i}{\sum_{i=1}^{N} m(z_i, z_j)}, \quad am_i = \frac{\sum_{j=1}^{N} m(z_i, z_j) x_j}{\sum_{j=1}^{N} m(z_i, z_j)}. \tag{24}$$

Table 3: Simulated Moments

	Simulated Moment	Data Moments	Simulation Result
	(1)	(2)	(3)
$sm(.)$	(i) Slope of regressing $\log(tm_j)$ on $\log(x_j)$	.223	.2234
	(ii) Slope of regressing $\log(tm_i)$ on $\log(x_i)$	.311	.3075
	(iii) Slope of regressing $\log(am_j)$ on $\log(x_j)$	.180	.1542
	(iv) Slope of regressing $\log(am_i)$ on $\log(x_i)$	.140	.1645
	$p75(\log x_i)/p50(\log x_i)$	1.0472	1.0472

Our targeted moments are summarized in Column (1) of Table 3. Column (2) of Table 3 presents the values of our targeted moments in the data. The first four moments in $sm(.)$ are used to identify (ξ_1, ξ_2, ρ). In particular, the first and second moments characterize, respectively, how in- and out-degree vary with respect to firm size, which aim at identifying (ξ_1, ξ_2). The third and fourth moments, instead, characterize to what extent large firms cite and are cited by large firms. These two moments are set to identify ρ.

The fifth moment in Table 3 is used to identify γ, since if $\gamma = 0$ then $\log x_i$ is normally distributed and thereby $p75(\log x_i)/p50(\log x_i) = \infty$.

Given the simulated $\{z_i\}_{i=1}^{N}$, the constrained simulated method of moments can be expressed as

$$\min_{\gamma, \xi_1, \xi_2, \rho, \{x_i\}_{i=1}^{N}} sm(\gamma, \xi_1, \xi_2, \rho, \{x_i\})' \Omega sm(\gamma, \xi_1, \xi_2, \rho, \{x_i\}),$$

s.t.

$$x_j^{\frac{1}{\sigma-1}} = z_j^{\frac{1}{\alpha-(\sigma-1)}} \left[x_j^{\frac{1}{\sigma-1}} + \delta \sum_{i=1}^{N} m(z_i, z_j; \gamma, \xi_1, \xi_2, \rho) x_i^{\frac{1}{\sigma-1}} \right]^{\frac{\alpha\beta}{\alpha-(\sigma-1)}}, \quad \forall j = 1, \ldots, N, \tag{25}$$

$$\frac{p75(\log x_i)}{p50(\log x_i)} = 1.0472,$$

where Ω is a positive definite weighting matrix.

We set $N = 100$ and use identity matrix as the weighting matrix. Let G be the Jacobian matrix of $sm(.)$ with respect to parameters and V_m be the variance-covariance matrix of the moments. Then by the property of extreme estimator, the variance-covariance matrix of the estimated parameters can be given by $\frac{1}{B}(G'G)^{-1}G'V_mG(G'G)^{-1}$ where V_m is computed by bootstrapping and B is the number of repetitions for bootstrapping.

Parameter		Value	Standard Error
Level of Matching Rate	γ	.2204	.103
Marginal effect of z'	ξ_1	1.2965	.1188
Marginal effect of z	ξ_2	.9711	.0879
Cross effect	ρ	4.2067	.1441

Table 4: Estimates on the Matching Function $m(z', z)$

Table 4 shows the estimation results for $(\gamma, \xi_1, \xi_2, \rho)$. Both ξ_1 and ξ_2 are significantly positive, confirming the empirical regularity that larger firms cite more and are cited by more firms. Moreover, ρ is positive and sizable, suggesting strong positive matching across firms. Based on Equation (19), our estimates on (ξ_1, ξ_2, ρ) suggest that large firms are well-connected with each other and lie at the center of inter-firm knowledge networks, whereas small firms can hardly get connected.

4.5 Sensitivity Matrix

How do our estimates depend on our estimation moments? Andrews et al. (2017) develop a sensitivity matrix that measures the dependence of estimates on moments. Following their methodology we first compute the Jacobian matrix of our simulated moments $sm(.)$ in Equation (25) with parameters (ξ_1, ξ_2, ρ), denoted as $\tilde{G}$. Then our estimates $\hat{\theta} := (\hat{\xi}_1, \hat{\xi}_2, \hat{\rho})$ has first-order asymptotic bias:

$$E(\tilde{\theta}) = \tilde{\Lambda}E(sm), \quad \tilde{\Lambda} = -\left(\tilde{G}'\Omega\tilde{G}\right)^{-1}\tilde{G}'\Omega. \tag{26}$$

The results are presented in Table 5. It suggests that our estimates of $\hat{\xi}_1$ and $\hat{\xi}_2$ strongly positively relate to data moments (i) and (ii) in Table 3, whereas $\hat{\rho}$ is sensitive to data moments (iii) and (iv) in Table 3. These results confirm that the positive matching assortativity in the patent citation data is crucial for identifying parameters of inter-firm knowledge networks.

	Moments			
	(i)	(ii)	(iii)	(iv)
$\hat{\xi}_1$	6.68	3.78	1.83	2.37
$\hat{\xi}_2$	8.88	-0.007	1.63	1.61
$\hat{\rho}$	12.56	1.99	14.96	16.23

Table 5: Sensitivity Matrix of Estimates with Simulated Moments

4.6 Model Fit

We have estimated four of our model's key parameters targeting on five moments listed in Table 3. As shown in Table 3, our model matches the targeted moments quite well. Our simulation generates strong positive correlation between firms' sales and their in-/out-degrees, which approximates the data tightly. Our model also generates strong positive correlation between firms' sales and the average sales of firms they cite from/are cited.

Moreover, our simulation exactly replicates the ratio of the 75th percentile of log sales over the median of log sales in the data. This result provides further evidence suggesting that although the exogenous innovation efficiency is assumed to be log-normal, the resulting firm sales distribution substantially deviates from log-normality because firms occupy heterogeneous positions in knowledge networks. Figure 9 shows the model's fit of the firm sales distribution. The model generates reasonable good approximation to the empirical firm size distribution.

