Trade and Technology Compatibility in General Equilibrium*

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

This Draft: October 2024
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 data. Quantification of the model shows that a semiconductor export embargo by the U.S. against China triggers technological decoupling between the two countries and realignment among other countries, doubling the global welfare loss caused by the embargo.

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

*We thank Jonathan Eaton, Thomas Chaney, 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 participants at Barcelona Summer Forum (geography, trade, and growth), Fudan Fanhai, Georgetown University, Hong Kong University, HKUST, Peking University (Guanghua; HSBC), the University of Nottingham, the 2023 Tsinghua Growth and Institutions Conference, the 2024 Asian Bureau of Economic Research (trade and growth), and the 2024 China International Conference in Macroeconomics for helpful comments.

1 Introduction

Recent global events have driven major economies to adopt strategies aimed at reducing their economic dependence on 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, such as high-end semiconductors. 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 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, increasing costs for firms serving multiple markets: 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, enabling us to assess whether specific policies lead to technological divergence or convergence. Second, we test and quantify the model using a new measure of technology similarity based on a textual analysis of global patents. We apply the model to assess the impacts of a U.S. embargo on Chinese 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 U.S. losses and about half of global losses.

In our model, firms differ in vertical efficiency and the endowment of a horizontal technology, denoted as a point $\bar{\theta}$ in the space of technologies $\mathbb{T}$, $\bar{\theta} \in \mathbb{T}$. Following the seminal Lancaster (1979), a point in the abstract $\mathbb{T}$ space represents a technological path—a specific combination of technological components and engineering specifications for a product.

[1] x86 is a CPU architecture originally developed by Intel. Chipmakers need a license to create CPUs compatible with x86, the foundation of many mainstream operating systems.

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

Nearby points represent similar, or compatible, combinations. These technologies are 'horizontal' in the sense that they are distinguished by their proximity to other technologies in the space, with no single technology being inherently better or worse. The endowment technology represents the path most suited to a firm's unique capabilities, shaped by factors such as accumulated know-how or the expertise of key personnel that are outside the scope of our model. The distribution from which firms draw their endowment varies across countries and sectors, reflecting the comparative advantage of different countries in specific technologies.

Given their endowment draw, firms choose a technology for production, represented by a point $\theta \in \mathbb{T}$ that can be different from $\bar{\theta}$. This decision balances adaptation costs, which increase with the distance between θ and $\bar{\theta}$, against the benefits of compatibility with key suppliers. By aligning their technology with that of potential suppliers, firms can reduce input costs by leveraging low-cost, compatible intermediate goods. Having chosen their technology, firms sample a set of suppliers, pick those that offer the lowest compatibility-adjusted prices for each intermediate input, and move forward with production. Products are then sold to final consumers and downstream firms. In equilibrium, given country-specific *exogenous* distributions from which firms draw their endowment technologies, firms anticipate the choices of all other firms and make their technology and supplier choices accordingly.

The equilibrium consists of a fixed point in the *endogenous* distribution of firms over their chosen technology in each country and industry—an infinite dimensional object. We establish that there generically exists an equilibrium in which firms' technology choice varies continuously with their endowment. Moreover, such an equilibrium is unique, provided that the cost of changing technologies is not too small and the benefit of compatibility is not too large. In the limiting case with degenerate endowment technology distributions and no benefit from compatibility, our model simplifies to the Caliendo and Parro (2015) model. Thus, the model retains the flexibility of canonical quantitative trade models.

Our model suggests novel relationship between trade and technology, which we characterize and later test. First, with compatibility incentives, firms with endowment technologies that are closer to a particular foreign country are more likely to adopt similar technologies and import from that country. This results in a *correlation* between a firm's decision to import from a country and technology similarity with that country. Second, a reduction in tariffs with a trading partner encourages firms in the importing country to shift their technology toward the partner, creating a *causal effect* of bilateral trade costs on technological similarity.

Our model also offers a fresh perspective on existing empirical patterns. For instance, firms that import from a certain country tend to also export to the same country. Exporters to a country are also likely to export to neighboring countries if those neighbors share similar technologies. These patterns, previously explained through state-dependent trade costs (Morales et al., 2019; Li et al., 2023), can be rationalized in our model through the lens of endogenous technological choice.

On the normative side, the model highlights a key externality in technology choice. As a firm selects its technology to maximize profit, it also affects the production costs of domestic and foreign downstream firms—an effect not internalized by the firm and amplified through input-output linkages. Because of these externalities, firms tend to underinvest in shifting away from their endowed technology. In a one-sector, two-symmetric-country special case of the model, we show that starting from the decentralized equilibrium, moving one country's technology choice toward that of the other *increases* welfare in *both* countries.

We test and quantify the importance of the model mechanisms. Crucial to our exercises is measures of horizontal technology similarity between firms and countries. We build a new measure using the text of global patents. Through a process known as 'embedding' by large language models, we obtain a numerical vector that summarises the meaning of a patent. We then construct bilateral similarities at different levels of aggregation using the cosine similarity between these vectors. We validate our measure by comparing it to other proxies of bilateral similarities. We also show that all our empirical results hold when we use patent citations to proxy for technology similarity.

We perform two tests. First, a premise of our model is that firms make joint technology and supplier choice with compatibility considerations. Using Chinese firm and customs data, we show that firms whose technology is more similar to a partner country is more likely to import from that country, a result robust to the inclusion of a rich set of controls and fixed effects. Additional evidence suggests that this is due to supplier technology compatibility, as opposed to technological diffusion or other information-based explanation. Second, our model implies that a decrease in bilateral trade barriers increase bilateral technology similarity. Using exogenous variations in HS-6 level bilateral tariffs arising from the changes in the most-favored-nation (MFN) tariffs, 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 in tariffs, offering support to the model.[3]

We parameterize the model to a sample of 29 major countries and 19 sectors. For tractability, we assume that $\mathbb{T}$ is the real line and firms in a country-sector cell draw their endowment technology from a respective Normal distribution. We show that, under a quadratic approximation to firms' technology choice problem, the distribution of technologies chosen by firms in a country-sector is a Normal distribution, with its mean and variance determined by the general equilibrium interactions. This enables us to solve the model efficiently.

Our model's calibration requires mapping the horizontal technologies of countries and sectors to the real line. We do this by targeting the bilateral technology similarities between pairs of country-sectors derived from patent texts.[4] To determine the economic impacts of

[3]The MFN tariffs could be chosen endogenously to increase trade. Following Boehm et al. (2023), we address this concern by excluding the largest trade partner for each importer-product from the regression.

[4]The goal of this quantitative exercise is not to determine why certain sectors, like automobiles, are closer to

technology choice, two parameters are key: the cost of incompatibility with suppliers and the cost of moving away from the endowment technology. We identify these parameters by matching the two reduced-form estimates: the firm-level correlation between importing and technology similarity and the elasticity of technology similarity in tariffs. Using these parameters and the calibrated technological distribution of countries, our model implies a trade cost component arising from technology incompatibility. We choose residual iceberg trade costs to match trade shares between countries.

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 lack of bilateral trade, trade costs need be well above 100%, often raising the question on the source of such high costs. Our model provides a partial explanation: technological incompatibility between countries.

We assess the impact of a trade conflict between the U.S. and China, focusing on a full embargo on Chinese semiconductor imports by the U.S. and allies. 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, firms in China diverge technologically from 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 an envelop theorem-based intuition that adjustments to negative shocks alleviate the loss, the divergence between China and major economies exacerbates the externality in technology choice and as a result, amplify the loss. Almost all of the U.S. welfare loss, and half of the global welfare loss, is due to endogenous technological changes.[5]

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 consider only two countries and two technologies (e.g., domestic vs. foreign), limiting their quantitative generalizability.[6,7] Instead of focusing on dichotomous technologies, we allow for multiple technologies and input-output

auto parts than to chairs. Instead, we focus on whether, for example, the U.S. auto industry are more similar to the German auto industry than to the Canadian one. Accordingly, we filtered out sector-pair fixed effects in our targets, focusing on capturing the variation in similarities between paris of countries.

[5] A recent literature has estimated the impacts of the recent trade war between the U.S. and China, focusing on the short-run effects (see e.g., Amiti et al., 2019; Fajgelbaum et al., 2020; Huang et al., 2018). Our finding suggests an important part of the welfare losses will appear gradually as technological adjustments take effect.

[6] 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) develop a two-country model to explore how horizontal standardization affects welfare. Given there are multiple countries in the world economy, each with potentially different technologies, mapping these technologies to the binary choice in the model can be subjective.

[7] 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) models technology transfer considering whether the technology is appropriate for a country.

linkages. Our approach builds on Lancaster (1979)'s work on endogenous product differentiation, but emphasizes compatibility with suppliers rather than consumers.

The network aspect of our model builds on recent works such as Jones (2011), Chaney (2014), Oberfield (2018), Lim (2018), Boehm and Oberfield (2020), Acemoglu and Azar (2020), Demir et al. (2021), Eaton et al. (2022), and Dhyne et al. (2023). 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 connects with Demir et al. (2021), which studies network formation under vertical quality choice. Our main contribution to this literature is to incorporate horizontal technology choice into a trade model with endogenous production networks.

Our paper also contributes to the quantitative trade literature (see Costinot and Rodriguez-Clare, 2014 for an early review). We generalize the canonical quantitative trade model with roundabout production to incorporate horizontal technology choice, which we show is important in shaping the effects of trade policies. To incorporate 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, because of market structure and distributional assumptions, the equilibrium can be characterized by aggregate quantities and prices only (e.g., Eaton and Kortum, 2002; Chaney, 2008; Caliendo and Parro, 2015; Lind and Ramondo, 2023; see Allen, Arkolakis and Li, 2023 for uniqueness results for some of these models).[8]

Finally, our paper connects with the literature on the relationship between trade and technology, which often uses patent citations to measure technology diffusion (e.g., Buera and Oberfield, 2020, Liu and Ma, 2021, Aghion et al., 2021, Cai et al., 2022, Lind and Ramondo, 2022). While most existing studies focus on how trade affects vertical technology, such as aggregate productivity, we focus on horizontal technology choices.[9] 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 offers a structural interpretation of the evidence, complementing existing studies emphasizing learning or other forms of technological diffusion.

[8]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 equilibrium existence and uniqueness. This approach does not apply to our model with externalities.

[9]Lind and Ramondo (2022) studies how trade affect the similarity of technologies among countries, modelling similarity through the correlation in productivity draws. In our model, technology similarity is an outcome of firms' deliberate choice.

2 Model

2.1 Environment.

There are N countries, denoted by d or o, and S sectors, denoted by i or j. In each country d, there is a representative household that supplies L_d units of labor, and a representative producer of a non-tradable final good with the following production technology:

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

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

In each country-sector (d, i) or (o, j), there is a unit mass of firms, denoted $\nu, \omega \in (0, 1)$, each producing one differentiated variety using labor and other intermediate goods. Firms charge marginal cost when selling to other intermediate producers but charge a monopolistically competitive markup when selling to the final good producer.

Intermediate good producers choose a horizontal technology, balancing the desire to use the technology their know best (their endowment 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 high-dimensional Euclidean space.[10] Let $\bar{\theta}(\nu) \in \mathbb{T}$ denote the technology firm ν knows the best (its endowment). Given $\bar{\theta}(\nu)$, firm ν first chooses a technology for production, denoted by $\theta(\nu) \in \mathbb{T}$. Choosing a distant technology from $\bar{\theta}(\nu)$ is costly but also brings benefits: the firm can use intermediate inputs more efficiently if its technology is close to that of the supplier. Having chosen its technology, each firm randomly samples a set of suppliers and choose the one with the lowest input-efficiency-adjusted price. Then all firms produce and sell their output to other firms and final good producers simultaneously.

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

2.2 Intermediate Firms Production and Sourcing

Firm ν in country-sector (d, i) who has chosen technology $\theta(\nu)$ can access a random set of production *techniques*, denoted by $R(\nu)$. A technique $r \in R(v)$ is characterized by (i) its

[10] 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.

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. Suppliers for technique r from (o, j) are denoted by $\Omega_o^j(\nu, r)$.

If firm ν adopts technique r, its 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 ν, and $m^j(\nu, r)$ is the intermediate input from sector j. The Cobb-Douglas structure implies that each input j has a revenue share of γ^{ij}. As there is no love-for-variety in production, for each input firm ν chooses one, and only one supplier.

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

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

where $c^j(\nu, r)$ is the minimum cost for input j among the available suppliers in r, i.e.,

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

Here, $\tilde{c}^j(\nu, \omega)$, the (compatibility-adjusted) unit cost of the intermediate good produced by supplier $\omega \in \Omega_o^j(\nu, r)$, is defined as

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

The first three components of the cost are standard: τ_{do}^j is the iceberg trade cost, $p(\omega)$ is the factory-gate price of supplier ω, $z(\nu, \omega)$ is an idiosyncratic match-specific efficiency draw. The novel component is $t(\theta(\nu), \theta(\omega))$, which is an increasing function in the difference between $\theta(\nu)$ and $\theta(\omega)$. It captures the idea that intermediate goods are more useful if the technologies they embody are consistent with that of the user.

Each firm chooses among $R(\nu)$ the technique that maximizes its profit, which accrues solely from final good producers. This is equivalent to minimizing the factory-gate price:

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

For tractable aggregation, we make the following assumption on the available techniques.

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

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

2. *For any production technique r, suppliers in $\Omega_o^j(\nu, r)$ receive independent match-specific efficiency draws with firm ν. For any $z_1 < z_2 \in (0, +\infty)$, the number of suppliers in $\Omega_o^j(\nu, r)$ for whom the match-specific efficiency draws $z(\nu, \omega) \in (z_1, z_2]$ follows a Poisson distribution*

with mean $z_1^{-\zeta} - z_2^{-\zeta}$, *where* $\zeta > 1$.

Assumption 1 specifies the distribution of productivity draws for techniques and that of match quality draws for suppliers. The Poisson assumption, which has been employed in Boehm and Oberfield (2020) in a closed-economy setting, can be viewed as an extension to the more familiar Pareto assumption (Chaney, 2008; Kortum, 1997).[11] In part (1), A_d^i determines the efficiency of techniques available to firms in (d, i)—with higher A_d^i implying more techniques with high efficiency; λ, only other hand, governs the heterogeneity these techniques.[12] Parameter ζ in part (2) plays a similar role to that of λ in part (1), governing the heterogeneity of match qualities among the set of suppliers available to a technique.

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

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

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

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

where Ξ^i is a sector-specific constant and Θ_o^j is the measure of firms in (o, j) choosing technology $\tilde{\theta}$.[13]

Proof. See Appendix A.1. □

To establish this result, we show that if the factory-gate prices of suppliers from each (o, i) follows a Weibull distribution, then the prices of input-using firms in (d, j) also follows a Weibull distribution. With the benefit from technology compatibility, firms choosing different θ have different factory-gate price distributions. We denote by $C_d^i(\theta)$ the location parameter of the distribution among the firms in (d, i) that have chosen technology θ. Since all firms are simultaneously users and producers of intermediate goods, in equilibrium, $C_d^i(\theta)$ also depends on $\{C_o^j(\theta) : \theta \in T\}_{o=1,j=1}^{N,S}$, forming a fixed point described by equation (3).

