Trade and Technology Compatibility in General Equilibrium*

Weili Chen — Penn State
Jingting Fan — Penn State
Wenlan Luo — Tsinghua University

November 2023

Abstract

We develop a model of endogenous production networks in which the horizontal proximity between a firm's own technology and the technology of its suppliers affects the input cost. Firms choose their technology balancing the benefit of moving in the technological space toward suppliers with the cost of moving away from the technology they know the best. By altering the relative costs of suppliers from different countries, trade policies shape technology choices, which in turn affect trade patterns and welfare. We parameterize the model to match the relationship between trade patterns and technology proximity that we construct from patent data. We use the model to assess the role of trade in shaping firms' technology choices in the global economy and to quantify the barriers to trade arising from technology incompatibility. Finally, we examine the welfare cost of the trade conflict between the U.S. and China in the semiconductor industry. We find that an embargo on semiconductor exports to China leads to a de-coupling in the technologies of the two countries and a re-alignment of the technologies of the rest, amplifying the welfare costs for both the U.S. and China.

*We thank Jonathan Eaton, Michael Zheng Song, and Stephen Yeaple for helpful conversations. We also thank participants at the 2023 Tsinghua Growth and Institute Conference for helpful comments.

1 Introduction

Firms source intermediate inputs from other firms. As intermediate inputs often embody the technology choice of the supplier, the best supplier of an input is not necessarily the one with the lowest cost; the compatibility between the firm's own technology and that of the supplier also plays a role. For example, a producer of internal combustion engines, no matter how efficient, does not stand much chance of becoming an engine supplier of Tesla; similarly, the leading producer of the power battery for electrical vehicles is unlikely to serve as an important supplier of gasoline car makers.

Guided by the compatibility incentive, in choosing a technology for making their product, firms consider the pool of suppliers they can tap into under each choice. In the global economy where firms with different technologies interact in a production network, firms' technology choice and international sourcing decision shape each other, resulting in a two-way relationship between trade and technology proximity of countries. On the one hand, firms tend to find compatible suppliers in countries with similar technologies; on the other hand, closer trade ties with a trade partner incentivize firms to choose technologies that are more aligned with the technologies of that trade partner. Following the example in the car industry, makers of electric vehicle (EV) source more inputs from countries specializing in EV parts, such as power batteries; conversely, freer trade with a country that is efficient at producing EV parts may incentivize more firms to make EV.

This interplay between trade and technology is illustrated in Figure 1. The figure plots log exports from sector j of country o to sector i of country d against the technology proximity between (d, i) and (o, j), controlling for $d - i$, $o - j$, and $i - j$ fixed effects, as well as bilateral distance metrics and income differences between d and o. We measure technology proximity using the cosine similarity of patent citation profiles in (d, i) and (o, j) following Jaffe (1986). The figure shows that country-sectors with more similar technologies tend to trade more with each other. This pattern cannot be rationalized by the standard Ricardian theory, which predicts that countries tend to trade more with partners that have dissimilar technologies, but is consistent with a theory where firms make technology and sourcing decisions jointly under compatibility incentives.

Building on these observations, this paper studies the relationship between trade and the technology proximity of countries. We answer the following questions: How do technology proximity shape trade? How important is technology choice under compatibility incentive in shaping the similarities and differences in the technologies of countries? How do such forces affect the welfare consequences of trade policy? In particular, the ongoing trade conflict between the U.S. and China has led to a discussion of trade and technological decoupling between the two countries. What are the general equilibrium effects of such decoupling?

To answer these questions, we construct a tractable model of trade with endogenous

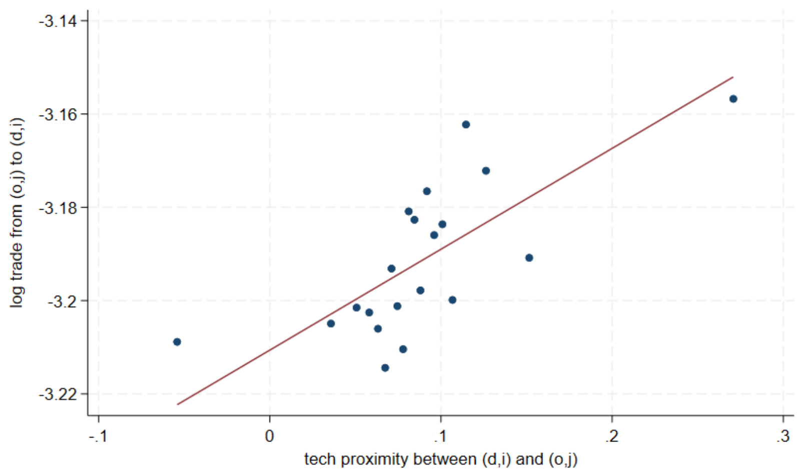


Figure 1: Technology Proximity and Trade

Note: the figure is a binned scatter plot of log exports from country o sector j to country d sector i against the technology proximity between (d, i) and (o, j). Technology proximity is measured as the cosine similarity (Jaffe, 1986) of patent citation profiles in (d, i) and (o, j). Trade data is from the World Input-Output Database (WIOD). Sample of countries are those covered by the WIOD. The regression controls for $d - i$, $o - j$, and $i - j$ fixed effects, as well as distance metrics (geographic, language, colonial ties) and income differences between o and d. The regression coefficient is 0.22 (s.e. 0.042).

production networks and technology choices under compatibility incentives. While a few recent studies have examined questions related to technology compatibility in trade, they do so in environments with two countries and two technologies differentiated from each other by nationality (e.g. domestic or foreign), which are unsuitable for general equilibrium quantification.[1] Rather than dichotomous technologies based on nationality, our model incorporates arbitrarily many technologies for firms to choose from and firm heterogeneity in comparative advantage across these technologies, while retaining the key features of the recent quantitative trade literature that make it amenable to the data—multiple countries, input-output linkages, and general equilibrium forces (Costinot and Rodriguez-Clare, 2014).

In the model, firms are heterogeneous in their (vertical) efficiency and the endowment of a (horizontal) technology, symbolized as a point $\bar{\theta} \in T$, where T is a complete metric space containing all technologies. Following the Hotelling (1929) tradition, a point in this space represents a combination of technical ingredients that can be used for production, which differs from vertical efficiency in that firms have different preferences over their supplier's horizontal technology. We interpret the endowment draw of a firm as the technology

[1]For example, Carluccio and Fally (2013) examine how the opportunity to supply foreign firms affects the availability of intermediate goods of domestic downstream firms. In their model, foreign firms use a 'modern' technology, and domestic firms in the upper-stream industry choose between this modern technology (which qualifies them to supply to foreign firms) and a domestic technology (which disqualifies them due to incompatibility). In a different setting, Costinot (2008) develops a model with two countries each with different technologies, and uses it to examine how the institutions governing horizontal standardization affect welfare. Given there are multiple countries in the world economy, each with potentially different technologies, mapping the technologies of countries in the data to the binary choice in the model can be subjective.

most fitting for a firm's unique capability, determined by factors such as the expertise of key personnel that are exogenous to our model. The distribution from which firms draw their endowment varies across countries and sectors, representing the comparative advantage of countries in the space of technology. Firms choose a point $\theta \in T$ with which they carry out production given their endowment draw. Although this choice comes with adaptation costs that increase with $\|\theta - \bar{\theta}\|$, i.e. the distance between θ and $\bar{\theta}$, it also opens doors to potential benefits. By aligning their technology more closely with those of their key suppliers, firms can leverage low-cost intermediate goods. This benefit is embodied in the assumption that the marginal cost for firm ν of using input from a supplier ω increases with $\|\theta(\nu) - \theta(\omega)\|$, namely, the distance between ν's own technology and that of supplier ω. Therefore, when choosing their technology, firms balance the benefit of moving in the technological space toward suppliers with the cost of moving away from the technology they are adept at.

After choosing their technology, firms randomly sample a set of suppliers, pick those that offer the lowest compatibility-adjusted prices for each intermediate input, and move forward with production. Products are then sold to final consumers and downstream firms. In equilibrium, given country-specific *exogenous* distributions from which firms draw their endowment technology, firms across all countries anticipate the choices of their suppliers and make their own choices accordingly. The equilibrium is a fixed point that is characterized by an *endogenous* distribution of firms across their chosen technology in each country—an infinite dimensional object. We show that the model aggregates tractably, that the equilibrium for technology choice exists, and that the equilibrium is unique if the cost of changing technology is not too small or if the benefit from compatibility is not too large. In the limiting case in which the benefit from compatibility is zero, our model specializes to a version of the Caliendo and Parro (2015) model with firm-specific trade cost. Therefore, it retains the versatility of the canonical quantitative trade models.

We characterize the positive and normative implications of the model. On the positive side, the model suggests a novel relationship between trade and firms' technological direction. For example, within a country-sector pair, firms with endowment draws that are closer to a foreign country are likely to choose a similar technology direction and import from it. This results in a *firm-level* correlation between technological proximity and imports. At the *country-sector* level, lower import tariffs from a trading partner can incentivize firms in the importing country to adopt technologies closer to those chosen by suppliers in that partner.

Our model also offers a fresh perspective on several established facts. For instance, firms that import from a certain origin country tend to also export to the same country, since both activities are shaped by technological proximity. Similarly, firms exporting to a destination country are more likely to export to neighboring countries if these neighbors share similar technologies with the destination. These patterns, previously explained via exogenous state-dependent trade costs (Morales, Sheu and Zahler, 2019; Li, Xu, Yeaple and Zhao, 2023), can be rationalized in our model through the lens of endogenous technological choice.

On the normative side, the model highlights a key externality in technology choice. As a firm selects technology to maximize its own profit, it also alters the production cost of all firms—both domestic and foreign—that use its output as intermediates, an effect not internalized by the firm. This effect can be amplified through input-output linkages. Because of these externalities, firms tend to underinvest in moving away from the technology they are endowed with. In particular, in a one-sector two-symmetric-country special case of the model, we show that starting from the decentralized equilibrium, moving one country's technology choice toward that of the other country increases the welfare of *both* countries. In other words, the externality affects not only the domestic economy but also other countries through trade.

We test and quantify the importance of the model mechanisms. We use bilateral citation statistics derived from the universe of the world's patents, obtained from PATSTAT, as a proxy for the proximity of technology. To give these statistics a structural interpretation, we embed a knowledge attribution decision into the model, in which each firm needs to attribute all other firms that use identical technologies as itself through citations. All else equal, two countries with similar technologies tend to cite each other's patents intensively.

We provide two sets of evidence. First, we show that among Chinese importing firms, there exists a statistically significant and economically meaningful correlation between importing from a country and being close to that country in the technological space, a finding that remains robust when controlling for firm-year and firm-origin country fixed effects. We view this correlation as capturing that among these importers, those whose endowment technology shifts closer to a foreign country choose a technology closer to that country and, as a result, import more from that country. Second, we show that at a country-sector level, exogenous changes in the most-favored-nation (MFN) tariffs cause importers to shift their technologies toward the exporter. In further support of the model mechanism, we find that it is input tariffs, rather than output tariffs, that drive the results. This differentiates our model from alternative explanations, such as learning by 'seeing' the products of competitors in the local market.

We parameterize the model for quantitative exercises. Our analysis includes 15 (N) major world regions, aggregated from the countries listed in the World Input-Output Database based on geographic proximity and political orientation. Each region is then divided into 19 (S) two-digit manufacturing sectors. Although our model allows the technology space to be metric spaces with any finite dimension, for simplicity and tractability we assume that T is the real line and that firms in a country-sector cell draw their endowment from a Normal distribution. Under these assumptions, we derive (quadratic) approximate solutions to firms' technology choice problems. We further show the distribution of technologies chosen by firms in a country-sector cell preserves a Normal distribution, with its mean and variance determined by the general equilibrium interactions.

We calibrate the model to match technological proximity between pairs of countries and sectors, trade flows, and output data. Two essential parameters in our model are the benefit of technological compatibility with suppliers and the cost of moving away from the technology a firm knows the best. We show that coefficients from the firm- and sector-level regressions discussed earlier identify these parameters and we target these estimates in indirect inference. Given these parameters and the calibrated distributions of technology, the model implies endogenous transaction costs due to incompatibility between any two firms. We then pick residual exogenous iceberg trade costs to exactly match the trade shares between countries.

Using the calibrated model, we assess the importance of technology proximity in explaining the data and in shaping firms' decisions in two ways. First, we examine the model's capacity to account for the technological proximity between pairs of country-sector cells. The model incorporates a total of 81,225 ($N^2 \times S^2$) pairs. Our calibration rationalizes the distance between each of these pairs by choosing $N \times S = 285$ mean location parameters, one for each country-sector cell, and S variance parameters, one for each sector so that the model-implied average technological proximity between firms matches the empirical counterpart. We find that these $N \times S + S$ parameters can explain 68.8% of the variations in bilateral technological proximity. This explanatory power comes from two forces: the calibrated difference/similarity in technologies between countries, and the size and sectoral composition of countries (i.e., counties with a larger stock of technology will be cited more often). If we eliminate the first force, the explanatory power decreases by more than a half, indicating that the calibrated positions of technological distributions play an important role in accounting for the data. With the model, we then back out the distributions of the endowment technology of countries. We find that almost half of the model's explanatory power is endogenous—from firms' choices due to the compatibility incentive, while the rest derives from differences in countries' *exogenous* distributions of endowment technologies. These results highlight the importance of both endowment distributions and endogenous changes shaped by compatibility incentives in explaining technological disparities between countries.

Second, we measure the iceberg-equivalent costs firms bear due to technology incompatibility. We find that on average, these costs total to around 3.4% of GDP. International trade accounts for approximately one-third of these costs, whereas domestic trade accounts for the remainder. The costs are bigger in countries whose technologies are distant from those of other countries. For example, in China, costs that arise from incompatibility with foreign suppliers amount to 2.2% of the GDP. These numbers show that technology compatibility is an important factor in firms' technology choice.

We use the model to investigate the welfare costs of the decoupling in the semiconductor industry between U.S. and China. Motivated by the recent events, we simulate a hypothetical trade conflict—a full embargo by the West on China's imports in the "Computer, Electronic, and Optical products" sector. We find that maintaining firms' technology choices constant during such an embargo would result in economic damage to both China (by 0.419%) and the U.S. (by 0.016%). If firms are allowed to adjust their technology, firms in both China and the U.S. would diverge in their technology choices, leading to increased input costs for downstream firms in both countries and, consequently, larger welfare losses—the cost of China increases to 0.795% and the U.S. to 0.081%. Thus, a significant portion of the cost of trade conflicts between the U.S. and China in the semiconductor industry materializes through the resulting decoupling in technology after trade linkages are severed.

In addition to the nascent literature on technology incompatibility and trade, our paper engages with four strands of the literature. First, our paper is related to the studies examining the network structure of production and trade, including work by Jones (2011), Chaney (2014), Oberfield (2018), Boehm and Oberfield (2020), Lim (2018), Acemoglu and Azar (2020), Dhyne, Kikkawa, Kong, Mogstad and Tintelnot (2023), and Demir, Fieler, Xu and Yang (2021), among many others. Within this literature, the closest papers are Boehm and Oberfield (2020) and Demir et al. (2021). In particular, our method of obtaining tractable aggregation through extreme-valued random efficiency draws shares similarities with the approaches of Boehm and Oberfield (2020); our focus on the 'sorting' of firms in the network is related to the work of Demir et al. (2021), who study network formation and vertical quality choice. Relative to these papers, the main contribution of our paper is to incorporate technology choice motivated by technology compatibility into a general equilibrium model with an endogenous buyer-seller network.

Second, our paper is also related to the quantitative trade literature (see Costinot and Rodriguez-Clare, 2014 for an early survey). Our contribution to this literature is twofold. First, we generalize the canonical trade model with input-output linkages to incorporate endogenous horizontal technology choice, which we show plays an important role in shaping the world economy. To incorporate this mechanism, in our model, heterogeneous firms choose among a continuum of options and then interact directly with one another—rather than play a mean field game as in canonical quantitative trade models (e.g. those based on Melitz, 2003; Chaney, 2008; Eaton and Kortum, 2002; Caliendo and Parro, 2015), in which firms' decision only consider equilibrium prices and aggregate quantities. Our second contribution to the literature is to develop intuition and tools for establishing the existence and uniqueness of equilibrium in this model.[2]

Third, our paper connects with the literature that examines the relationship between

[2]Other trade models that are not mean-field games include those based on the assignment model, e.g., Costinot and Vogel (2010); Costinot et al. (2015). Such models often imply efficient allocation, so welfare theorems can be invoked to establish the existence and uniqueness of equilibrium. This approach does not apply to our model due to the existence of externalities.

trade and technological spillovers. See e.g., Buera and Oberfield (2020), Cai, Li and Santacreu (2022), Lind and Ramondo (2023), Ayerst et al. (2023), Liu and Ma (2021), and Aghion, Bergeaud, Gigout, Lequien and Melitz (2021); see also Keller (2021) for a review. Most papers in this literature measure spillover using vertical measures, such as the total innovation or productivity of a country. Our empirical evidence, which relates trade to patent citations, complements this literature. However, our paper differs from most of the works by focusing on the relationship between trade and *directional* technology choices, captured by bilateral citations, and by offering a structural interpretation of this relationship.

Last but not least, our quantitative application on the recent trade conflict between the U.S. and China is related to a rapidly expanding body of literature on this topic, including, for example, Amiti, Redding and Weinstein (2019), Fajgelbaum, Goldberg, Kennedy and Khandelwal (2020), and Huang, Lin, Liu and Tang (2018), among many others. We show that the technology decoupling between U.S. and China, and the re-alignment of technologies in other countries, plays a crucial role in shaping the impacts of trade decoupling. Given the gradual nature of technology adaptation, our exercise suggests that a significant portion of the welfare costs may become apparent over a longer time frame.

2 Model

2.1 Environment.

There are N countries, denoted by d or o, and S sectors, denoted by i or j. In each country d, there is a representative household that inelastically supplies L_d units of labor, and a representative producer of a homogeneous final good made from intermediate inputs from all country-sector pairs with the following technology:

$$Q_d = \prod_{j=1}^{S} \Big[\sum_{o=1}^{N} \int_0^1 [q_{do}^j(\omega)]^{\frac{\eta-1}{\eta}} \mathrm{d}\omega \Big]^{\frac{\eta}{\eta-1}\cdot \rho_d^j}, \tag{1}$$

where $q_{do}^j(\omega)$ denotes the quantity of intermediate good ω produced in country o sector j that is used to produce final good in d, $\eta > 1$ is the elasticity of substitution between different varieties, and ρ_d^j is the share of sector j in final good in country d. Final goods are non-tradable and used for both household consumption and firms' innovation.

In each country-sector (d,i) or (o,j), there is a unit mass of firms that produces differentiated tradable intermediate goods $\nu, \omega \in (0,1)$, made of labor and other intermediate goods. Firms charge marginal cost when selling to other intermediate good producers, but charge a monopolistically competitive markup when selling to the representative final good producers. They differ from each other not only vertically in production efficiency but also horizontally in their technology choice. The latter is shaped endogenously by the firms' own technology expertise and their incentive to be technologically compatible with efficient suppliers.

Let $\bar{\theta}(\nu)$ denote the technology firm ν knows the best. Firm ν first decides what technology it would use for production, denoted by $\theta(\nu)$. Both $\bar{\theta}(\nu)$ and $\theta(\nu)$ are points in a metric space with a distance metric that we need not specify yet. Choosing a distant technology from the firm's expertise incurs adaptation costs; in return, firms reduce intermediate input costs by moving closer to the choice made by efficient suppliers. Given the choice of technology, each firm randomly samples a set of suppliers with different horizontal technology and vertical efficiency and chooses the ones with the best combination of the two components. Then firms produce and sell their output to other firms and final good-producers.

In the rest of this section, we start by describing firms' production and input sourcing decisions, taking as given the choice of technology. We then describe the technology choice and establish the condition for the existence and uniqueness of the equilibrium. Finally, we derive testable implications and discuss how the mechanisms in our model affect welfare.

2.2 Production and Sourcing Decisions: Intermediate Firms

Firm ν in country-sector (d, i) with technology $\theta(\nu)$ has access to a random set of production *techniques*. Each technique r in the set of the techniques available to firm ν, denoted by $R(\nu)$, is characterized by (i) a level of total factor productivity $A(\nu, r)$, and (ii) a set of potential suppliers drawn independently and uniformly from all firms in each country-sector pair (o, j). We denote the set of suppliers from o, j for technique r by $\Omega_o^j(\nu, r)$.

Under a technique r, the output of firm ν is given by

$$y(\nu, r) = A(\nu, r)[l(\nu)]^{\gamma^{iL}} \prod_j \left[m^j(\nu, r)\right]^{\gamma^{ij}}, \text{ with } \gamma^{iL} + \sum_j \gamma^{ij} = 1,$$

where $l(\nu)$ is the labor hired by ν, and $m^j(\nu, r)$ is the intermediate input sourced by ν from sector j. Denote the wage rate in d by w_d. The factory-gate price (marginal cost) of firm ν using production technique r is

$$p(\nu, r) = \frac{1}{A(\nu, r)} \cdot [w_d]^{\gamma^{iL}} \cdot \prod_j \left[c^j(\nu, r)\right]^{\gamma^{ij}},$$

where $c^j(\nu, r)$ is the lowest *effective* marginal cost for input j from the set of suppliers given by r:

$$c^j(\nu, r) = \min_o \min_{\omega \in \Omega_o^j(\nu, r)} \tilde{c}^j(\nu, \omega).$$

$\tilde{c}^j(\nu, \omega)$ is the effective marginal cost of the intermediate input produced by supplier $\omega \in \Omega_o^j(\nu, r)$, which depends on the trade cost between o and d in sector j, supplier ω's factory-gate price, the compatibility of technology between ν and ω, and an idiosyncratic match-specific efficiency draw. $\tilde{c}^j(\nu,\omega)$ is defined as

$$\tilde{c}^j(\nu,\omega) = \tau^j_{do} \cdot p(\omega) \cdot \frac{1}{z(\nu,\omega)} \cdot t(\theta(\nu),\theta(\omega)),$$

where τ^j_{do} is the sector-specific iceberg trade cost, $p(\omega)$ is the factory-gate price of supplier ω, $z(\nu,\omega)$ is the match-specific productivity, and $t(\theta(\nu),\theta(\omega))$ represents the cost due to technology incompatibility that is increasing in the difference between $\theta(\nu)$ and $\theta(\omega)$.

Each firm chooses the technique from its choice set to maximize profit—or equivalently, to minimize its factory-gate price:

$$p(\nu) = \min_{r\in R(\nu)} p(\nu,r).$$

The Cobb-Douglas production function implies that the share of input j is γ^{ij}. In each input sector, firm ν chooses only one supplier, who might be in a foreign country. Firms' choice of techniques and suppliers determine the trade in intermediate goods between countries. For tractable aggregation, we impose the following assumption on the available techniques.

Assumption 1. *For any firm ν in country-sector (d,i),*

1. *For any $a_1 < a_2 \in (0,+\infty)$, the number of production techniques with $A(\nu,r) \in (a_1,a_2)$ follows a Poisson distribution with mean $(a_1/A^i_d)^{-\lambda} - (a_2/A^i_d)^{-\lambda}$, where $\lambda > 1$.*

2. *For each production technique r, each firm in $\Omega^j_o(\nu,r)$ receives independent match-specific efficiency draws with firm ν. For any $z_1 < z_2 \in (0,+\infty)$, the number of suppliers in $\Omega^j_o(\nu,r)$ for whom the match-specific efficiency $z(\nu,\omega) \in (z_1,z_2]$ follows a Poisson distribution with mean $z_1^{-\zeta} - z_2^{-\zeta}$, where $\zeta > 1$.*

Assumption 1 describes the distribution of productivity and match-specific efficiency among the techniques available to a firm in (d,i). In part (1), A^i_d shapes the average productivity of techniques available to firms in (d,i) and λ governs the heterogeneity among the techniques. In part (2), ζ determines the heterogeneity in match-efficiency across suppliers. The Poisson assumption on productivity and match-specific efficiency distributions has been employed in Boehm and Oberfield (2020) in a closed-economy setting and can be viewed as an extension to the widely used Pareto assumption (Chaney, 2008; Kortum, 1997).[3] It allows us to characterize the distribution of factory-gate prices for firms in any (d,i) with any chosen technology $\theta(\nu)$.

Proposition 1. *Under Assumption 1, $p^i_d(\theta)$, the factory-gate price of a firm in (d,i) with technology*

[3] For example, for any $x > 0$, if the number of techniques with $A(v,r) > x$ follows Poisson distribution with mean $(\frac{x}{A^i_d})^{-\lambda}$ and these techniques draw their efficiency independently from a Pareto distribution with position parameter A^i_d and tail parameter λ, then the resulting distribution as $x \to 0$ would satisfy part (1).

location θ, follows a Weibull (inverse Fréchet) distribution with c.d.f.

$$F_d^i(p;\theta) = 1 - \exp\big(-[p/C_d^i(\theta)]^\lambda\big), \tag{2}$$

with $C_d^i(\theta)$ determined as the fixed point of

$$C_d^i(\theta) = \frac{\Xi^i}{A_d^i}[w_d]^{\gamma^{iL}} \prod_j \Big[\sum_o \int [\tau_{do}^j C_o^j(\tilde{\theta}) t(\theta,\tilde{\theta})]^{-\zeta} \mathrm{d}\Theta_o^j(\tilde{\theta})\Big]^{-\frac{\gamma^{ij}}{\zeta}}, \tag{3}$$

where Ξ^i's are sector-specific constants, [4] *and Θ_o^j is the measure of firms in (o,j) choosing technology $\tilde{\theta}$.*

Proof. See Appendix A.1. □

To establish this result, we show that if the distribution for the factory-gate price of suppliers from all (o,i) follows a Weibull distribution, then the price of the down-stream firm in (d,j) also follows a Weibull distribution. Since all firms both use inputs produced by others and at the same time supply goods to others, in equilibrium, the location parameters of these distributions, which we denote by $C_d^i(\theta)$ for the factory-gate price distribution of firms in (d,i) that choose technology θ, also depends on $\{C_o^j(\theta) : \theta \in T\}_{o=1,j=1}^{N,S}$. These location parameters form a fixed point described by equation (3).

Inspection of equation (3) reveals forces that shape the price distribution in (d,i). The standard forces, which appear in most trade models with input-output linkages, include the effective productivity of (d,i) that is embodied in $\frac{\Xi^i}{A_d^i}$, the wage in country d that is denoted by w_d, the iceberg cost of importing intermediate inputs τ_{do}^j for all o and j, and the parameters governing the price distribution of suppliers $\{C_o^j(\theta) : \theta \in T\}_{o=1,j=1}^{N,S}$. Distinct from the literature are two features that arise from technology compatibility. First, the cost of using a supplier with technology $\tilde{\theta}$ also depends on the importance of compatibility, captured in $t(\theta,\tilde{\theta})$. Second, because the degree of compatibility with suppliers differs across suppliers with different technologies, the entire distribution of supplier technology Θ_o^j and the cost for each $\tilde{\theta} \in \Theta_o^j$ matter.

These forces imply that firms in (d,i) who choose different technologies source their inputs from different suppliers. Moreover, the probability of a firm θ sourcing from a supplier with technology $\tilde{\theta}$ depends on the entire distribution $\Theta_o^j(\tilde{\theta})$. We characterize the sourcing decision as a corollary of Proposition 1:

Corollary 1. *For a firm in (d,i) with technology θ*

1. *the expenditure share of the firm's input in sector j that is supplied by firms in o with technology*

[4] Specifically, $\Xi^i \equiv \Big(\int_0^\infty \int_0^\infty ... \int_0^\infty \mathbb{I}[\prod_j [m^j]^{\frac{\gamma^{ij}}{\zeta}} \leq \kappa] \prod_j [\Gamma(1-\zeta/\lambda)]^{\frac{\gamma^{ij}}{\zeta}} \exp(-m^j)\lambda\kappa^{-\lambda-1} \mathrm{d}m^1...\mathrm{d}m^S \mathrm{d}\kappa\Big)^{-1/\lambda}$.

$\tilde{\theta}$ *is*

$$\chi^{j}_{do}(\theta,\tilde{\theta})\mathrm{d}\Theta^{j}_{o}(\tilde{\theta}) = \frac{[\tau^{j}_{do}C^{j}_{o}(\tilde{\theta})t(\theta,\tilde{\theta})]^{-\zeta}}{\sum_{o'}[\tau^{j}_{do'}\Lambda^{j}_{o'}(\theta)]^{-\zeta}}\mathrm{d}\Theta^{j}_{o}(\tilde{\theta}), \tag{4}$$

where $\Theta^{j}_{o}(\tilde{\theta})$ *denotes the measure of firms in* (o,j) *with technology* $\tilde{\theta}$ *and*

$$\Lambda^{j}_{o}(\theta) \equiv \Big(\int [C^{j}_{o}(\tilde{\theta})t(\theta,\tilde{\theta})]^{-\zeta}\mathrm{d}\Theta^{j}_{o}(\tilde{\theta})\Big)^{-1/\zeta}. \tag{5}$$

2. *the expenditure share on intermediate goods produced by firms in country* o *is*

$$\chi^{j}_{do}(\theta) = \int \chi^{j}_{do}(\theta,\tilde{\theta})\mathrm{d}\Theta^{j}_{o}(\tilde{\theta}) = \frac{[\tau^{j}_{do}\Lambda^{j}_{o}(\theta)]^{-\zeta}}{\sum_{o'}[\tau^{j}_{do'}\Lambda^{j}_{o'}(\theta)]^{-\zeta}}. \tag{6}$$

Proof. See Appendix A.1. □

In the first part of the corollary, $\Lambda^{j}_{o}(\theta)$ captures the overall competitiveness of country o as a supplier to a firm in d, and ζ governs the elasticity of sourcing decisions to cost differences across suppliers.[5] Intuitively, if firm with technology $\tilde{\theta}$ in o offers a lower incompatibility-adjusted input cost (smaller $C^{j}_{o}(\tilde{\theta})t(\theta,\tilde{\theta})$) or are more numerous (larger $\Theta^{j}_{o}(\tilde{\theta})$), then they have a higher chance of becoming a supplier. The second part of the corollary aggregate across firms in (o,j) to derive the probability that a firm in (d,i) with technology θ sources input j from any firm in country o.

2.3 Production and Sourcing Decisions: Final-Good Firms

Intermediate-good firms engage in monopolistic competition when selling to final-good producers, charging a constant markup of $\frac{\eta}{\eta-1}$. Facing the markups, the final good producer chooses input from each supplier to maximize their profits:

$$P_d Q_d - \sum_{j}\sum_{o}\int_0^1 [\frac{\eta}{\eta-1}\tau^{Uj}_{do}p^{j}_{do}(\omega)]q^{j}_{do}(\omega)\mathrm{d}\omega, \tag{7}$$

where the factory-gate price $p^{j}_{do}(\omega)$ follows the distribution characterized by (2), τ^{Uj}_{do} is the iceberg trade cost faced by final-good producers,[6] and P_d is the ideal price index for the final good in d,

$$P_d \equiv \prod_{j}(P^{j}_{d}/\rho^{j}_{d})^{\rho^{j}_{d}}, \quad \text{with } P^{j}_{d} \equiv \Big(\sum_{o}\int_0^1 [\frac{\eta}{\eta-1}\tau^{Uj}_{do}p^{j}_{do}(\omega)]^{1-\eta}\mathrm{d}\omega\Big)^{\frac{1}{1-\eta}}.$$

The sourcing decision of the final-good producer differs from that of the intermediate-good producers in two aspects. First, instead of purchasing only one input in each sector, the final

[5] $\Lambda^{j}_{o}(\theta)$ plays the role of $\exp(T)$ in the Eaton and Kortum (2002) model in shaping firms' sourcing decisions.

[6] 'U' in the superscript of τ^{Uj}_{do} is short for 'final user,' which we allow to be different from τ^{j}_{do}.

good producer benefits from the love for variety, hence purchasing from all producers. Second, they are not affected by technological compatibility considerations. We rationalize this assumption by the notion that, ultimately, consumers derive utility more significantly from the services and functionalities enabled by the products that meet their needs rather than from the underlying technology of these products. We characterize the sourcing decision of final-good producers as the second corollary of Proposition 1.

Corollary 2. *For final-good producers in country d, when sourcing sector-j goods,*

1. *the expenditure share allocated to goods produced by firms in country o with technology $\tilde{\theta}$ is*
$$\pi^j_{do}(\tilde{\theta})\mathrm{d}\Theta^j_o(\tilde{\theta}) = \frac{[\tau^{Uj}_{do} C^j_o(\tilde{\theta})]^{1-\eta}}{\sum_{o'}[\tau^{Uj}_{do'}\bar{\Lambda}^j_{o'}]^{1-\eta}}\mathrm{d}\Theta^j_o(\tilde{\theta}), \tag{8}$$
where
$$\bar{\Lambda}^j_o \equiv \left(\int [C^j_o(\tilde{\theta})]^{1-\eta}\mathrm{d}\Theta^j_o(\tilde{\theta})\right)^{1/(1-\eta)}. \tag{9}$$

2. *the expenditure share allocated to goods produced by firms in country o is*
$$\pi^j_{do} = \int \pi^j_{do}(\tilde{\theta})\mathrm{d}\Theta^j_o(\tilde{\theta}) = \frac{[\tau^{Uj}_{do}\bar{\Lambda}^j_o]^{1-\eta}}{\sum_{o'}[\tau^{Uj}_{do'}\bar{\Lambda}^j_{o'}]^{1-\eta}}. \tag{10}$$

Proof. See Appendix A.1. □

2.4 Technology Choice

Upon entry, each firm in country-sector (d, i) is endowed with a technology $\bar{\theta}$ that is randomly drawn from an exogenous probability measure $\bar{\Theta}^i_d(\bar{\theta})$, which we call the *ex-ante* technology distribution. Before making the production and sourcing decision, each firm chooses a desired technology θ and adapts its production process to θ.

Adapting to a different technology from the endowed technology is costly, but it can increase the firm's profit by enabling it to source input from efficient firms that choose similar technology. Recall that firms make profits only through their sales to final-good producers. Such sales depend on the firm's realized production cost, which in turn depends on its draws of production techniques and suppliers described in the previous subsections. Letting $X^i_d(\theta)$ be the expected sales to final-good producers by a firm from (d, i) that chooses technology θ, the technology adaption costs are given by
$$K^i_d(\theta; \bar{\theta}) \equiv \phi(\theta, \bar{\theta}) \cdot \frac{1}{\eta} X^i_d(\theta), \tag{11}$$
in which $\frac{1}{\eta}X^i_d(\theta)$ is the expected profits to the firm through the monopolistic markup and $\phi(\theta, \bar{\theta})$ denotes the fraction of the profits that must be expended for adapting technology.

$\phi(\theta, \bar{\theta})$ increases in the difference between θ and $\bar{\theta}$, capturing that to adapt to a more distinct technology, more investment is needed. [7]

Firms maximize profits netting technology adaption costs by choosing the technology θ that solves the following problem:

$$\max_{\theta} \Pi_d^i(\theta; \bar{\theta}) \equiv [1 - \phi(\theta, \bar{\theta})] \cdot \frac{1}{\eta} X_d^i(\theta), \tag{12}$$

Under monopolistic competition in sales to final-good producers, this is equivalent to

$$\max_{\theta} [1 - \phi(\bar{\theta}, \theta)] \cdot [C_d^i(\theta)]^{1-\eta}. \tag{13}$$

Denote the choice of the technology by $g_d^i(\bar{\theta})$. This choice implies a link between the *ex-post* technology distribution $\Theta_d^i(\theta)$ with the *ex-ante* distribution.

$$\Theta_d^i(\theta) = \int_{\bar{\theta} \in T} \mathbb{I}[g_d^i(\bar{\theta}) = \theta] \mathrm{d}\bar{\Theta}_d^i(\bar{\theta}). \tag{14}$$

We now define the equilibrium for technology choice, taking as given the wages of countries.

Definition 1. ***Equilibrium for Technology Choice.*** *Given the primitive of the economy and wages $\{w_d\}$, an equilibrium for firms' technological choice problem is the mapping from ex-ante to ex-post technology that describes firms' technology choice $g_d^i : T \to T$ and the cost functions $C_d^i : T \to R^+$ for all d, i, such that*

(1) Given $\{C_d^i\}$ and $\{g_d^i\}$, $g_d^i(\bar{\theta})$ solves equation problem (13) for all (d, i) and $\forall \bar{\theta} \in \Theta$ except for a zero measure set.

(2) $\{C_d^i\}$ and $\{g_d^i\}$ satisfy equations (3) and (14).

In this equilibrium, a continuum of ex-ante heterogeneous firms makes simultaneous choices among a continuum of options, with agents' choices interacting with those of all others directly instead of via a mean-field game, where interactions are summarized by a finite-dimensional vector of prices and quantities. We construct a procedure to characterize the existence and uniqueness of such an equilibrium. To simplify the characterization and to ease the computational burden in quantitative analysis, in the rest of the paper, we impose the following assumption:

Assumption 2. *1. The space of technology is the real line, i.e., $T \equiv \mathbb{R}$.*

2. The costs of technological incompatibility and adaptation are given by, respectively,

$$t(\theta, \tilde{\theta}) = \exp(\bar{t} \cdot (\theta - \tilde{\theta})^2) \text{ and } \phi(\bar{\theta}, \theta) = 1 - \exp(-\bar{\phi} \cdot (\bar{\theta} - \theta)^2), \quad \bar{t} \text{ and } \bar{\phi} > 0.$$

[7] We model adaption cost as a function of expected profit. This could be micro-founded in a model of bargaining, in which researchers bargain collectively with the firm to split the profit. Because the firm knows less about a technology that is more distant, the researchers can extract a higher share of the rent for that adaption.

The first part of Assumption 2 takes a one-dimensional interpretation of the technology space. The second part of the assumption implies that the 'iceberg' compatibility/adaptation costs both increases in the distance in technology with constant elasticities, $\bar{t}$, and $\bar{\phi}$, respectively.[8] Under Assumptions 1 and 2, we establish sufficient conditions for the existence and uniqueness of an equilibrium with continous policy functions.

Proposition 2. *Suppose wages $\{w_d\}$ are given.*

1. *Assume $\{\bar{\Theta}^i_d\}$ have bounded support that is contained in $[-M, M]$ for some $M > 0$ and have associated density functions $\{\bar{\varsigma}^i_d\}$. If $\zeta\bar{t} < 1/M^2$, then there exists an equilibrium with firms' technology choice $\{g^i_d\}$ being continuously differentiable functions. Moreover, in this equilibrium, the choice of firms from (d, i) with endowment technology $\bar{\theta}$ is characterized by the following first-order condition with a unique solution.*

$$g^i_d(\bar{\theta}) = \omega^i \sum_{j,o} \frac{\gamma^{ij}}{1-\gamma^{iL}} \int \chi^j_{do}(g^i_d(\bar{\theta}), \tilde{\theta}) \mathrm{d}\Theta^j_o(\tilde{\theta}) + (1-\omega^i)\bar{\theta}, \quad \forall \bar{\theta} \in [-M, M] \tag{15}$$

   where $\omega^i \equiv \frac{(\eta-1)(1-\gamma^{iL})\bar{t}}{(\eta-1)(1-\gamma^{iL})\bar{t}+\bar{\phi}} < 1$.