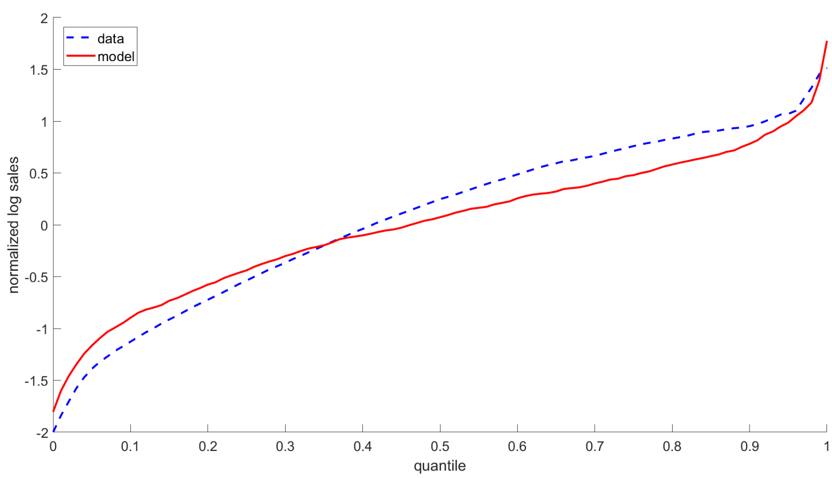
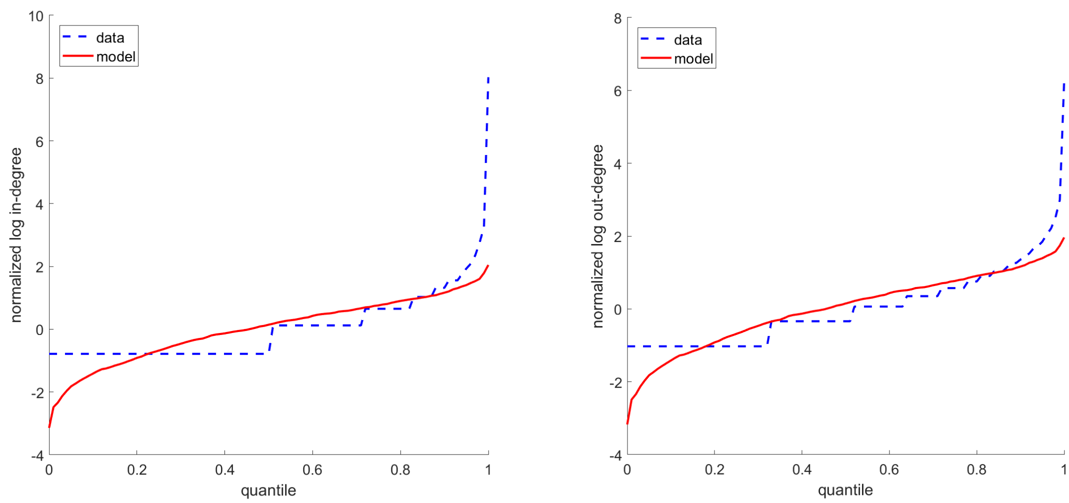


Figure 9: Firm Sales Distribution

(Notes: log sales are normalized by subtracting their mean and then dividing their standard deviation.)

Figure 10 shows that our model matches well the shape of normalized log degree distributions, although it underpredicts the connectivity of the firms that are most connected in the knowledge networks.



(a) In-degree (b) Out-degree

Figure 10: Firm Degree Distribution

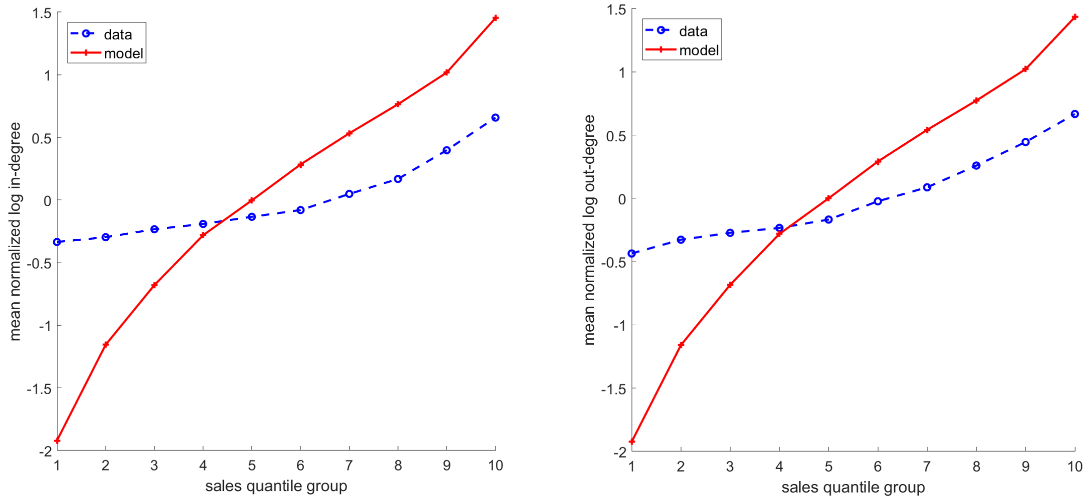
(Notes: log in-degrees and out-degrees are normalized by subtracting their mean and then dividing their standard deviation.)

Figure 11 shows the model's fit of the correlation between firm sales and degree. While the model over-predicts the connectivity of the largest firms, it is nonetheless consistent with the empirical pattern that larger firms tend to cite from and be cited by more firms.

Finally, Figure 12 and 13 illustrate the model's fit of the matching assortativity, which characterizes whether larger and more connected firms are connected to firms that are also larger and more connected (positive matching), or to firms that are smaller and less connected (negative matching). Figure 12 shows that larger firms indeed cite from larger firms and are cited by larger firms. Figure 13 shows that firms that cite from more firms tend to cite from firms that cite from more firms themselves. Similarly, firms that are cited by more firms tend to be cited by firms that are cited by more firms themselves. Therefore, the model replicates positive matching, both in terms of sales and degree, in the data.

5 Quantification

We apply our model to quantify the implications of inter-firm knowledge networks for the welfare gains from trade liberalization. We start by considering trade liberalization under



(a) In-degree

(b) Out-degree

Figure 11: Firm Sales and Degree

(Notes: log in-degrees and out-degrees are normalized by subtracting their mean and then dividing their standard deviation. Sales quantile group 1 refers to sales between the 0th and 10th percentile. Similarly, Sales quantile group 10 refers to sales between the 90th and 100th percentile.)

