Trade and Technology Compatibility in General Equilibrium*

Weili Chen† Jingting Fan‡ Wenlan Luo§
Penn State Penn State Tsinghua University

This Draft: March 2025
First Draft: November 2023

Abstract

We develop a trade model where the horizontal proximity between a firm's technology and that of its suppliers shapes input efficiency. Trade policies, by expanding or restricting firms' access to foreign suppliers, influence these horizontal technology choices and, through input-output linkages, affect the technological interdependence between countries. We characterize these effects and test them using technological similarity measures derived from patent texts. Quantification of the model shows that an electronics and semiconductor embargo by the U.S. against China triggers technological decoupling between the two countries and realignment among other countries. These endogenous technological shifts double the global welfare loss caused by the embargo.

Key words: technology proximity; trade conflict; technology decoupling; production network; geopolitics

---

*We thank Jonathan Eaton, Thomas Chaney, Agustín Gutiérrez, Andrei Levchenko, Nuno Limao, Ernest Liu, Dan Lu, Fernando Parro, Michael Zheng Song, Kjetil Storesletten, Chang Sun, Daniel Yi Xu, and Stephen Yeaple for helpful conversations. We also thank seminar and conference participants at Barcelona Summer Forum (geography, trade, and growth), CKGSB, EIIT (Purdue), Fudan Fanhai, Georgetown, Hong Kong University, HKUST, INSEAD, National University of Singapore, North Carolina State University, Peking University (Guanghua; HSBC), the University of Nottingham, Tsinghua Growth and Institutions Conference 2023, the Asian Bureau of Economic Research Conference 2024 (trade and growth), and China International Conference in Macroeconomics 2024 for helpful comments.

†The Pennsylvania State University, weilichen-econ@psu.edu
‡The Pennsylvania State University, jxf524@psu.edu
§Tsinghua University, luowl@sem.tsinghua.edu.cn

1 Introduction

Recent global events have driven major economies to adopt strategies aimed at reducing their economic engagement with geopolitical rivals. For instance, the European Union is diversifying its sources of essential goods to mitigate supply chain disruption risks, while the U.S. has implemented policies to limit China's access to advanced technology. Through initiatives like 'friend-shoring' and 'Made in China 2025,' both the U.S. and China are working to decrease reliance on critical goods from each other.

These shifts raise a growing concern: reduced access to, and dependence on, key foreign technologies may push firms toward alternative technological paths, potentially dividing the world into blocs of countries with incompatible technological systems. For example, restrictions on Chinese chipmakers' access to x86 instruction set licenses have led them to develop proprietary instruction sets or adopt open-source alternatives like RISC-V.[1] Similarly, in response to China's control over cobalt, a critical battery component, global manufacturers are developing cobalt-free batteries.[2] Both shifts force adaptations throughout the supply chain: new computer chips necessitate compatible operating systems and software adjustments, while cobalt-free batteries require updates to charging protocols and car designs. These examples illustrate how firms' technological choices are shaped by access to foreign inputs in their supply chains, and how policies affecting this access can either promote technological convergence across global supply chains or fragment them into incompatible systems.

This paper makes two contributions. First, we construct a tractable model to formalize the link between trade and the horizontal technological choices of firms. This enables us, for the first time in the literature, to assess whether a trade shock leads to technological divergence or convergence. Second, using a new measure of technology similarity based on the texts of global patents, we test and quantify the model. We apply the model to assess the impacts of a U.S. embargo on Chinese electronics and semiconductor imports. We find that the embargo leads to technological decoupling between the two countries and a realignment of others. These endogenous changes exacerbate—rather than alleviate—the welfare losses from the embargo, accounting for nearly all of the U.S. losses and half of the global losses.

In our model, firms differ in vertical efficiency and their endowed horizontal technology, denoted by a point $\bar{\theta}$ in technology space $\mathbb{T}$. Following the seminal work of Lancaster (1979), we interpret each point in $\mathbb{T}$ as a specific 'technological path' composed of particular technical components or engineering features. Points that are closer in $\mathbb{T}$ correspond to

[1] x86 is a CPU architecture developed jointly by Intel and AMD. Chipmakers need a license to create CPUs compatible with x86, which is essential for these CPUs to support some mainstream operating systems, such as Windows, and the software running on these systems. Since 2018, Chinese chipmakers' requests for license renewal have been denied due to national security concern.

[2] For example, see https://www.cnbc.com/2021/11/17/samsung-panasonic-and-tesla-embracing-cobalt-free-batteries-.html, accessed February, 2024.

more similar (and hence more compatible) technologies. We label technologies 'horizontal' because no technology is inherently superior to another; rather, they differ by their distance from one another. A firm's endowed technology $\bar{\theta}$ represents the path most suited to its intrinsic capabilities, shaped by factors outside our model such as accumulated know-how or key expertise of the firm. The distribution from which firms draw $\bar{\theta}$ varies across countries and sectors, reflecting countries' comparative advantage over different technologies.

Given its endowed $\bar{\theta}$, each firm chooses a production technology $\theta \in \mathbb{T}$. Choosing a technology closer to the ones used by key suppliers reduces effective input costs by enhancing compatibility, but deviating from $\bar{\theta}$ also entails an adaptation cost that grows with the distance $\|\theta - \bar{\theta}\|$. After deciding on θ, firms sample a set of suppliers and choose the one that offers the best compatibility-adjusted price for each intermediate input. They then produce and sell the output to final consumers and downstream firms. In equilibrium, given the country-specific *exogenous* distributions governing $\bar{\theta}$, firms anticipate the choices of all other firms and make their technology and supplier choices accordingly.

The equilibrium is characterized by a fixed point in the distributions of chosen technologies across firms in all countries and sectors—an infinite-dimensional object. We establish that there generically exists an equilibrium in which firms' technology choices vary continuously with their technology endowment, and that this equilibrium is unique provided the cost of technology adaptation is not too small and the benefit of compatibility is not too large. In the limiting case where endowment distributions are degenerate and compatibility yields no benefit, our model simplifies to the model of Caliendo and Parro (2015). Thus, the model retains the flexibility of canonical quantitative trade models.

Our model uncovers novel relationships between trade and technology, which we both formalize and empirically test. First, when compatibility drives firms' decisions, those whose endowed technologies are closer to a particular foreign country's technological profile are more likely to adopt similar technologies and import from that country. This results in a *correlation* between a firm's importing patterns and its technology similarity with exporting countries. Second, lowering tariffs with a specific trading partner incentivizes firms in the importing country to shift their technologies toward that partner's technologies, implying a *causal* link between bilateral trade costs and technological convergence.

Our model also offers a fresh perspective on existing empirical regularities. For instance, firms that import from a country often export to the same country, while exporters to one market are more likely to serve that market's neighboring markets if those neighbors share similar technologies. These patterns, previously rationalized through state-dependent trade costs (Morales et al., 2019; Li et al., 2024), can be explained here through technology compatibility.

On the normative side, the model highlights a key externality in technology choice. When a firm chooses its technology, it also affects the production costs of domestic and foreign downstream firms, who may face lower production costs if this upstream firm chooses a closer technology, an effect not internalized by the firm. Such externalities exist both within and across borders and are amplified by input-output linkages. In a one-sector, two-symmetric-country special case of the model, we show that starting from the decentralized equilibrium, shifting one country's technology toward the other's increases welfare in *both* countries. Thus, firms tend to stay too close to their endowed technology.

We test and quantify the importance of our model's mechanisms. Central to these exercises is a new measure of horizontal technology similarity across firms and countries. To construct this measure, we use the texts of worldwide patents and employ a state-of-the-art large language model to generate numerical vectors, or 'embeddings,' that summarize each patent's content. We then compute bilateral similarities at various levels of aggregation using the cosine similarity of these embeddings. We validate our measure by comparing it with other proxies of bilateral similarities. We also demonstrate that our empirical results remain robust when patent citations are used as a proxy for technology similarity.

We undertake two main empirical tests. First, a premise of our model is that firms make joint technology and supplier choices with compatibility considerations. Using Chinese firm and customs data, we show that firms whose technologies are more similar to a partner country's are more likely to import from that country—a relationship that remains robust to a range of controls, fixed effects, and the exclusion of foreign affiliates. Additional evidence supports technology compatibility as the primary channel rather than technological diffusion or other information-based explanations. Second, our model implies that lowering bilateral trade barriers increases bilateral technology similarity. Exploiting exogenous variation in HS-6-level tariffs arising from changes in most-favored-nation (MFN) rates, which affects some partners (those subject to the MFN tariffs) but not others, we estimate a negative and economically meaningful elasticity of bilateral technology similarity with respect to tariffs, in line with the model's prediction.[3]

We parameterize the model using data on 29 major countries and 19 sectors. For tractability, we let $\mathbb{T}$ be the real line and assume that firms in each country and sector draw their endowment technology from a respective Normal distribution. We show that under a quadratic approximation to firms' technology choice problem, the chosen technologies in each country-sector also follow a Normal distribution, with its mean and variance determined by general equilibrium forces. This structure makes the model computationally tractable.

Calibration of the model requires mapping the horizontal technologies of countries and sectors to the real line. We choose the technology positions by matching the model's implied pairwise technology similarities to those obtained from patent texts.[4] In addition, two

[3]MFN tariffs can be endogenously chosen to promote trade. Following Boehm et al. (2023), we address this concern by excluding from the sample the largest exporter in each importer-product pair.

[4]We do not seek to explain, for example, why the auto sector uses more similar technologies to the auto parts

parameters are key in shaping the impacts of technology choice: the cost of incompatibility with suppliers and the cost of deviating from a firm's endowed technology. We identify these parameters by matching the two reduced-form estimates: (i) the correlation between import behavior and technology similarity, and (ii) the elasticity of technology similarity with respect to tariffs. Using these parameters and the calibrated technological distribution of countries, the model implies a component of trade costs reflecting technology incompatibility. Residual iceberg trade costs are then chosen to match observed bilateral trade shares.

Our calibration suggests that technology compatibility is an important driver of sourcing decisions. The ad-valorem equivalent trade cost due to incompatibility is around 10%. Standard estimation suggests that to account for the observed bilateral trade (or the lack thereof), trade costs need to be well above 100%, often raising the question on the source of such high costs (Anderson and van Wincoop, 2004). Our model provides a partial explanation: technological incompatibility between countries.

We assess the impact of a trade conflict between the U.S. and China by implementing a counterfactual experiment, in which the U.S. and its allies impose an embargo on Chinese electronics and semiconductor imports. We find that the embargo leads to a 0.55% welfare loss for China, a 0.05% loss for the U.S., and a global loss of 0.23%. In response to the embargo, the technologies of firms in China diverge from those of other countries, leading to higher input costs for foreign downstream firms and triggering realignments of technologies in other countries, including the U.S. Contrary to the intuition that substitution adjustments to negative shocks alleviate welfare loss, the divergence between China and the U.S. amplifies the loss for both, due to the externality in technological choice. Almost all of the U.S. welfare loss and half of the global welfare loss are due to endogenous technological changes.

Our paper engages with several strands of the literature. The central idea—that firms' technology choices are shaped by compatibility requirements—has been explored in the trade literature. However, existing models often focus on stylized settings with two countries and two technologies (e.g., domestic vs. foreign), limiting their quantitative generalizability.[5,6] We allow for multiple technologies, countries, and input-output linkages. Our model builds on Lancaster (1979)'s work on endogenous product differentiation, but emphasizes

sector than to the apparel sector; rather, we are interested in, say, whether the U.S. auto sector is more similar to Germen or Canadian auto sector. Consequently, we remove sector-pair fixed effects from our calibration targets, isolating cross-country variation.

[5]For example, Carluccio and Fally (2013) examine how supplying foreign firms using a 'modern' technology can reduce input availability for domestic firms using a 'traditional' technology; Costinot (2008) develops a two-country model to explore how horizontal standardization affects welfare. Given that there are multiple countries in the world economy, each with potentially different technologies, mapping these technologies to a binary choice in the model can be subjective.

[6]The concept of compatibility-driven technology choice has also been studied outside of trade. For example, Akcigit et al. (2016) present a growth model where firms sell ideas that diverge from their core competence, Lazear (1999) models language choice driven by communication incentives, and Basu and Weil (1998) model technology transfer considering whether the technology is appropriate for a country.

compatibility with suppliers rather than consumers.

The network aspect of our model builds on recent works such as Jones (2013), Chaney (2014), Oberfield (2018), Lim (2018), Boehm and Oberfield (2020), Acemoglu and Azar (2020), Demir et al. (2024), Eaton et al. (2022), Dhyne et al. (2023) and Baqaee and Farhi (2024). In particular, we achieve tractable network formation through extreme-valued draws, similar to the sourcing model of Boehm and Oberfield (2020) and the diffusion model of Buera and Oberfield (2020). Our focus on technology choice in a network relates to Demir et al. (2024), which studies network formation under (vertical) quality choice. Our main contribution to this literature is to incorporate horizontal technology choice into a model with endogenous production networks.

The trade aspect of our model builds on the quantitative trade literature (e.g., Eaton and Kortum, 2002; Caliendo and Parro, 2015; see Costinot and Rodriguez-Clare, 2014 for an early review), with two main differences. First, it incorporates horizontal technology choice, which we show is important in shaping the effects of trade policies. To accommodate this mechanism, in our model, heterogeneous firms choose among a continuum of options and then interact with the entire distribution of other firms. We develop tools for establishing the existence and uniqueness of equilibrium in such environments, which differ from those in standard trade models where the equilibrium can be characterized solely by aggregate quantities and prices (e.g., Eaton and Kortum, 2002; Chaney, 2008; see Allen et al., 2024 for uniqueness results for some of these models).[7] Second, we provide a tractable extension of canonical quantitative trade models that allows firms to source inputs based on flexible interactions between buyer and seller characteristics.

Our paper also connects with the literature on the relationship between trade and technology (e.g., Eaton and Kortum, 1999; Buera and Oberfield, 2020; Liu and Ma, 2021; Aghion et al., 2021; Cai et al., 2022; Lind and Ramondo, 2023), which often uses patents to measure technology diffusion. While most existing studies focus on how trade affects vertical technology (e.g., aggregate or sectoral productivity), we focus on horizontal technology choices.[8] Motivated by this focus, we construct a new, theory-consistent, measure of technology similarity for empirical and quantitative analyses. We provide evidence for the compatibility mechanism and offer a structural interpretation of the evidence, complementing existing studies emphasizing learning or other forms of technology diffusion.

Finally, recent literature has estimated the impacts of the trade conflict between the U.S.

[7]Other trade models in which firms interact with the entire distribution of firms include those based on the assignment model, e.g., Costinot and Vogel (2010). 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 with externalities.

[8]Lind and Ramondo (2023) study how trade affects the similarity of technologies among countries, modeling similarity through the correlation in productivity draws. In our model, technology similarity is an outcome of firms' choices.

and China, focusing on the short-run effects (e.g., Amiti et al., 2019; Fajgelbaum et al., 2020; Huang et al., 2023). Our finding suggests that an important part of the welfare losses will appear as technological adjustments take effect. Our focus on the non-tariff aspects of trade conflicts also relates to works on the interaction between trade and international relations, e.g., see Kleinman et al. (2024), Bonadio et al. (2024), and Clayton et al. (2023).

2 Model

2.1 Environment.

There are N countries, denoted by d or o, and S sectors, denoted by i or j. When referring to directional flows, we use (d,i) to denote the buying country-sector pair and (o,j) to denote the selling country-sector pair.

In each country d, a representative household supplies L_d units of labor, and a representative producer makes a *non-tradable* final good with the following production 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 the final good in d. The parameter $\eta > 1$ is the elasticity of substitution between different varieties, and ρ_d^j is the share of sector j in the final good. This final good is used for domestic consumption and domestic firms' technology adaptation.

In each country-sector pair (d,i), there is a continuum of intermediate-good firms denoted by $\nu, \omega \in [0,1]$. Each firm produces a differentiated variety using labor and inputs purchased from other intermediate-good firms. We assume that intermediate-good firms engage in monopolistic competition when selling to final-good producers—thereby earning positive profits—and charge marginal production costs when selling to downstream intermediate-good firms. Before production, intermediate-good firms choose a horizontal technology, balancing the desire to use the technology they know best (their endowed technology) and the need to be compatible with key suppliers.

Let $\mathbb{T}$ be a metric space containing all technologies—such as a circle, the real line, or a finite-dimensional Euclidean space.[9] Each firm ν is endowed with a technology $\bar{\theta}(\nu) \in \mathbb{T}$. Given $\bar{\theta}(\nu)$, the firm selects a production technology $\theta(\nu) \in \mathbb{T}$. Deviating from $\bar{\theta}(\nu)$ incurs adaptation costs, yet may yield benefits by enhancing compatibility with certain suppliers. After choosing $\theta(\nu)$, the firm randomly samples a set of suppliers and selects those that offer the lowest compatibility-adjusted input cost. Firms then produce and sell the output to other

[9] For some of the theoretical characterizations and later quantitative analysis, we restrict to the real line. We discuss the implications of focusing on the real line in Section 2.4.

firms and final-good producers.

Technologies in our model have broad interpretations. They might represent distinct technical ecosystems—such as Android versus iOS for smartphones—or capture differences in marketing strategies, business networks, or political alignments among firms and countries. They may also embody product designs that rely on different scientific and engineering disciplines. For example, makers of internal combustion engine cars and their suppliers tend to rely predominantly on mechanical engineering, whereas makers of electric vehicles and their suppliers draw more on advances in chemistry and electronics. Although our theoretical framework does not require a precise definition of technology, our empirical and quantitative analyses operationalize technology as the knowledge embedded in patent texts, using patent similarity measures to assess technological relationships between countries.

In the remainder of this section, we first describe firms' production and input-sourcing decisions, taking their chosen technology as given. We then describe how firms select their technology and characterize the equilibrium for technology choice. Finally, we derive testable implications of the model and discuss its welfare implications.

2.2 Production and Sourcing Decisions: Intermediate-Good Firms

Consider a firm ν in country-sector (d, i) that has already chosen technology $\theta(\nu)$. This firm can access a random set of production *techniques*, denoted by $R(\nu)$. Each technique $r \in R(v)$ is characterized by (i) a production efficiency $A(\nu, r)$ and (ii) a set of potential suppliers drawn independently and uniformly from firms in country-sector pair (o, j) for each o and j. Denote the set of suppliers for technique r from (o, j) by $\Omega_o^j(\nu, r)$.

Under technique r, the firm's output 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 firm ν, and $m^j(\nu, r)$ is the sector-j intermediate good. Since there is no love-of-variety in input sourcing, suppliers of input j from around the world engage in perfect competition and the cheapest one wins the order from firm ν.

Denote the wage rate in d by w_d. The factory-gate price of firm ν under technique r is

$$p(\nu, r) = \frac{1}{A(\nu, r)} \cdot [w_d]^{\gamma^{iL}} \cdot \prod_j \left[c^j(\nu, r)\right]^{\gamma^{ij}}, \tag{2}$$

where $c^j(\nu, r)$ is the minimum cost for input j among the suppliers available for technique r:

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

$\tilde{c}^j(\nu,\omega)$, the (compatibility-adjusted) cost of input j from supplier $\omega \in \Omega_o^j(\nu,r)$ is

$$\tilde{c}^j(\nu,\omega) = \tau_{do}^j \cdot p(\omega) \cdot \frac{1}{z(\nu,\omega)} \cdot t(\theta(\nu),\theta(\omega)), \tag{3}$$

where τ_{do}^j is the iceberg trade cost, $p(\omega)$ is the factory-gate price of supplier ω, and $z(\nu,\omega)$ is an idiosyncratic match-specific efficiency draw. $t(\theta(\nu),\theta(\omega))$ is an increasing function of the distance between $\theta(\nu)$ and $\theta(\omega)$, capturing technology compatibility: intermediate goods perform better when the buyer and the suppliers use similar technologies.

Since firms sell their products to downstream firms at their marginal production cost, their profits accrue solely from sales to final good producers. Firm ν chooses among $R(\nu)$ the technique that maximizes these profits. Since firms charge a constant markup to final good producers under monopolistic competition, profit maximization is equivalent to minimizing the factory-gate price:

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

Because firms buy from and sell to each other simultaneously, the price distribution in any (d,i) depends endogenously on those in all other country–sector pairs. To characterize these distributions, we make the following assumption on the techniques available to firms.

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 an independent Poisson distribution with mean $(a_1/A_d^i)^{-\lambda} - (a_2/A_d^i)^{-\lambda}$, with $\lambda > 1$.*
2. *For any production technique r and each (o,j), suppliers in $\Omega_o^j(\nu,r)$ are drawn independently and uniformly from intermediate-good firms in country o sector j.*
   The match-specific efficiency draws of these suppliers, $z(\nu,\omega)$, satisfy the following: for any $z_1 < z_2 \in (0,+\infty)$, the number of suppliers in $\Omega_o^j(\nu,r)$ with $z(\nu,\omega) \in (z_1,z_2]$ follows an independent Poisson with mean $z_1^{-\zeta} - z_2^{-\zeta}$, where for integratability $\zeta \in \left(\max\{\max_{ij}\gamma^{ij}\lambda,1\},\lambda\right)$.

Assumption 1 specifies both the productivity draws for each technique and the match quality draws for suppliers. The Poisson assumption, which follows Boehm and Oberfield (2020), can be viewed as an extension to the more familiar Pareto assumption (Chaney, 2008; Kortum, 1997).[10] In Part 1, A_d^i determines the efficiency level of available techniques in (d,i)—higher A_d^i implying more high-efficiency techniques, which means country d has an advantage in i—while λ governs the heterogeneity among these techniques.[11] In Part 2, the

[10]Draws satisfying Part 1 of Assumption 1 can be generated as follows: take any $x > 0$, draw the number of techniques with $A(v,r) > x$ from a Poisson distribution with mean $(x/A_d^i)^{-\lambda}$; for each technique, draw its efficiency independently from a Pareto distribution with position parameter A_d^i and tail parameter λ. Based on the thinning properties of Poisson, the limit of these process when $x \to 0$ satisfies Part 1.

[11]An intuitive way to see this is to consider the limit as $a_2 \to \infty$, which reveals that A_d^i shapes the number of available techniques exceeding any threshold a_1 and λ governs how quickly that number decays with a_1.

uniform sampling of suppliers in $\Omega_o^j(\nu, r)$ implies that the technologies of potential suppliers follow those of all firms in (o, j). Parameter ζ then governs the heterogeneity of match quality among these suppliers, similar to how λ does for available techniques.

We now characterize the distribution of factory-gate prices of firms in (d, i) conditional on the technology they choose:

Proposition 1. *Under Assumption 1, $p_d^i(\theta)$, the factory-gate price of a firm in (d, i) that has chosen technology θ, follows a Weibull (inverse Fréchet) distribution with c.d.f.*

$$F_d^i(p; \theta) = 1 - \exp\big(- [p / C_d^i(\theta)]^\lambda \big). \tag{4}$$

Its location parameter $C_d^i(\theta)$ is determined as the solution to the following fixed point problem:

$$C_d^i(\theta) = \frac{\Xi^i}{A_d^i}[w_d]^{\gamma^{iL}} \prod_j \Big[\sum_o \int [\tau_{do}^j C_o^j(\tilde{\theta}) t(\theta, \tilde{\theta})]^{-\zeta} \mathrm{d}\Theta_o^j(\tilde{\theta}) \Big]^{-\frac{\gamma^{ij}}{\zeta}}, \tag{5}$$

where Ξ^i is a sector-specific constant, and Θ_o^j is the probability measure that describes the distribution of chosen technologies across firms in (o, j).[12]

Proof. See Appendix A.1. □

To prove this result, we show that if supplier price distributions in each (o, j) are Weibull, then the induced distribution of factory-gate prices for outputs is also Weibull. Because technology compatibility affects input sourcing, firms with different θ have different production cost distributions. We denote by $C_d^i(\theta)$ the location parameter of the cost distribution for firms in (d, i) which have chosen technology θ. In equilibrium, each $C_d^i(\theta)$ depends on all $\{C_o^j(\tilde{\theta}) : \tilde{\theta} \in \mathbb{T}\}_{o=1, j=1}^{N,S}$ through the fixed-point relationship (5).

Equation (5) reveals the forces shaping the price distribution in (d, i). Standard elements in trade models with input-output linkages appear: (d, i)'s effective sectoral productivity Ξ^i / A_d^i, country d's wage w_d, iceberg transport costs $\{\tau_{do}^j\}_{o=1, j=1}^{N,S}$, and the supplier price parameters $\{C_o^j(\tilde{\theta}) : \tilde{\theta} \in \mathbb{T}\}_{o=1, j=1}^{N,S}$. The novel compatibility effect enters via $t(\theta, \tilde{\theta})$. In the presence of match-specific efficiency draws, a firm with technology θ may end up buying from a supplier in any country o with any technology $\tilde{\theta}$. Consequently, $C_d^i(\theta)$ depends on the integral over the entire domain of Θ_o^j.

Because of the compatibility incentive, firms' sourcing decision is a function of their technology θ. We characterize the sourcing decision as a corollary of Proposition 1:

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

[12] Specifically, $\Xi^i \equiv \left(\int_0^\infty \int_0^\infty ... \int_0^\infty \mathbb{I}[\prod_j [m^j]^{\frac{\gamma^{ij}}{\zeta}} \leq \kappa] \prod_j [\Gamma(1 - \zeta/\lambda)]^{\frac{\gamma^{ij}}{\zeta}} \exp(-m^j) \lambda \kappa^{-\lambda-1} \mathrm{d}m^1 ... \mathrm{d}m^S \mathrm{d}\kappa \right)^{-1/\lambda}$, which is finite provided that $\max\{\max_{ij} \gamma^{ij}\lambda, 1\} < \zeta < \lambda$. As long as this inequality is satisfied, the exact value of λ does not affect our quantitative exercises.

1. *The expenditure share of the firm's input j produced by firms in o with technology $\tilde{\theta}$ is*

$$\chi^j_{do}(\theta,\tilde{\theta})\mathrm{d}\Theta^j_o(\tilde{\theta}) = \frac{[\tau^j_{do}C^j_o(\tilde{\theta})t(\theta,\tilde{\theta})]^{-\zeta}\cdot \mathrm{d}\Theta^j_o(\tilde{\theta})}{\sum_{o'}[\tau^j_{do'}\Lambda^j_{o'}(\theta)]^{-\zeta}}, \tag{6}$$

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

$$\Lambda^j_o(\theta) \equiv \left(\int [C^j_o(\tilde{\theta})t(\theta,\tilde{\theta})]^{-\zeta}\mathrm{d}\Theta^j_o(\tilde{\theta})\right)^{-1/\zeta}. \tag{7}$$

2. *The expenditure share of the firm's input j produced by firms in country o (across all $\tilde{\theta}$) is*

$$\chi^j_{do}(\theta) = \int \chi^j_{do}(\theta,\tilde{\theta})\mathrm{d}\Theta^j_o(\tilde{\theta}) = \frac{[\tau^j_{do}\Lambda^j_o(\theta)]^{-\zeta}}{\sum_{o'}[\tau^j_{do'}\Lambda^j_{o'}(\theta)]^{-\zeta}}. \tag{8}$$

Proof. See Appendix A.1. □

Equation (6) takes a similar form to the trade share in Eaton and Kortum (2002). The numerator reflects the likelihood that suppliers in (o,j) with technology $\tilde{\theta}$ are chosen by a firm in d with technology θ; suppliers are more likely to be chosen if they offer lower compatibility-adjusted costs (lower $\tau^j_{do}C^j_o(\tilde{\theta})t(\theta,\tilde{\theta})$) or if they are more numerous (larger $\mathrm{d}\Theta^j_o(\tilde{\theta})$). The denominator aggregates over all sector-j suppliers globally that can be accessed by a firm in (d,i) with technology θ, where $\Lambda^j_o(\theta)$ measures the competitiveness of suppliers in (o,j). The integration in (7) reflects that a country-sector (o,j) is more competitive as a source if it has more firms with low compatibility-adjusted costs. The second part of the corollary then aggregates across all firms in (o,j) to derive the overall probability that a firm in (d,i) with technology θ sources input j from country o.

The model described above generalizes Caliendo and Parro (2015) to incorporate flexible firm-level sourcing decisions. In particular, if we impose a single technology for all firms and countries (i.e., make θ exogenous and homogeneous), then $C^i_d(\theta)$ defined in (5) collapses to a scalar $\bar{C}^i_d$, $\Lambda^j_o(\theta)$ defined in (7) also reduces to $\bar{C}^j_o$, and the trade shares $\chi^j_{do}(\theta)$ no longer depend on θ—just as in Caliendo and Parro (2015). Thus, our model preserves the flexibility and tractability of canonical quantitative trade models while allowing for rich firm-level sourcing decisions that take into account compatibility requirements.[13]

Importantly, beyond technology compatibility, by letting θ represent a firm's attributes and $t(\theta,\tilde{\theta})$ capture interactions among different firm types, the trade block of our model can, in principle, accommodate any firm heterogeneity. This makes the framework applicable to other settings where allowing sourcing patterns to depend on firm-level traits is crucial.

[13]Tractability stems from Poisson assumptions on $A(\nu,r)$ and $z(\nu,\omega)$. The assumption on $A(\nu,r)$ ensures $p(\nu,r)$ in (2) is Weibull if all $c^j(\nu,r)$ are Weibull; likewise, the assumption on $z(\nu,\omega)$ ensures $\tilde{c}^j(\nu,\omega)$ in (3) is Weibull if $p(\omega)$ is Weibull. The Weibull distributions for prices are useful as they can be summarized by a small set of parameters (e.g., $C^i_d(\theta)$ in (4)) and deliver shares as in the standard Eaton and Kortum setup.