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

[12] One can see both aspects by setting α_2 to ∞. In this limit, for any α_1, A_d^i determines the number of draws that are above α_1; λ determines how fast these number decreases with α_1.

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

Equation (3) highlights the forces that determine the price distribution in country-sector (d, i). The typical forces in trade models with input-output linkages are present: the effective productivity of sector (d, i), represented by $\frac{\Xi^i}{A^i_d}$, the wage in country d denoted by w_d, iceberg transport cost for importing intermediate inputs $\{\tau^j_{do}\}^{N,S}_{o=1,j=1}$, and the parameters that govern the price distribution of suppliers $\{C^j_o(\theta) : \theta \in T\}^{N,S}_{o=1,j=1}$. The novel force arising from compatibility with suppliers is captured in $t(\theta, \tilde{\theta})$. Given the extreme-valued draws for match-specific efficiency, firms' best suppliers could come from any technology, so $C^i_d(\theta)$ depends on an integration over Θ^j_o.

Because of compatibility incentive, firms' sourcing decision depends on their technology choice. We characterize the sourcing decision as a corollary of Proposition 1:

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

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

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

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

$$\Lambda^j_o(\theta) \equiv \left(\int [C^j_o(\tilde{\theta}) t(\theta, \tilde{\theta})]^{-\zeta}\mathrm{d}\Theta^j_o(\tilde{\theta})\right)^{-1/\zeta}. \tag{5}$$

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

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

Proof. See Appendix A.1. □

In the first part of the corollary, $\Lambda^j_o(\theta)$ captures the overall competitiveness of country o as a supplier to a firm in d. Intuitively, if firms with technology $\tilde{\theta}$ in o offer incompatibility-adjusted input costs (smaller $C^j_o(\tilde{\theta})t(\theta, \tilde{\theta})$) or are more numerous (larger $\Theta^j_o(\tilde{\theta})$), then they have a higher chance of becoming a supplier. The integration in equation (5) reflects that a country-sector (o, j) is on average more competitive as a source, if it is competitive across the distribution Θ^j_o. The second part of the corollary aggregate across firms in (o, j) to derive the probability that a firm in (d, i) with technology θ sources input j from any firm in country o.

Our model is a generalization of the Caliendo and Parro (2015) model. To see the connection, when technology is fixed and homogeneous across firms and countries, $C^i_d(\theta)$ specializes to a scalar $\bar{C}^i_d$, reflecting the average output cost in (d, i); $\Lambda^j_o(\theta)$ reduces to $\bar{C}^j_o$; trade shares $\chi^j_{do}(\theta)$ also matches the form in Caliendo and Parro (2015). The model thus retains the flexibility of the canonical trade models while incorporating endogenous horizontal technology choices, which we describe in the rest of this section.

2.3 Production and Sourcing Decisions: Final-Good Firms

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

$$P_d Q_d - \sum_j \sum_o \int_0^1 [\frac{\eta}{\eta-1}\tau_{do}^{Uj} p_{do}^j(\omega)] q_{do}^j(\omega)\mathrm{d}\omega, \tag{7}$$

where the factory-gate price $p_{do}^j(\omega)$ follows the distribution characterized by (2), τ_{do}^{Uj} is the iceberg trade cost faced by final-good producers,[14] and P_d is the price of the final good in d,

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

The sourcing decision of the final-good producer differs from that of the intermediate-good producers in two aspects. First, instead of using only one input in each sector, the final good producer buy from all producers because they benefit from the love for variety. Second, they are unaffected by technological compatibility considerations. We rationalize this assumption by the notion that, ultimately, consumers derive utility more significantly from the services and functionalities enabled by the products that meet their needs rather than from the underlying technology of these products. We characterize the sourcing decision of final-good producers as the second corollary of Proposition 1.

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

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

$$\pi_{do}^j(\tilde{\theta})\mathrm{d}\Theta_o^j(\tilde{\theta}) = \frac{[\tau_{do}^{Uj} C_o^j(\tilde{\theta})]^{1-\eta}}{\sum_{o'}[\tau_{do'}^{Uj}\bar{\Lambda}_{o'}^j]^{1-\eta}}\mathrm{d}\Theta_o^j(\tilde{\theta}), \tag{8}$$

$$\textit{where} \quad \bar{\Lambda}_o^j \equiv \big(\int [C_o^j(\tilde{\theta})]^{1-\eta}\mathrm{d}\Theta_o^j(\tilde{\theta})\big)^{1/(1-\eta)}. \tag{9}$$

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

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

Proof. See Appendix A.1. □

Here, $\bar{\Lambda}_o^j$ measures the overall competitiveness of suppliers in (o, j). Because final-good producers do not care about comparability, $t(\theta, \tilde{\theta})$ does not appear in $\bar{\Lambda}_o^j$.

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

2.4 Technology Choice

Upon entry, each firm in country-sector (d, i) draws its technology endowment $\bar{\theta}$ from an exogenous probability measure $\bar{\Theta}_d^i$ on $\mathbb{T}$, referred to as the *ex-ante* technology distribution. The firm then selects a desired technology θ and adapts its production process accordingly before making production and sourcing decisions. Adapting to a technology different from the endowment is costly but can increase profit by facilitating input sourcing from efficient firms with similar technology.

Recall that firms earn profits solely through sales to final-good producers. Let $X_d^i(\theta)$ represent the *expected* sales to final-good producers by a firm in (d, i) with technology choice θ, with the expectation taken over potential draws of production techniques and suppliers. Firms' technology adaptation cost is:

$$K_d^i(\theta; \bar{\theta}) \equiv \phi(\theta, \bar{\theta}) \cdot \frac{1}{\eta} X_d^i(\theta), \tag{11}$$

where $\frac{1}{\eta} X_d^i(\theta)$ the firm's expected profits from the fixed markup, and $\phi(\theta, \bar{\theta})$ denotes the fraction of profits spent on technology adaptation. The function $\phi(\theta, \bar{\theta})$ increases with the difference between θ and $\bar{\theta}$, reflecting that adapting to a more distinct technology requires greater investment.[15]

Firms maximize profits net of technology adaptation costs by choosing θ to solve:

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

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

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

Denote the optimal technology choice by $g_d^i(\bar{\theta})$. This choice links the *ex-post* technology distribution $\Theta_d^i(\theta)$ to the *ex-ante* distribution.

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

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

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

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

[15] We model the adaptation cost as a function of expected profit. This could be micro-founded in a model of bargaining, in which researchers bargain collectively with the firm to split the profit. Since the firm knows less about more distant technologies, researchers can extract a larger share of rents for that adaptation.

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

The first condition ensures that the equilibrium production cost distributions, determined by C_d^i, are consistent with firms' technology choices. The second condition states that firms' choices are their best responses to the choices of all other firms. Unlike canonical trade models where similar conditions can be aggregated over firms and the equilibrium can be expressed solely in terms of aggregate prices and quantities,[16] in our setting, each firm's choice depends intricately on the entire distribution of other firms' choices.

We develop a procedure for characterizing conditions for existence and uniqueness, which can be adapted to other settings where interactions between firms' decisions cannot be summarized by aggregate outcomes. This procedure involves three key steps.

First, formulate equilibrium conditions in terms of policy functions (i.e., $\{g_d^i\}$), as in Definition 1, instead of distributions over technology choices (i.e., Θ_d^i). Constructing a fixed point in the space of policy functions circumvents the technical difficulties of applying fixed point theorems on spaces of probability measures. It also allows us to focus on equilibria where technology choices vary smoothly with endowment technologies.[17]

Second, stack the model outcomes to construct a joint mapping from the space of these outcomes (i.e., the space of $(\{g_d^i\}, \{C_d^i\})$) to itself. This differs from alternatives such as constructing a nested fixed point problem—e.g., mapping from g_d^i to itself, nesting inside a mapping for C_d^i given the policy functions (see, e.g., Alvarez and Lucas, 2007 for an example of nested fixed point approach in the context of the Eaton and Kortum model)—since our approach avoids the often challenging comparative statics in the inner problem. Our method shares the same spirit as that in Allen et al. (2023), but here the equilibrium outcomes 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 mapping from the space of (g_d^i, C_d^i) to itself. We then determine the conditions under which the Jacobian matrix of these Fréchet derivatives have a row sum (of absolute values) below 1, which allows us to apply the contraction mapping theorem. The procedure for deriving and bounding the derivatives can be applied to other models as well.

Our procedures work for generic metric space $\mathbb{T}$. To simplify the algebra, we assume:

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

[16]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 aggregated analytically into price indexes.

[17]Formally, we concentrate on equilibria with policy functions that are bounded, continuously differentiable, and have Lipschitz-continuous first derivatives.

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

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

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, $\bar{t}$, and $\bar{\phi}$.

The assumption $\mathbb{T} = \mathbb{R}$ has two key aspects: dimensionality and geometry. Regarding dimensionality, all theoretical results in this and subsequent subsections generalize to $\mathbb{T} = \mathbb{R}^n$ for a finite n.[18] In such a setting, one can view any technology 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 address the concern on this assumption for the interpretation of calibration outcomes and quantitative results in Section 4.

Regarding geometry, the technology space need not be limited to the real line—it could instead be a circle (or a sphere, a ball in higher dimensions). Different shape assumptions matter for firm choices.[19] Given the lack of evidence supporting one versus another, however, we choose a linear representation as it allows us to use Gaussian distributions to obtain closed-form solutions in quantification.

Under Assumptions 1 and 2, we establish the following results:

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

1. *Assume $\{\bar{\Theta}^i_d\}$ 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. Moreover, in this equilibrium, the choice of firms from (d, i) with endowment technology $\bar{\theta}$ is unique and characterized by the following first-order condition:*

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

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

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

Proof. See Appendix A.2. □

The first part of the proposition state the sufficient condition in $\zeta\bar{t}$ and M^2. Here, M^2 sets a limit on the variance of potential supplier technologies, which determines how much firms'

[18] In part (2) of the $\mathbb{T} = \mathbb{R}^n$ case, the cost functions take a quadratic form, e.g, $t(\theta, \tilde{\theta}) = (\theta - \tilde{\theta})^{tr} \cdot \bar{t} \cdot (\theta - \tilde{\theta})$, where x^{tr} is the transpose of a vector x and $\bar{t}$ here is a positive definite matrix.

[19] For example, on a real line, when a firm moves left, it gets closer to firms on its left while moving farther from those on its right. By contrast, on a circle, firms moving in opposite directions eventually meet.

choices can improve compatibility.[20] The parameter $\bar{t}$ governs the impact of compatibility on production cost, while ζ, the trade elasticity, captures how this affects sales and profits. For an equilibrium with continuous policy functions to exist, firms' policies must not be overly sensitive to their endowment, which is ensured if $\zeta\bar{t}M^2 < 1$.

The trade-off firms face in choosing their technology in such an equilibrium is summarized in the first-order condition: being close to their endowment, weighted by $1-\omega^i$, versus being close to the suppliers' technology, weighted by ω^i and the expenditure share of suppliers $\frac{\gamma^{ij}}{1-\gamma^{iL}}\chi^j_{do}(g^i_d(\bar{\theta}),\tilde{\theta})$.

For uniqueness, we establish the conditions under which the mapping described earlier is a contraction. The sufficient condition for uniqueness requires $\bar{t} < \frac{1}{2M}$, where $2M$ enters as the upper bound of the distance between a firm's technology and the average technology (weighted by expenditure shares) of its suppliers. Intuitively, $\bar{t}\cdot 2M$ determines the upper bound for the benefit from technology compatibility. If the benefit is not too large, and if the cost of technology adaptation $\bar{\phi}$ is not too small, then firms' technology would not be too sensitive to other firms' choice, in which case the mapping is a contraction.

The intuition behind the bound for $\bar{t}$ means that even if M is arbitrarily large, as long as the expenditure shares on firms with distant technology is not too large, uniqueness can be established without $\bar{t}\to 0$. Indeed, we show that under our calibration, the equilibrium under normal distributions on ex-ante distributions exists and is unique. More generally, Proposition 2 require bounded-valued technological space because we allow for arbitrary endowment distributions (as long as the density exists). This assumption can be relaxed by restricting the endowment distribution to particular shapes (e.g., Normal distributions).

It should also be clear that the equilibrium is not generically unique. In particular, if it $\bar{\phi}\to 0$ or if $\bar{t}\to\infty$—both would violate the condition in Proposition 2, then the incentive from compatibility will be so strong that all firms choose the same location—and any such locations would be an equilibrium.

2.5 Aggregation and General Equilibrium

We now embed firms' technology choice described above into the general equilibrium. Market clearing requires that for firms in country-sector (o,j) choosing technology $\tilde{\theta}$, the expected *sales to downstream firms*, $M^j_o(\tilde{\theta})$, should satisfy

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

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

[20] Larger variance means more heterogeneous suppliers technologies, implying larger room for firm to improve supplier comparability. By the Popoviciu's inequality on variances, $\mathrm{var}(\theta) < M^2$ for any bounded distribution for θ over $[-M,M]$.

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

$$X_o^j(\theta) \equiv \sum_d \rho_d^j P_d Q_d \pi_{do}^j(\theta), \tag{17}$$

where final goods demand include household consumption and firm adaptation costs:

$$P_d Q_d = I_d + \sum_i \int K_d^i(g_d^i(\bar{\theta}); \bar{\theta}) \mathrm{d}\bar{\Theta}_d^i(\bar{\theta}), \tag{18}$$

with technology adaptation costs $K_d^i(\theta; \bar{\theta})$ defined in (11).

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

$$I_d = w_d L_d + \sum_i \int \Pi_d^i(g_d^i(\bar{\theta}); \bar{\theta}) \mathrm{d}\bar{\Theta}_d^i(\bar{\theta}), \tag{19}$$

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

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

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

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

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

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 the ex-ante distribution in each (d, i) is degenerate, with a unit mass at $\bar{\theta}_d^i$. We characterize the symmetric equilibrium in which firms that are ex-ante identical, i.e. from the same (d, i), make the same technology choice and examine how these choice respond to changes in trade costs.[21]

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

[21] This setup can be viewed as a limit case of the premises described in Proposition 2 with the density functions for the ex-ante distribution approaching a Dirac delta function.

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

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

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

2. *Between two technology choice equilibria with different trade costs, the change in the location parameters of the factory-gate price distributions defined in (3),* $\bar{\mathbf{C}} \equiv (\bar{C}_1^1, \bar{C}_1^2, ..., \bar{C}_N^S)$,[22] *depends on the changes in trade cost and the endogenous response in technology,* $\{\mathrm{d}\ln \tau_{do}^j\}$ *and* $\{\mathrm{d}\theta_d^i\}$*, as below*

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

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

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

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

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

Proof. See Appendix A.3. □

Equation (21) is a special case of equation (15), which shows that firms' technology is weighted average between their endowment and the average technology of their suppliers.

In the second part of the proposition, $D_{\tilde{\gamma}}$ captures the importance of intermediate inputs for production costs. Ω is a Leontief inverse that captures the impact of the production cost in any country on that of any other country, accounting for the propagation via input-output and trade linkages. Inside the bracket are two components: the first captures the effect of changing technology distance, whereas the second captures the effect of the change in expenditure-weighted average change in import trade costs. Intuitively, factory-gate prices

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

of firms in a country increase if the distance between the country's technology to that of others increases (an increase in $\Lambda \mathrm{d}\boldsymbol{\theta}$) or if the import cost increases (an increase in $\mathrm{d}\ln\tilde{\boldsymbol{\tau}}$).