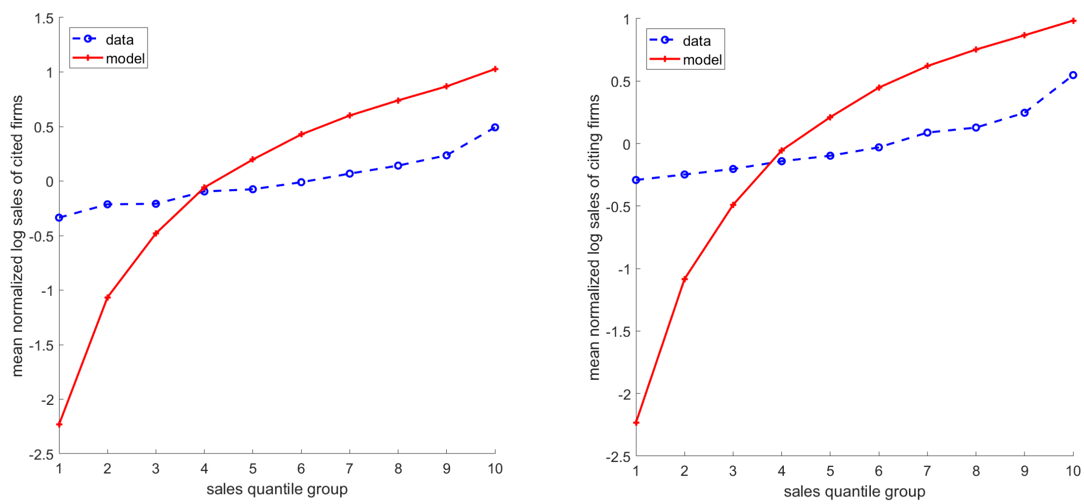


Figure 12: Firm Matching Assortativity (Sales)

(Notes: log sales of cited and citing firms are normalized by subtracting their mean and then dividing their standard deviation. Sales quantile group 1 refers to sales between the 0th and 10th percentile. Similarly, Sales quantile group 10 refers to sales between the 90th and 100th percentile.)

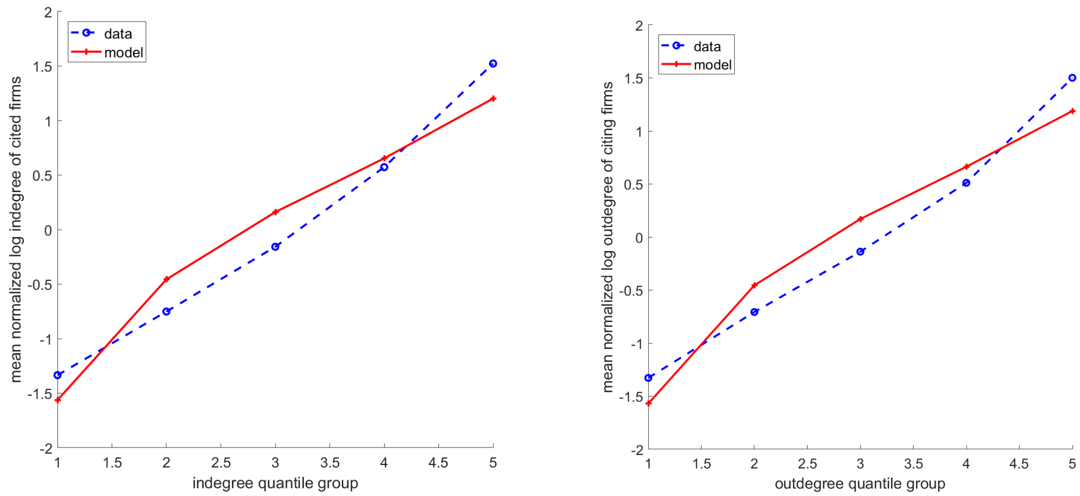


Figure 13: Firm Matching Assortativity (Degree)

(Notes: log in-degrees and out-degrees are normalized by subtracting their mean and then dividing their standard deviation. Sales quantile group 1 refers to sales between the 0th and 20th percentile. Similarly, Sales quantile group 5 refers to sales between the 80th and 100th percentile.)

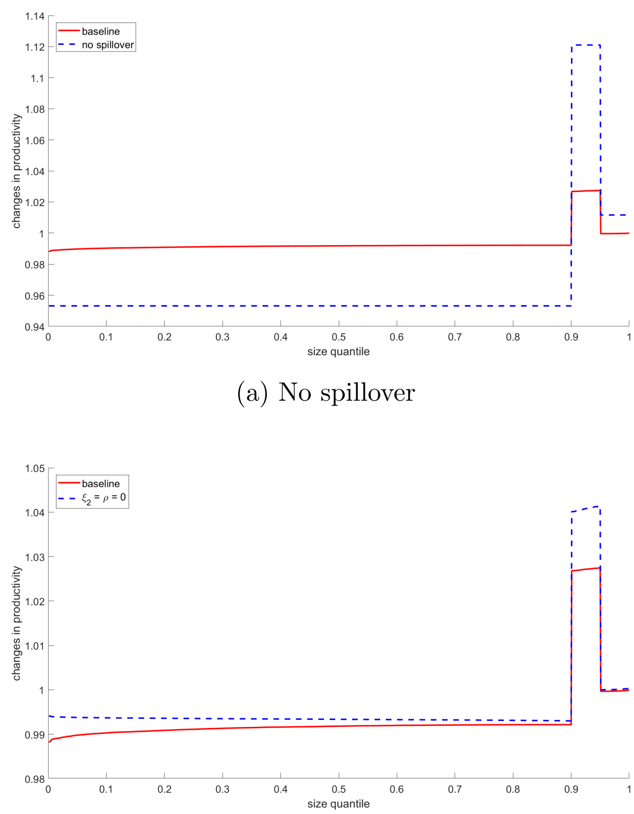
two symmetric countries as in Melitz (2003), to understand the mechanism through which inter-firm knowledge diffusion affects welfare effects of trade liberalization. Then we calibrate our dynamic model to the world that consists of China and the rest of the world (ROW) and quantify the importance of inter-firm knowledge networks to welfare gains from trade liberalization over 2001-2006.