2.3 Production and Sourcing Decisions: Final-Good Firms

Final-good producers face constant monopolistic competitive markups $\frac{\eta}{\eta-1}$ when purchasing from intermediate-good producers. Facing these markups, they choose input from each supplier to maximize their profits:

$$P_d Q_d - \sum_j \sum_o \int_0^1 [\frac{\eta}{\eta-1}\tau_{do}^{Uj} p_o^j(\omega)] q_{do}^j(\omega) \mathrm{d}\omega, \tag{9}$$

where the factory-gate price $p_o^j(\omega)$ follows the distribution described by (4), τ_{do}^{Uj} is the iceberg trade cost for final users (superscript 'U'), which may differ from the τ_{do}^j faced by intermediate firms, and P_d is the price of the final good in d,

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

The sourcing decision of final-good producers differ from that of intermediate-good producers in two aspects. First, instead of relying on a single supplier per sector, final-good producers purchase from all suppliers because they benefit from love-of-variety. Second, final-good producers are unaffected by technology compatibility. This latter assumption captures the idea that consumers ultimately care more about the functionality of products than the underlying technological details. It also aligns with our quantification strategy using patent data.

We characterize the final-good producer's sourcing decision in Corollary 2.

Corollary 2. *For the final-good producer in country d, when purchasing goods from sector j:*

1. *The expenditure share allocated to goods produced by firms in country o with technology $\tilde{\theta}$ is*

$$\pi_{do}^j(\tilde{\theta}) \mathrm{d}\Theta_o^j(\tilde{\theta}) = \frac{[\tau_{do}^{Uj} C_o^j(\tilde{\theta})]^{1-\eta}}{\sum_{o'} [\tau_{do'}^{Uj} \bar{\Lambda}_{o'}^j]^{1-\eta}} \mathrm{d}\Theta_o^j(\tilde{\theta}), \tag{10}$$

$$\textit{where} \quad \bar{\Lambda}_o^j \equiv \big(\int [C_o^j(\tilde{\theta})]^{1-\eta} \mathrm{d}\Theta_o^j(\tilde{\theta})\big)^{1/(1-\eta)}. \tag{11}$$

2. *The expenditure share allocated to goods produced by firms in country o is*

$$\pi_{do}^j = \int \pi_{do}^j(\tilde{\theta}) \mathrm{d}\Theta_o^j(\tilde{\theta}) = \frac{[\tau_{do}^{Uj} \bar{\Lambda}_o^j]^{1-\eta}}{\sum_{o'} [\tau_{do'}^{Uj} \bar{\Lambda}_{o'}^j]^{1-\eta}}. \tag{12}$$

Proof. See Appendix A.1. □

Here, $\bar{\Lambda}_o^j$ reflects the overall competitiveness of intermediate-good firms in (o, j) when serving final-good producers. Because final-good producers do not consider technology compatibility, $t(\theta, \tilde{\theta})$ does not appear in $\bar{\Lambda}_o^j$.

2.4 Technology Choice

Upon entry, each firm in country-sector (d, i) draws an endowment technology $\bar{\theta}$ from an exogenous distribution characterized by probability measure $\bar{\Theta}_d^i$ on $\mathbb{T}$, which we label as the *ex-ante* technology distribution. The firm then chooses a production technology θ and adapts its production process accordingly before making sourcing decisions. Adapting to a distant technology from the endowed one is costly but may increase profits through better compatibility with efficient suppliers.

Recall that firms earn profits solely through sales to final-good producers. Let $X_d^i(\theta)$ be the *expected* sales to final-good producers by a firm in (d, i) choosing technology θ, where the expectation is over possible draws of production techniques and suppliers. We specify the firm's technology adaptation cost as:

$$K_d^i(\theta; \bar{\theta}) \equiv \phi(\theta, \bar{\theta}) \cdot \frac{1}{\eta} X_d^i(\theta), \tag{13}$$

where $\frac{1}{\eta} X_d^i(\theta)$ represents the firm's expected profits under a monopolistic markup, and $\phi(\theta, \bar{\theta})$ denotes the fraction of profits spent on adapting to technology $\theta \neq \bar{\theta}$. The function $\phi(\theta, \bar{\theta})$ increases with the distance between θ and $\bar{\theta}$, capturing the idea that more distant technologies are costlier to adopt.[14]

Firms choose θ to maximize the expected net profits:

$$\Pi_d^i(\theta; \bar{\theta}) \equiv \max_{\theta} [1 - \phi(\theta, \bar{\theta})] \cdot \frac{1}{\eta} X_d^i(\theta). \tag{14}$$

Under monopolistic competition in sales to final-good producers, whose demand is given by (10), this is equivalent to:

$$\max_{\theta} [1 - \phi(\theta, \bar{\theta})] \times [C_d^i(\theta)]^{1-\eta}. \tag{15}$$

This formulation highlights the trade-off faced by a firm: choosing between a good technology (i.e., the one with low $C_d^i(\theta)$) and a close one. Although by themselves technologies are inherently horizontally differentiated, the ones shared by the most efficient suppliers end up vertically better. In the presence of productivity differences between countries, this means technologies adopted by leading countries tend to be vertically 'better.'

Assuming that the solution to (15) can be described by a policy function, $g_d^i(\bar{\theta})$ (for which Proposition 2 provides the existence condition), we can construct the following mapping between *ex-ante* technology distribution $\bar{\Theta}_d^i$ and the *ex-post* distribution:

$$\Theta_d^i(\mathcal{B}) = \int_{\bar{\theta} \in \mathbb{T}} \mathbb{I}[g_d^i(\bar{\theta}) \in \mathcal{B}] \mathrm{d}\bar{\Theta}_d^i(\bar{\theta}), \quad \text{for any measurable set } \mathcal{B}. \tag{16}$$

[14] We model the adaptation cost as a share of expected profit. This could be micro-founded in a model of bargaining, in which researchers bargain with the firm founder to split the profit. If the founder knows less about more distant technologies, researchers can extract a larger share of the rents. Since we do not formally introduce a research sector in the model, we treat adaptation costs as expenses.

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

Definition 1. ***Equilibrium for Technology Choice.*** *Given the model primitives and countries' wages $\{w_d\}_{d=1}^N$, an equilibrium for technology choice consists of the policy functions that describe firms' technology choice, $g_d^i : \mathbb{T} \to \mathbb{T}$, and the cost functions $C_d^i : \mathbb{T} \to \mathbb{R}^+$ for all (d, i), such that*

(1) Given $\{C_d^i\}$, $g_d^i(\bar{\theta})$ solves problem (15) for all (d, i) and $\forall \bar{\theta} \in \mathbb{T}$ except on a zero-measure set.

(2) Given $\{g_d^i\}$, $\{C_d^i\}$ satisfies equation (5), where $\{\Theta_d^i\}$ is given by (16).

The first condition ensures that firms' chosen technologies are optimal given the equilibrium distribution of factory-gate prices. The second condition guarantees that those choices, when aggregated, are consistent with the cost distributions C_d^i.

Unlike in standard quantitative trade models, where similar conditions can be aggregated over firms and expressed solely in terms of aggregate prices and quantities,[15] in our setting, due to the bilateral nature of compatibility requirement, each firm's choice depends on the entire distribution of other firms' choices. We develop a procedure to characterize the conditions for the existence and uniqueness of equilibrium in such environments, which can be adapted to other settings where interactions between firms' decisions cannot be summarized by aggregate sufficient statistics. Our characterization proceeds in three steps:

First, we formulate equilibrium conditions involving policy functions $\{g_d^i\}$—as in Definition 1—rather than involving distributions $\{\Theta_d^i\}$. Doing so sidesteps the technical difficulties of applying fixed-point theorems to spaces of probability measures. It also allows us to focus on equilibria where technology choices vary smoothly with endowment technologies.[16]

Second, we stack the equilibrium objects to construct a joint mapping from their space (i.e., the space of $(\{g_d^i\}, \{C_d^i\})$) to itself. This differs from alternative methods viewing the equilibrium as a nested fixed point system—e.g., mapping from the space of $\{g_d^i\}$ to itself, nesting inside a fixed-point mapping for $\{C_d^i\}$ (see, e.g., Alvarez and Lucas, 2007 for an example of the nested fixed point approach in trade models). Our approach avoids the often challenging characterization of comparative statistics in the inner mapping necessary for establishing properties of the outer mapping. It shares the same spirit as Allen et al. (2024), but here, the equilibrium objects are infinite-dimensional. We illustrate how to choose appropriate norms and define spaces suitable for applying the Schauder fixed-point theorem to establish equilibrium existence.

Third, to establish uniqueness, we derive the Fréchet derivatives of the constructed joint mapping defined on functional space. We then determine the conditions under which the

[15] For instance, in Caliendo and Parro (2015), the first condition reduces to a system of equations for production costs; in Chaney (2008), firms' export decisions depend on the entire distribution of firm choices, which can be analytically aggregated into price indexes.

[16] Our characterization focuses on equilibria with policy functions that are bounded, continuously differentiable, and have Lipschitz-continuous first derivatives.

Jacobian matrix of these Fréchet derivatives has a row sum norm bounded by a constant below 1, which allows us to apply the contraction mapping theorem. The procedure for deriving and bounding these derivatives is broadly applicable to other models featuring similar interdependencies among firm choices.

Our procedure works for any generic metric space $\mathbb{T}$. To simplify the algebra, however, we impose the following assumption:

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

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

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

The first part of Assumption 2 takes a one-dimensional representation of the technology space. The second part states that both the compatibility and adaptation costs increase in the technological distance $(\theta - \tilde{\theta})^2$ with constant elasticities, $\bar{t}/2$ and $\bar{\phi}/2$.

There are two aspects to the assumption $\mathbb{T} = \mathbb{R}$: dimensionality and geometry. Regarding dimensionality, our theoretical results generalize to $\mathbb{T} = \mathbb{R}^n$ for a finite n.[17] In an $n > 1$ setting, any technology can be viewed as a combination of different characteristics, with each dimension representing one characteristic. We focus on $\mathbb{T} = \mathbb{R}$ because it suffices in conveying the intuition while substantially reducing computational burden in quantification. We discuss additional implications of this assumption on the calibration outcomes and quantitative results in Section 4.

Regarding geometry, the technology space does not have to be a line—one could use a circle, a ball, or other higher-dimensional manifolds. Different shape assumptions matter for firm choices.[18] Given the lack of evidence supporting one versus another, we choose a linear representation because it allows Gaussian distributions, which, as shown later, yield tractable solutions in our quantitative exercise.

Under Assumptions 1 and 2, we establish the following results on existence and uniqueness of the technology choice equilibrium.

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

1. *Let* $\{\bar{\Theta}^i_d\}$ *have bounded support within* $[-M, M]$ *with 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. In this equilibrium, the choice of firms from* (d, i) *with endowment technology*

[17] For $\mathbb{T} = \mathbb{R}^n$, the cost functions in part (2) of Assumption 2 take a quadratic form in the exponent, e.g., $t(\theta, \tilde{\theta}) = \exp(\frac{1}{2}(\theta - \tilde{\theta})' \cdot \bar{t} \cdot (\theta - \tilde{\theta}))$, where x' is the transpose of a vector x and $\bar{t}$ here is a positive definite *matrix*.

[18] For example, on a line, when a firm moves left, it moves closer to firms on its left while moving away from those on its right. In contrast, on a circle, firms moving in opposite directions eventually meet.

$\bar{\theta}$ is characterized by the following first-order condition, which has a unique solution:

$$g_d^i(\bar{\theta}) = (1-\omega^i)\bar{\theta} + \omega^i \sum_{j,o} \frac{\gamma^{ij}}{1-\gamma^{iL}} \int \chi_{do}^j(g_d^i(\bar{\theta}), \tilde{\theta}) g_o^j(\tilde{\theta}) \mathrm{d}\Theta_o^j(\tilde{\theta}), \quad \forall \bar{\theta} \in [-M, M], \tag{17}$$

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

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

Proof. See Appendix A.2. □

For an equilibrium with continuous policy functions to exist, firms' technology choices should not be overly sensitive to their own endowment technologies, which is ensured if $\zeta\bar{t}M^2 < 1$. Here, M^2 sets a limit on the variance of potential supplier technologies, which determines how much firms can improve compatibility through technology choice.[19] $\bar{t}$ governs how fast production costs rise with technological distance, and ζ is the trade elasticity shaping how price advantages translate to profits. Together, these three parameters bound the sensitivity of firms' technology choice to their endowments.

The first order condition shows that firms balance staying close to their endowment $\bar{\theta}$ (weighted by $1-\omega^i$) against aligning with suppliers' average technology (weighted by ω^i and the importance of suppliers $\frac{\gamma^{ij}}{1-\gamma^{iL}}\chi_{do}^j(g_d^i(\bar{\theta}), \tilde{\theta})$).

The equilibrium for technology choice is not generically unique. In particular, in the limit as $\bar{\phi} \to 0$ or $\bar{t} \to \infty$, the incentive for compatibility outweighs the cost of adaptation, leading all firms to coordinate on the same location—and any such location can be an equilibrium. To establish uniqueness, we derive the conditions under which the joint mapping described earlier is a contraction. In the condition $\bar{t} < \frac{1}{2M}$, $2M$ enters as the upper bound of the distance between a firm's technology and the expenditure-weighted average technology of its suppliers. Intuitively, $\bar{t} \cdot 2M$ determines the scope for compatibility benefits. If these benefits are not too large, and if the cost of technology adaptation $\bar{\phi}$ is not too small, then firms' technology would not be overly sensitive to *other firms*' choices, thereby ensuring uniqueness.

This intuition behind the bound for $\bar{t}$ means that even if M is arbitrarily large, as long as firms' expenditure shares on distant suppliers are not too large, uniqueness can be ensured without $\bar{t} \to 0$. More generally, Proposition 2 requires bounded technological space because we allow for arbitrary endowment distributions. This assumption can be relaxed by restricting the endowment distribution to particular shapes. Indeed, under our calibration, the equilibrium under Normal distributions on ex-ante distributions exists and is unique.

[19] By Popoviciu's inequality on variances, $\text{var}(\theta) < M^2$ for any distribution of θ over $[-M, M]$. Larger variance means greater heterogeneity in supplier technologies, implying larger room for firms to improve supplier compatibility, hence leading to more sensitive technology choice with respect to endowed technology.

2.5 General Equilibrium

We now embed firms' technology choice equilibrium into a general equilibrium. For a firm in (o, j) using technology $\tilde{\theta}$, let $M_o^j(\tilde{\theta})$ denote its expected sales *to downstream firms*. Market clearing in intermediate goods, stated in value of goods, reads:[20]

$$M_o^j(\tilde{\theta}) = \sum_d \sum_i \int \gamma^{ij} \left[M_d^i(\theta) + (1 - \frac{1}{\eta}) X_d^i(\theta) \right] \chi_{do}^j(\theta, \tilde{\theta}) \mathrm{d}\Theta_d^i(\theta), \tag{18}$$

where $M_d^i(\theta)$ is the expected sales to other firms by a firm in (d, i) with technology θ, $X_d^i(\theta)$ is that firm's expected sales to final producers. The sum $\left[M_d^i(\theta) + (1 - \frac{1}{\eta}) X_d^i(\theta) \right]$ equals the expected production cost of that firm (given its $\frac{\eta}{\eta - 1}$ markup to final producers), γ^{ij} is that firm's expenditure share on input from sector j, and $\chi_{do}^j(\theta, \tilde{\theta})$ is its likelihood of purchasing from a firm in (o, j) with technology $\tilde{\theta}$.

The expected sales of a firm with technology $\tilde{\theta}$ to *final-good producers*, $X_o^j(\tilde{\theta})$, satisfy

$$X_o^j(\tilde{\theta}) \equiv \sum_d \rho_d^j P_d Q_d \pi_{do}^j(\tilde{\theta}). \tag{19}$$

Here, $P_d Q_d$ is the final good demand from household consumption and adaptation costs:

$$P_d Q_d = I_d + \sum_i \int K_d^i(g_d^i(\bar{\theta}); \bar{\theta}) \mathrm{d}\bar{\Theta}_d^i(\bar{\theta}), \tag{20}$$

with technology adaptation costs $K_d^i(\theta; \bar{\theta})$ defined in (13), and g_d^i the policy function for technology choice as part of the technology choice equilibrium (Definition 1).

Household income consists of wages and the net profits of domestic firms:

$$I_d = w_d L_d + \sum_i \int \Pi_d^i(g_d^i(\bar{\theta}); \bar{\theta}) \mathrm{d}\bar{\Theta}_d^i(\bar{\theta}), \tag{21}$$

where net profits $\Pi_d^i(\theta; \bar{\theta})$ are defined in (14), and wages are determined by the labor market clearing condition, stated in expenditure:

$$w_d L_d = \sum_i \gamma^{iL} \int [M_d^i(\theta) + (1 - \frac{1}{\eta}) X_d^i(\theta)] \mathrm{d}\Theta_d^i(\theta). \tag{22}$$

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 (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) policy function for technology choice $\{g_d^i(\bar{\theta})\}$ and production cost function $\{C_d^i(\theta)\}$, such that*

i. Given wages $\{w_d\}$, $\{g_d^i(\bar{\theta}), C_d^i(\theta)\}$ constitute a technology choice equilibrium defined in Defi-

[20] Or, in the integral form: for any set $\mathcal{B}$,

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

nition 1, and $\{\Theta_o^j\}$ is consistent with equation (16).

ii. *Goods and labor markets clear, i.e., (20), (21), (22) hold by d, and (18) and (19) hold by $(o, j, \tilde{\theta})$.*

Having defined the general equilibrium, we use special cases to analytically examine the interaction between technology choice and trade, and the resulting welfare implications.

2.6 Special Cases: Interaction between Technology Choice and Trade

Throughout this subsection, we assume that each country–sector pair (d, i) has a degenerate ex-ante technology distribution with a unit mass at $\bar{\theta}_d^i$. We focus on the equilibrium for technology choice (i.e., taking wages as given) that is symmetric, meaning that all firms from a country-sector pair (d, i) with the same endowment technology choose the same technology.

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

1. *The technology chosen by a firm in (d, i) is a weighted average between the firm's endowment technology $\bar{\theta}_d^i$ and the technologies chosen by its suppliers $\{\theta_o^j\}$:*

$$\theta_d^i = (1-\omega^i)\bar{\theta}_d^i + \omega^i \sum_{j,o} \frac{\gamma^{ij}}{1-\gamma^{iL}} \bar{\chi}_{do}^{ij} \theta_o^j, \tag{23}$$

where $\omega^i \equiv \frac{(\eta-1)(1-\gamma^{iL})\bar{t}}{(\eta-1)(1-\gamma^{iL})\bar{t}+\bar{\phi}}$, and $\bar{\chi}_{do}^{ij} \equiv \frac{\exp[-\zeta(\ln \tau_{do}^j + \ln \bar{C}_o^j + \frac{1}{2}\bar{t}(\theta_d^i - \theta_o^j)^2)]}{\sum_{o'} \exp[-\zeta(\ln \tau_{do'}^j + \ln \bar{C}_{o'}^j + \frac{1}{2}\bar{t}(\theta_d^i - \theta_{o'}^j)^2)]}$ is the share of sector-j input spent on country o by intermediate-good firms in (d, i).

2. *Let $\bar{\mathbf{C}} \equiv (\bar{C}_1^1, \bar{C}_1^2, ..., \bar{C}_N^S)$ denote the location parameters of factory-gate price distributions (i.e., the degenerate special case of the location parameters defined in (5)). Then between two equilibria with different trade costs and technology choices:*

$$\mathrm{d}\ln \bar{\mathbf{C}} = \mathbb{G}_C^\tau \cdot \mathrm{d}\ln \boldsymbol{\tau} + \mathbb{G}_C^\theta \cdot \mathrm{d}\boldsymbol{\theta},$$

where $\mathrm{d}\boldsymbol{\theta} \equiv (\mathrm{d}\theta_1^1, \mathrm{d}\theta_1^2, ..., \mathrm{d}\theta_N^S)$ and $\mathrm{d}\ln \boldsymbol{\tau} \equiv (\mathrm{d}\ln \tau_{11}^1, \mathrm{d}\ln \tau_{11}^2, ..., \mathrm{d}\ln \tau_{12}^1, \mathrm{d}\ln \tau_{12}^2, ..., \mathrm{d}\ln \tau_{NN}^S)$.

$\mathbb{G}_C^\tau$ is an $NS \times N^2S$ matrix that depends on the following sufficient statistics: ζ, the input-output parameters, and trade shares.

$\mathbb{G}_C^\theta$ is an $NS \times NS$ matrix that depends, additionally, on $\bar{t}$ and the expenditure-share-weighted average differences in $\{\theta_d^i\}$ across country-sectors.

3. *In response to changes in trade costs $\{\mathrm{d}\ln \tau_{do}^j\}$, firms' technology choice changes according to:*

$$\mathrm{d}\boldsymbol{\theta} = \mathbb{G}_\theta^\tau \cdot \mathrm{d}\ln \boldsymbol{\tau},$$

where $\mathbb{G}_\theta^\tau$ is a matrix that depends on $\eta, \bar{\phi}$, and all the sufficient statistics stated in part 2.

Proof. See Appendix A.3. The proof also presents these coefficient matrices explicitly. □

Equation (23) is a special case of (17). The difference here is that because firms within a (d,i) are homogeneous, they have the same trade shares $\bar{\chi}^{ij}_{do}$.

Part 2 of the proposition shows how changes in trade costs and technology distances influence the production cost distributions. Intuitively, firms' production costs increase either if their import costs increase (captured by $\mathrm{d}\ln\boldsymbol{\tau}$) or if they become more distant to suppliers in the technological space (captured by $\mathrm{d}\boldsymbol{\theta}$). Both have direct and indirect effects, captured by input expenditure shares and trade shares. Both effects are larger if inputs are less substitutable across suppliers—hence the dependence of $\mathbb{G}^{\tau}_{\theta}$ and $\mathbb{G}^{\tau}_{C}$ on intermediate-good trade elasticity ζ. The marginal effect of technological change further depends on $\bar{t}$—how fast incompatibility costs rise with technology distance—and existing technology configurations in the economy.

Part 3 shows that the effect of trade costs on $\boldsymbol{\theta}$ can be summarized by a coefficient matrix $\mathbb{G}^{\tau}_{\theta}$, which, in addition to the set of sufficient statistics stated in Part 2, also depends on the final demand elasticity η and adaptation cost parameter $\bar{\phi}$. Firms' technology choice balance the benefits of reducing production costs and the costs of adaptation. The benefits, according to Part 2, depend on all the sufficient statistics stated there and η, which determines how production costs translate into profits, whereas the costs depend crucially on $\bar{\phi}$.

Appendix A.3 presents all the coefficient matrices explicitly. To sharpen the intuition, we next consider a 'small' country-sector (d,i), whose technology choice does not affect the rest of countries and sectors.

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. Following an x % increase in the cost of (d,i) importing from (o,j):*

1. *The squared distance between θ^i_d and θ^j_o changes by:*

$$\Delta\|\theta^i_d-\theta^j_o\| = -\underbrace{\frac{\zeta\omega^i\gamma^{ij}\bar{\chi}^{ij}_{do}\|\theta^j_o-\vartheta^{ij}_d\|}{1-t\zeta\omega^i\sum_{j',o'}\gamma^{ij'}\bar{\chi}^{ij'}_{do'}\|\theta^{j'}_{o'}-\vartheta^{ij'}_d\|}}_{>0\text{ by firms' second-order condition}}\times\frac{\theta^i_d-\theta^j_o}{\theta^j_o-\vartheta^{ij}_d}\times x,$$

   where ω^i and $\bar{\chi}^{ij}_{do}$ are defined in Proposition 3, and $\vartheta^{ij}_d \equiv \sum_m \bar{\chi}^{ij}_{dm}\theta^j_m$ is the weighted average technology of sector-j suppliers to (d,i).

2. *$\|\theta^i_d-\theta^j_o\|$ increases relative to the expenditure-share weighted distance between θ^i_d and $\theta^j_{o'}$ across $o'=1,...,N$. Formally,*

$$\Delta\|\theta^i_d-\theta^j_o\| - \sum_{o'}\bar{\chi}^{ij}_{do'}\Delta\|\theta^i_d-\theta^j_{o'}\| = \frac{\zeta\omega^i\gamma^{ij}\bar{\chi}^{ij}_{do}\|\theta^j_o-\vartheta^{ij}_d\|}{1-t\zeta\omega^i\sum_{j',o'}\gamma^{ij'}\bar{\chi}^{ij'}_{do'}\|\theta^{j'}_{o'}-\vartheta^{ij'}_d\|}\times x>0. \quad (24)$$

Proof. See Appendix A.4. □

Part 1 of the proposition characterizes the impact of an increase in τ_{do}^{j} on $\|\theta_d^i - \theta_o^j\|$. As τ_{do}^{j} increases, firms in (d,i) increase their dependence on other suppliers, which increases the weight of these suppliers' average technology in θ_d^i. Whether this change increases or decreases $\|\theta_d^i - \theta_o^j\|$ depends on the sign of the ratio $(\theta_d^i - \theta_o^j)/(\theta_o^j - \vartheta_d^{ij})$, which in turn depends on the positions of θ_d^i and ϑ_d^{ij} relative to θ_o^j. If θ_d^i and ϑ_d^{ij} lie on the same side of θ_o^j, then (d,i) firms' putting more weights on other suppliers reduces their incentive to shift to θ_o^j, increasing $\|\theta_d^i - \theta_o^j\|$; conversely, $\|\theta_d^i - \theta_o^j\|$ may shrink.

Despite this ambiguity, Part 2 establishes that $\|\theta_d^i - \theta_o^j\|$ *always* increases *relative to* the expenditure-share-weighted average distance between θ_d^i and the technologies of all suppliers. The increase in the relative distance is larger if adaptation is less costly (smaller $\bar{\phi}$, which enters via ω^i), if (d,i) relies more heavily on (o,j) for inputs (large $\gamma^{ij}\bar{\chi}_{do}^{ij}$), or if θ_o^j is farther from the average technology of suppliers (larger $\|\theta_o^j - \vartheta_d^{ij}\|$). We will leverage this intuition to discipline the model for quantification.

Thus far, we have focused on how changes in trade costs influence technology choices. As discussed previously, firms' endowment technology also affects their choice, which in turn affects trade. To illustrate this mechanism, we consider a change in the endowed technology of a zero-measure set of firms from the rest of the firms in (d,i).

Proposition 5. *Suppose all firms in (d,i) share the same endowment technology $\bar{\theta}_d^i$, except for a zero measure set ν endowed with $\bar{\theta}(\nu)$. If $\bar{\theta}(\nu)$ moves closer to θ_o^j relative to $\theta_{o'}^j$, then*

1. *Firm ν chooses a technology closer to θ_o^j relative to $\theta_{o'}^j$, namely, $\|\theta_d^i(\nu) - \theta_o^j\| - \|\theta_d^i(\nu) - \theta_{o'}^j\|$ decreases.*

2. *Firm ν is more likely to source intermediate goods from (o,j) relative to (o',j); the change in the likelihood of sourcing from (o,j) relative to from (o',j) satisfies: $\Delta \log\left(\chi_{do}^{ij}(\nu)/\chi_{do'}^{ij}(\nu)\right) = -\frac{1}{2}\zeta\bar{t} \times \Delta\left(\|\theta_d^i(\nu) - \theta_o^j\| - \|\theta_d^i(\nu) - \theta_{o'}^j\|\right)$.*

Proof. See Appendix A.5. □

The proposition indicates that as a firm's endowed technology moves toward θ_o^j, so does its chosen technology. This, in turn, makes the firm more likely to source from (o,j). The response of sourcing patterns to the technology distance between trade partners is governed by $\bar{t}$ as well as the conventional trade elasticity ζ. Conditioning on ζ, this response identifies $\bar{t}$, an intuition we will use in quantification to pin down $\bar{t}$.

2.7 Special Cases: The Welfare Implications of Technology Choice

In our model, firms choose technology to maximize their own profits, but their choices impact the profits of downstream firms, too. Such externalities propagate through input-output and trade linkages, leading to potential inefficiencies. We illustrate these welfare implications using two special cases that separately highlight domestic and international spillovers. Throughout this subsection, each country-sector pair (d, i) has a degenerate endowment technology distribution.

Proposition 6. *Consider a closed economy (hence dropping country subscripts) with multiple sectors and each sector with an ex-ante endowment technology* $\{\bar{\theta}^i\}_{i=1}^S$.

1. *The marginal impact of increasing* θ^i *on household real income (denoted by* $U \equiv \frac{I}{P}$*) is given by*

$$\frac{\Delta \ln(U)}{\Delta \theta^i} = \rho^i \Big[\underbrace{\frac{\exp\big(-\frac{1}{2}\bar{\phi}(\theta^i - \bar{\theta}^i)^2\big)}{\eta - \sum_i \rho^i \exp\big(-\frac{1}{2}\bar{\phi}(\theta^i - \bar{\theta}^i)^2\big)} \bar{\phi}(\bar{\theta}^i - \theta^i)}_{\text{income effect}} - \underbrace{\bar{t} \sum_j \tilde{\gamma}^{ij}(\theta^i - \theta^j)}_{\text{sector-i price}} \Big] - \underbrace{\bar{t} \sum_{j \neq i} \rho^j \tilde{\gamma}^{ji}(\theta^i - \theta^j)}_{\text{other sector prices}},$$

   where $\tilde{\gamma}^{ij} \equiv \sum_m \Omega^{im} \gamma^{mj}$, *and* Ω^{im} *is the* (i, m)*-th element of* $(\mathbb{I}_{S \times S} - \Gamma)^{-1}$, *with* $\Gamma^{ij} \equiv \gamma^{ij}$.