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

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

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

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

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

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

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

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

   .

Proof. See Appendix A.4. □

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

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

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

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

1. *Firm ν moves closer to θ^j_o, namely $\|\theta^i_d(\nu) - \theta^j_o\|$ decreases*
2. *Firm ν is more likely to purchase from (o,j)*
3. $\Delta \log\left(\chi^{ij}_{do}(\nu)/\chi^{ii}_{dd}(\nu)\right) = -2\zeta\bar{t}\cdot\Delta\|\theta^i_d(\nu) - \theta^j_o\|$

Proof. See Appendix A.5. □

Parts 1 and 2 of the proposition indicate that as the technology endowment of these firms moves towards (o,j), their choice shift in the same direction. This, in turn, makes it more likely for them to source from (o,j). Part 3 shows the the elasticity of the odds ratio with respect to the change in the technology difference between the firm and (o,j) is the product of parameter t and the conventional trade elasticity ζ. Thus, we can pin down t by inspecting whether firms' proximity to the technology of a country is correlated with importing from that country, a result we will use later.

2.7 Special Cases: The Welfare Implications of Technology Choice

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

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

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

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

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

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

Proof. See Appendix A.6. □

The first part of the proposition characterizes the three channels through which $\Delta\theta^i$ affects social welfare. Without loss of generality, suppose $\bar{\theta}^i > \theta^i$, which means an increase θ^i reduces the adaptation of firms in i toward other sectors. This creates three effects. First, it saves adaptation cost, which are rebated to households for consumption (first term). Second, it causes a change in the price of goods j (second term), in which $\tilde{\gamma}^{ij}$ captures the GE effect of the distance to θ^j on the output price of sector i. Finally, when a firm ν shifts technologically towards their suppliers, their suppliers, which also uses the output of ν as input also benefit from lower input sourcing cost. This effect is stronger if the benefit of technology compatibility is large (large t), and is amplified the input-output linkages captured by $\frac{\tilde{\gamma}^{ij}}{\gamma^{ij}}$.

Contrast this with firms' first-order condition:

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

we can see that both the adaptation cost and sector-i price show up in firms' optimization problem, but in different forms. Whereas the compatibility with suppliers enters firms' problem with a coefficient of γ^{ij}, it enters the social welfare with a coefficient of $\tilde{\gamma}^{ij}$, reflecting the propagation via input-output linkages that individual firms do not consider. More importantly, in choosing technology, individual firms also do not internalize the fact that their choice affects their downstream firms.

For the above reasons, at the decentralization equilibrium, firms' technology choice generally do not maximize social welfare. Without additional restrictions, however, we cannot sign $\Delta \ln(U)$—as firms move in any direction, they move closer to some suppliers but away from other suppliers, so the aggregate effect on other firms is ambiguous. The second part of the proposition shows that when sectors have symmetric input-output structures, the positive externality always dominate and moving further away from the firm's endowment technology will improve the social welfare.[23]

We discuss the international spillovers of technology choice in the next proposition.

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

$$\frac{\Delta \ln U_2}{\Delta \theta_2} = \frac{2\exp(-\bar{\phi}(\theta_2-\bar{\theta}_2)^2)}{\eta-\exp(-\bar{\phi}(\theta_2-\bar{\theta}_2)^2)}\bar{\phi}(\bar{\theta}_2-\theta_2) + 2\bar{t}\frac{1-\gamma^L}{\gamma^L}\bar{\chi}_{12}(\theta_1-\theta_2) > 0,$$
$$\frac{\Delta \ln U_1}{\Delta \theta_2} = 2\bar{t}\frac{1-\gamma^L}{\gamma^L}\bar{\chi}_{12}(\theta_1-\theta_2) > 0.$$

Proof. See Appendix A.7. □

The result on $\ln(U_2)$ shows that the misalignment between private and social incentives in technology choice operating in a closed-economy described in Proposition 6 extends to an open economy. Intuitively, when firms in a country shift technology toward that of other countries, these firms benefit directly as the users of foreign inputs. Consumers benefit more because the reduction in the firms' production cost is amplified via the round-about production involving foreign firms. Thus, firms tend to under-invest in adapting to foreign technologies even from the home social planner perspective.

This mechanism has a global effect. By moving towards the technology position of country 1, firms in country 2 generate a positive externality as now firms in country 1 can source inputs more cheaply, a cost saving that also gets amplified by the production network.

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

This last result might appear surprising given the intuition, rooted in the envelop theorem, that endogenous response to negative shocks often mitigates the damage of the shock. In our model, however, because of the externalities discussed above, the equilibrium is not efficient. This inefficiency is exacerbated by the endogenous response, which is why endogenous technological changes amplify, rather than mitigate, the cost of trade conflicts.

3 Evidence

We now turn to the data, with two goals. First, the key premise of the model is that technology compatibility with suppliers increases the probability of sourcing from that supplier.

[23] When a firm's equilibrium choice is in the left or right of all other sector, there is only positive externality, in which case firms also under-invest in technological adaptation.

We provide evidence for this premise using firm-level Chinese customs data. Second, an implication of our model is that as tariff decreases, the technologies of importer and exporter become more compatible with each other. We test this mechanism using over-time changes in tariffs that affect MFN-bound exporting countries but not other countries. We will leverage these two estimates to discipline the model for quantitative exercises.

3.1 Measuring Technology Similarity Using Patent Texts

For both exercises, we need to measure the compatibility between pairs of technologies. In the model, technologies are compatible if they are close to each other. We proxy for compatibility by measuring the 'closeness' between technologies as reflected in patent texts.

Patent texts are a useful source of information for several reasons. Patents are a natural embodiment of new technologies. Patent texts are supposed to accurately summarize the features of a technology. Last but not least, patent data are available widely, enabling measurements at the global level.

Measure Construction. Our source of patent data is PATSTAT Global (2023 Fall version), an administrative database published by the European Patent Office (EPO) covering close to the universe of world patents. For consistent language processing, we focus on patents with an English language abstract. To avoid double counting patents in different offices (e.g. EPO, USPTO, etc.) corresponding to the same invention, we keep only one patent within each patent family. The final sample is approximately 45 million patent abstracts.[24]

We process patent abstracts using a leading large language model developed by Alibaba ('gte-base-en-v1.5'). The model takes as the input paragraphs of text and returns as the output a vector of 768 numbers—a processed called 'text embedding.'[25] The vector, representing a point in a space of 768 dimensions, summarizes the meaning of the input. Paragraphs with similar meanings are represented by points close to each other. In the context of patents, we can view the 768 dimensional space as representing all possible ways of combining technical ingredients/scientific discipline to achieve a function. Each patent, represented by one coordinate in the space, embodies one particular way.

We construct the proximity between firms and country in the space of technology using

[24] The EPO source abstracts from both national/regional patent offices and the Patent Cooperation Treaty, prioritizing abstracts in English whenever possible. As a result, approximately 90% of all abstracts in PATSTAT is in English. Our final sample keeps an invention as long as one of the patents in family is in English, so the coverage among *patent families* is even higher. We use abstracts, not full texts, because full texts are available in English in far fewer countries.

[25] See `https://huggingface.co/Alibaba-NLP/gte-base-en-v1.5` for descriptions of the model. The model strikes a good balance of efficiency and performance. Deployed on a desktop GeForce RTX 4090 GPU, it produces the embeddings of all our patents within two days. In terms of performance, as of August 2024, the model is ranked 36 out of 438 language models by the Massive Text Embedding Benchmark (MTEB) Leaderboard (Muennighoff et al., 2022). For reference, the S-BERT model ('all-mpnet-base-v2') is ranked 282; the latest text embedding model offered by OpenAI ('text-embedding-3-large-256'), is ranked 65.

the cosine similarity between the average coordinates of patents by a firm/country. For example, let $\bar{c}_{\omega t}$ be the average coordinates of all of the patents invented in firm ω in period t, and let $\bar{c}_{ot}$ be the average coordinates of patents in country o. Then the similarity between firm ω and country in period t is defined as:

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

where $||||$ denotes the norm of a vector.[26, 27] We construct similarity measures at other level aggregation, such as between pairs of countries or sectors, analogously. By construction, these measures are bounded between -1 and 1.

Validation and Summary Statistics. In Appendix B.2, we validate the text-based similarity measures by correlating them with other factors likely associated with technology similarity. Specifically, we show that patents within the same class are more similar than those from different classes. Likewise, sectors with stronger input-output linkages exhibit greater similarity than those with weaker connections. Lastly, countries in close geographical proximity tend to have more similar patents. These patterns indicate that patent texts, when processed by language models, provide useful insights into technology proximity.

Our empirical analyses use three text-based measures. In Section 3.2, we exploit firm-level variations to show that technology proximity and importing decisions are correlated. For that, we measure technology similarity between Chinese firms and foreign regions ($\text{Similarity}_{\omega ot}$) and between Chinese firms and pairs of foreign region-sector ($\text{Similarity}_{\omega ojt}$). In Section 3.3, we use import-level HS-6 product MFN tariffs, interacted with an exporters' MFN status, to examine the causal effect of trade barriers on technology proximity. Tailoring the measure to the level of external variation, we measure technology similarity between importing country d and exporting country o and product $HS6$ ($\text{Similarity}_{do,HS6,t}$).

For both analyses, our data span the periods from 2000 to 2014, which we group to 5 three-year periods. We aggregate countries in the World Input-Output Tables (WIOT) to 29 regions, including one for the rest of the world, based on their geographic and political proximity.[28] We outline the key steps involved in constructing the datasets for each exercise below, with more detailed information provided in Appendix B.1. Table 1 provide summary statistics for the key measures used in our empirical analysis.

Citation as Alternative Measure. Researchers often use patent citations as an alternative measure of technology similarity. Although citations are a natural way to measure technol-

[26]We assign patents to industries and (later) HS6 products using the crosswalks between the international patent classification (IPC) and standard industry and product classification by Lybbert and Zolas (2014).

[27]An alternative to calculating similarity between average coordinates is to calculate similarity between individual patents and then aggregate patent-level similarity to the firm and country-sector levels. Given the large number of patents, this alternative is computationally expensive. Using a small sample of patents, we verify that our similarity metric is strongly correlated with a metric aggregated from patent-level similarity.

[28]See Table B.1 for the list of all regions.

Table 1: Summary Statistics: Similarity between Firms

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

Note: Similarity$_{\omega ot}$ is the cosine similarity between a Chinese firm ω and a foreign region o in period t. Similarity$_{\omega ot}$ is the cosine similarity between a Chinese firm ω and a region-industry pair (o, j) in period t. Similarity$_{do,HS6,t}$ is the cosine similarity between patents in country d and patents in a particular HS-6 category in foreign region o and period t.

ogy proximity, they are also frequently associated with knowledge diffusion between the cited and citing patents (Liu and Ma, 2021; Jaffe et al., 1993). For this reason, we have chosen to use a 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 citation as the measure of 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 and sector level data. The vertical axis is the log exports from sector j of country o to sector i of country d; the horizontal axis is the similarity of technologies 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 $d - i$, by $o - j$, and by $i - j$.

The figure reveals that buyers source more from sellers with technologies that are similar to their own. This pattern cannot be rationalized by the Ricardian trade theory with exogenous technologies, which predicts that countries tend to trade more with partners that have dissimilar technologies, but is consistent with a technology similarities are shaped by trade.

In the rest of this section, we first look into such correlation at the firm-level, which allows us to better isolate the source of the correlation; we then examine how technological proximity respond to exogenous change in bilateral trade costs.

3.2 Firm-Level Correlation

Proposition 5 suggests that within a country-sector (d, i), firms with an endowment draw closer to a foreign country o are more likely to choose a similar technology to firms in o and import from o, resulting in a *firm-level* correlation between imports and technology proximity.[29] We use Chinese data to test this relationship.

Data. Our sample includes manufacturing firms in the Annual Survey of Industrial En-

[29]Our model is static, so firms' endowment draws 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.

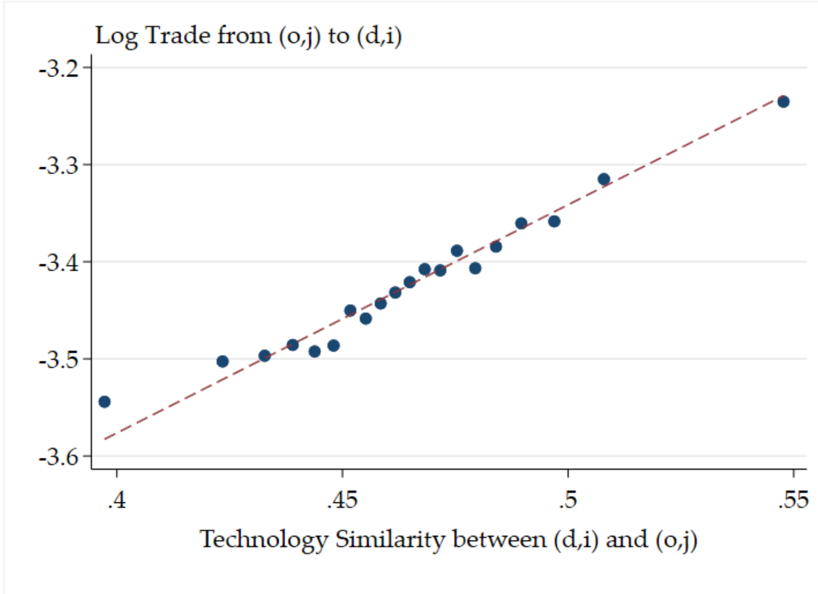


Figure 1: Technology Proximity and Trade

Note: the figure is a binned scatter plot of log exports from country o sector j to country d sector i against the technology similarity between (d, i) and (o, j). Technology proximity is measured as the cosine similarity of textual embedding of patent abstracts in (d, i) and (o, j). Trade data is from the World Input-Output Database (WIOD). The regression controls for $d - i$, $o - j$, and $i - j$ fixed effects, as well as distance metrics (geographic, language, colonial ties) and income differences between o and d. The regression coefficient is 1.9 (s.e. 0.2 clustered by country pair).

terprise maintained by the National Bureau of Statistics of China (NBSC). The dataset offers detailed accounting information for all state-owned firms and all non-state firms with sales greater than 5 million Chinese Yuan (≈ US\$600,000).

We link these firms to the PATSTAT Global database to retrieve the families of patents owned by these firms. The abstracts of these patents are then used to construct technology similarity measures, $\text{Similarity}_{\omega o t}$ and $\text{Similarity}_{\omega o j t}$. Our final dataset includes all manufacturing firms with patents. 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. We obtain importing information of these firms from China's General Administration of Customs, which provides detailed records on the universe of all Chinese trade transactions for the years 2000-2014. We aggregate the import records to construct a panel at the $\omega - o - t$ level, with each observation indicating whether a firm ω in period t imports from o.