2. *If, in addition, $\bar{t} < \frac{1}{2M}$ and $\bar{\phi} > \underline{\phi}$, where $\underline{\phi} > 0$ is a constant determined by parameters $(\zeta, \bar{t}, \eta, M, \gamma^{iL})$ as detailed in the proof, then such an equilibrium is unique.*

Proof. See Appendix A.2. □

For existence, we formulate a fixed point problem in firms' policy functions. Let $\boldsymbol{g} : [-M, M] \to [-M, M]^{N\times S}$ be stacked policy functions whose elements are $g^i_d(\cdot)$. Let $\mathcal{G}$ be the set of $\boldsymbol{g}$ that is bounded with Lipschitz-continuous first derivative. We endow $\mathcal{G}$ with the C^1 norm: $\|\boldsymbol{g}\|_{\mathcal{G}} = \|\boldsymbol{g}\|_\infty + \|\boldsymbol{g}'\|_\infty$. We devise a mapping from the resulting normed space $(\mathcal{G}, \|\cdot\|_{\mathcal{G}})$ to itself. We characterize the condition under which the mapping is continuous, and the normed is space convex, compact, and complete (under the C^1 norm but not the infinity norm, hence our choice of the C^1 norm), so the Schauder fixed-point theorem applies.

These conditions are stated in terms of $\zeta\bar{t}$ and M^2. M^2 bounds the variance of technology of potential suppliers, which determines the scope for a firm's technology choice to affect the improvement in supplier compatibility.[9] The parameter $\bar{t}$ captures the impact of compatibility on production cost; and the trade elasticity ζ captures how such cost maps into sales, hence the firms' profits. Intuitively, for an equilibrium with smooth policy function to exist,

[8] A natural extension is for T to be $\mathbb{R}^n$, in which case an intuitive generalization for incompatibility and adaptation costs is for them to take a quadratic form. For example, $t(\theta, \tilde{\theta}) \equiv (\theta - \tilde{\theta})^{tr} \cdot \bar{t} \cdot (\theta - \tilde{\theta})$, where x^{tr} is the transpose of a vector x and $\bar{t}$ here is a positive definite matrix. Our theoretical results presented in this section and the analytical solution under the Gaussian distribution in the quantitative section can both be extended to $T = \mathbb{R}^n$ for a finite n. We focus on the one-dimensional case because, as we show later, with $n = 1$, our model can already fit well the measured technological distance between countries.

[9] By the Popoviciu's inequality on variances, $\mathrm{var}(\theta) < M^2$ for any bounded distribution for θ over $[-M, M]$.

we need firms' policy to be not too sensitive to their own technology endowment, hence the restriction $\zeta \bar{t} M^2 < 1$. The first-order condition clarifies the trade-off firms face in choosing technology: between being close to their endowment technology with a weight of $1-\omega^i$ and being close to the technology of suppliers $\tilde{\theta}$, weighted by ω^i and the expenditure shares of $\tilde{\theta}$ given by $\frac{\gamma^{ij}}{1-\gamma^{iL}} \chi^j_{do}(g^i_d(\bar{\theta}), \tilde{\theta})$.

For uniqueness, we establish the conditions under which the mapping described earlier is a contraction mapping.[10] The sufficient condition for uniqueness requires $\bar{t} < \frac{1}{2M}$, in which $2M$ enters as the upper bound of the technology distance between any firm and the expenditure-share weighted average supplier technology. If the product of this distance and $\bar{t}$ is not too large and if the cost of technology adaption $\bar{\phi}$ relative to $\bar{t}$ and other parameters is not too small, then firms' technology would not be too sensitive to other firms' choice, in which case the mapping is contractive.[11] In the limit case of $\bar{\phi} \to \infty$ or $\bar{t} \to 0$, the unique equilibrium is simply all firms choosing their endowment technology.

2.5 Aggregation and General Equilibrium

We embed firms' technology choice problem described above into the general equilibrium. Market clearing requires that for each firm in country-sector (o, j) with chosen technology location around $\tilde{\theta}$, the total *sales to downstream firms*, $M^j_o(\tilde{\theta})$, should satisfy

$$M^j_o(\tilde{\theta}) = \sum_d \sum_i \gamma^{ij} \int \left[M^i_d(\theta) + (1 - \frac{1}{\eta}) X^i_d(\theta) \right] \chi^j_{do}(\theta, \tilde{\theta}) \mathrm{d}\Theta^i_d(\theta), \tag{16}$$

in which a fraction $\frac{1}{\eta}$ of sales to final-good producers accrue to firms' profits.

Meanwhile, the total *sales to final-good producers*, $X^j_o(\tilde{\theta})$, should satisfy

$$X^j_o(\theta) \equiv \sum_d \rho^j_d P_d Q_d \pi^j_{do}(\theta), \tag{17}$$

where final goods are demanded by household consumption and firms' innovation,

$$P_d Q_d = I_d + \sum_i \int K^i_d(g^i_d(\bar{\theta}); \bar{\theta}) \mathrm{d}\bar{\Theta}^i_d(\bar{\theta}), \tag{18}$$

with technology adaption costs $K^i_d(\theta; \bar{\theta})$ defined in (11).

Household income consists of wages and the net profits of domestic firms, which gives

$$I_d = w_d L_d + \sum_i \int \Pi^i_d(g^i_d(\bar{\theta}); \bar{\theta}) \mathrm{d}\bar{\Theta}^i_d(\bar{\theta}), \tag{19}$$

[10] This is achieved by characterizing the condition under which the Fréchet derivatives of the constructed mapping with respect to the functions describing firms' technology choice (e.g., $\boldsymbol{g}$) have a row sum (of absolute values) below 1.

[11] The interpretation behind the bound for $\bar{t}$ also means that even if M is arbitrarily large, as long as the expenditure shares on firms with distant technology is too large, uniqueness can be established without $\bar{t} \to 0$. Indeed, we show that under our calibration, the equilibrium under normal distributions on ex-ante distributions exists and is unique.

where net profits $\Pi_d^i(\theta;\bar{\theta})$ are defined in (12), and wages are determined by the labor market clearing condition:

$$w_d L_d = \sum_i \gamma^{iL} \int [M_d^i(\theta) + (1-\frac{1}{\eta})X_d^i(\theta)]\mathrm{d}\Theta_d^i(\theta). \tag{20}$$

Definition 2. ***Competitive Equilibrium.*** *Given parameters on geography $\{\tau_{do}^j, \tau_{do}^{Uj}, L_d\}$, preference $\{\rho_d^j, \eta\}$, production technology $\{\gamma^{ij}, \gamma^{iL}, A_d^i, \lambda, \zeta\}$, and the ex-ante technology distribution $\{\bar{\Theta}_o^j\}$, a competitive equilibrium is defined as sets of (i) wages, price index and income $\{w_d, P_d, I_d\}$, (ii) ex-post technology distribution $\{\Theta_o^j\}$, (iii) sales characterized by $\{X_o^j(\theta), M_o^j(\theta)\}$, and (iv) distribution of production costs characterized by $\{C_o^j(\theta)\}$, such that*

i. Given wages, (3) and (14) constitute an equilibrium in firms' technology choice.

ii. Goods and labor market clear, i.e., (16), (17), (18), (19), (20) are solved.

Relationship to existing trade models. Note that if we eliminate the incentive for technology adaption by setting $\bar{t} = 0$ and impose that firms charge marginal cost in selling to final good producers,[12] the model specializes to a generalized Caliendo and Parro (2015) model with firm-specific trade costs that are determined by the difference between the seller's and buyer's ex-ante technologies. Thus, our model generalizes workhorse trade models used in a growing quantitative trade literature to incorporate endogenous horizontal technology choices arising from compatibility incentives. In the rest of this section, we examine analytically the interaction between technology choice and trade and the implications of this interaction for welfare through special cases.

2.6 Special Cases: Interaction between Technology Choice and Trade

Throughout this subsection, we assume that the ex-ante distribution in each (d,i) is degenerate, with a unit mass at $\bar{\theta}_d^i$. We characterize the symmetric equilibrium in which firms that are ex-ante identical, i.e. from the same (d,i), make the same technology choice and examine how these choice respond to changes in trade costs. [13]

Proposition 3. *In the equilibrium of technology choice under degenerate endowment distributions*

(1) The technology choice of firms in (d,i) is

$$\theta_d^i = \omega^i \sum_{j,o} \frac{\gamma^{ij}}{1-\gamma^{iL}} \bar{\chi}_{do}^{ij} \theta_o^j + (1-\omega^i)\bar{\theta}_d^i \tag{21}$$

[12]This can be achieved by assuming that each $\omega \in [0,1]$ varieties are not differentiated by country of origin and the producers of the same variety from different countries engage in perfect competition.

[13]This setup can be viewed as a limit case of the premises described in Proposition 2 with the density functions for the ex-ante distribution approaches a Dirac delta function. Because in this limit case, there are only finite types of firms, the existence and uniqueness of equilibrium can be alternatively established using fixed point theorems on finite-dimensional space; see the supplementary appendix for details.

where $\omega^i \equiv \frac{(\eta-1)(1-\gamma^{iL})\bar{t}}{(\eta-1)(1-\gamma^{iL})\bar{t}+\bar{\phi}}$, *and* $\bar{\chi}^{ij}_{do} \equiv \frac{\exp[-\zeta(\ln\tau^j_{do}+\ln\bar{C}^j_o+\bar{t}(\theta^i_d-\theta^j_o)^2)]}{\sum_{o'}\exp[-\zeta(\ln\tau^j_{do'}+\ln\bar{C}^j_{o'}+\bar{t}(\theta^i_d-\theta^j_{o'})^2)]}$ *is the share of spending by firms in* (d,i) *on the intermediate goods produce by* o *when sourcing input* j*.*

(2) *Given changes in trade costs and technology choices,* $\{\mathrm{d}\ln\tau^j_{do}\}$ *and* $\{\mathrm{d}\theta^i_d\}$*, the location parameters of the distribution for factory-gate prices defined in (3),* $\bar{\boldsymbol{C}} \equiv (\bar{C}^1_1, \bar{C}^2_1, ..., \bar{C}^S_N)$ *change according to:* [14]

$$\mathrm{d}\ln\bar{\boldsymbol{C}} = D_{\tilde{\boldsymbol{\gamma}}}\Omega[2\bar{t}\Lambda\mathrm{d}\boldsymbol{\theta} + \mathrm{d}\ln\tilde{\boldsymbol{\tau}}],$$

where D_x *is an* $NS \times NS$ *diagonal matrix with the diagonal elements being* NS *repetitions of* $NS \times 1$ *vector* x*;* $\tilde{\boldsymbol{\gamma}}$ *is an* $NS \times 1$ *vector whose* $d \times i$*-th element is* $\tilde{\gamma}^i_d = 1-\gamma^{iL}$*;* $\Omega \equiv [\mathbb{I}_{NS\times NS} - D_{\tilde{\boldsymbol{\gamma}}}\Gamma]^{-1}$ *is the Leontief inverse of the expenditure share of sourcing, with* $\Gamma^{ij}_{do} \equiv \frac{\gamma^{ij}}{1-\gamma iL}\bar{\chi}^{ij}_{do}$*;* $\mathrm{d}\ln\tilde{\boldsymbol{\tau}}$ *is an* $NS \times 1$ *vector stacked from* $\mathrm{d}\ln\tilde{\tau}^i_d \equiv \sum_{oj}\Gamma^{ij}_{do}\mathrm{d}\ln\tau^j_{do}$*, the expenditure-weighted average changes in import trade costs of* (d,i)*; and* Λ *is an* $NS \times NS$ *matrix of expenditure-weighted average distance in technology:* $\Lambda \equiv D_{(\mathbb{I}_{NS\times NS}-\Gamma)\theta} + \Gamma D_\theta - D_\theta\Gamma$*.*

(3) *In response to exogenous changes in trade costs* $\{\mathrm{d}\ln\tau^j_{do}\}$*, the change in firms' technology choice is:*

$$\mathrm{d}\boldsymbol{\theta} = -\zeta[\mathbb{I}_{NS\times NS} - D_\omega(\Gamma - 2\zeta\bar{t}\widetilde{\Lambda}D_{\tilde{\boldsymbol{\gamma}}}\Omega\Lambda - 2\zeta\bar{t}\widehat{\Lambda})]^{-1}\left[D_\omega\widetilde{\Lambda}D_{\tilde{\boldsymbol{\gamma}}}\Omega\mathrm{d}\ln\tilde{\boldsymbol{\tau}} + D_\omega\mathrm{d}\ln\widehat{\boldsymbol{\tau}}\right],$$

where $\boldsymbol{\omega}$ *is an* $NS \times 1$ *vector stacked from* ω^i*;* $\widetilde{\Lambda}$ *is an* $NS \times NS$ *matrix stacked from* $\widetilde{\Lambda}^{ij}_{do} \equiv \Gamma^{ij}_{do}[\theta^j_o - \sum_{\tilde{o}}\chi^{ij}_{d\tilde{o}}\theta^j_{\tilde{o}}]$*, the expenditure-weighted average distance between* (o,j) *and all suppliers that* (d,i) *source from;* $\widehat{\Lambda} \equiv -D_{\widetilde{\Lambda}\theta} + \widetilde{\Lambda}D_\theta - D_\theta\widetilde{\Lambda}$*; and* $d\ln\widehat{\boldsymbol{\tau}}$ *is an* $NS \times 1$ *vector stacked from* $\mathrm{d}\ln\widehat{\tau}^i_d \equiv \sum_{jo}\widetilde{\Lambda}^{ij}_{do}\ln\tau^j_{do}$*.*

Proof. See Appendix A.3. □

The first part of the proposition is a special case of equation (A.5) and shows the trade-off firms face between the proximity to the technology of their major suppliers and the proximity to their own endowment technology.

In the second part of the proposition, $D_{\tilde{\gamma}}$ captures the importance of intermediate inputs for production costs. Ω is a Leontief inverse that captures the impact of the production cost in any country on that of any other country, accounting for the propagation via input-output and trade linkages. Inside the bracket are two components: the first captures the effect of changing technology distance, whereas the second captures the effect of the change in expenditure-weighted average change in import trade costs. Intuitively, factory-gate prices of firms in a country increase if the distance between the country's technology to that of others increases (an increase in $\Lambda\mathrm{d}\boldsymbol{\theta}$) or if the import cost increases (an increase in $\mathrm{d}\ln\tilde{\boldsymbol{\tau}}$).

[14]We use $\{\bar{C}^i_d\}$ instead of $\{C^i_d\}$ here to highlight the fact that in this case, the parameter for factory-gate price distributions in (d,i) is a scaler instead of a function.

The third part of the proposition characterizes how equilibrium technology reacts to changes in trade costs. It shows that up to the first order, the change in technology are entirely summarized by: trade and input-output linkages and the technology choice for each (d,i) in the observed equilibrium, and structural elasticities including $\bar{t}$, $\bar{\phi}$, and ζ. Loosely speaking, the terms in the second bracket measure how changes in tariffs affect the technology choice of each country, taking other countries' choices as given; the Leiontief inverse in the first component captures the equilibrium amplification of the technology choices between countries.

To sharpen the intuition, we consider an increase in the cost of importing from (o,j) by a country-sector (d,i) that is small relative to the rest of the sector and countries.

Proposition 4. *Consider a country-sector (d,i) that is small in the sense that its input and output account for a negligible share of all countries and sectors, including sectors in country d. Then after an x % increase in the cost of (d,i) importing from (o,j):*

1. *The distance between θ_d^i and θ_o^j change by:*

$$\Delta\|\theta_d^i - \theta_o^j\| = -\frac{\zeta\omega^i\gamma^{ij}\bar{\chi}_{do}^{ij}\|\theta_o^j - \vartheta_d^{ij}\|}{1 - 2t\zeta\omega^i\sum_{j',o'}\gamma^{ij'}\bar{\chi}_{do'}^{ij'}\|\theta_{o'}^{j'} - \vartheta_d^{ij'}\|} \times \frac{\theta_d^i - \theta_o^j}{\theta_o^j - \vartheta_d^{ij}} \times x,$$

where $\vartheta_d^{ij} \equiv \sum_m \bar{\chi}_{dm}^{ij}\theta_m^j$ is the average location of the suppliers of (d,i) that is in sector j.

2. *$\|\theta_d^i - \theta_o^j\|$ increases relative to the expenditure-share weighted distance between θ_d^i and $\theta_{o'}^j$ across $o' = 1,...,N$ increases. More precisely,*

$$\Delta\|\theta_d^i - \theta_o^j\| - \sum_{o'}\bar{\chi}_{do'}^{ij}\Delta\|\theta_d^i - \theta_{o'}^j\| = \frac{\zeta\omega^i\gamma^{ij}\bar{\chi}_{do}^{ij}\|\theta_o^j - \vartheta_d^{ij}\|}{1 - 2t\zeta\omega^i\sum_{j',o'}\gamma^{ij'}\bar{\chi}_{do'}^{ij'}\|\theta_{o'}^{j'} - \vartheta_d^{ij'}\|} \times x > 0 \qquad (22)$$

Proof. See Appendix A.4. □

Part 1 of the proposition describes how the cost of sourcing input from (o,j) affects $\|\theta_d^i - \theta_o^j\|$. As the first term on the right-hand side of the equation is positive, whether $\Delta\|\theta_d^i - \theta_o^j\|$ is positive or negative depends on the sign of $(\theta_d^i - \theta_o^j)/(\theta_o^j - \vartheta_d^{ij})$, which reflect a subtle third-country effect. An increase in the cost of importing from (o,j) leads firms in (d,i) to depend more heavily on other suppliers, pushing their technology in the direction of these suppliers. Somewhat subtle, how this change affects $\|\theta_d^i - \theta_o^j\|$ depends on the relative position of θ_d^i, θ_o^j, and ϑ_d^{ij}. If θ_d^i and ϑ_d^{ij} are on the same side of θ_o^j, then as firms switch to other suppliers, θ_d^i moves away from θ_o^j; conversely, firms' move towards ϑ_d^{ij} end up reducing $\|\theta_d^i - \theta_o^j\|$.

Despite this ambiguity, Part 2 shows that θ_d^i moves away from θ_o^j *relative to* other trade partners. The size of the increase in relative distance is governed by the first term on the right-hand side of equation (22). All else equal, the increase is larger if $\bar{\phi}$ is small, if (d,i) rely more heavily on (o,j) for inputs (larger $\gamma^{ij}\bar{\chi}_{do}^{ij}$), and if θ_o^j is further away from the (import share-weighted) average position of the suppliers of $d-i$ (larger $\|\theta_o^j - \vartheta_d^{ij}\|$). In quantification, we will use this relationship to calibrate the model.

Proposition 4 illustrates how trade costs shape firms' technology choices. As discussed previously, firms' own endowment technology also affects their choice, which can in turn affect trade. To illustrate this mechanism, we consider a change in the endowment technology of a zero-measure set of firms from the endowment common to all other firms in (d,i). Because these firms account for zero share of the sector output, their change does not affect aggregate outcomes, which simplifies the exposition. We establish the following:

Proposition 5. *Suppose firms in (d,i) have an endowment technology of $\bar{\theta}_d^i$ with probability 1 but a zero-measure of set of firms in (d,i), denoted by ν, have an endowment of $\bar{\theta}(\nu)$. Then in response to a change in $\bar{\theta}(\nu)$ that reduces $\|\bar{\theta}(\nu) - \theta_o^j\|$,*

1. *Firm ν moves closer to θ_o^j, namely $\|\theta_d^i(\nu) - \theta_o^j\|$ decreases*

2. *Firm ν is more likely to purchase from (o,j)*

3. $\Delta \log\left(\chi_{do}^{ij}(\nu)/\chi_{dd}^{ii}(\nu)\right) = -2\zeta\bar{t} \cdot \Delta\|\theta_d^i(\nu) - \theta_o^j\|$

Proof. See Appendix A.5. □

Parts 1 and 2 of the proposition indicate that as the endowment of these firms moves towards (o,j) in the technology space, their choice shift in the same direction. This, in turn, leads to an increase in the probability of purchasing from (o,j). Part 3 shows the the elasticity of the odds ratio with respect to the change in the technology distance between the firm and (o,j) is the product of the conventional trade elasticity ζ and parameter t which governs the importance of technology compatibility on input efficiency. The proposition shows that we can pin down t by inspecting among firms facing the same trade environment, whether firms' proximity to the technology of a country is correlated with importing from that country, a result that we will exploit for identification later.

2.7 Special Cases: The Welfare Implications of Technology Choice

In this subsection, we discuss the welfare effect of endogenous technology choice. In our model, firms make optimal technology choices to maximize their own profits. Their decision, however, affect the profit of downstream users. Such an externality propagates to other sectors and countries through input-output linkages. We zoom into the nature of the externality through two special cases, highlighting the roles of domestic and international spillovers, respectively. As in Section 2.6, we assume that the ex-ante distribution in each (d,i) pair is degenerate.

Proposition 6. *Consider a closed economy with multiple sectors and each sector with an ex-ante endowment location $\bar{\theta}^i$, $i = 1, .., S$.*

i. The marginal impact of increasing θ^i on the social welfare, $\frac{\Delta \ln(U)}{\Delta \theta^i}$, is given by

$$2\rho^i \Big[\underbrace{\frac{\exp\big(-\bar{\phi}(\theta^i - \bar{\theta}^i)^2\big)}{\eta - \sum_i \rho^i \exp\big(-\bar{\phi}(\theta^i - \bar{\theta}^i)^2\big)} \bar{\phi}(\bar{\theta}^i - \theta^i)}_{\text{income effect}} - \underbrace{\bar{t} \sum_j \tilde{\gamma}^{ij}(\theta^i - \theta^j)}_{\text{sector-}i\text{ price}} \Big] - \underbrace{2\,\bar{t} \sum_{j \neq i} \rho^j \tilde{\gamma}^{ji}(\theta^i - \theta^j)}_{\text{other sector prices}},$$

where the three terms capture the income effect, the price effect in sector i, and the price effect in all other sectors; $\tilde{\gamma}^{ij}$ is the general equilibrium impact of sector j price on sector i price, defined as $\tilde{\gamma}^{ij} \equiv \sum_m \Omega^{im} \gamma^{mj}$, where Ω^{im} is the (i, m)-th element of $(\mathbb{I}_{NS \times NS} - \Gamma)^{-1}$.

ii. If sectors have the same weights in the final consumption and symmetric input-output structure, i.e., for all $i \neq j \neq j'$, $\rho^i = \rho^j$, $\gamma^{ii} = \gamma^{jj}$ and $\gamma^{ij} = \gamma^{ij'} = \gamma^{jj'}$, then the equilibrium $||\theta^i - \bar{\theta}^i||$ is too small. In other words, firms under-invest in technological adaption.

Proof. See Appendix A.6. □

The first part of the proposition characterizes the three channels through which a change in θ^i affect social welfare. Without loss of generality, suppose $\bar{\theta}^i > \theta^i$, which means an increase θ^i reduces the adaption of firms in i toward other sectors. This creates three effects. First, it saves adaption cost, which are rebated to households for consumption (first term). Second, it causes a change in the price of goods j (second term), in which $\tilde{\gamma}^{ij}$ captures the GE effect of the distance to θ^j on the price of sector i. Compare this to the the optimal condition characterizing θ^i. Finally, when firms in a sector shifts technology towards the technologies of their suppliers, their suppliers, who are also down-stream users of the firms also benefit through lower cost of sourcing. Because of this, firms under-invest in adaption, i.e., they stay too closer to their endowment technology. When the benefit of moving away is large (large t), this effect is stronger. This effect is also amplified dy the input-output linkage of the economy, measured by $\frac{\tilde{\gamma}^{ij}}{\gamma^{ij}}$.

Contrasting this with the first order condition that characterize firms' private choice:

$$\rho^i \big[\frac{1}{\eta - 1} \bar{\phi}(\bar{\theta}^i - \theta^i) - \bar{t} \sum_j \gamma^{ij}(\theta^i - \theta^j)\big] = 0, \tag{23}$$

we can see that both the adaption cost and sector-i price show up in firms' optimization problem in different forms. In particular, wheras the compatibility with suppliers enters firms' problem with a coefficient of γ^{ij}, it enters the social welfare with a coefficient of $\tilde{\gamma}^{ij}$, reflecting the propagation via input-output linkages that individual firms do not consider. In addition, individual firms also do not consider the spillover effects to other sectors, which affects welfare.

For the above reasons, at the decentralization equilibrium, firms' technology choice generally do not maximize social welfare. Without additional restrictions, however, we cannot sign $\Delta \ln(U)$. Intuitively, as firms move away from $\bar{\theta}^i$, they move closer to some suppliers but away from other suppliers, so this movement generate both positive and negative externalises on other firms. The second part of the proposition shows that when sectors have symmetric input-output structures, the positie externality always dominate and moving further away from the firm's endowment technology will improve the social welfare.[15]

Proposition 7. *Consider an open economy with one sector with roundabout production and two symmetric countries, country* 1 *and* 2*. Assume WOLG that in equilibrium,* $\theta_2 < \theta_1$*. Then the effect of a move of country* 2*'s technology towards country* 1 *from the equilibrium on welfare is:*

$$\frac{\Delta \ln U_2}{\Delta \theta_2} = \frac{2\exp(-\bar{\phi}(\theta_2 - \bar{\theta}_2)^2)}{\eta - \exp(-\bar{\phi}(\theta_2 - \bar{\theta}_2)^2)}\bar{\phi}(\bar{\theta}_2 - \theta_2) + 2\bar{t}\frac{1-\gamma^L}{\gamma^L}\bar{\chi}_{12}(\theta_1 - \theta_2) > 0$$

$$\frac{\Delta \ln U_1}{\Delta \theta_2} = 2\bar{t}\frac{1-\gamma^L}{\gamma^L}\bar{\chi}_{12}(\theta_1 - \theta_2) > 0$$

Proof. See Appendix A.7. □

The result on $\ln(U_2)$ highlights two departures of the social welfare between firms' incentive in our model, which already appears in Proposition 6: the different importance of the household and the firm place on the innovation expense, and the price effect. Here, even though there is only one sector, the GE price impact due to a change in technology location is different from the effect internalized by the firm. Intuitively, when firms in a country shift technology toward that of other countries, these firms benefit directly as the users of foreign inputs. Consumers benefit more because the reduction in the firms' production cost is amplified via the round-about production and ultimately benefits the consumer. Thus, due to within-country externality, firms tend to under-invest in adapting to foreign technologies relative to the local social planner.

In addition to this within-country externality, the proposition also highlights an international externality. By moving towards the technology position of country 1, firms in country 2 generate a positive externality as now firms in country 1 can source inputs more cheaply, a cost saving that also gets amplified by the production network.

This mechanism has important implications for the welfare effects of trade openness. If, in response to trade liberalization, countries' technology converges, then this endogenous change in technology amplifies the welfare gains from trade liberalization. Similarly, if after a trade war, countries' technologies drift apart, then the drift in technology again amplifies the welfare losses from the trade war.

2.8 From Theory to Data

Until now, we have been agnostic about the nature of 'technology' and the distance between technologies. Indeed, the core mechanism in our model encompasses several views about

[15] When a firm's decentralized equilibrium choice is in the left or right of all other sector, there is only positive externality, in which case we can also show firms under invest in technological adaption.

why firms might choose different technologies and interact with different customers and suppliers. One is to view firms in an industry with distinct technology as belonging to different global value chains. Consider the broader category of the 'Manufacture of machinery and equipment'. Within this sector, we find both tractor manufacturers and construction machinery producers. Although they operate under the same umbrella sector, these groups cater to distinct downstream users—agriculture and construction, respectively—and likely source from different suppliers. Even among tractor manufacturers, differences in product offerings can lead to distinct suppliers and customers. Through the lens of our model, such differences can be viewed as firms belonging to different supply chains embodied in θ that standard sectoral-level trade data do not capture.

We take a different view. Instead of treating θ as the residual that explains firms' deviations from sectoral-level averages, we map θ to the technology embodied by a country's patents, and the distance between different technologies of firms to the similarity between these patents, which we measure using patent citations.

To derive a measure of citations from the model, we embed a problem of knowledge attribution through (backward) citation into the model. Consider a firm in (d, i) with technology θ. This technology is potentially related to technologies of all other firms around the world. We assume that firms are required to acknowledge through citations all technologies that are sufficiently close to their own. Let H_o^j be the stock of technologies in (o, j). Then this requirement implies that the total citations to technologies in (o, j) is $H_o^j d\Theta_o^j(\theta)$, in which $d\Theta_o^j(\theta)$ is the frequency of the technologies in (o, j) that is around θ. We can derive the composition of a citation made by a firm or a country to other firms and countries. For example, the share of citation made to technologies in country o by this firm is $\sum_j H_o^j d\Theta_o^j(\theta) / [\sum_{o',j'} H_{o'}^{j'} d\Theta_{o'}^{j'}(\theta)]$; the aggregate share across all firms from (d, i) is $\int_T \sum_j H_o^j d\Theta_o^j(\theta) / [\sum_{o',j'} H_{o'}^{j'} d\Theta_{o'}^{j'}(\theta)] d\Theta_d^i(\theta)$.

This formulation suggests that, if among the universe of world patents, a firm cites heavily those invented in country o, then it can be either because a higher fraction of country o's technology overlaps with the technology of the firm (large $d\Theta_o^j(\theta)$) or because the stock of technologies there is large or of high quality(large H_o^j). Therefore, to isolate the variations in citation patterns it is important to account for the difference in the stock and quality of technologies across countries. Guided by this observation, when we test Proposition 4 and 5 in Section 3, we control for the stock (quality) of the technology of countries through fixed effects; when we parameterize the model in Section 4, we adjust for the quality/stock of countries' technologies using citation data.

3 Reduced-Form Evidence

In this section, we provide reduced-form evidence for the mechanisms discussed in Section 2.6. The estimates will help us pin down key structural parameters of the model.

3.1 Tariff Variation and Technological Choice

Proposition 4 suggests that a decrease in the cost of importing from (o, j) by firms in d would incentivize these firms to adopt technologies closer to those of the firms in (o, j). To test this prediction, we examine how technology proximity, measured by patent citations, is shifted by exogenous variations in import costs prompted by the changes in the Most-Favored-Nation (MFN) tariffs.

Data. We assemble a dataset on bilateral patent citations and tariffs across different countries and industries. The dataset is a balanced panel that spans 28 geo-political regions (d and o), 127 manufacturing industries (j), over 5 three-year periods (t) from 2000 to 2014.[16]

Each observation d-o-j-t records the average tariff faced by the *output* of region-industry (o, j) in destination market d in period t, alongside the total number of patents invented in region-industry (o, j) that are cited by patents invented in region d in period t.

Bilateral tariff data are sourced from UN TRAINS, where we obtain both the effectively applied and the MFN tariff rates at the country-industry-year level. To facilitate analysis, we aggregate the tariff rates from country pairs into region pairs and from yearly figures into three-year periods by computing simple averages, while maintaining the granularity of industry (j) at the 4-Digit ISIC (Rev. 3) level.

Bilateral citation data are obtained from PATSTAT, where each citation record allows us to identify the unique patent identities for both the citing and cited patents. We designate d as the inventor region of the citing patent, o as the inventor region of the cited patent, and t as the period in which the citing patent was initially applied. To link cited patents with relevant industry classifications, we map the 4-Digit International Patent Classification (IPC) symbols to the 4-Digit ISIC (Rev. 3) industry categories (j) using the crosswalk provided by Lybbert and Zolas (2014). We aggregate all citation records to form a panel at the d-o-j-t level, which we then merge with the tariff data. The appendix provides detailed information on our dataset construction process.

Specification. To investigate how import tariffs impact the direction of technology choice, we employ the following specification:

$$\ln \text{Citation}_{dojt} = \beta \ln \tau_{dojt} + FE_{ojt} + FE_{doj} + FE_{d(j)t} + \epsilon_{dojt}, \tag{24}$$

where Citation_{dojt} is the total number of patents invented in (o, j) that are cited by patents in d invented in t, and τ_{dojt} is the applied gross ad-valorem tariff. Our identification strategy crucially relies on a set of fixed effects. Firstly, the inclusion of origin-industry-time fixed effects, FE_{ojt}, controls for the vertical aspect of the citation probability, accounting for factors like patent quality and quantity of (o, j). This adjustment enables us to use patent citations as a measure of technology proximity. Secondly, we incorporate the destination-origin-time

[16] See the appendix for details on the aggregation of geo-political regions from countries.

Table 1: Tariff Variation and Technological Choice

	ln Citation$_{dojt}$			
	(1)	(2)	(3)	(4)
$\ln \tau^{MFN}_{dojt}$	-0.793***		-0.285**	
	(0.118)		(0.139)	
$\ln \tau_{dojt}$		-0.822***		-0.296**
		(0.123)		(0.144)
FE o-j-t	Yes	Yes	Yes	Yes
FE d-o-j	Yes	Yes	Yes	Yes
FE d-t	Yes	Yes		
FE d-j-t			Yes	Yes
Observations	243010	243010	242799	242799

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the Importer(d)-Exporter(o)-Industry(j) level. Industries are at a level of 4-digit ISIC (Rev. 3). Columns (1) and (3) report the reduced-form regression, and columns (2) and (4) report the second stage of 2SLS using MFN tariffs as the instrumental variable.

fixed effects, FE_{doj}, which absorb time-invariant components in sector-specific bilateral trade costs. Moreover, to address potential endogeneity and measurement error in applied tariffs, we instrument τ_{dojt} with τ^{MFN}_{dojt}, the tariff rates under the Most Favored Nation (MFN) principle, which mandates that a country applies the same tariffs to all its WTO trade partners. Finally, the destination-time fixed effects, FE_{dt}, allow us to interpret the citation variable in the context of within-destination citation shares. To further account for potential endogeneity stemming from destination-industry-time components, such as demand shifters like aggregate consumption, we also present results that include the destination-industry-time fixed effects, FE_{djt}.

Results. Table 1 reports the results. In column 1, we present the reduced-form outcome, regressing the logged number of citations on the logged MFN tariffs. In column 2, we provide the result of the second stage of the 2SLS method, using MFN tariffs as instruments for the applied tariffs. As expected, we observe a negative shift in technology proximity due to an exogenous increase in import costs. This impact is mitigated but remains statistically significant after controlling for destination-industry-time fixed effects (FE_{djt}), as demonstrated in columns 3 and 4. Notably, a 1% increase in import tariffs leads to a 0.3% decrease in total patent citations. As indicated in Proposition 4, all else equal, this elasticity is monotonically decreasing with respect to $\bar{\phi}$. Therefore, we rely on this coefficient for calibrating $\bar{\phi}$ in our quantification analysis.

Our finding is related to a large body of literature, starting with Jaffe et al. (1993), that uses citation flows to measure the extent of knowledge spillover. This research typically observes that a citation to a patent declines significantly with geographic distance, which is interpreted as evidence of localized knowledge spillovers. The rationale is that inventors in close proximity are more likely to encounter each other or a common intellectual stimulus, which leads to more closely related patents (that cite each other). Thus, even in the digital age where patent records are readily accessible online, geographically proximate inventors tend to cite each other more frequently. Following this reasoning, changes that bolster the interactions among inventors or their exposure to a common source of inspiration can lead to an increase in citations. In particular, Aghion et al. (2021) shows that after a French firm begins exporting to a destination country, its patents receive more citations from that country. We differ from Aghion et al. (2021) in two aspects. First, we use exogenous MFN tariffs to show the causal effect of trade costs on patent citations. Second, our model interprets the findings through the lens of the compatibility of imported goods with the technology chosen by domestic firms, complementing existing views of the literature based on knowledge spillovers via trade in goods. In support of this interpretation, in the appendix, we show that when looking at the data by importing sector, it is the input tariffs, rather than output tariffs, that affect technology choice.

3.2 Firm-Level Correlation between Trade and Technology Choice

Proposition 5 suggests that within a country-sector pair, (d, i), firms with an endowment draw closer to a foreign country o are more likely to choose a similar technology to firms in o and import from o. This results in a *firm-level* correlation between imports and technology proximity. We use Chinese firm-level data to test this implication.

Data. We compile a firm(ω)-region(o)-period(t) panel dataset that encompasses data on citations and imports, spanning the 27 geo-political regions (excluding China) and 5 three-year periods from 2000 to 2014. We outline the key steps involved in constructing our dataset here, with more detailed information provided in the appendix.

Our analysis focuses on Chinese manufacturing firms in the Annual Survey of Industrial Enterprise maintained by the National Bureau of Statistics of China (NBSC). The dataset offers detailed accounting information for all Chinese manufacturing firms with annual sales greater than US$800,000 over 1998–2014. Crucial to our analysis is a firm's identity (ω) and the prime industry (i) it belongs to. To link each firm consistently over time, we employ a procedure following Brandt et al. (2017) to create a unique identifier. We manually map the industry codes to the 2017 version to accommodate the constant changes in the Chinese Industry Classification (CIC) codes during the specified period.

We link the NBSC Database to patent data provided by China's State Intellectual Property Office (SIPO), which we further merge with citation data from the PATSTAT Global using unique patent identifiers. We extract the regions of invention of the patents cited by the patents of Chinese firms, which allows us to construct a ω-o-t panel detailing the citation patterns of firm ω's inventions in different periods.

Finally, we obtain information on firms' imports from China's General Administration of

Table 2: Firm-Level Correlation between Trade and Technology Choice

	$\text{IMPORT}_{\omega ot}$			$\ln(\text{Import})_{\omega ot}$		
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{CITATION}_{\omega ot}$	0.024***	0.023***	0.022***	0.054**	0.056**	0.051**
	(0.001)	(0.001)	(0.001)	(0.023)	(0.023)	(0.025)
FE ω-t	Yes	Yes	Yes	Yes	Yes	Yes
FE ω-o	Yes	Yes	Yes	Yes	Yes	Yes
FE o-t	Yes	Yes		Yes	Yes	
X_{oit}		Yes			Yes	
FE i-o-t-province			Yes			Yes
Observations	9108423	8771074	9080046	250659	249939	220814

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered by firm. Industries are Input-Output Industries equivalent to 3-digit CIC. X_{oit} are the average applied tariffs of importing the goods in sector j from world region o, weighted by the input share of sector j in sector i production.

Customs, which provides detailed records on the universe of all Chinese trade transactions by both importing and exporting firms at the HS eight-digit level for the years 2000-2014. We aggregate all import records to construct a ω-o-t panel, where each observation is the total value of goods imported by firm ω from region o during period t. This panel is then merged with the citation data.

Our final dataset includes all manufacturing firms with patents, regardless of whether they import goods from abroad. The panel is unbalanced with the number of firms increasing from 32,293 in the period 2000-2002 to 95,813 in the period 2012-2014.