2. *Suppose that all sectors have the same weight in final-good consumption and a symmetric input-output structure, i.e., for all* $i \neq j \neq j'$, $\rho^i = \rho^j$, $\gamma^{ii} = \gamma^{jj}$ *and* $\gamma^{ij} = \gamma^{ij'} = \gamma^{jj'}$. *In equilibrium,* $||\theta^i - \bar{\theta}^i||$ *is too small compared to the social planner's solution. In other words, firms under-invest in technology adaptation.*

Proof. See Appendix A.6. □

Part 1 illustrates the mechanisms through which $\Delta\theta^i$ affects welfare. Without loss of generality, suppose that in industry i, firms shift to the left from their endowed technology, i.e., $\theta^i < \bar{\theta}^i$, presumably because they would like to use inputs from other sectors with technologies further to the left. A marginal increase in θ^i from the equilibrium choice creates three effects. First, it saves adaptation costs, which can be used by households for consumption (first term). Second, it causes a change in the price of sector i (second term), in which $\tilde{\gamma}^{ij}$ captures the general-equilibrium effect of the distance to sector-j's technology θ^j. Finally, as θ^i shifts, other sectors using sector-i output as their input may benefit or lose depending on whether their distance to θ^i increases or decreases. This effect is stronger if the benefit of technology compatibility is larger (larger $\bar{t}$) or if the linkage is stronger (larger $\tilde{\gamma}^{ji}$).

Contrasting this with firms' first-order condition that shapes their technology choice:

$$\rho^i \Big[\underbrace{\frac{1}{\eta - 1} \bar{\phi}(\bar{\theta}^i - \theta^i)}_{\text{adaptation costs}} - \underbrace{\bar{t} \sum_j \gamma^{ij}(\theta^i - \theta^j)}_{\text{production cost}} \Big] = 0, \tag{25}$$

we can see that, compared to the social effects of adaptation, individual firms (i) weigh the income effect differently (because firms' profits are only a share of household income), and (ii) weigh the sector i price effect differently (γ^{ij} instead of $\tilde{\gamma}^{ij}$ because they do not take general equilibrium into account). Moreover, individual firms do not internalize how their technology choice benefits downstream users (the third term in household welfare).

Because of these externalities, the decentralized equilibrium typically fails to maximize household welfare. Without further restrictions, the net effect of moving θ^i in any direction is ambiguous—when firms move closer to some suppliers, they also move away from others. The second part of the proposition shows that under symmetric input-output structures, the positive externality dominates, and firms always under-invest in adaptation and stay too close to their endowed technology.[21]

We next consider the international spillovers of technology choice.

Proposition 7. *Consider an open economy with one sector (hence dropping sector superscripts) with roundabout production and two symmetric countries, country* 1 *and country* 2*. Assume, without loss of generality, that in equilibrium* $\theta_2 < \theta_1$*. Then, the effect of increasing* θ_2 *from its equilibrium level (moving it closer to* θ_1*) on welfare is:*

$$\frac{\Delta \ln U_2}{\Delta \theta_2} = \frac{\exp(-\frac{1}{2}\bar{\phi}(\theta_2 - \bar{\theta}_2)^2)}{\eta - \exp(-\frac{1}{2}\bar{\phi}(\theta_2 - \bar{\theta}_2)^2)}\bar{\phi}(\bar{\theta}_2 - \theta_2) + \bar{t}\frac{1-\gamma^L}{\gamma^L}\bar{\chi}(\theta_1 - \theta_2) > 0,$$
$$\frac{\Delta \ln U_1}{\Delta \theta_2} = \bar{t}\frac{1-\gamma^L}{\gamma^L}\bar{\chi}(\theta_1 - \theta_2) > 0,$$

where $\bar{\chi} \equiv \frac{(\tau \exp(\frac{1}{2}\bar{t}(\theta_2 - \theta_1)^2))^{-\zeta}}{(\tau \exp(\frac{1}{2}\bar{t}(\theta_2 - \theta_1)^2))^{-\zeta} + 1}$ *is each country's expenditure share on foreign intermediate goods.*

Proof. See Appendix A.7. □

As in the closed-economy setting, firms' private incentive for technology adaptation does not fully internalize the broader benefits of compatibility, which here extends to the foreign country. Starting from the decentralized equilibrium, moving θ_2 closer to θ_1 lowers the production costs of both countries due to stronger compatibility, as countries use each other's output as input. Beyond this direct effect, lower production costs in country 1 benefit producers in country 2, and vice versa. As a result, *both* countries benefit from such a move. Because firms do not internalize the full extent of these benefits, they under-invest in adapting toward foreign technologies from the perspective of both foreign and domestic welfare.

This mechanism has implications for the impact of trade policies. If trade liberalization leads to technology convergence between the two countries, that convergence amplifies the resulting welfare gains. Conversely, if a trade conflict causes countries' technologies to diverge, that divergence amplifies the welfare losses.

This last result might seem surprising given the intuition that, when facing adverse shocks, agents' endogenous substitutions often mitigate the damage of the shock for them.

[21] We assume that intermediate-good firms sell to downstream producers under perfect competition. We conjecture that even under alternative arrangements (e.g., fixed markups or bilateral bargaining), technology choices would remain inefficient. This inefficiency stems from the ex-ante nature of the technology decision, which creates a hold-up problem during downstream transactions. Thus, unless firms can fully extract the rents from the gains of improved compatibility or offer ex-ante contracts governing both technology choice and sales, they will not fully internalize the externalities.

However, in our model, externalities lead to an inefficient initial equilibrium, where there is a first-order gap between social and private benefits of technology adaptation. In the presence of this gap, endogenous technological divergence leads to social losses. As a result, it exacerbates, rather than mitigates, the costs of trade conflicts.

3 Evidence

We now turn to the data, with two goals in mind. First, we test a key premise of our model, that compatibility with suppliers' technologies increases the likelihood of sourcing from those suppliers. We investigate this premise using firm-level Chinese customs data. Second, we examine the model's implication that declines in tariffs make importer and exporter technologies more similar. To test this, we exploit time-series changes in tariffs that affect certain MFN-bound exporting countries but not others. We leverage these two estimates to discipline the model for quantitative exercises.

3.1 Measuring Technology Similarity Using Patent Texts

Both of our empirical exercises require a measure of compatibility between technology pairs. In the model, technologies are compatible if they are close to each other in the technology space. We use a proxy for this by constructing similarity measures from patent texts. Patent texts offer two advantages for our purpose. First, patents are embodiments of new technologies by design, and their texts are intended to accurately summarize the technical features of inventions. Second, patent data are available globally, facilitating comprehensive measurement across countries and sectors.

Measure Construction. We use the 2023 Fall release of *PATSTAT Global*, a database maintained by the European Patent Office (EPO) that covers close to the universe of world patents. To ensure consistency in language processing, we focus on patents with English abstracts. To avoid double counting inventions patented in multiple offices (e.g. European, U.S. etc.), we keep only one patent per patent family. This yields around 45 million patent abstracts.[22]

We process these English abstracts using a leading large language model developed by Alibaba ('gte-base-en-v1.5'), which transforms each paragraph into a 768-dimensional numerical vector—a process called 'text embedding.'[23] The embedding of a paragraph sum-

[22]Because the EPO prioritizes abstracts in English when sourcing the data, approximately 90% of all abstracts in PATSTAT are in English. Our final sample keeps a patent family as long as one of the patents in the family is in English, so the coverage is even higher.

[23]See `https://huggingface.co/Alibaba-NLP/gte-base-en-v1.5` for model descriptions. The model balances efficiency and quality. Deployed on a desktop GeForce RTX 4090 GPU, it produces the embeddings for 45 million patents in 2 days. In terms of quality, as of August 2024, it ranks 36 of 438 on the Massive Text Embedding Benchmark Leaderboard (Muennighoff et al., 2023). For comparison, the default model of SBERT ('all-mpnet-base-v2') ranks 282, and OpenAI's latest embedding model ('text-embedding-3-large-256') ranks 65.

marizes its meaning as a point in a 768-dimensional space. In the context of patents, we can view this space as representing all possible combination of technical ingredients and insights from different scientific disciplines to achieve a function. Closer points in the space represent patents that are more similar to each other.

To measure similarity between firms and countries, we take the average embedding of all patents produced by a given unit (firm or country) and compute the cosine similarity between units.[24] For example, let $\bar{c}_{\omega t}$ be the average embedding of all patents invented in firm ω in period t, and let $\bar{c}_{ot}$ be the average embedding of patents invented in country o. The similarity between firm ω and country o in period t is defined as:

$$\text{Similarity}_{\omega o t} = \frac{\bar{c}_{\omega t} \cdot \bar{c}_{ot}}{\|\bar{c}_{ot}\| \cdot \|\bar{c}_{\omega t}\|},$$

where the numerator is the inner product of the two vectors, and the denominator is the product of their norms. We construct analogous measures at other levels of aggregation, such as between country–sector pairs, and between firm and country-sector pairs. We assign patents to sectors and products (HS-6) using the crosswalks between the international patent classification and standard industry and product classifications by Lybbert and Zolas (2014).

Validation and Summary Statistics. In Appendix B.2, we provide examples to demonstrate that the embedding-based similarities capture technological compatibility between firms along the production chain. We also show that our measures correlate with plausible predictors of technology closeness. In particular, patents within the same class are more similar than those in different classes, sectors with stronger input-output linkages exhibit greater similarity, and geographically proximate countries also have more similar patents. These patterns suggest that the embeddings indeed capture meaningful variations in technology proximity.

Our empirical analyses use measures at three different levels. In Section 3.2, we examine whether firms are more likely to import from countries with technologies similar to their own. To do this, we compute technology similarity between Chinese firms (denoted ω) and either foreign countries ($\text{Similarity}_{\omega o t}$) or foreign country-sector pairs ($\text{Similarity}_{\omega o j t}$). In Section 3.3, we investigate the causal effect of trade barriers on technology proximity by exploiting changes in HS-6-level MFN tariffs. We tailor our empirical designs to the variations in such tariffs, measuring similarity at the level of the importing country, exporting country, HS-6 product, and time ($\text{Similarity}_{do,HS6,t}$).

Both exercises use data from 2000–2014, which we group into five three-year intervals. We aggregate countries in the World Input-Output Tables (WIOT) into 29 regions based on geographic and political proximity; see Table B.1 for the full list of regions. (In the rest of the

[24] An alternative would be to calculate pairwise similarity across individual patents and then aggregate. Given the large number of patents, this alternative is computationally infeasible; we verify on a smaller sample that the two methods produce strongly correlated metrics.

Table 1: Summary Statistics: Similarity between Firms

	Mean	Std. Dev.	Min	Max	Obs.
$\text{Similarity}_{\omega ot}$	0.564	0.090	0.071	0.965	5,147,016
$\text{Similarity}_{\omega ojt}$	0.506	0.110	-0.011	0.978	13,781,232
$\text{Similarity}_{do,HS6,t}$	0.694	0.112	0.171	0.977	5,720,642

Notes: $\text{Similarity}_{\omega ot}$ is the cosine similarity between the text embeddings of patents in period t from a Chinese firm ω and those from a foreign region o. $\text{Similarity}_{\omega ojt}$ and $\text{Similarity}_{do,HS6,t}$ are defined analogously. $\text{Similarity}_{\omega ojt}$ compares firm ω with a region-sector pair (o, j); $\text{Similarity}_{do,HS6,t}$ compares country d with patents in a specific HS-6 category in region o. Regions are classified as in Table B.1. Sectors are classified at the two-digit level of the International Standard Industrial Classification.

paper, we use 'regions' and 'countries' interchangeably.) Table 1 reports summary statistics for the key similarity measures used in our analysis.

Citation as an Alternative Measure. An often-used measure for technology similarity is patent citations. Because citation-based measures are also frequently associated with knowledge diffusion between the cited and citing patents (Jaffe et al., 1993; Liu and Ma, 2021), we have chosen to use the text-based measure as our preferred approach. Nevertheless, in Appendix B.2, we show that our text-based measure correlates with bilateral citations. Moreover, all empirical results hold when we use citations to measure technology proximity.

Relationship between Technology and Bilateral Trade. Figure 1 takes a first look at the relationship between bilateral trade and technology similarity using country-sector-level data. The vertical axis shows log exports from sector j of country o to sector i of country d. The horizontal axis plots the text-based technology similarity between (d, i) and (o, j). Included in the controls are bilateral distance metrics and income differences between d and o, and fixed effects by d-i, o-j, and i-j.

The figure shows that buyers source more from sellers with similar technologies. This pattern contrasts with the pure Ricardian view of trade—where exogenous technology differences drive trade and greater differences lead to more trade—but is consistent with our model. In the rest of this section, we first analyze this correlation at the firm level to isolate specific mechanisms and then examine how technology proximity responds to exogenous shocks in bilateral tariffs. We briefly describe the additional datasets used in each exercise; further details on data preparation are provided in Appendix B.1.

3.2 Firm-Level Correlation

Proposition 5 suggests that within a country-sector (d, i), firms whose endowed technology is closer to a foreign country o tend to choose similar technologies to those of o and import

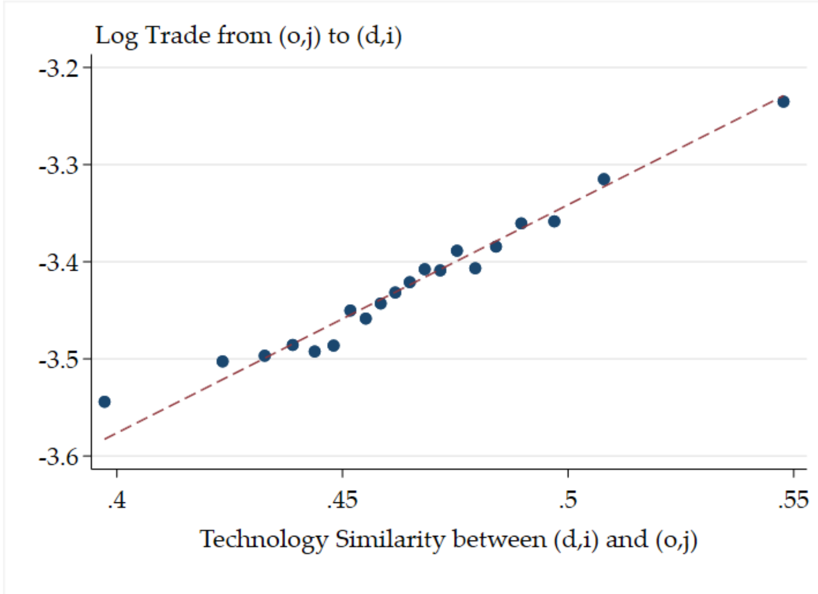


Figure 1: Technology Proximity and Trade

Notes: The figure is a binned scatter plot of log exports from country o sector j to country d sector i against the technology similarity between (d,i) and (o,j). Technology proximity is measured as the cosine similarity of the average textual embeddings of patent abstracts between (d,i) and (o,j). Trade data is from the World Input-Output Database (WIOD). The plot 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 1.9 (s.e. 0.2 clustered by country pair).

from o. This leads to a *firm-level* correlation between imports and technology proximity.[25]

Data. We use data on manufacturing firms from the Annual Survey of Industrial Enterprises (ASIE) of China's National Bureau of Statistics (NBS). This survey covers all state-owned firms and non-state firms with sales above 5 million RMB (approximately $\$600,000$) before year 2011 and with sales above 20 million RMB after 2011. We match these firms to *PATSTAT Global* to retrieve their patent families and then use the abstracts to construct $\text{Similarity}_{\omega ot}$ and $\text{Similarity}_{\omega ojt}$. Our final sample contains all manufacturing firms with patents, forming an unbalanced panel that grows from 57,465 firms in the first period (2000–2002) to 102,153 in the last period (2012–2014). We match these firms to Chinese Customs Data (available to us over 2000-2014), with which we construct an indicator for whether firm ω imported from region o in period t.

Specification. We estimate the following regression:

$$\mathbb{I}[\text{Import}_{\omega ot} > 0] = \beta\,\text{Similarity}_{\omega ot} + \beta_2 X_{\omega ot} + FE_{\omega t} + FE_{\omega o} + FE_{ot} + \epsilon_{\omega ot}, \tag{26}$$

where $\mathbb{I}[\text{Import}_{\omega ot} > 0]$ is an indicator for whether firm ω imports from o in period t and

[25] Our model is static, so firms' endowed technologies are permanent. In the empirical specification, we control for firm-region fixed effects to absorb the permanent component of firm-country connections, exploiting only the variations from the over-time shifts in technology endowments. If any, the empirical correlation between technology proximity and imports becomes stronger if firm-region fixed effects are excluded.

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

	$\mathbb{I}[\text{Import}_{\omega ot} > 0]$			
	(1)	(2)	(3)	(4)
$\text{Similarity}_{\omega ot}$	0.017***	0.017***	0.021***	0.022***
	(0.005)	(0.005)	(0.005)	(0.005)
FE ω-t	Yes	Yes	Yes	Yes
FE ω-o	Yes	Yes	Yes	Yes
FE o-t	Yes	Yes	Yes	Yes
$X_{i(\omega)ot}$		Yes		
FE $p(\omega)$-$i(\omega)$-o-t			Yes	Yes
Exclude Foreign Firms				Yes
Observations	3,326,764	3,326,764	3,262,280	2,431,940
R^2	0.748	0.748	0.766	0.729

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered by firm. $i(\omega)$ and $p(\omega)$ represent the industry and province of firm ω, respectively. $X_{i(\omega)ot}$ includes the applied tariffs of industry i's inputs and outputs. Industry i's input tariff is defined by the average applied tariff on imports of all goods from world region o, weighted by each input's share in industry i's production. Industries are classified at the three-digit level of China's Industry Classification.

the key regressor is $\text{Similarity}_{\omega ot}$. We include the following controls: firm-time fixed effects $FE_{\omega t}$, which absorb time-varying firm characteristics that affect firms' tendency to import or use foreign technologies (e.g., productivity, openness to foreign ideas); firm-region fixed effects $FE_{\omega o}$, which capture time-invariant linkages between a specific firm and a region. These fixed effects also alleviate the concern that our findings may reflect differences in firms' English proficiency or the way they write patent applications. FE_{ot} represents region-time fixed effects, accounting for changes over time in region o. Lastly, $X_{\omega ot}$ includes additional controls for other determinants of import and technology choices.

Main Results. Table 2 reports our findings. Column 1 shows that firms with technologies more similar to region o are more likely to import from o. One may be concerned that the correlation is driven by tariffs, which may be correlated with both trade and other unobserved linkages between firms and foreign regions. However, controlling for input and output tariffs in the firm's industry barely changes the coefficient (Column 2).

One may also be concerned that the correlation reflects differences among firms from different provinces or industries in China, which face different barriers when interacting with foreign suppliers due to either their geographic locations or historical connections with these countries. While the firm-country fixed effects absorb the time-invariant aspect of such connections, they do not capture the time-varying aspect. In Column 3, our preferred specification, we absorb all such variations by controlling for $i(\omega)$-o-$p(\omega)$-t fixed effects, in which $p(\omega)$ is the province of firm ω and $i(\omega)$ is the industry of ω. If any, the coefficient for $\text{Similarity}_{\omega ot}$ is larger. The estimate, 0.021, means that a one-standard-deviation rise in similarity with o is associated with a 0.19-percentage-point increase in the probability of importing from o, or 4% of the sample mean.

Finally, foreign-owned firms might drive this result if they import from their parent companies and adopt their parent's technology. In Column 4, we exclude foreign affiliates and joint ventures between foreign and Chinese firms. The sample shrinks by a third, but the estimate remains unchanged. Since our model applies to foreign affiliates as well as domestic firms, in the rest of this subsection, we use the full sample of firms. All results are robust if we exclude foreign firms and joint ventures.

Heterogeneity by Industry. The correlation documented in Table 2 is consistent with our model's premise that firms choose technology and suppliers jointly for compatibility reasons. An alternative information-based explanation is that firms may discover foreign suppliers and technologies through correlated sources. For example, the firm owner may meet a contact from a region o, who bring the technologies as well as the suppliers of o. Although we include a rich set of fixed effects, we cannot fully rule out such mechanisms. To support our interpretation of the correlation, we provide auxiliary evidence by looking into two implications of the model on how the correlation varies across industries.

First, if the correlation is indeed driven by a desire for compatible inputs, it should be stronger in sector pairs with more intensive input-output linkages. To test this implication, we use $\text{Similarity}_{\omega ojt}$, the similarity between firm ω and region-sector (d, j), as the explanatory variable. As the outcome variable, we use an indicator for whether a firm ω imports *goods in sector j* from country o, denoted by $\mathbb{I}[\text{Import}_{\omega ojt} > 0]$. Column 1 of Table 3 reports a baseline specification where we simply regress $\text{Similarity}_{\omega ojt}$ on $\mathbb{I}[\text{Import}_{\omega ojt} > 0]$. Given that the data have a j dimension, we are able to include as controls the interactions between the fixed effects in specification (26) and j. The estimate is 0.02, similar to those in Table 2. Column 2 adds the interaction between $\text{Similarity}_{\omega ojt}$ and $\gamma_{i(\omega)j}$, the input share of sector j in firm ω's sector $i(\omega)$. This interaction is sizable: when $\gamma_{i(\omega)j}$ increases by 10 percentage points, the correlation between similarity and importing increases by 0.0053—one quarter of the average correlation.

Second, the importance of input compatibility can vary across industries. For instance, electronics and telecommunications equipment often require stricter compatibility than apparel. If our estimates indeed capture technology compatibility, we should see a stronger correlation in industries where compatibility is more critical.

To test this hypothesis, we construct a measure of compatibility intensity using the full text of U.S. patents in the Orbis Intellectual Property Database, which links patents to firms worldwide. For each three-digit industry, we compute the fraction of patenting firms with at least one patent containing in its text either the word 'compatibility/compatible' or the word 'inter-operability.' Treating patents with these key words as an indicator of technologies where compatibility is important, this fraction, which we denote as $\text{Compatibility}_{i(\omega)}$,

Table 3: Firm-Level Correlation: Mechanisms and Alternative Explanations

	$\mathbb{I}[\text{Import}_{\omega ojt} > 0]$			
	(1)	(2)	(3)	(4)
$\text{Similarity}_{\omega ojt}$	0.020***	0.017***	0.018***	0.015***
	(0.004)	(0.004)	(0.004)	(0.004)
$\text{Similarity}_{\omega ojt} \times \gamma_{i(\omega)j}$		0.053*		0.056*
		(0.029)		(0.029)
$\text{Similarity}_{\omega ojt} \times \text{Compatibility}_{i(\omega)}$			0.083**	0.084**
			(0.035)	(0.035)
FE ω-j-t	Yes	Yes	Yes	Yes
FE ω-o-j	Yes	Yes	Yes	Yes
FE o-j-t	Yes	Yes	Yes	Yes
Observations	8,656,925	8,570,998	8,572,704	8,570,998
R^2	0.750	0.750	0.750	0.750

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered by firm. $\gamma_{i(\omega)j}$ is the input share of sector j in firm ω's sector $i(\omega)$. $\text{Compatibility}_{i(\omega)}$ measures the compatibility intensity of sector $i(\omega)$.

captures how often firms in industry i require compatible technologies.[26]

Column 3 of Table 3 adds the interaction between $\text{Compatibility}_{i(\omega)}$ and $\text{Similarity}_{\omega ojt}$. The coefficient on the interaction, 0.083, implies that a one-standard-deviation increase in industry compatibility intensity (0.11) raises the correlation by 0.9 percentage points—about 8.9% of the mean correlation. In Column 4, we further include the interaction with the input-output parameter, and the coefficients for both interactive terms remain similar. Together, these results offer support for our hypothesis that compatibility along the input–output network drives the correlation.

Extended Gravity. Another implication of our model is an 'extended-gravity'-like effect (Morales et al., 2019) in the relationship between technology similarity and trade: if region o' has technology similar to o, then stronger similarity between firm ω and o' is associated with a higher likelihood of ω importing from o. Thus, if we regress whether ω imports from o on the technological similarity between ω and a region o' that shares similar technologies as o, the resulting coefficient should be positive. This relationship can hold even after controlling for $\text{Similarity}_{\omega ot}$ unless the similarity measure is perfectly accurate.

In Appendix B.4.2, we show that, indeed, the similarity between firm ω and other regions o', which are technology neighbors of o, is strongly correlated with ω importing from o.

One might suspect an information-based alternative explanation—perhaps firm ω meets a new contact, who introduces to ω *both* the suppliers in o *and* the technologies of o'. While such knowledgeable and helpful experts certainly exist, given the language and cultural

[26]Table B.5 reports this compatibility intensity measure by industry. Our results remain robust if we refine the context in which these keywords appear in patent texts. See Appendix B.3 for details.

barriers between countries, it is more likely than not that their knowledge domain extends to country pairs that are linguistically and geographically close to each other. We therefore tease out the influence of such experts by controlling for the technology similarity between firm ω and the linguistic and geographic neighbors of region o. The coefficients for these controls are small and statistically insignificant. More importantly, they do not diminish the coefficient for the technology similarity between ω and the technology neighbors of o. This pattern further supports our interpretation that technology closeness, rather than purely informational channels, drives the correlation between imports and technology similarity.

3.3 Tariff Shocks and Technology Similarity

The second part of our empirical analysis tests another key implication of our model: a decline in bilateral tariffs between country pairs o and d increases technology similarity between o and d (relative to other country pairs).

Data. We construct a panel of bilateral technology similarity and tariffs at the level of $(d, o, HS6, t)$, covering 28 regions for both d and o (excluding ROW in Table B.1) and five three-year intervals from 2000-2014. Tariff data come from the UN TRAINS database, which records the ad valorem tariffs faced by each $(o, HS6)$ product in destination d at time t. In addition to these applied tariffs, we also collect most-favored-nation (MFN) tariffs. As described previously, we measure technology similarity, $\text{Similarity}_{do,HS6,t}$, between patents in country d and patents associated with $HS6$ products in country o at time t.

Specification. We estimate the following regression:

$$\text{Similarity}_{do,HS6,t} = \beta \ln \tau_{do,HS6,t} + FE_{d,HS6,t} + FE_{o,HS6,t} + FE_{do,HS6} + \epsilon_{do,HS6,t}, \tag{27}$$

where $\tau_{do,HS6,t}$ is the applied tariff imposed by importer d on an HS-6 product from o. To address endogeneity in $\tau_{do,HS6,t}$ (e.g., countries might lower tariffs on partners from whom they seek to adopt foreign technologies), we follow Boehm et al. (2023) and instrument $\ln \tau_{do,HS6,t}$ with the interaction between (i) the change in MFN tariff rates on the product in d ($\ln \tau^{MFN}_{d,HS6,t}$) and (ii) an indicator of whether this MFN rate applies to exporter o in period t. This design, explained in detail in the appendix, exploits the changes in MFN tariffs that affect some of an importer's partners but not others. We exclude the largest exporter for each importer-product pair to mitigate concerns that MFN tariffs are set with the major exporters in mind.

We include three sets of fixed effects to account for time-invariant and time-varying confounding factors. $FE_{d,HS6,t}$ absorbs changes in the importer d's similarity to all producers of that HS-6 good in each period t. $FE_{o,HS6,t}$ captures the average similarity of all importers with a particular exporter-product pair $(o, HS6)$ at time t. Finally, $FE_{do,HS6}$ absorbs any pair-specific aspects of $(d, o, HS6)$ that correlate with both tariffs and technology similarity.

Results. Table 4 reports the results. Columns 1 and 2 use the log of HS-6-level trade flow

Table 4: Tariff Shocks and Technology Similarity

	$\ln \text{Trade}_{do,HS6,t}$		$\text{Similarity}_{do,HS6,t}$	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
$\ln \tau^{MFN}_{do,HS6,t}$	-0.956***		-0.007***	
	(0.098)		(0.001)	
$\ln \tau^{Applied}_{do,HS6,t}$		-1.379***		-0.009***
		(0.141)		(0.001)
FE *d-o-HS6*	Yes	Yes	Yes	Yes
FE *o-HS6-t*	Yes	Yes	Yes	Yes
FE *d-HS6-t*	Yes	Yes	Yes	Yes
Observations	5,216,306	5,216,306	5,652,028	5,652,028
R^2	0.900	0.900	0.994	0.994

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the *d-o-HS6* level. Columns 1 and 3 report reduced-form regressions using OLS, while columns 2 and 4 report the second stage of 2SLS regressions using variations from MFN tariffs to instrument for applied tariffs.

as the outcome variable and log tariffs as the explanatory variable. Column 1 estimates the reduced-form specification, while Column 2 estimates the 2SLS specification, instrumenting $\ln \tau^{Applied}_{do,HS6,t}$ with the IV constructed above. In both cases, the estimate is around 1, consistent with the fixed-effect specification elasticity estimate in Boehm et al. (2023).

Columns 3 and 4 replace the outcome variable in the first two columns with technology similarity. According to the 2SLS estimate (Column 4), doubling the applied tariff reduces technology similarity by 0.009, or about 8% of the standard deviation of bilateral similarity. Since tariffs are only one of many barriers to bilateral trade (such as informational and cultural barriers), this estimate suggests that trade barriers can significantly influence technology choices.

In Appendix B.5, we show that using citations as a measure of bilateral similarity leads to similar findings. We also show that the estimates remain robust if we focus on goods primarily intended for intermediate use. This is reassuring given our model's emphasis on intermediate input and supplier compatibility.