Specification. Our regression specification is as follows:

$$\mathbb{I}[\text{Import}_{\omega o t} > 0] = \beta \, \text{Similarity}_{\omega o t} + \beta_2 X_{\omega o t} + FE_{\omega t} + FE_{\omega o} + FE_{o t} + \epsilon_{\omega o t}. \tag{24}$$

The outcome $\mathbb{I}[\text{Import}_{\omega o t} > 0]$ is the indicator for whether firm ω imports from o in period t. In addition to $\text{Similarity}_{\omega o t}$, the regression includes several control variables. $FE_{\omega t}$ is the firm-time fixed effects, which absorb the factors that affect the overall tendency of a firm to import or use foreign technologies, such as productivity, openness to foreign ideas, etc. $FE_{\omega o}$

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

	$\mathbb{I}[\text{Import}_{\omega ot} > 0]$			
	(1)	(2)	(3)	(4)
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

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered by firm. $X_{i(\omega)ot}$ include both the output applied tariffs of industry i and the input tariffs, defined by the average applied tariffs of importing the goods in industry j from world region o, weighted by the input share of industry j in industry i production. Industries are at the three digit level.

is firm-region fixed effect, which control for time-invariant connection between a firm and a region. It also ensures that any finding we have is not due to technology similarity measure picking up the styles in which firms write their patent applications. FE_{ot} is the region-time fixed effects, which account for factors that shape the average influence/quality of patents in region o period t. $X_{\omega ot}$ controls for other determinants of import and technology choices.

Main Results. Table 2 reports the results. Column 1 shows that firms with similar technologies to foreign region o are more likely to import from o. One may be concerned that this correlation is driven by the input and output tariffs the firm face, which not only affect imports but also can be correlated with technology similarities due to unobserved (time-varying) connections between Chinese firms and foreign regions. We control for input and output tariffs in the (3-digit) industry of the firm in Column 2. The estimate is the same.

Another source of concern is that firms from different provinces/industries in China face different barriers to interact with foreign suppliers due to either their geographic locations or historic connections with these countries. The firm-origin fixed effects absorb the time-invariant part of such connections but not the time-varying part. 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. Under this specification, Similarity$_{\omega ot}$ has a coefficient of 0.021, meaning that a one-standard-deviation increase in similarity with o is associated a 0.19 percentage point increase in the probability of importing from o, or 4% of the mean value of the outcome variable.

Given the importance of multinational firms in global trade, one may wonder whether this correlation is driven by affiliates of foreign firms, which may both import from their parents and use the technology of their parents. 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 essentially the same. Since our model applies to foreign affiliates as well as domestic firms, in the rest of this section, we use the full sample of firms. The results are robust if we exclude foreign firms and joint ventures from the sample.

Heterogeneity by Industry. The correlation documented in Table 2 is consistent with the key premise of our model: firms make interdependent technology and supplier choices in order to use compatible inputs. An alternative explanation for this correlation is that firms may learn about foreign suppliers and technologies through a correlated source. For example, the firm owner may meet a contact from a region o, who introduce the technologies as well as the suppliers of region o to the firm. Although we control for a rich set of fixed effects, such possibilities cannot be ruled out. To support our interpretation of the mechanism, we provide auxiliary evidence by looking into two implications of the model on how the correlation varies across industries.

First, if the correlation between technology similarity and importing decision arises because firms value supplier with compatible inputs more, we should expect that the correlation to be higher for industry pairs with more intensive input-output linkages. To test this, we use the technology similarity measure between firm ω and country-sector $d-j$ as the key explanatory variable. We construct an indicator for whether a firm ω imports in industry j from country o, denoted by $\mathbb{I}[\text{Import}_{\omega ojt} > 0]$, as the outcome variable.

Column 1 of Table 3 reports a simple specification that correlates similarity and importing decision. Given that the data have a j dimension, we can control the interactions between the full sets of fixed effects in equation (24) with j. The point 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 share of industry j's output in production of the industry of firm ω. The interaction is statistically significant and sizable: if industry j's share in industry $i(\omega)$ input increases by 10 percentage points, then the correlation between similarity and importing increases by 0.0053, or a quarter of the average correlation.

Second, the importance of input compatibility can vary across firms. For example, firms in the telecommunication and electronics industries may have much higher compatibility requirement than firms in food manufacturing. If the estimated coefficient is indeed due to technology compatibility, we should expect the correlation to be higher for industries in which compatibility is more important.

To test this hypothesis, we construct a novel measure of compatibility intensity based on the full text of U.S. patents from the Orbis Intellectual Property Database, which link full list of patents to firms worldwide. For each three-digit industry, we count the number of patenting firms, and the share among these firms owning patents contain the word 'compatibility/compatible' or 'inter-operability' *and* one of the words that put 'compatibility' in

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

	$\mathbb{I}[\text{Import}_{\omega o j t} > 0]$			
	(1)	(2)	(3)	(4)
$\text{Similarity}_{\omega o j t}$	0.020***	0.017***	0.018***	0.015***
	(0.004)	(0.004)	(0.004)	(0.004)
$\text{Similarity}_{\omega o j t} \times \text{Input-Output } \gamma_{i(\omega)j}$		0.053*		0.056*
		(0.029)		(0.029)
$\text{Similarity}_{\omega o j t} \times \text{Compatibility}_{i(\omega)}$			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

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered by firm. Industries are Input-Output Industries equivalent to 3-digit CIC. $X_{i(\omega)ot}$ include both the output applied tariffs of industry i and the input tariffs, defined by the average applied tariffs of importing the goods in industry j from world region o, weighted by the input share of industry j in industry i production.

appropriate contexts.[30] Viewing patents with these key words as technologies for which compatibility is important, the share of patenting firms in an industry with such patents is the likelihood that for a random firm in this industry, technology compatibility is important. We denote this industry-specific measure of compatibility intensity as $\text{Compatibility}_{i(\omega)}$, see Table B.3 for this measure for all manufacturing industries.

In Column 3 of Table 3, we add the interaction between compatibility intensity of the industry and the similarity measure. The coefficient for the interaction term, 0.083, implies that increasing industry compatibility intensity by one standard deviation (0.11) is associated with a 0.9 percentage point increase in the correlation, or 8.9% of the average correlation. In Column 4, we add both interaction terms jointly. The coefficients are similar. Together, the results based on input and compatibility intensities of industries provide support that the correlation between importing and technology similarity is driven by compatibility along the input-output network.

Extended Gravity. Our model features a novel 'extended-gravity' like prediction (Morales et al., 2019; Alfaro-Urena et al., 2023) on the relationship between technology similarity and importing decisions. Suppose regions o' and o are close to each other in the space of technologies. Then technology similarity between a firm ω and o' could be positively correlated

[30]Examples of such words include interface, system, protocol, plug-and-play, standard. See Appendix B.3 for the full list, generated by ChatGPT. The presence of these additional words help exclude instances where 'compatibility' refers to things other than technology, e.g., 'water compatibility', 'skin compatibility.' Results are robust if we do not require additional key words.

with the firm importing from o. This can be the case even if we control for the measure of technology similarity between ω and o unless the measure is perfect.

We show in Appendix B.4.2 that indeed, a firm ω's being technologically similar to the technology neighbors of a region o, defined as countries with similar technologies to o, is strongly correlated with ω importing from o. Of course, this correlation may be 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'. While such knowledgeable contacts do exist, 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 linguistically and geographically close to each other. We tease out the influence of such experts by controlling for the technology similarity between ω and the linguistic and geographic neighbors of o. The coefficients for these controls are small and statistically insignificant. More importantly, they do not diminish the coefficient for technology similarity between ω and technology neighbors of o. These results offer additional support to our technology-based interpretation of the correlation.

3.3 Tariff Shocks and Technology Similarity

The second part of our empirical analysis tests another key implication of our model: a decrease in bilateral tariff between o and d leads to an increase in the similarity between o and d (relative to other countries).

Data. We assemble a dataset on bilateral technology similarity and tariffs across different countries and industries. The dataset is a balanced panel that spans 28 geo-political regions (d and o) over 5 three-year periods (t) from 2000 to 2014.

The observations are at the d-o-$HS6$-t level. We source from the UN TRAINS Database the tariff faced by the region-product $(o, HS6)$ in destination market d in period t. The data include both the applied tariff rates and the most-favored-nation (MFN) tariff rates. As described previously, we also construct technology similarity between country d and the producers of an $HS6$ product in d in period t, denoted $\text{Similarity}_{do,HS6,t}$.

Specification. We estimate the following regression specification:

$$\text{Similarity}_{do,HS6,t} = \beta \ln \tau_{do,HS6,t} + FE_{d,HS6,t} + FE_{o,HS6,t} + FE_{do,HS6} + \epsilon_{do,HS6,t}, \tag{25}$$

where $\tau_{do,HS6,t}$ is the applied ad-valorem tariffs. To address potential endogeneity and measurement error in applied tariffs, following Boehm et al. (2023), we instrument $\tau_{do,HS6,t}$ with the interaction between $\tau^{MFN}_{do,HS6,t}$, the tariff rates under the Most Favored Nation (MFN) principle that applies to all its WTO trade partners (that are not covered by other agreements such as a free trade agreement), and an indicator for whether the MFN tariffs of d applies to o in period t. For each importer and product, we will exclude the biggest exporter to alleviate the concern that countries choose MFN tariffs with the major exporters in mind.

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

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. 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 $\ln \tau^{Applied}_{do,HS6,t}$.

As in the long-run specification in Boehm et al. (2023), we control for a wide array of fixed effects. In particular, the inclusion of the exporter-HS6-time fixed effects ($FE_{o,HS6,t}$) accounts for the average similarity of patents invented by firms in o, $HS6$ with other patents. Similarly, the importer-HS6-time fixed effects $FE_{do,HS6}$ accounts for the average similarity of patents invented by firms in d with patents invented by all world-wide firms producing the $HS6$ product. Thus, any variations in similarity due to writing style is absorbed. Importantly, we control for $d - o - HS6$ fixed effects, so the identification comes from the changes in tariffs.

Results. Table 4 reports the results. Columns 1 and 2 replicate the results in Boehm et al. (2023), with the log of HS6 level trade flow as the outcome variable. Column 1 is the reduced-form specification; Column 2 is the two-stage least squares. These estimates are around the same order of magnitude as Boehm et al. (2023).

Columns 3 and 4 replace the outcome variables in the first two columns with technology similarity. Focusing on the IV estimate, we find that a doubling of importing tariff leads to a decrease in similarity of 0.009. This is approximately 8% of the standard deviation of technology similarity. Since tariffs typically account for a small share in all frictions impede trade, the estimate reflect a potentially large impact of trade barriers on technology choice.

In Appendix B.5, we show that using citation as a measure of bilateral similarity lead to similar findigs. We also 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 findings here, and the appendix results using patent citations as the measure, are related to a large body of literature that uses citation flows to measure the extent of knowledge spillover pioneered by Jaffe et al. (1993). Most closely related, Aghion et al. (2021) shows that after a French firm begins exporting to a destination country, its patents receive more citations from that country. We differ from Aghion et al. (2021) in two aspects. First, we use exogenous MFN tariffs to show the causal effect of trade costs on patent citations. Based on the industry heterogeneous effects and extended gravity result, we provide suggestive evidence that technology compatibility is at least part of the story. Second, our model interprets the findings through the lens of the compatibility of imported goods with the technology chosen by domestic firms, complementing existing works focusing on other channels of technological diffusion.

4 Quantification

We use the model to quantify the impacts of trade on technology choice and, through that channel, on welfare. We begin by outlining the functional form assumptions and approximations used in this section, which enable us to derive closed-form solutions for firms' technology choices. We then explain how we parameterize the model, and finally, we conduct counterfactual analyses.

4.1 Functional-Form Assumptions and Numerical Algorithm

In our model, heterogeneous firms interact with each other directly (see e.g., equation (3)), rather than through aggregate prices as in canonical quantitative trade models. Consequently, solving the model requires tracking the entire distribution of Θ_d^i across all (d, i), not just their first moments. To maintain tractability, we impose the following structure on the technology endowment distributions:

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, $\bar{\Theta}_d^i \sim \mathcal{N}(\bar{\mu}_d^i, (\bar{\sigma}^i)^2)$.*

Given this assumption and by applying quadratic approximations to $\ln C_o^j(\theta)$—the log of production cost parameter in (o, j) for a firm choosing technology θ—around the mean ex-post technology of (o, j) for each (o, j), we derive closed-form solutions for firms' technology choice problems and the ex-post technology distributions, summarized below:

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

1. *$\ln C_o^j(\theta)$ is a quadratic function of θ characterized by:*

$$\ln C_o^j(\theta) = k_{A,o}^j + m_A^j(\theta - n_{A,o}^j)^2; \tag{26}$$

where $\{m_A^j, n_{A,o}^j, k_{A,o}^j\}$ depend only on model primitives and wages $\{w_o\}$.

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

$$g_o^j(\bar{\theta}) \equiv \alpha_o^j + \beta^j \bar{\theta}, \tag{27}$$

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

3. *The ex-post technology distribution of* (o, j) *is*

$$\Theta_o^j \sim \mathcal{N}(\mu_o^j, (\sigma^j)^2), \text{ with } \mu_o^j = \alpha_o^j + \beta^j \bar{\mu}_o^j \text{ and } \sigma^j = \beta^j \bar{\sigma}^j. \tag{28}$$

Proof. See Appendix C.1. □

Proposition 8 provides closed-form expressions for the cost function and firms' policy function.[31] In particular, equation (27) shows that firms' technology choices are linear functions of their endowment technologies. This has two attractive implications for quantitative implementation. First, with the analytical expressions, we no longer need to solve problem (13) numerically. Second, under the normality assumption for endowment technologies, this linear policy function implies that ex-post technology distributions are also normal, with their parameters depending on equilibrium outcomes and endowment distributions as described in equation (28). This result will prove useful for model calibration.

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

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

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

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

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

4.2 Calibration Procedure

This subsection explains how the parameters are calibrated.

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

Parameters Externally Calibrated. We set the elasticity of substitution $\eta = 5$, which implies a cost elasticity of 4 (Simonovska and Waugh, 2014) and 20% markup for sales to final-good producers. We set the shape parameter of match-specific productivity $\zeta = 4$. This value means that the conditional on their technology, the trade elasticity between any pairs of intermediate-good producers is 4, as shown in equation (6).

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

Technology Distribution. We recover the ex-ante distribution, $\{\bar{\mu}^i_d, \bar{\sigma}^i\}$, in two steps. In the first step, we choose the parameters governing the ex-post distributions, $\{\mu^i_d, \sigma^i\}$, to match the patent abstract similarity between and within country-sectors. This can be implemented 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—without the knowledge of $\bar{\phi}$, the parameter that governs the cost of technology adaptation, among other parameters.

In the second step, we calibrate $\bar{\phi}$ and recover the ex-ante distributions using equation (28). Note that $\bar{\phi}$ only affects how firms' markups are divided between profits and technology adaptation costs, leaving equilibrium wages and prices unchanged once the ex-post distributions and other general equilibrium outcomes (such as wages and prices) have been calibrated. This allows us to adjust $\bar{\phi}$ without altering earlier steps in the calibration. Additional details are available in Appendix C.4.