5.1 Trade Liberalization under Two Symmetric Countries

Consider two countries $\ell = 1, 2$ with $L_{\ell,t} = 10$, $\mu_{z,\ell} = 0$, and $f^X_{\ell n,t} = 1.4$ for all (ℓ, t) and $\ell \neq n$. We embed time-invariant parameters estimated in Section 4. The world is in a steady-state initially. We assume that $\tau_{\ell n,t} = \tau_t$ for all t and $\ell \neq n$. We choose τ_t so that in the initial steady-state 5 percent of firms export.

Now we consider a symmetric trade liberalization: reducing τ_t permanently so that in the new steady-state the share of exporters is 10%. We then compute firm productivities in the new steady-state and compare them with those in the initial steady-state. Relative changes in firm productivities are illustrated by red lines in Figure 14. The results suggest that trade liberalization substantially expands the market shares of new exporters and therefore promotes their innovation and productivities. For incumbent exporters, the productivity improvements are negligible. For non-exporters, trade liberalization bids the wage up and reduces their innovation incentives.

To isolate the role of inter-firm knowledge networks, we set $\delta = 0$ and re-conduct the trade liberalization that increases the steady-state share of exporters from 5% to 10%. Relative changes in firm productivities are illustrated by the blue line in Panel (a) of Figure 14. The results suggest that without knowledge spillovers, trade liberalization reduces the non-exporters' steady-state productivities by much larger extent. This is consistent with Term (3) and (4) in Equation (17): in the presence of inter-firm knowledge networks, non-exporters could learn from new exporters whose innovation is dramatically promoted by trade liberalization. As a result, inter-firm knowledge networks considerably mitigate the innovation reallocation effect of trade liberalization emphasized by Aghion et al. (2018).



(a) No spillover

(b) $\xi_2 = \rho = 0$

Figure 14: Productivity Effects of Trade Liberalization: the Steady-State

(Note: we fix $f^X_{\ell n,t} = 1.4$ for all t and $\ell \neq n$.)

We further investigate the implications of the structure of inter-firm knowledge networks. We depart from our baseline case by setting $\xi_2 = \rho = 0$, i.e. all firms have the same capability to learn from other firms. Panel (b) of Figure 14 suggests that in this case, non-exporters could benefit more from exporters' knowledge spillovers. Consequently, the extent to which small non-exporters could learn from exporters is quantitatively important to the productivity and welfare effects of trade liberalization.[8]

5.2 Calibrating Trade Costs

In this subsection, we calibrate changes in trade costs to replicate the observed changes in trade shares between China and the rest of the world over 2001-2006. In our baseline calibration, we take our estimates of σ_z^2 and $m(z', z)$ from Section 4. We assume that (i) the economy is in the steady-state in the year of 2001, and (ii) country size (L_ℓ), average innovation efficiency ($\mu_{z,\ell}$), and the fixed trade cost $\left(f_{\ell n}^X\right)$ are all time-invariant for $\ell, n \in \{\text{CHN}, \text{ROW}\}$. In other words, the only exogenous shock considered in this counterfactual exercise is the change in iceberg trade costs. The targeted moments and the calibrated parameters are shown in Table 6.

Parameter Values	Definition	Target Moments
$L_{\text{CHN}} = 20$, $L_{\text{ROW}} = 77.5$	Population	Population in World Development Index (2001)
$\mu_{z,\text{CHN}} = 0$, $\mu_{z,\text{ROW}} = 0.18$	Mean of innovation efficiency	Chinese TFP = $\frac{2}{3}$ of the ROW TFP in 2001
$f_{\text{CHN}}^X = 1$, $f_{\text{ROW}}^X = 6.25$	Fixed trade costs	Exporter Shares in 2001: 20% in China and 10% in the ROW
$(\tau_{\text{CHN,ROW}}, \tau_{\text{ROW,CHN}})$	Iceberg trade costs over 2001-2006	Export Shares over 2001-2006

Table 6: Targeted Moments and Calibrated Parameters

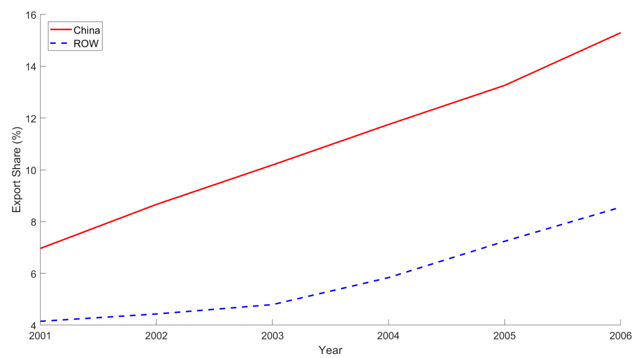
The calibration of iceberg trade costs over 2001-2006 is illustrated by Figure 15. It suggests that the iceberg trade costs between China and the rest of the world decrease steadily after China's accession to the WTO. The cumulative decrease in Chinese export costs is about 15% over this period.

5.3 Welfare Implications of China's Trade Liberalization

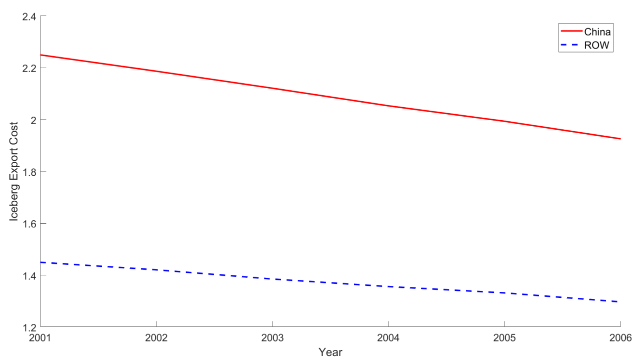
Armed with the calibrated model, we isolate and quantify the impacts of trade liberalization between China and the rest of the world over 2001-2006, highlighting the role of inter-firm knowledge networks. To this end, we start from the steady-state in 2001, embedding changes in trade costs calibrated in Section 5.2 and computing the entire transitional path over 2001-2006. To quantify the effects of trade liberalization, we compare the transitional path with the counterfactual path in which trade costs are fixed to their 2001 levels.

For each year between 2001 and 2006, we calculate the time-discounted total real income from 2001 to that year, using time-discounting factor 0.97, and compare it to the time-discounted total real income without changes in trade costs. The red line in Figure 16

[8] The impacts of having $\xi_1 = \rho = 0$ are much smaller. The results are presented in the appendix.