Our findings here, and the appendix results using patent citations as a measure of similarity, relates to a literature pioneered by Jaffe et al. (1993), which uses citation flows to track technology diffusion. Most closely related, Aghion et al. (2021) show 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) and other existing works in two aspects. First, we exploit exogenous MFN tariff variation to isolate the causal effect of trade costs on bilateral patent similarity. We provide evidence on heterogeneous effects and extended-gravity results to support technology compatibility as one of the mechanisms. Second, our model interprets these findings through the lens of horizontal technology choice in production networks, complementing existing works that focus on other channels of technological diffusion.

4 Quantification

We use the model to quantify how trade influences firms' technology choices and how these choices affect welfare. We begin by outlining the functional-form assumptions and approximations that enable closed-form solutions for firms' technology choice. We then discuss our parametrization strategy, followed by counterfactual exercises.

4.1 Functional-Form Assumptions and Numerical Algorithm

In our model, firms directly interact with the distribution of heterogeneous firms rather than indirectly through aggregate prices and quantities. Solving for the equilibrium thus requires tracking the entire distribution of Θ_d^i across all (d, i). To maintain tractability, we assume:

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 a quadratic approximation to the log of $C_o^j(\theta)$ (defined in (5)) around the mean ex-post technology in each (o, j), we derive closed-form expressions for firms' technology choices and the resulting ex-post technology distributions:

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

1. *$\ln C_o^j(\theta)$ is given by the following function:*

$$\ln C_o^j(\theta) = k_{A,o}^j + m_A^j(\theta - n_{A,o}^j)^2, \tag{28}$$

   where coefficients $\{m_A^j, n_{A,o}^j, k_{A,o}^j\}$ can be solved as a function of primitives and wages $\{w_o\}$.

2. *Firms' technology choice $g_o^j(\cdot)$ is characterized by*

$$g_o^j(\bar{\theta}) \equiv \alpha_o^j + \beta^j \bar{\theta}, \tag{29}$$

   with $\alpha_o^j = \frac{(\eta-1)m_A^j}{\frac{1}{2}\bar{\phi}+(\eta-1)m_A^j} n_{A,o}^j$ and $\beta^j = \frac{\frac{1}{2}\bar{\phi}}{\frac{1}{2}\bar{\phi}+(\eta-1)m_A^j}$.

3. *Hence, the ex-post technology distribution Θ_o^j is also Normal:*

$$\Theta_o^j \sim \mathcal{N}(\mu_o^j, (\sigma^j)^2), \text{ with } \mu_o^j = \alpha_o^j + \beta^j \bar{\mu}_o^j \text{ and } \sigma^j = \beta^j \bar{\sigma}^j. \tag{30}$$

Proof. See Appendix C.1. □

Proposition 8 expresses in closed-form the parameters governing production costs and firms' technology choices.[27] In particular, equation (28) shows that we can recover the function $C_o^j(\theta)$ by solving for a small number of coefficients. Equation (29) is the policy function for technology choice, which circumvents the need to numerically solve the technology choice problem (15). The linearity of the policy in $\bar{\theta}$, together with Normal ex-ante technologies, also implies Normally distributed ex-post technologies, whose means and variances follow directly from the closed-form expressions in (30).

Building on this proposition, we devise a tractable algorithm to solve for the equilibrium. First, given wages $\{w_o\}$ and model fundamentals, solve for firms' technology choices $\{g_o^j(\bar{\theta})\}$ and cost parameters $\{C_o^j(\theta)\}$. This amounts to solving a contraction mapping problem for the coefficients in (28). Second, compute the sourcing decisions for both intermediate firms $\{\chi_{do}^j(\theta,\tilde{\theta})\}$ and final-good producers $\{\pi_{do}^j(\theta)\}$. Third, solve for industry aggregates, $\{X_o^j(\theta)\}$ and $\{M_o^j(\theta)\}$, using equations (18) to (21). Finally, check the labor market clearing condition (22). If it is satisfied, the equilibrium is found; if not, update $\{w_o\}$ and iterate from the first step. Appendix C.2 provides details for this algorithm.

4.2 Calibration Procedure

We now describe how we calibrate the model parameter, targeting the 2014 world economy.

Parameters Calibrated Externally. We set the elasticity of substitution for final-good producers $\eta = 5$, which implies a 20% markup for their purchase from intermediate firms. We set the shape parameter for match-specific efficiency $\zeta = 4$. This value means that the trade elasticity for intermediate firms (conditional on technology) is 4.[28]

We take labor endowments $\{L_d\}$ from the Penn World Table population data. We derive production-function coefficients $\{\rho_d^i, \gamma^{ij}, \gamma^{iL}\}$ from the World Input-Output Database (WIOD), matching sectoral consumption shares in final-good production (ρ_d^i) and sectoral input shares (γ^{ij} and γ^{iL}). All these values are taken over the average of 2012-2014.

Technology Distribution. We recover the (unobserved) *ex-ante* technology distributions, $\{\bar{\mu}_d^i, \bar{\sigma}^i\}$, along with other model primitives, in two steps. In the first step, we choose the parameters governing the *ex-post* distributions, $\{\mu_d^i, \sigma^i\}$, to match the similarities between country-sector pairs and their variances as reflected in patent text embeddings. In the second step, we calibrate the production and sourcing blocks of the model by treating firms' technology as exogenously fixed at the calibrated ex-post distribution, exploiting the result

[27] 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 Gaussian ex-ante technology distributions. In practice, the one-dimensional setup is sufficient to capture the key features of bilateral technology similarity data, as we show below.

[28] See equation (8). The unconditional trade elasticity may be larger than 4 due to endogenous technology. In the calibrated model, we find that the unconditional trade elasticity is close to 4 for small changes in bilateral trade costs.

that, conditional on the chosen technology, firms' sourcing decisions are independent of their endowment technology. Once this is complete, we calibrate $\bar{\phi}$ (which governs the cost of technology adaptation) and then recover the ex-ante distributions by inverting the firms' policy functions.[29]

To implement the first step, we map the technological distances in our model to a similarity metric bounded between 0 and 1, facilitating the comparison between model and empirical bilateral similarity.[30] We define the similarity between two firms with technologies θ and $\tilde{\theta}$ as:

$$sim(\theta, \tilde{\theta}) = \exp(-(\theta - \tilde{\theta})^2). \tag{31}$$

Analogously, the similarity between the mean technologies of two country-sector pairs is:

$$sim(\mu_d^i, \mu_o^j) = \exp(-(\mu_d^i - \mu_o^j)^2). \tag{32}$$

We choose $\{\mu_d^i\}$ by solving the following problem:

$$\min_{\{\mu_d^i\}} \sum_{d,i,o,j} \left[\ln sim(\mu_d^i, \mu_o^j) - \ln sim_{do}^{ij,Data}\right]^2, \tag{33}$$

where $\ln sim_{do}^{ij,Data}$ is the log of empirical technology similarity between country-sector pairs (d,i) and (o,j), averaged over the period 2012-2014, residualized by *i*-*j* fixed effects. With this residualization, our calibration does not attempt to explain why, for example, the auto industry is inherently closer to the auto parts industry than to agriculture. Instead, it focuses on the differences between pairs of countries within the same industry pair, e.g., whether the German auto sector is closer to the Chinese or American auto sector.[31]

Problems of the form (33) is known to be challenging.[32] We show in Appendix C.3 that by fitting the logarithms of the similarities, it can be transformed into a classical multidimensional scaling problem, which has unique solutions and can be efficiently solved for $\{\mu_d^i\}$.

To infer the standard deviation of technology, $\{\sigma^i\}$, we match the standard deviation of similarity between technologies within each (d,i). For each country-sector pair, we ran-

[29] Note that $\bar{\phi}$ affects only the allocation of firms' markups between profits and technology adaptation costs. For this reason, once the ex-post distributions and other general equilibrium outcomes (such as wages and prices) have been calibrated, $\bar{\phi}$ can be adjusted without altering earlier calibration steps (see Appendix C.4 for details).

[30] Although cosine similarity theoretically ranges from $[-1, 1]$, the entries of the empirical similarity matrix (at the country-sectoral level) are all positive, making the logarithm transformation possible.

[31] This approach is justified by two key reasons. First, our counterfactual analysis centers on how international trade shocks shift the technological positions of countries. Although our one-dimensional technology space may not capture all variations across sectors, the most important variation for our study is the difference between countries within a sector pair. By matching the residualized data, we match precisely these cross-country differences. Second, under the Cobb-Douglas production function of intermediate-good producers, adding exogenous technological distance between pairs of sectors amounts to adding exogenous mutilative cost shifters. Therefore, enriching the model (or its calibration) to match exogenous sectoral distance will not affect the counterfactual findings. Our calibration and counterfactuals have the interpretation of holding constant the *sectoral* technological differences while focusing on country differences.

[32] See e.g., Chapter 13 of Borg and Groenen (2007).

Table 5: Summary of Model Parameters

Parameters	Descriptions	Value	Target/Source
A. Externally Calibrated			
$\gamma^{ij}, \gamma^{iL}, \alpha^{j}$	IO Structure and Consumption Share	-	WIOT; $N = 29, S = 19$
L_d	Labor Endowment	-	PWT
$\eta - 1, \zeta$	Trade Elasticity	4	Simonovska and Waugh (2014)
B. Just-Identified			
A^i_d	Productivity	-	Countries' output shares in each sector
$\tau^j_{do}, \tau^{Uj}_{do}$	Iceberg Trade Costs	-	Bilateral trade shares (WIOT)
$\bar{t}$	Compatibility Incentive	12.2	Firm-level import-similarity corr: 0.021
$\bar{\phi}$	Adaptation Costs	0.58	Country-sector-level similarity-tariff elas.: -0.007
C. Best Fit			
μ^i_d, σ^i	Ex-post Technology Distribution	-	Tech similarity between countries and sectors
$\bar{\mu}^i_d, \bar{\sigma}^i$	Ex-ante Technology Distribution	-	Inverted from calibrated ex-post distributions

domly sample 1,000 patents from year 2014, calculate the pairwise similarity of their abstract embeddings, and compute the standard deviation of these similarities by sector. We then average this statistic across countries to infer $\{\sigma^i\}$.

The second step of our calibration parameterizes the trade and production aspects of the model to match the data across a set of moments, as explained below.

Productivity and Trade Costs. We calibrate country-sectors' productivity $\{A^i_d\}$ to match the output share of (d, i) in each sector i, and trade costs $\{\tau^j_{do}, \tau^{Uj}_{do}\}$ to match the trade shares of intermediate and final goods.[33]

Technology Compatibility $\bar{t}$. The parameter $\bar{t}$ governs firms' input-sourcing decisions given their chosen technology. We calibrate $\bar{t}$ by matching the correlation between import indicator and technology similarity (i.e., Column 3 of Table 2) using the simulated method of moments. Since the regression focuses on Chinese firms, our simulation centers on China too. For each sector i, we draw $10,000$ firms from the ex-post distribution Θ^i_{CHN}. We simulate whether a firm with technology θ imports from (o, j) using equation (8). We then regress the import indicator on the technology similarity between θ and (o, j), defined as $sim(\theta, \mu^j_o)$, controlling for o-i and firm fixed effects. This exercise sets $\bar{t} = 12.2$, under which the model-simulated regression matches the empirical coefficient of 0.021.

Adaptation Costs $\bar{\phi}$. We choose the adaptation cost parameter $\bar{\phi}$ to match the elasticity of technology similarity with respect to tariffs, as estimated in Column 3 of Table 4. As Proposition 4 illustrates, conditional on $\bar{t}$ and other model fundamentals, this elasticity identifies $\bar{\phi}$. To calculate the model-based elasticity, we simulate changes to trade costs $\{\tau^j_{do}\}$, mimicking the size of the actual MFN tariff changes observed over time for each (d, j). This produces a counterfactual equilibrium. We then treat the calibrated equilibrium and the counterfactual

[33] Trade and output shares do not separately identify all $\{A^i_d\}$ and all $\{\tau^j_{do}, \tau^{Uj}_{do}\}$ jointly (Burstein and Vogel, 2017). Our counterfactual results are invariant to normalization choice in the calibration of these parameters.

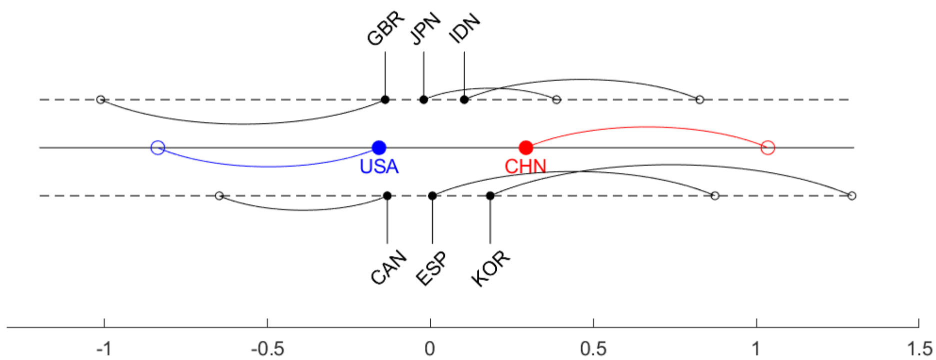


Figure 2: Estimated Technology Distributions for Selected Economies

Notes: Filled dots are the means of firms' ex-post technology distributions, $\{\mu_d^i\}$, averaged across sectors for each d for selected countries. Unfilled circles are the means of their endowment technology $\{\bar{\mu}_d^i\}$.

as two periods and regress the model's technology similarity between d and (o, j) on the logarithm of τ_{do}^j, controlling for a range of fixed effects mirroring the empirical specification. We adjust $\bar{\phi}$ until the model's elasticity matches the empirical coefficient of -0.007, yielding a calibrated value of $\bar{\phi} = 0.58$.

Numerical Implementation. We implement the calibration as follows. First, we estimate the ex-post distribution parameters $\{\mu_d^i, \sigma^i\}$ to best match the empirical technology similarity within and between country-sectors, as previously described. Given these parameters, we employ a nested algorithm for the second step. In the outer layer, we adjust $\bar{t}$ to match the firm-level extensive margin import-similarity regression coefficient. In the inner loop, we solve for the competitive equilibrium, while simultaneously calibrating $\{\tau_{do}^j, \tau_{do}^{Uj}, \Xi^i / A_d^i\}$ to match trade shares. We then simulate the model and feed the regression coefficient back into the outer loop. Once this process is complete, we calibrate $\bar{\phi}$ to match the country-sector similarity-tariff elasticity and use the results from Proposition 8 to back out the ex-ante distribution parameters $\{\bar{\mu}_d^i, \bar{\sigma}^i\}$ from $\{\mu_d^i, \sigma^i\}$. Appendix C.4 reports the detailed procedure.

Table 5 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) of the mean locations of ex-ante and ex-post technology distributions for a subset of countries. The filled dots depict the locations of ex-post technologies. Among major economies, the technologies of the U.S. and China fall at the two ends of the spectrum, whereas the technologies of other major economies, such as Japan, Canada, and Western European countries, fall in the middle. This

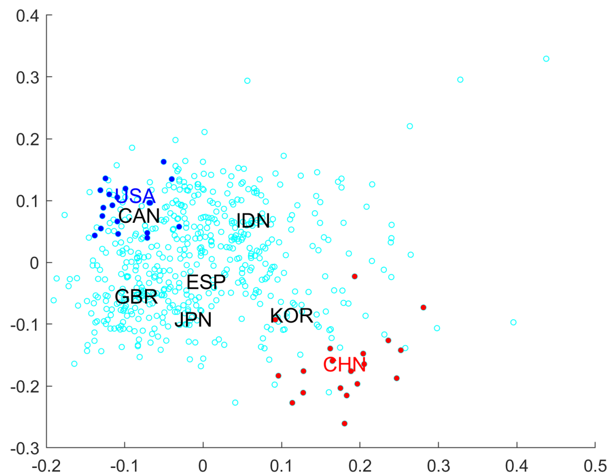


Figure 3: First Two Principal Coordinates of the Similarity Matrix

Notes: Plotted are the first two principal coordinates of the similarity matrix between country-sector pairs. The similarity measure is defined as the log of cosine similarity of the average embeddings of patent abstracts between two country-sectors (d, i) and (o, j), netting out sector-pair (i.e., i-j) fixed effects. Each dot corresponds to the principal coordinates of a country-sector pair, obtained by solving a problem analogous to (33) in the two-dimensional space for technology. Country name indicates the average coordinates of the country across sectors. Sectors in China and the U.S. are indicated in red and blue, respectively.

pattern highlights a key feature of bilateral technology similarity statistics derived from the textual analysis of patent abstracts: within sector pairs, American and Chinese technologies tend to be less similar to each other than they are to those of other countries.

The circles in Figure 2 depict the average locations of ex-ante (endowment) technologies for these countries. Two patterns are noteworthy. First, ex-ante technologies tend to cluster spatially, with Asian economies grouped on the right and other economies clustered on the left. This clustering may reflect the influence of geographic, historical, or cultural factors on initial technological endowments. Second, with technology compatibility incentives, international trade brings countries' technologies closer. Indeed, if countries do not engage in trade, then firms will move toward the average technologies of their own countries.

Assessing Fit of Technology Similarity. Our model allows technology to be represented in a high-dimensional space. For simplicity and transparency, however, we restrict our calibration to a one-dimensional space. To demonstrate that the calibrated technology positions are not an artifact of our one-dimensional representation, Figure 3 plots the first two principal coordinates of the similarity matrix for country-sector pairs in a two-dimensional technology space, which correspond to the calibration results (up to a scaling factor) if the two-dimensional space were used. The figure shows that the divergence between U.S. and Chinese technologies, with other countries positioned in between, remains a salient feature of the data when viewed in this space.

Table 6: Bilateral Similarity Between Country-Sectors: Model v.s Data

Log Similarity in Model	Log Similarity in Data (1)	(2)
at Ex-post Tech. Dist.	0.363***	
	(0.001)	
at Ex-ante Tech. Dist.		0.001***
		(0.000)
Obs.	302,664	302,664
Within R^2	0.212	0.128
Fixed effects	d-i, o-j	d-i, o-j

Notes: This table assesses the goodness of fit in the model. Each column reports the regression of data log similarity between country-sector pairs on model log similarity. Column 1 uses the ex-post technology distributions $\{\mu_o^j, \sigma^j\}$ to construct model similarity. Column 2 uses the ex-ante distributions $\{\bar{\mu}_o^j, \bar{\sigma}^j\}$. Both columns control for d-i and o-j fixed effects. The reported within R^2 excludes the variation explained by the fixed effects.

To evaluate more systematically how well the calibration captures the cross-sectional technology similarities between country-sectors, we regress the log of empirical technology similarity between country-sectors on the log of model-predicted similarity, controlling for country-sector fixed effects. Column 1 of Table 6 reports the result. Approximately 21% of the variation in bilateral similarities is explained by the model. Note that fully describing the bilateral similarity matrix between country-sectors would typically require $\frac{(NS)^2-NS}{2}$ parameters. Our model captures a substantial share of the variation by choosing $\{\mu_d^i\}$, totaling only $NS-1$ parameters. As another benchmark, the explanatory power of bilateral distance for bilateral trade, one of the most successful empirical model, is around 25%-50%.

Relevance in Shaping Technology. To further unpack the model's fit and highlight the role of compatibility incentives in shaping technology interdependence through trade, Column 2 of Table 6 regresses the log of data similarity on the log of model similarity implied by the calibrated endowment technologies. Here, the regression coefficient declines to almost zero, and the R^2 also falls from 21% to 13%. This comparison suggests that two-fifths of the model's explanatory power (8% out of 21%) can be attributed to endogenous technology choices driven by compatibility incentives, whereas the remaining comes from countries' endowment technology distributions.

Relevance in Shaping Trade Costs. Under our calibration, on average (weighted by trade flows), costs arising from technology incompatibility is equivalent to 9.5% ad valorem trade costs. The literature has found that, to account for the low level of international trade, iceberg trade costs must be substantially higher than the level of tariffs (Anderson and van Wincoop, 2004). Our model suggests that a modest but non-negligible part of these unobserved trade barriers could be technological incompatibility.

Table 7: Welfare Costs of a Semiconductor and Electronics Embargo on China

	Import share in targeted China's sector (%)	$\Delta Sim_{US,China}$	Δ Welfare (%) China	U.S.	Others	World
U.S. only	1.18	0.000	-0.006	-0.002	-0.001	-0.002
All countries	100	-0.014	-0.546	-0.056	-0.167	-0.232

Notes: This table reports the impact of an embargo on China's imports of Computer, Electronic, and Optical Products from the U.S. and other countries. $Sim_{US,China}$ is the technology similarity between the U.S. and China as defined by equation (32), averaged across sectors. Welfare is defined as household real income.

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 U.S. and—due to the long-arm jurisdiction of U.S. sanctions—from other countries as well. In our data, this sector accounts for approximately 19.7% of China's total imports.

To implement this counterfactual, we shut down exports in this sector to China by raising the corresponding trade costs to infinity. We investigate two cases: (i) only the U.S. imposes the embargo, and (ii) all countries impose the embargo. In both cases, countries' endowment distributions are set to be the same as their endowment distribution in the calibrated equilibrium. The comparison between these counterfactuals and the baseline equilibrium reveals what would happen when the economy starts from the same technological primitives but operates under different trade environments.[34]

Table 7 reports the main findings. Since the U.S. accounts for only 1.2% of China's imports in this sector, when only the U.S. imposes the embargo, the impact on the technology distance between the two countries is minimal. While the embargo imposes more harm on China than the U.S., the overall economic damage is limited.

The second row of the table reports the results when all countries impose the embargo at the same time. China now sustains a welfare loss of 0.55%, about 100 times the loss 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 the direct U.S. sanctions are the same in both cases. This occurs because, among major economies, China and the U.S.

[34]By holding endowments in these counterfactuals fixed at the ex-ante distribution in the calibrated equilibrium, we adopt the view that countries' technology endowments are persistent. According to this view, sustained investment in technology adaptation is necessary for firms to deviate from their endowed technology, so different counterfactuals start from the same endowment.

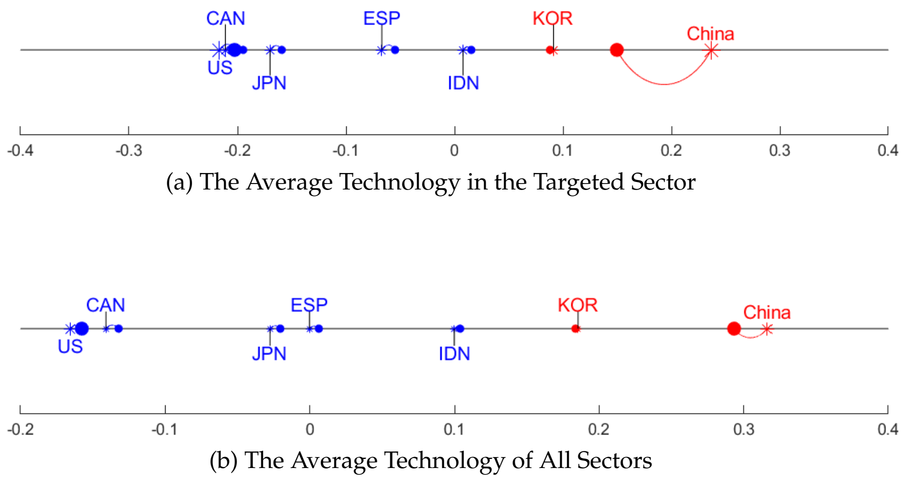


Figure 4: Technology Distributions Before and After the Embargo for Selected Economies

Notes: Dots are the ex-post means in the baseline equilibrium, and stars are the equilibrium with the embargo. Blue indicates countries moving toward the U.S. and red indicates countries moving toward China.

occupy the two ends of the technology spectrum. As Chinese technologies shift away from those of other countries due to the embargo, they also shift away from U.S. technologies. Second, the welfare cost to the U.S. is larger than before. This amplification occurs for two reasons: first, production costs in China increase due to the direct effect of the embargo, and the rest of the world must now 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 Technology Decoupling. To shed light on the role of endogenous technology, Figure 4 plots the changes in countries' technologies due to the embargo. The upper panel depicts the change in the mean technology of the targeted industry. As discussed earlier, the embargo leads to a divergence of technologies between China and the U.S.[35] This divergence results in a realignment of other countries' technologies. We indicate a country in blue if its technology moves toward the U.S. and in red if it moves toward China. Among others, Japan, Mexico, and European nations move closer to the U.S., while Korea is one of the few that shift toward China. However, even as most countries' technologies move toward the U.S., the gap between their technologies and U.S. technology may still widen because the change in technology choice is smaller than that of the U.S. This widening technology gap contributes to amplified welfare costs for the U.S.

Because of the compatibility incentives, the divergence in technologies in semiconduc-

[35]The embargo directly impacts the technology choices of downstream sectors. For the Computer, Electronic, and Optical Products sector, the most affected downstream sector is itself, as it relies heavily on its own inputs.

Table 8: Mechanism Decomposition

	$\Delta Sim_{US,China}$	Δ Welfare (%)			
		China	U.S.	Others	World
Fixed Technology	0.000	-0.43	-0.01	-0.04	-0.12
+ response of targeted China's sector	-0.002	-0.33	-0.02	-0.06	-0.11
+ response of all sectors in China	-0.010	-0.51	-0.03	-0.10	-0.18
+ response of all country-sectors	-0.014	-0.55	-0.06	-0.17	-0.23

Notes: The rows show the effects of allowing different degrees of changes in firms' technology choices under the embargo. The 'Fixed Technology' row represents the scenario where technology choices remain fixed at their calibrated equilibrium values.

tors spreads to all other sectors via input-output linkages. The lower panel plots the mean technology for each country across all sectors, including those not directly affected by the embargo. The changes are qualitatively similar to those in the target sector.

To underscore the importance of endogenous technology decoupling on welfare, we rerun the embargo scenarios under the assumption that firms cannot change their technology choice after the shock, effectively freezing technology choices at the pre-embargo equilibrium ('Fixed Technology'). The results, shown in the first row of Table 8, reveal that when technology choice is restricted, welfare losses are significantly smaller for China, the U.S., and the world. In contrast, when firms can adapt their technology, technological responses amplify welfare losses—by 28% for China (from 0.43% to 0.55%), by 500% for the U.S. (from 0.01% to 0.06%), and by 92% for the global average (from 0.12% to 0.23%). This amplification occurs because, while individual firms benefit from the ability to choose technologies most suitable for the embargo trade environment, the broader economy suffers from the negative externalities arising from technology incompatibility across sectors and countries. These externalities compound via input-output linkages, increasing the overall welfare costs of the embargo.

We decompose the role of various margins of technology choice in the remaining rows of Table 8, gradually introducing technology changes—from the most directly affected sector to all sectors in China, and to all other countries. Row 2 shows that when firms in the directly targeted sector in China adjust their technology, China's welfare losses actually decrease. These firms can re-optimize by balancing sourcing efficiency against adaptation costs in response to changes in sourcing origins, as reflected in Figure 4a. However, these unilateral adjustments create externalities for downstream sectors in other countries, contributing to additional welfare losses abroad.

Row 3 shows that as firms in non-targeted sectors in China adjust their technology choice, the negative externalities outweigh the benefits, leading to greater welfare losses for China and foreign countries compared to the 'Fixed Technology' scenario. Finally, Row 4 demonstrates that when firms in other countries adjust their technologies, technological decoupling becomes more pronounced, further increasing global welfare losses. Almost all of the U.S. loss and half of the global loss from the embargo arise because of endogenous technological decoupling.

These counterfactual findings echo the externality mechanisms discussed in Propositions 7 in the two-country setting. It is worth noting that in an environment with multiple countries, endogenous technological divergence between any two does not necessarily reduce welfare. For instance, if a trade conflict between Korea and Japan leads to a technological divergence between the two, welfare may improve—if the positive externalities of Japan aligning more closely with U.S. technology and Korea with Chinese technology outweigh the negative externalities of divergence between Korea and Japan. This possibility highlights the value of developing a quantitative framework and disciplining it with data on technology similarity.

5 Conclusion

In this paper, we construct a quantitative trade model that incorporates endogenous production networks and horizontal technology choice. In our framework, firms jointly decide on their technology and suppliers under compatibility incentives, with both decisions shaped by general equilibrium forces. We establish sufficient conditions for equilibrium existence and uniqueness and provide tractable aggregation results, making the model suitable for quantitative applications.

Using patent and trade data, we provide novel evidence at both the firm and country levels supporting the core mechanisms of our model. Our quantitative analyses reveal three main findings. First, trade-induced endogenous technology plays a significant role in shaping global technologies, accounting for two-fifths of the variation in technological proximity between countries, while differences in endowment distributions explain the remainder. Second, technology incompatibility imposes an average cost equivalent to 9.5% in ad valorem terms in international trade, underscoring the importance of this mechanism. Last, trade conflicts between the U.S. and China can lead to technological decoupling between them and trigger realignments among other countries. These endogenous responses amplify the losses from the conflict, accounting for almost all of the U.S. welfare loss and half of the global loss.

Our framework can be extended in two directions. First, in reality, firms and countries develop institutions to manage the externalities highlighted in our model. Examples include supply chain integration to internalize externalities and collaborative efforts to establish common standards. Investigating the equilibrium impacts of such mechanisms is a promising research avenue. Second, our model is static and therefore abstracts from dynamics and vertical innovation. Integrating our model into a dynamic setting can illuminate how compatibility incentive in firms' innovation decisions and supplier choices shape both industry dynamics and aggregate growth.

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," 2021. Unpublished.

Akcigit, Ufuk, Murat Alp Celik, and Jeremy Greenwood, "Buy, Keep, or Sell: Economic Growth and the Market for Ideas," *Econometrica*, 2016, *84* (3), 943–984.