Specification. Our regression specification is as follows:

$$\text{IMPORT}_{\omega ot} = \beta \text{CITATION}_{\omega ot} + \beta_2 X_{iot} + FE_{\omega t} + FE_{\omega o} + FE_{ot} + \epsilon_{\omega ot}, \tag{25}$$

where $\text{IMPORT}_{\omega ot}$ and $\text{CITATION}_{\omega ot}$ are binary variables, taking the value of one when firm ω imports or cites patents from region o during period t. Due to the granularity of firm-level data, we focus on the extensive-margin variation of citations, but we also explore the relationship with the intensive margin variation of imports. We control for firm-time ($FE_{\omega t}$) fixed effects, which account for firm-level shifters that affect the overall import and citation of a firm, and firm-region ($FE_{\omega o}$) fixed effects, which account for the time-invariant relationship between a firm and a region. We include region-time fixed effects (FE_{ot}) to address the vertical aspect of the citation probability, enabling us to use patent citations as a proxy for technology proximity. Lastly, X_{iot} stands for shocks that affect the imports of sector i from o in period t. In practice, we will either control for the average applied tariffs faced by firms in sector i across different input sectors or absorb X_{iot} with additional fixed effects.

Results. Table 2 reports the results. Column 1 shows that conditional on the fixed effects, firms citing patents from a specific region exhibit a 2 percentage point higher probability of importing from that same region. This correlation holds its significance and robustness as we control for industry-region-time-specific tariffs, as evidenced in Column 2. In Column 3, we further control for industry-region-time-province fixed effects. This set of fixed effects addresses the concern that firms' decisions regarding input sourcing and knowledge acquisition might be influenced by the behavior of their suppliers or peer competitors. The coefficient remains unchanged.

Columns 4 to 6 shift the focus to the logged value of imports as the dependent variable. The noticeable reduction in our sample size can be attributed to the fact that, in our dataset, only approximately 28% of firms engage in importing goods from abroad each year and on average, each firm imports from fewer than two distinct regions. The results reveal that conditional on the fixed effects, citing patents from a region corresponds to a substantial 6% increase in firm-level imports.

Through the lens of Proposition 5, this correlation captures the combined impacts of a firm's technology endowment on its endogenous technology choice and subsequent sourcing decisions. Intuitively, as firms' technology choice becomes more aligned with a particular partner, they source more inputs from that partner. Firms' technology can change due to a shift in their technology endowment—-through the recruitment of researchers with expertise in areas closer to a foreign country, for example.[17] While the estimate does not reflect a causal relationship—both choices are driven by unobserved technology endowment—Proposition 5 suggests that the estimate maps directly into the structural parameters of the model. We therefore rely on this estimate to calibrate $\bar{t}$ in quantification.

4 Quantification

In this section, we use the model to quantify the impacts of trade on technology choice, and the joint impacts on welfare. We first describe the functional form assumption and approximation invoked in this section that enable closed-form solutions to firms' technology choice; we then explain how we parameterize the model; lastly, we conduct counterfactual exercises.

4.1 Functional-Form Assumptions and Numerical Algorithm

In our model, heterogeneous firms interact with each other directly (see e.g., equation (3)), rather than through aggregate prices as in many quantitative trade models (e.g., Eaton and Kortum, 2002). This means that to solve the model, we need to keep track of the entire distribution of Θ_d^i rather than its first moment for all (d, i). For tractability, we impose the following structure on technology endowment distributions:

[17] Our model is static, so firms only make one-shot technology decisions. In a dynamic extension of the model, in each period firms can build on the choice of the previous period and continue to adapt. In such an extension, the variation in endowment technology can be due to firms' own past choices.

Assumption 3. *For each (d,i), the endowment technology distribution is Normal with a mean of $\bar{\mu}_d^i$ and a variance of $(\bar{\sigma}^i)^2$. That is, $\bar{\Theta}_d^i \sim \mathcal{N}(\bar{\mu}_d^i, (\bar{\sigma}^i)^2)$*

Under this assumption and quadratic approximations of $\ln C_o^j(\theta)$—the log production cost in (o,j) for a firm choosing technology θ—around the mean ex-post technology of (o,j), we obtain closed-form solutions to firms' technology choice problem and the ex-post technology distribution in each (o,j), which we summarize in the following proposition:

Proposition 8. *Under Assumptions 3, up to a second-order approximation of $\ln C_o^j(\theta)$:*

1. *$\ln C_o^j(\theta)$ is a quadratic function of θ characterized by:*
$$\ln C_o^j(\theta) = k_{A,o}^j + m_A^j(\theta - n_{A,o}^j)^2; \tag{26}$$
where $\{m_A^j, n_{A,o}^j, k_{A,o}^j\}$ depend only on model primitives and wages $\{w_o\}$.

2. *Firms' technological choice $g_o^j(\cdot)$ is characterized by*
$$g_o^j(\bar{\theta}) \equiv \alpha_o^j + \beta^j \bar{\theta}, \tag{27}$$
with $\alpha_o^j = \frac{(\eta-1)m_A^j}{\bar{\phi}+(\eta-1)m_A^j} n_{A,o}^j$ and $\beta^j = \frac{\bar{\phi}}{\bar{\phi}+(\eta-1)m_A^j}$.

3. *The ex-post technology distribution of (o,j) is*
$$\Theta_o^j \sim \mathcal{N}(\mu_o^j, (\sigma^j)^2), \text{ with } \mu_o^j = \alpha_o^j + \beta^j \bar{\mu}_o^j \text{ and } \sigma^j = \beta^j \bar{\sigma}^j. \tag{28}$$

Proof. See Appendix C.1. □

Proposition 8 expresses the cost function and firms' policy function in closed form.[18] In particular, equation (27) shows that firms' technology choice is a linear function of its endowment technology, which has two attractive implications for quantitative implementation. First, with the analytical expression in hand, we no longer need to solve problem (13) numerically. Second, under the normal assumption of endowment technology distribution, this policy function implies normal distributions for ex-post technology distributions, with the parameters mapped one-to-one to the ex-ante distribution as described in equation (28). This result will prove useful for model calibration, as we discuss below.

Building on this proposition, we devise the following algorithm to solve the model:

1. Given the model fundamentals and wages $\{w_o\}$, solve for $\{m_A^j, n_{A,o}^j, k_{A,o}^j\}$, which delivers the cost functions $\{C_o^j(\theta)\}$, firms' technology choice, and the ex-post technology distributions. We show in the appendix that $\{m_A^j, n_{A,o}^j, k_{A,o}^j\}$ can be solved using a contraction mapping algorithm as a function of fundamentals and $\{w_o\}$.

[18]Proposition 8 can be extended to allow the variance of the ex-ante technology distribution to vary not only by sector but also by country (i.e., $\bar{\sigma}_d^i$ instead of $\bar{\sigma}^i$); it can also be generalized to multi-dimensional θ with multivariate normal ex-ante technology distribution. It turns out that the parsimonious one-dimensional setup in Assumption 3 can already fit the data well.

2. With $C_o^j(\theta)$ at hand, evaluate the sourcing decisions of intermediate firms $\chi_{do}^j(\theta, \tilde{\theta})$ and final-good producers $\pi_{do}^j(\theta)$ for all $\theta, \tilde{\theta} \in T$.

3. Combine equations (16) to (19) to arrive at a system of equations of $\{X_o^j(\theta)\}$ and $\{M_o^j(\theta)\}$, taking as given $\{w_o\}$. We discretize the domain of θ, in which case the system of equations is linear in $\{X_o^j(\theta)\}$ and $\{M_o^j(\theta)\}$ and can be easily solved.

4. Evaluate if equation (20) is satisfied under $\{X_o^j(\theta)\}$ and $\{M_o^j(\theta)\}$. If yes, then we have found an equilibrium of the model; if not, update $\{w_o\}$ and return to step 1.

4.2 Calibration Procedure

This subsection explains how the parameters are calibrated.

Parameters externally calibrated. We set the elasticity of substitution $\eta = 5$, which implies a cost elasticity of 4 and 20% markup for sales to final-good producers. We set the shape parameter of match-specific productivity $\zeta = 4$. This value means that the conditional on their technology, the trade elasticity between any pairs of intermediate-good producers is 4, as shown in equation (6).[19]

Labor Endowments $\{L_d\}$ are calibrated to population of each country from Penn World Table. Production function parameters are calibrated with data from World Input-Output Database (WIOD). The country-specific sector shares in final-good production ρ_d^i are calibrated to the household consumption shares of sector i in country d. The input shares in intermediate-good production γ^{ij} and γ^{iL} are calibrated to the input-output weights of sector-j input and value-added in the production of sector i, respectively. All these values are taken over the average of 2010-2014.

Technology distribution. The ex-ante technology distribution of countries are by assumption not observed. To calibrate $\{\bar{\mu}_d^i, \bar{\sigma}^i\}$, we use two pieces of information: the ex-post technology distribution, and the one-to-one mapping from the ex-ante to the ex-post distributions characterized in Proposition 8. As the mapping depends on all model primitives (such as trade costs) and the equilibrium wage, the ex-ante distributions cannot be recovered independent of the rest of the model.

For transparency and tractability, we recover the ex-ante distribution in two steps. In the first step, we choose the parameters governing the ex-post distributions, $\{\mu_d^i, \sigma^i\}$, to match patent citation statistics. This step can be carried out without knowing the primitives of the model. Conditional on their own technology, firms' sourcing decisions only depend on the ex-post distributions. We can therefore calibrate the primitives of the model governing trade using only the ex-post distributions and other data. Importantly, our model allows

[19] Aggregate trade elasticity could be because of endogenous technology and changes in the composition of firms. In the calibrated model, we find that trade elasticity is close to 4 for small changes in bilateral trade costs.

us to calibrate these parameters without specifying $\bar{\phi}$, the parameter that governs the extent to which the calibrated ex-post distributions are shaped by ex-ante distributions versus by firms' endogenous technology choices. In the second step, we calibrate $\bar{\phi}$ and recover the ex-ante distributions using equation (28).[20]

To calibrate the ex-post distribution, we use the model-implied citation shares grounded in the knowledge-source attribution problem described in Section 2.8, extended to account for the fact that each output sector relies differently on the knowledge of different input sectors. Concretely, for any firm from sector i that chooses technology location θ, we calculate the share of citations it makes that goes to (o, j) as

$$\psi_o^{ij}(\theta) \equiv \delta^{ij} \cdot \frac{H_o^j \cdot \mathrm{d}\Theta_o^j(\theta)}{\sum_{o'} H_{o'}^{j'} \cdot \mathrm{d}\Theta_{o'}^{j'}(\theta)}, \tag{29}$$

where δ^{ij} is the sectoral technology proximity, measured by the share of patent citations made to sector j by sector i, and H_o^j is the total number of citations received by patents invented in (o, j) in data. By using citation counts to measure H_o^j, this share accounts for the difference in the vertical quality of patents across (o, j); by using δ^{ij} to weight supplier sectors, this share allows for the possibility that (o, j) receive more citation from (d, i) because sector i is technically more dependent on j than other sectors.

We integrate $\psi_o^{ij}(\theta)$ over the technology distribution in (d, i) to obtain aggregate bilateral citation shares:

$$\Psi_{do,model}^{ij} \equiv \int \psi_o^{ij}(\theta) \mathrm{d}\Theta_d^i(\theta),$$

where $\sum_{o,j} \Psi_{do,model}^{ij} = 1$. Plugging in δ^{ij} and H_o^j, for any $\{\Theta_o^j\}$, this expression delivers the model-implied share of citations made by (d, i) to (o, j). We construct the citation data using the universe of patents from PATSTAT aggregated over 2010-2014. We then find $\{\mu_d^i, \sigma^i\}$ by solving the following nonlinear least square problem:

$$(\mu_d^i, \sigma^i) = \arg\min \sum_{o,d,i,j} (\Psi_{od,model}^{ij} - \Psi_{od,data}^{ij})^2.$$

Trade costs and distribution of production techniques. With the ex-post technology distributions $\{\Theta_o^j\}$ at hand, we jointly calibrate the parameters $\{\Xi^i / A_d^i\}$, which determines the productivity of (d, i), to match the output share of (d, i) in industry i, and calibrate $\{\tau_{do}^j, \tau_{do}^{Uj}\}$ to match the trade shares of intermediate and final goods, respectively.

Technology compatibility $\bar{t}$. In the model, the parameter $\bar{t}$ on technology compatibility governs the input-sourcing decisions of firms given their chosen technology locations. We

[20]Conditional on calibrated ex-post distributions and other general equilibrium outcomes (such as wages and prices), parameter $\bar{\phi}$ only affects the split of firms' markup between profit to the representative consumer and the expense on adaption. Therefore, changing $\bar{\phi}$ after the rest of the model has been calibrated affects only the welfare of agents but not equilibrium wages or prices. For this reason, we can calibrate $\bar{\phi}$ at the end of the algorithm without interfering with the earlier steps. See Appendix C.3 for additional details.

Table 3: Summary of Model Parameters

Parameters	Descriptions	Value	Target/Source
A. Externally Calibrated			
$\gamma^{ij}, \gamma^{iL}, \alpha^{j}$	IO Structure and Consumption Share	-	WIOT; $N = 15, S = 19$
L_d	Labor Endowment	-	PWT
$\eta, \zeta - 1$	Trade Elasticity	4	Literature
B. Just-Identified			
$\tau^{j}_{do}, \tau^{Uj}_{do}$	Iceberg trade costs		Bilateral trade shares
$\bar{\phi}$	Adaption cost	0.005	Country-sector-level citation-tariff elas.: -0.296
$\bar{t}$	Compatibility incentive	0.05	Firm-level Import-citation corr: 0.022
C. Nonlinear Least Square			
$\bar{\mu}^{i}_{d}, \bar{\sigma}^{i}$	Ex-ante Technology Distribution	-	Bilateral Citation Shares

calibrate $\bar{t}$ by matching the extensive-margin import-citation elasticity (Column 3 of Table 2), using the simulated method of moments. Given the ex-post technology distribution $\{\Theta^{i}_{d}\}$ and the equilibrium coefficients, a firm with technology location θ from (d, i) would source input from (o, j) with probability $\chi^{ij}_{do}(\theta)$ given by equation (6) and would cite patents in (o, j) with probability $\psi^{ij}_{o}(\theta)$ given by (29). For each value of $\bar{t}$, we simulate 190,000 Chinese firms ($d = CHN$), 10,000 from each sector i from the ex-post distribution Ψ^{i}_{d}. Then, we regress the realized extensive margin of importing from each country o on the extensive margin of citing patents from o, controlling for firm and country-industry fixed effects. We calibrate $\bar{t} = 0.05$ to match this coefficient to 0.022, the extensive-margin import-citation correlation obtained from the regression results in Column 3 of Table 2.

Innovation costs $\bar{\phi}$. We calibrate the parameter $\bar{\phi}$ on innovation costs to match the citation elasticity of tariffs (Column 2 of Table 1). As shown in Proposition 4, conditional on trade shares and γ^{ij}, this elasticity identifies $\bar{\phi}$. To obtain the elasticities in the model, we conduct counterfactuals that decrease the trade costs $\{\tau^{j}_{do}\}$ by 5% for all (d, o, j)'s with $d \neq o$ and χ^{j}_{do} greater than 5%, one (d, o, j) at a time. For each counterfactual, we calculate the implied elasticity for the change in the share of citations made by country d that goes to (o, j). We then adjust the value of $\bar{\phi}$ such that the mean of the elasticities calculated across these simulations matches the regression coefficient -0.296. The calibrated value is $\bar{\phi} = 0.005$.

Numerical implementation. The parameters are calibrated using a two-step procedure. We summarize the key steps here, delegating all details to the appendix. In the first step, we search for the ex-post distribution parameters $\{\mu^{i}_{d}, \sigma^{i}\}$ that best match the empirical patent citation shares. Given $\{\mu^{i}_{d}, \sigma^{i}\}$, the second step involves a nested algorithm. In the outer layer, we choose $\bar{t}$ to match the regression coefficient on the extensive-margin import-citation elasticity. In the inner loop, we solve for the competitive equilibrium, while simultaneously choosing $\{\tau^{j}_{do}, \tau^{Uj}_{do}, \Xi^{i}/A^{i}_{d}\}$ to match the trade shares. We then simulate the model and obtain the regression coefficient to feed into the outer loop. Once the procedure is complete, we

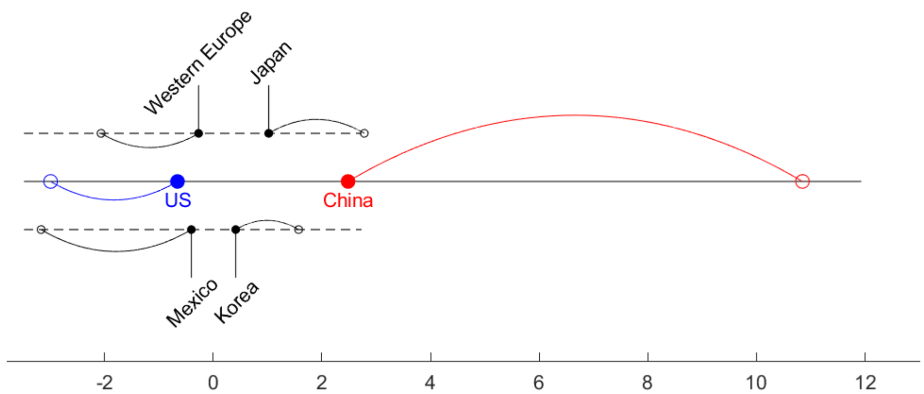


Figure 2: Average Mean of Technology Distributions

Note: Circles are ex-ante mean, and dots are ex-post mean.

choose $\bar{\phi}$ to match the country-sector-level citation-tariff elasticity and invert the equations in Proposition 8 to obtain the corresponding ex-ante distribution parameters $\{\bar{\mu}^i_d, \bar{\sigma}^i\}$.

Table 3 summarizes the calibration. We discuss the implications of our calibration in the next subsection.

4.3 Calibrated Technologies and Model Fit

Calibrated technologies. Figure 2 shows the average (across sectors) locations of ex-ante and ex-post technology distributions for a few sample countries. The dots depict the locations of ex-post technology. Among the major economies, the technologies of the U.S. and China fall at the two ends of the specturm, whereas the technologies of Japan, Korea, and Western Europe fall in the middle. This reflects the fact that, controlling for patent stock and quality (H^j_o) and sectoral differences in the compositions of knowledge source (δ^{ij}), Chinese and American patents still cite each other less intensively than they cite European, Korean, and Japanese patents.

The circles depict the average locations of ex-ante technologies for these countries. Two patterns are worth discussing. First, countries' ex-ante technologies are clustered spatially, with East Asian countries falling on the right and Western economies clustered on the left. This could reflect the role of geographic or cultural factors in shaping the endowment technology.[21] Second, technology compatibility incentives bring countries' technologies closer, with countries that are farther out—China and Mexico—moving more. International trade plays an important role here. Indeed, if countries do not engage in trade, then firms will

[21] Part of the difference could be due to spatial diffusion of ideas. One can view the ex-ante technology distribution in our model as capturing endowments shaped by various factors of countries, including idea diffusion.

Table 4: Bilateral Citation Shares: Model v.s Data

	Citation Share in Data			
Citation Share in Model	(1)	(2)	(3)	(4)
at Ex-post Tech. Dist.	0.855			
	(0.002)			
with Identical Tech.		0.657		
		(0.003)		
at Ex-ante Tech. Dist.				0.709
				(0.001)
Fixed Effects	-	-	ij, oj	-
Observations	81,225	81,225	81,225	81,225
Adjusted R^2	0.688	0.303	0.198	0.377

Note: This table assess the goodness of fit in the model. Each column reports the regression of the citation share in data on a measure of citation share in the model. Column (1) uses the ex-post technology distribution $\{\mu_o^j, \sigma^j\}$. Column (2) restricts to the case where $\mu_o^j = 0$ and $\sigma^j = 0$ for all (o, j). Column (3) regresses the data on o-j and i-j fixed effects. Column (4) restricts the technology distribution to the ex-ante distribution $\{\bar{\mu}_o^j, \bar{\sigma}^j\}$.

move towards the average technologies of their own countries.

Assessing the fit of the model. We assess the fit of the calibration. Recall that the bilateral citations are defined at $(d, i) \times (o, j)$ level. To fully account for bilateral distance between these many points through the choice of points in a metric space, one needs the metric space to have up to $\frac{1}{2}N \times S^2 - 1$ dimensions.[22] Our calibration of ex-post technologies attempts to do so through the choice of locations in the real line. To evaluate the fit, in the first column of Table 4, we regress the data on the values from the calibrated model. The coefficient is 0.855 and the R^2 is 0.688.

We compare the fit of our calibration to two alternative models. First, we consider a calibration with all firms having the same technology, i.e., $\mu_o^j = 0, \sigma^j = 0$ for all (o, j). In this case, variations in bilateral citation shares arise only from δ^{ij} and H_o^j. Regression of the data on this measure, reported in Column 2 of Table 4, has an R^2 of 0.303, which suggests that most of the model's explanatory power comes from the horizontal differences in technologies between countries and sectors, not from the measured δ^{ij} and H_o^j. Second, we consider a statistical exercise, in which we regress the data on o-j and i-j fixed effects, capturing the variations within δ^{ij} and H_o^j. This specification allows more flexible variations across δ^{ij} and H_o^j but did not use the model's implied multiplicative structure. As a result, despite there being $N \times S + S^2$ flexible parameters—more than the number of free parameters in the calibration—the R^2 of this regression, reported in Column 3, is only a third of that in the first

[22] For example, to account for the distance between two points, we can choose the points on a real line; to account for bilateral distance between three points, we need a two-dimensional space; to account for bilateral distance between four points, we need a three-dimensional space.

Table 5: Innovation Costs and Technology Compatibility Costs as Shares of GDP

Country/Region	**Tech Compat. Costs (t_d)**	**Tech Compat. Costs for Foreign Inputs**
BRA	2.64	0.66
CAN	2.31	0.96
CEU	2.58	1.03
CHN	6.60	2.19
IND	2.75	0.72
IDN	3.17	1.06
JPN	3.04	1.25
KOR	3.23	1.52
MEX	2.96	1.26
OCE	2.11	0.87
ROW	3.08	1.54
RUS	2.20	0.57
TUR	2.60	0.83
USA	2.27	0.67
WEU	2.20	0.55
World	**3.41**	**1.16**

column. Together, these two comparisons suggest that our model, though parsimonious, can explain most of the variations in the data.

Decomposition of model fit. We decompose the fit of the calibration to examine the importance of the endogenous technology choice in explaining the data. In Column 4 of Table 4, we regress the data on the model-implied citation shares if all country-sector pairs have their endowment technology. The R^2 decreases to 0.377. Comparing the findings between columns, about $\frac{30.3\%}{68.8\%}$ of the explanatory power comes from H_o^j and δ^{ij}, about $\frac{37.7-30.3=7.4\%}{68.8\%}$ comes from the differences in the endowment technologies of countries, and the remaining $\frac{68.8-37.7=31.1\%}{68.8\%}$ comes from the endogenous change in technology due to the compatibility incentive.

The cost of being incompatible with suppliers. To provide another view on the importance of compatibility incentives, we calculate the total costs firms bear due to not being perfectly compatible with their suppliers. Recall that a firm with technology θ pays an iceberg cost of $t(\theta, \tilde{\theta})$ when importing from a supplier with technology $\tilde{\theta}$. Aggregating across all firm-to-firm trades, we calculate the total cost that arises from incompatibility.

Column 1 of Table 5 reports the share of this cost in GDP. On average, frictions arising from incompatibility amounts to 4.2% of world GDP. Column 2 shows that approximately one-third of these costs occur on importing transactions, while the remaining occur on within-country transactions. Both the overall size of the cost and its composition between within-country and importing transactions differ across countries. For example, China, both distant in the technological space from its major trade partners and featuring large within-country heterogeneity among the technologies of different sectors, bears the highest cost on domestic as well as importing transactions. On the other hand, Western Europe and the U.S.

Table 6: The Technology Decoupling Effect of a Semi-Conductor Embargo

Embargo Origin	Share of Imports (%)	Δ Cite from US (%)	Endo. Tech. ($\Delta \ln U$ %)		Fixed Tech. ($\Delta \ln U$ %)	
			CHN	USA	CHN	USA
USA Only	2.1	-1.321	-0.016	-0.004	-0.010	-0.002
All Western-Aligned Countries	99.9	-50.516	-0.795	-0.081	-0.419	-0.016

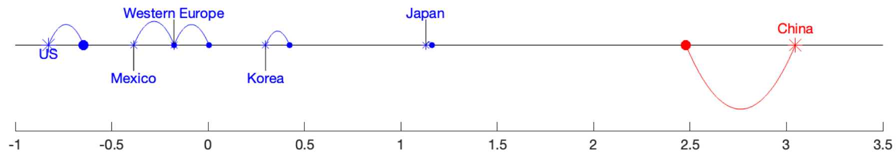
are much less affected by this channel.

4.4 Counterfactual: Trade Conflict and Technology Decoupling

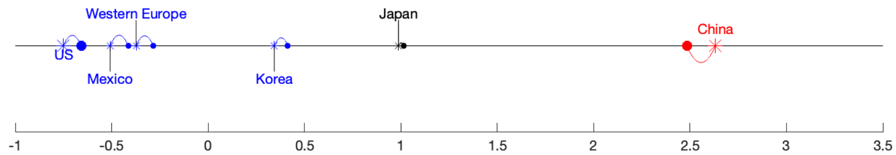
We use our model to evaluate the welfare costs of trade conflicts, with an emphasis on the role of endogenous technology choice. To this end, we consider a counterfactual that is intended to speak to the recent silicon blockade of the U.S. and its allies against China. In this counterfactual, we shut down China's imports of intermediate goods in the sector of Computer, Electronic, and Optical Products from the United States and other Western-aligned economies (which in our sample is all but Russia). In our data, this sector takes up roughly 17.6% in the total imports of China. Within the sector, 2.1% of the total imports are from the United States, and over 99.9% are from all Western-aligned countries.

To implement this counterfactual, we shut down exports of this sector to China by raising the corresponding trade costs to infinity. We investigate two cases, one in which China is embargoed by the United States only, and the other by all Western-aligned countries. Table 6 reports the main findings. When only the U.S. imposes the embargo, 2.1% of China's imports in this sector are directly affected. This leads to a small divergence of the technologies in the two countries. Technology proximity, measured using the model-implied U.S. share of the citations made by Chinese patients, decreases by 1.321%. This embargo inflicts more damage on China than on the U.S., but since the U.S. is not an important direct supplier of semiconductors to China, the damages in both countries are minor.

The second row of the table reports the results when all Western-aligned countries impose the embargo at the same time. These countries essentially account for all of Chinese imports in this sector, so the impacts are substantially larger. China now sustains a welfare loss of 0.8%, about 50 times when only the U.S. imposes the export ban. In addition to the larger loss for China, two other results are noteworthy. First, the distance between Chinese and U.S. technologies increases substantially more than in the first experiment despite the fact that in both cases the sanction originating from the U.S. is the same. This occurs because, among major economies, China and the U.S. occupy the two ends of the technology spectrum. As Chinese technologies shift away from the technologies of other countries due to export ban, they also shift away from the technologies of the U.S. Second, the welfare cost to the U.S. is amplified. This amplification occurs for two reasons: first, the production cost in China increases due to the direct effect of the embargo and the rest of the world now



(a) The Average Technology in the Targeted Sector



(b) The Average Technology of Other Sectors

Figure 3: Average Mean of Technology Distributions

Note: Dots are the ex-post mean in the baseline equilibrium, and stars are the equilibrium with the embargo. Blue indicates countries with distance to the USA relative to China decreases by more than 5%.

has to pay higher prices for Chinese products. Second, the endogenous divergence in the technologies around the globe makes souring intermediate inputs even more costly.

The role of endogenous technology. To shed light on the role of endogenous technology, Figure 3 plots the changes in countries' technology due to the embargo. The upper panel depicts the change in the average technology of the targeted industry. As discussed earlier, the embargo leads to a divergence of technologies between the U.S. and China. This divergence results in a re-alignment of the technology of other countries. We indicate using blue if a country's technological distance to the U.S. relative to its technological distance to China increases significantly (by more than 5%). We find that West Europe, Japan, Korea, and Mexico, all gravitate toward the U.S. Importantly, even though these countries' technology moves in the direction of the U.S. technology, the gap between their technology and the U.S. technology could increase because their technology shifts less than the U.S. technology. This widening technological difference contributes to the welfare costs of the embargo on the U.S. economy, as discussed earlier.

The lower panel plots the average technologies of other sectors that are not directly affected by the embargo. There, the changes are less apparent but qualitatively similar to the changes in the targeted sector. Because of the compatibility incentives, the divergence in technologies in semiconductors trickles down to all other sectors via input-output linkages.

To uncover the importance of endogenous technology for welfare, we consider the same embargo experiments under the restrictions that firms are stuck with the ex-post technology distribution ('fixed technology'). The last two columns of Table 6 report the results.

Table 7: Mechanism Decomposition

	$\Delta \ln U_{CHN}$ (%)	$\Delta \ln U_{USA}$ (%)
No Response of **Direction of Technology**	-0.419	-0.016
+ Response from the targeted Chinese Sector	-0.576	-0.030
+ Response from All Chinese Sectors	-0.692	-0.069
+ Response from All Countries	-0.795	-0.081

Compared with when firms can choose any technologies they wish, the welfare losses are substantially lower under the 'fixed technology' scenario. Even for individual firms, having a choice is clearly beneficial, for the world economy as a whole, the externality in technology compatibility between sectors and countries dominates, resulting in an amplification of the welfare losses.

We decompose the importance of various margins of technology adjustment in Table 7 by gradually allowing the technological responses from the mostly directly affected sector to the remaining sectors in China to other countries. Table 7 shows the results. Loosely speaking, the response in the targeted sector in China is more important than China's other sectors which, in turn, is more important than the response in all other countries.

Remarks. These counterfactual findings echo the mechanisms discussed in Propositions 6 and 7. However, in our general equilibrium model with multiple countries, endogenous divergence in technologies between any two does not have to be welfare decreasing. For example, a divergence in the technologies of Korea and Japan that arises from a trade conflict between these two countries *can* improve the welfare, if the positive externality of Korea being closer to the U.S. and Japan being closer to China exceeds the negative externality of these two countries being closer to each other. This possibility highlights the value of the general equilibrium framework and discipline from the data.

5 Conclusion

In this paper, we develop a model of production networks with endogenous horizontal technologies. In the model, firms choose both their technology and suppliers, two decisions that interact with each other because of the compatibility incentives we introduce. This interaction is further shaped by general equilibrium forces that play out in our multi-country, multi-sector trade model, which we show can lead to important externalities in firms' technology choices. We characterize sufficient conditions for equilibrium existence and uniqueness, and provide aggregation results for the model that make it tractable to take to the data.

Using patent data and trade data, we document novel firm- and country-level evidence that supports the model's key mechanisms. Using the calibrated model, we obtain three main findings. First, endogenous technology due to trade plays an important role in shaping global technologies, accounting for *two-thirds* of the variations in technological proximity between countries, whereas the differences in endowment distributions explain the rest. Second, costs that firms bear due to technology incompatibility account for approximately 3.4% of the World's GDP, suggesting the importance of this mechanism. Lastly, trade conflicts between the U.S. and China can lead to the decoupling of their technologies and the re-alignment of technologies by other countries. These technological changes substantially increase the welfare costs of the trade conflicts for both countries.

Our framework can be extended in a few directions. First, in reality, firms and countries use different institutions to manage the externalities highlighted in our model. For example, firms can integrate their supply chains to internalize the externality. Firms can collaborate with other firms to develop a standard to which all of them comply. Investigating the equilibrium impacts of such mechanisms is a promising avenue for future research. Second, our model is static and therefore abstracts from dynamics and vertical innovation. Integrating our model into a dynamic model requires additional work, but such an effort can illuminate how the compatibility incentive introduced in this model interacts with growth and how such interactions play out in the production network.

References

Acemoglu, Daron and Pablo D Azar, "Endogenous production networks," *Econometrica*, 2020, *88* (1), 33–82.

Aghion, Philippe, Antonin Bergeaud, Timothee Gigout, Matthieu Lequien, and Marc Melitz, "Exporting ideas: Knowledge flows from expanding trade in goods," in "Working Paper" 2021.

Amiti, Mary, Stephen J Redding, and David E Weinstein, "The Impact of the 2018 Tariffs on Prices and Welfare," *Journal of Economic Perspectives*, 2019, *33* (4), 187–210.

Ayerst, Stephen, Faisal Ibrahim, Gaelan MacKenzie, and Swapnika Rachapalli, "Trade and Diffusion of Embodied Technology: An Empirical Analysis," *Journal of Monetary Economics*, 2023.

Boehm, Johannes and Ezra Oberfield, "Misallocation in the Market for Inputs: Enforcement and the Organization of Production," *The Quarterly Journal of Economics*, 2020, *135* (4), 2007–2058.

Brandt, Loren, Johannes Van Biesebroeck, Luhang Wang, and Yifan Zhang, "WTO Accession and Performance of Chinese Manufacturing Firms," *American Economic Review*, 2017, *107* (9), 2784–2820.

Buera, Francisco J and Ezra Oberfield, "The Global Diffusion of Ideas," *Econometrica*, 2020, *88* (1), 83–114.

Cai, Jie, Nan Li, and Ana Maria Santacreu, "Knowledge Diffusion, Trade, and Innovation across Countries and Sectors," *American Economic Journal: Macroeconomics*, 2022, *14* (1), 104–145.

Caliendo, Lorenzo and Fernando Parro, "Estimates of the Trade and Welfare Effects of NAFTA," *The Review of Economic Studies*, 2015, *82* (1), 1–44.

Carluccio, Juan and Thibault Fally, "Foreign entry and spillovers with technological incompatibilities in the supply chain," *Journal of International Economics*, 2013, *90* (1), 123–135.

Chaney, Thomas, "Distorted Gravity: the Intensive and Extensive Margins of International Trade," *The American Economic Review*, 2008, *98* (4), 1707–1721.

_, "The Network Structure of International Trade," *American Economic Review*, 2014, *104* (11), 3600–3634.

Costinot, Arnaud, "A comparative institutional analysis of agreements on product standards," *Journal of International Economics*, 2008, *75* (1), 197–213.

_ **and Andres Rodriguez-Clare**, "Trade Theory with Numbers: Quantifying the Consequences of Globalization," *Handbook of International Economics*, 2014, *4*, 197.

_ **and Jonathan Vogel**, "Matching and Inequality in the World Economy," *Journal of Political Economy*, 2010, *118* (4).

_ **,** _ **et al.**, "Beyond Ricardo: Assignment Models in International Trade," *Annual Review of Economics*, 2015, *7* (1), 31–62.

Demir, Banu, Ana Cecilia Fieler, Daniel Yi Xu, and Kelly Kaili Yang, "O-ring Production Networks," 2021.

Dhyne, Emmanuel, Ayumu Ken Kikkawa, Xianglong Kong, Magne Mogstad, and Felix Tintelnot, "Endogenous production networks with fixed costs," *Journal of International Economics*, 2023, p. 103841.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography, and Trade," *Econometrica*, 2002, *70* (5), 1741–1779.

Fajgelbaum, Pablo D, Pinelopi K Goldberg, Patrick J Kennedy, and Amit K Khandelwal, "The Return to Protectionism," *The Quarterly Journal of Economics*, 2020, *135* (1), 1–55.

Fan, Haichao, Yao Amber Li, and Stephen R Yeaple, "Trade Liberalization, Quality, and Export Prices," *Review of Economics and Statistics*, 2015, *97* (5), 1033–1051.

Hotelling, Harold, "Stability in Competition," *The Economic Journal*, 1929, *39* (153), 41–57.

Huang, Yi, Chen Lin, Sibo Liu, and Heiwai Tang, "Trade Linkages and Firm Value: Evidence from the 2018 US-China" Trade War"," 2018.

Jaffe, A, "Technological Opportunity and Spillovers of R&D: Evidence from Firms, Patents, Profits and Market Value," *American Economic Review*, 1986, *76*, 984–1001.

Jaffe, Adam B, Manuel Trajtenberg, and Rebecca Henderson, "Geographic localization of knowledge spillovers as evidenced by patent citations," *the Quarterly journal of Economics*, 1993, *108* (3), 577–598.

Jones, Charles I, "Misallocation, Economic growth, and Input-Output Economics," *National* *bureau of economic research Working Paper*, 2011.

Keller, Wolfgang, "Knowledge Spillovers, Trade, and FDI," Technical Report, National Bureau of Economic Research 2021.

Kortum, Samuel S, "Research, Patenting, and Technological Change," *Econometrica: Journal of the Econometric Society*, 1997, pp. 1389–1419.

Li, Amber, Sichuang Xu, Stephen Yeaple, and Tengyu Zhao, "Bilateral Economies of Scope," *Miemo*, 2023.

Lim, Kevin, "Firm-to-firm Trade in Sticky Production Networks," 2018.

Lind, Nelson and Natalia Ramondo, "Trade with Correlation," *American Economic Review*, 2023, *113* (2), 317–353.

Liu, Ernest and Song Ma, "Innovation Networks and R&D Allocation," Technical Report, National Bureau of Economic Research 2021.

Lybbert, Travis J and Nikolas J Zolas, "Getting Patents and Economic Data to Speak to Each Other: An 'Algorithmic Links with Probabilities' Approach for Joint Analyses of Patenting and Economic Activity," *Research Policy*, 2014, *43* (3), 530–542.

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

Morales, Eduardo, Gloria Sheu, and Andrés Zahler, "Extended gravity," *The Review of economic studies*, 2019, *86* (6), 2668–2712.

Oberfield, Ezra, "A Theory of Input–Output Architecture," *Econometrica*, 2018, *86* (2), 559–589.

Online Appendix: Trade and Technology Compatibility in General Equilibrium

Weili Chen
Penn State

Jingting Fan
Penn State

Wenlan Luo
Tsinghua University

Appendix A Theory

A.1 Proof of Proposition 1 and Corollaries 1 and 2

This subsection provides the proof of Proposition 1 and its two corollaries. We start with introducing two lemmas.

Lemma A.1. *Suppose random variable X follows a Weibull (inverse Fréchet) distribution with c.d.f.*

$$F(x) = 1 - \exp[-(x/C)^{\lambda}],$$

where $C > 0$ and $\lambda > 1$ are parameters. Then, for any $\varepsilon > 0$,

$$\mathbb{E}[X^{\varepsilon}] = C^{\varepsilon} \cdot \Gamma(1 + \frac{\varepsilon}{\lambda}),$$

where $\Gamma(\cdot)$ denotes the Gamma function.