(a) Export Shares



(b) Calibrated Iceberg Trade Costs

Figure 15: Export Shares and Calibrated Iceberg Trade Costs

(Notes: Export Share is defined as the export value as a share of total expenditure.)

summarizes the result. It suggests that the Chinese welfare gains from trade liberalization gradually accumulated to 1.8% in 2006. This result is in line with the estimates of the Chinese gains from trade liberalization in the literature.

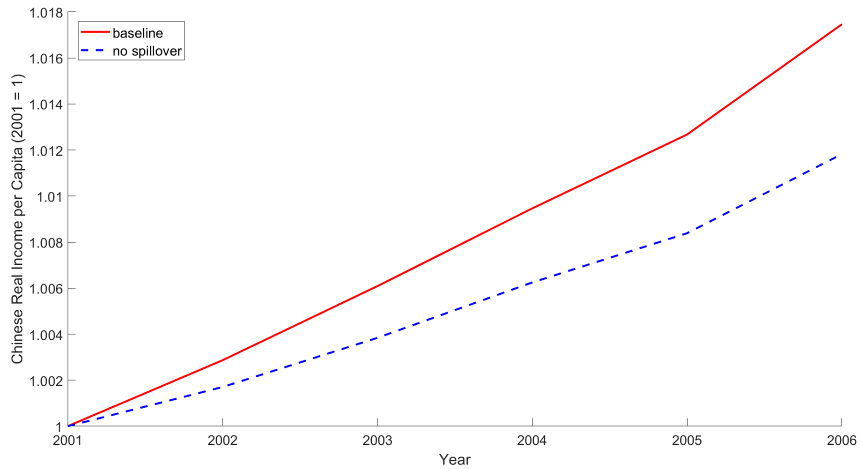


Figure 16: Welfare Gains from Trade Liberalization: Baseline vs. No Spillover

(Notes: In the "no spillover" scenario, we fix all parameters estimated and calibrated in the baseline case, except having $\delta = 0$. Each point represents the accumulated changes in time-discounted real income up to that year. We set time discounting factor as 0.97.)

To highlight the quantitative importance of inter-firm knowledge networks, we compute the Chinese gains from trade liberalization in the model with $\delta = 0$. Illustrated by the blue line in Figure 16, the Chinese welfare gains from trade liberalization accumulated to 1.2% in 2006. As a result, inter-firm knowledge networks account for about one third of the Chinese welfare gains from trade liberalization over 2001-2006. In the presence of inter-firm knowledge networks, non-exporters could improve their technologies by learning from exporters whose innovation activities are promoted by trade liberalization. Our counterfactual exercises show that this indirect productivity effect is quantitatively important to understanding welfare gains from trade liberalization.

6 Conclusion

Our model aims at characterizing firms' heterogeneous innovation responses to export liberalization and drawing aggregate implications. By emphasizing the role of inter-firm knowledge networks, we propose a novel indirect gain from export liberalization: non-exporters could adopt technologies from exporters. We utilize the unique data on patent citations across Chinese manufacturing firms as a proxy for inter-firm knowledge networks and structurally estimate our model. Our estimated model is able to replicate rich heterogeneity in inter-firm knowledge diffusion observed in the data. Simulations of our model suggest that inter-firm knowledge networks could mitigate the reallocation effect between exporters and non-exporters emphasized by Melitz (2003) and Aghion et al. (2018), and therefore increase welfare gains from export liberalization.

In this paper, we regard inter-firm knowledge networks as exogenous and parameterize the matching function flexibly to match the data. One important direction for the future exploration is to understand the formation of inter-firm knowledge networks. However, a message from this paper is that any mechanism aiming at rationalizing inter-firm knowledge networks should be able to reproduce key features in the patent citation data such as strong positive assortativity.

References

[1] Acemoglu, Daron, Ufuk Akcigit, and William Kerr (2016) "Innovation Network" *Proceedings of the National Academy of Sciences*, 113 (41): 11483-11488.

[2] Acemoglu, Daron, Pablo D. Azar (2017) "Endogenous Production Networks" *NBER Working Paper*, No.24116.

[3] Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza TahbazSalehi (2012) "The Network Origins of Aggregate Fluctuations" *Econometrica*, 80(5):1977-2016.

[4] Aghion, Philippe, Antonin Bergeaud, Matthieu Lequien, and Marc J. Melitz (2018) "The Impact of Exports on Innovation: Theory and Evidence" *NBER Working Paper*, No. 24600.

[5] Alvarez, Fernando,Francisco J. Buera, and Robert E. Lucas (2014) "Idea Flows, Economic Growth, and Trade" *NBER Working Paper*, No.19667.

[6] Andrews, Isaiah, Matthew Gentzkow, and Jesse M. Shapiro (2017) "Measuring the Sensitivity of Parameter Estimates to Estimation Moments" *The Quarterly Journal of Economics*, 132 (4): 1553-1592.

[7] Arkolakis, Costas, Arnaud Costinot, and Andres Rodriguez-Clare (2011) "New Trade Models, Same Old Gains?" *American Economic Review*, 102 (1): 94-130.

[8] Autor, David, David Dorn , and Gordon H. Hanson (2013) "The China syndrome: Local labor market impacts of import competition in the united states" *American Economic Review*, 103 (6): 2121-2168.

[9] Bernard, Andrew, Andreas Moxnes, and Yukiko U. Saito (2019) "Production Networks, Geography and Firm Performance" *Journal of Political Economy*, 127 (2):639-688.

[10] Bernard, Andrew, Andreas Moxnes, and Karen Helene Ulltveit-Moe (2018) "Two-sided Heterogeneity and Trade" *The Review of Economics and Statistics*,100(3):424-439.

[11] Bloom, Nicholas, Mark Schankerman,and John Van Reenen (2013) "Identifying Technology Spillovers and Product Market Rivalry" *Econometrica*, 81(4): 1347-1393.

[12] Borusyak, Kirill, Peter Hull, and Xavier Jaravel (2018) "Quasi-experimental shift-share research designs" *NBER Working Paper*, No.24997.

[13] Buera, Francisco J. , Ezra Oberfield (2016) "The Global Diffusion of Ideas" *NBER Working Paper*, No. 21844.

[14] Cai, Jie, Nan Li (2018) "Growth Through Inter-sectoral Knowledge Linkages" *The Review of Economic Studies*, 86(5):18271866.