To implement the first step, we map the technological distances in our model into a similarity metric valued between 0 and 1 to facilitate comparison with the empirical similarity measure. We define the similarity between two technologies as:

$$sim(\theta, \tilde{\theta}) = \exp(-(\theta - \tilde{\theta})^2), \tag{29}$$

where θ and $\tilde{\theta}$ represent the technologies of two firms. An analogous measure is constructed for technology across country-sector pairs:

$$sim(\mu^i_d, \mu^j_o) = \exp(-(\mu^i_d - \mu^j_o)^2).$$

We choose $\{\mu^i_d\}$ to fit the similarity of mean technologies between country-sectors, based on patent abstract embeddings. Specifically, we solve the following optimization problem:

$$\min_{\{\mu_d^i\}} \sum_{d,i,o,j} \left[\ln sim(\mu_d^i, \mu_o^j) - \ln sim_{do}^{ij,Data} \right]^2, \tag{30}$$

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 (the last year group of our empirical exercise) and residualized by $i-j$ fixed effects. With this residualization, our calibration does not attempt to explain why the auto industry is technologically closer to the auto parts industry than to agriculture. Instead, it focuses on accounting for variations between pairs of countries *within* a sector pair.

Two rationales motivate this choice and justify the one dimensional technology space. First, with the one-dimensional representation of the technology space, we are bound to miss important variations in technological distances between sectors. Since our counterfactuals focus on how shocks to international trade costs will shift the technologies of countries, our calibration concentrates on what matters most for these scenarios. Second, and more importantly, under the Cobb-Douglas production function for intermediate producers, adding exogenous average sectoral technological distances (i,j) to each producer in each country has no impact on firms' choices. Thus, our calibration and counterfactuals have the interpretation of holding sectoral distances constant while focusing on country differences in technologies.

Problems of the form (30) is known to be challenging. By fitting the logarithms of the similarities, however, the problem can be transformed into a classical multidimensional scaling problem, which has unique solutions and can be efficiently solved (see Appendix C.2).[32]

To infer the standard deviation of technology for each sector i, σ^i, we match the standard deviation of similarity between two technologies randomly drawn from the same (d,i). For each country-sector, we randomly 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 statistics across countries to infer $\{\sigma^i\}$.

Trade Costs and Distribution of Production Techniques. The remaining model parameters are calibrated jointly in equilibrium to match the data across a set of moments. The number of parameters equals the number of moments, so these parameters are just identified. We calibrate the parameters $\{\Xi^i/A_d^i\}$, which determines the productivity of (d,i), to match the output share of (d,i) in industry i, and calibrate $\{\tau_{do}^j, \tau_{do}^{Uj}\}$ to match the trade shares of intermediate and final goods, respectively.

Technology Compatibility $\bar{t}$. In our model, the parameter $\bar{t}$ governs firms' input-sourcing decisions given their chosen technology locations. We calibrate $\bar{t}$ by matching the extensive-

[32]Although cosine similarity theoretically ranges from $[-1,1]$, the entries of the actual similarity matrix are all positive, making the logarithm transformation possible.

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			
$\tau^j_{do}, \tau^{Uj}_{do}$	Iceberg trade costs		Bilateral trade shares
$\bar{\phi}$	Adaptation cost	0.29	Country-sector-level similarity-tariff elas.: -0.007
$\bar{t}$	Compatibility incentive	6.1	Firm-level import-similarity corr: 0.021
C. Best Fit			
$\bar{\mu}^i_d, \bar{\sigma}^i$	Ex-ante Technology Distribution	-	Tech. similarity between and within country-sectors

margin import-similarity correlation (as shown in Column 3 of Table 2) using the simulated method of moments. Given the ex-post technology distribution $\{\Theta^i_d\}$ and equilibrium prices, a firm with technology θ in (d, i) will source inputs from (o, j) with probability $\chi^{ij}_{do}(\theta)$, according to equation (6). The similarity between the firm's technology and that of (o, j) is $sim(\theta, \mu^j_o)$, as defined in (29).

To perform the calibration, we simulate 190,000 Chinese firms ($d = CHN$), drawing 10,000 firms from each sector i based on the ex-post distribution Θ^i_d. We then regress the realized extensive margin of importing from each country o on the technology similarity with o, averaged across sectors j, while controlling for firm and $o - i$ fixed effects. The parameter $\bar{t}$ is set at 6.1 to ensure the model regression coefficient matches the empirical estimate of 0.021 (Column 3, Table 2).

Adaptation Costs $\bar{\phi}$. We calibrate the adaptation cost parameter $\bar{\phi}$ to match the semi-elasticity of similarity to tariffs, as estimated in Column 3 of Table 4. According to Proposition 4, conditional on trade shares and γ^{ij}, this semi-elasticity identifies $\bar{\phi}$. To calculate the model-based semi-elasticity, we replicate the variation in MFN tariffs observed in the data, scaling the model's trade cost $\{\tau^j_{do}\}$ by the standard deviation of the MFN tariff for each (d, j) over time. This produces a counterfactual equilibrium. 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^j_{do}, controlling for $d - o - j$, $o - j - t$, and $d - t$ fixed effects, mirroring the empirical specification.[33] We adjust $\bar{\phi}$ until the model's semi-elasticity matches the empirical coefficient of -0.007, yielding a calibrated value of $\bar{\phi} = 0.29$.

Numerical Implementation. The calibration proceeds in two steps. First, we estimate the ex-post distribution parameters $\{\mu^i_d, \sigma^i\}$ to best match the empirical technology similarity within and between country-sectors, as previously described. Given these parameters, we

[33]Given the MFN variation at the importer-sector (d, j) level, we control for $d - t$ fixed effects in the model instead of the $d - j - t$ fixed effects.

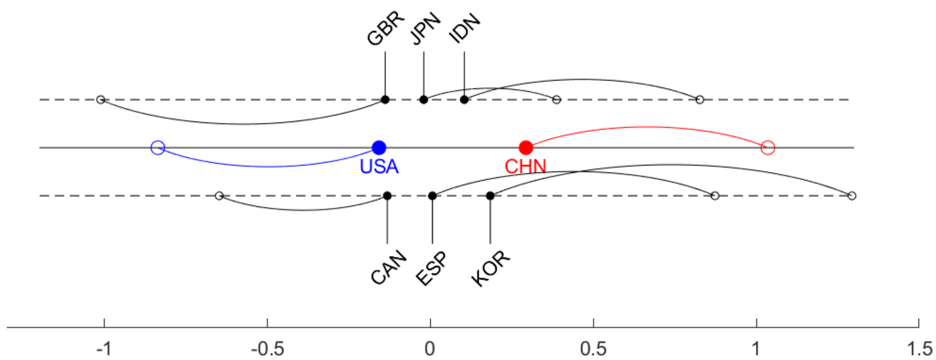


Figure 2: Estimated Technology Distributions for Selected Economies

Note: Dots are the means of technologies distributions at firms' optimal technology choice, μ^i_d, averaged across sectors for each d. Circles are the means of their endowment technology.

employ a nested algorithm for the second step. In the outer layer, we adjust $\bar{t}$ to match the firm-level extensive import- similarity regression coefficient. In the inner loop, we solve for the competitive equilibrium, while simultaneously calibrating $\{\tau^j_{do}, \tau^{Uj}_{do}, \Xi^i / A^i_d\}$ to match trade shares. The model is then simulated, and the regression coefficient is fed back into the outer loop. Once this process is complete, we calibrate $\bar{\phi}$ to match the country-sector similarity-tariff semi-elasticity and use the results from Proposition 8 to infer the ex-ante distribution parameters $\{\bar{\mu}^i_d, \bar{\sigma}^i\}$.

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) locations of ex-ante and ex-post technology distributions for a subset of countries. The dots depict the locations of ex-post technology. Among the major economies, the technologies of the U.S. and China fall at the two ends of the spectrum, whereas the technologies of other major economies such as Japan, Canada, and Western European countries fall in the middle. This pattern highlights a key feature of the bilateral technology similarity statistics, which are derived from the semantic analysis of patent abstracts: within any sector pair, U.S. and Chinese technologies are 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 noteworthy patterns emerge. First, ex-ante technologies tend to cluster spatially, with Asian and developing countries grouped on the right, while Western and developed economies are clustered on the left. This clustering may reflect the influence

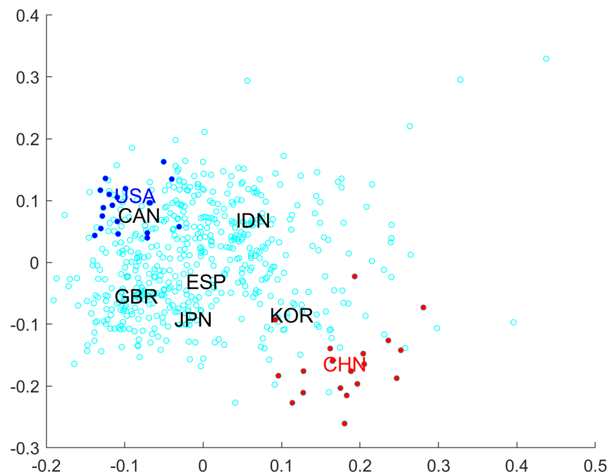


Figure 3: First Two Principal Coordinates of the Similarity Matrix

Note: Plotted are the first two principal coordinates of the similarity matrix between country sectors. The similarity is the cosine similarity of the mean embeddings of patent abstracts between two country-sectors (d, i) and (o, j), netting out of the sector pair (i.e., $i - j$) fixed effects. Each dot corresponds to the principal coordinate of a country-sector. Country name indicates the average coordinate of the country across sectors. Blue dots correspond to US country-sectors and red dots correspond to Chinese country-sectors.

of geographic or cultural factors on initial technological endowments.[34] Second, technology compatibility incentives bring countries' technologies closer. International trade plays an important role here. Indeed, if countries do not engage in trade, then firms will move towards the average technologies of their own countries.

To demonstrate the calibrated technology distance is 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 two-dimensional space, which correspond to the calibration results in the two-dimensional space. As the figure shows, the divergence between U.S. and Chinese technologies, with other countries positioned in between, remains robust when viewed in this two-dimensional space.

Assessing Fit of Technology Similarity. Our structural model allows for technology to be represented in a high-dimensional space. For simplicity and transparency, however, we have restricted this representation to a one-dimensional technology space. Figure 3 provides a visual confirmation that the calibration captures the salient patterns that are evidence in the two-dimensional case. To evaluate more systematically how well this simplification captures the cross-sectional technology similarities between country-sectors, we regress the log of data similarity between country-sectors on the log of model-predicted similarity in

[34]Some of these differences could stem from the spatial diffusion of ideas, as our model's ex-ante technology distribution captures various country-specific factors, including the spread of ideas.

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

	Log Similarity in Data	
Log Similarity in Model	(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	di, oj	di, oj

Note: 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 distribution $\{\mu_o^j, \sigma^j\}$. Column (2) restricts the technology distribution to the ex-ante distribution $\{\bar{\mu}_o^j, \bar{\sigma}^j\}$. Both columns control for $d-i$ and $o-j$ fixed effects. The reported within R^2 exclude the variation explained by the fixed effects.

Column (1) of Table 6, controlling for unilateral country-sector fixed effects. The R-squared statistic shows that approximately 21% of the variation in bilateral similarities is explained by the model—the same order of magnitude as the explanatory power of bilateral distance for bilateral trade flows (25%-50%, depending on sample). Notably, fully describing the bilateral similarity matrix between country-sectors would typically require $(NS)^2 - NS$ parameters. In contrast, our model achieves this with $NS - 1$ parameters, $\{\mu_d^i\}$, reducing the dimensionality by two orders of magnitude. If the data similarity showed no systematic structure, it is unlikely we observe this level of fit.

Relevance in Shaping Technology. To further unpack the model's performance and highlight the role of compatibility incentives in shaping technology interdependence through international trade, Column (2) of Table 6 regresses the log of data similarity on model-predicted similarity, this time at the level of the inferred endowment technologies. Here, the regression coefficient declines significantly, with the economic significance nearly vanishing. The R^2 also falls, from 21% to 13%. This comparison suggests that 38% of the model's explanatory power (8% out of 21%) can be attributed to the 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, costs arising from technology incompatibility account for 9.5% of foreign imports. The literature has found that to account for the low level of international trade, the iceberg trade costs must be substantially higher than the level of tariffs (Anderson and Van Wincoop, 2004). Our model suggests that a small, through non-negligible part of these unobserved barriers to trade could be technological incompatibility.

Table 7: Welfare Costs and Technology Decoupling of a Semi-Conductor Embargo

	Import share in targeted China's sector (%)	$\Delta Sim_{US,China}$	Δ Welfare (%) China	US	Others	World
US 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

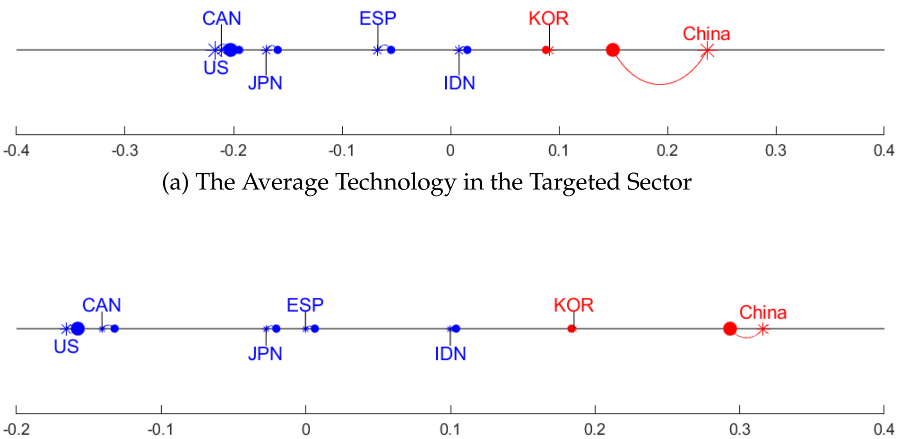
4.4 Counterfactual: Trade Conflict and Technology Decoupling

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, also other countries. In our data, this sector represents approximately 19.7% of China's total imports, with 1.2% coming directly from the U.S.

To implement this counterfactual, we shut down exports of this sector to China by raising the corresponding trade costs to infinity. We investigate two cases, one in which China is embargoed by the United States only, and the other by all countries. Table 7 reports the main findings. When only the U.S. imposes the embargo, 1.2% of China's imports in this sector are directly affected. The impact on the technology distance between the two countries is minimal, and while the embargo imposes more harm on China than the U.S., the overall economic damage is limited—reflecting the relatively small role of the U.S. as a direct supplier of semiconductors to China.

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 when only the U.S. imposes the export ban. In addition to the larger loss for China, two other results are noteworthy. First, the distance between Chinese and U.S. technologies increases substantially more than in the first experiment despite the fact that in both cases the sanction originating from the U.S. is the same. This occurs because, among major economies, China and the U.S. occupy the two ends of the technology spectrum. As Chinese technologies shift away from the technologies of other countries due to export ban, they also shift away from the technologies of the U.S. Second, the welfare cost to the U.S. is amplified. This amplification occurs for two reasons: first, the production cost in China increases due to the direct effect of the embargo and the rest of the world now has to pay higher prices for Chinese products. Second, the endogenous divergence in the technologies around the globe makes souring intermediate inputs even more costly.