Proof.

$$\begin{aligned}
\mathbb{E}[X^{\varepsilon}] &= \int_0^{+\infty} x^{\varepsilon} \cdot \exp[-(x/C)^{\lambda}] \cdot C^{-\lambda} \lambda x^{\lambda-1} \mathrm{d}x \\
&= C^{\varepsilon} \cdot \int_0^{+\infty} \kappa^{\varepsilon/\lambda} \exp(-\kappa) \mathrm{d}\kappa \\
&= C^{\varepsilon} \cdot \Gamma(1 + \frac{\varepsilon}{\lambda}).
\end{aligned}$$

□

Lemma A.2. *Let $\{X_z\}_z$ be a collection of random variables indexed by $z \in (0, \infty)$ such that*

$$F(x, z) \equiv \Pr(X_z \leq x \mid z)$$

is jointly continuous in (x, z). Let $\mathcal{X}$ be a random set of uniformly drawn X_z such that for any $z_1 < z_2 \in (0, +\infty)$, $\mid \{X_z \in \mathcal{X} : z_1 < z \leq z_2\} \mid$ follows a Poisson distribution with mean $H(z_1) - H(z_2)$, where $\mid \cdot \mid$ denotes the number of elements in a set, and $H(z)$ is a decreasing function on $(0, +\infty)$. Then, $\mid \{X_z \in \mathcal{X} : X_z \leq x\} \mid$ follows a Poisson distribution with mean

$$\int_0^{\infty} F(x, z) \mathrm{d}(-H(z)),$$

provided that the integral exists.

Proof. Since X_z is uniformly drawn, $\mid \{X_z \in \mathcal{X} : X_z \leq x, z_1 < z \leq z_2\} \mid$ follows a Poisson distribution with mean $G(x, z_1, z_2)$ that satisfies

$$\inf_{\tilde{z} \in (z_1, z_2]} F(x, \tilde{z})[H(z_1) - H(z_2)] \leq G(x, z_1, z_2) \leq \sup_{\tilde{z} \in (z_1, z_2]} F(x, \tilde{z})[H(z_1) - H(z_2)].$$

Taking any monotonically increasing sequence $\{z_i\}_{i=1}^{\infty}$ with $z_1 = 0$ and $\lim_{i \to \infty} z_i = \infty$ we have

$$\sum_{i=1}^{\infty} \inf_{\tilde{z} \in (z_i, z_{i+1}]} F(x, \tilde{z})[H(z_i) - H(z_{i+1})] \leq \lim_{\tilde{z} \to \infty} G(x, 0, \tilde{z}) \leq \sum_{i=1}^{\infty} \sup_{\tilde{z} \in (z_i, z_{i+1}]} F(x, \tilde{z})[H(z_i) - H(z_{i+1})].$$

Therefore,

$$\lim_{\tilde{z}\to\infty} G(x,0,\tilde{z}) = \int_0^\infty F(x,z)\mathrm{d}(-H(z)),$$

provided that the integral exists (in the sense of Riemann integration). $\square$

Proof of Proposition 1. We prove by guess and verification. Suppose $p_d^i(\theta)$, the factory-gate price of a firm in (d,i) with technology location θ, follows a Weibull (inverse Fréchet) distribution with c.d.f. specified in (2). We verify that this is consistent with firms' behaviors in the equilibrium.

Proof. The first step to prove the proposition is to characterize the distribution of $c^j(\nu,r)$ for a firm ν in country-sector (d,i). Following Lemma A.2 and Assumption 1, given the targeted technology location $\theta(\nu)=\theta$, in any sourcing country o, the number of suppliers with effective marginal cost $\tilde{c}^j(\nu,\omega)$ less or equal to any level $c>0$ and supplier technology location $\theta(\omega)=\tilde{\theta}$ follows a Poisson distribution with mean

$$\begin{aligned}
&\int_0^\infty F_o^j[\frac{z\cdot c}{\tau_{do}^j\cdot t(\theta,\tilde{\theta})};\tilde{\theta}]\zeta z^{-\zeta-1}\mathrm{d}z\cdot \mathrm{d}\Theta_o^j(\tilde{\theta})\\
=&\int_0^\infty F_o^j(\kappa;\tilde{\theta})\zeta\kappa^{-\zeta-1}[\frac{c}{\tau_{do}^j\cdot t(\theta,\tilde{\theta})}]^\zeta \mathrm{d}\kappa\cdot \mathrm{d}\Theta_o^j(\tilde{\theta})\\
=&\int_0^\infty [t(\theta,\tilde{\theta})\kappa]^{-\zeta}\mathrm{d}F_o^j(\kappa;\tilde{\theta})\cdot(\frac{c}{\tau_{do}^j})^\zeta\cdot \mathrm{d}\Theta_o^j(\tilde{\theta})\\
=&\,\Gamma(1-\zeta/\lambda)\cdot[t(\theta,\tilde{\theta})C_o^j(\tilde{\theta})]^{-\zeta}\cdot(\frac{c}{\tau_{do}^j})^\zeta\cdot \mathrm{d}\Theta_o^j(\tilde{\theta}). \qquad \text{(A.1)}
\end{aligned}$$

Integrating over $\tilde{\theta}$, the number of suppliers with effective marginal cost $\tilde{c}^j(\nu,\omega)$ less or equal to any level $c>0$ follows a Poisson distribution with mean

$$\Gamma(1-\zeta/\lambda)\cdot(\frac{c}{\tau_{do}^j\cdot\Lambda_o^j(\theta)})^\zeta,$$

where $\Lambda_o^j(\cdot)$ is defined in (5),

$$\Lambda_o^j(\theta)\equiv\left(\int[C_o^j(\tilde{\theta})t(\theta,\tilde{\theta})]^{-\zeta}\mathrm{d}\Theta_o^j(\tilde{\theta})\right)^{-1/\zeta}.$$

The probability that no such supplier arrives is

$$\Pr\left[\min_{\omega\in\Omega_o^j(\nu,r)}\tilde{c}^j(\nu,\omega)>c\right]=\exp\left[-\Gamma(1-\zeta/\lambda)\cdot(\frac{c}{\tau_{do}^j\cdot\Lambda_o^j(\theta)})^\zeta\right].$$

Therefore, the distribution of $c^j(\nu,r)$ is characterized by

$$\begin{aligned}
\Pr[c^j(\nu,r)>c]&=\exp[-\Gamma(1-\zeta/\lambda)\cdot\sum_o(\frac{c}{\tau_{do}^j\cdot\Lambda_o^j(\theta)})^\zeta]\\
&=\exp[-\tilde{\Lambda}_d^j(\theta)c^\zeta], \qquad \text{(A.2)}
\end{aligned}$$

where $\tilde{\Lambda}_d^j(\theta)\equiv\Gamma(1-\zeta/\lambda)\sum_o(\tau_{do}^j\Lambda_o^j(\theta))^{-\zeta}$.

Next, we derive the distribution of the factory-gate price given by (2). Following Lemma A.2 and Assumption 1, for a firm ν in (d,i) with technology location $\theta(\nu)$, the number of techniques such that the factory-gate price is weakly less than p follows a Poisson distribution with mean

$$\int_0^\infty \int_0^\infty ... \int_0^\infty \mathbb{I}\left[\frac{1}{a}[w_d]^{\gamma^i}\prod_j [c^j]^{\gamma^{ij}} \le p\right]$$
$$\cdot \prod_j \zeta [c^j]^{\zeta-1}\tilde{\Lambda}_d^j(\theta)\exp[-\underbrace{\tilde{\Lambda}_d^j(\theta)[c^j]^\zeta}_{\equiv m^j}]\lambda[A_d^i]^\lambda a^{-\lambda-1}\mathrm{d}c^1...\mathrm{d}c^S\mathrm{d}a$$
$$= \int_0^\infty \int_0^\infty ... \int_0^\infty \mathbb{I}[\frac{1}{a}[w_d]^{\gamma^i}\prod_j(\frac{m^j}{\tilde{\Lambda}_d^j(\theta)})^{\frac{\gamma^{ij}}{\zeta}} \le p]\prod_j \exp(-m^j)\lambda[A_d^i]^\lambda a^{-\lambda-1}\mathrm{d}m^1...\mathrm{d}m^S\mathrm{d}a$$
$$= \int_0^\infty \int_0^\infty ... \int_0^\infty \mathbb{I}[\prod_j [m^j]^{\frac{\gamma^{ij}}{\zeta}} \le \underbrace{ap[w_d]^{-\gamma^i}\prod_j \tilde{\Lambda}_d^j(\theta)^{\frac{\gamma^{ij}}{\zeta}}}_{\equiv\kappa}]\prod_j \exp(-m^j)\lambda[A_d^i]^\lambda a^{-\lambda-1}\mathrm{d}m^1...\mathrm{d}m^S\mathrm{d}a$$
$$= \int_0^\infty \int_0^\infty ... \int_0^\infty \mathbb{I}[\prod_j [m^j]^{\frac{\gamma^{ij}}{\zeta}} \le \kappa]\prod_j \exp(-m^j)[A_d^i]^\lambda (p[w_d]^{-\gamma^i}\prod_j \tilde{\Lambda}_d^j(\theta)^{\frac{\gamma^{ij}}{\zeta}})^\lambda \lambda\kappa^{-\lambda-1}\mathrm{d}m^1...\mathrm{d}m^S\mathrm{d}\kappa$$
$$= \int_0^\infty \int_0^\infty ... \int_0^\infty \mathbb{I}[\prod_j [m^j]^{\frac{\gamma^{ij}}{\zeta}} \le \kappa]\prod_j [\Gamma(1-\zeta/\lambda)]^{\frac{\gamma^{ij}}{\zeta}}\exp(-m^j)\lambda\kappa^{-\lambda-1}\mathrm{d}m^1...\mathrm{d}m^S\mathrm{d}\kappa$$
$$\cdot [A_d^i]^\lambda (p[w_d]^{-\gamma^i}\prod_j[\sum_o(\tau_{do}^j\Lambda_o^j(\theta))^{-\zeta}]^{\frac{\gamma^{ij}}{\zeta}})^\lambda$$
$$= [\Xi^i]^{-\lambda}[A_d^i]^\lambda([w_d]^{-\gamma^i}\prod_j[\sum_o(\tau_{do}^j\Lambda_o^j(\theta))^{-\zeta}]^{\frac{\gamma^{ij}}{\zeta}})^\lambda p^\lambda,$$

where

$$\Xi^i \equiv (\int_0^\infty \int_0^\infty ... \int_0^\infty \mathbb{I}[\prod_j [m^j]^{\frac{\gamma^{ij}}{\zeta}} \le \kappa]\prod_j [\Gamma(1-\zeta/\lambda)]^{\frac{\gamma^{ij}}{\zeta}}\exp(-m^j)\lambda\kappa^{-\lambda-1}\mathrm{d}m^1...\mathrm{d}m^S\mathrm{d}\kappa)^{-1/\lambda} \quad \text{(A.3)}$$

is a sector-specific constant that depends on the technology parameters only.

This implies that $F_d^i(p;\theta)$ satisfies

$$1 - F_d^i(p;\theta) = \exp(-[p/C_d^i(\theta)]^\lambda),$$

where

$$C_d^i(\theta) = \frac{\Xi_i}{A_d^i}[w_d]^{\gamma^i}\prod_j(\sum_o[\tau_{do}^j\Lambda_o^j(\theta)]^{-\zeta})^{-\frac{\gamma^{ij}}{\zeta}}$$
$$= \frac{\Xi_i}{A_d^i}[w_d]^{\gamma^i}\prod_j(\sum_o\int[\tau_{do}^j C_o^j(\tilde{\theta})t(\theta,\tilde{\theta})]^{-\zeta}d\Theta_o^j(\tilde{\theta}))^{-\frac{\gamma^{ij}}{\zeta}}.$$

This finishes the proof of Proposition 1.

□

Proof of Corollary 1. As is shown in (A.1), for any firm in (d,i) with targeted technology location θ, in any sourcing country o, the number of suppliers from sector j with technology location $\tilde{\theta}$ and effective marginal cost less or equal to any level $c > 0$ follows a Poisson distribution with mean

$$\Gamma(1-\zeta/\lambda)\cdot[t(\theta,\tilde{\theta})C_o^j(\tilde{\theta})]^{-\zeta}\cdot(\frac{c}{\tau_{do}^j})^{\zeta}\mathrm{d}\Theta_o^j(\tilde{\theta}).$$

This implies that the effective cost of sourcing sector-j input from firms with technology location around $\tilde{\theta}$ from o, denoted $\tilde{c}_{do}^j(\theta,\tilde{\theta})$, is distributed with c.d.f.

$$\tilde{G}_{do}^j(c;\theta,\tilde{\theta}) = 1-\exp[-\Gamma(1-\zeta/\lambda)\cdot[\tau_{do}^j t(\theta,\tilde{\theta})C_o^j(\tilde{\theta})]^{-\zeta}\cdot c^{\zeta}\mathrm{d}\Theta_o^j(\tilde{\theta})].$$

Therefore, the probability of sourcing sector-j inputs from firms in country o with technology location around $\tilde{\theta}$ is

$$\begin{aligned}
&\Pr[(o,\tilde{\theta}) = \arg\min_{o',\tilde{\theta}'}\tilde{c}_{do'}^j(\theta,\tilde{\theta}')] \\
=&\int_0^{+\infty}\Pr[\tilde{c}_{do}^j(\theta,\tilde{\theta}) = c\cap\tilde{c}_{do'}^j(\theta,\tilde{\theta}) > c,\forall(o',\tilde{\theta}')\neq(o,\tilde{\theta})]\mathrm{d}c \\
=&\int_0^{+\infty}\prod_{(o',\tilde{\theta}')\neq(o,\tilde{\theta})}[1-\tilde{G}_{do'}^j(c;\theta,\tilde{\theta}')]\mathrm{d}\tilde{G}_{do}^j(c;\theta,\tilde{\theta}) \\
=&\int_0^{+\infty}\exp[-\Gamma(1-\zeta/\lambda)\cdot\sum_o\int[\tau_{do}^j t(\theta,\tilde{\theta})C_o^j(\tilde{\theta})]^{-\zeta}\mathrm{d}\Theta_o^j(\tilde{\theta})\cdot c^{\zeta}] \\
&\qquad\cdot\Gamma(1-\zeta/\lambda)\cdot[\tau_{do}^j t(\theta,\tilde{\theta})C_o^j(\tilde{\theta})]^{-\zeta}\mathrm{d}\Theta_o^j(\tilde{\theta})\cdot\zeta c^{\zeta-1}\mathrm{d}c \\
=&\{\int_0^{+\infty}\exp[-\Gamma(1-\zeta/\lambda)\cdot\sum_o[\tau_{do}^j\Lambda_o^j(\theta)]^{-\zeta}\cdot c^{\zeta}]\Gamma(1-\zeta/\lambda)\cdot\sum_o[\tau_{do}^j\Lambda_o^j(\theta)]^{-\zeta}\cdot\zeta c^{\zeta-1}\mathrm{d}c\} \\
&\qquad\cdot\frac{[\tau_{do}^j t(\theta,\tilde{\theta})C_o^j(\tilde{\theta})]^{-\zeta}\mathrm{d}\Theta_o^j(\tilde{\theta})}{\sum_o[\tau_{do}^j\Lambda_o^j(\theta)]^{-\zeta}} \\
=&\frac{[\tau_{do}^j t(\theta,\tilde{\theta})C_o^j(\tilde{\theta})]^{-\zeta}}{\sum_o[\tau_{do}^j\Lambda_o^j(\theta)]^{-\zeta}}\mathrm{d}\Theta_o^j(\tilde{\theta}) \\
\equiv&\chi_{do}^j(\theta,\tilde{\theta})\mathrm{d}\Theta_o^j(\tilde{\theta}).
\end{aligned}$$

Integrating over $\tilde{\theta}$, the probability of sourcing from firms in country o is

$$\chi_{do}^j(\theta) = \int\chi_{do}^j(\theta,\tilde{\theta})d\Theta_o^j(\tilde{\theta}) = \frac{[\tau_{do}^j\Lambda_o^j(\theta)]^{-\zeta}}{\sum_{o'}[\tau_{do'}^j\Lambda_{o'}^j(\theta)]^{-\zeta}}.$$

Proof of Corollary 2. Since firms engage in monopolistic competition when selling to final-good producers, they charge a monopolistic markup $\eta/(\eta-1)$. Final-good producers maximize their profits

$$P_dQ_d-\sum_j\sum_o\int_0^1[\frac{\eta}{\eta-1}\tau_{do}^{Uj}p_{do}^j(\omega)]q_{do}^j(\omega)\mathrm{d}\omega$$

subject to (1), where the factory-gate price $p_{do}^j(\omega)$ follows the distribution characterized by $F_d^i(p;\theta) = 1-\exp(-[p/C_d^i(\theta)]^{\lambda})$, and the ideal price index for the final good in d,

$$P_d = \prod_j(\frac{P_d^j}{\rho_d^j})^{\rho_d^j},\quad\text{with } P_d^j\equiv(\sum_o\int_0^1[\frac{\eta}{\eta-1}\tau_{do}^{Uj}p_{do}^j(\omega)]^{1-\eta}\mathrm{d}\omega)^{\frac{1}{1-\eta}}$$

Optimization of the final-good production implies that in country d, the market share of any good ω in sector j over all goods in the sector is given by

$$\frac{\frac{\eta}{\eta-1}\tau_{do}^{Uj} p_{do}^{j}(\omega) x_{do}^{j}(\omega)}{\rho_d^j P_d Q_d} = \frac{[\frac{\eta}{\eta-1}\tau_{do}^{Uj} p_{do}^{j}(\omega)]^{1-\eta}}{(P_d^j)^{1-\eta}},$$

where the price index of sector-j goods in country d, denoted as P_d^j, satisfies

$$\begin{aligned}
(P_d^j)^{1-\eta} &= \sum_o \int_0^1 [\frac{\eta}{\eta-1}\tau_{do}^{Uj} p_{do}^{j}(\omega)]^{1-\eta} \mathrm{d}\omega \\
&= \sum_o \int [\frac{\eta}{\eta-1}\tau_{do}^{Uj} p_{do}^{j}(\tilde{\theta})]^{1-\eta} \mathrm{d}\Theta_o^j(\tilde{\theta}) \\
&= \sum_o [\frac{\eta}{\eta-1}\tau_{do}^{Uj}]^{1-\eta} \int \mathbb{E}[p_{do}^{j}(\tilde{\theta})]^{1-\eta} \mathrm{d}\Theta_o^j(\tilde{\theta}) \\
&= \sum_o \Gamma(1+\frac{1-\eta}{\lambda}) \cdot [\frac{\eta}{\eta-1}\tau_{do}^{Uj}]^{1-\eta} \cdot (\bar{\Lambda}_o^j)^{1-\eta}, \qquad \text{(A.4)}
\end{aligned}$$

where $\bar{\Lambda}_o^j$ is defined in (9),

$$\bar{\Lambda}_o^j \equiv (\int [C_o^j(\tilde{\theta})]^{1-\eta} \mathrm{d}\Theta_o^j(\tilde{\theta}))^{1/(1-\eta)}.$$

Meanwhile, by (2) and Lemma A.1, the expected sales of goods produced by firms in country o with technology location around $\tilde{\theta}$ is

$$\Gamma(1+\frac{1-\eta}{\lambda}) \cdot [\frac{\eta}{\eta-1}\tau_{do}^{Uj}]^{1-\eta} \cdot C_o^j(\tilde{\theta})^{1-\eta} \cdot \mathrm{d}\Theta_o^j(\tilde{\theta}).$$

Therefore, the expected expenditure share allocated to goods produced by firms in country o with technology location around $\tilde{\theta}$ is

$$\pi_{do}^j(\tilde{\theta})\mathrm{d}\Theta_o^j(\tilde{\theta}) \equiv \frac{\mathbb{E}[\frac{\eta}{\eta-1}\tau_{do}^{Uj} p_o^j(\tilde{\theta})]^{1-\eta}}{(P_d^j)^{1-\eta}} \mathrm{d}\Theta_o^j(\tilde{\theta}) = \frac{[\tau_{do}^{Uj} C_o^j(\tilde{\theta})]^{1-\eta}}{\sum_{o'}[\tau_{do'}^{Uj}\bar{\Lambda}_{o'}^j]^{1-\eta}} \mathrm{d}\Theta_o^j(\tilde{\theta}).$$

Integrating over $\tilde{\theta}$, the total expenditure share allocated to goods produced by firms in country o is

$$\pi_{do}^j = \int \pi_{do}^j(\tilde{\theta})\mathrm{d}\Theta_o^j(\tilde{\theta}) = \frac{[\tau_{do}^{Uj}\bar{\Lambda}_o^j]^{1-\eta}}{\sum_{o'}[\tau_{do'}^{Uj}\bar{\Lambda}_{o'}^j]^{1-\eta}}.$$

A.2 Proof of Proposition 2

We restate Proposition 2 below:

Proposition A.1. *Suppose wages $\{w_d\}$ are given.*

1. *Assume $\{\bar{\Theta}_d^i\}$ have bounded support that is contained in $[-M, M]$ for some $M > 0$ and have associated density functions $\{\bar{\varsigma}_d^i\}$. If $\zeta\bar{t} < 1/M^2$, then there exists an equilibrium with firms' technology choice $\{g_d^i\}$ being continuously differentiable functions. Moreover, in this equilibrium, the choice of firms from (d, i) with endowment technology $\bar{\theta}$ is characterized by the*

following first-order condition with a unique solution.

$$g_d^i(\bar{\theta}) = \omega^i \sum_{j,o} \frac{\gamma^{ij}}{1-\gamma^{iL}} \int \chi_{do}^j(g_d^i(\bar{\theta}), \tilde{\theta}) \mathrm{d}\Theta_o^j(\tilde{\theta}) + (1-\omega^i)\bar{\theta}, \quad \forall \bar{\theta} \in [-M, M] \tag{A.5}$$

where $\omega^i \equiv \frac{(\eta-1)(1-\gamma^{iL})\bar{t}}{(\eta-1)(1-\gamma^{iL})\bar{t}+\bar{\phi}} < 1$.

2. *If, in addition,* $\bar{t} < \frac{1}{2M}$ *and* $\bar{\phi} > \underline{\phi}$*, where* $\underline{\phi} > 0$ *is a constant determined by parameters* $(\zeta, \bar{t}, \eta, M, \gamma^{iL})$ *as detailed in the proof, then such an equilibrium is unique.*

For convenience in taking derivatives, throughout the proof, we redefine $\bar{t}$ as twice the $\bar{t}$ in the main text, and $\bar{\phi}$ as twice the actual $\bar{\phi}$.

Definition A.1. *Given wages* w_d *and parameters* $M, \eta, \gamma^{iL}, \zeta, \bar{t}, \Xi^i, A_d^i, \tau_{do}^j$ *of the model, define constants*

- $\underline{\gamma}^L \equiv \min_i \gamma^{iL}$.
- $\omega^i \equiv \frac{(\eta-1)(1-\gamma^{iL})\bar{t}}{(\eta-1)(1-\gamma^{iL})\bar{t}+\bar{\phi}}$, $\overline{\omega} \equiv \max_i \omega^i$.
- $\xi_d^i \equiv \Xi^i (w_d)^{\gamma^{iL}} / A_d^i$.
- $\overline{M}' \equiv \max_i\{1 - \omega^i(1-\zeta\bar{t}M^2)\}$, $\underline{M}' \equiv 1 - \overline{\omega}$.
- $\overline{M}'' \equiv \frac{3\overline{\omega}\zeta\bar{t}M^3}{1-\overline{\omega}\zeta\bar{t}M^2}$.
- $M^{\mathcal{C}} \equiv \max_{d,i,j} \left| \frac{1}{\gamma^{iL}} \left[\ln \xi_d^i + (1-\gamma^{iL}) \ln\{\sum_o [\tau_{do}^j]^{-\zeta}\}^{-\frac{1}{\zeta}} \right] \right|$.

<u>Outline of Proofs</u>

- **For existence**, we formulate a fixed point problem for the policy function $\boldsymbol{g}$. We work on the space of function defined over $[-M, M]$ with uniformly bounded value and Lipschitz continuous first derivative. Define $\mathcal{G} = \{\boldsymbol{g} : [-M, M] \to \mathbb{R}^{N\times S}, \boldsymbol{g}$ is differentiable, $\|\boldsymbol{g}\|_\infty \leq M$; $[g_d^i]'(\bar{\theta}) \in [\underline{M}', \overline{M}']$; $\boldsymbol{g}'$ is Lipschitz continuous with a Lipschitz constant $\overline{M}''$ $\}$. Equip $\mathcal{G}$ with the C^1 norm: $\|\boldsymbol{g}\|_{\mathcal{G}} = \|\boldsymbol{g}\|_\infty + \|\boldsymbol{g}'\|_\infty$. The reason for having to work with the C^1 norm is that $\boldsymbol{g}'$ enters the operator defined below.
- Define operator $\mathcal{T}$ on $\mathcal{G}$ as below. The fixed point of $\mathcal{T}$, if exists, solves the first order condition of the technology adaptation problem.

$$[\mathcal{T}\boldsymbol{g}]_d^i(\bar{\theta}) = \omega^i \sum_{j,o} \frac{\gamma^{ij}}{1-\gamma^{iL}} \int \chi_{do}^{ij}(\bar{\theta}, \tilde{\theta}; \boldsymbol{g}, \boldsymbol{g}') g_o^j(\tilde{\theta}) \mathrm{d}\tilde{\theta} + (1-\omega^i)\bar{\theta}, \tag{A.6}$$

with $\omega^i = \frac{(\eta-1)(1-\gamma^{iL})\bar{t}}{(\eta-1)(1-\gamma^{iL})\bar{t}+\bar{\phi}} < 1$, where, to slightly abuse notations and make them dependent on $\boldsymbol{g}$ and $\boldsymbol{g}'$ explicitly, $\boldsymbol{\chi}$ and $\boldsymbol{\mathcal{C}}$ satisfy

$$\mathcal{C}^i_d(\bar{\theta};\boldsymbol{g},\boldsymbol{g}') \equiv \ln \xi^i_d - \sum_j \frac{\gamma^{ij}}{\zeta} \ln \left(\sum_o \int [\tau^j_{do}]^{-\zeta} \exp\left(-\zeta \mathcal{C}^j_o(\tilde{\theta};\boldsymbol{g},\boldsymbol{g}')\right) \exp\left(-\frac{1}{2}\zeta \bar{t}(g^i_d(\bar{\theta}) - g^j_o(\tilde{\theta}))^2\right) [g^j_o]'(\tilde{\theta}) \bar{\varsigma}^j_o(\tilde{\theta}) \mathrm{d}\tilde{\theta} \right), \tag{A.7}$$

$$\chi^{ij}_{do}(\bar{\theta},\tilde{\theta};\boldsymbol{g},\boldsymbol{g}') \equiv \frac{[\tau^j_{do}]^{-\zeta} \exp\left(-\zeta \mathcal{C}^j_o(\tilde{\theta};\boldsymbol{g},\boldsymbol{g}')\right) \exp\left(-\frac{1}{2}\zeta \bar{t}(g^i_d(\bar{\theta}) - g^j_o(\tilde{\theta}))^2\right) [g^j_o]'(\tilde{\theta}) \bar{\varsigma}^j_o(\tilde{\theta})}{\sum_m \int [\tau^j_{dm}]^{-\zeta} \exp\left(-\zeta \mathcal{C}^j_m(\tilde{\tilde{\theta}};\boldsymbol{g},\boldsymbol{g}')\right) \exp\left(-\frac{1}{2}\zeta \bar{t}(g^i_d(\bar{\theta}) - g^j_m(\tilde{\tilde{\theta}}))^2\right) [g^j_m]'(\tilde{\tilde{\theta}}) \bar{\varsigma}^j_o(\tilde{\tilde{\theta}}) \mathrm{d}\tilde{\tilde{\theta}}}. \tag{A.8}$$

- Lemma A.3 establishes the existence and uniqueness of $\boldsymbol{\mathcal{C}}$ given $\boldsymbol{g},\boldsymbol{g}'$ by formulating (A.7) as a fixed point problem for $\boldsymbol{\mathcal{C}}$. It also establishes that $\boldsymbol{\mathcal{C}}$ is continuous in $\boldsymbol{g},\boldsymbol{g}'$ and is differentiable in $\bar{\theta}$ with bounded derivative for any given $\boldsymbol{g},\boldsymbol{g}'$. It follows that $\boldsymbol{\chi}$ is also continuous in $\boldsymbol{g},\boldsymbol{g}'$, and that $\mathcal{T}$ is continuous in $\boldsymbol{g}$ with the norm $\|\cdot\|_{\mathcal{G}}$.
- Lemma A.4 further characterizes the uniform bounds for $\mathcal{T}\boldsymbol{g}$, $[\mathcal{T}\boldsymbol{g}]'$ and the Lipschitz continuity of $[\mathcal{T}\boldsymbol{g}]'$ which are used to show $\mathcal{T}\boldsymbol{g} \in \mathcal{G}$.
- Since $\boldsymbol{g} \in \mathcal{G}$ which is closed, and has uniformly bounded values and Lipschitz continuous first derivatives, by the Arzelà-Ascoli theorem, $\mathcal{G}$ is compact under the norm $\|\cdot\|_{\mathcal{G}}$. That $\mathcal{T}\boldsymbol{g} \in \mathcal{G}$ and $\mathcal{T}$ is continuous under the norm $\|\cdot\|_{\mathcal{G}}$ thus ensures the existence of a fixed point of $\mathcal{T}$ in $\mathcal{G}$ by the Schauder fixed-point theorem. These arguments are formalized in Proposition A.2 and its proof.
- **For uniqueness**, we treat (A.6) and (A.7) as a joined fixed point problem for $(\boldsymbol{g},\boldsymbol{\mathcal{C}})$. We show that with $\bar{t}$ small enough and $\bar{\phi}$ large enough as stated in the Proposition, the joint operator defined by the left hand side of the equations is a contraction mapping under the C^1 norm of the space of $\boldsymbol{g}$ combined with the C^0 norm of the space of $\boldsymbol{\mathcal{C}}$. This is done by showing the induced matrix norm (maximum absolute row sum norm) of the Jacobian matrix that contains the Frechet derivatives of the operator with respect to $(\boldsymbol{g},\boldsymbol{\mathcal{C}})$ is uniformly bounded by a number below one, when $\bar{t}$ is small and $\bar{\phi}$ is large enough. The estimates of the Jacobian entries are presented in Lemma A.5 and the uniqueness proof is formally stated in Proposition 2.

Intuitions for the existence/uniqueness results. For (A.6) restated below

$$[\mathcal{T}\boldsymbol{g}]^i_d(\bar{\theta}) = \omega^i \sum_{j,o} \frac{\gamma^{ij}}{1-\gamma^{iL}} \int \chi^{ij}_{do}(\bar{\theta},\tilde{\theta};\boldsymbol{g},\boldsymbol{g}') g^j_o(\tilde{\theta}) \mathrm{d}\tilde{\theta} + (1-\omega^i)\bar{\theta},$$

we see that if χ^{ij}_{do} does not vary across $\bar{\theta}$ or change by $\boldsymbol{g}$, then we have already established the unique existence of $\boldsymbol{g}$ by the contraction mapping. (To see this, denote $\bar{\chi}^{ij}_{do}(\tilde{\theta}) = \chi^{ij}_{do}(\bar{\theta},\tilde{\theta};\boldsymbol{g})$ under the premise, then the equation above is reduced to $[\mathcal{T}\boldsymbol{g}]^i_d(\bar{\theta}) = \omega^i \sum_{j,o} \frac{\gamma^{ij}}{1-\gamma^{iL}} \int \bar{\chi}^{ij}_{do}(\tilde{\theta}) g^j_o(\tilde{\theta}) \mathrm{d}\tilde{\theta} + (1-\omega^i)\bar{\theta}$, with $\mathcal{T}$ trivially satisfying Blackwell's sufficiency conditions for contraction with a modulus $\max_i \omega^i$.)

Part 1 of the proposition says that if $\zeta\bar{t}$ is not too large relative to the variation in $\bar{\theta}$, then the derivative of $\chi^{ij}_{do}(\bar{\theta},\tilde{\theta};\boldsymbol{g},\boldsymbol{g}')$ in $\bar{\theta}$ is uniformly bounded. The first and second derivative of $\boldsymbol{g}$ thus exist and are uniformly bounded. The fact that the first derivative of $\boldsymbol{g}$ is bounded also implies that the value of $\boldsymbol{g}$ is bounded on the bounded domain $[-M, M]$. This is sufficient to establish compactness of the space of $\boldsymbol{g}$ under and ensures existence.

Part 2 of the proposition says that if further $\bar{\phi}$ is large enough, then the variation of $\chi^{ij}_{do}(\bar{\theta},\tilde{\theta};\boldsymbol{g},\boldsymbol{g}')$ in $(\boldsymbol{g},\boldsymbol{g}')$ is bounded uniformly to ensure that $\mathcal{T}$ is a contraction.

The proofs are technically complicated by the fact that $\boldsymbol{g}'$ enters the mapping $\mathcal{T}$, so we have to work with the C^1 norm, and look for compactness or contraction property under the C^1 norm. The intuition behind the proof can be gained from the uniqueness proof for the degenerate prior case, stated in Proposition A.2.

Lemma A.3. *For $\boldsymbol{g} \in \mathcal{G}$, there uniquely exists a $\boldsymbol{C}(\bar{\theta};\boldsymbol{g},\boldsymbol{g}')$ defined by (A.7) that is bounded and continuous in $(\bar{\theta},\boldsymbol{g},\boldsymbol{g}')$. Further, $\boldsymbol{C}$ satisfies*

(1) $\|\boldsymbol{C}\|_\infty \leq M^{\mathcal{C}}$, in which $M^{\mathcal{C}}$ is defined in Definition A.1.

(2) $\forall(\boldsymbol{g},\boldsymbol{g}')$, $\boldsymbol{C}(\bar{\theta};\boldsymbol{g},\boldsymbol{g}')$ is differentiable in $\bar{\theta}$ and $\|\boldsymbol{C}'\|_\infty \leq (1-\underline{\gamma}^L)2tM$.

Proof. Denote $\widetilde{\mathcal{G}} = \{g \in C^0([-M,M] \to \mathbb{R}^{N\times S}) : \|g\|_\infty \leq M\}$, $\widetilde{\mathcal{G}}' = \{g' \in C^0([-M,M] \to \mathbb{R}^{N\times S}) : \underline{M}' \leq g' \leq \overline{M}'\}$, and $\mathcal{M} = [-M,M]$, where $\underline{M}'$ and $\overline{M}'$ are the constants defined in Definition A.1.

Define $\mathbb{C} = \{\boldsymbol{C} : \mathcal{M} \times \mathcal{G} \times \widetilde{\mathcal{G}} \to \mathbb{R}^{N\times S}, \|\boldsymbol{C}\|_\infty \leq M^{\mathcal{C}}, \boldsymbol{C} \text{ is continuous}\}$, for $M^{\mathcal{C}}$ defined in the proposition. It can be shown that $\mathbb{C}$ is complete under the infinity norm.

We now prove operator $\mathcal{T}^{\mathcal{C}}$ mapping from $\mathbb{C}$ defined below has an image contained in $\mathbb{C}$:

$$[\mathcal{T}^{\mathcal{C}}\boldsymbol{C}]^i_d(\bar{\theta};\boldsymbol{g},\boldsymbol{g}') = \ln \xi^i_d - \sum_j \frac{\gamma^{ij}}{\zeta} \ln\Big(\sum_o \int [\tau^j_{do}]^{-\zeta} \exp\left(-\zeta C^j_o(\tilde{\theta};\boldsymbol{g},\boldsymbol{g}')\right) \exp\left(-\frac{1}{2}\zeta\bar{t}(g^i_d(\bar{\theta}) - g^j_o(\tilde{\theta}))^2\right) [g^j_o]'(\tilde{\theta})\bar{\xi}^j_o(\tilde{\theta})\mathrm{d}\tilde{\theta}\Big).$$

Since $\boldsymbol{C} \in \mathbb{C}$ is continuous, $\mathcal{T}^{\mathcal{C}}\boldsymbol{C}$ is also continuous. Since $\exp\left(-\zeta C^j_o(\tilde{\theta};\boldsymbol{g},\boldsymbol{g}')\right) \in [\exp(-\zeta M^{\mathcal{C}}), \exp(\zeta M^{\mathcal{C}})]$, $\|\boldsymbol{g}'\|_\infty < \overline{M}' < 1$, we have

$$\|[\mathcal{T}^{\mathcal{C}}\boldsymbol{C}]^i_d\|_\infty \leq \ln \xi^i_d + (1-\gamma^{iL})\ln\{\sum_o [\tau^j_{do}]^{-\zeta}\}^{-\frac{1}{\zeta}} + (1-\gamma^{iL})M^{\mathcal{C}}$$

$$\Rightarrow \|\mathcal{T}^{\mathcal{C}}\boldsymbol{C}\|_\infty \leq \max_{d,i,j}\left|\left[\ln \xi^i_d + (1-\gamma^{iL})\ln\{\sum_o [\tau^j_{do}]^{-\zeta}\}^{-\frac{1}{\zeta}}\right]\right| + (1-\gamma^{iL})M^{\mathcal{C}} \leq M^{\mathcal{C}},$$

in which the last inequality applies the definition of $M^{\mathcal{C}}$.

We now verify $\mathcal{T}^{\mathcal{C}}$ satisfies Blackwell's sufficiency conditions for contraction. For any $\boldsymbol{C},\widehat{\boldsymbol{C}} \in \mathbb{C}$, such that $\boldsymbol{C} \leq \widehat{\boldsymbol{C}}$ point-wisely, it trivially holds that $[\mathcal{T}^{\mathcal{C}}\boldsymbol{C}]^i_d \leq [\mathcal{T}^{\mathcal{C}}\widehat{\boldsymbol{C}}]^i_d$. And it holds that $\mathcal{T}^{\mathcal{C}}[\boldsymbol{C}(\bar{\theta};\boldsymbol{g},\boldsymbol{g}') + c]^i_d = [\mathcal{T}^{\mathcal{C}}\boldsymbol{C}]^i_d(\bar{\theta};\boldsymbol{g},\boldsymbol{g}') + (1-\gamma^{iL})c \leq [\mathcal{T}^{\mathcal{C}}\boldsymbol{C}]^i_d(\bar{\theta};\boldsymbol{g},\boldsymbol{g}') + (1-\underline{\gamma}^L)c$ for $c \geq 0$. $\mathcal{T}^{\mathcal{C}}$ is thus a contraction mapping with a modulus $1-\underline{\gamma}^L$. $\mathcal{T}^{\mathcal{C}}$ thus has a unique fixed point that is contained in $\mathbb{C}$ by the contraction mapping theorem.

Next, consider

$$[\mathcal{T}^{\mathcal{C}}\boldsymbol{C}]^{i\,\prime}_d(\bar{\theta};\boldsymbol{g},\boldsymbol{g}') = \sum_j \gamma^{ij}\sum_o \int \chi^{ij}_{do}(\bar{\theta},\tilde{\theta};\boldsymbol{g},\boldsymbol{g}')\left[\bar{t}(g^i_d(\bar{\theta}) - g^j_o(\tilde{\theta}))\right]\mathrm{d}\tilde{\theta}$$

$$\Rightarrow \left|[\mathcal{T}^{\mathcal{C}}\boldsymbol{C}]^{i\,\prime}_d(\bar{\theta};\boldsymbol{g},\boldsymbol{g}')\right| \leq (1-\underline{\gamma}^L)2tM,$$

where the last line applies that $\sum_o \int \chi^{ij}_{do}(\bar{\theta},\tilde{\theta};\boldsymbol{g},\boldsymbol{g}')\mathrm{d}\tilde{\theta} = 1$ and $|g^i_d(\bar{\theta})| \leq M$. Note that the derivation does not rely on that the starting point $\boldsymbol{C}$ is differentiable in $\bar{\theta}$. Thus, $\boldsymbol{C}$ is differentiable in $\bar{\theta}$ and $\|\boldsymbol{C}'\|_\infty \leq (1-\underline{\gamma}^L)2tM$.