[15] Cohen, Wesley M, Steven Klepper (1996) "Firm Size and the Nature of Innovation within Industries: the Case of Process and Product R&D" *The Review of Economics and Statistics*, 78(2):232-243.

[16] Costinot, Arnaud, Andres Rodriguez-Clare (2014) "Trade Theory with Numbers: Quantifying the Consequences of Globalization" *Handbook of International Economics*, vol.4.

[17] Desmet, Klaus, Dvid Krisztin Nagy ,and Esteban Rossi-Hansberg (2018) "The Geography of Development" *Journal of Political Economy*, 126(3):903-983.

[18] Desmet, Klaus, Esteban Rossi-Hansberg (2014) "Spatial Development" *American Economic Review*, 104(4):1211-1243.

[19] Giuliani, Elisa (2007) "The selective nature of knowledge networks in clusters: Evidence from the wine industry" *Journal of Economic Geography*, 7(2):139-168.

[20] Hulten, Charles R. (1978) "Growth accounting with intermediate inputs" *The Review of Economic Studies*, 45(3): 511-518.

[21] Hummels, David, Rasmus Jorgensen, Jakob Munch ,and Chong Xiang (2014) "The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data" *American Economic Review*, 104(6) :1597-1629.

[22] Jaffe, Adam B. , Manuel Trajtenberg ,and Michael S Fogarty (2000) "Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors" *American Economic Review, Papers and Proceedings*, 90(2):215-218.

[23] Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson (1993) "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations" *The Quarterly Journal of Economics*, 108(3):577598.

[24] Klette, J. Jakob, Samuel Kortum (2004) "Innovating Firms and Aggregate Innovation" *Journal of Political Economy*, 112(5): 986-1018.

[25] Lim, Kevin. (2017) “Firm-to-firm Trade in Sticky Production Networks” *Society for Economic Dynamics 2017 Meeting Papers*,280.

[26] Melitz, Marc. (2003). “The impact of trade on intra-industry reallocations and aggregate industry productivity”. *Econometrica*, 71(6): 1695-1725.

[27] Oberfield, Ezra. (2018) “A Theory of Input-Output Architecture” *Econometrica*, 86(2):559-589.

[28] Tintelnot,Felix, Ayumu Ken Kikkawa, Magne Mogstad, and Emmanuel Dhyne (2018) “Trade and Domestic Production Networks” *NBER Working Paper*, No. 25120.

Appendix A Data

A.1 Data Sources

A.1.1 Annual Survey of Industrial Firms (ASIF)

The main data set used in this study comes from the Annual Survey of Industrial Firms (ASIF), conducted by the National Bureau of Statistics of China from 1998 to 2013. This is the most comprehensive firm-level data set in China, as it covers all state-owned enterprises and all non-state-owned enterprises with annual sales above five million Renminbi (around US$650,000). The number of firms varies from more than 140,000 in the late 1990s to more than 243,000 in 2007. The data set spans all 31 provinces or province-equivalent municipalities, and all manufacturing industries, which ensures its invaluable national representativeness.

The data set provides detailed firm information, including industry affiliation, location, and all operation and performance items from the accounting statements, such as exports, book value and net value of fixed assets, employment and wage rate. We depreciate all pecuniary variables with 2-digit price deflators constructed by Brandt et al.(2012). However, one drawback of this data set is that it does not directly provide information on fixed investment. To obtain data on fixed investment, we follow Song and Wu (2012) in using book values of fixed assets(reported in the ASIF data set) and assuming a constant depreciation rate of 5%.

A.1.2 Patent Records

The second data source is a comprehensive dataset of patents granted at the Chinese National Intellectual Property Administration (CNIPA). Similar to the patent data provided by the United States Patent and Trademark Office(USPTO), the CNIPA dataset contains detailed information on each patent filing since 1985, including date of filing, official name and address of the applicant, name of the patent, and type of patent classified according to Chinas Patent Law, i.e., whether the application is for an invention patent, a utility model patent, or a design patent. Some remarks on the use of CNIPA data are in order. In general, measuring innovation activities is difficult. The report of OECD (2009) provides a good description of the advantages and drawbacks of using patent as a measure of innovation. In addition to patent filing data, other types of data to measure innovations exist, but some are almost impossible to obtain while some are even less satisfactory than patent filing.

First, R&D expenditure measures innovation input, but R&D data is available only for years 20012003 and 20052007, and thus is not good for our DID estimation. Second, Chinas patent filing and/or granting abroad is a good alternative because it may reflect more genius innovation, as argued by Holmes et al. (2015). Such data can be found in WIPO dataset. However, linking the WIPO data to the NBS data is almost impossible.[9]

A.1.3 Patent Citations

Our third data source is Google Patent database.[10] Google Patents was launched in 2006. It covers patent information from 17 patent offices around the world. In the case of Chinese patents, information is extracted from China's State Intellectual Property Office (CNIPA) and is translated into English. Google Patents not only provides information on names of inventor(s), current and original assignee(s), priority and publication date, application date and number, the country where the patent was granted and the status of the patent, but also lists detailed forward and backward citation.

A.2 Merging ASIF with Patent Records

We merge ASIF with patent records by the following steps:

1. Pre-processing the firm and assignee names to get a standardized expression, where we dealing with the punctuations, full-width letters and Chinese numbers;
2. Identifying and removing various firm-form designators to get stem names;
3. Conducting an exact matching based on standardized stem names, regardless of temporal information ;
4. Conducting manual check matching quality.

In the end, we get nearly 3 million matches between firms and patent applications, involving about 100 thousand ASIF firms and 2.7 million patents. For comparison, He et al. (2018) results in 653,360 matches from 1998-2009 and our result covers over 91.67% of them.

[9]The two datasets have no common identifier; WIPO contains English names only while NBS contains Chinese names only.

[10]See https://patents.google.com/

A.3 Descriptive Statistics for Our Constructed Database

In this subsection, we present descriptive statistics for the database constructed in Section 2.1.