Allen, Treb, Costas Arkolakis, and Xiangliang Li, "On the Equilibrium Properties of Spatial Models," *American Economic Review: Insights*, 2024, *6* (4), 472–89.

Alvarez, Fernando and Robert E Lucas, "General Equilibrium Analysis of the Eaton–Kortum Model of International Trade," *Journal of Monetary Economics*, 2007, *54* (6), 1726–1768.

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.

Anderson, James E. and Eric van Wincoop, "Trade Costs," *Journal of Economic Literature*, 2004, *42* (3), 691–751.

Baqaee, David Rezza and Emmanuel Farhi, "Networks, Barriers, and Trade," *Econometrica*, 2024, *92* (2), 505–541.

Basu, Susanto and David N. Weil, "Appropriate Technology and Growth," *The Quarterly Journal of Economics*, 1998, *113* (4), 1025–1054.

Boehm, Christoph E., Andrei A. Levchenko, and Nitya Pandalai-Nayar, "The Long and Short (Run) of Trade Elasticities," *American Economic Review*, 2023, *113* (4), 861–905.

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.

Bonadio, Barthélémy, Zhen Huo, Elliot Kang, Andrei A. Levchenko, Nitya Pandalai-Nayar, Hiroshi Toma, and Petia Topalova, "Playing with Blocs: Quantifying Decoupling," 2024. Unpublished.

Borg, Ingwer and Patrick JF Groenen, *Modern Multidimensional Scaling: Theory and Applications*, Springer Science & Business Media, 2007.

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

Burstein, Ariel and Jonathan Vogel, "International Trade, Technology, and the Skill Premium," *Journal of Political Economy*, 2017, *125* (5), 1356–1412.

Cai, Jie, Nan Li, and Ana Maria Santacreu, "Knowledge Diffusion, Trade, and Innovation across Countries and Sectors," *American Economic Journal: Macroeconomics*, 2022, *14* (1), 104–45.

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.

Clayton, Christopher, Matteo Maggiori, and Jesse Schreger, "A Framework for Geoeconomics," Working Paper 31852, National Bureau of Economic Research November 2023.

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).

Demir, Banu, Ana Cecília Fieler, Daniel Yi Xu, and Kelly Kaili Yang, "O-Ring Production Networks," *Journal of Political Economy*, 2024, *132* (1), 200–247.

Dhyne, Emmanuel, Ayumu Ken Kikkawa, Xianglong Kong, Magne Mogstad, and Felix Tintelnot, "Endogenous Production Networks with Fixed Costs," *Journal of International Economics*, 2023, *145*, 103841.

Eaton, Jonathan and Samuel Kortum, "International Technology Diffusion: Theory and Measurement," *International Economic Review*, 1999, *40* (3), 537–570.

__ **and** __ , "Technology, Geography, and Trade," *Econometrica*, 2002, *70* (5), 1741–1779.

__ **, Samuel S Kortum, and Francis Kramarz**, "Firm-to-Firm Trade: Imports, Exports, and the Labor Market," Working Paper 29685, National Bureau of Economic Research 2022.

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.

Huang, Yi, Chen Lin, Sibo Liu, and Heiwai Tang, "Trade Networks and Firm Value: Evidence from the US-China Trade War," *Journal of International Economics*, 2023, *145*, 103811.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *The Quarterly Journal of Eco-*

nomics, 1993, *108* (3), 577–598.

Jones, Charles I., "Misallocation, Economic Growth, and Input–Output Economics," in Daron Acemoglu, Manuel Arellano, and Eddie Dekel, eds., *Advances in Economics and Econometrics: Tenth World Congress*, Econometric Society Monographs, Cambridge University Press, 2013, p. 419–456.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy, "Measuring Technological Innovation over the Long Run," *American Economic Review: Insights*, 2021, *3* (3), 303–20.

Kleinman, Benny, Ernest Liu, and Stephen J. Redding, "International Friends and Enemies," *American Economic Journal: Macroeconomics*, 2024, *16* (4), 350–85.

Kortum, Samuel S., "Research, Patenting, and Technological Change," *Econometrica*, 1997, *65* (6), 1389–1419.

Lancaster, Kelvin, *Variety, Equity, and Efficiency: Product Variety in an Industrial Society*, New York Chichester, West Sussex: Columbia University Press, 1979.

Lazear, Edward P., "Culture and Language," *Journal of Political Economy*, 1999, *107* (S6), S95–S126.

Li, Yao Amber, Sichuang Xu, Stephen R. Yeaple, and Tengyu Zhao, "Bilateral Economies of Scope," *The Review of Economics and Statistics*, 2024, pp. 1–45.

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

Lind, Nelson and Natalia Ramondo, "Global Innovation and Knowledge Diffusion," *American Economic Review: Insights*, 2023, *5* (4), 494–510.

Liu, Ernest and Song Ma, "Innovation Networks and R&D Allocation," Working Paper 29607, 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.

Morales, Eduardo, Gloria Sheu, and Andrés Zahler, "Extended Gravity," *The Review of Economic Studies*, 2019, *86* (6), 2668–2712.

Muennighoff, Niklas, Nouamane Tazi, Loïc Magne, and Nils Reimers, "MTEB: Massive Text Embedding Benchmark," 2023.

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

Simonovska, Ina and Michael E Waugh, "The Elasticity of Trade: Estimates and Evidence," *Journal of International Economics*, 2014, *92* (1), 34–50.

Timmer, Marcel P, Erik Dietzenbacher, Bart Los, Robert Stehrer, and Gaaitzen J De Vries, "An Illustrated User Guide to the World Input–Output Database: the Case of Global Automotive Production," *Review of International Economics*, 2015, *23* (3), 575–605.

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

We restate Proposition 1 below:

Proposition A.1. *Under Assumption 1, $p_d^i(\theta)$, the factory-gate price of a firm in (d,i) with technology location θ, follows a Weibull (inverse Fréchet) distribution with c.d.f.*

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

The location parameter $C_d^i(\theta)$ is determined as the fixed point of

$$C_d^i(\theta) = \frac{\Xi^i}{A_d^i}[w_d]^{\gamma^{iL}} \prod_j \Big[\sum_o \int [\tau_{do}^j C_o^j(\tilde{\theta}) t(\theta,\tilde{\theta})]^{-\zeta} \mathrm{d}\Theta_o^j(\tilde{\theta})\Big]^{-\frac{\gamma^{ij}}{\zeta}}, \tag{A.2}$$

where Ξ^i is a sector-specific constant, and Θ_o^j is the probability measure that describes the distribution of chosen technologies across firms in (o,j).

We prove Proposition 1 by guess and verification. We start with introducing a lemma. The lemma employs the Poisson assumption in the draws of suppliers to characterize the unconditional distribution of supplier efficiency.

Lemma A.1. *Let $\{X_z\}_z$ be a collection of random variables indexed by $z \in (0,\infty)$ such that*

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

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

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

Taking any monotonically increasing sequence $\{z_i\}_{i=1}^\infty$ with $z_1 = 0$ and $\lim_{i\to\infty} z_i = \infty$, we have: (since the number draws in any interval follow independent Poisson, the sum of draws across these intervals follow a Poisson)

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

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). □

Proof of Proposition 1. We now prove Proposition 1.

Proof. Suppose $p_d^i(\theta)$, the factory-gate price of a firm in (d,i) with technology location θ, follows a Weibull distribution with c.d.f. specified in (4). We verify that this is consistent with firms' behaviors in the equilibrium.

Consider a firm ν in (d,i) with technology location $\theta = \theta(\nu)$. This firm chooses among all techniques one that minimizes its production cost. For each technique r, the firm chooses one supplier in each sector j to source input.

We first characterize the distribution of production cost for different techniques, which we denote by $c^j(\nu,r)$. Recall that

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

where $\tilde{c}^j(\nu,\omega)$, the (compatibility-adjusted) cost of for input j from supplier $\omega \in \Omega_o^j(\nu,r)$ is

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

Assumption 1 specifies the distribution of the match-specific efficiency of potential suppliers, $z(\nu,\omega)$. Following Lemma A.1, from any origin country-sector pair (o,j), the number of suppliers such that each supplier ω has a technology location $\theta(\omega)$ around $\tilde{\theta}$ and an effective marginal cost $\tilde{c}^j(\nu,\omega)$ less or equal to $c>0$ follows a Poisson distribution with mean

$$\begin{aligned}
&\int_0^\infty F_o^j[\frac{z\cdot c}{\tau_{do}^j\cdot t(\theta,\tilde{\theta})};\tilde{\theta}]\zeta z^{-\zeta-1}\mathrm{d}z\cdot \mathrm{d}\Theta_o^j(\tilde{\theta})\\
=&\int_0^\infty F_o^j(\kappa;\tilde{\theta})\zeta\kappa^{-\zeta-1}[\frac{c}{\tau_{do}^j\cdot t(\theta,\tilde{\theta})}]^\zeta \mathrm{d}\kappa\cdot \mathrm{d}\Theta_o^j(\tilde{\theta})\\
=&\int_0^\infty [t(\theta,\tilde{\theta})\kappa]^{-\zeta}\mathrm{d}F_o^j(\kappa;\tilde{\theta})\cdot(\frac{c}{\tau_{do}^j})^\zeta\cdot \mathrm{d}\Theta_o^j(\tilde{\theta})\\
=&\,\Gamma(1-\zeta/\lambda)\cdot[t(\theta,\tilde{\theta})C_o^j(\tilde{\theta})]^{-\zeta}\cdot(\frac{c}{\tau_{do}^j})^\zeta\cdot \mathrm{d}\Theta_o^j(\tilde{\theta}). \qquad \text{(A.3)}
\end{aligned}$$

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 (7),

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

Therefore, for each (o,j), the probability that no such supplier arrives is

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

The cost distribution for each j when choosing technique r, $c^j(\nu,r)$, which is the minimum over all countries, is hence characterized by

$$\Pr[c^j(\nu,r) > c] = \exp[-\Gamma(1-\zeta/\lambda)\cdot\sum_o(\frac{c}{\tau^j_{do}\cdot\Lambda^j_o(\theta)})^\zeta]$$
$$= \exp[-\tilde{\Lambda}^j_d(\theta)c^\zeta], \tag{A.4}$$

where $\tilde{\Lambda}^j_d(\theta) \equiv \Gamma(1-\zeta/\lambda)\sum_o(\tau^j_{do}\Lambda^j_o(\theta))^{-\zeta}$.

Among all available techniques, the firm chooses one to minimize its factory-gate price:

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

Assumption 1 specifies the distribution of production efficiency $A(\nu,r)$. By applying Lemma A.1 again, the number of techniques such that the factory-gate price is weakly less than p follows a Poisson distribution with mean

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

where

$$\Xi^i \equiv (\int_0^\infty\int_0^\infty ...\int_0^\infty \mathbb{I}[\prod_j[m^j]^{\frac{\gamma^{ij}}{\zeta}} \le \kappa]\prod_j[\Gamma(1-\zeta/\lambda)]^{\frac{\gamma^{ij}}{\zeta}}\exp(-m^j)\lambda\kappa^{-\lambda-1}\mathrm{d}m^1...\mathrm{d}m^S\mathrm{d}\kappa)^{-1/\lambda} \tag{A.5}$$

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

The probability that no such technique arrives is

$$\Pr[p(\nu) > p] = \exp(-[p/C^i_d(\theta)]^\lambda),$$

where

$$C^i_d(\theta) = \frac{\Xi_i}{A^i_d}[w_d]^{\gamma^{iL}}\prod_j(\sum_o[\tau^j_{do}\Lambda^j_o(\theta)]^{-\zeta})^{-\frac{\gamma^{ij}}{\zeta}} = \frac{\Xi_i}{A^i_d}[w_d]^{\gamma^{iL}}\prod_j(\sum_o\int[\tau^j_{do}C^j_o(\tilde{\theta})t(\theta,\tilde{\theta})]^{-\zeta}\mathrm{d}\Theta^j_o(\tilde{\theta}))^{-\frac{\gamma^{ij}}{\zeta}}.$$

Therefore, for each destination country-sector pair (d,i), the factory-gate price of a firm with technology location θ is characterized by

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

This completes the proof of Proposition 1. $\square$

Proof of Corollary 1. As is shown in (A.3), 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

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

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

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

This completes the proof of Corollary 1.

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

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

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

$$P_d = \prod_j (\frac{P^j_d}{\rho^j_d})^{\rho^j_d}, \quad \text{with } P^j_d \equiv (\sum_o \int_0^1 [\frac{\eta}{\eta-1}\tau^{Uj}_{do} p^j_{do}(\omega)]^{1-\eta} \mathrm{d}\omega)^{\frac{1}{1-\eta}}$$

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

$$\frac{\frac{\eta}{\eta-1}\tau^{Uj}_{do} p^j_{do}(\omega) x^j_{do}(\omega)}{\rho^j_d P_d Q_d} = \frac{[\frac{\eta}{\eta-1}\tau^{Uj}_{do} p^j_{do}(\omega)]^{1-\eta}}{(P^j_d)^{1-\eta}}.$$

Taking expectation overs firms with technology location around $\tilde{\theta}$, we have the expected expenditure share allocated to these firms:

$$\begin{aligned}
\pi^j_{do}(\tilde{\theta})\mathrm{d}\Theta^j_o(\tilde{\theta}) &\equiv \frac{\mathbb{E}[(\frac{\eta}{\eta-1}\tau^{Uj}_{do} p^j_o(\tilde{\theta}))^{1-\eta}]}{(P^j_d)^{1-\eta}} \cdot \mathrm{d}\Theta^j_o(\tilde{\theta}) \\
&= \frac{[\frac{\eta}{\eta-1}\tau^{Uj}_{do}]^{1-\eta}\mathbb{E}[p^j_{do'}(\tilde{\theta})^{1-\eta}]}{\sum_{o'}[\frac{\eta}{\eta-1}\tau^{Uj}_{do'}]^{1-\eta}\int \mathbb{E}[p^j_{do'}(\tilde{\theta})^{1-\eta}]\mathrm{d}\Theta^j_{o'}(\tilde{\theta})} \cdot \mathrm{d}\Theta^j_o(\tilde{\theta}) \\
&= \frac{\Gamma(1+\frac{1-\eta}{\lambda}) \cdot [\frac{\eta}{\eta-1}\tau^{Uj}_{do}]^{1-\eta} \cdot C^j_o(\tilde{\theta})^{1-\eta}}{\sum_{o'}\Gamma(1+\frac{1-\eta}{\lambda}) \cdot [\frac{\eta}{\eta-1}\tau^{Uj}_{do'}]^{1-\eta} \cdot (\bar{\Lambda}^j_{o'})^{1-\eta}} \cdot \mathrm{d}\Theta^j_o(\tilde{\theta}),
\end{aligned}$$

where the last equation follows the property of Weibull distribution, [1] and

$$\bar{\Lambda}^j_o \equiv (\int [C^j_o(\tilde{\theta})]^{1-\eta}\mathrm{d}\Theta^j_o(\tilde{\theta}))^{1/(1-\eta)}.$$

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

$$\pi^j_{do} = \int \pi^j_{do}(\tilde{\theta})\mathrm{d}\Theta^j_o(\tilde{\theta}) = \frac{[\tau^{Uj}_{do}\bar{\Lambda}^j_o]^{1-\eta}}{\sum_{o'}[\tau^{Uj}_{do'}\bar{\Lambda}^j_{o'}]^{1-\eta}}.$$

This completes the proof of Corollary 2.

A.2 Proof of Proposition 2

We restate Proposition 2 below:

Proposition A.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*

[1] Suppose random variable X follows a Weibull distribution with c.d.f. $F(x) = 1 - \exp[-(x/C)^\lambda]$, where $C > 0$ and $\lambda > 1$ are parameters. Then, for any $\varepsilon > 0$,

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

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

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.

$$g_d^i(\bar{\theta}) = (1-\omega^i)\bar{\theta} + \omega^i \sum_{j,o} \frac{\gamma^{ij}}{1-\gamma^{iL}} \int \chi_{do}^j(g_d^i(\bar{\theta}),\tilde{\theta}) g_o^j(\tilde{\theta}) \mathrm{d}\Theta_o^j(\tilde{\theta}), \quad \forall \bar{\theta} \in [-M, M] \quad \text{(A.6)}$$

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.*

We outline the proof and explain the intuition behind each step, with technical details deferred to subsequent lemmas. To simplify notation, we begin by introducing the following constants.

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*

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

We first characterize the optimality conditions for technology choices. Under Assumption 2, with the objective function specified in (15), the first-order condition with respect to technology choice θ is

$$\bar{\phi}(\bar{\theta}-\theta) = (\eta-1)\frac{\mathrm{d}}{\mathrm{d}\theta} \ln C_d^i(\theta) = (\eta-1)\bar{t} \cdot \sum_{j,o} \gamma^{ij} \int \chi_{do}^j(\theta,\tilde{\theta})(\theta-\tilde{\theta}) \mathrm{d}\Theta_o^j(\tilde{\theta}),$$

which can be rearranged into

$$\theta = \omega^i \sum_{j,o} \frac{\gamma^{ij}}{1-\gamma^{iL}} \int \chi_{do}^j(\theta,\tilde{\theta})\tilde{\theta} \mathrm{d}\Theta_o^j(\tilde{\theta}) + (1-\omega^i)\bar{\theta}, \quad \text{(A.7)}$$

with $\omega^i < 1$. The second-order condition requires that $-\bar{\phi} - (\eta-1)\frac{\mathrm{d}^2}{\mathrm{d}\theta^2} \ln C_d^i(\theta) < 0$.

To explicitly work with the policy function, we rewrite (A.7) into

$$g_d^i(\bar{\theta}) = \omega^i \sum_{j,o} \frac{\gamma^{ij}}{1-\gamma^{iL}} \int \chi_{do}^{ij}(\bar{\theta},\tilde{\theta};\boldsymbol{g},\boldsymbol{g}') g_o^j(\tilde{\theta}) \mathrm{d}\tilde{\theta} + (1-\omega^i)\bar{\theta}, \quad \text{(A.8)}$$

where, to slightly abuse notations and make them dependent on $\boldsymbol{g}$ and $\boldsymbol{g}'$ explicitly, [2]

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

$$\mathcal{C}^i_d(\bar{\theta};\boldsymbol{g},\boldsymbol{g}') \equiv \ln\xi^i_d - \sum_j\frac{\gamma^{ij}}{\zeta}\ln\left(\sum_o\int[\tau^j_{do}]^{-\zeta}\exp\left(-\zeta\mathcal{C}^j_o(\tilde{\theta};\boldsymbol{g},\boldsymbol{g}')\right)\exp\left(-\frac{1}{2}\zeta\bar{t}(g^i_d(\bar{\theta})-g^j_o(\tilde{\theta}))^2\right)[g^j_o]'(\tilde{\theta})\bar{\zeta}^j_o(\tilde{\theta})\mathrm{d}\tilde{\theta}\right). \tag{A.10}$$

Note that the derivative introduced by the change of variable in the integration has been incorporated into the definition of χ^{ij}_{do} in (A.9) (via the term $[g^j_o]'(\tilde{\theta})\bar{\zeta}^j_o(\tilde{\theta})$). Our goal is to prove the existence and uniqueness of an equilibrium with the policy function characterized by (A.8).

To show 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

$$\mathcal{G} = \{\boldsymbol{g}:[-M,M]\to\mathbb{R}^{N\times S}, \boldsymbol{g}\text{ is differentiable}, \|\boldsymbol{g}\|_\infty \le M;$$
$$[g^i_d]'(\bar{\theta}) \in [\underline{M}',\overline{M}']\text{ and }\boldsymbol{g}'\text{ is Lipschitz continuous with a Lipschitz constant }\overline{M}''\}.$$

Define an operator $\mathcal{T}$ on $\mathcal{G}$ as below (viewing $\boldsymbol{g}'$ as an operator mapping $\boldsymbol{g}$ to its first derivative):

$$[\mathcal{T}\boldsymbol{g}]^i_d(\bar{\theta}) = \omega^i\sum_{j,o}\frac{\gamma^{ij}}{1-\gamma^{iL}}\int\chi^{ij}_{do}(\bar{\theta},\tilde{\theta};\boldsymbol{g},\boldsymbol{g}')g^j_o(\tilde{\theta})\mathrm{d}\tilde{\theta} + (1-\omega^i)\bar{\theta}, \tag{A.11}$$

with $\omega^i < 1$. The fixed point of operator $\mathcal{T}$, if exists, solves the first-order condition of the technology adaptation problem, i.e. equation (A.8).

Intuitively, 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 $\bar{\chi}^{ij}_{do}(\tilde{\theta}) = \chi^{ij}_{do}(\bar{\theta},\tilde{\theta})$ 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 $\overline{\omega} \equiv \max_i\omega^i$.

Following this intuition, 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 first derivative of $\boldsymbol{g}$ being bounded also implies that the value of $\boldsymbol{g}$ is bounded on the domain $[-M,M]$. This is sufficient to establish compactness of the space of $\boldsymbol{g}$ and hence ensures existence.

The formal proof of existence consists of three steps. Firstly, in Lemma A.2, we establish the existence and uniqueness of $\boldsymbol{\mathcal{C}}$ given $\boldsymbol{g},\boldsymbol{g}'$ by formulating (A.10) 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 derivatives for any given $\boldsymbol{g},\boldsymbol{g}'$. Since $\boldsymbol{g}'$ enters the operator $\mathcal{T}$, we equip $\mathcal{G}$ with the C^1 norm:

[2] Note that $\chi^{ij}_{do}(\bar{\theta},\tilde{\theta};\boldsymbol{g},\boldsymbol{g}')$ takes the *endowed* technologies as the first two arguments, which is different from $\chi^j_{do}(\cdot,\cdot)$ in (A.6) that takes the chosen technologies as the arguments. Also, $\mathcal{C}^i_d(\bar{\theta};\boldsymbol{g},\boldsymbol{g}') = \ln C^i_d(g^i_d(\bar{\theta}))$.

$\|g\|_{\mathcal{G}} = \|g\|_\infty + \|g'\|_\infty$. It follows that χ is also continuous in g, g', and that $\mathcal{T}$ is continuous in g with the norm $\|\cdot\|_{\mathcal{G}}$.

Secondly, Lemma A.3 further characterizes the uniform bounds for $\mathcal{T}g$, $[\mathcal{T}g]'$ and the Lipschitz continuity of $[\mathcal{T}g]'$. This implies that $\mathcal{T}g \in \mathcal{G}$.

Lastly, in Proposition A.3, we show that since $g \in \mathcal{G}$ 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 $\mathcal{T}g \in \mathcal{G}$ 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.

For uniqueness, Part 2 of the proposition says that if further $\bar{\phi}$ is large enough, then the variation of $\chi^{ij}_{do}(\bar{\theta}, \tilde{\theta}; g, g')$ in (g, g') is bounded uniformly to ensure that $\mathcal{T}$ is a contraction. To show this, we treat (A.8) and (A.10) as a joint fixed point problem for $(g, \mathcal{C})$. We show that with $\bar{t}$ small enough and $\bar{\phi}$ large enough, the joint operator defined by the left-hand side of these equations forms a contraction mapping under the C^1 norm of the space of g combined with the C^0 norm of the space of $\mathcal{C}$. This is achieved by proving that the induced matrix norm—specifically, the maximum absolute row sum norm—of the Jacobian matrix containing the Fréchet derivatives of the operator with respect to $(g, \mathcal{C})$ is uniformly bounded by a constant less than one, provided that $\bar{t}$ is sufficiently small and $\bar{\phi}$ is sufficiently large. The bounds on the Jacobian entries are detailed in Lemma A.4, and the formal proof of uniqueness is presented in Proposition A.4.

Lemma A.2. *For $g \in \mathcal{G}$, there uniquely exists a $\mathcal{C}(\bar{\theta}; g, g')$ defined by (A.10) that is bounded and continuous in $(\bar{\theta}, g, g')$. Further, $\mathcal{C}$ satisfies*

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

(2) $\forall (g, g')$, $\mathcal{C}(\bar{\theta}; g, g')$ is differentiable in $\bar{\theta}$ and $\|\mathcal{C}'\|_\infty \le (1-\gamma^L) 2\bar{t} M \overline{M}'$.

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

Define $\mathbb{C} = \{\mathcal{C} : \mathcal{M} \times \mathcal{G} \times \widetilde{\mathcal{G}} \to \mathbb{R}^{N\times S}, \|\mathcal{C}\|_\infty \le M^{\mathcal{C}}, \mathcal{C} \text{ is continuous}\}$, 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 $\mathbb{C}$:

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

Since $\mathcal{C} \in \mathbb{C}$ is continuous, $\mathcal{T}^{\mathcal{C}}\mathcal{C}$ is also continuous. Since $\|g'\|_\infty < \overline{M}' < 1$ and

$$\exp\left(-\zeta \mathcal{C}^j_o(\tilde{\theta}; g, g')\right) \in [\exp(-\zeta M^{\mathcal{C}}), \exp(\zeta M^{\mathcal{C}})],$$

we have

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

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

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

We now verify that $\mathcal{T}^{\mathcal{C}}$ satisfies Blackwell's sufficiency conditions for contraction. For any $\boldsymbol{\mathcal{C}}, \widehat{\boldsymbol{\mathcal{C}}} \in \mathbb{C}$, such that $\boldsymbol{\mathcal{C}} \leq \widehat{\boldsymbol{\mathcal{C}}}$ point-wisely, it trivially holds that $[\mathcal{T}^{\mathcal{C}}\boldsymbol{\mathcal{C}}]_d^i \leq [\mathcal{T}^{\mathcal{C}}\widehat{\boldsymbol{\mathcal{C}}}]_d^i$. And it holds that $\mathcal{T}^{\mathcal{C}}[\boldsymbol{\mathcal{C}}(\bar{\theta};\boldsymbol{g},\boldsymbol{g}') + c]_d^i = [\mathcal{T}^{\mathcal{C}}\boldsymbol{\mathcal{C}}]_d^i(\bar{\theta};\boldsymbol{g},\boldsymbol{g}') + (1-\gamma^{iL})c \leq [\mathcal{T}^{\mathcal{C}}\boldsymbol{\mathcal{C}}]_d^i(\bar{\theta};\boldsymbol{g},\boldsymbol{g}') + (1-\underline{\gamma}^L)c$ 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

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

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

where the last line applies that $\sum_o \int \chi_{do}^{ij}(\bar{\theta},\tilde{\theta};\boldsymbol{g},\boldsymbol{g}')\mathrm{d}\tilde{\theta} = 1$ and $|g_d^i(\bar{\theta})| \leq M$. Note that the derivation for $[\mathcal{T}^{\mathcal{C}}\boldsymbol{\mathcal{C}}]_d^{i\,\prime}$ does not rely on the starting point $\boldsymbol{\mathcal{C}}$ being differentiable in $\bar{\theta}$. Thus, $\boldsymbol{\mathcal{C}}$ is differentiable in $\bar{\theta}$ and $\|\boldsymbol{\mathcal{C}}'\|_\infty \leq (1-\underline{\gamma}^L)2\bar{t}M\overline{M}'$.

□

Lemma A.3. *For $\boldsymbol{g} \in \mathcal{G}$ and $\mathcal{T}\boldsymbol{g}$ defined by (A.8), restated below*

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

(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.2, $\boldsymbol{\mathcal{C}}$ is continuous in $\boldsymbol{g}, \boldsymbol{g}'$ under the infinity norm. By the definition of $\boldsymbol{\chi}$ in (A.9), $\boldsymbol{\chi}$ 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. This proves part (1).

Next, 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 $\forall \bar{\theta} \in (\bar{\theta}^*, M]$. Since $\overline{M}' < 1$, we have that

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

Similarly, $\forall \bar{\theta} \in [-M, \bar{\theta}^*)$

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

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

Lastly, consider

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

where $cov_d^{ij}[\cdot,\cdot]$ is the covariance between the two arguments across $(o,\tilde\theta)$ under the distribution characterized by $\chi_{do}^{ij}(\bar\theta,\tilde\theta)$ and the second line applies Lemma A.5 stated below. Since $0 < [g_d^i]'(\bar\theta) < 1$ and by Popoviciu's inequality on variances $cov_d^{ij}[g_o^j(\tilde\theta), g_o^j(\tilde\theta)] \le M^2$, we have that

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

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

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

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 $\omega^i\zeta\bar{t}M^2\overline{M}'' + 3\omega^i\zeta\bar{t}M^3 \le \overline{M}''$ by the definition of $\overline{M}''$. This proves part (3).

□

Proposition A.3. *Given wages, assume $\{\bar\Theta_o^j\}$ have bounded support that is contained in $[-M,M]$ for a positive constant $M>0$ and have associated density functions $\bar\varsigma_o^j$. 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; $\|\boldsymbol{g}\|_\infty \le M$; $[g_d^i]'(\bar\theta) \in [\underline{M}', \overline{M}'], \forall(\bar\theta,d,i)$; $\boldsymbol{g}'$ 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, $\|\boldsymbol{g}\|_\infty \le M$; $[g_d^i]'(\bar\theta) \in [\underline{M}', \overline{M}']$, $\boldsymbol{g}'$ 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.3, $\mathcal{T}$ defined by (A.8) 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 $\mathcal{T}$.