□

Lemma A.4. *For $\boldsymbol{g} \in \mathcal{G}$ and $\mathcal{T}\boldsymbol{g}$ defined by (A.6), restated below*

$$[\mathcal{T}\boldsymbol{g}]_d^i(\bar{\theta}) = \omega^i \sum_{j,o} \frac{\gamma^{ij}}{1-\gamma^{iL}} \int \chi_{do}^{ij}(\bar{\theta}, \tilde{\theta}; \boldsymbol{g}, \boldsymbol{g}') g_o^j(\tilde{\theta}) \mathrm{d}\tilde{\theta} + (1-\omega^i)\bar{\theta}, \tag{A.6}$$

(1) $\mathcal{T}\boldsymbol{g}$ is continuous in $\boldsymbol{g}$ under the C^1 norm, $\|\boldsymbol{g}\|_\infty + \|\boldsymbol{g}'\|_\infty$.

(2) $\|\mathcal{T}\boldsymbol{g}\|_\infty \leq M$.

(3) $\forall \boldsymbol{g} \in \mathcal{G}$, $\mathcal{T}\boldsymbol{g}(\bar{\theta})$ is twice differentiable in $\bar{\theta}$ with $[\mathcal{T}\boldsymbol{g}]_d^{i\,\prime}(\bar{\theta}) \in [\underline{M}', \overline{M}']$, for $\underline{M}'$ and $\overline{M}'$ the constants defined in Definition A.1; $[\mathcal{T}\boldsymbol{g}]_d^{i\,\prime}(\bar{\theta})$ is Lipschitz continuous with a Lipschitz constant $\omega^i \zeta \bar{t} M^2 \overline{M}'' + 3\omega^i \zeta \bar{t} M^3 \leq \overline{M}''$.

Proof. From Lemma A.3, $\boldsymbol{\mathcal{C}}$ is continuous in $\boldsymbol{g}, \boldsymbol{g}'$ under the infinity norm and by the definition of $\boldsymbol{\chi}$ in (A.8), $\boldsymbol{\chi}$ is continuous in $(\boldsymbol{g}, \boldsymbol{g}')$ and so is $\mathcal{T}\boldsymbol{g}$. $\mathcal{T}\boldsymbol{g}$ is thus continuous in $\boldsymbol{g}$ under the C^1 norm.

Next, consider

$$\begin{aligned}
[\mathcal{T}\boldsymbol{g}]_d^{i\,\prime}(\bar{\theta}) &= \omega^i \sum_j \frac{\gamma^{ij}}{1-\gamma^{iL}} \sum_o \int \partial_{\bar{\theta}} \chi_{do}^{ij}(\bar{\theta}, \tilde{\theta}) g_o^j(\tilde{\theta}) \mathrm{d}\tilde{\theta} + (1-\omega^i) \\
&= \omega^i \sum_j \frac{\gamma^{ij}}{1-\gamma^{iL}} \zeta \bar{t} [g_d^i]'(\bar{\theta}) cov_d^{ij}[g_o^j(\tilde{\theta}), g_o^j(\tilde{\theta})] + (1-\omega^i),
\end{aligned}$$

in which the second line applies part (1) of Lemma A.6. Since $0 < [g_d^i]'(\bar{\theta}) < 1$ and by Popoviciu's inequality on variances $cov_d^{ij}[g_o^j(\tilde{\theta}), g_o^j(\tilde{\theta})] \leq M^2$ we have that

$$1-\omega^i \leq [\mathcal{T}\boldsymbol{g}]_d^{i\,\prime}(\bar{\theta}; \boldsymbol{g}, \boldsymbol{g}') \leq 1-\omega^i(1-\zeta\bar{t}M^2).$$

Further, observe that $\forall \bar{\theta}, \hat{\theta} \in [-M, M]$,

$$\begin{aligned}
&\left| [\mathcal{T}\boldsymbol{g}]_d^{i\,\prime}(\bar{\theta}) - [\mathcal{T}\boldsymbol{g}]_d^{i\,\prime}(\hat{\theta}) \right| = \left| \omega^i \sum_j \frac{\gamma^{ij}}{1-\gamma^{iL}} \zeta\bar{t} \left\{ [g_d^i]'(\bar{\theta}) cov_{d,\bar{\theta}}^{ij}[g_o^j(\tilde{\theta}), g_o^j(\tilde{\theta})] - [g_d^i]'(\hat{\theta}) cov_{d,\hat{\theta}}^{ij}[g_o^j(\tilde{\theta}), g_o^j(\tilde{\theta})] \right\} \right| \\
&= \left| \omega^i \sum_j \frac{\gamma^{ij}}{1-\gamma^{iL}} \zeta\bar{t} \left\{ ([g_d^i]'(\bar{\theta}) - [g_d^i]'(\hat{\theta})) cov_{d,\bar{\theta}}^{ij}[g_o^j(\tilde{\theta}), g_o^j(\tilde{\theta})] + [g_d^i]'(\hat{\theta}) \left[cov_{d,\bar{\theta}}^{ij}[g_o^j(\tilde{\theta}), g_o^j(\tilde{\theta})] - cov_{d,\hat{\theta}}^{ij}[g_o^j(\tilde{\theta}), g_o^j(\tilde{\theta})] \right] \right\} \right| \\
&\leq \omega^i \zeta\bar{t} \sum_j \frac{\gamma^{ij}}{1-\gamma^{iL}} \Big(\mathcal{L}([g_d^i]') \left| \bar{\theta} - \hat{\theta} \right| \left| cov_{d,\bar{\theta}}^{ij}[g_o^j(\tilde{\theta}), g_o^j(\tilde{\theta})] \right| \\
&\quad + [g_d^i]'(\bar{\theta}) \left| cov_{d,\hat{\hat{\theta}}}^{ij}([g_o^j(\tilde{\theta})]^2, g_o^j(\tilde{\theta})) - 2cov_{d,\hat{\hat{\theta}}}^{ij}[g_o^j(\tilde{\theta}), g_o^j(\tilde{\theta})] (\sum_o \int \chi_{do}^{ij}(\hat{\hat{\theta}}, \tilde{\theta}) g_o^j(\tilde{\theta}) \mathrm{d}\tilde{\theta}) \right| \left| \bar{\theta} - \hat{\theta} \right| \Big) \\
&\leq [\omega^i \zeta\bar{t} M^2 \overline{M}'' + 3\omega^i \zeta\bar{t} M^3] \left| \bar{\theta} - \hat{\theta} \right|,
\end{aligned}$$

where $\mathcal{L}(\cdot)$ denotes the Lipschitz constant of a Lipschitz continuous function $(\cdot)$, $\hat{\hat{\theta}}$ is between $\bar{\theta}$ and $\hat{\theta}$, and the third inequality applies the definition of Lipschitz continuity of $\boldsymbol{g}'$ and the mean value theorem. Therefore, $[\mathcal{T}\boldsymbol{g}]_d^{i\,\prime}(\bar{\theta})$ is Lipschitz continuous with a Lipschitz constant $\omega^i \zeta\bar{t} M^2 \overline{M}'' + 3\omega^i \zeta\bar{t} M^3$ which satisfies $\omega^i \zeta\bar{t} M^2 \overline{M}'' + 3\omega^i \zeta\bar{t} M^3 \leq \overline{M}''$ by the definition of $\overline{M}''$. These prove part (3).

Since $g_o^j(\tilde{\theta}) \in [-M, M], \forall \tilde{\theta}$, we have that $[\mathcal{T}\boldsymbol{g}]_d^i(\bar{\theta}) - \bar{\theta} > 0$ for $\bar{\theta} = -M$ and $[\mathcal{T}\boldsymbol{g}]_d^i(\bar{\theta}) - \bar{\theta} < 0$ for $\bar{\theta} = M$. By the intermediate value theorem, $\exists \bar{\theta}^* \in [-M, M]$, $[\mathcal{T}\boldsymbol{g}]_d^i(\bar{\theta}^*) = \bar{\theta}^*$. Now consider $\forall\bar{\theta} \in (\bar{\theta}^*, M]$, we have that

$$[\mathcal{T}\boldsymbol{g}]^i_d(\bar{\theta}) \leq \bar{\theta}^* + \overline{M}'(\bar{\theta} - \bar{\theta}^*) \leq \bar{\theta}^* + (\bar{\theta} - \bar{\theta}^*) \leq \bar{\theta} \leq M,$$

since $\overline{M}' < 1$. Similarly, $\forall\bar{\theta} \in [-M, \bar{\theta}^*)$

$$[\mathcal{T}\boldsymbol{g}]^i_d(\bar{\theta}) \geq \bar{\theta}^* - \overline{M}'(\bar{\theta}^* - \bar{\theta}) \geq \bar{\theta}^* - (\bar{\theta}^* - \bar{\theta}) \geq \bar{\theta} \geq -M.$$

We thus have $[\mathcal{T}\boldsymbol{g}]^i_d(\bar{\theta}) \in [-M, M], \forall\bar{\theta}$. This proves part (2).

□

Proposition A.2. *Given wages, assume $\{\overline{\Theta}^j_o\}$ have bounded support that is contained in $[-M, M]$ for a positive constant $M > 0$ and have associated density functions $\bar{\varsigma}^j_o$. If $\bar{\phi} > 0$ and $\zeta\bar{t} < 1/M^2$, then an equilibrium exists which satisfies that the policy function $\boldsymbol{g}$ is twice differentiable; $\|\boldsymbol{g}\|_\infty \leq M$; $[g^i_d]'(\bar{\theta}) \in [\underline{M}', \overline{M}'], \forall(\bar{\theta}, d, i)$; $\boldsymbol{g}'$ is Lipschitz continuous with a Lipschitz constant $\overline{M}''$; for constants $\underline{M}', \overline{M}', \overline{M}''$ defined in Definition A.1. Moreover, under such an equilibrium, the first order condition of the technology direction choice problem has a unique solution that characterizes the optimal decision.*

Proof. Define $\mathcal{G} = \{\boldsymbol{g} : [-M, M] \to \mathbb{R}^{N\times S}, \boldsymbol{g}$ is differentiable, $\|\boldsymbol{g}\|_\infty \leq M$; $[g^i_d]'(\bar{\theta}) \in [\underline{M}', \overline{M}']$, $\boldsymbol{g}'$ is Lipschitz continuous with a Lipschitz constant $\overline{M}''$ $\}$, for $\underline{M}', \overline{M}', \overline{M}''$ defined in the proposition. Equip $\mathcal{G}$ with the C^1 norm: $\|\boldsymbol{g}\|_{\mathcal{G}} = \|\boldsymbol{g}\|_\infty + \|\boldsymbol{g}'\|_\infty$. It can be shown that $\mathcal{G}$ is closed. By the Arzelà-Ascoli theorem, $\mathcal{G}$ is compact under the norm $\|\cdot\|_{\mathcal{G}}$.

From Lemma A.4, $\mathcal{T}$ defined by (A.6) is continuous under the C^1 norm of $\mathcal{G}$ and $\mathcal{T}\boldsymbol{g} \in \mathcal{G}$. By the Schauder fixed-point theorem, $\mathcal{G}$ contains a fixed point of $\mathcal{T}$.

Now for a fixed point $\boldsymbol{g}$, we verify the second order optimality condition holds for the technology direction choice problem. From the proof in Lemma A.3 we have (to save notations we use the original definition of $\mathcal{C}^i_d$ and χ^{ij}_{do} that takes θ instead of $\bar{\theta}$ as their first arguments)

$$\begin{aligned}
[\mathcal{C}^i_d]'(\theta) &= \sum_j \gamma^{ij} \sum_o \int \chi^{ij}_{do}(\theta, \tilde{\theta}) \left[\bar{t}(\theta - g^j_o(\tilde{\theta}))\right] \mathrm{d}\tilde{\theta} \\
\Rightarrow \mathcal{C}^{i\,\prime\prime}_d(\theta) &= \sum_{jo} \gamma^{ij} \int \frac{\partial \chi^{ij}_{do}(\theta, \tilde{\theta})}{\partial \theta} \left[\bar{t}(\theta - g^j_o(\tilde{\theta}))\right] \mathrm{d}\tilde{\theta} + \underbrace{\sum_{jo} \gamma^{ij} \int \chi^{ij}_{do}(\theta, \tilde{\theta}) \left[\bar{t}\right] \mathrm{d}\tilde{\theta}}_{\bar{t}(1-\gamma^{iL})}, \\
&= -\zeta\bar{t}^2 \sum_j \gamma^{ij} var^i_{do}(g^j_o(\tilde{\theta})) + \bar{t}(1 - \gamma^{iL}) \\
&\in [\bar{t}(1 - \gamma^{iL})(1 - \zeta\bar{t}M^2), \bar{t}(1 - \gamma^{iL})].
\end{aligned}$$

in which the third line applies part (1) of Lemma A.6 and the last line applies that $\|\mathcal{T}\boldsymbol{g}\|_\infty \leq M$. Therefore, the second derivative of the objective with respect to choice θ

$$-\bar{\phi} - (\eta - 1)[\mathcal{C}^i_d]''(\theta) < 0$$

globally under the premise that $1 - \zeta\bar{t}M^2 \geq 0$. □

Lemma A.5. *Denote $\widetilde{\mathcal{G}} = \{\boldsymbol{g} \in C^0([-M, M] \to \mathbb{R}^{N\times S}) : \|\boldsymbol{g}\|_\infty \leq M\}$, $\widetilde{\mathcal{G}}' = \{\boldsymbol{g}' \in C^0([-M, M] \to \mathbb{R}^{N\times S}) : \underline{M}' \leq \boldsymbol{g}' \leq \overline{M}'\}$, and $\mathbb{C} = \{\boldsymbol{\mathcal{C}} \in C^0([-M, M] \to \mathbb{R}^{N\times S}) : \|\boldsymbol{\mathcal{C}}\|_\infty \leq M^{\mathcal{C}}\}$, for $\underline{M}'$, $\overline{M}'$,* *and $M^{\mathcal{C}}$ defined in Definition A.1. Define $\mathcal{T}^g, \mathcal{T}^{g'}, \mathcal{T}^{\mathcal{C}}$ mapping from $\widetilde{\mathcal{G}} \times \widetilde{\mathcal{G}}' \times \mathbb{C}$, given below*

$$[\mathcal{T}^g(\boldsymbol{g},\boldsymbol{g}',\boldsymbol{\mathcal{C}})]^i_d(\bar\theta) = \omega^i \sum_{j,o} \frac{\gamma^{ij}}{1-\gamma^{iL}} \int \chi^{ij}_{do}(\bar\theta,\tilde\theta;\boldsymbol{g},\boldsymbol{g}',\boldsymbol{\mathcal{C}}) g^j_o(\tilde\theta)\mathrm{d}\tilde\theta + (1-\omega^i)\bar\theta,$$

$$[\mathcal{T}^{g'}(\boldsymbol{g},\boldsymbol{g}',\boldsymbol{\mathcal{C}})]^i_d(\bar\theta) = \omega^i \sum_{j} \frac{\gamma^{ij}}{1-\gamma^{iL}} \sum_o \int \partial_{\bar\theta}\chi^{ij}_{do}(\bar\theta,\tilde\theta;\boldsymbol{g},\boldsymbol{g}',\boldsymbol{\mathcal{C}}) g^j_o(\tilde\theta)\mathrm{d}\tilde\theta + (1-\omega^i),$$

$$[\mathcal{T}^{\mathcal{C}}(\boldsymbol{g},\boldsymbol{g}',\boldsymbol{\mathcal{C}})]^i_d(\bar\theta) = \ln\xi^i_d - \sum_j \frac{\gamma^{ij}}{\zeta} \ln\Big(\sum_o \int [\tau^j_{do}]^{-\zeta} \exp\Big(-\zeta\mathcal{C}^j_o(\tilde\theta)\Big) \exp\Big(-\frac{1}{2}\zeta\bar t(g^i_d(\bar\theta)-g^j_o(\tilde\theta))^2\Big) [g^j_o]'(\tilde\theta)\bar\varsigma^j_o(\tilde\theta)\mathrm{d}\tilde\theta\Big),$$

where to slightly abuse notations, $\boldsymbol{\chi}$ is the one defined in (A.8), but also highlighting the dependence on $\boldsymbol{\mathcal{C}}$:

$$\chi^{ij}_{do}(\bar\theta,\tilde\theta;\boldsymbol{g},\boldsymbol{g}',\boldsymbol{\mathcal{C}}) \equiv \frac{[\tau^j_{do}]^{-\zeta} \exp\Big(-\zeta\mathcal{C}^j_o(\tilde\theta)\Big) \exp\Big(-\frac{1}{2}\zeta\bar t(g^i_d(\bar\theta)-g^j_o(\tilde\theta))^2\Big) [g^j_o]'(\tilde\theta)\bar\varsigma^j_o(\tilde\theta)}{\sum_m \int [\tau^j_{dm}]^{-\zeta} \exp\Big(-\zeta\mathcal{C}^j_m(\tilde{\tilde\theta})\Big) \exp\Big(-\frac{1}{2}\zeta\bar t(g^i_d(\bar\theta)-g^j_m(\tilde{\tilde\theta}))^2\Big) [g^j_m]'(\tilde{\tilde\theta})\bar\varsigma^j_o(\tilde{\tilde\theta})\mathrm{d}\tilde{\tilde\theta}}. \tag{A.9}$$

Then we have

(1) $\mathcal{T}^g(\boldsymbol{g},\boldsymbol{g}',\boldsymbol{\mathcal{C}}) \in \widetilde{\mathcal{G}}$, $\mathcal{T}^{g'}(\boldsymbol{g},\boldsymbol{g}',\boldsymbol{\mathcal{C}}) \in \widetilde{\mathcal{G}}'$, $\mathcal{T}^{\mathcal{C}}(\boldsymbol{g},\boldsymbol{g}',\boldsymbol{\mathcal{C}}) \in \mathbb{C}$.

(2) $\sum_{m,k} \|\partial^k_{g,m}\mathcal{T}^g\| \le \overline{\omega}\Big[3\zeta\bar t M^2+1\Big]$. $\sum_{m,k} \|\partial^k_{g',m}\mathcal{T}^g\| \le \frac{\overline{\omega}\zeta M}{\underline{M}'}$. $\sum_{m,k} \|\partial^k_{\mathcal{C},m}\mathcal{T}^g\| \le \overline{\omega}\zeta M$.

(3) $\sum_{m,k} \|\partial^k_{g,m}\mathcal{T}^{g'}\| \le 3\overline{\omega}\zeta^2\bar t^2 M^3$. $\sum_{m,k} \|\partial^k_{g',m}\mathcal{T}^{g'}\| \le 3\overline{\omega}\zeta^2\bar t\frac{M^2}{\underline{M}'} + \overline{\omega}\zeta t M^2$. $\sum_{m,k} \|\partial^k_{\mathcal{C},m}\mathcal{T}^{g'}\| \le 3\overline{\omega}\zeta^2\bar t M^2$.

(4) $\sum_{m,k} \|\partial^k_{g,m}\mathcal{T}^{\mathcal{C}}\| \le (1-\underline{\gamma}^L)2tM$. $\sum_{m,k} \|\partial^k_{g',m}\mathcal{T}^{\mathcal{C}}\| \le (1-\underline{\gamma}^L)\frac{1}{\underline{M}'}$. $\sum_{m,k} \|\partial^k_{\mathcal{C},m}\mathcal{T}^{\mathcal{C}}\| \le 1-\underline{\gamma}^L$.

Proof. Combining the proofs for part (2) and (3) of Lemma A.4, and the proof for part (1) of Lemma A.3 we have proved part (1).

For part (2), suppressing the argument $(\boldsymbol{g},\boldsymbol{g}',\boldsymbol{\mathcal{C}})$ to simplify notations, consider

$$[\partial^k_{g,m}\mathcal{T}^g]^i_d(\bar\theta) = \omega^i \sum_j \frac{\gamma^{ij}}{1-\gamma^{iL}} \Big\{ \sum_o \int \partial^k_{g,m}\chi^{ij}_{do}(\bar\theta,\tilde\theta) g^j_o(\tilde\theta)\mathrm{d}\tilde\theta + \sum_o \int \chi^{ij}_{do}(\bar\theta,\tilde\theta)\partial^k_{g,m}g^j_o(\tilde\theta)\mathrm{d}\tilde\theta \Big\}.$$

Apply part (1) of Lemma A.7:

$$\sum_o \int \partial^k_{g,m}\chi^{ij}_{do}(\bar\theta,\tilde\theta) g^j_o(\tilde\theta)\mathrm{d}\tilde\theta = -\zeta\Big\{ -tg^i_d(\bar\theta)cov^{ij}_d(\partial^k_{g,m}g^j_o(\tilde\theta), g^j_o(\tilde\theta))$$
$$-\bar t[\partial^k_{g,m}g^i_d(\bar\theta)]cov^{ij}_d(g^j_o(\tilde\theta), g^j_o(\tilde\theta)) + tcov^{ij}_d(g^j_o(\tilde\theta)\partial^k_{g,m}g^j_o(\tilde\theta), g^j_o(\tilde\theta)\Big\}.$$

By Popoviciu's inequality on variances and note that $\partial^k_{g,m}g^i_d = 1$ for $(m,k) = (d,i)$ and zero otherwise, we have

$$\sum_{m,k} \Big| \sum_o \int \partial^k_{g,m}\chi^{ij}_{do}(\bar\theta,\tilde\theta) g^j_o(\tilde\theta)\mathrm{d}\tilde\theta \Big| \le \zeta\Big\{\bar t M^2 + tM^2 + tM^2\Big\} = 3\zeta\bar t M^2.$$

Therefore,

$$\sum_{m,k} \Big| [\partial^k_{g,m}\mathcal{T}^g]^i_d(\bar\theta) \Big| \le \omega^i\Big[3\zeta\bar t M^2+1\Big].$$

Consider

$$[\partial^k_{g',m}\mathcal{T}^g]^i_d(\bar\theta) = \omega^i \sum_j \frac{\gamma^{ij}}{1-\gamma^{iL}} \sum_o \int \partial^k_{g',m}\chi^{ij}_{do}(\bar\theta,\tilde\theta) g^j_o(\tilde\theta)\mathrm{d}\tilde\theta$$
$$= -\omega^i \sum_j \frac{\gamma^{ij}}{1-\gamma^{iL}} \zeta cov\Big(\partial^k_{g',m}\ln[g^j_o]'(\tilde\theta), g^j_o(\tilde\theta)\Big),$$

which applies part (2) of Lemma A.7. Note that $\partial^k_{g',m}\ln[g^j_o]' = 1/[g^j_o]'$ for $(m,k) = (d,i)$ and zero otherwise; we have

$$\sum_{m,k}\Big|[\partial^k_{g',m}\mathcal{T}^g]^i_d(\bar\theta)\Big| \le \frac{\omega^i\zeta M}{\underline{M}'}.$$

Consider

$$[\partial^k_{\mathcal{C},m}\mathcal{T}^g]^i_d(\bar\theta) = \omega^i \sum_j \frac{\gamma^{ij}}{1-\gamma^{iL}} \sum_o \int \partial^k_{\mathcal{C},m}\chi^{ij}_{do}(\bar\theta,\tilde\theta) g^j_o(\tilde\theta)\mathrm{d}\tilde\theta$$
$$= -\omega^i \sum_j \frac{\gamma^{ij}}{1-\gamma^{iL}} \zeta cov\Big(\partial^k_{\mathcal{C},m}\mathcal{C}^j_o(\tilde\theta), g^j_o(\tilde\theta)\Big)$$
$$\Rightarrow \sum_{m,k}\Big|[\partial^k_{\mathcal{C},m}\mathcal{T}^g]^i_d(\bar\theta)\Big| \le \omega^i\zeta M.$$

For part (3), expand

$$[\mathcal{T}^{g'}]^i_d(\bar\theta) = \omega^i \sum_j \frac{\gamma^{ij}}{1-\gamma^{iL}} \zeta\bar t[g^i_d]'(\bar\theta)\Big[\sum_o \int \chi^{ij}_{do}(\bar\theta,\tilde\theta)[g^j_o(\tilde\theta)]^2\mathrm{d}\tilde\theta - \Big(\sum_o \int \chi^{ij}_{do}(\bar\theta,\tilde\theta)g^j_o(\tilde\theta)\mathrm{d}\tilde\theta\Big)^2\Big] + (1-\omega^i)$$

Apply part (1) of Lemma A.7:

$$[\partial^k_{g,m}\mathcal{T}^{g'}]^i_d(\bar\theta) = \omega^i \sum_j \frac{\gamma^{ij}}{1-\gamma^{iL}} \zeta\bar t[g^i_d]'(\bar\theta)\Big[-\zeta\Big\{-tg^i_d(\bar\theta)cov^{ij}_d(\partial^k_{g,m}g^j_o(\tilde\theta),[g^j_o(\tilde\theta)]^2)$$
$$-\bar t\partial^k_{g,m}g^i_d(\bar\theta)cov^{ij}_d(g^j_o(\tilde\theta),[g^j_o(\tilde\theta)]^2) + tcov^{ij}_d(g^j_o(\tilde\theta)\partial^k_{g,m}g^j_o(\tilde\theta),[g^j_o(\tilde\theta)]^2\Big\}$$
$$+2\zeta\Big(\sum_o \int \chi^{ij}_{do}(\bar\theta,\tilde\theta;\boldsymbol{g},\boldsymbol{g}')g^j_o(\tilde\theta)\mathrm{d}\tilde\theta\Big)\Big\{-tg^i_d(\bar\theta)cov^{ij}_d(\partial^k_{g,m}g^j_o(\tilde\theta),g^j_o(\tilde\theta))$$
$$-\bar t\partial^k_{g,m}g^i_d(\bar\theta)cov^{ij}_d(g^j_o(\tilde\theta),g^j_o(\tilde\theta)) + tcov^{ij}_d(g^j_o(\tilde\theta)\partial^k_{g,m}g^j_o(\tilde\theta),g^j_o(\tilde\theta)\Big\}\Big]$$
$$\Rightarrow \sum_{m,k}\|[\partial^k_{g,m}\mathcal{T}^{g'}]^i_d\|_\infty \le \omega^i\zeta^2\bar t^2\Big[\Big\{M^3+M^3+M^3\Big\} + 2M\Big\{M^2+M^2+M^2\Big\}\Big] = 3\omega^i\zeta^2\bar t^2 M^3$$

Apply part (2) of Lemma A.7:

$$[\partial^k_{g',m}\mathcal{T}^{g'}]^i_d(\bar\theta) = \omega^i \sum_j \frac{\gamma^{ij}}{1-\gamma^{iL}} \zeta\bar t[g^i_d]'(\bar\theta)\Big[-\zeta\Big\{cov^{ij}_d(\partial^k_{g',m}\ln[g^j_o]'(\tilde\theta),[g^j_o(\tilde\theta)]^2)\Big\}$$
$$+2\zeta\Big(\sum_o \int \chi^{ij}_{do}(\bar\theta,\tilde\theta)g^j_o(\tilde\theta)\mathrm{d}\tilde\theta\Big)\Big\{cov^{ij}_d(\partial^k_{g',m}\ln[g^j_o]'(\tilde\theta),g^j_o(\tilde\theta))\Big\}\Big] + \omega^i\sum_j \frac{\gamma^{ij}}{1-\gamma^{iL}}\zeta\bar t cov^{ij}_d[g^j_o(\tilde\theta),g^j_o(\tilde\theta)]$$
$$\Rightarrow \sum_{m,k}\|[\partial^k_{g',m}\mathcal{T}^{g'}]^i_d\|_\infty \le \omega^i\zeta\bar t\Big[\zeta\Big\{\frac{M^2}{\underline{M}'}\Big\} + 2\zeta M\Big\{\frac{M}{\underline{M}'}\Big\}\Big] + \omega^i\zeta\bar t M^2 = 3\omega^i\zeta^2\bar t\frac{M^2}{\underline{M}'} + \omega^i\zeta t M^2$$

Apply part (3) of Lemma A.7:

$$[\partial^k_{\mathcal{C},m}\mathcal{T}^{g'}]^i_d(\bar\theta) = \omega^i \sum_j \frac{\gamma^{ij}}{1-\gamma^{iL}}\zeta\bar t[g^i_d]'(\bar\theta)\Big[-\zeta\Big\{cov^{ij}_d(\partial^k_{\mathcal{C},m}\mathcal{C}^j_o(\tilde\theta),[g^j_o(\tilde\theta)]^2)\Big\}$$
$$+2\zeta\Big(\sum_o\int\chi^{ij}_{do}(\bar\theta,\tilde\theta)g^j_o(\tilde\theta)\mathrm{d}\tilde\theta\Big)\Big\{cov^{ij}_d(\partial^k_{\mathcal{C},m}\mathcal{C}^j_o(\tilde\theta),g^j_o(\tilde\theta))\Big\}\Big]$$
$$\Rightarrow \sum_{m,k}\|[\partial^k_{\mathcal{C},m}\mathcal{T}^{g'}]^i_d\|_\infty \le \omega^i\zeta\bar t\Big[\zeta\Big\{M^2\Big\}+2\zeta M\Big\{M\Big\}\Big] = 3\omega^i\zeta^2\bar t M^2$$

For part (4). Directly apply the chain rule

$$\partial^k_{g,m}[\mathcal{T}^{\mathcal{C}}]^i_d(\bar\theta) = \sum_j\gamma^{ij}\sum_o\int\chi^{ij}_{do}(\bar\theta,\tilde\theta)\Big[\bar t(g^i_d(\bar\theta)-g^j_o(\tilde\theta))(\partial^k_{g,m}g^i_d(\bar\theta)-\partial^k_{g,m}g^j_o(\tilde\theta))\Big]\mathrm{d}\tilde\theta$$
$$\Rightarrow \sum_{m,k}\|\partial^k_{g,m}[\mathcal{T}^{\mathcal{C}}]^i_d\|_\infty \le (1-\underline{\gamma}^L)2tM.$$
$$\partial^k_{g',m}[\mathcal{T}^{\mathcal{C}}]^i_d(\bar\theta) = \sum_j\gamma^{ij}\sum_o\int\chi^{ij}_{do}(\bar\theta,\tilde\theta)\Big[\partial^k_{g',m}\ln[g^j_o]'(\tilde\theta)\Big]\mathrm{d}\tilde\theta$$
$$\Rightarrow \sum_{m,k}\|\partial^k_{g',m}[\mathcal{T}^{\mathcal{C}}]^i_d\|_\infty \le (1-\underline{\gamma}^L)\frac{1}{\underline{M}'}.$$
$$\partial^k_{\mathcal{C},m}[\mathcal{T}^{\mathcal{C}}]^i_d(\bar\theta) = \sum_j\gamma^{ij}\sum_o\int\chi^{ij}_{do}(\bar\theta,\tilde\theta)\partial^k_{\mathcal{C},m}\mathcal{C}^j_o(\tilde\theta)\mathrm{d}\tilde\theta$$
$$\Rightarrow \sum_{m,k}\|\partial^k_{\mathcal{C},m}[\mathcal{T}^{\mathcal{C}}]^i_d\|_\infty \le 1-\underline{\gamma}^L.$$

□

Proposition A.3. *Given wages, $\{w_d\}$, assume $\{\overline{\Theta}^j_o\}$ have bounded support that is contained in $[-M,M]$ for some $M>0$ and have associated density functions $\bar\varsigma^j_o$. If $\zeta\bar t < 1/M^2$, $\bar t < \frac{1}{2M}$ and $\bar\phi > \underline{\phi}$, where $\underline{\phi} > 0$ is a constant determined by parameters $(\zeta,\bar t,\eta,M,\underline{\gamma}^L)$ detailed in the proof, then an equilibrium uniquely exists and satisfies all properties stated in Proposition A.2.*

Proof. Denote $\mathcal{G} = \{\boldsymbol{g} : [-M,M] \to \mathbb{R}^{N\times S}; \boldsymbol{g}$ is differentiable; $\|\boldsymbol{g}\|_\infty \le M$; $[g^i_d]'(\bar\theta) \in [\underline{M}',\overline{M}'], \forall d,i,\bar\theta\}$. Denote $\mathbb{C} = \{\boldsymbol{\mathcal{C}} \in C^0([-M,M] \to \mathbb{R}^{N\times S}) : \|\boldsymbol{\mathcal{C}}\|_\infty \le M^{\mathcal{C}}\}$. Denote $\mathcal{X} = \mathcal{G}\times\mathbb{C}$. Endow $\mathcal{X}$ with the C^1 norm of $\boldsymbol{g}$ combined with the C^0 norm of $\boldsymbol{\mathcal{C}}$: $\|(\boldsymbol{g},\boldsymbol{\mathcal{C}})\|_{\mathcal{X}} = \|\boldsymbol{g}\|_\infty + \|\boldsymbol{g}'\|_\infty + \|\boldsymbol{\mathcal{C}}\|_\infty$. It can be verified that $\mathcal{X}$ is a complete metric space with the norm $\|\cdot\|_{\mathcal{X}}$.

Define $\widetilde{\mathcal{T}} = (\widetilde{\mathcal{T}}^g,\widetilde{\mathcal{T}}^{\mathcal{C}})$ mapping from $\mathcal{X}$ given by

$$[\widetilde{\mathcal{T}}^g(\boldsymbol{g},\boldsymbol{\mathcal{C}})]^i_d(\bar\theta) = \omega^i\sum_{j,o}\frac{\gamma^{ij}}{1-\gamma^{iL}}\int\chi^{ij}_{do}(\bar\theta,\tilde\theta;\boldsymbol{g},\boldsymbol{\mathcal{C}})g^j_o(\tilde\theta)\mathrm{d}\tilde\theta + (1-\omega^i)$$

$$[\widetilde{\mathcal{T}}^{\mathcal{C}}(\boldsymbol{g},\boldsymbol{\mathcal{C}})]^i_d(\bar\theta) = \ln\xi^i_d - \sum_j\frac{\gamma^{ij}}{\zeta}\ln\Big(\sum_o\int[\tau^j_{do}]^{-\zeta}\exp\Big(-\zeta\mathcal{C}^j_o(\tilde\theta)\Big)\exp\Big(-\frac{1}{2}\zeta\bar t(g^i_d(\bar\theta)-g^j_o(\tilde\theta))^2\Big)[g^j_o]'(\tilde\theta)\bar\varsigma^j_o(\tilde\theta)\mathrm{d}\tilde\theta\Big),$$

where $\boldsymbol{g}'$ is viewed as an operator applied to $\boldsymbol{g}$, and to slightly abuse notations, $\boldsymbol{\chi}$ is redefined below to highlight its dependence on $(\boldsymbol{g},\boldsymbol{\mathcal{C}})$:

$$\chi^{ij}_{do}(\bar{\theta},\tilde{\theta};\boldsymbol{g},\boldsymbol{\mathcal{C}}) \equiv \frac{[\tau^{j}_{do}]^{-\zeta}\exp\left(-\zeta\mathcal{C}^{j}_{o}(\tilde{\theta})\right)\exp\left(-\frac{1}{2}\zeta\bar{t}(g^{i}_{d}(\bar{\theta})-g^{j}_{o}(\tilde{\theta}))^{2}\right)[g^{j}_{o}]'(\tilde{\theta})\overline{\varsigma}^{j}_{o}(\tilde{\theta})}{\sum_{m}\int[\tau^{j}_{dm}]^{-\zeta}\exp\left(-\zeta\mathcal{C}^{j}_{m}(\tilde{\tilde{\theta}})\right)\exp\left(-\frac{1}{2}\zeta\bar{t}(g^{i}_{d}(\bar{\theta})-g^{j}_{m}(\tilde{\tilde{\theta}}))^{2}\right)[g^{j}_{m}]'(\tilde{\tilde{\theta}})\overline{\varsigma}^{j}_{o}(\tilde{\tilde{\theta}})\mathrm{d}\tilde{\tilde{\theta}}}.$$

Part (1) of Lemma A.5 shows that $\widetilde{\mathcal{T}}(\boldsymbol{g},\boldsymbol{\mathcal{C}}) \in \mathcal{X}$.
Consider $\forall(\boldsymbol{g},\boldsymbol{\mathcal{C}}) \in \mathcal{X}, (\hat{\boldsymbol{g}},\hat{\boldsymbol{\mathcal{C}}}) \in \mathcal{X}$

$$\begin{aligned}
&\|\widetilde{\mathcal{T}}(\boldsymbol{g},\boldsymbol{\mathcal{C}})-\widetilde{\mathcal{T}}(\hat{\boldsymbol{g}},\hat{\boldsymbol{\mathcal{C}}})\|_{\mathcal{X}}\\
=&\|\widetilde{\mathcal{T}}^{g}(\boldsymbol{g},\boldsymbol{\mathcal{C}})-\widetilde{\mathcal{T}}^{g}(\hat{\boldsymbol{g}},\hat{\boldsymbol{\mathcal{C}}})\|_{\infty}+\|[\widetilde{\mathcal{T}}^{g}(\boldsymbol{g},\boldsymbol{\mathcal{C}})]'-[\widetilde{\mathcal{T}}^{g}(\hat{\boldsymbol{g}},\hat{\boldsymbol{\mathcal{C}}})]'\|_{\infty}+\|\widetilde{\mathcal{T}}^{\mathcal{C}}(\boldsymbol{g},\boldsymbol{\mathcal{C}})-\widetilde{\mathcal{T}}^{\mathcal{C}}(\hat{\boldsymbol{g}},\hat{\boldsymbol{\mathcal{C}}})\|_{\infty}\\
\leq&(\sum_{m,k}\|\partial^{k}_{g,m}\mathcal{T}^{g}\|+\sum_{m,k}\|\partial^{k}_{g,m}\mathcal{T}^{g'}\|+\sum_{m,k}\|\partial^{k}_{g,m}\mathcal{T}^{\mathcal{C}}\|)\|\boldsymbol{g}-\hat{\boldsymbol{g}}\|_{\infty}\\
&+(\sum_{m,k}\|\partial^{k}_{g',m}\mathcal{T}^{g}\|+\sum_{m,k}\|\partial^{k}_{g',m}\mathcal{T}^{g'}\|+\sum_{m,k}\|\partial^{k}_{g',m}\mathcal{T}^{\mathcal{C}}\|)\|\boldsymbol{g}'-\hat{\boldsymbol{g}}'\|_{\infty}\\
&+(\sum_{m,k}\|\partial^{k}_{\mathcal{C},m}\mathcal{T}^{g}\|+\sum_{m,k}\|\partial^{k}_{\mathcal{C},m}\mathcal{T}^{g'}\|+\sum_{m,k}\|\partial^{k}_{\mathcal{C},m}\mathcal{T}^{\mathcal{C}}\|)\|\boldsymbol{\mathcal{C}}-\hat{\boldsymbol{\mathcal{C}}}\|_{\infty},
\end{aligned}$$

where the second line applies the definition of $\|\cdot\|_{\mathcal{X}}$, the third line applies the mean value theorem for Frechet derivatives,[1] and $\mathcal{T}^{g},\mathcal{T}^{g'},\mathcal{T}^{\mathcal{C}}$ are the operators defined in Lemma A.5. From the estimates in part (2)-(4) of Lemma A.5

$$\|\widetilde{\mathcal{T}}(\boldsymbol{g},\boldsymbol{\mathcal{C}})-\widetilde{\mathcal{T}}(\hat{\boldsymbol{g}},\hat{\boldsymbol{\mathcal{C}}})\|_{\mathcal{X}} \leq \Omega_{g}\|\boldsymbol{g}-\hat{\boldsymbol{g}}\|_{\infty}+\Omega_{g'}\|\boldsymbol{g}'-\hat{\boldsymbol{g}}'\|_{\infty}+\Omega_{\mathcal{C}}\|\boldsymbol{\mathcal{C}}-\hat{\boldsymbol{\mathcal{C}}}\|_{\infty},$$

where $\Omega_{g}=\overline{\omega}\left[3\zeta\bar{t}M^{2}+1\right]+3\overline{\omega}\zeta^{2}\bar{t}^{2}M^{3}+(1-\underline{\gamma}^{L})2tM$, $\Omega_{g'}=\frac{\overline{\omega}\zeta M}{\underline{M'}}+3\overline{\omega}\zeta^{2}\bar{t}\frac{M^{2}}{\underline{M'}}+\overline{\omega}\zeta tM^{2}+(1-\underline{\gamma}^{L})\frac{1}{1-\overline{\omega}}$, $\Omega_{\mathcal{C}}=\overline{\omega}\zeta M+3\overline{\omega}\zeta^{2}\bar{t}M^{2}+1-\underline{\gamma}^{L}$. Since it is assumed that $2tM<1$, $\overline{\Omega}\equiv\max\{\Omega_{g},\Omega_{g'},\Omega_{\mathcal{C}}\}$ is thus increasing in $\overline{\omega}$ and $\overline{\Omega}\Big|_{\overline{\omega}=0}=(1-\underline{\gamma}^{L})<1$. Choose any $\overline{\Omega}^{*}\in(1-\underline{\gamma}^{L},1)$. Since $\overline{\omega}=\max_{i}\frac{(\eta-1)(1-\gamma^{iL})\bar{t}}{(\eta-1)(1-\gamma^{iL})\bar{t}+\bar{\phi}}$ which is decreasing in $\bar{\phi}$, $\exists\underline{\phi}>0$ such that for all $\bar{\phi}>\underline{\phi}$, $\overline{\Omega}<\overline{\Omega}^{*}$. We thus have found $\underline{\phi}$ such that for all $\bar{\phi}>\underline{\phi}$, $\widetilde{\mathcal{T}}$ is a contraction mapping with the norm $\|\cdot\|_{\mathcal{X}}$ with a modulus $\overline{\Omega}^{*}$. The existence and uniqueness of the equilibrium then follows from the contraction mapping theorem.