Appendix B Theory

B.1 Proof to Proposition 2

Proof. In the steady-state, we have $\phi_t(z) = \phi(z)$ for all t. Inserting Equation (3) into Equation (6), we have

$$\phi(z) = \tilde{D}^{\beta}\left[z^{\frac{1}{\alpha-(\sigma-1)}}\phi(z)^{\frac{\alpha}{\alpha-(\sigma-1)}} + \delta\int_{S_z} m(z', z)\,(z')^{\frac{1}{\alpha-(\sigma-1)}}\,\phi(z')^{\frac{\alpha}{\alpha-(\sigma-1)}}dG(z')\right]^{\beta}, \quad \forall z \in S_z. \tag{27}$$

Rearranging Equation (27), we have

$$\phi(z)^{\frac{1}{\beta}} = \int_{S_z}\tilde{D}\left[\mathbf{1}(z'=z) + \delta m(z', z)\right](z')^{\frac{1}{\alpha-(\sigma-1)}}\,\phi(z')^{\frac{\alpha}{\alpha-(\sigma-1)}}dG(z'). \tag{28}$$

Let $\tilde{\phi}(z) = \left[\tilde{D}^{\frac{1}{\frac{\alpha\beta}{\alpha-(\sigma-1)}-1}}\phi(z)\right]^{\frac{1}{\beta}}$.

$$\tilde{\phi}(z) = \int_{S_z}\left[\mathbf{1}(z'=z) + \delta m(z', z)\right](z')^{\frac{1}{\alpha-(\sigma-1)}}\,\tilde{\phi}(z')^{\frac{\alpha\beta}{\alpha-(\sigma-1)}}dG(z'). \tag{29}$$

Since $1 - \frac{\alpha\beta}{\alpha-(\sigma-1)} > 0$, by Theorem 1 of Allen, Arkolakis, and Li (2017), there exists a unique solution to Equation (29) and the solution can be computed by a simple iteration procedure. Notice that price index P can be computed by Equation (10) and the total expenditure X is exogenous. Given the unique $\{\tilde{\phi}(z)\}$, the price index P is unique and so is $\{\phi(z)\}$. Therefore, our steady-state equilibrium is unique. ■

B.2 Two Symmetric Countries

Under two symmetric countries, we take wage as the numeraire. Firm z exports at period t if and only if

$$\tilde{\Lambda}_1\left[z^{\sigma-1}\phi_t(z)^{\alpha(\sigma-1)}\right]^{\frac{1}{\alpha-(\sigma-1)}}\left[\left(1+\tau_t^{1-\sigma}\right)^{\frac{\alpha}{\alpha-(\sigma-1)}} - 1\right]D_t^{\frac{\alpha}{\alpha-(\sigma-1)}} \geq f_t^X. \tag{30}$$

Suppose that $\phi_t(z)$ is non-decreasing with z. The marginal exporter $\tilde{z}_t$ satisfies

$$\tilde{\Lambda}_1 \left[\tilde{z}_t^{\sigma-1}\phi_t(\tilde{z}_t)^{\alpha(\sigma-1)}\right]^{\frac{1}{\alpha-(\sigma-1)}} \left[\left(1+\tau_t^{1-\sigma}\right)^{\frac{\alpha}{\alpha-(\sigma-1)}} - 1\right] D_t^{\frac{\alpha}{\alpha-(\sigma-1)}} = f_t^X. \tag{31}$$

The equilibrium innovation thereby can be expressed as

$$\kappa_t^*(z) = \begin{cases} \left[\frac{(\sigma-1)\tilde{\sigma}}{\alpha}\left(1+\tau_t^{1-\sigma}\right) D_t\right]^{\frac{1}{\alpha-(\sigma-1)}} \left[z\phi_t(z)^{\sigma-1}\right]^{\frac{1}{\alpha-(\sigma-1)}}, & \text{if } z \geq \tilde{z}_t. \\ \left[\frac{(\sigma-1)\tilde{\sigma}}{\alpha} D_t\right]^{\frac{1}{\alpha-(\sigma-1)}} \left[z\phi_t(z)^{\sigma-1}\right]^{\frac{1}{\alpha-(\sigma-1)}}, & \text{if } z < \tilde{z}_t. \end{cases} \tag{32}$$

Then the productivity $\phi_{t+1}(z)$ is given by Equation (6). The aggregate price index can be expressed as

$$P_t = \frac{\sigma}{\sigma-1}\left[\int_{S_z} \left[\kappa_t^*(z)\phi_t(z)\right]^{\sigma-1} dG(z) + \int_{\tilde{z}_t}^{\infty} \tau_t^{1-\sigma}\left[\kappa_t^*(z)\phi_t(z)\right]^{\sigma-1} dG(z)\right]^{\frac{1}{1-\sigma}}. \tag{33}$$

Total expenditure satisfies

$$X_t = L + \frac{1}{\sigma}\left(1 - \frac{\sigma-1}{\alpha}\right) X_t - \left[1 - G(\tilde{z}_t)\right] f_t^X. \tag{34}$$

Steady-state equilibrium in the two-symmetric country world consists of $(P, X, \phi(z), \tilde{z})$ such that

$$\kappa^*(z) = \begin{cases} \left[\frac{(\sigma-1)\tilde{\sigma}}{\alpha}\left(1+\tau^{1-\sigma}\right) D\right]^{\frac{1}{\alpha-(\sigma-1)}} \left[z\phi(z)^{\sigma-1}\right]^{\frac{1}{\alpha-(\sigma-1)}}, & \text{if } z \geq \tilde{z}. \\ \left[\frac{(\sigma-1)\tilde{\sigma}}{\alpha} D\right]^{\frac{1}{\alpha-(\sigma-1)}} \left[z\phi(z)^{\sigma-1}\right]^{\frac{1}{\alpha-(\sigma-1)}}, & \text{if } z < \tilde{z}. \end{cases}, \quad D = P^{\sigma-1}X. \tag{35}$$

$$\tilde{\Lambda}_1 \left[\tilde{z}^{\sigma-1}\phi(\tilde{z})^{\alpha(\sigma-1)}\right]^{\frac{1}{\alpha-(\sigma-1)}} \left[\left(1+\tau^{1-\sigma}\right)^{\frac{\alpha}{\alpha-(\sigma-1)}} - 1\right] D^{\frac{\alpha}{\alpha-(\sigma-1)}} = f^X. \tag{36}$$

$$\phi(z) = \left[\kappa^*(z)\phi(z) + \delta \int_{S_z} m(z', z)\kappa^*(z')\phi(z')\, dG(z')\right]^{\beta}. \tag{37}$$

$$P = \frac{\sigma}{\sigma-1}\left[\int_{S_z} \left[\kappa^*(z)\phi(z)\right]^{\sigma-1} dG(z) + \int_{\tilde{z}}^{\infty} \tau^{1-\sigma}\left[\kappa^*(z)\phi(z)\right]^{\sigma-1} dG(z)\right]^{\frac{1}{1-\sigma}}. \tag{38}$$

$$X = L + \frac{1}{\sigma}\left(1 - \frac{\sigma-1}{\alpha}\right) X - \left[1 - G(\tilde{z})\right] f^X. \tag{39}$$