The Role of Endogenous Technology. To shed light on the role of endogenous tech-



(a) The Average Technology in the Targeted Sector

(b) The Average Technology of All Sectors

Figure 4: Average Mean of Technology Distributions

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

nology, Figure 4 plots the changes in countries' technology 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 the U.S. and China.[35] This divergence results in a re-alignment of the technology of other countries. We indicate using blue if a country's technology moves in the direction of the U.S. Countries like Japan, Mexico, and European nations move closer to the U.S., while Korea is one of the few that shifts 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 their shifts are smaller than that of the U.S. This widening technology gap contributes to amplified welfare costs for the U.S. and the rest of the world.

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

To underscore the importance of endogenous technology choice, we rerun the embargo scenarios under the assumption that firms cannot adapt their technologies after the shock, effectively freezing technology choices at pre-embargo levels ('Fixed Technology'). The results, shown in the first row of Table 8, reveal that when technology adaptation is restricted, welfare losses are significantly smaller for China, the U.S., and the world. In contrast, when

[35] The theory suggests the embargo will directly impact the technology choices of downstream sectors. For the Computer, Electronic, and Optical Products sector, which is highly interconnected, the most affected downstream sector is itself, as it relies heavily on its own inputs, according to the input-output table.

Table 8: Mechanism Decomposition

	$\Delta Sim_{US,China}$	Δ Welfare (%)			
		China	US	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

firms can adapt their technology, technology responses amplify the 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 adapting their technology to the shock, the broader economy suffers from the negative externalities arising from technology incompatibility across sectors and countries. As firms adapt, these misalignments compound, increasing the overall welfare costs of the embargo.

We decompose the role of various margins of technology adjustment in Rows 2-4 of Table 8, gradually introducing technology responses—from the mostly directly affected sector to the remaining sectors in China, and to other countries. Row 2 reveals that when firms in the directly targeted sector in China adjust their technology, it actually reduces China's welfare losses. These firms can re-optimize by balancing sourcing efficiency with adaptation costs in response to changes in sourcing origins, as reflected in Figure 4(a). 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 technologies, the externalities for China outweigh the benefits of flexibility in the targeted sector, leading to greater welfare losses for China compared to the 'Fixed Technology' scenario, where no adaptation is allowed. Finally, Row 4 demonstrates that when firms in other countries adjust their technologies, technological decoupling becomes more pronounced, further increasing welfare losses globally. An important finding from this decomposition is that a significant portion of the welfare loss to the U.S. is driven by the endogenous technology decoupling that occurs as countries respond to the embargo by adjusting their technologies.

Remarks. These counterfactual findings echo the mechanisms discussed in Propositions 6 and 7. However, in a general equilibrium model with multiple countries, endogenous technological divergence between any two does not always reduce welfare. For instance, if a trade conflict between Korea and Japan leads to a technological divergence, this could potentially improve welfare—provided that the positive externalities of Korea aligning more closely with U.S. technology, and Japan with Chinese technology, outweigh the negative externalities of divergence between Korea and Japan. This possibility highlights the value of the general equilibrium framework and the discipline from the data on the technology of countries.

5 Conclusion

In this paper, we present a model of production networks with endogenous horizontal technologies. In our framework, firms make joint decisions about their technology and suppliers, with these choices interacting due to the compatibility incentives we introduce. General equilibrium forces in a multi-country, multi-sector trade context further shape this interaction, compounding the externalities in firms' technology choices. We establish sufficient conditions for the existence and uniqueness of equilibrium and provide aggregation results that make the model suitable for quantitative application.

Using patent and trade data, we provide novel firm- and country-level evidence that supports the core mechanisms of our model. Our quantitative analyses reveal three main findings. First, endogenous technology due to trade plays an important role in shaping global technologies, accounting for *two-fifth* of the variations in technological proximity between countries, whereas the differences in endowment distributions explain the remainder. Second, technology incompatibility poses an average cost equivalent to 9.5% in ad valorem terms in international trade, underscoring the importance of this mechanism. Lastly, trade conflicts between the U.S. and China can lead to technological decoupling between them and trigger realignments by other countries. These endogenous responses magnify 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 a few 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 develop common standards. Investigating the equilibrium impacts of such mechanisms represents a promising research avenue. Second, our model is static and therefore abstracts from dynamics and vertical innovation. Integrating our model into a dynamic setting requires additional work, but such an effort can illuminate how the compatibility incentive introduced in this model interacts with growth and how such interactions play out in the production network.

References

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

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

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.

Alfaro-Urena, Alonso, Juanma Castro-Vincenzi, Sebastián Fanelli, and Eduardo Morales, "Firm Export Dynamics in Interdependent Markets," Technical Report, National Bureau of Economic Research 2023.

Allen, Treb, Costas Arkolakis, and Xiangliang Li, "On the equilibrium properties of spatial models," *Am. Econ. Rev.: Insights (Forthcoming)*, 2023.

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.

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.

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.

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

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

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

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

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

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

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

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

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

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

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

— **, Samuel S Kortum, and Francis Kramarz**, "Firm-to-Firm Trade: Imports, exports, and the labor market," Technical Report, 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 Linkages and Firm Value: Evidence from the 2018 US-China" Trade War"," 2018.

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

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

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

Lancaster, Kelvin, *Variety, equity, and efficiency: product variety in an industrial society*, Columbia University Press, 1979.

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

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

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

Lind, Nelson and Natalia Ramondo, "Global Innovation and Knowledge Diffusion," *NBER Working Paper Series*, January 2022.

— **and** — , "Trade with correlation," *American Economic Review*, 2023, *113* (2), 317–353.

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

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

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," *arXiv preprint arXiv:2210.07316*, 2022.

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

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

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

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

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

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

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

Proof.

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

□

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

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

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

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

provided that the integral exists.

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

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

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

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

Therefore,

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

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

Proof of Proposition 1. We prove by guess and verify. 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 (2). We verify that this is consistent with firms' behaviors in the equilibrium.

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

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

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

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

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

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

The probability that no such supplier arrives is

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

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

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

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

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

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

where

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

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

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

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

where

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

This finishes the proof of Proposition 1.

□

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

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

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

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

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

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

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

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

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

$$P_dQ_d - \sum_j\sum_o\int_0^1[\frac{\eta}{\eta-1}\tau^{Uj}_{do}p^{j}_{do}(\omega)]q^{j}_{do}(\omega)\mathrm{d}\omega$$

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 in sector j over all goods in the sector is given by

$$\frac{\frac{\eta}{\eta-1}\tau_{do}^{Uj}p_{do}^{j}(\omega)x_{do}^{j}(\omega)}{\rho_{d}^{j}P_{d}Q_{d}}=\frac{[\frac{\eta}{\eta-1}\tau_{do}^{Uj}p_{do}^{j}(\omega)]^{1-\eta}}{(P_{d}^{j})^{1-\eta}},$$

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

$$\begin{aligned}(P_{d}^{j})^{1-\eta} &= \sum_{o}\int_{0}^{1}[\frac{\eta}{\eta-1}\tau_{do}^{Uj}p_{do}^{j}(\omega)]^{1-\eta}\mathrm{d}\omega \\ &= \sum_{o}\int[\frac{\eta}{\eta-1}\tau_{do}^{Uj}p_{do}^{j}(\tilde{\theta})]^{1-\eta}\mathrm{d}\Theta_{o}^{j}(\tilde{\theta}) \\ &= \sum_{o}[\frac{\eta}{\eta-1}\tau_{do}^{Uj}]^{1-\eta}\int\mathbb{E}[p_{do}^{j}(\tilde{\theta})]^{1-\eta}\mathrm{d}\Theta_{o}^{j}(\tilde{\theta}) \\ &= \sum_{o}\Gamma(1+\frac{1-\eta}{\lambda})\cdot[\frac{\eta}{\eta-1}\tau_{do}^{Uj}]^{1-\eta}\cdot(\bar{\Lambda}_{o}^{j})^{1-\eta}, \end{aligned} \tag{A.4}$$

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

$$\bar{\Lambda}_{o}^{j}\equiv(\int[C_{o}^{j}(\tilde{\theta})]^{1-\eta}\mathrm{d}\Theta_{o}^{j}(\tilde{\theta}))^{1/(1-\eta)}.$$

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

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

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

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

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

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

A.2 Proof of Proposition 2

We restate Proposition 2 below:

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

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

following first-order condition with a unique solution.

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

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

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

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

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

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

Outline of Proofs

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

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

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

$$\mathcal{C}_d^i(\bar{\theta};\boldsymbol{g},\boldsymbol{g}') \equiv \ln \xi_d^i - \sum_j \frac{\gamma^{ij}}{\zeta} \ln\left(\sum_o \int [\tau_{do}^j]^{-\zeta} \exp\left(-\zeta \mathcal{C}_o^j(\tilde{\theta};\boldsymbol{g},\boldsymbol{g}')\right) \exp\left(-\frac{1}{2}\zeta\bar{t}(g_d^i(\bar{\theta}) - g_o^j(\tilde{\theta}))^2\right) [g_o^j]'(\tilde{\theta})\bar{\varsigma}_o^j(\tilde{\theta})\mathrm{d}\tilde{\theta}\right), \tag{A.7}$$

$$\chi_{do}^{ij}(\bar{\theta},\tilde{\theta};\boldsymbol{g},\boldsymbol{g}') \equiv \frac{[\tau_{do}^j]^{-\zeta} \exp\left(-\zeta \mathcal{C}_o^j(\tilde{\theta};\boldsymbol{g},\boldsymbol{g}')\right) \exp\left(-\frac{1}{2}\zeta\bar{t}(g_d^i(\bar{\theta}) - g_o^j(\tilde{\theta}))^2\right) [g_o^j]'(\tilde{\theta})\bar{\varsigma}_o^j(\tilde{\theta})}{\sum_m \int [\tau_{dm}^j]^{-\zeta} \exp\left(-\zeta \mathcal{C}_m^j(\tilde{\tilde{\theta}};\boldsymbol{g},\boldsymbol{g}')\right) \exp\left(-\frac{1}{2}\zeta\bar{t}(g_d^i(\bar{\theta}) - g_m^j(\tilde{\tilde{\theta}}))^2\right) [g_m^j]'(\tilde{\tilde{\theta}})\bar{\varsigma}_o^j(\tilde{\tilde{\theta}})\mathrm{d}\tilde{\tilde{\theta}}}. \tag{A.8}$$

- Lemma A.3 establishes the existence and uniqueness of $\boldsymbol{\mathcal{C}}$ given $\boldsymbol{g},\boldsymbol{g}'$ by formulating (A.7) as a fixed point problem for $\boldsymbol{\mathcal{C}}$. It also establishes that $\boldsymbol{\mathcal{C}}$ is continuous in $\boldsymbol{g},\boldsymbol{g}'$ and is differentiable in $\bar{\theta}$ with bounded derivative for any given $\boldsymbol{g},\boldsymbol{g}'$. It follows that $\boldsymbol{\chi}$ is also continuous in $\boldsymbol{g},\boldsymbol{g}'$, and that $\mathcal{T}$ is continuous in $\boldsymbol{g}$ with the norm $\|\cdot\|_{\mathcal{G}}$.

- Lemma A.4 further characterizes the uniform bounds for $\mathcal{T}\boldsymbol{g}$, $[\mathcal{T}\boldsymbol{g}]'$ and the Lipschitz continuity of $[\mathcal{T}\boldsymbol{g}]'$ which are used to show $\mathcal{T}\boldsymbol{g} \in \mathcal{G}$.

- Since $\boldsymbol{g} \in \mathcal{G}$ which is closed, and has uniformly bounded values and Lipschitz continuous first derivatives, by the Arzelà-Ascoli theorem, $\mathcal{G}$ is compact under the norm $\|\cdot\|_{\mathcal{G}}$. That $\mathcal{T}\boldsymbol{g} \in \mathcal{G}$ and $\mathcal{T}$ is continuous under the norm $\|\cdot\|_{\mathcal{G}}$ thus ensures the existence of a fixed point of $\mathcal{T}$ in $\mathcal{G}$ by the Schauder fixed-point theorem. These arguments are formalized in Proposition A.2 and its proof.

- **For uniqueness**, we treat (A.6) and (A.7) as a joined fixed point problem for $(\boldsymbol{g},\boldsymbol{\mathcal{C}})$. We show that with $\bar{t}$ small enough and $\bar{\phi}$ large enough as stated in the Proposition, the joint operator defined by the left hand side of the equations is a contraction mapping under the C^1 norm of the space of $\boldsymbol{g}$ combined with the C^0 norm of the space of $\boldsymbol{\mathcal{C}}$. This is done by showing the induced matrix norm (maximum absolute row sum norm) of the Jacobian matrix that contains the Frechet derivatives of the operator with respect to $(\boldsymbol{g},\boldsymbol{\mathcal{C}})$ is uniformly bounded by a number below one, when $\bar{t}$ is small and $\bar{\phi}$ is large enough. The estimates of the Jacobian entries are presented in Lemma A.5 and the uniqueness proof is formally stated in Proposition 2.

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

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

we see that if χ_{do}^{ij} 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 $\bar{\chi}_{do}^{ij}(\tilde{\theta}) = \chi_{do}^{ij}(\bar{\theta},\tilde{\theta};\boldsymbol{g})$ under the premise, then the equation above is reduced to $[\mathcal{T}\boldsymbol{g}]_d^i(\bar{\theta}) = \omega^i \sum_{j,o} \frac{\gamma^{ij}}{1-\gamma^{iL}} \int \bar{\chi}_{do}^{ij}(\tilde{\theta}) g_o^j(\tilde{\theta})\mathrm{d}\tilde{\theta} + (1-\omega^i)\bar{\theta}$, with $\mathcal{T}$ trivially satisfying Blackwell's sufficiency conditions for contraction with a modulus $\max_i \omega^i$.)

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

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

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

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

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

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

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

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

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

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

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

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

We now verify $\mathcal{T}^{\mathcal{C}}$ satisfies Blackwell's sufficiency conditions for contraction. For any $\boldsymbol{C},\widehat{\boldsymbol{C}} \in \mathbb{C}$, such that $\boldsymbol{C} \le \widehat{\boldsymbol{C}}$ point-wisely, it trivially holds that $[\mathcal{T}^{\mathcal{C}}\boldsymbol{C}]^i_d \le [\mathcal{T}^{\mathcal{C}}\widehat{\boldsymbol{C}}]^i_d$. And it holds that $\mathcal{T}^{\mathcal{C}}[\boldsymbol{C}(\bar{\theta};\boldsymbol{g},\boldsymbol{g}') + c]^i_d = [\mathcal{T}^{\mathcal{C}}\boldsymbol{C}]^i_d(\bar{\theta};\boldsymbol{g},\boldsymbol{g}') + (1-\gamma^{iL})c \le [\mathcal{T}^{\mathcal{C}}\boldsymbol{C}]^i_d(\bar{\theta};\boldsymbol{g},\boldsymbol{g}') + (1-\underline{\gamma}^L)c$ for $c \ge 0$. $\mathcal{T}^{\mathcal{C}}$ 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

$$\begin{aligned}
&[\mathcal{T}^{\mathcal{C}}\boldsymbol{C}]^{i\,\prime}_d(\bar{\theta};\boldsymbol{g},\boldsymbol{g}') = \sum_j \gamma^{ij} \sum_o \int \chi^{ij}_{do}(\bar{\theta},\tilde{\theta};\boldsymbol{g},\boldsymbol{g}')\left[\bar{t}(g^i_d(\bar{\theta}) - g^j_o(\tilde{\theta}))\right]\mathrm{d}\tilde{\theta} \\
\Rightarrow\ &\left|[\mathcal{T}^{\mathcal{C}}\boldsymbol{C}]^{i\,\prime}_d(\bar{\theta};\boldsymbol{g},\boldsymbol{g}')\right| \le (1-\underline{\gamma}^L)2tM,
\end{aligned}$$

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

□

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

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

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

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

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

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

Next, consider

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

in which the second line applies part (1) of Lemma A.6. Since $0 < [g^i_d]'(\bar{\theta}) < 1$ and by Popoviciu's inequality on variances $cov^{ij}_d[g^j_o(\tilde{\theta}), g^j_o(\tilde{\theta})] \le M^2$ we have that

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

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

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

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

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

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

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

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

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

□

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Then we have

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

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

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

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

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

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

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

Apply part (1) of Lemma A.7:

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

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

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

Therefore,

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

Consider

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

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

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

Consider

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

For part (3), expand

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

Apply part (1) of Lemma A.7:

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

Apply part (2) of Lemma A.7:

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

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

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

□

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

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

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

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

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

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

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

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

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

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

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

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

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

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

□

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

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

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

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

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

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

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

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

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

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

□

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

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

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

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

Proof. For part (1),

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

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

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

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

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

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

For part (2),

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

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

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

For part (3),

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

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

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

□

A.3 Proof of Proposition 3

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

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

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

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

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

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

FOC w.r.t. θ reads

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

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

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

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

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

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

Rearranging gives

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

where

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

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

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

with

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

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

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

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

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

Therefore,

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

with Λ defined in the proposition.