Since the conditions stated in Proposition A.2 also hold, we have that the unique fixed point also satisfies the properties stated in Proposition A.2.

□

Lemma A.6. *For $\boldsymbol{\chi}$ defined in (A.8) and any function $f_{o}(\tilde{\theta})$ we have*

(1) $\sum_{o}\int\partial_{\bar{\theta}}\chi^{ij}_{do}(\bar{\theta},\tilde{\theta};\boldsymbol{g},\boldsymbol{g}')f_{o}(\tilde{\theta})\mathrm{d}\tilde{\theta}=\zeta\bar{t}[g^{i}_{d}]'(\bar{\theta})cov^{ij}_{d}(g^{j}_{o}(\tilde{\theta}),f_{o}(\tilde{\theta}))$, where cov^{ij}_{d} is the variance taken under the distribution $\chi^{ij}_{do}(\bar{\theta},\tilde{\theta};\boldsymbol{g},\boldsymbol{g}')$ across $(o,\tilde{\theta})$.

[1] This can be shown as a corollary of the Hahn-Banach Theorem, see e.g., Theorem 1.8 of Ambrosetti and Prodi (1995).

Proof. For part (1), omitting $\boldsymbol{g}, \boldsymbol{g}'$ in arguments

$$\sum_o \int \partial_{\bar{\theta}} \chi^{ij}_{do}(\bar{\theta},\tilde{\theta}) f_o(\tilde{\theta}) \mathrm{d}\tilde{\theta} = \sum_o \int \partial_{\bar{\theta}} \ln \chi^{ij}_{do}(\bar{\theta},\tilde{\theta}) \cdot \chi^{ij}_{do}(\bar{\theta},\tilde{\theta}) f_o(\tilde{\theta}) \mathrm{d}\tilde{\theta}$$

$$= \sum_o \int \Big\{ -\zeta \bar{t}(g^i_d(\bar{\theta}) - g^j_o(\tilde{\theta})) + \sum_{\tilde{o}} \int \chi^{ij}_{d\tilde{o}}(\bar{\theta},\tilde{\tilde{\theta}}) \Big[\zeta \bar{t}(g^i_d(\bar{\theta}) - g^j_{\tilde{o}}(\tilde{\tilde{\theta}}) \Big] \mathrm{d}\tilde{\tilde{\theta}} \Big\} \cdot [g^i_d]'(\bar{\theta}) \cdot \chi^{ij}_{do}(\theta,\tilde{\theta}) f_o(\tilde{\theta}) \mathrm{d}\tilde{\theta}$$

$$= -\zeta \bar{t} [g^i_d]'(\bar{\theta}) \Big\{ \sum_o \int (g^i_d(\bar{\theta}) - g^j_o(\tilde{\theta})) f_o(\tilde{\theta}) \chi^{ij}_{do}(\bar{\theta},\tilde{\theta}) \mathrm{d}\tilde{\theta}$$

$$- \Big(\sum_o \int (g^i_d(\bar{\theta}) - g^j_o(\tilde{\theta})) \chi^{ij}_{do}(\bar{\theta},\tilde{\theta}) \mathrm{d}\tilde{\theta} \Big) \Big(\sum_o \int f_o(\tilde{\theta}) \chi^{ij}_{do}(\bar{\theta},\tilde{\theta}) \mathrm{d}\tilde{\theta} \Big) \Big\}$$

$$= -\zeta \bar{t} [g^i_d]'(\bar{\theta}) cov^{ij}_d [g^i_d(\bar{\theta}) - g^j_o(\tilde{\theta}), f_o(\tilde{\theta})]$$

$$= \zeta \bar{t} [g^i_d]'(\bar{\theta}) cov^{ij}_d [g^j_o(\tilde{\theta}), f_o(\tilde{\theta})].$$

□

Lemma A.7. *For $\boldsymbol{\chi}$ defined in (A.9) and any function $f_o(\tilde{\theta})$ we have*

(1) $\sum_o \int \partial^k_{g,m} \chi^{ij}_{do}(\bar{\theta},\tilde{\theta}) f_o(\tilde{\theta}) \mathrm{d}\tilde{\theta} = -\zeta \Big\{ -t g^i_d(\bar{\theta}) cov^{ij}_d(\partial^k_{g,m} g^j_o(\tilde{\theta}), f_o(\tilde{\theta})) - \bar{t} \partial^k_{g,m} g^i_d(\bar{\theta}) cov^{ij}_d(g^j_o(\tilde{\theta}), f_o(\tilde{\theta})) + t cov^{ij}_d(g^j_o(\tilde{\theta}) \partial^k_{g,m} g^j_o(\tilde{\theta}), f_o(\tilde{\theta}) \Big\}$.

(2) $\sum_o \int \partial^k_{g',m} \chi^{ij}_{do}(\bar{\theta},\tilde{\theta}) f_o(\tilde{\theta}) \mathrm{d}\tilde{\theta} = -\zeta cov^{ij}_d \Big(\partial^k_{g',m} \ln [g^j_o]'(\tilde{\theta}), f_o(\tilde{\theta}) \Big)$.

(3) $\sum_o \int \partial^k_{\mathcal{C},m} \chi^{ij}_{do}(\bar{\theta},\tilde{\theta}) f_o(\tilde{\theta}) \mathrm{d}\tilde{\theta} = -\zeta cov^{ij}_d \Big(\partial^k_{\mathcal{C},m} \mathcal{C}^j_o(\tilde{\theta}), f_o(\tilde{\theta}) \Big)$.

Proof. For part (1),

$$\sum_o \int \partial^k_{g,m} \chi^{ij}_{do}(\bar{\theta},\tilde{\theta}) f_o(\tilde{\theta}) \mathrm{d}\tilde{\theta} = \sum_o \int \partial^k_{g,m} \ln \chi^{ij}_{do}(\bar{\theta},\tilde{\theta}) \cdot \chi^{ij}_{do}(\bar{\theta},\tilde{\theta}) f_o(\tilde{\theta}) \mathrm{d}\tilde{\theta}$$

$$= -\zeta \sum_o \int \Big\{ \bar{t}(g^i_d(\bar{\theta}) - g^j_o(\tilde{\theta}))(\partial^k_{g,m} g^i_d(\bar{\theta}) - \partial^k_{g,m} g^j_o(\tilde{\theta}))$$

$$- \sum_{\tilde{o}} \int \chi^{ij}_{d\tilde{o}}(\bar{\theta},\tilde{\tilde{\theta}}) \Big[\bar{t}(g^i_d(\bar{\theta}) - g^j_{\tilde{o}}(\tilde{\tilde{\theta}}))(\partial^k_{g,m} g^i_d(\bar{\theta}) - \partial^k_{g,m} g^j_{\tilde{o}}(\tilde{\tilde{\theta}})) \Big] \mathrm{d}\tilde{\tilde{\theta}} \Big\} \chi^{ij}_{do}(\theta,\tilde{\theta}) f_o(\tilde{\theta}) \mathrm{d}\tilde{\theta}$$

$$= -\zeta cov^{ij}_d \Big(\bar{t}(g^i_d(\bar{\theta}) - g^j_o(\tilde{\theta}))(\partial^k_{g,m} g^i_d(\bar{\theta}) - \partial^k_{g,m} g^j_o(\tilde{\theta})), f_o(\tilde{\theta}) \Big)$$

$$= -\zeta \Big\{ -t g^i_d(\bar{\theta}) cov^{ij}_d(\partial^k_{g,m} g^j_o(\tilde{\theta}), f_o(\tilde{\theta}))$$

$$- \bar{t} \partial^k_{g,m} g^i_d(\bar{\theta}) cov^{ij}_d(g^j_o(\tilde{\theta}), f_o(\tilde{\theta})) + t cov^{ij}_d(g^j_o(\tilde{\theta}) \partial^k_{g,m} g^j_o(\tilde{\theta}), f_o(\tilde{\theta}) \Big\}.$$

For part (2),

$$\sum_o \int \partial^k_{g',m} \chi^{ij}_{do}(\bar{\theta},\tilde{\theta}) f_o(\tilde{\theta}) \mathrm{d}\tilde{\theta} = \sum_o \int \partial^k_{g',m} \ln \chi^{ij}_{do}(\bar{\theta},\tilde{\theta}) \cdot \chi^{ij}_{do}(\bar{\theta},\tilde{\theta}) f_o(\tilde{\theta}) \mathrm{d}\tilde{\theta}$$

$$= -\zeta \sum_o \int \Big\{ \partial^k_{g',m} \ln [g^j_o]'(\tilde{\theta}) - \sum_{\tilde{o}} \int \chi^{ij}_{d\tilde{o}}(\bar{\theta},\tilde{\tilde{\theta}}) \Big[\partial^k_{g',m} \ln [g^j_{\tilde{o}}]'(\tilde{\theta}) \Big] \mathrm{d}\tilde{\tilde{\theta}} \Big\} \chi^{ij}_{do}(\theta,\tilde{\theta}) f_o(\tilde{\theta}) \mathrm{d}\tilde{\theta}$$

$$= -\zeta cov^{ij}_d \Big(\partial^k_{g',m} \ln [g^j_o]'(\tilde{\theta}), f_o(\tilde{\theta}) \Big).$$

For part (3),

$$\sum_o \int \partial^k_{\mathcal{C},m} \chi^{ij}_{do}(\bar{\theta},\tilde{\theta}) f_o(\tilde{\theta}) \mathrm{d}\tilde{\theta} = \sum_o \int \partial^k_{\mathcal{C},m} \ln \chi^{ij}_{do}(\bar{\theta},\tilde{\theta}) \cdot \chi^{ij}_{do}(\bar{\theta},\tilde{\theta}) f_o(\tilde{\theta}) \mathrm{d}\tilde{\theta}$$

$$= -\zeta \sum_o \int \left\{ \partial^k_{\mathcal{C},m} \mathcal{C}^j_o(\tilde{\theta}) - \sum_{\tilde{o}} \int \chi^{ij}_{d\tilde{o}}(\bar{\theta},\tilde{\tilde{\theta}}) \left[\partial^k_{\mathcal{C},m} \mathcal{C}^j_{\tilde{o}}(\tilde{\tilde{\theta}}) \right] \mathrm{d}\tilde{\tilde{\theta}} \right\} \chi^{ij}_{do}(\theta,\tilde{\theta}) f_o(\tilde{\theta}) \mathrm{d}\tilde{\theta}$$

$$= -\zeta cov^{ij}_d \left(\partial^k_{\mathcal{C},m} \mathcal{C}^j_o(\tilde{\theta}), f_o(\tilde{\theta}) \right).$$

□

A.3 Proof of Proposition 3

Proof. For part 1, since all firms in (d,i) has ex-ante technology $\bar{\theta}^i_d$, they solve

$$\max_\theta [1 - \phi(\theta, \bar{\theta}^i_d)][C^i_d(\theta)]^{1-\eta},$$

where, with $\bar{C}^j_o \equiv C^j_o(\theta^j_o)$,

$$C^i_d(\theta) \propto \prod_j (\sum_o \left[\tau^j_{do} \bar{C}^j_o \exp(\bar{t}(\theta - \theta^j_o)^2) \right]^{-\zeta})^{-\frac{\gamma^{ij}}{\zeta}}.$$

Plugging in the functional form and taking log on the objective gives

$$-\bar{\phi}(\theta - \bar{\theta}^i_d)^2 - \frac{1-\eta}{\zeta} \sum_j \gamma^{ij} \ln(\sum_o \exp[-\zeta(\ln \tau^j_{do} + \ln \bar{C}^j_o + \bar{t}(\theta - \theta^j_o)^2)]).$$

FOC w.r.t. θ reads

$$-\bar{\phi}(\theta - \bar{\theta}^i_d) + (1-\eta) \sum_{j,o} \gamma^{ij} \frac{\exp[-\zeta(\ln \tau^j_{do} + \ln \bar{C}^j_o + \bar{t}(\theta - \theta^j_o)^2)]}{\sum_{o'} \exp[-\zeta(\ln \tau^j_{do'} + \ln \bar{C}^j_{o'} + \bar{t}(\theta - \theta^j_{o'})^2)]} \cdot \bar{t}(\theta - \theta^j_o) = 0.$$

Therefore, the technology choice of firms in (d,i), θ^i_d, should satisfy

$$\bar{\phi}(\theta^i_d - \bar{\theta}^i_d) = (1-\eta)\bar{t} \sum_{j,o} \gamma^{ij} \bar{\chi}^{ij}_{do} (\theta^i_d - \theta^j_o)$$

$$= (1-\eta)\bar{t}(1-\gamma^{iL})\theta^i_d - (1-\eta)\bar{t} \sum_{j,o} \gamma^{ij} \bar{\chi}^{ij}_{do} \theta^j_o,$$

where the share of spending by firms in (d,i) on o when sourcing j

$$\bar{\chi}^{ij}_{do} \equiv \frac{\exp[-\zeta(\ln \tau^j_{do} + \ln \bar{C}^j_o + \bar{t}(\theta^i_d - \theta^j_o)^2)]}{\sum_{o'} \exp[-\zeta(\ln \tau^j_{do'} + \ln \bar{C}^j_{o'} + \bar{t}(\theta^i_d - \theta^j_{o'})^2)]}.$$

Rearranging gives

$$\theta^i_d = \omega^i \sum_{j,o} \frac{\gamma^{ij}}{1-\gamma^{iL}} \bar{\chi}^{ij}_{do} \theta^j_o + (1-\omega^i)\bar{\theta}^i_d,$$

where

$$\omega^i \equiv \frac{(\eta-1)(1-\gamma^{iL})\bar{t}}{(\eta-1)(1-\gamma^{iL})\bar{t} + \bar{\phi}}.$$

For part 2, holding wages constant, total differentiating $\ln \bar{C}^i_d$ gives

$$\begin{aligned}
\mathrm{d}\ln \bar{C}^i_d &= \sum_{j,o} \gamma^{ij}\chi^{ij}_{do}\Big[\mathrm{d}\ln \bar{C}^j_o + 2\bar{t}(\theta^i_d - \theta^j_o)\mathrm{d}(\theta^i_d - \theta^j_o) + \mathrm{d}\ln \tau^j_{do}\Big] \\
&= 2\bar{t}(1-\gamma^{iL})[\Omega\iota]^i_d + (1-\gamma^{Li})[\Omega \mathrm{d}\ln\tilde{\tau}]^i_d,
\end{aligned}$$

with

$$\iota^i_d = \sum_{j,o} \Gamma^{ij}_{do}(\theta^i_d - \theta^j_o)\mathrm{d}(\theta^i_d - \theta^j_o),$$

and $\mathrm{d}\ln\tilde{\tau}$ defined in the proposition. Consider

$$[\Omega\iota]^j_o = \sum_{mk} \Omega^{jk}_{om} \sum_{nh} \Gamma^{kh}_{mn}(\theta^k_m - \theta^h_n)(\mathrm{d}\theta^k_m - \mathrm{d}\theta^h_n)$$

Utilizing $\boldsymbol{\theta} \circ \mathrm{d}\boldsymbol{\theta} = D_{\boldsymbol{\theta}} \cdot \mathrm{d}\boldsymbol{\theta}$, we write the four terms:

$$\begin{aligned}
\sum_{mk} \Omega^{jk}_{om} \sum_{nh} \Gamma^{kh}_{mn}\theta^k_m \mathrm{d}\theta^k_m &= [\Omega D_\theta \mathrm{d}\theta]^j_o \\
\sum_{mk} \Omega^{jk}_{om} \sum_{nh} \Gamma^{kh}_{mn}\theta^h_n \mathrm{d}\theta^k_m &= [\Omega D_{\Gamma\theta} \mathrm{d}\theta]^j_o \\
\sum_{mk} \Omega^{jk}_{om} \sum_{nh} \Gamma^{kh}_{mn}\theta^k_m \mathrm{d}\theta^h_n &= [\Omega D_\theta \Gamma \mathrm{d}\theta]^j_o \\
\sum_{mk} \Omega^{jk}_{om} \sum_{nh} \Gamma^{kh}_{mn}\theta^h_n \mathrm{d}\theta^h_n &= [\Omega \Gamma D_\theta \mathrm{d}\theta]^j_o.
\end{aligned}$$

Therefore,

$$\mathrm{d}\ln \bar{\boldsymbol{C}} = D_{\tilde{\boldsymbol{\gamma}}}\Omega[2\bar{t}\Lambda \mathrm{d}\boldsymbol{\theta} + \mathrm{d}\ln\tilde{\boldsymbol{\tau}}],$$

with Λ defined in the proposition.

For part 3, total differentiating $d\ln\bar{\chi}^{ij}_{do}$:

$$\begin{aligned}
\mathrm{d}\ln\bar{\chi}^{ij}_{do} = -\zeta\Big\{&\mathrm{d}\ln\tau^j_{do} + \mathrm{d}\ln\bar{C}^j_o + 2\bar{t}(\theta^i_d - \theta^j_o)\mathrm{d}(\theta^i_d - \theta^j_o) \\
&- \sum_{\tilde{o}} \bar{\chi}^{ij}_{d\tilde{o}}[\mathrm{d}\ln\tau^j_{d\tilde{o}} + \mathrm{d}\ln\bar{C}^j_{\tilde{o}} + 2\bar{t}(\theta^i_d - \theta^j_{\tilde{o}})\mathrm{d}(\theta^i_d - \theta^j_{\tilde{o}})]\Big\}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\sum_{j,o} \frac{\gamma^{ij}}{1-\gamma^{iL}}\bar{\chi}^{ij}_{do}\theta^j_o \mathrm{d}\ln\bar{\chi}^{ij}_{do} = -\zeta\Big\{&\sum_{j,o}\frac{\gamma^{ij}}{1-\gamma^{iL}}\bar{\chi}^{ij}_{do}\theta^j_o \mathrm{d}\ln\bar{C}^j_o - \sum_{j,\tilde{o}}\frac{\gamma^{ij}}{1-\gamma^{iL}}\bar{\chi}^{ij}_{d\tilde{o}}[\sum_o \bar{\chi}^{ij}_{do}\theta^j_o]\mathrm{d}\ln\bar{C}^j_{\tilde{o}} \\
&+ \sum_{j,o}\frac{\gamma^{ij}}{1-\gamma^{iL}}\bar{\chi}^{ij}_{do}\theta^j_o \mathrm{d}\ln\tau^j_{do} - \sum_{j,\tilde{o}}\frac{\gamma^{ij}}{1-\gamma^{iL}}\bar{\chi}^{ij}_{d\tilde{o}}[\sum_o \bar{\chi}^{ij}_{do}\theta^j_o]\mathrm{d}\ln\tau^j_{d\tilde{o}} \\
&+ \sum_{j,o}\frac{\gamma^{ij}}{1-\gamma^{iL}}\bar{\chi}^{ij}_{do}\theta^j_o \cdot 2\bar{t}(\theta^i_d - \theta^j_o)\mathrm{d}(\theta^i_d - \theta^j_o) \\
&- \sum_{j,\tilde{o}}\frac{\gamma^{ij}}{1-\gamma^{iL}}\bar{\chi}^{ij}_{d\tilde{o}}[\sum_o \bar{\chi}^{ij}_{do}\theta^j_o] \cdot 2\bar{t}(\theta^i_d - \theta^j_{\tilde{o}})\mathrm{d}(\theta^i_d - \theta^j_{\tilde{o}})\Big\} \\
= [-\zeta\widetilde{\Lambda}\mathrm{d}\ln\bar{\boldsymbol{C}} &- \zeta \mathrm{d}\ln\widehat{\boldsymbol{\tau}} - 2\zeta\bar{t}\widehat{\Lambda}\mathrm{d}\boldsymbol{\theta}]^i_d,
\end{aligned}$$

for $\widetilde{\Lambda}$, $\mathrm{d}\ln\widehat{\boldsymbol{\tau}}$, and $\widehat{\Lambda}$ defined in the proposition. Hence,

$$\begin{aligned}
\mathrm{d}\theta_d^i &= \omega^i \sum_{j,o} \frac{\gamma^{ij}}{1-\gamma^{iL}} \bar{\chi}_{do}^{ij} (\theta_o^j \mathrm{d}\ln\bar{\chi}_{do}^{ij} + \mathrm{d}\theta_o^j) \\
&= \omega^i [-\zeta\widetilde{\Lambda}\mathrm{d}\ln\bar{\boldsymbol{C}} - \zeta\mathrm{d}\ln\widehat{\boldsymbol{\tau}} - 2\zeta\bar{t}\widehat{\Lambda}\mathrm{d}\boldsymbol{\theta}]_d^i + \omega^i \sum_{j,o} \Gamma_{do}^{ij}\mathrm{d}\theta_o^j \\
&= -\zeta\omega^i \left[\widetilde{\Lambda}D_{\tilde{\boldsymbol{\gamma}}}\Omega[2\bar{t}\Lambda\mathrm{d}\boldsymbol{\theta} + \mathrm{d}\ln\tilde{\boldsymbol{\tau}}] + \mathrm{d}\ln\widehat{\boldsymbol{\tau}} + 2\bar{t}\widehat{\Lambda}\mathrm{d}\boldsymbol{\theta}\right]_d^i + \omega^i \sum_{j,o} \Gamma_{do}^{ij}\mathrm{d}\theta_o^j.
\end{aligned}$$

Therefore,

$$\mathrm{d}\boldsymbol{\theta} = -\zeta[I - D_\omega(\Gamma - 2\zeta\bar{t}\widetilde{\Lambda}D_{\tilde{\boldsymbol{\gamma}}}\Omega\Lambda - 2\zeta\bar{t}\widehat{\Lambda})]^{-1}\left[D_\omega\widetilde{\Lambda}D_{\tilde{\boldsymbol{\gamma}}}\Omega\mathrm{d}\ln\tilde{\boldsymbol{\tau}} + D_\omega\mathrm{d}\ln\widehat{\boldsymbol{\tau}}\right].$$

□

A.4 Proof of Proposition 4

Proof. Slightly abusing notation, we denote the change in the cost as $\mathrm{d}\ln\tau_{do}^{ij}$.

$$\begin{aligned}
\mathrm{d}\theta_d^i &= \omega^i \sum_{j',o'} \gamma^{ij'}\bar{\chi}_{do'}^{ij'}(\mathrm{d}\theta_o^j + \theta_o^j\mathrm{d}\ln\chi_{do}^{ij}) \\
&= \underbrace{\omega^i \sum_{j',o'} \gamma^{ij'}\bar{\chi}_{do'}^{ij'}\mathrm{d}\theta_o^j}_{=0 \text{ noting } \bar{\chi}_{dd}^{ii}=0} - \zeta\omega^i \sum_{j',o'} \gamma^{ij'}\bar{\chi}_{do'}^{ij'}\theta_{o'}^{j'}[\underbrace{\mathrm{d}\ln\bar{C}_{o'}^{j'}}_{=0 \text{ noting } \bar{\chi}_{dd}^{ii}=0} + \mathrm{d}\ln\tau_{do'}^{ij'} + 2\bar{t}(\theta_d^i - \theta_{o'}^{j'})\mathrm{d}\theta_d^i - \underbrace{2\bar{t}(\theta_d^i - \theta_{o'}^{j'})\mathrm{d}\theta_o^j}_{=0}] \\
&\quad + \zeta\omega^i \sum_{j',o'} \gamma^{ij'}\bar{\chi}_{do'}^{ij'}\theta_{o'}^{j'} \sum_m \chi_{dm}^{ij'}[\underbrace{\mathrm{d}\ln\bar{C}_m^{j'}}_{=0} + \mathrm{d}\ln\tau_{dm}^{ij'} + 2\bar{t}(\theta_d^i - \theta_m^{j'})\mathrm{d}\theta_d^i - \underbrace{2\bar{t}(\theta_d^i - \theta_m^{j'})\mathrm{d}\theta_m^{j'}}_{=0}] \\
&= -\zeta\omega^i\gamma^{ij}\bar{\chi}_{do}^{ij}\theta_o^j\mathrm{d}\ln\tau_{do}^{ij} - 2\zeta\bar{t}\omega^i[\sum_{j',o'} \gamma^{ij'}\bar{\chi}_{do'}^{ij'}\theta_{o'}^{j'}(\theta_d^i - \theta_{o'}^{j'})]\mathrm{d}\theta_d^i \\
&\quad + \zeta\omega^i\gamma^{ij}\bar{\chi}_{do}^{ij}(\sum_{o'} \bar{\chi}_{do'}^{ij}\theta_{o'}^j)\mathrm{d}\ln\tau_{do}^{ij} + 2\zeta\bar{t}\omega^i[\sum_{j',o'} \gamma^{ij'}\bar{\chi}_{do'}^{ij'}\theta_{o'}^{j'}(\theta_d^i - \sum_m \bar{\chi}_{dm}^{ij'}\theta_m^{j'})]\mathrm{d}\theta_d^i \\
&= -\zeta\omega^i\gamma^{ij}\bar{\chi}_{do}^{ij}[\theta_o^j - \sum_m \bar{\chi}_{dm}^{ij}\theta_m^j]\mathrm{d}\ln\tau_{do}^{ij} + \underbrace{2\zeta\bar{t}\omega^i \sum_{j',o'} \gamma^{ij'}\bar{\chi}_{do'}^{ij'}\theta_{o'}^{j'}[\theta_{o'}^{j'} - \sum_m \bar{\chi}_{dm}^{ij'}\theta_m^{j'}]\mathrm{d}\theta_d^i}_{\text{effect due to the change in } \theta_d^i \text{ through trade shares}}
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathrm{d}\theta_d^i &= \frac{-\zeta\omega^i\gamma^{ij}\bar{\chi}_{do}^{ij}[\theta_o^j - \sum_m \bar{\chi}_{dm}^{ij}\theta_m^j]\mathrm{d}\ln\tau_{do}^{ij}}{1 - 2\zeta\bar{t}\omega^i\sum_{j',o'}\gamma^{ij'}\bar{\chi}_{do'}^{ij'}\theta_{o'}^{j'}[\theta_{o'}^{j'} - \sum_m \bar{\chi}_{dm}^{ij'}\theta_m^{j'}]} \\
&= -\frac{\zeta\omega^i\gamma^{ij}\bar{\chi}_{do}^{ij}\|\theta_o^j - \vartheta_d^{ij}\|}{1 - 2\bar{t}\zeta\omega^i\sum_{j',o'}\gamma^{ij'}\bar{\chi}_{do'}^{ij'}\|\theta_{o'}^{j'} - \vartheta_d^{ij'}\|} \times \frac{\mathrm{d}\ln\tau_{do}^{ij}}{\theta_o^j - \vartheta_d^{ij}}.
\end{aligned}$$

The second equality holds because

$$
\begin{aligned}
&\sum_{j',o'} \gamma^{ij'} \chi^{ij'}_{do'} \sum_m \bar{\chi}^{ij'}_{dm} \theta^{j'}_m [\theta^{j'}_{o'} - \sum_m \bar{\chi}^{ij'}_{dm} \theta^{j'}_m] \\
&= \sum_{j'} \gamma^{ij'} \sum_{o'} \chi^{ij'}_{do'} \theta^{j'}_{o'} \sum_m \bar{\chi}^{ij'}_{dm} \theta^{j'}_m - \sum_{j'} \gamma^{ij'} \sum_{o'} \chi^{ij'}_{do'} \Big[\sum_m \bar{\chi}^{ij'}_{dm} \theta^{j'}_m\Big]^2 \\
&= 0.
\end{aligned}
$$

Recalling that $\Delta\|\theta^i_d - \theta^j_{o'}\| \approx \frac{1}{2}(\theta^i_d - \theta^j_{o'})(\mathrm{d}\theta^i_d - \mathrm{d}\theta^j_{o'})$, we have $\forall o', o$

$$
\begin{aligned}
\Delta\|\theta^i_d - \theta^j_o\| &= -\frac{\zeta \omega^i \gamma^{ij} \bar{\chi}^{ij}_{do} \|\theta^j_o - \vartheta^{ij}_d\|}{1 - 2\bar{t}\zeta\omega^i \sum_{j',o'} \gamma^{ij'} \bar{\chi}^{ij'}_{do'} \|\theta^{j'}_{o'} - \vartheta^{ij'}_d\|} \times \frac{\theta^i_d - \theta^j_o}{\theta^j_o - \vartheta^{ij}_d} \times x, \\
\Delta\|\theta^i_d - \theta^j_{o'}\| &= -\frac{\zeta \omega^i \gamma^{ij} \bar{\chi}^{ij}_{do} \|\theta^j_o - \vartheta^{ij}_d\|}{1 - 2\bar{t}\zeta\omega^i \sum_{j',o'} \gamma^{ij'} \bar{\chi}^{ij'}_{do'} \|\theta^{j'}_{o'} - \vartheta^{ij'}_d\|} \times \frac{\theta^i_d - \theta^j_{o'}}{\theta^j_o - \vartheta^{ij}_d} \times x.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&\Delta\|\theta^i_d - \theta^j_o\| - \sum_{o'} \bar{\chi}^{ij}_{do'} \Delta\|\theta^i_d - \theta^j_{o'}\| \\
&= -\frac{\zeta \omega^i \gamma^{ij} \bar{\chi}^{ij}_{do} \|\theta^j_o - \vartheta^{ij}_d\|}{1 - 2\bar{t}\zeta\omega^i \sum_{j',o'} \gamma^{ij'} \bar{\chi}^{ij'}_{do'} \|\theta^{j'}_{o'} - \vartheta^{ij'}_d\|} \times \frac{x}{\theta^j_o - \vartheta^{ij}_d} [(\theta^i_d - \theta^j_o) - \sum_{o'} \bar{\chi}^{ij}_{do'} (\theta^i_d - \theta^j_o)] \\
&= -\frac{\zeta \omega^i \gamma^{ij} \bar{\chi}^{ij}_{do} \|\theta^j_o - \vartheta^{ij}_d\|}{1 - 2\bar{t}\zeta\omega^i \sum_{j',o'} \gamma^{ij'} \bar{\chi}^{ij'}_{do'} \|\theta^{j'}_{o'} - \vartheta^{ij'}_d\|} \times \frac{x}{\theta^j_o - \vartheta^{ij}_d} \times (\vartheta^{ij}_d - \theta^j_o) \\
&= \frac{\zeta \omega^i \gamma^{ij} \bar{\chi}^{ij}_{do} \|\theta^j_o - \vartheta^{ij}_d\|}{1 - 2\bar{t}\zeta\omega^i \sum_{j',o'} \gamma^{ij'} \bar{\chi}^{ij'}_{do'} \|\theta^{j'}_{o'} - \vartheta^{ij'}_d\|} \times x,
\end{aligned}
$$

where the denominator is positive by the second-order condition of θ^i_d. □

A.5 Proof of Proposition 5

Proof. From the first-order condition, for

$$
\begin{aligned}
\theta^i_d(\nu) &= \frac{(\eta-1)\bar{t}}{(\eta-1)(1-\gamma^{iL})\bar{t} + \bar{\phi}} \sum_{j,o} \gamma^{ij} \bar{\chi}^{ij}_{do}(\nu) \theta^j_o + \frac{\bar{\phi}}{(\eta-1)(1-\gamma^{iL})\bar{t} + \bar{\phi}} \bar{\theta}(\nu) \\
&= \frac{(\eta-1)\bar{t}}{(\eta-1)(1-\gamma^{iL})\bar{t} + \bar{\phi}} \times \\
&\qquad \sum_{j,o} \gamma^{ij} \frac{\exp[-\zeta(\ln \tau^j_{do} + \ln \bar{C}^j_o + \bar{t}(\theta^i_d(\nu) - \theta^j_o)^2)]}{\sum_{o'} \exp[-\zeta(\ln \tau^j_{do'} + \ln \bar{C}^j_{o'} + \bar{t}(\theta^i_d(\nu) - \theta^j_{o'})^2)]} \cdot \theta^j_o + \frac{\bar{\phi}}{(\eta-1)(1-\gamma^{iL})\bar{t} + \bar{\phi}} \bar{\theta}(\nu).
\end{aligned}
$$

Totally differentiate θ and $\bar{\chi}^{ij}_{do}(\theta)$ w.r.t $\bar{\theta}(\nu)$ around θ^i_d. Since only one firm is deviating, all aggregate outcomes will not change.

$$\begin{aligned}
\mathrm{d}\theta^i_d(\nu) &= \frac{(\eta-1)\bar{t}}{(\eta-1)(1-\gamma^{iL})\bar{t}+\bar{\phi}}\sum_{j,o}\gamma^{ij}\theta^j_o\cdot\bar{\chi}^{ij}_{do}[\sum_m\bar{\chi}^{ij}_{dm}(\theta^j_o-\theta^j_m)]\cdot\mathrm{d}\theta^i_d(\nu)\\
&\quad+\frac{\bar{\phi}}{(\eta-1)(1-\gamma^{iL})\bar{t}+\bar{\phi}}\mathrm{d}\bar{\theta}(\nu)\\
\Rightarrow \mathrm{d}\theta^i_d(\nu) &= \Big(\frac{\frac{\bar{\phi}}{(\eta-1)(1-\gamma^{iL})\bar{t}+\bar{\phi}}}{1-\frac{(\eta-1)\bar{t}}{(\eta-1)(1-\gamma^{iL})\bar{t}+\bar{\phi}}\sum_{j,o}\gamma^{ij}\bar{\chi}^{ij}_{do}\|\theta^j_o-\vartheta^{ij}_d\|}\Big)\mathrm{d}\bar{\theta}(\nu),
\end{aligned}$$

where the denominator is positive by the second order condition of firms' optimal $\theta^i_d(\nu)$.

The rest of the proposition follows from the following expression:

$$\bar{\chi}^{ij}_{do}(\nu) = \frac{\exp[-\zeta(\ln\tau^j_{do}+\ln\bar{C}^j_o+\bar{t}(\theta^i_d(\nu)-\theta^j_o)^2)]}{\sum_{o'}\exp[-\zeta(\ln\tau^j_{do'}+\ln\bar{C}^j_{o'}+\bar{t}(\theta^i_d(\nu)-\theta^j_{o'})^2)]}.$$

□

A.6 Proof of Proposition 6

Proof. We suppress the location index for now. Normalizing wage to 1, household nominal income X in the decentralized equilibrium satisfies:

$$\begin{aligned}
X &= 1+\frac{1}{\eta}\sum_i\rho^i X\exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)\\
\Longrightarrow\quad X &= \frac{1}{1-\frac{1}{\eta}\sum_i\rho^i\exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}
\end{aligned}$$

And the household utility in the decentralized equilibrium is

$$U\propto\frac{X}{P},$$

where

$$\begin{aligned}
\ln(P) &\propto \sum_i\rho^i\ln(\bar{C}^i),\\
\ln(\bar{C}^i) &= \sum_j\gamma^{ij}\ln(\bar{C}^j)+\sum_j\bar{t}\gamma^{ij}(\theta^i-\theta^j)^2\\
\Rightarrow \ln(\bar{C}^i) &= \bar{t}\sum_m\Omega^{im}[\sum_j\gamma^{mj}(\theta^i-\theta^j)^2]\\
\Rightarrow \ln(P) &\propto \bar{t}\sum_i\rho^i\sum_m\Omega^{im}[\sum_j\gamma^{mj}(\theta^i-\theta^j)^2]
\end{aligned}$$

where Ω^{im} is the element of the matrix $(I-\Gamma)^{-1}$, which characterizes the GE influence of $\sum_j\gamma^{mj}(\theta^i-\theta^j)^2$ on $\ln(\bar{C}^i)$.