For a fixed point $\boldsymbol{g}$, i.e. $\mathcal{T}\boldsymbol{g} = \boldsymbol{g}$, it solves the first-order condition of the technology adaptation problem. We now verify that the second-order optimality condition holds for the technology direction choice problem. With the objective function specified in (15), the second-order derivative of the objective with respect to choice θ is

$$-\bar{\phi} - (\eta - 1)\frac{\mathrm{d}^2}{\mathrm{d}\theta^2} \ln C_d^i(\theta)$$

where

$$\begin{aligned}
\frac{\mathrm{d}^2}{\mathrm{d}\theta^2} \ln C_d^i(\theta) &= \bar{t} \cdot \sum_{j,o} \gamma^{ij} \int \partial_\theta \chi_{do}^j(\theta, \tilde{\theta})(\theta - \tilde{\theta}) + \chi_{do}^j(\theta, \tilde{\theta}) \mathrm{d}\Theta_o^j(\tilde{\theta}) \\
&= \bar{t} \cdot \sum_{j,o} \gamma^{ij} \left[\int \partial_\theta \chi_{do}^j(\theta, \tilde{\theta})(\theta - \tilde{\theta}) \mathrm{d}\Theta_o^j(\tilde{\theta}) + 1 \right] \\
&= -\zeta \bar{t}^2 \sum_j \gamma^{ij} var_d^j(\tilde{\theta}) + \bar{t}(1 - \gamma^{iL}) \\
&\in [\bar{t}(1 - \gamma^{iL})(1 - \zeta \bar{t} M^2), \bar{t}(1 - \gamma^{iL})].
\end{aligned}$$

Here, $var_d^j(\tilde{\theta})$ in the third line is the variance taken under distribution $\chi_{do}^j(\theta, \tilde{\theta})$ across $(o, \tilde{\theta})$, and the line follows Lemma A.6. The last line applies that $\|\boldsymbol{g}\|_\infty \leq M$ and the Popoviciu's inequality. Therefore, the second-order condition of the technology choice problem

$$-\bar{\phi} - (\eta - 1)\frac{\mathrm{d}^2}{\mathrm{d}\theta^2} \ln C_d^i(\theta) < 0$$

holds globally, provided that $1 - \zeta \bar{t} M^2 \geq 0$. □

Lemma A.4. *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*

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

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

$$[\mathcal{T}^{\mathcal{C}}(\boldsymbol{g}, \boldsymbol{g}', \boldsymbol{\mathcal{C}})]_d^i(\bar{\theta}) = \ln \xi_d^i - \sum_j \frac{\gamma^{ij}}{\zeta} \ln \left(\sum_o \int [\tau_{do}^j]^{-\zeta} \exp\left(-\zeta \mathcal{C}_o^j(\tilde{\theta})\right) \exp\left(-\frac{1}{2}\zeta \bar{t}(g_d^i(\bar{\theta}) - g_o^j(\tilde{\theta}))^2\right) [g_o^j]'(\tilde{\theta}) \bar{\varsigma}_o^j(\tilde{\theta}) \mathrm{d}\tilde{\theta} \right),$$

where to slightly abuse notations, $\boldsymbol{\chi}$ *is the one defined in (A.9), but also highlighting the dependence on* $\boldsymbol{\mathcal{C}}$:

$$\chi_{do}^{ij}(\bar{\theta}, \tilde{\theta}; \boldsymbol{g}, \boldsymbol{g}', \boldsymbol{\mathcal{C}}) \equiv \frac{[\tau_{do}^j]^{-\zeta} \exp\left(-\zeta \mathcal{C}_o^j(\tilde{\theta})\right) \exp\left(-\frac{1}{2}\zeta \bar{t}(g_d^i(\bar{\theta}) - g_o^j(\tilde{\theta}))^2\right) [g_o^j]'(\tilde{\theta}) \bar{\varsigma}_o^j(\tilde{\theta})}{\sum_m \int [\tau_{dm}^j]^{-\zeta} \exp\left(-\zeta \mathcal{C}_m^j(\tilde{\tilde{\theta}})\right) \exp\left(-\frac{1}{2}\zeta \bar{t}(g_d^i(\bar{\theta}) - g_m^j(\tilde{\tilde{\theta}}))^2\right) [g_m^j]'(\tilde{\tilde{\theta}}) \bar{\varsigma}_o^j(\tilde{\tilde{\theta}}) \mathrm{d}\tilde{\tilde{\theta}}}. \tag{A.12}$$

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

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

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

Proof. Combining the proofs for part (2) and (3) of Lemma A.3, and the proof for part (1) of Lemma A.2 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 stated below:

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

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

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

Therefore,

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

Consider

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

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

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

Similarly,

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

which implies

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

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:

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

Apply part (2) of Lemma A.7:

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

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

$$\partial^k_{g,m}[\mathcal{T}^{\mathcal{C}}]^i_d(\bar\theta) = \sum_j \gamma^{ij} \sum_o \int \chi^{ij}_{do}(\bar\theta,\tilde\theta) \left[\bar t(g^i_d(\bar\theta) - g^j_o(\tilde\theta))(\partial^k_{g,m} g^i_d(\bar\theta) - \partial^k_{g,m} g^j_o(\tilde\theta))\right] \mathrm{d}\tilde\theta$$

$$\Rightarrow \sum_{m,k} \|\partial^k_{g,m}[\mathcal{T}^{\mathcal{C}}]^i_d\|_\infty \le (1-\underline{\gamma}^L) 2\bar t M.$$

$$\partial^k_{g',m}[\mathcal{T}^{\mathcal{C}}]^i_d(\bar\theta) = \sum_j \gamma^{ij} \sum_o \int \chi^{ij}_{do}(\bar\theta,\tilde\theta) \left[\partial^k_{g',m} \ln [g^j_o]'(\tilde\theta)\right] \mathrm{d}\tilde\theta$$

$$\Rightarrow \sum_{m,k} \|\partial^k_{g',m}[\mathcal{T}^{\mathcal{C}}]^i_d\|_\infty \le (1-\underline{\gamma}^L) \frac{1}{\underline{M}'}.$$

$$\partial^k_{\mathcal{C},m}[\mathcal{T}^{\mathcal{C}}]^i_d(\bar\theta) = \sum_j \gamma^{ij} \sum_o \int \chi^{ij}_{do}(\bar\theta,\tilde\theta) \partial^k_{\mathcal{C},m} \mathcal{C}^j_o(\tilde\theta) \mathrm{d}\tilde\theta$$

$$\Rightarrow \sum_{m,k} \|\partial^k_{\mathcal{C},m}[\mathcal{T}^{\mathcal{C}}]^i_d\|_\infty \le 1-\underline{\gamma}^L.$$

□

Proposition A.4. *Given wages, $\{w_d\}$, assume $\{\bar\Theta^j_o\}$ 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.3.*

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

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

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

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 $(\boldsymbol{g},\boldsymbol{\mathcal{C}})$:

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

Part (1) of Lemma A.4 shows that $\widetilde{\mathcal{T}}(\boldsymbol{g},\boldsymbol{\mathcal{C}}) \in \mathcal{X}$.

Consider $\forall (\boldsymbol{g}, \boldsymbol{\mathcal{C}}) \in \mathcal{X}, (\hat{\boldsymbol{g}}, \hat{\boldsymbol{\mathcal{C}}}) \in \mathcal{X}$

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

where the second line applies the definition of $\|\cdot\|_{\mathcal{X}}$, the third line applies the mean value theorem for Fréchet derivatives,[3] and $\mathcal{T}^g, \mathcal{T}^{g'}, \mathcal{T}^{\mathcal{C}}$ are the operators defined in Lemma A.4. From the estimates in part (2)-(4) of Lemma A.4

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

where $\Omega_g = \overline{\omega}\Big[3\zeta\bar{t}M^2 + 1\Big] + 3\overline{\omega}\zeta^2\bar{t}^2M^3 + (1 - \underline{\gamma}^L)2tM$, $\Omega_{g'} = \frac{\overline{\omega}\zeta M}{\underline{M}'} + 3\overline{\omega}\zeta^2\bar{t}\frac{M^2}{\underline{M}'} + \overline{\omega}\zeta t M^2 + (1 - \underline{\gamma}^L)\frac{1}{1-\overline{\omega}}$, $\Omega_{\mathcal{C}} = \overline{\omega}\zeta M + 3\overline{\omega}\zeta^2\bar{t}M^2 + 1 - \underline{\gamma}^L$. 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.3 also hold, we have that the unique fixed point also satisfies the properties stated in Proposition A.3.

□

Lemma A.5. *For $\boldsymbol{\chi}$ defined in (A.9) and any function $f_o(\tilde{\theta})$ we have*

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

where cov^{ij}_d is the covariance taken under the distribution $\chi^{ij}_{do}(\bar{\theta}, \tilde{\theta}; \boldsymbol{g}, \boldsymbol{g}')$ across $(o, \tilde{\theta})$.

[3] 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 simplicity we omit $\boldsymbol{g}, \boldsymbol{g}'$ in the arguments.

$$
\begin{aligned}
&\sum_o \int \partial_{\bar\theta} \chi^{ij}_{do}(\bar\theta, \tilde\theta) f_o(\tilde\theta) \mathrm{d}\tilde\theta \\
&= \sum_o \int \partial_{\bar\theta} \ln \chi^{ij}_{do}(\bar\theta, \tilde\theta) \cdot \chi^{ij}_{do}(\bar\theta, \tilde\theta) f_o(\tilde\theta) \mathrm{d}\tilde\theta \\
&= \sum_o \int \Big\{ -\zeta \bar t (g^i_d(\bar\theta) - g^j_o(\tilde\theta)) + \sum_{\tilde o} \int \chi^{ij}_{d\tilde o}(\bar\theta, \tilde{\tilde\theta}) \Big[\zeta \bar t (g^i_d(\bar\theta) - g^j_{\tilde o}(\tilde{\tilde\theta})) \Big] \mathrm{d}\tilde{\tilde\theta} \Big\} \cdot [g^i_d]'(\bar\theta) \cdot \chi^{ij}_{do}(\bar\theta, \tilde\theta) f_o(\tilde\theta) \mathrm{d}\tilde\theta \\
&= -\zeta \bar t [g^i_d]'(\bar\theta) \Big\{ \sum_o \int (g^i_d(\bar\theta) - g^j_o(\tilde\theta)) f_o(\tilde\theta) \chi^{ij}_{do}(\bar\theta, \tilde\theta) \mathrm{d}\tilde\theta \\
&\qquad - \Big(\sum_o \int (g^i_d(\bar\theta) - g^j_o(\tilde\theta)) \chi^{ij}_{do}(\bar\theta, \tilde\theta) \mathrm{d}\tilde\theta \Big) \Big(\sum_o \int f_o(\tilde\theta) \chi^{ij}_{do}(\bar\theta, \tilde\theta) \mathrm{d}\tilde\theta \Big) \Big\} \\
&= -\zeta \bar t [g^i_d]'(\bar\theta) cov^{ij}_d [g^i_d(\bar\theta) - g^j_o(\tilde\theta), f_o(\tilde\theta)] \\
&= \zeta \bar t [g^i_d]'(\bar\theta) cov^{ij}_d [g^j_o(\tilde\theta), f_o(\tilde\theta)].
\end{aligned}
$$

□

Lemma A.6. *For $\chi^j_{do}(\theta, \tilde\theta)$ defined in equation (6) and any function $f_o(\theta, \tilde\theta)$ we have*

$$
\sum_o \int \partial_\theta \chi^j_{do}(\theta, \tilde\theta) f_o(\theta, \tilde\theta) \mathrm{d}\Theta^j_o(\tilde\theta) = -\zeta \bar t \cdot cov^j_d(\theta - \tilde\theta, f_o(\theta, \tilde\theta)),
$$

where cov^j_d is the covariance under the distribution $\chi^j_{do}(\theta, \tilde\theta)$ across $(o, \tilde\theta)$.

Proof. Consider

$$
\begin{aligned}
&\sum_o \int \partial_\theta \chi^j_{do}(\theta, \tilde\theta) f_o(\theta, \tilde\theta) \mathrm{d}\Theta^j_o(\tilde\theta) \\
&= \sum_o \int \partial_\theta \ln \chi^j_{do}(\theta, \tilde\theta) \cdot \chi^j_{do}(\theta, \tilde\theta) f_o(\theta, \tilde\theta) \mathrm{d}\Theta^j_o(\tilde\theta) \\
&= \sum_o \int \Big\{ -\zeta \bar t (\theta - \tilde\theta) + \sum_{\tilde o} \int \chi^j_{d\tilde o}(\theta, \tilde{\tilde\theta}) \Big[\zeta \bar t (\theta - \tilde{\tilde\theta}) \Big] \mathrm{d}\Theta^j_{\tilde o}(\tilde{\tilde\theta}) \Big\} \cdot \chi^j_{do}(\theta, \tilde\theta) f_o(\theta, \tilde\theta) \mathrm{d}\Theta^j_o(\tilde\theta) \\
&= -\zeta \bar t \Big\{ \sum_o \int (\theta - \tilde\theta) f_o(\theta, \tilde\theta) \chi^j_{do}(\theta, \tilde\theta) \mathrm{d}\Theta^j_o(\tilde\theta) \\
&\quad - \Big(\sum_{\tilde o} \int (\theta - \tilde{\tilde\theta}) \chi^j_{d\tilde o}(\theta, \tilde{\tilde\theta}) \mathrm{d}\Theta^j_{\tilde o}(\tilde{\tilde\theta}) \Big) \Big(\sum_o \int f_o(\theta, \tilde\theta) \chi^j_{do}(\theta, \tilde\theta) \mathrm{d}\Theta^j_o(\tilde\theta) \Big) \Big\} \\
&= -\zeta \bar t \cdot cov^j_d(\theta - \tilde\theta, f_o(\theta, \tilde\theta))
\end{aligned}
$$

□

Lemma A.7. *For $\boldsymbol{\chi}$ defined in (A.12) and any function $f_o(\tilde\theta)$ we have*

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

(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),

$$
\begin{aligned}
&\sum_o \int \partial^k_{g,m} \chi^{ij}_{do}(\bar\theta,\tilde\theta) f_o(\tilde\theta) \mathrm{d}\tilde\theta \\
&= \sum_o \int \partial^k_{g,m} \ln \chi^{ij}_{do}(\bar\theta,\tilde\theta) \cdot \chi^{ij}_{do}(\bar\theta,\tilde\theta) f_o(\tilde\theta) \mathrm{d}\tilde\theta \\
&= -\zeta \sum_o \int \Big\{ \bar t (g^i_d(\bar\theta) - g^j_o(\tilde\theta))(\partial^k_{g,m} g^i_d(\bar\theta) - \partial^k_{g,m} g^j_o(\tilde\theta)) \\
&\quad - \sum_{\tilde o} \int \chi^{ij}_{d\tilde o}(\bar\theta,\tilde{\tilde\theta}) \Big[\bar t (g^i_d(\bar\theta) - g^j_{\tilde o}(\tilde{\tilde\theta}))(\partial^k_{g,m} g^i_d(\bar\theta) - \partial^k_{g,m} g^j_{\tilde o}(\tilde{\tilde\theta})) \Big] \mathrm{d}\tilde{\tilde\theta} \Big\} \chi^{ij}_{do}(\theta,\tilde\theta) f_o(\tilde\theta) \mathrm{d}\tilde\theta \\
&= -\zeta cov^{ij}_d \Big(\bar t (g^i_d(\bar\theta) - g^j_o(\tilde\theta))(\partial^k_{g,m} g^i_d(\bar\theta) - \partial^k_{g,m} g^j_o(\tilde\theta)), f_o(\tilde\theta) \Big) \\
&= -\zeta \Big\{ - \bar t g^i_d(\bar\theta) cov^{ij}_d (\partial^k_{g,m} g^j_o(\tilde\theta), f_o(\tilde\theta)) \\
&\quad - \bar t \partial^k_{g,m} g^i_d(\bar\theta) cov^{ij}_d (g^j_o(\tilde\theta), f_o(\tilde\theta)) + \bar t \cdot cov^{ij}_d (g^j_o(\tilde\theta) \partial^k_{g,m} g^j_o(\tilde\theta), f_o(\tilde\theta) \Big\}.
\end{aligned}
$$

For part (2),

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

For part (3),

$$
\begin{aligned}
&\sum_o \int \partial^k_{\mathcal{C},m} \chi^{ij}_{do}(\bar\theta,\tilde\theta) f_o(\tilde\theta) \mathrm{d}\tilde\theta \\
&= \sum_o \int \partial^k_{\mathcal{C},m} \ln \chi^{ij}_{do}(\bar\theta,\tilde\theta) \cdot \chi^{ij}_{do}(\bar\theta,\tilde\theta) f_o(\tilde\theta) \mathrm{d}\tilde\theta \\
&= -\zeta \sum_o \int \Big\{ \partial^k_{\mathcal{C},m} \mathcal{C}^j_o(\tilde\theta) - \sum_{\tilde o} \int \chi^{ij}_{d\tilde o}(\bar\theta,\tilde{\tilde\theta}) \Big[\partial^k_{\mathcal{C},m} \mathcal{C}^j_{\tilde o}(\tilde{\tilde\theta}) \Big] \mathrm{d}\tilde{\tilde\theta} \Big\} \chi^{ij}_{do}(\theta,\tilde\theta) f_o(\tilde\theta) \mathrm{d}\tilde\theta \\
&= -\zeta cov^{ij}_d \Big(\partial^k_{\mathcal{C},m} \mathcal{C}^j_o(\tilde\theta), f_o(\tilde\theta) \Big).
\end{aligned}
$$

□

A.3 Proof of Proposition 3

We state the proposition in full by writing out the coefficient matrix explicitly.

Proposition 3 (Full). *In the equilibrium of technology choice under degenerate endowment distributions:*

1. *The technology chosen by firms in (d,i) is*

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

where $\omega^i \equiv \frac{(\eta-1)(1-\gamma^{iL})\bar{t}}{(\eta-1)(1-\gamma^{iL})\bar{t}+\bar{\phi}}$, *and* $\bar{\chi}^{ij}_{do} \equiv \frac{\exp[-\zeta(\ln\tau^j_{do}+\ln\bar{C}^j_o+\frac{1}{2}\bar{t}(\theta^i_d-\theta^j_o)^2)]}{\sum_{o'}\exp[-\zeta(\ln\tau^j_{do'}+\ln\bar{C}^j_{o'}+\frac{1}{2}\bar{t}(\theta^i_d-\theta^j_{o'})^2)]}$ *is the share of sector-j input spent on country o by intermediate-good firms from (d,i). Thus, θ^i_d is a weighted average between the firm's endowment technology $\bar{\theta}^i_d$ and the endogenous technologies of its suppliers $\{\theta^j_o\}$.*

2. *Let $\bar{\boldsymbol{C}} \equiv (\bar{C}^1_1, \bar{C}^2_1, ..., \bar{C}^S_N)$ denote the location parameters of factory-gate price distributions (i.e., the degenerate special case of the location parameters defined in (5)). Then between two equilibria with different trade costs and technology choices:*

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

We use D_x to represent an $NS\times NS$ diagonal matrix with the diagonal elements being the entries of the $NS\times 1$ vector x.

$\tilde{\boldsymbol{\gamma}} \equiv (\tilde{\gamma}^1_1, \tilde{\gamma}^2_1, ..., \tilde{\gamma}^S_N)$ with $\tilde{\gamma}^i_d \equiv 1-\gamma^{iL}$. $\Omega \equiv [\mathbb{I}_{NS\times NS} - D_{\tilde{\boldsymbol{\gamma}}}\Gamma]^{-1}$ is the Leontief inverse of the expenditure share of sourcing, with $\Gamma^{ij}_{do} \equiv \frac{\gamma^{ij}}{1-\gamma^{iL}}\bar{\chi}^{ij}_{do}$. $\mathrm{d}\boldsymbol{\theta} \equiv (\mathrm{d}\theta^1_1, \mathrm{d}\theta^2_1, ..., \mathrm{d}\theta^S_N)$. $\mathrm{d}\ln\tilde{\boldsymbol{\tau}} \equiv (\mathrm{d}\ln\tilde{\tau}^1_1, \mathrm{d}\ln\tilde{\tau}^2_1, ..., \mathrm{d}\ln\tilde{\tau}^S_N)$ with $\mathrm{d}\ln\tilde{\tau}^i_d \equiv \sum_{oj}\Gamma^{ij}_{do}\mathrm{d}\ln\tau^j_{do}$ being 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 between trading pairs: $\Lambda \equiv D_{(\mathbb{I}_{NS\times NS}-\Gamma)\boldsymbol{\theta}} + \Gamma D_{\boldsymbol{\theta}} - D_{\boldsymbol{\theta}}\Gamma$.

3. *In response to changes in trade costs $\{\mathrm{d}\ln\tau^j_{do}\}$, firms change their technologies according to:*

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

where $\boldsymbol{\omega} = (\omega^1_1, \omega^2_1, ..., \omega^S_N)$ with $\omega^i_d = \omega^i$ defined in (23). $\widetilde{\Lambda}$ is an $NS\times NS$ matrix whose elements (indexed by $((d,i),(o,j))$) are $\widetilde{\Lambda}^{ij}_{do} \equiv \Gamma^{ij}_{do}[\theta^j_o - \sum_{\tilde{o}}\bar{\chi}^{ij}_{d\tilde{o}}\theta^j_{\tilde{o}}]$, the expenditure-weighted average distance between (o,j) and all suppliers of (d,i). $\widehat{\Lambda} \equiv -D_{\widetilde{\Lambda}\boldsymbol{\theta}} + \widetilde{\Lambda}D_{\boldsymbol{\theta}} - D_{\boldsymbol{\theta}}\widetilde{\Lambda}$. And $\mathrm{d}\ln\widehat{\boldsymbol{\tau}} \equiv (\mathrm{d}\ln\widehat{\tau}^1_1, \mathrm{d}\ln\widehat{\tau}^2_1, ..., \mathrm{d}\ln\widehat{\tau}^S_N)$ with $\mathrm{d}\ln\widehat{\tau}^i_d \equiv \sum_{jo}\widetilde{\Lambda}^{ij}_{do}\ln\tau^j_{do}$.

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

$$C^i_d(\theta) \propto \prod_j\Big(\sum_o\Big[\tau^j_{do}\bar{C}^j_o\exp(\tfrac{1}{2}\bar{t}(\theta-\theta^j_o)^2)\Big]^{-\zeta}\Big)^{-\frac{\gamma^{ij}}{\zeta}}.$$

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

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

FOC w.r.t. θ reads

$$-\bar{\phi}(\theta-\bar{\theta}_d^i)+(1-\eta)\sum_{j,o}\gamma^{ij}\frac{\exp[-\zeta(\ln\tau_{do}^j+\ln\bar{C}_o^j+\frac{1}{2}\bar{t}(\theta-\theta_o^j)^2)]}{\sum_{o'}\exp[-\zeta(\ln\tau_{do'}^j+\ln\bar{C}_{o'}^j+\frac{1}{2}\bar{t}(\theta-\theta_{o'}^j)^2)]}\cdot\bar{t}(\theta-\theta_o^j)=0.$$

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

$$\begin{aligned}\bar{\phi}(\theta_d^i-\bar{\theta}_d^i)&=(1-\eta)\bar{t}\sum_{j,o}\gamma^{ij}\bar{\chi}_{do}^{ij}(\theta_d^i-\theta_o^j)\\&=(1-\eta)\bar{t}(1-\gamma^{iL})\theta_d^i-(1-\eta)\bar{t}\sum_{j,o}\gamma^{ij}\bar{\chi}_{do}^{ij}\theta_o^j,\end{aligned}$$

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

$$\bar{\chi}_{do}^{ij}\equiv\frac{\exp[-\zeta(\ln\tau_{do}^j+\ln\bar{C}_o^j+\frac{1}{2}\bar{t}(\theta_d^i-\theta_o^j)^2)]}{\sum_{o'}\exp[-\zeta(\ln\tau_{do'}^j+\ln\bar{C}_{o'}^j+\frac{1}{2}\bar{t}(\theta_d^i-\theta_{o'}^j)^2)]}.$$

Rearranging gives

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

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, totally differentiating $\ln\bar{C}_d^i$ gives

$$\begin{aligned}\mathrm{d}\ln\bar{C}_d^i&=\sum_{j,o}\gamma^{ij}\chi_{do}^{ij}\Big[\mathrm{d}\ln\bar{C}_o^j+\bar{t}(\theta_d^i-\theta_o^j)\mathrm{d}(\theta_d^i-\theta_o^j)+\mathrm{d}\ln\tau_{do}^j\Big]\\&=\bar{t}(1-\gamma^{iL})[\Omega\iota]_d^i+(1-\gamma^{Li})[\Omega\mathrm{d}\ln\tilde{\tau}]_d^i,\end{aligned}$$

with

$$\iota_d^i=\sum_{j,o}\Gamma_{do}^{ij}(\theta_d^i-\theta_o^j)\mathrm{d}(\theta_d^i-\theta_o^j),$$

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

$$[\Omega\iota]_o^j=\sum_{mk}\Omega_{om}^{jk}\sum_{nh}\Gamma_{mn}^{kh}(\theta_m^k-\theta_n^h)(\mathrm{d}\theta_m^k-\mathrm{d}\theta_n^h)$$

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

$$\begin{aligned}\sum_{mk}\Omega_{om}^{jk}\sum_{nh}\Gamma_{mn}^{kh}\theta_m^k\mathrm{d}\theta_m^k&=[\Omega D_\theta\mathrm{d}\theta]_o^j\\\sum_{mk}\Omega_{om}^{jk}\sum_{nh}\Gamma_{mn}^{kh}\theta_n^h\mathrm{d}\theta_m^k&=[\Omega D_{\Gamma\theta}\mathrm{d}\theta]_o^j\\\sum_{mk}\Omega_{om}^{jk}\sum_{nh}\Gamma_{mn}^{kh}\theta_m^k\mathrm{d}\theta_n^h&=[\Omega D_\theta\Gamma\mathrm{d}\theta]_o^j\\\sum_{mk}\Omega_{om}^{jk}\sum_{nh}\Gamma_{mn}^{kh}\theta_n^h\mathrm{d}\theta_n^h&=[\Omega\Gamma D_\theta\mathrm{d}\theta]_o^j.\end{aligned}$$

Therefore,

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

with Λ defined in the proposition.