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

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

Therefore,

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

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

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

Therefore,

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

□

A.4 Proof of Proposition 4

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

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

Hence,

$$\begin{aligned}
\mathrm{d}\theta^i_d &= \frac{-\zeta\omega^i\gamma^{ij}\bar{\chi}^{ij}_{do}[\theta^j_o - \sum_m\bar{\chi}^{ij}_{dm}\theta^j_m]\mathrm{d}\ln\tau^{ij}_{do}}{1 - 2\zeta\bar{t}\omega^i\sum_{j',o'}\gamma^{ij'}\bar{\chi}^{ij'}_{do'}\theta^{j'}_{o'}[\theta^{j'}_{o'} - \sum_m\bar{\chi}^{ij'}_{dm}\theta^{j'}_m]} \\
&= -\frac{\zeta\omega^i\gamma^{ij}\bar{\chi}^{ij}_{do}\|\theta^j_o - \vartheta^{ij}_d\|}{1 - 2\bar{t}\zeta\omega^i\sum_{j',o'}\gamma^{ij'}\bar{\chi}^{ij'}_{do'}\|\theta^{j'}_{o'} - \vartheta^{ij'}_d\|} \times \frac{\mathrm{d}\ln\tau^{ij}_{do}}{\theta^j_o - \vartheta^{ij}_d}.
\end{aligned}$$

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

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

It follows that

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

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

A.5 Proof of Proposition 5

Proof. From the first-order condition, for

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

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

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

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

The rest of the proposition follows from the following expression:

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

□

A.6 Proof of Proposition 6

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

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

And the household utility in the decentralized equilibrium is

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

where

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

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

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

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

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

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

$$=\rho^i[\frac{\exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}{\eta-\sum_i\rho^i\exp(-\bar{\phi}(\theta^i-\bar{\theta}^i)^2)}\bar{\phi}(\bar{\theta}^i-\theta^i)-\bar{t}\sum_j\tilde{\gamma}^{ij}(\theta^i-\theta^j)]-\bar{t}\sum_{j\neq i}\rho_j\tilde{\gamma}^{ji}(\theta^i-\theta^j) \tag{A.10}$$

The decentralized θ^i satisfies

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

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

Plugging this to equation (A.10) delivers

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

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

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

Since input-output coefficients are symmetric across all sectors,

$$\forall j\neq i,\ \frac{\gamma^{ij}}{\sum_{j\neq i}\gamma^{ij}}=\frac{1}{J-1}=\frac{\tilde{\gamma}^{ij}}{\sum_{j\neq i}\tilde{\gamma}^{ij}}$$

$$\begin{aligned}\sum_{j\neq i}\gamma^{ij}(\theta^i-\theta^j)&=\sum_{j\neq i}\gamma^{ij}\theta^i-\sum_{j\neq i}\gamma^{ij}\theta^j=(\sum_{j\neq i}\gamma^{ij})(\theta^i-\sum_{j\neq i}\frac{\gamma^{ij}}{\sum_{j\neq i}\gamma^{ij}}\theta^j)\\&=\frac{(\sum_{j\neq i}\gamma^{ij})}{(\sum_{j\neq i}\tilde{\gamma}^{ij})}(\sum_{j\neq i}\tilde{\gamma}^{ij})(\theta^i-\sum_{j\neq i}\frac{\tilde{\gamma}^{ij}}{\sum_{j\neq i}\tilde{\gamma}^{ij}}\theta^j)\\&=\frac{\gamma^{ij}}{\tilde{\gamma}^{ij}}\sum_{j\neq i}\tilde{\gamma}^{ij}(\theta^i-\theta^j)\end{aligned}$$

$$\implies sign(\sum_{j\neq i}\gamma^{ij}(\theta^i-\theta^j))=sign(\sum_{j\neq i}\tilde{\gamma}^{ij}(\theta^i-\theta^j))\text{ and }|\sum_{j\neq i}\tilde{\gamma}^{ij}(\theta^i-\theta^j)|>|\sum_{j\neq i}\gamma^{ij}(\theta^i-\theta^j)|$$

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 $\bar{\theta}^i > \theta^i > \frac{\sum \gamma^{ij}}{1-\gamma^{iL}}\theta^j$, we can prove that decreasing θ^i increases the social welfare analogously. □

A.7 Proof of Proposition 7

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

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

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

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

For $i = 1, 2$,

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

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

The technology choice of firms in country 2 is

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

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

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

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

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

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

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

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

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

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

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

Also,

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

□

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

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

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

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

with C_d^i determined as the fixed point of

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

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

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

The probability that no such supplier arrives is

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

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

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

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

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

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

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

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

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

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

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

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

where the sector-level price index is defined by

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

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

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

and sales to final-good producers should satisfy

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

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

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

Finally, labor market clearing condition requires

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

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

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

Appendix B Reduced-Form Evidence

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

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

tion, sales, employment, etc.

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.

We link the NBSC Database to patent data provided by China's State Intellectual Property Office (SIPO). This link is established by exactly matching the firms' names (in Chinese). For each patent filed by the Chinese firm, we establish link with the citation data from PATSTAT by matching the unique patent application number with that of the citing patent. To mitigate the concern on missing patents in the match with the SIPO dataset, we supplement with patents in PATSTAT applied by exactly the same firm. This increases the number of patents from 1,987,313 to 2,810,083 (a nearly 30% increase).

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

Table B.2: 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

Note: 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.

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.

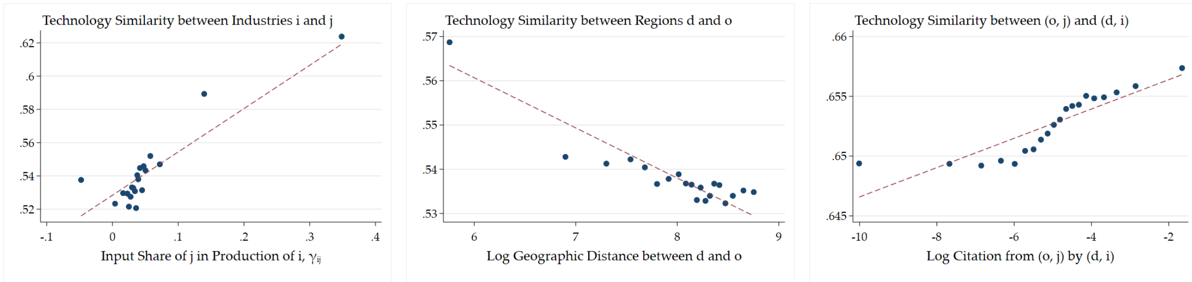
B.2 Technology Similarity Measure Based on Patent Texts

In Section 3.1, we construct a measure of technology similarity based on the closeness of patent texts. In this subsection, we validate this patent-text-based similarity measure by correlating with other factors that are likely associated with technology proximity.

Firstly, we show in Table B.2 that patent abstracts within the same technology classes tend to have higher average similarity, with finer classification levels showing even greater similarity.

Secondly, 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 $\gamma^i j$, 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.

Thirdly, 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



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

Figure B.1: Validation Checks for Technology Similarity

Note: 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.

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.

Lastly, 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 firms in different 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 link full list of patents to global firms. For each 3-digit ISIC industry, We count the number of patenting firms, and among these, how many have patents that contain keywords 'compatibility/compatible' or 'interoperability.' Since 'compatibility' can appear with other meanings such as 'water compatibility', 'skin compatibility', 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. We note that all our results are robust if we do not impose this restriction.

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.3 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 $\text{Compatibility}_{i(\omega)}$.

Table B.3: Compatibility Intensity by Industries

ISIC	Industry	Compatibility Intensity
21	Manufacture of Basic Pharmaceutical Products	0.52
12	Manufacture of Tobacco Products	0.46
19	Manufacture of Coke and Refined Petroleum Products	0.45
20	Manufacture of Chemicals and Chemical Products	0.41
26	Manufacture of Computer, Electronic and Optical Products	0.35
30	Manufacture of Other Transport Equipment	0.32
18	Printing and Reproduction of Recorded Media	0.29
32	Other Manufacturing	0.27
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 Plastics Products	0.24
11	Manufacture of Beverages	0.23
10	Manufacture of Food Products	0.22
23	Manufacture of Other Non-Metallic Mineral Products	0.21
13	Manufacture of Textiles	0.21
33	Repair and Installation of Machinery and Equipment	0.20
28	Manufacture of Machinery and Equipment N.E.C.	0.18
16	Manufacture of Wood and of Products of Wood and Cork, Except Furniture	0.16
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
15	Manufacture of Leather and Related Products	0.10

Note: This table lists the compatibility intensity for 2-digit manufacturing industries in ISIC (Rev. 4). Compatibility intensity is defined as the fraction of patenting firms with patents containing keywords similar to 'compatibility'.

B.4 Additional Details on Firm-Level Correlation

B.4.1 Input and Output Tariffs

In one of the specifications, 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 region d imposed on products in industry i from region o in period t:

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

where applied tariffs at a region-pair level are aggregated from the country-pair level by weighing with initial trade shares

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

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:

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

where $\gamma_{i,j(HS6)}$ denotes the industry-level input-output weight derived from the 2007 Chinese Input-Output Table. Throughout 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.

Table B.4: Extended Gravity

	$\mathbb{I}[\text{Import}_{\omega ojt} > 0]$			
	(1)	(2)	(3)	(4)
$\text{Similarity}_{\omega ot}$	0.014***	0.013***	0.024***	0.024***
	(0.005)	(0.005)	(0.009)	(0.009)
$\text{Similarity}_{\omega,\mathcal{R}^{tech}_t(o),t}$	0.083***	0.095***	0.156***	0.158***
	(0.031)	(0.030)	(0.046)	(0.046)
$\text{Similarity}_{\omega,\mathcal{R}^{tech}_t(o)\cap\mathcal{R}^{geo\|lang}(o),t}$				
$\text{Similarity}_{\omega,\mathcal{R}^{geo}(o),t}$		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
$\mathbb{I}[\text{Import}_{\omega,\mathcal{R}^{tech}_t(o),t} > 0]$	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

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered by firm. Industries are Input-Output Industries equivalent to 3-digit CIC. $X_{i(\omega)ot}$ include both the output applied tariffs of industry i and the input tariffs, defined by the average applied tariffs of importing the goods in industry j from world region o, weighted by the input share of industry j in industry i production.

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 $R_t^{tech}(d)$, as the regions o's with the averaged residual ranking at the top 10% for d during period t.

In column (1) of Table B.4, we extend our baseline specification by including an indicator for whether firm ω cites patents from the technological neighbours of o. Because this indicator is likely correlated with firm ω's importing from these technological neighbours, and because importing from different countries can be substitutes, we control for whether ω imports from these technological neighbours. 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 neighbours as those sharing a common official language. We find that the coefficients associated with proximity to language and geographic neighbours do not matter on themselves; more importantly, they do not diminish the role of similarity with technology neighbours.

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.5 shows that our empirical results in Section 3.2 hold if we use citations to measure technology similarity.

Table B.5: Firm-Level Correlation between Trade and Citation

	$\mathbb{I}[\text{Import}_{\omega ot} > 0]$			
	(1)	(2)	(3)	(4)
$\mathbb{I}[\text{Citation}_{\omega ot} > 0]$	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

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered by firm. $X_{i(\omega)ot}$ include both the output applied tariffs of industry i and the input tariffs, defined by the average applied tariffs of importing the goods in industry j from world region o, weighted by the input share of industry j in industry i production. Industries are at the three digit level.

B.5 Additional Details on Tariff Variation and Technological Choice

B.5.1 MFN-Binding Tariffs

We instrument the effectively applied tariffs with the exogenous tariff changes from binding MFN tariffs. To construct this instrument, we first trace a series of MFN-changing applied tariff at the country-product level:

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

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 $t-1$:

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

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

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

The MFN-binding tariff rates $\tau^{MFN}_{do,HS6,t}$ are then used as the instrument for the effectively applied tariff rates $\tau^{Applied}_{do,HS6,t}$. 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.

Table B.6: 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

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. 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 $\ln \tau^{Applied}_{do,HS6,t}$.