The first order condition of $\ln(U)$ w.r.t. θ^i reads,

$$\frac{\frac{1}{\eta}\rho^i \exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}{1-\frac{1}{\eta}\sum_i \rho^i \exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}\bar{\phi}(\bar{\theta}^i-\theta^i) - \bar{t}\rho^i \sum_m \Omega^{im}[\sum_j \gamma^{mj}(\theta^i-\theta^j)] - \bar{t}\sum_{j\neq i}\alpha_j \sum_m \omega^{jm}[\gamma^{mi}(\theta^i-\theta^j)]$$
$$= \frac{\frac{1}{\eta}\rho^i \exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}{1-\frac{1}{\eta}\sum_i \rho^i \exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}\bar{\phi}(\bar{\theta}^i-\theta^i) - \bar{t}\rho^i \sum_j (\theta^i-\theta^j)\underbrace{\sum_m \Omega^{im}\gamma^{mj}}_{\equiv\tilde{\gamma}^{ij}} - \bar{t}\sum_{j\neq i}\alpha_j(\theta^i-\theta^j)\underbrace{\sum_m \Omega^{jm}\gamma^{mi}}_{\equiv\tilde{\gamma}^{ji}}$$
$$= \frac{\frac{1}{\eta}\rho^i \exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}{1-\frac{1}{\eta}\sum_i \rho^i \exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}\bar{\phi}(\bar{\theta}^i-\theta^i) - \bar{t}\rho^i \sum_j \tilde{\gamma}^{ij}(\theta^i-\theta^j) - \bar{t}\sum_{j\neq i}\alpha_j\tilde{\gamma}^{ji}(\theta^i-\theta^j)$$
$$= \rho^i[\frac{\exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}{\eta-\sum_i \rho^i \exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}\bar{\phi}(\bar{\theta}^i-\theta^i) - \bar{t}\sum_j \tilde{\gamma}^{ij}(\theta^i-\theta^j)] - \bar{t}\sum_{j\neq i}\alpha_j\tilde{\gamma}^{ji}(\theta^i-\theta^j) \tag{A.10}$$

The decentralized θ^i satisfies

$$\frac{1}{\eta-1}\bar{\phi}(\bar{\theta}^i-\theta^i) = \bar{t}\sum_j \gamma^{ij}(\theta^i-\theta^j), \tag{A.11}$$

which implies that θ^i falls between $\bar{\theta}^i$ and $\frac{\sum\gamma^{ij}}{1-\gamma^{iL}}\theta^j$. WOLG, assume $\bar{\theta}^i < \theta^i < \frac{\sum\gamma^{ij}}{1-\gamma^{iL}}\theta^j$

Plugging this to equation (A.10) delivers

$$\rho^i\bar{t}[\frac{\exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}{\eta-\sum_i \rho^i \exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}(\eta-1)\sum_j \gamma^{ij}(\theta^i-\theta^j) - \sum_j \tilde{\gamma}^{ij}(\theta^i-\theta^j)] - \bar{t}\sum_{j\neq i}\alpha_j\tilde{\gamma}^{ji}(\theta^i-\theta^j).$$

Noting that $\frac{\exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}{\eta-\sum_i \rho^i \exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)} < \frac{1}{\eta-\sum_i \rho^i \exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)} < \frac{1}{\eta-1}$, under the assumption that $\bar{\theta}^i < \theta^i < \frac{\sum\gamma^{ij}}{1-\gamma^{iL}}\theta^j$, we have

$$\rho^i\bar{t}[\frac{\exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}{\eta-\sum_i \rho^i \exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}(\eta-1)\sum_j \gamma^{ij}(\theta^i-\theta^j) - \sum_j \tilde{\gamma}^{ij}(\theta^i-\theta^j)] - \bar{t}\sum_{j\neq i}\alpha_j\tilde{\gamma}^{ji}(\theta^i-\theta^j)$$
$$> \rho^i\bar{t}[\sum_{j\neq i}\gamma^{ij}(\theta^i-\theta^j) - \sum_{j\neq i}\tilde{\gamma}^{ij}(\theta^i-\theta^j)] - \bar{t}\sum_{j\neq i}\alpha_j\tilde{\gamma}^{ji}(\theta^i-\theta^j)$$

Since input-output coefficients are symmetric across all sectors,

$$\forall j\neq i,\ \frac{\gamma^{ij}}{\sum_{j\neq i}\gamma^{ij}} = \frac{1}{J-1} = \frac{\tilde{\gamma}^{ij}}{\sum_{j\neq i}\tilde{\gamma}^{ij}}$$
$$\sum_{j\neq i}\gamma^{ij}(\theta^i-\theta^j) = \sum_{j\neq i}\gamma^{ij}\theta^i - \sum_{j\neq i}\gamma^{ij}\theta^j = (\sum_{j\neq i}\gamma^{ij})(\theta^i - \sum_{j\neq i}\frac{\gamma^{ij}}{\sum_{j\neq i}\gamma^{ij}}\theta^j)$$
$$= \frac{(\sum_{j\neq i}\gamma^{ij})}{(\sum_{j\neq i}\tilde{\gamma}^{ij})}(\sum_{j\neq i}\tilde{\gamma}^{ij})(\theta^i - \sum_{j\neq i}\frac{\tilde{\gamma}^{ij}}{\sum_{j\neq i}\tilde{\gamma}^{ij}}\theta^j)$$
$$= \frac{\gamma^{ij}}{\tilde{\gamma}^{ij}}\sum_{j\neq i}\tilde{\gamma}^{ij}(\theta^i-\theta^j)$$
$$\Longrightarrow sign(\sum_{j\neq i}\gamma^{ij}(\theta^i-\theta^j)) = sign(\sum_{j\neq i}\tilde{\gamma}^{ij}(\theta^i-\theta^j)) \text{ and } |\sum_{j\neq i}\tilde{\gamma}^{ij}(\theta^i-\theta^j)| > |\sum_{j\neq i}\gamma^{ij}(\theta^i-\theta^j)|$$

From $\bar{\theta}^i < \theta^i < \frac{\sum \gamma^{ij}}{1-\gamma^{iL}}\theta^j$ and the symmetry in input-output coefficients, we have

$$sign(\sum_j \gamma^{ij}(\theta^i - \theta^j)) = sign(\sum_{j\neq i}(\theta^i - \theta^j)) = sign(\sum_{j\neq i}\alpha_j\tilde{\gamma}^{ji}(\theta^i - \theta^j)) < 0.$$

It follows that

$$\rho^i\bar{t}[\sum_{j\neq i}\gamma^{ij}(\theta^i - \theta^j) - \sum_{j\neq i}\tilde{\gamma}^{ij}(\theta^i - \theta^j)] - \bar{t}\sum_{j\neq i}\alpha_j\tilde{\gamma}^{ji}(\theta^i - \theta^j) > 0.$$

Thus, the marginal effect of increasing θ_i on the social welfare is positive. In the case of $\bar{\theta}^i > \theta^i > \frac{\sum \gamma^{ij}}{1-\gamma^{iL}}\theta^j$, we can prove that decreasing θ^i increases the social welfare analogously. □

A.7 Proof of Proposition 7

Proof. Suppose there are two symmetric countries, denominated by 1 and 2, and only one sector. We suppress industry indexes i and j. Denote $\tau_{12} = \tau_{21} = \tau$. Impose $\tau_{11} = \tau_{22} = 1$. WOLG, assume $\bar{\theta}_2 < 0 < \bar{\theta}_1$ and that $|\bar{\theta}_1| = |\bar{\theta}_2|$.

The symmetric setup implies that in the decentralized equilibrium, the two countries have the same nominal wage, which we normalize to 1. Moreover, $\theta_2 < 0 < \theta_1$, and $|\theta_2| = |\theta_1|$.

A marginal increase in θ_2 affects the economy through two channels. First, some of the net profit in country 2 is now expended as innovation cost, which affects the welfare of country 2; second, the distance between θ_2 and θ_1 decreases, which reduces the production cost in *both countries*. Note that the innovation expense and household consumption has the same composition of domestic versus imported goods, so if the wage and production cost in both countries remain the same, the demand for the goods produced in the two countries will be the same. Further notice that if the wages are the same, the reduction in production cost due to the decrease in distance between the two countries will be the same, which means symmetric wage also clear the market after the change. Therefore, throughout the subsequent analysis, we can normalize the wage of both countries to 1.

Below we first derive analytically household welfare under the decentralized equilibrium. We then show how it varies with a shift in the location choice of one of the countries.

For $i = 1, 2$,

$$Q_i = 1 + \frac{Q_i}{\eta}\exp(-\bar{\phi}(\theta^i - \bar{\theta}^i)^2)$$

$$\implies X = \frac{1}{1 - \frac{1}{\eta}\sum_i \rho^i \exp(-\bar{\phi}(\theta^i - \bar{\theta}^i)^2)}$$

The technology choice of firms in country 2 is

$$\theta_2 = \frac{(\eta-1)\bar{t}}{(\eta-1)(1-\gamma^{iL})\bar{t} + \bar{\phi}}(1-\gamma^L)[\bar{\chi}_{21}\theta_1 + (1-\bar{\chi}_{21})\theta_2] + \frac{\bar{\phi}}{(\eta-1)(1-\gamma^{iL})\bar{t} + \bar{\phi}}\bar{\theta}_2,$$

with analogous expression for country i and $\bar{\chi}_{21} = \bar{\chi}_{12}$ being defined by

$$\bar{\chi}_{21} = \bar{\chi}_{12} = \frac{[(\tau\exp(\bar{t}(\theta_2 - \theta_1)^2))^{-\zeta}]}{[(\tau\exp(\bar{t}(\theta_2 - \theta_1)^2))^{-\zeta} + 1]},$$

It follows from the symmetric assumption that $\bar{C}_d$ and P_d are also common across the two countries.

$$\bar{C}_1 = \bar{C}_2 \equiv \bar{C} \propto (w)^{\gamma^L} \cdot \bar{C}^{1-\gamma^L}[(\tau \exp(\bar{t}(\theta_2-\theta_1)^2))^{-\zeta}+1]^{-\frac{1-\gamma^L}{\zeta}}$$

$$\Rightarrow \bar{C} = [(\tau \exp(\bar{t}(\theta_2-\theta_1)^2))^{-\zeta}+1]^{-\frac{1-\gamma^L}{\gamma^L}\frac{1}{\zeta}} \text{ and } P \propto \bar{C}[1+\tau^{1-\eta}]^{\frac{1}{1-\eta}}.$$

The welfare of household in country i in the symmetric setup is

$$U_i = \frac{\frac{1}{1-\frac{1}{\eta}\exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}}{[(\tau\exp(\bar{t}(\theta_i-\theta_j)^2))^{-\zeta}+1]^{-\frac{1-\gamma^L}{\gamma^L}\frac{1}{\zeta}}\cdot(1+\tau^{1-\eta})^{\frac{1}{1-\eta}}}.$$

The welfare effect of a marginal increase in θ_2 on U_2 is

$$\begin{aligned}
\frac{\partial \ln U_2}{\partial \theta_2} &= \frac{\frac{1}{\eta}\exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}{1-\frac{1}{\eta}\exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)} 2\bar{\phi}(\bar{\theta}_2-\theta_2) + 2\bar{t}\frac{1-\gamma^L}{\gamma^L}\bar{\chi}_{12}(\theta_1-\theta_2) \\
&\quad (\text{noting that } \frac{\frac{1}{\eta}\exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}{1-\frac{1}{\eta}\exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)} < \frac{1}{\eta-1}) \\
&> \frac{1}{\eta-1}2\bar{\phi}(\bar{\theta}_2-\theta_2) + 2\bar{t}\frac{1-\gamma^L}{\gamma^L}\bar{\chi}_{12}(\theta_1-\theta_2) \\
&> \frac{1}{\eta-1}2\bar{\phi}(\bar{\theta}_2-\theta_2) + 2\bar{t}(1-\gamma^L)\bar{\chi}_{12}(\theta_1-\theta_2) \\
&= 0,
\end{aligned}$$

where the last inequality follows from in equilibrium $\frac{1}{\eta-1}\bar{\phi}(\theta_2-\bar{\theta}_2) = \bar{t}(1-\gamma^L)\bar{\chi}_{12}(\theta_1-\theta_2)$.

Also,

$$\frac{\partial \ln U_1}{\partial \theta_2} = 2\bar{t}\frac{1-\gamma^L}{\gamma^L}\bar{\chi}_{12}(\theta_1-\theta_2) > 0.$$

□

A.8 The Special Case with no Compatibility Incentive ($\bar{t} = 0$)

When $\bar{t} = 0$, the incentive for endogenous technological choice is eliminated, and our model becomes a version of the Caliendo and Parro (2015) model in which firms charge a fixed markup for selling to households.

Price Distribution. We first show that in this case, the factory-gate price of any firm (regardless of its technology location θ) in (d,i), denoted as p_d^i, follows a Weibull (inverse Fréchet) distribution with c.d.f.

$$F_d^i(p) = 1 - \exp(-[p/C_d^i]^\lambda), \tag{A.12}$$

with C_d^i determined as the fixed point of

$$C_d^i = \frac{\Xi_i}{A_d^i}[w_d]^{\gamma^i}\prod_j(\sum_o[\tau_{do}^j C_o^j]^{-\zeta})^{-\frac{\gamma^{ij}}{\zeta}}. \tag{A.13}$$

To see this, note that following Lemma A.2, by Assumption 1.2 and Assumption 1.3, for a firm ν in (d,i), from any sourcing country o, the number of suppliers ω with effective marginal cost $\tilde{c}^j(\nu,\omega)$ less or equal to any level $c > 0$ follows a Poisson distribution with mean

$$\int_0^\infty F_o^j(\frac{z \cdot c}{\tau_{do}^j})\zeta z^{-\zeta-1}\mathrm{d}z = \Gamma(1-\zeta/\lambda)\cdot(\frac{c}{\tau_{do}^j \cdot C_o^j})^\zeta \tag{A.14}$$

The probability that no such supplier arrives is

$$\Pr[\tilde{c}_{do}^j > c] = \exp[-\Gamma(1-\zeta/\lambda)\cdot(\frac{c}{\tau_{do}^j \cdot C_o^j})^\zeta]. \tag{A.15}$$

Taking the minimum over sourcing countries o, the distribution of input j in country d is characterized by

$$\Pr[c_d^j > c] = \exp[-\Gamma(1-\zeta/\lambda)\cdot\sum_o(\frac{c}{\tau_{do}^j \cdot C_o^j})^\zeta]. \tag{A.16}$$

Next, consider the distribution of the factory-gate price. Following Lemma A.2 and by Assumption 1, it can be shown that for any firm in (d,i), the number of techniques such that the factory-gate price is weakly less than p follows a Poisson distribution with mean

$$[\Xi^i]^{-\lambda}[A_d^i]^\lambda([w_d]^{-\gamma^i}\prod_j[\sum_o(\tau_{do}^j C_o^j)^{-\zeta}]^{\frac{\gamma^{ij}}{\zeta}})^\lambda p^\lambda,$$

where Ξ^i is the same sector-specific constant defined in (A.3). This implies that $F_d^i(p)$ satisfies

$$1 - F_d^i(p) = \exp(-[p/C_d^i]^\lambda). \tag{A.17}$$

Sourcing Strategies. Now consider the sourcing strategies of firms and final-good producers.

For any firm producing intermediate good in (d,i), the probability of sourcing input j from country o is

$$\chi_{do}^j \equiv \Pr[o = \arg\min_o \tilde{c}_{do}^j] = \frac{[\tau_{do}^j C_o^j]^{-\zeta}}{\sum_{o'}[\tau_{do'}^j C_{o'}^j]^{-\zeta}}. \tag{A.18}$$

For final-good producers in d, when sourcing input j, the expected total expenditure share on goods from country o is

$$\pi_{do}^j \equiv \frac{\mathbb{E}[\frac{\eta}{\eta-1}\tau_{do}^{Uj}p_o^j]^{1-\eta}}{(P_d^j)^{1-\eta}} = \frac{[\tau_{do}^{Uj}C_o^j]^{1-\eta}}{\sum_{o'}[\tau_{do'}^{Uj}C_{o'}^j]^{1-\eta}}, \tag{A.19}$$

where the sector-level price index is defined by

$$P_d^j \equiv (\sum_o \int_0^1 [\frac{\eta}{\eta-1}\tau_{do}^{Uj}p_{do}^j(\omega)]^{1-\eta}\mathrm{d}\omega)^{\frac{1}{1-\eta}} = (\sum_o \Gamma(1+\frac{1-\eta}{\lambda})\cdot[\frac{\eta}{\eta-1}\tau_{do}^{Uj}]^{1-\eta}\cdot(C_o^j)^{1-\eta})^{\frac{1}{1-\eta}}. \tag{A.20}$$

Market-Clearing Conditions. When market clears, sales to downstream firms should satisfy

$$M_o^j = \sum_d\sum_i \gamma^{ij}[M_d^i(\theta) + (1-\frac{1}{\eta})X_d^i]\cdot\chi_{do}^j, \tag{A.21}$$

and sales to final-good producers should satisfy

$$X_o^j \equiv \sum_d \rho_d^j P_d Q_d \cdot \pi_{do}^j. \tag{A.22}$$

Without gains from input compatibility, firms do not perform directed innovation. Therefore, final goods would only be consumed by household, with a market-clearing condition given by

$$P_d Q_d = I_d = w_d L_d + \Pi_i = w_d L_d + \sum_i \frac{1}{\eta} X_d^i. \tag{A.23}$$

Finally, labor market clearing condition requires

$$w_d L_d = \sum_i \gamma^{iL} [M_d^i + (1 - \frac{1}{\eta}) X_d^i]. \tag{A.24}$$

Equilibrium. The equilibrium is characterized by the following set of equations:

$$\begin{aligned}
\ln C_d^i &= \ln(\frac{\Xi^i}{A_d^i}) + \gamma^{iL} \ln(w_d) - \zeta^{-1} \sum_j \gamma^{ij} \ln(\sum_o \exp[-\zeta(\ln \tau_{do}^j + \ln C_o^j)]) \\
X_o^j &= \sum_d \sum_i \rho_d^j \pi_{do}^j [(\frac{1}{\eta} + \gamma^{iL}(1 - \frac{1}{\eta})) X_d^i + \gamma^{iL} M_d^i] \\
M_o^j &= \sum_d \sum_i \gamma^{ij} \chi_{do}^j [(1 - \frac{1}{\eta}) X_d^i + M_d^i] \\
w_d &= \frac{1}{L_d} \sum_i \gamma^{iL} [M_d^i + (1 - \frac{1}{\eta}) X_d^i] \\
\chi_{do}^j &= \frac{[\tau_{do}^j C_o^j]^{-\zeta}}{\sum_{o'} [\tau_{do'}^j C_{o'}^j]^{-\zeta}} = \frac{\exp[-\zeta(\ln \tau_{do}^j + \ln C_o^j)]}{\sum_{o'} \exp[-\zeta(\ln \tau_{do'}^j + \ln C_{o'}^j)]} \\
\pi_{do}^j &= \frac{[\tau_{do}^{Uj} C_o^j]^{1-\eta}}{\sum_{o'} [\tau_{do'}^{Uj} C_{o'}^j]^{1-\eta}} = \frac{\exp[(1-\eta)(\ln \tau_{do}^{Uj} + \ln C_o^j)]}{\sum_{o'} \exp[(1-\eta)(\ln \tau_{do'}^{Uj} + \ln C_{o'}^j)]}
\end{aligned}$$

Appendix B Reduced-Form Evidence

In this section, we first explain how we aggregate countries around the world into 29 geo-political regions to facilitate our empirical and quantitative analyses. We then discuss in detail on data sources, dataset construction, and cleaning procedures. Finally, we present additional robustness checks to our reduced-form evidence.

B.1 Geo-Political Regions

We aggregate countries around the world into 29 regions by further classifying countries in the World Input-Output Tables (WIOTs) based on their geographic and political proximity. Table B.1 lists the regions. In our reduced-form analysis, we exclude the region ROW due to systematic missing values of tariff data.

B.2 Dataset Construction

Our two datasets consist of data from three main sources, patent citation data from the PATSTAT Global, tariff data from the TRAINS database, and Chinese firm-level data. We explain in detail where we source the data and how we construct the final datasets.

Citation Data. We rely on patent citation data sourced from the PATSTAT Global, a comprehensive database comprising bibliographical information on over 100 million patent documents. This database encompasses patent records from 90 patent-issuing authorities, including all major national (e.g. USPTO), regional (e.g. EPO), and global (e.g. the Patent Cooperative Treaty) patent offices. To focus on production technologies, we narrow our focus to invention patents and utility models while excluding design patents from our analysis.

Throughout our analysis, we define a patent as a *patent family*, which typically protects an invention, which can be a new product, a new process to produce, or a new technical solution. A patent family can involve multiple patent *applications* when this technology seeks patent protection from different authorities. For each patent, we denote its *year of invention* as the year in which the first application is filed.

Each patent record contains details about the inventors of the patent (the individuals who invent the patent, though not necessarily the applicant or the owner of the patent) and their countries of residence. We first map the countries into the geo-political regions and take all these regions as the *regions of invention* of that patent, assigning weights to each region based on the number of inventors from that specific location. For instance, if a patent involves five inventors—two from China and three from the US—we consider that China holds 2/5 of the patent, while the US holds 3/5. In case where the inventor information is missing, we designate the region of its first application as the region of invention. This assumption is grounded in the idea that a patent would typically be filed domestically before seeking international protection.

We categorize each patent into specific industries using a similar weighting approach. The database provides the International Patent Classification (IPC) symbols for each patent, extending up to an 8-digit subgroup level. Given that a patent's technology may span multiple fields, it can be associated with multiple IPC symbols. We consolidate the 8-digit symbols into a 4-digit subclass level and assign weights based on the frequency of the corresponding 8-digit symbols. We then map the 4-digit IPC symbols to 4-digit ISIC (Rev. 3) industries us-

Region Code	Region	ISO3 in WIOTs
AUS	Australia	AUS
AUT	Austria	AUT
BLK	Balkans	BGR, HRV, GRC
BLT	Baltic States	EST, LVA, LTU
BNE	Benelux	BEL, LUX, NLD
BRA	Brazil	BRA
CAN	Canada	CAN
CHE	Switzerland	CHE
CHN	China (Mainland)	CHN
CNE	Central Europe	CZE, HUN, POL, SVK, SVN
DEU	Germany	DEU
ESP	Spain	ESP
FRA	France	FRA
GBR	United Kingdom	GBR
IDN	Indonesia	IDN
IND	India	IND
IRL	Ireland	IRL
ITA	Italy	ITA
JPN	Japan	JPN
KOR	South Korea	KOR
MEX	Mexico	MEX
NRD	Nordic Countries	DNK, FIN, NOR, SWE
PRT	Portugal	PRT
ROU	Romania	ROU
ROW	Rest of the World	CYP, MLT, ROW
RUS	Russia	RUS
TUR	Turkey	TUR
TWN	Taiwan	TWN
USA	United States	USA

Table B.1: Geo-Political Regions

ing the crosswalk provided by Lybbert and Zolas (2014), which provides the mapping from each IPC symbol to a set of industries with specific weights.

The mapping of each patent to a set of region-industry pairs allows us to convert each family-to-family citation in the data into a number of citation flows between region-industry pairs. Specifically, suppose the citing patent is invented in period t and is mapped to region d with weight $v_{1,d}$ and industry i with weight v_1^i, and the cited patent is mapped to region o with weight $v_{2,o}$ and industry j with weight v_2^j. Then, we convert this citation record into citation flows from (o, j) to (d, i) in period t with intensity $v_{do,t}^{ij} = v_{1,d} \times v_1^i \times v_{2,o} \times v_2^j$.

Our analyses involve different levels of aggregation for the citation data. In the reduced-form analysis in Section 3.1, we aggregate all the citation flows between 2000-2014 into a balanced panel at the destination(d)-origin(o)-industry(j)-period(t) level. In the subsequent reduced-form analysis in Section 3.2, we establish links between each Chinese patent and the citing patent using their unique patent application numbers. We then aggregate these citations into a firm (ω)-origin (o)-period (t) level. For the quantification phase in Section 4, we aggregate the citations between 2010-2014 to construct a cross-sectional citation matrix between (d, i) and (o, j) pairs.

Tariff Data. We source tariff data from UN TRAINS, which are downloaded from `https://wits.worldbank.org` for each year between 2000-2014. The raw data include the effectively applied tariff rates and MFN tariff rates at the importer-exporter-industry level, where importers and exporters are in three-letter ISO country codes, and industries are at the level of 4-digit ISIC (Rev. 3). We use the reported tariff rates in simple averages and restrict to manufacturing industries.

When cleaning the data, we drop all the observations where either the importer or the exporter is unspecified. For both the applied and MFN tariffs, when the observations are missing, we impute them with the first non-missing preceding tariff. If no earlier observation is available, we leave them missing and drop these observations in our regressions.

To facilitate analysis, we aggregate the tariff rates from country pairs into region pairs and from yearly figures into three-year periods by computing simple averages, while maintaining the granularity of industry (j) at the 4-Digit ISIC (Rev. 3) level. In case where observations are missing at the initial country-year level, we preserve these as missing entries at the corresponding region-period level, and they are not used in our regression analyses.

In one of the specifications in Section 3.2, we control for the import tariff faced by an industry i in China when *sourcing inputs* from abroad. We construct this variable by weighting the effectively applied tariffs on industry j, denoted as $\tau_{CHN,ojt}$, with the share of industry j as inputs to industry i. This share is derived from China's 2002 Input-Output Tables.

Chinese Firm-Level Data. Our firm-level analysis focuses on Chinese manufacturing firms in the Annual Survey of Industrial Enterprise maintained by the National Bureau of Statistics of China (NBSC). The dataset offers detailed accounting information for all Chinese manufacturing firms with annual sales greater than US$800,000 over 1998–2014, including plant-level information on industry, location, sales, employment, etc.

Crucial to our analysis is a firm's identity (ω) and the prime industry (i) it belongs to. To link each firm consistently over time, we employ a procedure following Brandt et al. (2017) to create a unique identifier. The algorithm establishes firm linkages over time using information on the NBS ID, firm name, the name of legal person representative, phone number, address, name of main products, founding year, etc. We manually map the industry codes to the 2017 version to accommodate the constant changes in the Chinese Industry Classification (CIC) codes during the specified period.

We link the NBSC Database to patent data provided by China's State Intellectual Property Office (SIPO). For each patent filed by the Chinese firm, we establish link with the citation data from the PATSTAT Global by matching the unique patent application number with that of the citing patent.

Finally, we obtain information on firms' imports from China's General Administration of Customs, which provides detailed records on the universe of all Chinese trade transactions by both importing and exporting firms at the HS eight-digit level for the years 2000-2014. We follow the matching process in Fan et al. (2015) to merge the import data with the NBSC manufacturing firm survey data. The matching procedure consists of three main steps: (1) match by company names (in Chinese); (2) match by phone number and zip code; (3) match by phone number and the name of contact person.

Our final dataset includes all manufacturing firms with patents, regardless of whether they import goods from abroad. The panel is unbalanced with the number of firms increasing from 32,293 in the period 2000-2002 to 95,813 in the period 2012-2014.

Appendix C Quantification

C.1 Proof of Proposition 8

We prove the proposition by guess and verification. Suppose that the distribution of production cost is characterized by (26), and the ex-post technology distribution is characterized by (28). Then, under Assumption 3, by taking the log of (3), we have

$$\ln C_d^i(\theta) = \ln(\frac{\Xi^i}{A_d^i}) + \gamma^{iL}\ln(w_d) - \zeta^{-1}\sum_j \gamma^{ij}\ln(\sum_o \int [\tau_{do}^j]^{-\zeta}\cdot$$
$$\exp[-\zeta(k_{A,o}^j + m_A^j(\tilde{\theta}-n_{A,o}^j)^2) - \zeta\bar{t}(\theta-\tilde{\theta})^2]\cdot\frac{1}{\sqrt{2\pi(\sigma^j)^2}}\exp[-\frac{1}{2}(\frac{\tilde{\theta}-\mu_o^j}{\sigma^j})^2]\mathrm{d}\tilde{\theta})$$

where

$$\begin{aligned}
& k_{A,o}^j + m_A^j(\tilde{\theta}-n_{A,o}^j)^2 + \bar{t}(\theta-\tilde{\theta})^2 \\
={}& k_{A,o}^j + m_A^j\tilde{\theta}^2 - 2m_A^j n_{A,o}^j\tilde{\theta} + m_A^j(n_{A,0}^j)^2 + \bar{t}\theta^2 - 2\bar{t}\theta\tilde{\theta} + \bar{t}\tilde{\theta}^2 \\
={}& [m_A^j + \bar{t}]\tilde{\theta}^2 - 2(m_A^j n_{A,o}^j + \bar{t}\theta)\tilde{\theta} + k_{A,o}^j + m_A^j(n_{A,o}^j)^2 + \bar{t}\theta^2 \\
={}& [m_A^j + \bar{t}](\tilde{\theta}^2 - 2\frac{m_A^j n_{A,o}^j + \bar{t}\theta}{m_A^j + \bar{t}}\tilde{\theta}) + k_{A,o}^j + m_A^j(n_{A,o}^j)^2 + \bar{t}\theta^2 \\
={}& [m_A^j + \bar{t}](\tilde{\theta} - \frac{m_A^j n_{A,o}^j + \bar{t}\theta}{m_A^j + \bar{t}})^2 + k_{A,o}^j + \frac{m_A^j\bar{t}(\theta - n_{A,o}^j)^2}{m_A^j + \bar{t}}.
\end{aligned}$$

Lemma A.8. *Suppose $X \sim N(\mu, \sigma^2)$. Then for $m < \frac{1}{2\sigma^2}$,*

$$\mathbb{E}[\exp(mX^2)] = \exp(\frac{m\mu^2}{1-2m\sigma^2})(1-2m\sigma^2)^{-1/2}.$$

Proof.

$$\begin{aligned}
& \mathbb{E}\exp(mX^2) \\
={}& \int \exp(mx^2)\frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{1}{2\sigma^2}(x-\mu)^2)\mathrm{d}x \\
={}& \int \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{1}{2\sigma^2}[(1-2m\sigma^2)x^2 - 2\mu x + \mu^2])\mathrm{d}x \\
={}& \int \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{1-2m\sigma^2}{2\sigma^2}[x^2 - \frac{2}{1-2m\sigma^2}\mu x] - \frac{\mu^2}{2\sigma^2})\mathrm{d}x \\
={}& \int \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{1-2m\sigma^2}{2\sigma^2}[x^2 - \frac{2}{1-2m\sigma^2}\mu x + (\frac{1}{1-2m\sigma^2})^2\mu^2] + \frac{1}{2\sigma^2(1-2m\sigma^2)}\mu^2 - \frac{\mu^2}{2\sigma^2})\mathrm{d}x \\
={}& \int \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{1-2m\sigma^2}{2\sigma^2}[x - \frac{1}{1-2m\sigma^2}\mu]^2 + \frac{m}{1-2m\sigma^2}\mu^2)\mathrm{d}x \\
={}& \exp(\frac{m\mu^2}{1-2m\sigma^2})(1-2m\sigma^2)^{-1/2}\int \frac{1}{\sqrt{2\pi\sigma^2(1-2m\sigma^2)^{-1}}}\exp(-\frac{1}{2\sigma^2(1-2m\sigma^2)^{-1}}[x - \frac{1}{1-2m\sigma^2}\mu]^2)\mathrm{d}x \\
={}& \exp(\frac{m\mu^2}{1-2m\sigma^2})(1-2m\sigma^2)^{-1/2},
\end{aligned}$$

where the last line applies that the latter integral is a density. □

Since $\tilde{\theta} \sim N(\mu_o^j, [\sigma^j]^2), \tilde{\theta} - \frac{m_A^j n_{A,o}^j + \bar{t}\theta}{m_A^j + \bar{t}} \sim N(\mu_o^j - \frac{m_A^j n_{A,o}^j + \bar{t}\theta}{m_A^j + \bar{t}}, [\sigma^j]^2)$. Apply Lemma A.8 and we have

$$\begin{aligned}
&\int \exp[-\zeta(k_{A,o}^j + m_A^j(\tilde{\theta} + n_{A,o}^j)^2 + \bar{t}(\theta - \tilde{\theta})^2)] \cdot \frac{1}{\sqrt{2\pi(\sigma^j)^2}} \exp[-\frac{1}{2}(\frac{\tilde{\theta} - \mu_o^j}{\sigma^j})^2]\mathrm{d}\tilde{\theta} \\
&= \int \exp[-\zeta((m_A^j + \bar{t})(\tilde{\theta} - \frac{m_A^j n_{A,o}^j + \bar{t}\theta}{m_A^j + \bar{t}})^2 + k_{A,o}^j + \frac{m_A^j \bar{t}(\theta - n_{A,o}^j)^2}{m_A^j + \bar{t}})] \cdot \frac{1}{\sqrt{2\pi(\sigma^j)^2}} \exp[-\frac{1}{2}(\frac{\tilde{\theta} - \mu_o^j}{\sigma^j})^2]\mathrm{d}\tilde{\theta} \\
&= \exp[-\zeta(k_{A,o}^j + \frac{m_A^j \bar{t}(\theta - n_{A,o}^j)^2}{m_A^j + \bar{t}})] \cdot \mathbb{E} \exp[-\zeta(m_A^j + \bar{t})(\tilde{\theta} - \frac{m_A^j n_{A,o}^j + \bar{t}\theta}{m_A^j + \bar{t}})^2] \\
&= \exp[-\zeta(k_{A,o}^j + \frac{m_A^j \bar{t}(\theta - n_{A,o}^j)^2}{m_A^j + \bar{t}})] \cdot [1 + 2\zeta(m_A^j + \bar{t})(\sigma^j)^2]^{-1/2} \cdot \exp(\frac{-\zeta(m_A^j + \bar{t})}{1 + 2\zeta(m_A^j + \bar{t})(\sigma^j)^2}[\mu_o^j - \frac{m_A^j n_{A,o}^j + \bar{t}\theta}{m_A^j + \bar{t}}]^2) \\
&= \exp[-\zeta(k_{A,o}^j + \frac{m_A^j \bar{t}(\theta - n_{A,o}^j)^2}{m_A^j + \bar{t}})] \cdot [1 + 2\zeta(m_A^j + \bar{t})(\sigma^j)^2]^{-1/2} \cdot \exp(\frac{-\zeta(m_A^j + \bar{t})}{1 + 2\zeta(m_A^j + \bar{t})(\sigma^j)^2}[\frac{\bar{t}}{m_A^j + \bar{t}}\theta - (\mu_o^j - \frac{m_A^j n_{A,o}^j}{m_A^j + \bar{t}})]^2) \\
&= \exp(-\zeta k_{A,o}^j) \cdot [1 + 2\zeta(m_A^j + \bar{t})(\sigma^j)^2]^{-1/2} \\
&\quad \cdot \exp[-\zeta(\frac{m_A^j \bar{t}}{m_A^j + \bar{t}}(\theta - n_{A,o}^j)^2 + \frac{\bar{t}^2}{[1 + 2\zeta(m_A^j + \bar{t})(\sigma^j)^2](m_A^j + \bar{t})}[\theta - (\frac{m_A^j + \bar{t}}{\bar{t}}\mu_o^j - \frac{m_A^j n_{A,o}^j}{\bar{t}})]^2)] \\
&= \exp(-\zeta k_{A,o}^j) \cdot [1 + 2\zeta(m_A^j + \bar{t})(\sigma^j)^2]^{-1/2} \\
&\quad \cdot \exp[-\zeta(\frac{\bar{t}[1 + 2\zeta m_A^j(\sigma^j)^2]}{1 + 2\zeta(m_A^j + \bar{t})(\sigma^j)^2}\theta^2 - \frac{2\bar{t}[\mu_o^j + 2\zeta m_A^j n_{A,o}^j(\sigma^j)^2]}{1 + 2\zeta(m_A^j + \bar{t})(\sigma^j)^2}\theta + \frac{m_A^j \bar{t}(n_{A,o}^j)^2}{m_A^j + \bar{t}} + \frac{[(m_A^j + \bar{t})\mu_o^j - m_A^j n_{A,o}^j]^2}{[1 + 2\zeta(m_A^j + \bar{t})(\sigma^j)^2](m_A^j + \bar{t})})] \\
&= \exp(-\zeta k_{A,o}^j) \cdot [1 + 2\zeta(m_A^j + \bar{t})(\sigma^j)^2]^{-1/2} \cdot \exp(\frac{-\zeta\bar{t}[1 + 2m_A^j(\sigma^j)^2]}{1 + 2\zeta(m_A^j + \bar{t})(\sigma^j)^2}[\theta - \frac{\mu_o^j + 2\zeta m_A^j n_{A,o}^j(\sigma^j)^2}{1 + 2\zeta m_A^j(\sigma^j)^2}]^2) \\
&\quad \cdot \exp[-\zeta(\frac{\bar{t}[\mu_o^j + 2\zeta m_A^j n_{A,o}^j(\sigma^j)^2]^2}{[1 + 2\zeta(m_A^j + \bar{t})(\sigma^j)^2][1 + 2\zeta m_A^j(\sigma^j)^2]} + \frac{m_A^j \bar{t}(n_{A,o}^j)^2}{m_A^j + \bar{t}} + \frac{[(m_A^j + \bar{t})\mu_o^j - m_A^j n_{A,o}^j]^2}{[1 + 2\zeta(m_A^j + \bar{t})(\sigma^j)^2](m_A^j + \bar{t})})] \\
&= \exp(-\zeta k_{A,o}^j) \cdot [1 + 2\zeta(m_A^j + \bar{t})(\sigma^j)^2]^{-1/2} \cdot \exp(\frac{-\zeta m_A^j[(\mu_o^j)^2 - 2\mu_o^j n_{A,o}^j + (n_{A,o}^j)^2]}{1 + 2\zeta m_A^j(\sigma^j)^2}) \\
&\quad \cdot \exp(\frac{-\zeta\bar{t}[1 + 2\zeta m_A^j(\sigma^j)^2]}{1 + 2\zeta(m_A^j + \bar{t})(\sigma^j)^2}[\theta - \frac{\mu_o^j + 2\zeta m_A^j n_{A,o}^j(\sigma^j)^2}{1 + 2\zeta m_A^j(\sigma^j)^2}]^2) \\
&= \exp(-\zeta k_{A,o}^j - \frac{1}{2}\log[1 + 2\zeta(m_A^j + \bar{t})(\sigma^j)^2] - \frac{\zeta m_A^j(\mu_o^j - n_{A,o}^j)^2}{1 + 2\zeta m_A^j(\sigma^j)^2} - \frac{\zeta\bar{t}[1 + 2\zeta m_A^j(\sigma^j)^2]}{1 + 2\zeta(m_A^j + \bar{t})(\sigma^j)^2}[\theta - \frac{\mu_o^j + 2\zeta m_A^j n_{A,o}^j(\sigma^j)^2}{1 + 2\zeta m_A^j(\sigma^j)^2}]^2)
\end{aligned}$$

Therefore,

$$\begin{aligned}
\ln C_d^i(\theta) &= \ln(\frac{\Xi^i}{A_d^i}) + \gamma^{iL}\ln(w_d) - \zeta^{-1}\sum_j \gamma^{ij}\ln(\sum_o [\tau_{do}^j]^{-\zeta}\exp[-\zeta(k_{B,o}^j + m_B^j(\theta - n_{B,o}^j)^2)]), \\
&= \ln(\frac{\Xi^i}{A_d^i}) + \gamma^{iL}\ln(w_d) - \zeta^{-1}\sum_j \gamma^{ij}\ln(\sum_o \exp[-\zeta(\ln\tau_{do}^j + k_{B,o}^j + m_B^j(\theta - n_{B,o}^j)^2)]),
\end{aligned} \tag{C.1}$$

where

$$\begin{aligned}
k^j_{B,o} &= k^j_{A,o} + \frac{1}{2\zeta}\log[1+2\zeta(m^j_A+\bar{t})(\sigma^j)^2] + \frac{m^j_A(\mu^j_o - n^j_{A,o})^2}{1+2\zeta m^j_A(\sigma^j)^2},\\
m^j_B &= \frac{\bar{t}[1+2\zeta m^j_A(\sigma^j)^2]}{1+2\zeta(m^j_A+\bar{t})(\sigma^j)^2},\\
n^j_{B,o} &= \frac{\mu^j_o + 2\zeta m^j_A(\sigma^j)^2 n^j_{A,o}}{1+2\zeta m^j_A(\sigma^j)^2}.
\end{aligned}$$

Consider

$$\begin{aligned}
\frac{\mathrm{d}\ln C^i_d(\theta)}{\mathrm{d}\theta} &= \sum_j \gamma^{ij}\sum_o \chi^j_{do}(\theta)2m^j_B[\theta - n^j_{B,o}]\\
\frac{\mathrm{d}^2\ln C^i_d(\theta)}{\mathrm{d}\theta^2} &= \sum_j \gamma^{ij}\sum_o([\chi^j_{do}]'(\theta)2m^j_B[\theta - n^j_{B,o}] + \chi^j_{do}(\theta)2m^j_B),
\end{aligned}$$

where

$$\chi^j_{do}(\theta) \equiv \frac{\exp[-\zeta(\ln\tau^j_{do} + k^j_{B,o} + m^j_B(\theta - n^j_{B,o})^2)]}{\sum_{o'}\exp[-\zeta(\ln\tau^j_{do'} + k^j_{B,o'} + m^j_B(\theta - n^j_{B,o'})^2)]},$$

with $\sum_o \chi^j_{do}(\theta) = 1$.