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

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

Therefore,

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

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

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

Therefore,

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

□

A.4 Proof of Proposition 4

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

$$\begin{aligned}
\mathrm{d}\theta^i_d &= \omega^i \sum_{j',o'} \gamma^{ij'} \bar{\chi}^{ij'}_{do'} (\mathrm{d}\theta^{j'}_{o'} + \theta^{j'}_{o'} \mathrm{d}\ln\chi^{ij'}_{do'}) \\
&= \underbrace{\omega^i \sum_{j',o'} \gamma^{ij'} \bar{\chi}^{ij'}_{do'} \mathrm{d}\theta^{j'}_{o'}}_{=0 \text{ noting the assumption implies } \bar{\chi}^{ii}_{dd}=0} \\
&\quad - \zeta\omega^i \sum_{j',o'} \gamma^{ij'} \bar{\chi}^{ij'}_{do'} \theta^{j'}_{o'} [\underbrace{\mathrm{d}\ln\bar{C}^{j'}_{o'}}_{=0} + \mathrm{d}\ln\tau^{ij'}_{do'} + \bar{t}(\theta^i_d - \theta^{j'}_{o'})\mathrm{d}\theta^i_d - \underbrace{\bar{t}(\theta^i_d - \theta^{j'}_{o'})\mathrm{d}\theta^{j'}_{o'}}_{=0}] \\
&\quad + \zeta\omega^i \sum_{j',o'} \gamma^{ij'} \bar{\chi}^{ij'}_{do'} \theta^{j'}_{o'} \sum_m \chi^{ij'}_{dm} [\underbrace{\mathrm{d}\ln\bar{C}^{j'}_{m}}_{=0} + \mathrm{d}\ln\tau^{ij'}_{dm} + \bar{t}(\theta^i_d - \theta^{j'}_{m})\mathrm{d}\theta^i_d - \underbrace{\bar{t}(\theta^i_d - \theta^{j'}_{m})\mathrm{d}\theta^{j'}_{m}}_{=0}] \\
&= -\zeta\omega^i\gamma^{ij}\bar{\chi}^{ij}_{do}\theta^j_o \mathrm{d}\ln\tau^{ij}_{do} - \zeta\bar{t}\omega^i [\sum_{j',o'} \gamma^{ij'} \bar{\chi}^{ij'}_{do'} \theta^{j'}_{o'} (\theta^i_d - \theta^{j'}_{o'})] \mathrm{d}\theta^i_d \\
&\quad + \zeta\omega^i\gamma^{ij}\bar{\chi}^{ij}_{do} (\sum_{o'} \bar{\chi}^{ij}_{do'} \theta^j_{o'}) \mathrm{d}\ln\tau^{ij}_{do} + \zeta\bar{t}\omega^i [\sum_{j',o'} \gamma^{ij'} \bar{\chi}^{ij'}_{do'} \theta^{j'}_{o'} (\theta^i_d - \sum_m \bar{\chi}^{ij'}_{dm} \theta^{j'}_m)] \mathrm{d}\theta^i_d \\
&= -\zeta\omega^i\gamma^{ij}\bar{\chi}^{ij}_{do}[\theta^j_o - \sum_m \bar{\chi}^{ij}_{dm}\theta^j_m] \mathrm{d}\ln\tau^{ij}_{do} + \underbrace{\zeta\bar{t}\omega^i \sum_{j',o'} \gamma^{ij'} \bar{\chi}^{ij'}_{do'} \theta^{j'}_{o'} [\theta^{j'}_{o'} - \sum_m \bar{\chi}^{ij'}_{dm} \theta^{j'}_m] \mathrm{d}\theta^i_d}_{\text{effect due to the change in } \theta^i_d \text{ through trade shares}}
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathrm{d}\theta^i_d &= \frac{-\zeta\omega^i\gamma^{ij}\bar{\chi}^{ij}_{do}[\theta^j_o - \sum_m \bar{\chi}^{ij}_{dm}\theta^j_m]\mathrm{d}\ln\tau^{ij}_{do}}{1 - \zeta\bar{t}\omega^i \sum_{j',o'} \gamma^{ij'}\bar{\chi}^{ij'}_{do'}\theta^{j'}_{o'}[\theta^{j'}_{o'} - \sum_m \bar{\chi}^{ij'}_{dm}\theta^{j'}_m]} \\
&= -\frac{\zeta\omega^i\gamma^{ij}\bar{\chi}^{ij}_{do}\|\theta^j_o - \vartheta^{ij}_d\|}{1 - \bar{t}\zeta\omega^i \sum_{j',o'} \gamma^{ij'}\bar{\chi}^{ij'}_{do'}\|\theta^{j'}_{o'} - \vartheta^{ij'}_d\|} \times \frac{\mathrm{d}\ln\tau^{ij}_{do}}{\theta^j_o - \vartheta^{ij}_d}.
\end{aligned}$$

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 $\Delta\|\theta^i_d - \theta^j_{o'}\| \approx \frac{1}{2}(\theta^i_d - \theta^j_{o'})(\mathrm{d}\theta^i_d - \mathrm{d}\theta^j_{o'}) \times x$, and $\mathrm{d}\theta^j_{o'} = 0$ by (d, i) being small, we have $\forall o', o$

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

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

It follows that

$$\begin{aligned}
&\Delta\|\theta_d^i - \theta_o^j\| - \sum_{o'}\bar{\chi}_{do'}^{ij}\Delta\|\theta_d^i - \theta_{o'}^j\| \\
&= -\frac{\zeta\omega^i\gamma^{ij}\bar{\chi}_{do}^{ij}\|\theta_o^j - \vartheta_d^{ij}\|}{1 - \bar{t}\zeta\omega^i\sum_{j',o'}\gamma^{ij'}\bar{\chi}_{do'}^{ij'}\|\theta_{o'}^{j'} - \vartheta_d^{ij'}\|} \times \frac{x}{\theta_o^j - \vartheta_d^{ij}}[(\theta_d^i - \theta_o^j) - \sum_{o'}\bar{\chi}_{do'}^{ij}(\theta_d^i - \theta_o^j)] \\
&= -\frac{\zeta\omega^i\gamma^{ij}\bar{\chi}_{do}^{ij}\|\theta_o^j - \vartheta_d^{ij}\|}{1 - \bar{t}\zeta\omega^i\sum_{j',o'}\gamma^{ij'}\bar{\chi}_{do'}^{ij'}\|\theta_{o'}^{j'} - \vartheta_d^{ij'}\|} \times \frac{x}{\theta_o^j - \vartheta_d^{ij}} \times (\vartheta_d^{ij} - \theta_o^j) \\
&= \frac{\zeta\omega^i\gamma^{ij}\bar{\chi}_{do}^{ij}\|\theta_o^j - \vartheta_d^{ij}\|}{1 - \bar{t}\zeta\omega^i\sum_{j',o'}\gamma^{ij'}\bar{\chi}_{do'}^{ij'}\|\theta_{o'}^{j'} - \vartheta_d^{ij'}\|} \times x,
\end{aligned}$$

where the denominator is positive by the second-order condition of θ_d^i. □

A.5 Proof of Proposition 5

Proof. From the first-order condition, for

$$\begin{aligned}
\theta_d^i(\nu) &= \frac{(\eta-1)\bar{t}}{(\eta-1)(1-\gamma^{iL})\bar{t} + \bar{\phi}}\sum_{j,o}\gamma^{ij}\bar{\chi}_{do}^{ij}(\nu)\theta_o^j + \frac{\bar{\phi}}{(\eta-1)(1-\gamma^{iL})\bar{t} + \bar{\phi}}\bar{\theta}(\nu) \\
&= \frac{(\eta-1)\bar{t}}{(\eta-1)(1-\gamma^{iL})\bar{t} + \bar{\phi}} \times \\
&\quad \sum_{j,o}\gamma^{ij}\frac{\exp[-\zeta(\ln\tau_{do}^j + \ln\bar{C}_o^j + \frac{1}{2}\bar{t}(\theta_d^i(\nu) - \theta_o^j)^2)]}{\sum_{o'}\exp[-\zeta(\ln\tau_{do'}^j + \ln\bar{C}_{o'}^j + \frac{1}{2}\bar{t}(\theta_d^i(\nu) - \theta_{o'}^j)^2)]}\cdot\theta_o^j + \frac{\bar{\phi}}{(\eta-1)(1-\gamma^{iL})\bar{t} + \bar{\phi}}\bar{\theta}(\nu).
\end{aligned}$$

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

$$\begin{aligned}
\mathrm{d}\theta_d^i(\nu) &= \frac{(\eta-1)\bar{t}}{(\eta-1)(1-\gamma^{iL})\bar{t} + \bar{\phi}}\sum_{j,o}\gamma^{ij}\theta_o^j\cdot\bar{\chi}_{do}^{ij}[\sum_m\bar{\chi}_{dm}^{ij}(\theta_o^j - \theta_m^j)]\cdot\mathrm{d}\theta_d^i(\nu) \\
&\quad + \frac{\bar{\phi}}{(\eta-1)(1-\gamma^{iL})\bar{t} + \bar{\phi}}\mathrm{d}\bar{\theta}(\nu) \\
\Rightarrow \mathrm{d}\theta_d^i(\nu) &= \left(\frac{\frac{\bar{\phi}}{(\eta-1)(1-\gamma^{iL})\bar{t} + \bar{\phi}}}{1 - \frac{(\eta-1)\bar{t}}{(\eta-1)(1-\gamma^{iL})\bar{t} + \bar{\phi}}\sum_{j,o}\gamma^{ij}\bar{\chi}_{do}^{ij}\|\theta_o^j - \vartheta_d^{ij}\|}\right)\mathrm{d}\bar{\theta}(\nu),
\end{aligned}$$

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

Therefore,

$$\begin{aligned}
&\mathrm{d}[\|\theta_d^i(\nu) - \theta_o^j\| - \|\theta_d^i(\nu) - \theta_{o'}^j\|] \\
&\quad = -2\mathrm{d}\theta_d^i(\nu)[\theta_o^j - \theta_{o'}^j],
\end{aligned}$$

whose sign agrees with $-2\mathrm{d}\bar{\theta}(\nu)[\theta_o^j - \theta_{o'}^j] = \mathrm{d}[\|\bar{\theta}(\nu) - \theta_o^j\| - \|\bar{\theta}(\nu) - \theta_{o'}^j\|]$.

The rest of the proposition follows from the following expression:

$$\bar{\chi}_{do}^{ij}(\nu) = \frac{\exp[-\zeta(\ln \tau_{do}^j + \ln \bar{C}_o^j + \frac{1}{2}\bar{t}(\theta_d^i(\nu) - \theta_o^j)^2)]}{\sum_{o'} \exp[-\zeta(\ln \tau_{do'}^j + \ln \bar{C}_{o'}^j + \frac{1}{2}\bar{t}(\theta_d^i(\nu) - \theta_{o'}^j)^2)]}.$$

□

A.6 Proof of Proposition 6

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

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

And the household utility in the decentralized equilibrium is

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

where

$$\begin{aligned}
\ln(P) &= cons + \sum_i \rho^i \ln(\bar{C}^i), \\
\ln(\bar{C}^i) &= \sum_j \gamma^{ij} \ln(\bar{C}^j) + \sum_j \tfrac{1}{2}\bar{t}\gamma^{ij}(\theta^i - \theta^j)^2 \\
\Rightarrow \ln(\bar{C}^i) &= \tfrac{1}{2}\bar{t}\sum_m \Omega^{im}[\sum_j \gamma^{mj}(\theta^i - \theta^j)^2] \\
\Rightarrow \ln(P) &= cons + \tfrac{1}{2}\bar{t}\sum_i \rho^i \sum_m \Omega^{im}[\sum_j \gamma^{mj}(\theta^i - \theta^j)^2]
\end{aligned}$$

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

The first-order derivative of $\ln(U)$ w.r.t. θ^i is given by

$$\frac{\frac{1}{\eta}\rho^i\exp(-\frac{1}{2}\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}{1-\frac{1}{\eta}\sum_i\rho^i\exp(-\frac{1}{2}\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}\bar{\phi}(\bar{\theta}^i-\theta^i)-\bar{t}\rho^i\sum_m\Omega^{im}[\sum_j\gamma^{mj}(\theta^i-\theta^j)]-\bar{t}\sum_{j\neq i}\rho^j\sum_m\Omega^{jm}[\gamma^{mi}(\theta^i-\theta^j)]$$

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

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

$$=\rho^i[\frac{\exp(-\frac{1}{2}\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}{\eta-\sum_i\rho^i\exp(-\frac{1}{2}\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}\bar{\phi}(\bar{\theta}^i-\theta^i)-\bar{t}\sum_j\tilde{\gamma}^{ij}(\theta^i-\theta^j)]-\bar{t}\sum_{j\neq i}\rho^j\tilde{\gamma}^{ji}(\theta^i-\theta^j) \quad \text{(A.13)}$$

This proves part (1).

For part (2), note that the decentralized θ^i satisfies

$$\frac{1}{\eta-1}\bar{\phi}(\bar{\theta}^i-\theta^i)=\bar{t}\sum_j\gamma^{ij}(\theta^i-\theta^j), \quad \text{(A.14)}$$

which implies that θ^i falls between $\bar{\theta}^i$ and θ^j. WOLG, assume $\bar{\theta}^i<\theta^i<\theta^j$.

Plugging (A.14) to (A.13) delivers

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

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

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

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

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

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

It follows that

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

Thus, the marginal effect of increasing θ^i on the social welfare is positive. In the case of $\bar{\theta}^i > \theta^i > \theta^j$, 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 have 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, meaning that symmetric wages will 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 $d = 1, 2$, household nominal income I_d in the decentralized equilibrium is given by

$$I_d = 1 + \frac{1}{\eta} I_d \exp(-\tfrac{1}{2}\bar{\phi}(\theta_d - \bar{\theta}_d)^2)$$
$$\implies I_d = \frac{1}{1 - \frac{1}{\eta}\exp(-\frac{1}{2}\bar{\phi}(\theta_d - \bar{\theta}_d)^2)}$$

Following Proposition 3, the technology choice of firms in country 2 is

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

with analogous expression for country 1 and

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

The symmetry of $\bar{\chi}$ stems from the symmetry of $\bar{C}_d$, following the symmetric assumption of the two countries.

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

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

$$U_d \propto \frac{\frac{1}{1-\frac{1}{\eta}\exp(-\frac{1}{2}\bar{\phi}(\theta_d-\bar{\theta}_d)^2)}}{[(\tau\exp(\frac{1}{2}\bar{t}(\theta_d-\theta_{-d})^2))^{-\zeta}+1]^{-\frac{1-\gamma^L}{\gamma^L}\frac{1}{\zeta}} \cdot (1+\tau^{1-\eta})^{\frac{1}{1-\eta}}}.$$

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

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

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

Also,

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

□

Appendix B Reduced-Form Evidence

We discuss the details for our reduced-form evidence in this section. Section B.1 introduces the data sources, dataset construction, and cleaning procedures. Section B.2 provides validation exercises for our text-based measure on technology similarity. Section B.3 introduces details in constructing the measure of compatibility intensity for each industry. Finally, Sections B.4 and B.5 present additional robustness checks for our reduced-form results.

B.1 Data

Our dataset is built from four main sources, the World Input-Output Tables, patent data from the PATSTAT Global, Chinese firm-level data, and tariff data from the TRAINS database. We explain in detail where we source the data and how we construct the final datasets.

World Input-Output Tables (WIOTs). Our multi-country analyses center around countries in the WIOTs (Timmer et al., 2015) sourced from the World Input-Output Database (2016 Release). To facilitate analysis, we classify all countries in the WIOTs into 29 regions based on their geographic and political proximity. Table B.1 lists the regions.

Patent Data. Our patent data are sourced from the PATSTAT Global (the 2023 Fall version), 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 one invention. An invention 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. We keep those patents with at least one application granted.

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 example, 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.

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 a yearly census of all state-owned manufacturing firms and all non-state manufacturing firms with sales greater than RMB 5 million (approx. US$600,000) over 1998–2014, including plant-level information on industry, location, sales, employment, etc.

Table B.1: Geo-Political Regions

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

Crucial to our analysis are a firm's identity (ω) and the prime industry (i) it belongs to. We link each firm consistently 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. To investigate the input-output linkage, we manually map the industry codes to the 3-digit industry classification used in the 2007 Chinese Input-Output Table.

To link Chinese patents to their patent portfolios, we use as a bridge patent application data provided by China's State Intellectual Property Office (SIPO). For each firm in the SIPO data, we link it to firms in the NBSC database by matching the names (in Chinese). We then trace individual patents in SIPO to patents in PATSTAT using application numbers that can be linked between the two databases. With this link, for each Chinese firm, we observe its global portfolio of patents, and their (forward and backward) citations with other worldwide patents.[4]

[4]For this bridge to work, our assumption is that when Chinese firms file a patent in any foreign office, they also file for a patent for the same invention in the SIPO. The reason we use PATSTAT, instead of the SIPO data only is that the former allows us to measure citations and similarity between Chinese firms and worldwide

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 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 firm-level 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 57,465 in the period 2000-2002 to 102,153 in the period 2012-2014.

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-product level, where importers and exporters are in three-letter ISO country code, and products are at the level of 6-digit HS code (2007).

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. Finally, we aggregate the tariff rates from yearly figures into three-year periods by computing simple averages.

B.2 Technology Similarity Measure Based on Patent Texts

Case Studies. In Section 3.1, we construct a measure of technology similarity based on the closeness of patent texts. To validate that our patent-based semantic similarity captures the level of compatibility between related technologies, we examine several cases of production of final products that have similar functions but follow different technological paths, which thus require different upstream technologies to be compatible.

Table B.2 shows the similarity scores between the IPC subclass (IPC4) of patents used by the production of a final product and that of its intimate upstream industry, averaged over 2000-2014. For the first example, the automotive sector, the technology class associated with electronic vehicles (B60L) has a similarity score of 0.81 with that of battery technology (H01M), while it has a similarity score of 0.78 with that of combustion engines (F02M). In contrast, the technology class associated with traditional fuel vehicles (B60K) has a similarity score of 0.74 with that of battery technology (H01M), and a higher similarity score of 0.87 with that of combustion engines (F02M). This matches the natural technological fit between different energy-based vehicles and the fundamental technology for converting the corresponding energy to motion.

The other example involves the technology transition from Cathode Ray Tube (CRT)-based display screens to flat panel displays (LCD and OLED). CRT technology (H01J) has a high similarity score (0.92) with plasma physics (H05H), reflecting CRT's core dependence on electron beams. In contrast, flat panel displays (G02F) show lower similarity with plasma physics (0.67) but higher similarity with organic electronic devices (H10K, 0.85). This aligns with the fundamental technological differences between CRT and flat panel displays: CRT

patents that are not covered by SIPO.

Table B.2: Similarity Scores Between Selected Technology Pairs

IPC4 of input using technology	IPC4 of input supplying technology	Similarity
B60L (Electric Vehicle)	H01M (Battery)	0.81
B60L (Electric Vehicle)	F02M (Fuel System)	0.78
B60K (Fuel Vehicle)	H01M (Battery)	0.74
B60K (Fuel Vehicle)	F02M (Fuel System)	0.87
G02F (Flat Panel Display)	H10K (Organic Display)	0.85
G02F (Flat Panel Display)	H05H (Plasma Tech)	0.67
H01J (CRT Display)	H10K (Organic Display)	0.78
H01J (CRT Display)	H05H (Plasma Tech)	0.92

Notes: This table presents similarity scores between embeddings of downstream technologies (first column) and their corresponding upstream technologies (second column). The table is divided into two panels demonstrating different technological paths. The upper panel contrasts electric vs. fuel vehicles in relation to their key input technologies (batteries vs. fuel systems). The lower panel contrasts display technologies (CRT vs. flat panel) and their associated manufacturing processes (plasma technology vs. organic electronic devices). Each similarity score represents the cosine similarity between the average embeddings (over 2000-2014) of the technology pairs.

Table B.3: Examples of Technology Description and Embedding-based Similarity

Text1 (Input using)	Text2 (Input supplying)	Sentence Embedding Similarity	TF-IDF Similarity
Electric vehicles need batteries, not gas engines	Batteries power electric cars, not gas cars	0.77	0.21
Electric vehicles need batteries, not gas engines	Gas engines power gas cars, not electric cars	0.76	0.21
Gas cars need engines, not batteries	Batteries power electric cars, not gas cars	0.70	0.21
Gas cars need engines, not batteries	Gas engines power gas cars, not electric cars	0.75	0.21

Notes: This table presents examples of technology descriptions and their similarity scores calculated using two different methods: sentence embeddings and TF-IDF. The descriptions represent technologies related to electric and gasoline vehicles, along with their upstream components (batteries and engines). The last two columns report the computed similarity measures for different technology pairs under the two methods.

relies on electron beams in vacuum tubes, while flat panel displays use solid-state electronic control. These examples demonstrate that our semantic similarity measure captures meaningful technological compatibility relationships along production chains.

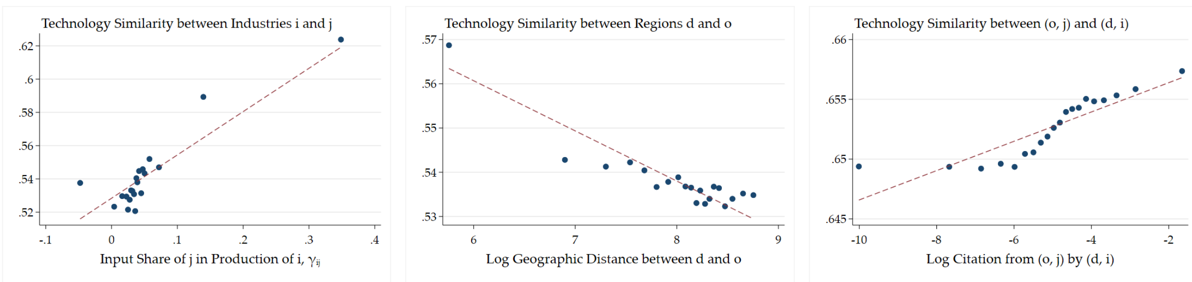
To illustrate what information the sentence embedding method encodes in terms of abstract semantic relationships, and to demonstrate its advantages over traditional word-count based measures like TF-IDF (used by e.g., Kelly, Papanikolaou, Seru and Taddy, 2021) for capturing degree of compatibility between technologies,[5] we construct an example of technology descriptions of electric vehicles (EVs) and gasoline vehicles, along with their respective upstream technologies: batteries and combustion engines. Table B.3 reports the example descriptions, and the similarity scores calculated for different technology pairs under differ-

[5]TF-IDF (Term Frequency-Inverse Document Frequency) is a numerical statistic that reflects how important a word is to a document. It combines how frequently a term appears in a document (TF) with how unique that term is across all documents (IDF). Words that appear frequently in one document but rarely in others receive higher TF-IDF scores.

Table B.4: Average Similarity Across Different IPC Levels

Classification Level	Average Similarity
Within same IPC subclass (IPC4)	0.40
Within same IPC class (IPC2)	0.33
Within same IPC section (IPC1)	0.31
All sample	0.29

Notes: This table is based on a sample of 10 patent abstracts from each IPC4 subclass, drawn from U.S. patents in 2014, totaling 5,616 patents. Bilateral cosine similarity scores were calculated for each pair of patents (excluding self-similarity), and the average similarity was computed for patents within different IPC levels.



(a) Input-Output Linkages (b) Geographic Proximity (c) Correlation with Citation

Figure B.1: Validation Checks for Technology Similarity

Notes: All figures are binned scatter plots. Figure B.1a plots the average similarity against the input-output weights between 2-digit industries i and j, controlling for i and j fixed effects. Figure B.1b plots the average similarity against the log of geographic distance between countries d and o, controlling for d and o fixed effects. Figure B.1c plots the average similarity against the log of citation flows between (d, i) and (o, j), controlling for d-i, o-j, d-o and i-j fixed effects.

ent methods.

The example descriptions intentionally include both positive and negative statements, which are common in patent abstracts to delineate a technology's applicability. The results show that sentence embeddings successfully capture the technological compatibility relationships: they assign higher similarity scores to naturally compatible pairs (EVs with batteries, gasoline vehicles with engines). In contrast, the TF-IDF method assigns identical similarity scores across all technology pairs as it relies solely on word overlap rather than semantic meaning.

These results show that sentence embeddings are particularly effective for understanding technological compatibility, as they can process and interpret complete descriptions rather than just individual terms, while word-count based methods remain valuable for other analytical purposes, such as identifying emerging technical terminology.

Statistical Validation. We now further validate this patent-text-based similarity measure by correlating it with other factors that are likely associated with technology proximity.

First, we show in Table B.4 that patent abstracts within the same technology class tend to have higher average similarity, with finer classification levels showing greater similarity.

Second, we show that sectors with stronger input-output linkages exhibit greater similarity than those with weaker connections. Figure B.1a shows a binned scatter plot where we regress the average similarity between industries i and j against γ^{ij}, the input weight from i in the production of j, controlling for the fixed effects of i and j. We observe a positive correlation that is statistically significant.

Third, we confirm that countries in close geographical proximity tend to have more similar patents. We show this with a binned scatter plot in Figure B.1b, where the average similarity between countries d and o is regressed on the logged geographic distance between the two countries, controlling for the fixed effects of d and o. There is a significantly negative correlation.

Last, Figure B.1c demonstrates a strong correlation between our text-based similarity measure and citation flows between country-industry pairs (d, i) and (o, j), even conditional on $d - i$, $o - j$, $d - o$ and $i - j$ fixed effects. This also motivates our robustness checks using citation as an alternative measure.

B.3 Measuring Compatibility Intensity

Since the importance of input compatibility may vary across industries, we construct a novel measure of compatibility intensity for each industry based on the full text of U.S. patents from the Orbis Intellectual Property Database, which a link full list of patents to global firms. For each 3-digit ISIC industry, We count the number of patenting firms, and the number of patenting firms with patents containing keywords 'compatibility/compatible' or 'interoperability.'

We use the share of patenting firms in an industry with such patents as a proxy for the likelihood that for a random firm in this industry, technology compatibility is important. Table B.5 lists the measure for all manufacturing industries at a 2-digit level. In the regressions, we further map this industry-specific measure of compatibility intensity from 3-digit ISIC to 3-digit CIC industries, denoted as $\text{Compatibility}_{i(\omega)}$.

A concern with this measure is that sometimes the two keywords can appear under a different meaning. For example, 'compatibility' can appear with other meanings such as 'water compatibility', 'skin compatibility.' As a robustness, we require at least one of the following words to appear in a patent containing 'compatibility' for it to be counted as about technological compatibility: technology, interface, system, software, hardware, standard, compliance, protocol, input, output, firmware, plug-and-play, backward, network, modular. This list of key words is generated by ChatGPT. Imposing this restriction does not materially alter either the intensity measure or the regression results based on the measure.

Table B.5: Compatibility Intensity by Industries

ISIC	Industry	Compatibility Intensity
21*	Manufacture of basic pharmaceutical products	0.52
60	Programming and broadcasting activities	0.48
12*	Manufacture of tobacco products	0.46
19*	Manufacture of coke and refined petroleum products	0.45
61	Telecommunications	0.42
20*	Manufacture of chemicals and chemical products	0.41
51	Air transport	0.36
63	Information service activities	0.35
26*	Manufacture of computer, electronic and optical products	0.35
58	Publishing activities	0.34
62	Computer programming, consultancy and related activities	0.33
53	Postal and courier activities	0.32
30*	Manufacture of other transport equipment	0.32
18*	Printing and reproduction of recorded media	0.29
32*	Other manufacturing	0.27
35	Electricity, gas, steam and air conditioning supply	0.26
52	Warehousing and support activities for transportation	0.25
29*	Manufacture of motor vehicles, trailers and semi-trailers	0.25
27*	Manufacture of electrical equipment	0.25
24*	Manufacture of basic metals	0.24
17*	Manufacture of paper and paper products	0.24
22*	Manufacture of rubber and plastic products	0.24
59	Motion picture, video and television programme production	0.23
11*	Manufacture of beverages	0.23
10*	Manufacture of food products	0.22
23*	Manufacture of other non-metallic mineral products	0.21
50	Water transport	0.21
13*	Manufacture of textiles	0.21
37	Sewerage	0.20
49	Land transport and transport via pipelines	0.20
33*	Repair and installation of machinery and equipment	0.20
46	Wholesale trade, except of motor vehicles and motorcycles	0.20
36	Water collection, treatment and supply	0.19
39	Remediation activities and other waste management services	0.19
56	Food and beverage service activities	0.19
42	Civil engineering	0.19
28*	Manufacture of machinery and equipment nec	0.18
47	Retail trade, except of motor vehicles and motorcycles	0.18
45	Wholesale and retail trade and repair of motor vehicles and motorcycles	0.18
16*	Manufacture of wood and of products of wood and cork, except furniture	0.16
41	Construction of buildings	0.16
43	Specialised construction activities	0.16
38	Waste collection, treatment and disposal activities; materials recovery	0.15
14*	Manufacture of wearing apparel	0.15
25*	Manufacture of fabricated metal products, except machinery and equipment	0.15
31*	Manufacture of furniture	0.12
55	Accommodation	0.10
15*	Manufacture of leather and related products	0.10

Notes: This table lists the compatibility intensity for two-digit industries in manufacturing and related services (Sections C-K). Compatibility intensity is defined as the fraction of patenting firms with patents containing keywords related to 'compatibility'. An ISIC code with a star suffix indicates a manufacturing industry.

B.4 Additional Details on Firm-Level Correlation

B.4.1 Input and Output Tariffs

In one of the specifications in Table 2, we control for the input and output tariffs faced by an industry i in China when sourcing inputs from abroad. We construct these variables from the effectively applied tariffs.

In general, output tariffs τ_{doit}^{Output} are defined as the weighted average of applied tariffs that region d (a collection of countries) imposes on products in industry i from region o (another collection of countries) in period t:

$$\tau_{doit}^{Output} \equiv \sum_{HS6 \in i} \tau_{do,HS6,t}^{Applied} \times \frac{X_{d,HS6,initial}}{\sum_{HS6' \in i} X_{d,HS6',initial}},$$

where $\tau_{do,HS6,t}^{Applied}$ are aggregated from the country-pair level by weighing with each pair's initial trade shares in the region

$$\tau_{do,HS6,t}^{Applied} \equiv \sum_{ISOd \in d} \sum_{ISOo \in o} \tau_{ISOd,ISOo,HS6,t}^{Applied} \times \frac{X_{ISOd,ISOo,HS6,initial}}{\sum_{ISOd' \in d} \sum_{ISOo' \in o} X_{ISOd',ISOo',HS6,initial}}.$$

For industry i, the input tariffs τ_{doit}^{Input} are defined as the averaged tariffs a firm in i faces when *sourcing inputs* from abroad. We construct this measure by weighing the import tariffs of products from other industries with their input-output weights and initial trade shares:

$$\tau_{doit}^{Input} \equiv \sum_{HS6 \notin i} \tau_{do,HS6,t}^{Applied} \times \gamma_{i,j(HS6)} \times \frac{X_{d,HS6,initial}}{\sum_{HS6' \in j(HS)} X_{d,HS6',initial}},$$

where $\gamma_{i,j(HS6)}$ denotes the industry-level input-output weight derived from the 2007 Chinese Input-Output Table. Throughout our firm-level regressions (Section 3.2), region d is restricted to China (CHN).

B.4.2 Extended Gravity

In this section, we provide empirical evidence for the 'extended-gravity'-like implication on the relationship between technology similarity and importing decisions. Our theory predicts that if regions o' and o are close to each other in the space of technologies, then the technology similarity between a firm ω and o' would be positively correlated with the firm importing from o.

To test this implication, we first define the set of technology neighbors of each region d, which are regions with technologies similar to those of region d. Specifically, we run the following regression:

$$\text{Similarity}_{doijt} = FE_{dij} + \varepsilon_{doijt},$$

where $\text{Similarity}_{doijt}$ denotes the technology similarity between region-industry pairs (d, i) and (o, j) in period t. We obtain the residuals from the regression, and calculate its mean over all industries i and j, which essentially capture the similarity of technologies in region o to d in period t, conditional on the fixed effects. Then, we define the *'technology neighbors'* of region d, denoted with $\mathcal{R}_t^{tech}(d)$, as the regions o's with the averaged residual ranking at the top 10% for d during period t.

Table B.6: Extended Gravity

	$\mathbb{I}[\text{Import}_{\omega ot} > 0]$			
	(1)	(2)	(3)	(4)
$\text{Similarity}_{\omega ot}$	0.014***	0.013***	0.024***	0.024***
	(0.005)	(0.005)	(0.009)	(0.009)
$\text{Similarity}_{\omega,\mathcal{R}_t^{tech}(o),t}$	0.083***	0.095***	0.156***	0.158***
	(0.031)	(0.030)	(0.046)	(0.046)
$\text{Similarity}_{\omega,\mathcal{R}^{geo}(o),t}$		0.003		-0.005
		(0.009)		(0.012)
$\text{Similarity}_{\omega,\mathcal{R}^{lang}(o),t}$			0.001	0.003
			(0.010)	(0.011)
FE ω-t	Yes	Yes	Yes	Yes
FE ω-o	Yes	Yes	Yes	Yes
FE o-t	Yes	Yes	Yes	Yes
$X_{i(\omega)ot}$	Yes	Yes	Yes	Yes
$\mathbb{I}[\text{Import}_{\omega,\mathcal{R}_t^{tech}(o),t} > 0]$	Yes	Yes	Yes	Yes
Observations	3,326,764	3,207,951	2,019,821	2,019,821
R^2	0.751	0.749	0.758	0.758

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered by firm. $R_t^{tech}(d)$, $\mathcal{R}^{geo}(o)$, and $\mathcal{R}^{lang}(o)$ represent the technological, geographical, and linguistic neighbors of region o, respectively. $X_{i(\omega)ot}$ includes the applied tariffs of industry i's inputs and outputs. Industry i's input tariff is defined by the average applied tariff on imports of all goods from world region o, weighted by each input's share in industry i's production. Industries are classified at the three-digit level of China's Industry Classification.