In the transitional paths, the innovation of exporting firms can be expressed as

$$\kappa_t^*(z) = \left[\frac{(\sigma-1)\tilde{\sigma}}{\alpha}\left(1+\tau_t^{1-\sigma}\right)D_t\right]^{\frac{1}{\alpha-(\sigma-1)}}\left[z\phi_t(z)^{\sigma-1}\right]^{\frac{1}{\alpha-(\sigma-1)}}. \tag{40}$$

We now consider the effects of trade liberalization on non-exporting firms' innovation. Under two symmetric countries, the innovation of non-exporting firms can be expressed as

$$\kappa_t^*(z) = \left[\frac{(\sigma-1)\tilde{\sigma}}{\alpha}D_t\right]^{\frac{1}{\alpha-(\sigma-1)}}\left[z\phi_t(z)^{\sigma-1}\right]^{\frac{1}{\alpha-(\sigma-1)}}. \tag{41}$$

Equation (41) shows that non-exporting firms cannot benefit from the direct market size of trade liberalization. Their innovation depends on the aggregate demand shifter D_t.

B.2.1 Proof to Lemma 3

Proof. By Equation (37), the steady-state $\phi(z)$ is non-decreasing with z. So we can define $\tilde{z}_t$ as in Equation (12).

Equation (12) implies that $\tilde{z}_t$ is decreasing with D_t. Equation (15) implies that D_t is increasing with $\tilde{z}_t$. As a result, there is a unique equilibrium combination of $(D_t, \tilde{z}_t)$. Similar logic applies to all $(D_k, \tilde{z}_k)_{k>t}$. ∎

B.2.2 Proof to Proposition 4

Proof. By Equation (41), $\kappa_t^*(z)$ is increasing with D_t for non-exporters. The equilibrium $(D_t, \tilde{z}_t)$ is determined by Equation (12) and (15). As illustrated by Figure 8, D_t is increasing with τ. So we have $-\frac{d\kappa_t^*(z)}{d\tau} < 0$. ∎

Appendix C Quantification

C.1 Calibrating the initial $\left(\mu_{z,\ell}, L_\ell, \tau_{\ell n}, f_{\ell n}^X\right)$

We consider two economies: China and the rest of the world (ROW). We calibrate L_ℓ to population in China and the ROW. We normalize $\mu_{z,\text{CHN}} = 0$. Then we solve the steady-state to match the following moments: (i) the relative welfare $\frac{X_{\text{CHN}}/P_{\text{CHN}}}{X_{\text{ROW}}/P_{\text{ROW}}}$ is equal to the relative real GDP per capita; (ii) the export share, $\frac{X_\ell^e}{X_\ell}$, is equal to the export share in the data; and (iii) the share of exporters, $1-G_\ell(\tilde{z}_\ell)$, is equal to the share of exporters in the data.

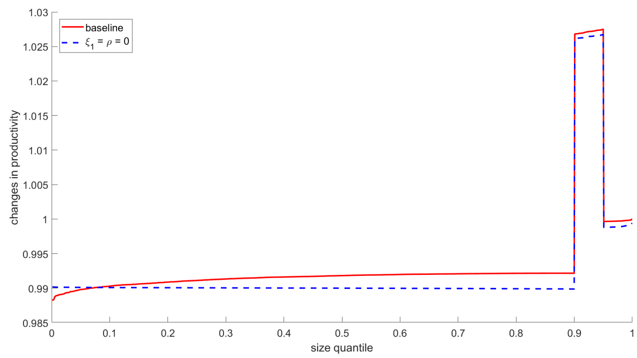
C.2 Computing the Steady-State of the Trade Economy

Given $(\mu_{z,\ell}, L_\ell, \tau_{\ell n}, f_{\ell n}^X)$ and $(\alpha, \beta, \gamma, \xi_1, \xi_2, \rho)$, we solve for the steady-state $(w_\ell, P_\ell, X_\ell, \tilde{z}_\ell)$ with the numeraire $w_1 = 1$.

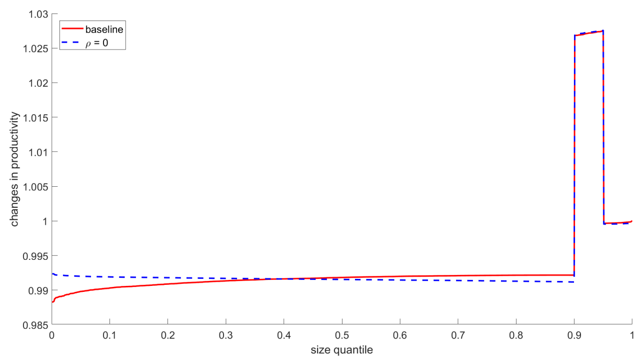
Algorithm 5 *We solve the steady-state $(w_\ell, P_\ell, X_\ell, \tilde{z}_\ell)$ as follows:*

1. *Initial guess $(w_\ell^{(0)}, P_\ell^{(0)}, X_\ell^{(0)}, \tilde{z}_\ell^{(0)})$.*
2. *Compute κ_ℓ^* by Equation (3).*
3. *Compute $\phi_\ell(z)$ by iterating Equation (6).*
4. *Update to $P_\ell^{(1)}$ by Equation (10).*
5. *Update to $w_\ell^{(1)}$ by Equation (8) and (7).*
6. *Update to $X_\ell^{(1)}$ by Equation (8) and (9).*
7. *Update to $\tilde{z}_\ell^{(1)}$ by Equation (5).*
8. *Repeat until convergence.*

C.3 Trade Liberalization under Two Symmetric Countries



(a) $\xi_1 = \rho = 0$



(b) $\rho = 0$

Figure 17: Productivity Effects of Trade Liberalization: the Steady-State

(Note: we fix $f^X_{\ell n,t} = 1.4$ for all t and $\ell \neq n$.)