We consider a second-order approximation with respect to θ around a fixed $\hat{\theta}^i_d$ by ignoring the terms with $[\chi^j_{do}]'(\theta)$, which is

$$\begin{aligned}
&\ln C^i_d(\theta)\\
&\approx \ln C^i_d(\hat{\theta}^i_d) + \sum_j\gamma^{ij}\sum_o \chi^j_{do}(\hat{\theta}^i_d)2m^j_B(\hat{\theta}^i_d - n^j_{B,o})(\theta-\hat{\theta}^i_d) + \frac{1}{2}\sum_j\gamma^{ij}\sum_o\chi^j_{do}(\hat{\theta}^i_d)2m^j_B(\theta-\hat{\theta}^i_d)^2\\
&= (\sum_j\gamma^{ij}m^j_B)[(\theta-\hat{\theta}^i_d)^2 + 2\sum_{j,o}\frac{\gamma^{ij}m^j_B\chi^j_{do}(\hat{\theta}^i_d)}{\sum_{j'}\gamma^{ij'}m^{j'}_B}(\hat{\theta}^i_d - n^j_{B,o})(\theta-\hat{\theta}^i_d)] + \ln C^i_d(\hat{\theta}^i_d)\\
&= (\sum_j\gamma^{ij}m^j_B)[\theta-\hat{\theta}^i_d + \sum_{j,o}\hat{\chi}^{ij}_{do}(\hat{\theta}^i_d - n^j_{B,o})]^2 - (\sum_j\gamma^{ij}m^j_B)[\sum_{j,o}\hat{\chi}^{ij}_{do}(\hat{\theta}^i_d - n^j_{B,o})]^2 + \ln C^i_d(\hat{\theta}^i_d)\\
&= (\sum_j\gamma^{ij}m^j_B)[\theta - \sum_{j,o}\hat{\chi}^{ij}_{do}n^j_{B,o}]^2 - (\sum_j\gamma^{ij}m^j_B)[\sum_{j,o}\hat{\chi}^{ij}_{do}(\hat{\theta}^i_d - n^j_{B,o})]^2 + \ln C^i_d(\hat{\theta}^i_d),
\end{aligned}$$

where

$$\hat{\chi}^{ij}_{do} \equiv \frac{\gamma^{ij}m^j_B\chi^j_{do}(\hat{\theta}^i_d)}{\sum_{j'}\gamma^{ij'}m^{j'}_B},$$

with $\sum_{o,j}\hat{\chi}^{ij}_{do} = 1$.

This verifies the functional form in (26) with m^i_A, $n^i_{A,d}$ and $k^i_{A,d}$ being

$$
\begin{aligned}
m^i_A &= \sum_j \gamma^{ij} m^j_B = \sum_j \gamma^{ij} \frac{\bar{t}[1+2\zeta m^j_A(\sigma^j)^2]}{1+2\zeta(m^j_A+\bar{t})(\sigma^j)^2}, \\
n^i_{A,d} &= \sum_{j,o} \hat{\chi}^{ij}_{do} n^j_{B,o} = \sum_{j,o} \hat{\chi}^{ij}_{do} \frac{\mu^j_o + 2\zeta m^j_A(\sigma^j)^2 n^j_{A,o}}{1+2\zeta m^j_A(\sigma^j)^2}, \\
k^i_{A,d} &= \ln C^i_d(\hat{\theta}^i_d) - m^i_A[\sum_{j,o} \hat{\chi}^{ij}_{do}(\hat{\theta}^i_d - n^j_{B,o})]^2.
\end{aligned}
$$

Consider the innovation decision in (o, j) given by (13). Under Assumption 3, this is equivalent to

$$
\max_\theta \exp[-\bar{\phi}(\bar{\theta}-\theta)^2] \cdot [C^j_o(\theta)]^{1-\eta}.
$$

Taking log and apply the quadratic approximation, this is

$$
\max_\theta (1-\eta) m^j_A(\theta - n^j_{A,o})^2 - \bar{\phi}(\bar{\theta}-\theta)^2.
$$

The first-order condition implies that

$$
(1-\eta) m^j_A(\theta - n^j_{A,o}) = \bar{\phi}(\bar{\theta}-\theta),
$$

which gives the policy function (27)

$$
\theta = g^j_o(\bar{\theta}) \equiv \alpha^j_o + \beta^j \bar{\theta},
$$

where

$$
\begin{aligned}
\alpha^j_o &= \frac{(\eta-1)m^j_A}{\bar{\phi}+(\eta-1)m^j_A} n^j_{A,o}, \\
\beta^j &= \frac{\bar{\phi}}{\bar{\phi}+(\eta-1)m^j_A}.
\end{aligned}
$$

Since the ex-ante technology distribution $\bar{\Theta}^j_o$ is Normal with mean $\bar{\mu}^j_o$ and variance $(\bar{\sigma}^j)^2$, the ex-post technology distribution Θ^j_o is also Normal with mean μ^j_o and variance $(\sigma^j)^2$, where

$$
\mu^j_o = \alpha^j_o + \beta^j \bar{\mu}^j_o \quad \text{and} \quad \sigma^j = \beta^j \bar{\sigma}^j.
$$

This completes the proof of Proposition 8.

C.2 Algorithm to Solve the Equilibrium

Building on Proposition 8, we develop the following algorithm to solve the model.

Step 1. Given wages $\{w_d\}$ and parameters on geography $\{\tau^j_{do}\}$, preference η, production technology $\{\gamma^{ij}, \gamma^{iL}, \Xi^i, A^i_d, \zeta, \bar{t}, \bar{\phi}\}$, and the ex-ante technology distribution $\{\bar{\mu}^i_d, \bar{\sigma}^i\}$, we solve for $\{k^i_{A,d}, m^i_A, n^i_{A,d}\}$ and $\{\mu^i_d, \sigma^i\}$ to obtain the cost functions $\{C^j_o(\cdot)\}$ and the ex-post technology distributions. This involves simultaneously solving the following system of equations:

$$m^i_A = \sum_j \gamma^{ij} m^j_B \tag{C.2}$$

$$n^i_{A,d} = \sum_{j,o} \hat{\chi}^{ij}_{do} n^j_{B,o} \tag{C.3}$$

$$k^i_{A,d} = \ln C^i_d(\mu^i_d) - m^i_A [\sum_{j,o} \hat{\chi}^{ij}_{do} (\mu^i_d - n^j_{B,o})]^2 \tag{C.4}$$

$$\mu^i_d = \alpha^i_d + \beta^i \bar{\mu}^i_d \tag{C.5}$$

$$\sigma^i = \beta^i \bar{\sigma}^i \tag{C.6}$$

where

$$m^j_B = \frac{\bar{t}[1 + 2\zeta m^j_A (\sigma^j)^2]}{1 + 2\zeta (m^j_A + \bar{t})(\sigma^j)^2}, \tag{C.7}$$

$$n^j_{B,o} = \frac{1}{1 + 2\zeta m^j_A (\sigma^j)^2} \mu^j_o + \frac{2\zeta m^j_A (\sigma^j)^2}{1 + 2\zeta m^j_A (\sigma^j)^2} n^j_{A,o} \tag{C.8}$$

$$k^j_{B,o} = k^j_{A,o} + \frac{1}{2\zeta} \log[1 + 2\zeta (m^j_A + \bar{t})(\sigma^j)^2] + \frac{m^j_A (\mu^j_o - n^j_{A,o})^2}{1 + 2\zeta m^j_A (\sigma^j)^2} \tag{C.9}$$

$$\hat{\chi}^{ij}_{do} \equiv \frac{\gamma^{ij} m^j_B}{\sum_{j'} \gamma^{ij'} m^{j'}_B} \times \frac{\exp[-\zeta(\ln \tau^j_{do} + k^j_{B,o} + m^j_B (\mu^i_d - n^j_{B,o})^2)]}{\sum_{o'} \exp[-\zeta(\ln \tau^j_{do'} + k^j_{B,o'} + m^j_B (\mu^i_d - n^j_{B,o'})^2)]} \tag{C.10}$$

$$\ln C^i_d(\mu^i_d) = \ln(\frac{\Xi^i}{A^i_d}) + \gamma^{iL} \ln(w_d) - \zeta^{-1} \sum_j \gamma^{ij} \ln(\sum_o \exp[-\zeta(\ln \tau^j_{do} + k^j_{B,o} + m^j_B (\mu^i_d - n^j_{B,o})^2)]) \tag{C.11}$$

$$\alpha^i_d = \frac{(\eta - 1) m^i_A}{\bar{\phi} + (\eta - 1) m^i_A} n^i_{A,d} = (1 - \beta^i) n^i_{A,d} \tag{C.12}$$

$$\beta^i = \frac{\bar{\phi}}{\bar{\phi} + (\eta - 1) m^i_A}. \tag{C.13}$$

Here, given the ex-ante standard deviation $\{\bar{\sigma}^i\}$, parameters $\{\bar{t}, \bar{\phi}\}$ and $\{\gamma^{ij}, \zeta, \eta\}$ only, (C.2), (C.6), (C.7), and (C.13) form a contraction mapping. With $\{m^i_A\}$ and $\{\sigma^j\}$, given the ex-ante mean $\{\bar{\mu}^i_d\}$, (C.3), (C.5), (C.8), and (C.12) form another contraction mapping. In this contraction mapping, one of the weights (C.10) depends on (C.4), (C.9) and (C.11), which should be simultaneously determined, depending on other parameters $\{\Xi^i, A^i_d\}$ and wages $\{w_d\}$.

Step 2. With $\{C^j_o(\theta)\}$ at hand, we can explicitly evaluate the sourcing decisions of intermediate firms $\chi^j_{do}(\theta,\tilde{\theta})$ and final-good producers $\pi^j_{do}(\theta)$ for all $\theta,\tilde{\theta} \in T$.

For any firm in (d,i) with technology location θ to source input j, by (C.1),

$$
\begin{aligned}
[\tau^j_{do}\Lambda^j_o(\theta)]^{-\zeta} &= (\tau^j_{do})^{-\zeta}\int [C^j_o(\tilde{\theta})t(\theta,\tilde{\theta})]^{-\zeta}\mathrm{d}\Theta^j_o(\tilde{\theta}) \\
&= \exp[-\zeta(\ln\tau^j_{do} + k^j_{B,o} + m^j_B(\theta - n^j_{B,o})^2)].
\end{aligned}
$$

Then, the probability density of sourcing from firms in country o with $\tilde{\theta}$ in (4) is

$$
\chi^j_{do}(\theta,\tilde{\theta}) = \frac{[\tau^j_{do}C^j_o(\tilde{\theta})t(\theta,\tilde{\theta})]^{-\zeta}}{\sum_{o'}[\tau^j_{do'}\Lambda^j_{o'}(\theta)]^{-\zeta}} = \frac{\exp[-\zeta(\ln\tau^j_{do} + k^j_{A,o} + m^j_A(\tilde{\theta} - n^j_{A,o})^2 + \bar{t}(\theta - \tilde{\theta})^2)]}{\sum_{o'}\exp[-\zeta(\ln\tau^j_{do'} + k^j_{B,o'} + m^j_B(\theta - n^j_{B,o'})^2)]}. \tag{C.14}
$$

For final-good producers in country d to consume sector-j goods, by Lemma A.8,

$$
\begin{aligned}
(\bar{\Lambda}^j_o)^{1-\eta} &= \int C^j_o(\tilde{\theta})^{1-\eta}\mathrm{d}\Theta^j_o(\tilde{\theta}) \\
&= \int \exp[(1-\eta)(k^j_{A,o} + m^j_A(\tilde{\theta} - n^j_{A,o})^2)]\cdot\frac{1}{\sqrt{2\pi(\sigma^j)^2}}\exp[-\frac{1}{2}(\frac{\tilde{\theta}-\mu^j_o}{\sigma^j})^2]\mathrm{d}\tilde{\theta} \\
&= \exp[(1-\eta)k^j_{A,o}] \times [1 - 2(1-\eta)m^j_A(\sigma^j)^2]^{-1/2} \times \exp[\frac{(1-\eta)m^j_A(\mu^j_o - n^j_{A,o})^2}{1-2(1-\eta)m^j_A(\sigma^j)^2}] \\
&= \exp[(1-\eta)\underbrace{(k^j_{A,o} - \frac{1}{2(1-\eta)}\log[1-2(1-\eta)m^j_A(\sigma^j)^2] + \frac{m^j_A(\mu^j_o - n^j_{A,o})^2}{1-2(1-\eta)m^j_A(\sigma^j)^2})}_{\equiv k^j_{C,o}}]
\end{aligned}
$$

Then, the expenditure density allocated to goods from country o with $\tilde{\theta}$ in (8) is

$$
\pi^j_{do}(\tilde{\theta}) = \frac{[\tau^{Uj}_{do}C^j_o(\tilde{\theta})]^{1-\eta}}{\sum_{o'}[\tau^{Uj}_{do'}\bar{\Lambda}^j_{o'}]^{1-\eta}} = \frac{\exp[(1-\eta)(\ln\tau^{Uj}_{do} + k^j_{A,o} + m^j_A(\tilde{\theta} - n^j_{A,o})^2)]}{\sum_{o'}\exp[(1-\eta)(\ln\tau^{Uj}_{do'} + k^j_{C,o'})]}. \tag{C.15}
$$

Step 3. With the sourcing decisions specified, we can combine the market-clearing conditions (16) to (19) to arrive at a system of equations of $\{X^j_o(\theta)\}$ and $\{M^j_o(\theta)\}$, taking as given $\{w_d\}$. We discretize the domain of θ, in which case the system of equations is linear in $\{X^j_o(\theta)\}$ and $\{M^j_o(\theta)\}$ and can be easily solved.

Specifically, since the policy function (27) is invertible, summing over the market-clearing conditions (18), (19) and (20), we get

$$
\begin{aligned}
P_dQ_d &= w_dL_d + \sum_i\int \Pi^i_d(g^i_d(\bar{\theta});\bar{\theta})\mathrm{d}\bar{\Theta}^i_d(\bar{\theta}) + \sum_i\int K^i_d(g^i_d(\bar{\theta});\bar{\theta})\mathrm{d}\bar{\Theta}^i_d(\bar{\theta}) && \text{(C.16)} \\
&= w_dL_d + \sum_i\int [1 - \phi(\theta;(g^i_d)^{-1}(\theta)) + \phi(\theta;(g^i_d)^{-1}(\theta))]\frac{1}{\eta}X^i_d(\theta)\mathrm{d}\Theta^i_d(\theta) && \text{(C.17)} \\
&= w_dL_d + \sum_i\int \frac{1}{\eta}X^i_d(\theta)\mathrm{d}\Theta^i_d(\theta). && \text{(C.18)}
\end{aligned}
$$

Substituting (C.14), (C.15) and (C.18) back to (16) and (17), we get

$$X_o^j(\tilde{\theta}) = \sum_d \sum_i \int \rho_d^j \pi_{do}^j(\tilde{\theta})[(\frac{1}{\eta} + \gamma^{iL}(1-\frac{1}{\eta}))X_d^i(\theta) + \gamma^{iL} M_d^i(\theta)]\mathrm{d}\Theta_d^i(\theta),$$

$$M_o^j(\tilde{\theta}) = \sum_d \sum_i \int \gamma^{ij} \chi_{do}^j(\theta, \tilde{\theta})[(1-\frac{1}{\eta})X_d^i(\theta) + M_d^i(\theta)]\mathrm{d}\Theta_d^i(\theta).$$

To numerically approximate this system of equations, we discretize the domain of θ into $\theta \in \widetilde{T} \equiv \{\vartheta_1, \vartheta_2, ..., \vartheta_{N_\theta}\}$ and have

$$d\theta \in \{[\vartheta_1, \vartheta_2 - \frac{\vartheta_2 - \vartheta_1}{2}), [\vartheta_2 - \frac{\vartheta_2 - \vartheta_1}{2}, \vartheta_2 + \frac{\vartheta_3 - \vartheta_2}{2}), ..., [\vartheta_{N_\theta} - \frac{\vartheta_{N_\theta} - \vartheta_{N_\theta - 1}}{2}, \vartheta_{N_\theta})\}.$$

This transforms the system of equations into

$$X_o^j(\tilde{\vartheta})|\mathrm{d}\tilde{\vartheta}| = \sum_d \sum_i \sum_\vartheta \rho_d^j \pi_{do}^j(\tilde{\vartheta})|\mathrm{d}\tilde{\vartheta}| \times [(\frac{1}{\eta} + \gamma^{iL}(1-\frac{1}{\eta}))X_d^i(\vartheta)|\mathrm{d}\vartheta| + \gamma^{iL} M_d^i(\vartheta)|\mathrm{d}\vartheta|],$$

$$M_o^j(\tilde{\theta})|\mathrm{d}\tilde{\vartheta}| = \sum_d \sum_i \sum_\vartheta \gamma^{ij} \chi_{do}^j(\vartheta, \tilde{\vartheta})|\mathrm{d}\tilde{\vartheta}| \times [(1-\frac{1}{\eta})X_d^i(\vartheta)|\mathrm{d}\vartheta| + M_d^i(\vartheta)|\mathrm{d}\vartheta|],$$

where $|\cdot|$ denotes the length of an interval, which is a linear system of equation for $X_d^i(\vartheta)|\mathrm{d}\vartheta|$ and $M_d^i(\vartheta)|\mathrm{d}\vartheta|$, two vectors of real numbers of length $N \times S \times N_\theta$. In matrix form, this is

$$\begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{M} \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}_{X\to X} & \boldsymbol{B}_{M\to X} \\ \boldsymbol{B}_{X\to M} & \boldsymbol{B}_{M\to M} \end{bmatrix} \begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{M} \end{bmatrix}, \tag{C.19}$$

where $\boldsymbol{B}$'s are matrixes of coefficients depending on $\{k_{A,d}^i, m_A^i, n_{A,d}^i, \mu_d^i, \sigma^i\}$ and parameters $\{\rho_d^i, \tau_{do}^j, \tau_{do}^{Uj}, \gamma^{ij}, \gamma^{iL}, \zeta, \eta\}$.[2]

The linear system (C.19) can be easily solved to obtain $\{X_o^j(\theta)\}$ and $\{M_o^j(\theta)\}$.

Step 4. Finally, given the expenditures $\{X_o^j(\theta)\}$ and $\{M_o^j(\theta)\}$, the ex-post technology distribution $\{\mu_d^i, \sigma^i\}$, and parameters $\{L_d, \gamma^{iL}, \eta\}$, we can evaluate whether the labor-market clearing condition, equation (20), is satisfied, i.e.,

$$w_d = \frac{1}{L_d} \sum_i \sum_\vartheta \gamma^{iL}[M_d^i(\vartheta)|\mathrm{d}\vartheta| + (1-\frac{1}{\eta})X_d^i(\vartheta)|\mathrm{d}\vartheta|].$$

If yes, then we have found an equilibrium; if not, update wages $\{w_d\}$ and return to step 1.

Equilibrium Characterization. Once the model is solved, we can characterize the equilibrium explicitly with a number of statistics. These would then allow us to evaluate both consumer welfare and economic efficiency.

We define consumer welfare for each country d as

$$U_d \equiv \frac{w_d L_d + \Pi_d}{P_d},$$

where $w_d L_d$ are the total outputs (GDP) of country d, Π_d are the total profits earned by domestic firms, and P_d the aggregate price index.

[2] Note that (C.19) is homogeneous of degree 1. This can be verified by summing over o, j, and $\tilde{\vartheta}$ on both sides.

The total profits earned by domestic firms can be easily calculated as

$$\Pi_d \equiv \sum_i \int \Pi_d^i(\theta) \mathrm{d}\Theta_d^i(\theta) = \sum_i \int \exp[-\bar{\phi}(\frac{\theta - \alpha_d^i}{\beta^i} - \theta)^2] \cdot \frac{1}{\eta} X_d^i(\theta) \mathrm{d}\Theta_d^i(\theta).$$

To calculate the price index, recall that by (A.4), the sectoral-level price indexes are

$$(P_d^j)^{1-\eta} = \Gamma(1 + \frac{1-\eta}{\lambda}) \cdot [\frac{\eta}{\eta - 1}]^{1-\eta} \cdot \sum_o \exp[(1-\eta)(\ln \tau_{do}^{Uj} + k_{C,o}^j)].$$

This gives the aggregate price index for each country as

$$\begin{aligned}
\ln P_d &= \sum_j \rho_d^j (\ln P_d^j - \ln \rho_d^j) \\
&= \sum_j \rho_d^j [\frac{1}{1-\eta} \ln(\sum_o \exp[(1-\eta)(\ln \tau_{do}^{Uj} + k_{C,o}^j)]) \\
&\qquad . + \frac{1}{1-\eta} \ln(\Gamma(1 + \frac{1-\eta}{\lambda}) \cdot [\frac{\eta}{\eta-1}]^{1-\eta}) - \ln \rho_d^j] \\
&= \sum_j \frac{\rho_d^j}{1-\eta} \ln(\sum_o \exp[(1-\eta)(\ln \tau_{do}^{Uj} + k_{C,o}^j)]) + \sum_j \ln(\frac{\Gamma(1 + \frac{1-\eta}{\lambda})^{\frac{1}{1-\eta}} \cdot (\frac{\eta}{\eta-1})}{\rho_d^j})^{\rho_d^j}.
\end{aligned}$$

In this economy, firms adopt new technologies in order to reduce the costs of being incompatible with suppliers. In equilibrium, the total costs due to technology incompatibility incurred in each country can be calculated as

$$t_d \equiv \sum_i \sum_o \sum_j \int \int \frac{t(\theta, \tilde{\theta}) - 1}{t(\theta, \tilde{\theta})} \cdot \hat{M}_{do}^{ij}(\theta, \tilde{\theta}) \ \mathrm{d}\Theta_d^i(\theta) \mathrm{d}\Theta_o^j(\tilde{\theta}), \tag{C.20}$$

where the imports from $(o, j, \tilde{\theta})$ by (d, i, θ) are

$$\hat{M}_{do}^{ij}(\theta, \tilde{\theta}) \equiv \left[M_d^i(\theta) + (1 - \frac{1}{\sigma}) X_d^i(\theta) \right] \cdot \gamma^{ij} \cdot \chi_{do}^j(\theta, \tilde{\theta}).$$

Correspondingly, the total technology adoption costs spent by firms in each country are

$$K_d \equiv \sum_i \int K_d^i(\theta) \mathrm{d}\Theta_d^i(\theta) = \sum_i \int (1 - \exp[-\bar{\phi}(\frac{\theta - \alpha_d^i}{\beta^i} - \theta)^2]) \cdot \frac{1}{\eta} X_d^i(\theta) \mathrm{d}\Theta_d^i(\theta).$$

C.3 Numerical Implementation of Calibration

This subsection discusses the details of calibration to recover the primitive of the model. Aside the parameters calibrated externally, we jointly determine the remaining parameters on technology distributions $\{\bar{\mu}_o^j, \bar{\sigma}^j\}$, technology adoption costs $\bar{\phi}$, input incompatibility costs $\bar{t}$, and those determining production and trade $\{\tau_{do}^j, \tau_{do}^{Uj}, \Xi^i / A_d^i\}$, leaning on the equilibrium conditions of the model.

The ex-ante technology distribution of countries are by assumption not observed. To calibrate $\{\bar{\mu}_d^i, \bar{\sigma}^i\}$, we use two pieces of information: the ex-post technology distribution, and the one-to-one mapping from the ex-ante to the ex-post distributions characterized in Proposition 8. As the mapping depends on all model primitives (such as trade costs) and the equilibrium wage, the ex-ante distributions cannot be recovered independent of the rest of the model. Instead, we recover the ex-ante distribution in two steps.

In the first step, we choose the parameters governing the ex-post distributions, $\{\mu_d^i, \sigma^i\}$, to match patent citation statistics. This step can be carried out without knowing the primitives of the model. Conditional on their own technology, firms' sourcing decisions only depend on the ex-post distributions. We can therefore calibrate the primitives of the model governing trade using only the ex-post distributions and other data. In the second step, we calibrate $\bar{\phi}$ and recover the ex-ante distributions using equation (28).

Ex-post technology distribution. To calibrate the ex-post distribution, we use the model-implied citation shares grounded in the knowledge-source attribution problem described in Section 2.8, extended to account for the fact that each output sector relies differently on the knowledge of different input sectors. Concretely, for any firm from sector i that chooses technology location θ, we calculate the share of citations it makes that goes to (o, j) as

$$\psi_o^{ij}(\theta) \equiv \delta^{ij} \cdot \frac{H_o^j \cdot \mathrm{d}\Theta_o^j(\theta)}{\sum_{o'} H_{o'}^{j'} \cdot \mathrm{d}\Theta_{o'}^{j'}(\theta)}$$
$$\text{with} \quad d\Theta_o^j(\theta) = \frac{1}{\sqrt{2\pi(\sigma^j)^2}} \exp\left[-\frac{1}{2(\sigma^j)^2}\left(\theta - \mu_o^j\right)^2\right].$$

where δ^{ij} is the sectoral technology proximity, measured by the share of patent citations made to sector j by sector i, and H_o^j is the total number of citations received by patents invented in (o, j) in data. By using citation counts to measure H_o^j, this share accounts for the difference in the vertical quality of patents across (o, j); by using δ^{ij} to weight supplier sectors, this share allows for the possibility that (o, j) receive more citation from (d, i) because sector i is technically more dependent on j than other sectors.

We integrate $\psi_o^{ij}(\theta)$ over the technology distribution in (d, i) to obtain aggregate bilateral citation shares:

$$\Psi_{do,model}^{ij} \equiv \int \psi_o^{ij}(\theta)\mathrm{d}\Theta_d^i(\theta),$$

where $\sum_{o,j} \Psi_{do,model}^{ij} = 1$. Plugging in δ^{ij} and H_o^j, for any $\{\Theta_o^j\}$, this expression delivers the model-implied share of citations made by (d, i) to (o, j). We then find $\{\mu_d^i, \sigma^i\}$ by solving the following nonlinear least square problem:

$$(\mu_d^i, \sigma^i) = \arg\min \sum_{o,d,i,j} (\Psi_{od,model}^{ij} - \Psi_{od,data}^{ij})^2.$$

We construct the empirical measure of patent citation shares $\Psi_{od,data}^{ij}$, total number of citations H_o^j, and the sectoral shares of patent citations δ^{ij} from the PATSTAT Global, all aggregated over 2010-2014.

Trade costs and distribution of production techniques. With the ex-post technology distributions $\{\Theta_o^j\}$ at hand, we design a nested algorithm to jointly calibrate the parameters $\{\Xi^i / A_d^i\}$, which determines the productivity of (d, i), to match the output share of (d, i) in industry i, and calibrate $\{\tau_{do}^j, \tau_{do}^{Uj}\}$ to match the trade shares of intermediate and final goods, respectively. We lay out the algorithm as follows before discussing several details.

The nested algorithm.

(A) Choose a $\bar{t}$

(a) Choose a set of parameters $\{\tau^j_{do}, \tau^{Uj}_{do}, \Xi^i / A^i_d\}$

(b) Solve the equilibrium given the parameters and the ex-post technology distributions $\{\mu^i_d, \sigma^i\}$

(c) Evaluate the trade shares of intermediate and final goods at the equilibrium. If they match their data counterparts, proceed to Step (B); if not, return to Step (A)(b).

(B) Simulate 10,000 Chinese firms from each sector and regress the extensive margin of importing from each country o on the extensive margin of citing patents from o, conditional on firm an country-industry fixed effects

(C) Compare the model-based regression coefficient in Step 2 to its data counterpart (Column 3 of Table 2). If they are close enough, exit; if not, return to Step (A).

In Step (A)(b), we need to solve the equilibrium given the ex-post distribution and other parameters. This requires some modifications on the algorithm developed in C.2. Specifically, in Step 1 of the algorithm to solve the equilibrium, the system of equations are simplified into

$$
\begin{aligned}
m^i_A &= \sum_j \gamma^{ij} m^j_B \\
n^i_{A,d} &= \sum_{j,o} \hat{\chi}^{ij}_{do} n^j_{B,o} \\
k^i_{A,d} &= \ln C^i_d(\mu^i_d) - m^i_A [\sum_{j,o} \hat{\chi}^{ij}_{do} (\mu^i_d - n^j_{B,o})]^2
\end{aligned}
$$

where

$$
\begin{aligned}
m^j_B &= \frac{\bar{t}[1 + 2\zeta m^j_A (\sigma^j)^2]}{1 + 2\zeta(m^j_A + \bar{t})(\sigma^j)^2}, \\
n^j_{B,o} &= \frac{1}{1 + 2\zeta m^j_A (\sigma^j)^2} \mu^j_o + \frac{2\zeta m^j_A (\sigma^j)^2}{1 + 2\zeta m^j_A (\sigma^j)^2} n^j_{A,o} \\
k^j_{B,o} &= k^j_{A,o} + \frac{1}{2\zeta} \log[1 + 2\zeta(m^j_A + \bar{t})(\sigma^j)^2] + \frac{m^j_A (\mu^j_o - n^j_{A,o})^2}{1 + 2\zeta m^j_A (\sigma^j)^2} \\
\hat{\chi}^{ij}_{do} &\equiv \frac{\gamma^{ij} m^j_B}{\sum_{j'} \gamma^{ij'} m^{j'}_B} \times \frac{\exp[-\zeta(\ln \tau^j_{do} + k^j_{B,o} + m^j_B (\mu^i_d - n^j_{B,o})^2)]}{\sum_{o'} \exp[-\zeta(\ln \tau^j_{do'} + k^j_{B,o'} + m^j_B (\mu^i_d - n^j_{B,o'})^2)]} \\
\ln C^i_d(\mu^i_d) &= \ln(\frac{\Xi^i}{A^i_d}) + \gamma^{iL} \ln(w_d) - \zeta^{-1} \sum_j \gamma^{ij} \ln(\sum_o \exp[-\zeta(\ln \tau^j_{do} + k^j_{B,o} + m^j_B (\mu^i_d - n^j_{B,o})^2)])
\end{aligned}
$$

Given wages $\{w_d\}$ and parameters $\{\tau^j_{do}, \tau^{Uj}_{do}, \Xi^i / A^i_d\}$, this is a contraction mapping and can be efficiently solved.

The remaining steps to solve the equilibrium follow exactly with

$$\chi^j_{do}(\theta,\tilde{\theta}) = \frac{\exp[-\zeta(\ln\tau^j_{do} + k^j_{A,o} + m^j_A(\tilde{\theta} - n^j_{A,o})^2 + \bar{t}(\theta-\tilde{\theta})^2)]}{\sum_{o'}\exp[-\zeta(\ln\tau^j_{do'} + k^j_{B,o'} + m^j_B(\theta - n^j_{B,o'})^2)]}$$

$$\pi^j_{do}(\tilde{\theta}) = \frac{\exp[(1-\eta)(\ln\tau^{Uj}_{do} + k^j_{A,o} + m^j_A(\tilde{\theta} - n^j_{A,o})^2)]}{\sum_{o'}\exp[(1-\eta)(\ln\tau^{Uj}_{do'} + k^j_{C,o'})]}$$

$$X^j_o(\tilde{\vartheta})|\mathrm{d}\tilde{\vartheta}| = \sum_d\sum_i\sum_\vartheta \rho^j_d\pi^j_{do}(\tilde{\vartheta})|\mathrm{d}\tilde{\vartheta}| \times [(\frac{1}{\eta} + \gamma^{iL}(1-\frac{1}{\eta}))X^i_d(\vartheta)|\mathrm{d}\vartheta| + \gamma^{iL}M^i_d(\vartheta)|\mathrm{d}\vartheta|]$$

$$M^j_o(\tilde{\vartheta})|\mathrm{d}\tilde{\vartheta}| = \sum_d\sum_i\sum_\vartheta \gamma^{ij}\chi^j_{do}(\vartheta,\tilde{\vartheta})|\mathrm{d}\tilde{\vartheta}| \times [(1-\frac{1}{\eta})X^i_d(\vartheta)|\mathrm{d}\vartheta| + M^i_d(\vartheta)|\mathrm{d}\vartheta|].$$

The equilibrium is reached when the wages $\{w_d\}$ satisfies the labor-market clearing condition given by

$$w_d = \frac{1}{L_d}\sum_i\sum_\vartheta \gamma^{iL}[M^i_d(\vartheta)|\mathrm{d}\vartheta| + (1-\frac{1}{\eta})X^i_d(\vartheta)|\mathrm{d}\vartheta|].$$

In equilibrium, we calculate the trade shares in final and intermediate goods as

$$\frac{\hat{M}^j_{do}}{\sum_{o'}\hat{M}^j_{do'}} \text{ with } \hat{M}^j_{do} \equiv \sum_{\tilde{\vartheta}}\sum_i\sum_\vartheta \gamma^{ij}\chi^j_{do}(\vartheta,\tilde{\vartheta})|\mathrm{d}\tilde{\vartheta}| \times [(1-\frac{1}{\eta})X^i_d(\vartheta)|\mathrm{d}\vartheta| + M^i_d(\vartheta)|\mathrm{d}\vartheta|]$$

$$\frac{\hat{X}^j_{do}}{\sum_{o'}\hat{X}^j_{do'}} \text{ with } \hat{X}^j_{do} \equiv \sum_{\tilde{\vartheta}}\sum_i\sum_\vartheta \rho^j_d\pi^j_{do}(\tilde{\vartheta})|\mathrm{d}\tilde{\vartheta}| \times [(\frac{1}{\eta} + \gamma^{iL}(1-\frac{1}{\eta}))X^i_d(\vartheta)|\mathrm{d}\vartheta| + \gamma^{iL}M^i_d(\vartheta)|\mathrm{d}\vartheta|]$$

We match these two statistics to their data counterparts constructed from WIOTs, both averaged over 2010-2014.

In Step (B), for each value of $\bar{t}$, we simulate 190,000 Chinese firms ($d = CHN$), 10,000 from each sector i with θ drawn from the calibrated ex-post distribution Ψ^i_d. With parameters $\{\tau^j_{do}, \tau^{Uj}_{do}, \Xi^i/A^i_d\}$ calibrated, we can explicitly determine the input sourcing and patent citation patterns for each firm. Specifically, a firm with technology location θ from (d,i) would source input from (o,j) with probability $\chi^j_{do}(\theta)$ given by

$$\chi^j_{do}(\theta,\tilde{\theta}) = \frac{\exp[-\zeta(\ln\tau^j_{do} + k^j_{A,o} + m^j_A(\tilde{\theta} - n^j_{A,o})^2 + \bar{t}(\theta-\tilde{\theta})^2)]}{\sum_{o'}\exp[-\zeta(\ln\tau^j_{do'} + k^j_{B,o'} + m^j_B(\theta - n^j_{B,o'})^2)]},$$

and would cite patents in (o,j) with probability $\psi^{ij}_o(\theta)$ given by

$$\psi^{ij}_o(\theta) \equiv \delta^{ij}\cdot\frac{H^j_o\cdot\mathrm{d}\Theta^j_o(\theta)}{\sum_{o'}H^{j'}_{o'}\cdot\mathrm{d}\Theta^{j'}_{o'}(\theta)} \text{ with } d\Theta^j_o(\theta) = \frac{1}{\sqrt{2\pi(\sigma^j)^2}}\exp\left[-\frac{1}{2(\sigma^j)^2}\left(\theta-\mu^j_o\right)^2\right].$$

Then, we regress the realized extensive margin of importing from each country o on the extensive margin of citing patents from o, controlling for firm and country-industry fixed effects. We calibrate $\bar{t} = 0.05$ to match this coefficient to 0.022, the extensive-margin import-citation correlation obtained from the regression results in Column 3 of Table 2.

Notice that the above algorithm does not rely on the value of parameter $\bar{\phi}$. This is because that conditional on the ex-post distribution, varying $\bar{\phi}$ only affects the split of firms' markup between profit to the representative consumer and the expense on adaption. In other words, $\bar{\phi}$ only affects the welfare of agents but not the equilibrium wages or prices. For this reason, we can calibrate $\bar{\phi}$ separately.

Calibrate $\bar{\phi}$ and the ex-ante distributions. With all other model primitives calibrated, the last step is to calibrate $\bar{\phi}$, the parameter on innovation costs. We calibrate it by matching the citation elasticity of tariffs (Column 2 of Table 1). As shown in Proposition 4, conditional on trade shares and γ^{ij}, this elasticity identifies $\bar{\phi}$. To obtain the elasticities in the model, we conduct counterfactuals that decrease the trade costs $\{\tau_{do}^{j}\}$ by 5% for all (d, o, j)'s with $d \neq o$ and χ_{do}^{j} greater than 5%, one (d, o, j) at a time. For each counterfactual, we calculate the implied elasticity for the change in the share of citations made by country d that goes to (o, j). We then adjust the value of $\bar{\phi}$ such that the mean of the elasticities calculated across these simulations matches the regression coefficient -0.296. The calibrated value is $\bar{\phi} = 0.005$.

With $\bar{\phi}$ specified, we can then recover the ex-ante technology distribution by equations (27) and (28), i.e.,

$$\begin{aligned}
\bar{\mu}_d^i &= \frac{1}{\beta^i} \cdot \mu_d^i - \frac{1-\beta^i}{\beta^i} \cdot n_{A,d}^i \\
\bar{\sigma}^i &= \frac{1}{\beta^i} \cdot \sigma^i \\
\text{with} \quad \beta^i &= \frac{\bar{\phi}}{\bar{\phi} + (\eta - 1) m_A^i}.
\end{aligned}$$

This completes the calibration.