In column (1) of Table B.6, we extend our baseline specification (26) by including the similarity measure between firm ω and the technological neighbors of o. Because this similarity is likely correlated with firm ω's importing from these technological neighbors, and because importing from different regions can be substitutes of each other, we control for whether ω imports from these technological neighbors. The result shows that indeed, a firm ω's being technologically similar to the technology neighbors of a region o is strongly correlated with ω importing from o.

Still, it is possible that this result is driven by an information-based story: Firm ω meets a new contact, who introduces to ω both the suppliers in o and the technologies of o'. Such knowledgeable contacts certainly exist, but given the language and cultural barriers between countries, it is more likely than not that their knowledge domain extends to pairs of countries that are geographically or linguistically close to each other. In columns (2) to (4), we tease out the influence of such experts by controlling for the technology similarity between ω and the geographic or linguistic neighbors of o. We define geographic neighbors for a region as those with the closest 10 percentile distance, and language neighbors as those sharing a common official language. We find that the coefficients associated with proximity to language and geographic neighbors do not matter on themselves; more importantly, they do not diminish the role of similarity with technology neighbors.

Table B.7: Firm-Level Correlation between Trade and Citation

	$\mathbb{I}[\text{Import}_{\omega o t} > 0]$			
	(1)	(2)	(3)	(4)
$\mathbb{I}[\text{Citation}_{\omega o t} > 0]$	0.010***	0.010***	0.009***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)
FE ω-t	Yes	Yes	Yes	Yes
FE ω-o	Yes	Yes	Yes	Yes
FE o-t	Yes	Yes	Yes	Yes
$X_{i(\omega)ot}$		Yes		
FE $p(\omega)$-$i(\omega)$-o-t			Yes	Yes
Exclude Foreign Firms				Yes
Observations	3,426,052	3,349,472	3,361,232	2,517,536
R^2	0.749	0.748	0.766	0.729

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered by firm. $i(\omega)$ and $p(\omega)$ represent the industry and province of firm ω, respectively. $X_{i(\omega)ot}$ includes the applied tariffs of industry i's inputs and outputs. Industry i's input tariff is defined by the average applied tariff on imports of all goods from world region o, weighted by each input's share in industry i's production. Industries are classified at the three-digit level of China's Industry Classification.

B.4.3 Using Citation as an Alternative Measure

Researchers often use patent citations as an alternative measure of technology similarity. Although we have chosen the text-based measure of similarity as our preferred approach, Table B.7 replicates the baseline result in Section 3.2 using citations, showing its robustness to this alternative measure.

B.5 Additional Details on Tariff Variation and Technological Choice

B.5.1 MFN-Binding Tariffs

In Section 3.3, we instrument the effectively applied tariffs with the exogenous tariff changes from binding MFN tariffs. To construct this instrument, we first construct 'pseudo' MFN tariffs using the applied tariffs in the initial periods and *changes* in MFN rates at the country-product level:

$$\hat{\tau}^{MFN}_{ISOd,ISOo,HS6,t} \equiv \tau^{Applied}_{ISOd,ISOo,HS6,initial} + \sum_{t'=initial+1}^{t} \widehat{\Delta\tau}^{MFN}_{ISOd,ISOo,HS6,t'}$$

where the changes are defined as the interaction between the changes in MFN tariffs and an indicator for whether the MFN tariffs of d applies to o in both period t and $t-1$:

$$\widehat{\Delta\tau}^{MFN}_{ISOd,ISOo,HS6,t} \equiv \mathbb{I}\left[\tau^{Applied}_{ISOd,ISOo,HS6,t} = \tau^{MFN}_{ISOd,HS6,t}\right] \times \mathbb{I}\left[\tau^{Applied}_{ISOd,ISOo,HS6,t-1} = \tau^{MFN}_{ISOd,HS6,t-1}\right] \times \left(\tau^{MFN}_{ISOd,HS6,t} - \tau^{MFN}_{ISOd,HS6,t-1}\right).$$

Table B.8: Additional Evidence on Tariff Shocks and Technology Similarity

	$\ln \text{Citation}_{do,HS6,t}$		$\text{Similarity}_{do,HS6,t}$	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
$\ln \tau^{MFN}_{do,HS6,t}$	-0.153**		-0.010***	
	(0.076)		(0.001)	
$\ln \tau^{Applied}_{do,HS6,t}$		-0.187**		-0.013***
		(0.093)		(0.002)
FE *d-o-HS6*	Yes	Yes	Yes	Yes
FE *o-HS6-t*	Yes	Yes	Yes	Yes
FE *d-HS6-t*	Yes	Yes	Yes	Yes
Exclude Consumption Goods			Yes	Yes
Observations	4,240,509	4,240,509	4,338,370	4,338,370
R^2	0.937	0.937	0.994	0.994

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the *d-o-HS6* level. Columns 1 and 3 report reduced-form regressions using OLS, while columns 2 and 4 report the second stage of 2SLS regressions using $\ln \tau^{MFN}_{do,HS6,t}$ as an instrument for $\ln \tau^{Applied}_{do,HS6,t}$.

Then, we aggregate both the applied and 'pseudo' MFN tariffs into the region-product level by weighing on initial trade shares:

$$\tau^{MFN}_{do,HS6,t} \equiv \sum_{ISOd \in d} \sum_{ISOo \in o} \widehat{\tau}^{MFN}_{ISOd,ISOo,HS6,t} \times \frac{X_{ISOd,ISOo,HS6,initial}}{\sum_{ISOd' \in d} \sum_{ISOo' \in o} X_{ISOd',ISOo',HS6,initial}}$$

$$\tau^{Applied}_{do,HS6,t} \equiv \sum_{ISOd \in d} \sum_{ISOo \in o} \tau^{Applied}_{ISOd,ISOo,HS6,t} \times \frac{X_{ISOd,ISOo,HS6,initial}}{\sum_{ISOd' \in d} \sum_{ISOo' \in o} X_{ISOd',ISOo',HS6,initial}}.$$

The (log of) MFN-binding tariff rates $\tau^{MFN}_{do,HS6,t}$ are then used as the instrument for the (log of) effectively applied tariff rates $\tau^{Applied}_{do,HS6,t}$. Since our regressions control for d-o-HS6 fixed effects, this IV exploits the changes in the MFN rates for identification. When running the regressions, we exclude the largest exporter for each importer-product pair to address the concern that countries choose MFN tariffs with the major exporters in mind.

B.5.2 Citation as Alternative Measure and Excluding Final Goods

Table B.8 provides robustness checks for our results in Section 3.3. In the first two columns, we show that we obtain similar findings using citation as an alternative measure of bilateral similarity. In the last two columns, we show that the estimates are essentially the same when goods for final consumers are excluded. This is reassuring because our model centers on firms' choice of intermediate suppliers and technologies. Our definition of final consumption goods rely on the Classification by the Broad Economic Categories (BEC) and its mapping with HS products provided by the UNSD.

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 (28), and the ex-post technology distribution is characterized by (30). Then, under Assumption 3, by taking the log of (5), we have

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

where

$$\begin{aligned}
&k_{A,o}^j + m_A^j(\tilde{\theta}-n_{A,o}^j)^2 + \tfrac{1}{2}\bar{t}(\theta-\tilde{\theta})^2 \\
&= [m_A^j+\bar{t}](\tilde{\theta} - \frac{m_A^j n_{A,o}^j + \frac{1}{2}\bar{t}\theta}{m_A^j+\frac{1}{2}\bar{t}})^2 + k_{A,o}^j + \frac{m_A^j\frac{1}{2}\bar{t}(\theta-n_{A,o}^j)^2}{m_A^j+\frac{1}{2}\bar{t}},
\end{aligned}$$

which collects the quadratic term with respect to $\tilde{\theta}$ and then θ.

Since $\tilde{\theta} \sim N(\mu_o^j, [\sigma^j]^2), \tilde{\theta} - \frac{m_A^j n_{A,o}^j+\frac{1}{2}\bar{t}\theta}{m_A^j+\bar{t}} \sim N(\mu_o^j - \frac{m_A^j n_{A,o}^j+\frac{1}{2}\bar{t}\theta}{m_A^j+\bar{t}}, [\sigma^j]^2)$, apply Lemma A.8 stated below and we have

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

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

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

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

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

Consider

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

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

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 $\hat{\theta}^i_d$,[6] which gives

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

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

[6] The derivation works for any $\hat{\theta}^i_d$. In the numerical implementation, we choose $\hat{\theta}^i_d$ to be the mean of the ex-post technology distribution, μ^i_d.

This verifies the functional form in (28) with m^i_A, $n^i_{A,d}$ and $k^i_{A,d}$ satisfying

$$m^i_A = \sum_j \gamma^{ij} m^j_B = \sum_j \gamma^{ij} \frac{\frac{1}{2}\bar{t}[1 + 2\zeta m^j_A (\sigma^j)^2]}{1 + 2\zeta(m^j_A + \frac{1}{2}\bar{t})(\sigma^j)^2},$$

$$n^i_{A,d} = \sum_{j,o} \hat{\chi}^{ij}_{do} n^j_{B,o} = \sum_{j,o} \hat{\chi}^{ij}_{do} \frac{\mu^j_o + 2\zeta m^j_A (\sigma^j)^2 n^j_{A,o}}{1 + 2\zeta m^j_A (\sigma^j)^2},$$

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

Consider the innovation decision in (o, j) given by (15). Under Assumption 3, this is equivalent to

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

Taking log and apply the quadratic approximation to $\ln C^j_o(\theta)$, this is

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

The first-order condition implies that

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

which gives the policy function (29)

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

where

$$\alpha^j_o = \frac{(\eta - 1) m^j_A}{\frac{1}{2}\bar{\phi} + (\eta - 1) m^j_A} n^j_{A,o},$$

$$\beta^j = \frac{\frac{1}{2}\bar{\phi}}{\frac{1}{2}\bar{\phi} + (\eta - 1) m^j_A}.$$

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.

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.

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

where the last line applies that the integrand is a probability density function. □

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:

$$m^i_A = \sum_j \gamma^{ij} m^j_B \tag{C.3}$$

$$n^i_{A,d} = \sum_{j,o} \hat{\chi}^{ij}_{do} n^j_{B,o} \tag{C.4}$$

$$k^i_{A,d} = \ln C^i_d(\mu^i_d) - m^i_A[\sum_{j,o}\hat{\chi}^{ij}_{do}(\mu^i_d - n^j_{B,o})]^2 \tag{C.5}$$

$$\mu^i_d = \alpha^i_d + \beta^i \bar{\mu}^i_d \tag{C.6}$$

$$\sigma^i = \beta^i \bar{\sigma}^i \tag{C.7}$$

where

$$m_B^j = \frac{\frac{1}{2}\bar{t}[1 + 2\zeta m_A^j (\sigma^j)^2]}{1 + 2\zeta(m_A^j + \frac{1}{2}\bar{t})(\sigma^j)^2}, \tag{C.8}$$

$$n_{B,o}^j = \frac{1}{1 + 2\zeta m_A^j (\sigma^j)^2}\mu_o^j + \frac{2\zeta m_A^j (\sigma^j)^2}{1 + 2\zeta m_A^j (\sigma^j)^2} n_{A,o}^j \tag{C.9}$$

$$k_{B,o}^j = k_{A,o}^j + \frac{1}{2\zeta}\log[1 + 2\zeta(m_A^j + \tfrac{1}{2}\bar{t})(\sigma^j)^2] + \frac{m_A^j(\mu_o^j - n_{A,o}^j)^2}{1 + 2\zeta m_A^j (\sigma^j)^2} \tag{C.10}$$

$$\hat{\chi}_{do}^{ij} \equiv \frac{\gamma^{ij} m_B^j}{\sum_{j'} \gamma^{ij'} m_B^{j'}} \times \frac{\exp[-\zeta(\ln \tau_{do}^j + k_{B,o}^j + m_B^j(\mu_d^i - n_{B,o}^j)^2)]}{\sum_{o'} \exp[-\zeta(\ln \tau_{do'}^j + k_{B,o'}^j + m_B^j(\mu_d^i - n_{B,o'}^j)^2)]} \tag{C.11}$$

$$\ln C_d^i(\mu_d^i) = \ln(\frac{\Xi^i}{A_d^i}) + \gamma^{iL}\ln(w_d) - \zeta^{-1}\sum_j \gamma^{ij}\ln(\sum_o \exp[-\zeta(\ln \tau_{do}^j + k_{B,o}^j + m_B^j(\mu_d^i - n_{B,o}^j)^2)]) \tag{C.12}$$

$$\alpha_d^i = \frac{(\eta - 1)m_A^i}{\frac{1}{2}\bar{\phi} + (\eta - 1)m_A^i} n_{A,d}^i = (1 - \beta^i) n_{A,d}^i \tag{C.13}$$

$$\beta^i = \frac{\frac{1}{2}\bar{\phi}}{\frac{1}{2}\bar{\phi} + (\eta - 1)m_A^i}. \tag{C.14}$$

Here, (C.3) and (C.8) form a contraction mapping for $\{m_A^i\}$ with $\{\sigma^i, \beta^i\}$ determined by (C.7) and (C.14). With $\{m_A^i\}$ and $\{\sigma^j\}$ solved, (C.4) and (C.5) form another contraction mapping for $\{n_{A,d}^i, k_{A,d}^i\}$ when $\bar{t}$ is not too large, with $\{\mu_d^i, n_{B,o}^j, k_{B,o}^j, \chi_{do}^{ij}, \ln C_d^i(\mu_d^i), \alpha_d^i\}$ determined by (C.6), (C.9), (C.10), (C.11), (C.12) and (C.13), which can be evaluated directly.

Step 2. With objects solved in Step 1, the cost function $\{C_o^j(\theta)\}$ can be constructed from (C.2), and we can then explicitly evaluate the sourcing decisions of intermediate firms $\chi_{do}^j(\theta, \tilde{\theta})$ and final-good producers $\pi_{do}^j(\theta)$ for all $\theta, \tilde{\theta} \in \mathbb{T}$.

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

$$\begin{aligned}[\tau_{do}^j \Lambda_o^j(\theta)]^{-\zeta} &= (\tau_{do}^j)^{-\zeta}\int [C_o^j(\tilde{\theta})t(\theta, \tilde{\theta})]^{-\zeta} \mathrm{d}\Theta_o^j(\tilde{\theta}) \\ &= \exp[-\zeta(\ln \tau_{do}^j + k_{B,o}^j + m_B^j(\theta - n_{B,o}^j)^2)].\end{aligned}$$

Then, the probability density of sourcing from firms in country o with $\tilde{\theta}$ in (6) is

$$\chi_{do}^j(\theta, \tilde{\theta}) = \frac{[\tau_{do}^j C_o^j(\tilde{\theta})t(\theta, \tilde{\theta})]^{-\zeta}}{\sum_{o'}[\tau_{do'}^j \Lambda_{o'}^j(\theta)]^{-\zeta}} = \frac{\exp[-\zeta(\ln \tau_{do}^j + k_{A,o}^j + m_A^j(\tilde{\theta} - n_{A,o}^j)^2 + \frac{1}{2}\bar{t}(\theta - \tilde{\theta})^2)]}{\sum_{o'} \exp[-\zeta(\ln \tau_{do'}^j + k_{B,o'}^j + m_B^j(\theta - n_{B,o'}^j)^2)]}. \tag{C.15}$$

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

$$
\begin{aligned}
(\bar{\Lambda}_o^j)^{1-\eta} &= \int C_o^j(\tilde{\theta})^{1-\eta} \mathrm{d}\Theta_o^j(\tilde{\theta}) \\
&= \int \exp[(1-\eta)(k_{A,o}^j + m_A^j(\tilde{\theta} - n_{A,o}^j)^2)] \cdot \frac{1}{\sqrt{2\pi(\sigma^j)^2}} \exp[-\frac{1}{2}(\frac{\tilde{\theta}-\mu_o^j}{\sigma^j})^2] \mathrm{d}\tilde{\theta} \\
&= \exp[(1-\eta)k_{A,o}^j] \times [1 - 2(1-\eta)m_A^j(\sigma^j)^2]^{-1/2} \times \exp[\frac{(1-\eta)m_A^j(\mu_o^j - n_{A,o}^j)^2}{1-2(1-\eta)m_A^j(\sigma^j)^2}] \\
&= \exp[(1-\eta)\underbrace{(k_{A,o}^j - \frac{1}{2(1-\eta)}\log[1-2(1-\eta)m_A^j(\sigma^j)^2] + \frac{m_A^j(\mu_o^j - n_{A,o}^j)^2}{1-2(1-\eta)m_A^j(\sigma^j)^2})}_{\equiv k_{C,o}^j}]
\end{aligned}
$$

Then, the expenditure density allocated to goods from country o with $\tilde{\theta}$ in (10) is

$$
\pi_{do}^j(\tilde{\theta}) = \frac{[\tau_{do}^{Uj} C_o^j(\tilde{\theta})]^{1-\eta}}{\sum_{o'}[\tau_{do'}^{Uj}\bar{\Lambda}_{o'}^j]^{1-\eta}} = \frac{\exp[(1-\eta)(\ln \tau_{do}^{Uj} + k_{A,o}^j + m_A^j(\tilde{\theta} - n_{A,o}^j)^2)]}{\sum_{o'} \exp[(1-\eta)(\ln \tau_{do'}^{Uj} + k_{C,o'}^j)]}. \tag{C.16}
$$

Step 3. With the sourcing decisions specified, we can combine the market-clearing conditions (18) to (21) to arrive at a system of equations for $\{X_o^j(\theta)\}$ and $\{M_o^j(\theta)\}$, taking as given $\{w_d\}$. 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.

Specifically, since the policy function (29) is invertible, summing over the market-clearing conditions (20), (21) and (22), we get

$$
P_d Q_d = w_d L_d + \sum_i \int \Pi_d^i(g_d^i(\bar{\theta}); \bar{\theta}) \mathrm{d}\bar{\Theta}_d^i(\bar{\theta}) + \sum_i \int K_d^i(g_d^i(\bar{\theta}); \bar{\theta}) \mathrm{d}\bar{\Theta}_d^i(\bar{\theta}) \tag{C.17}
$$

$$
= w_d L_d + \sum_i \int [1 - \phi(\theta; (g_d^i)^{-1}(\theta)) + \phi(\theta; (g_d^i)^{-1}(\theta))] \frac{1}{\eta} X_d^i(\theta) \mathrm{d}\Theta_d^i(\theta) \tag{C.18}
$$

$$
= w_d L_d + \sum_i \int \frac{1}{\eta} X_d^i(\theta) \mathrm{d}\Theta_d^i(\theta). \tag{C.19}
$$

Substituting (C.15), (C.16) and (C.19) back to (18) and (19), 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 $\theta \in \widetilde{\mathbb{T}} \equiv \{\vartheta_1, \vartheta_2, ..., \vartheta_{N_\theta}\}$ 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. This 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

$$\begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{M} \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}_{X\to X} & \boldsymbol{B}_{M\to X} \\ \boldsymbol{B}_{X\to M} & \boldsymbol{B}_{M\to M} \end{bmatrix} \begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{M} \end{bmatrix}, \tag{C.20}$$

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 $\{\rho_d^i, \tau_{do}^j, \tau_{do}^{Uj}, \gamma^{ij}, \gamma^{iL}, \zeta, \eta\}$. [7]

The linear system (C.20) is homogeneous of degree one. By adding a normalized equation that total expenditures equal one, it can be 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 (22), is satisfied, i.e.,

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

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

Other statistics in equilibrium. 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.

The total profits earned by domestic firms can be calculated as

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

To calculate the price index, recall that the sectoral-level price indexes are

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

[7] Note that (C.20) is homogeneous of degree 1. This can be verified by summing over o, j, and $\tilde{\vartheta}$ on both sides.

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

$$
t_d \equiv \sum_i \sum_o \sum_j \int \int \frac{t(\theta, \tilde{\theta}) - 1}{t(\theta, \tilde{\theta})} \cdot \hat{M}_{do}^{ij}(\theta, \tilde{\theta}) \, \mathrm{d}\Theta_d^i(\theta) \mathrm{d}\Theta_o^j(\tilde{\theta}), \tag{C.21}
$$

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 adaptation costs spent by firms in each country are

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

C.3 Fitting the Posterior Technology Distributions

Before discussing the details of the calibration procedure, in this subsection, we explain how we solve the problem in (33) as the first step. Under the assumption that model similarity follows $sim(\mu_d^i, \mu_o^j) = \exp(-(\mu_d^i - \mu_o^j)^2)$, the fitting problem for $\{\mu_d^i\}$ solves

$$
\begin{aligned}
&\min_{\{\mu_d^i\}} \sum_{d,i,o,j} \left[\ln(sim(\mu_d^i, \mu_o^j) - \ln(sim_{do}^{ij,Data}) \right]^2 \\
&\qquad = \min_{\{\mu_d^i\}} \sum_{d,i,o,j} \left[(\mu_d^i - \mu_o^j)^2 - D_{do}^{ij} \right]^2
\end{aligned}
$$

where $sim_{do}^{ij,Data}$ is the empirical technology similarity between country-sector pairs (d, i) and (o, j), averaged over the period 2012-2014, and $D_{do}^{ij} \equiv -\ln(sim_{do}^{ij,Data})$ is a measure of bilateral distance. Instead of solving this original problem, which is known to be difficult for locating the global optimum, we transform it into a classical multidimensional scaling problem. This transformation recognizes the fact that if D_{do}^{ij} is a true bilateral distance matrix in Euclidean space, then by writing the summand of the objective into

$$
(\mu_d^i - \bar{\mu})^2 + (\mu_o^j - \bar{\mu})^2 - 2(\mu_d^i - \bar{\mu})(\mu_o^j - \bar{\mu}) - D_{do}^{ij},
$$

one can first demean the rows and columns of D_{do}^{ij}, and fit

$$\min_{\{\tilde{\mu}_d^i\}} \sum_{d,i,o,j} [\tilde{\mu}_d^i \tilde{\mu}_o^j - B_{do}^{ij}]^2, \quad s.t. \quad \sum_{d,i} \tilde{\mu}_d^i = 0 \tag{C.22}$$

where $\tilde{\mu}_d^i = \mu_d^i - \bar{\mu}$, and B_{do}^{ij} is the corresponding entry in the demeaned matrix B, defined as $B = -\frac{1}{2}CDC$. Here, $D = (D_{do}^{ij})$ is the bilateral distance matrix, and $C = I - \frac{1}{NS}J_{NS}$ is the centering matrix, with J_{NS} being an $NS \times NS$ matrix of ones.

The solution to the transformed problem (C.22) is uniquely given by the eigenvector of B associated with the largest eigenvalue, up to a scaling factor.[8] We further scale this eigenvector so that the standard deviation of the off-diagonal elements of the generated similarity matrix between country-sectors matches the data counterpart.

C.4 Algorithm for 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 adaptation costs $\bar{\phi}$, input incompatibility costs $\bar{t}$, and those determining production and trade $\{\tau_{do}^j, \tau_{do}^{Uj}, A_d^i\}$, 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 bilateral similarities between country-sectors, as in Appendix C.3. 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 (30).

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 $\{A_d^i\}$, 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. We describe the nested calibration algorithm below:

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

  (a) Choose a set of parameters $\{\tau_{do}^j, \tau_{do}^{Uj}, A_d^i\}$

[8] See e.g., Chapter 12 of Borg and Groenen (2007).

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

(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, adjust $\{\tau_{do}^j, \tau_{do}^{Uj}, A_d^i\}$ and return to Step (A)(b).

(B) Simulate 10,000 Chinese firms from each of the 19 sectors and regress the extensive margin of importing from each country o on the similarity between the technology of a firm and the technology of country o, controlling for firm and 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, adjust $\bar{t}$ and return to Step (A).

In Step (A)(b), we need to solve the trade equilibrium *conditional on* the ex-post distribution and parameters. This differs from the algorithm developed in Section C.2, which solves the general equilibrium (in which technology choice is endogenous). We modify the algorithm accordingly. Specifically, the system of equations from Step 1 of the algorithm in Section C.2 can be now simplified into

$$m_A^i = \sum_j \gamma^{ij} m_B^j$$

$$n_{A,d}^i = \sum_{j,o} \hat{\chi}_{do}^{ij} n_{B,o}^j$$

$$k_{A,d}^i = \ln C_d^i(\mu_d^i) - m_A^i [\sum_{j,o} \hat{\chi}_{do}^{ij} (\mu_d^i - n_{B,o}^j)]^2$$

where

$$m_B^j = \frac{\frac{1}{2}\bar{t}[1 + 2\zeta m_A^j (\sigma^j)^2]}{1 + 2\zeta (m_A^j + \frac{1}{2}\bar{t})(\sigma^j)^2},$$

$$n_{B,o}^j = \frac{1}{1 + 2\zeta m_A^j (\sigma^j)^2} \mu_o^j + \frac{2\zeta m_A^j (\sigma^j)^2}{1 + 2\zeta m_A^j (\sigma^j)^2} n_{A,o}^j$$

$$k_{B,o}^j = k_{A,o}^j + \frac{1}{2\zeta} \log[1 + 2\zeta (m_A^j + \tfrac{1}{2}\bar{t})(\sigma^j)^2] + \frac{m_A^j (\mu_o^j - n_{A,o}^j)^2}{1 + 2\zeta m_A^j (\sigma^j)^2}$$

$$\hat{\chi}_{do}^{ij} \equiv \frac{\gamma^{ij} m_B^j}{\sum_{j'} \gamma^{ij'} m_B^{j'}} \times \frac{\exp[-\zeta(\ln \tau_{do}^j + k_{B,o}^j + m_B^j (\mu_d^i - n_{B,o}^j)^2)]}{\sum_{o'} \exp[-\zeta(\ln \tau_{do'}^j + k_{B,o'}^j + m_B^j (\mu_d^i - n_{B,o'}^j)^2)]}$$

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

Given wages $\{w_d\}$ and parameters $\{\tau_{do}^j, \tau_{do}^{Uj}, A_d^i\}$, this is a contraction mapping with $\bar{t}$ being not too large and can be efficiently solved.

The remaining steps to solve the equilibrium follow exactly with

$$\chi^j_{do}(\theta,\tilde{\theta}) = \frac{\exp[-\zeta(\ln\tau^j_{do} + k^j_{A,o} + m^j_A(\tilde{\theta}-n^j_{A,o})^2 + \frac{1}{2}\bar{t}(\theta-\tilde{\theta})^2)]}{\sum_{o'}\exp[-\zeta(\ln\tau^j_{do'} + k^j_{B,o'} + m^j_B(\theta-n^j_{B,o'})^2)]} \tag{C.23}$$

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

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

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

The equilibrium is reached when wages $\{w_d\}$ satisfy 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 2012-2014.

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 Θ^i_d. With parameters $\{\tau^j_{do}, \tau^{Uj}_{do}, A^i_d\}$ 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 $\tilde{\chi}^j_{do}(\theta)$ given by

$$\tilde{\chi}^j_{do} = \int \chi^j_{do}(\theta,\tilde{\theta})\mathrm{d}\Theta^j_o(\tilde{\theta}),$$

with $\chi^j_{do}(\theta,\tilde{\theta})$ given by (C.23), and the average similarity between the technology of the firm and that of (o,j) is

$$\frac{1}{S}\sum_j sim(\theta,\mu^j_o) = \frac{1}{S}\sum_j \exp(-(\theta-\mu^j_o)^2).$$

Then, we regress the realized extensive margin of importing from each country o on the average similarity constructed above, controlling for firm and country-industry fixed effects. We calibrate $\bar{t} = 12.2$ to match this coefficient to 0.021, the extensive-margin import-similarity 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 adaptation. In other words, $\bar{\phi}$ only affects the welfare of agents but not other equilibrium prices and allocations. 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 adaptation costs. We calibrate it by matching the elasticity of similarity to tariffs (Column 3 of Table 4). 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 replicate the variation in MFN tariffs observed in the data, scaling the model's trade cost $\{\tau_{do}^{j}\}$ by the standard deviation of the MFN tariff for each (d, j) over time. We then treat the calibrated equilibrium and the counterfactual as two periods and regress the model's technology similarity between d and (o, j) on the logarithms of τ_{do}^{j}, controlling for d-o-j, o-j-t, and d-t fixed effects, mirroring the empirical specification. We adjust $\bar{\phi}$ until the model's elasticity matches the empirical coefficient of -0.007, yielding a calibrated value of $\bar{\phi} = 0.58$.

With $\bar{\phi}$ specified, we can then recover the ex-ante technology distribution by equations (29) and (30), i.e.,

$$\bar{\mu}_d^i = \frac{1}{\beta^i} \cdot \mu_d^i - \frac{1-\beta^i}{\beta^i} \cdot n_{A,d}^i$$

$$\bar{\sigma}^i = \frac{1}{\beta^i} \cdot \sigma^i$$

$$\text{with} \quad \beta^i = \frac{\frac{1}{2}\bar{\phi}}{\frac{1}{2}\bar{\phi} + (\eta - 1)m_A^i}.$$

This completes the calibration.