B.5.2 Additional Evidence

Table B.6 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 Broad Economic Categories (BEC) and its mapping with HS products provided by 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 (26), and the ex-post technology distribution is characterized by (28). Then, under Assumption 3, by taking the log of (3), we have

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

where

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

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

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

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

Therefore,

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

where

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

Consider

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

where

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

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

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

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

where

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

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

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

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

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

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

Taking log and apply the quadratic approximation, this is

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

The first-order condition implies that

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

which gives the policy function (27)

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

where

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

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 latter integrand is a density. □

C.2 Fitting the Posterior Technology Distributions

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

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

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

$$\begin{aligned}\min_{\{\tilde{\mu}_d^i\}} \sum_{d,i,o,j} [\tilde{\mu}_d^i \tilde{\mu}_o^j - B_{do}^{ij}]^2, \\ s.t. \quad \sum_{d,i} \tilde{\mu}_d^i = 0\end{aligned} \tag{C.3}$$

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.3) is uniquely given by the eigenvector of B associated with the largest eigenvalue, up to a scaling factor.[2] 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.3 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 $\{\tau_{do}^j\}$, preference η, production technology $\{\gamma^{ij}, \gamma^{iL}, \Xi^i, A_d^i, \zeta, \bar{t}, \bar{\phi}\}$, and the ex-ante technology distribution $\{\bar{\mu}_d^i, \bar{\sigma}^i\}$, we solve for $\{k_{A,d}^i, m_A^i, n_{A,d}^i\}$ and $\{\mu_d^i, \sigma^i\}$ to obtain the cost functions $\{C_o^j(\cdot)\}$ and the ex-post technol-

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

ogy distributions. This involves simultaneously solving the following system of equations:

$$m^i_A = \sum_j \gamma^{ij} m^j_B \tag{C.4}$$

$$n^i_{A,d} = \sum_{j,o} \hat{\chi}^{ij}_{do} n^j_{B,o} \tag{C.5}$$

$$k^i_{A,d} = \ln C^i_d(\mu^i_d) - m^i_A[\sum_{j,o} \hat{\chi}^{ij}_{do}(\mu^i_d - n^j_{B,o})]^2 \tag{C.6}$$

$$\mu^i_d = \alpha^i_d + \beta^i \bar{\mu}^i_d \tag{C.7}$$

$$\sigma^i = \beta^i \bar{\sigma}^i \tag{C.8}$$

where

$$m^j_B = \frac{\bar{t}[1 + 2\zeta m^j_A(\sigma^j)^2]}{1 + 2\zeta(m^j_A + \bar{t})(\sigma^j)^2}, \tag{C.9}$$

$$n^j_{B,o} = \frac{1}{1 + 2\zeta m^j_A(\sigma^j)^2}\mu^j_o + \frac{2\zeta m^j_A(\sigma^j)^2}{1 + 2\zeta m^j_A(\sigma^j)^2} n^j_{A,o} \tag{C.10}$$

$$k^j_{B,o} = k^j_{A,o} + \frac{1}{2\zeta}\log[1 + 2\zeta(m^j_A + \bar{t})(\sigma^j)^2] + \frac{m^j_A(\mu^j_o - n^j_{A,o})^2}{1 + 2\zeta m^j_A(\sigma^j)^2} \tag{C.11}$$

$$\hat{\chi}^{ij}_{do} \equiv \frac{\gamma^{ij} m^j_B}{\sum_{j'} \gamma^{ij'} m^{j'}_B} \times \frac{\exp[-\zeta(\ln \tau^j_{do} + k^j_{B,o} + m^j_B(\mu^i_d - n^j_{B,o})^2)]}{\sum_{o'} \exp[-\zeta(\ln \tau^j_{do'} + k^j_{B,o'} + m^j_B(\mu^i_d - n^j_{B,o'})^2)]} \tag{C.12}$$

$$\ln C^i_d(\hat{\theta}^i_d) = \ln(\frac{\Xi^i}{A^i_d}) + \gamma^{iL}\ln(w_d) - \zeta^{-1}\sum_j \gamma^{ij}\ln(\sum_o \exp[-\zeta(\ln \tau^j_{do} + k^j_{B,o} + m^j_B(\hat{\theta}^i_d - n^j_{B,o})^2)]) \tag{C.13}$$

$$\alpha^i_d = \frac{(\eta - 1)m^i_A}{\bar{\phi} + (\eta - 1)m^i_A} n^i_{A,d} = (1 - \beta^i)n^i_{A,d} \tag{C.14}$$

$$\beta^i = \frac{\bar{\phi}}{\bar{\phi} + (\eta - 1)m^i_A}. \tag{C.15}$$

Here, (C.4) and (C.9) form a contraction mapping for $\{m^i_A\}$ with $\{\sigma^i, \beta^i\}$ determined by (C.8) and (C.15). With $\{m^i_A\}$ and $\{\sigma^j\}$ solved, (C.5) and (C.6) form another contraction mapping for $\{n^i_{A,d}, k^i_{A,d}\}$ when $\bar{t}$ is not too large, with $\{\mu^i_d, n^j_{B,o}, k^j_{B,o}, \chi^{ij}_{do}, \ln C^i_d(\hat{\theta}^i_d), \alpha^i_d\}$ determined by (C.7), (C.10), (C.11), (C.12), (C.13) and (C.14) and can be evaluated directly.

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

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

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

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

$$\chi^j_{do}(\theta,\tilde{\theta}) = \frac{[\tau^j_{do} C^j_o(\tilde{\theta}) t(\theta,\tilde{\theta})]^{-\zeta}}{\sum_{o'}[\tau^j_{do'}\Lambda^j_{o'}(\theta)]^{-\zeta}} = \frac{\exp[-\zeta(\ln\tau^j_{do} + k^j_{A,o} + m^j_A(\tilde{\theta}-n^j_{A,o})^2 + \bar{t}(\theta-\tilde{\theta})^2)]}{\sum_{o'}\exp[-\zeta(\ln\tau^j_{do'} + k^j_{B,o'} + m^j_B(\theta - n^j_{B,o'})^2)]}. \tag{C.16}$$

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

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

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

$$\pi^j_{do}(\tilde{\theta}) = \frac{[\tau^{Uj}_{do} C^j_o(\tilde{\theta})]^{1-\eta}}{\sum_{o'}[\tau^{Uj}_{do'}\bar{\Lambda}^j_{o'}]^{1-\eta}} = \frac{\exp[(1-\eta)(\ln\tau^{Uj}_{do} + k^j_{A,o} + m^j_A(\tilde{\theta}-n^j_{A,o})^2)]}{\sum_{o'}\exp[(1-\eta)(\ln\tau^{Uj}_{do'} + k^j_{C,o'})]}. \tag{C.17}$$

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

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

$$\begin{aligned}
P_d Q_d &= w_d L_d + \sum_i \int \Pi^i_d(g^i_d(\bar{\theta});\bar{\theta})\mathrm{d}\bar{\Theta}^i_d(\bar{\theta}) + \sum_i \int K^i_d(g^i_d(\bar{\theta});\bar{\theta})\mathrm{d}\bar{\Theta}^i_d(\bar{\theta}) && \text{(C.18)} \\
&= w_d L_d + \sum_i \int [1-\phi(\theta;(g^i_d)^{-1}(\theta)) + \phi(\theta;(g^i_d)^{-1}(\theta))]\frac{1}{\eta} X^i_d(\theta)\mathrm{d}\Theta^i_d(\theta) && \text{(C.19)} \\
&= w_d L_d + \sum_i \int \frac{1}{\eta} X^i_d(\theta)\mathrm{d}\Theta^i_d(\theta). && \text{(C.20)}
\end{aligned}$$

Substituting (C.16), (C.17) and (C.20) back to (16) and (17), we get

$$\begin{aligned}
X^j_o(\tilde{\theta}) &= \sum_d \sum_i \int \rho^j_d \pi^j_{do}(\tilde{\theta})[(\frac{1}{\eta} + \gamma^{iL}(1-\frac{1}{\eta}))X^i_d(\theta) + \gamma^{iL}M^i_d(\theta)]\mathrm{d}\Theta^i_d(\theta), \\
M^j_o(\tilde{\theta}) &= \sum_d \sum_i \int \gamma^{ij}\chi^j_{do}(\theta,\tilde{\theta})[(1-\frac{1}{\eta})X^i_d(\theta) + M^i_d(\theta)]\mathrm{d}\Theta^i_d(\theta).
\end{aligned}$$

To numerically approximate this system of equations, we discretize the domain of θ into

$\theta \in \widetilde{T} \equiv \{\vartheta_1, \vartheta_2, ..., \vartheta_{N_\theta}\}$ and have

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

This transforms the system of equations into

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

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

where $|\cdot|$ denotes the length of an interval. 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

$$\begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{M} \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}_{X\to X} & \boldsymbol{B}_{M\to X} \\ \boldsymbol{B}_{X\to M} & \boldsymbol{B}_{M\to M} \end{bmatrix} \begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{M} \end{bmatrix}, \tag{C.21}$$

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 $\{\rho_d^i, \tau_{do}^j, \tau_{do}^{Uj}, \gamma^{ij}, \gamma^{iL}, \zeta, \eta\}$. [3]

The linear system (C.21) 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 (20), is satisfied, i.e.,

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

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

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

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

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

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

[3]Note that (C.21) 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

$$t_d \equiv \sum_i \sum_o \sum_j \int \int \frac{t(\theta, \tilde{\theta}) - 1}{t(\theta, \tilde{\theta})} \cdot \hat{M}_{do}^{ij}(\theta, \tilde{\theta}) \; \mathrm{d}\Theta_d^i(\theta) \mathrm{d}\Theta_o^j(\tilde{\theta}), \tag{C.22}$$

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

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

C.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 adoption costs $\bar{\phi}$, input incompatibility costs $\bar{t}$, and those determining production and trade $\{\tau_{do}^j, \tau_{do}^{Uj}, \Xi^i / A_d^i\}$, leaning on the equilibrium conditions of the model.

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

In the first step, we choose the parameters governing the ex-post distributions, $\{\mu_d^i, \sigma^i\}$, to match bilateral similarities between country-sectors, as in Appendix C.2. This step can be carried out without knowing the primitives of the model. Conditional on their own technology, firms' sourcing decisions only depend on the ex-post distributions. We can therefore calibrate the primitives of the model governing trade using only the ex-post distributions and other data. In the second step, we calibrate $\bar{\phi}$ and recover the ex-ante distributions using equation (28).

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

The nested algorithm. We describe the nested calibration algorithm below:

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

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

  (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, 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, return to Step (A).

In Step (A)(b), we need to solve the equilibrium given the ex-post distribution and other parameters. This requires some modifications on the algorithm developed in C.3. Specifically, in Step 1 of the algorithm to solve the equilibrium, the system of equations are simplified into

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

where

$$
\begin{aligned}
m_B^j &= \frac{\bar{t}[1+2\zeta m_A^j(\sigma^j)^2]}{1+2\zeta(m_A^j+\bar{t})(\sigma^j)^2}, \\
n_{B,o}^j &= \frac{1}{1+2\zeta m_A^j(\sigma^j)^2}\mu_o^j + \frac{2\zeta m_A^j(\sigma^j)^2}{1+2\zeta m_A^j(\sigma^j)^2} n_{A,o}^j \\
k_{B,o}^j &= k_{A,o}^j + \frac{1}{2\zeta}\log[1+2\zeta(m_A^j+\bar{t})(\sigma^j)^2] + \frac{m_A^j(\mu_o^j - n_{A,o}^j)^2}{1+2\zeta m_A^j(\sigma^j)^2} \\
\hat{\chi}_{do}^{ij} &\equiv \frac{\gamma^{ij} m_B^j}{\sum_{j'}\gamma^{ij'} m_B^{j'}} \times \frac{\exp[-\zeta(\ln\tau_{do}^j + k_{B,o}^j + m_B^j(\mu_d^i - n_{B,o}^j)^2)]}{\sum_{o'}\exp[-\zeta(\ln\tau_{do'}^j + k_{B,o'}^j + m_B^j(\mu_d^i - n_{B,o'}^j)^2)]} \\
\ln C_d^i(\mu_d^i) &= \ln(\frac{\Xi^i}{A_d^i}) + \gamma^{iL}\ln(w_d) - \zeta^{-1}\sum_j \gamma^{ij}\ln(\sum_o \exp[-\zeta(\ln\tau_{do}^j + k_{B,o}^j + m_B^j(\mu_d^i - n_{B,o}^j)^2)])
\end{aligned}
$$

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

The remaining steps to solve the equilibrium follow exactly with

$$
\begin{aligned}
\chi_{do}^j(\theta,\tilde{\theta}) &= \frac{\exp[-\zeta(\ln\tau_{do}^j + k_{A,o}^j + m_A^j(\tilde{\theta} - n_{A,o}^j)^2 + \bar{t}(\theta-\tilde{\theta})^2)]}{\sum_{o'}\exp[-\zeta(\ln\tau_{do'}^j + k_{B,o'}^j + m_B^j(\theta - n_{B,o'}^j)^2)]} \qquad \text{(C.23)} \\
\pi_{do}^j(\tilde{\theta}) &= \frac{\exp[(1-\eta)(\ln\tau_{do}^{Uj} + k_{A,o}^j + m_A^j(\tilde{\theta} - n_{A,o}^j)^2)]}{\sum_{o'}\exp[(1-\eta)(\ln\tau_{do'}^{Uj} + k_{C,o'}^j)]} \\
X_o^j(\tilde{\vartheta})|\mathrm{d}\tilde{\vartheta}| &= \sum_d\sum_i\sum_\vartheta \rho_d^j \pi_{do}^j(\tilde{\vartheta})|\mathrm{d}\tilde{\vartheta}| \times [(\frac{1}{\eta} + \gamma^{iL}(1-\frac{1}{\eta}))X_d^i(\vartheta)|\mathrm{d}\vartheta| + \gamma^{iL}M_d^i(\vartheta)|\mathrm{d}\vartheta|] \\
M_o^j(\tilde{\theta})|\mathrm{d}\tilde{\vartheta}| &= \sum_d\sum_i\sum_\vartheta \gamma^{ij}\chi_{do}^j(\vartheta,\tilde{\vartheta})|\mathrm{d}\tilde{\vartheta}| \times [(1-\frac{1}{\eta})X_d^i(\vartheta)|\mathrm{d}\vartheta| + M_d^i(\vartheta)|\mathrm{d}\vartheta|].
\end{aligned}
$$

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

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

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

$$
\begin{aligned}
&\frac{\hat{M}_{do}^j}{\sum_{o'}\hat{M}_{do'}^j} \text{ with } \hat{M}_{do}^j \equiv \sum_{\tilde{\vartheta}}\sum_i\sum_\vartheta \gamma^{ij}\chi_{do}^j(\vartheta,\tilde{\vartheta})|\mathrm{d}\tilde{\vartheta}| \times [(1-\frac{1}{\eta})X_d^i(\vartheta)|\mathrm{d}\vartheta| + M_d^i(\vartheta)|\mathrm{d}\vartheta|] \\
&\frac{\hat{X}_{do}^j}{\sum_{o'}\hat{X}_{do'}^j} \text{ with } \hat{X}_{do}^j \equiv \sum_{\tilde{\vartheta}}\sum_i\sum_\vartheta \rho_d^j\pi_{do}^j(\tilde{\vartheta})|\mathrm{d}\tilde{\vartheta}| \times [(\frac{1}{\eta} + \gamma^{iL}(1-\frac{1}{\eta}))X_d^i(\vartheta)|\mathrm{d}\vartheta| + \gamma^{iL}M_d^i(\vartheta)|\mathrm{d}\vartheta|]
\end{aligned}
$$

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

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

$$\tilde{\chi}^j_{do} = \int \chi^j_{do}(\theta,\tilde{\theta})\Theta^j_o(\tilde{\theta}),$$

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 $\bar{t} = 6.1$ 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 the equilibrium wages or prices. For this reason, we can calibrate $\bar{\phi}$ separately.

Calibrate $\bar{\phi}$ and the ex-ante distributions. With all other model primitives calibrated, the last step is to calibrate $\bar{\phi}$, the parameter on innovation costs. We calibrate it by matching the semi-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 $\{\tau^j_{do}\}$ 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^j_{do}, 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 semi-elasticity matches the empirical coefficient of -0.007, yielding a calibrated value of $\bar{\phi} = 0.29$.

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

$$\begin{aligned}
\bar{\mu}^i_d &= \frac{1}{\beta^i}\cdot\mu^i_d - \frac{1-\beta^i}{\beta^i}\cdot n^i_{A,d}\\
\bar{\sigma}^i &= \frac{1}{\beta^i}\cdot\sigma^i\\
\text{with}\quad \beta^i &= \frac{\bar{\phi}}{\bar{\phi}+(\eta-1)m^i_A}.
\end{aligned}$$

This completes the calibration.