The Macroeconomic Implications of Endogenous Production Networks

Yimei Zou

TESI DOCTORAL UPF / ANY 2018

DIRECTOR DE LA TESI
Jaume Ventura

Departament d'Economia i Empresa

Universitat
Pompeu Fabra
Barcelona

To female researchers in economics – nevertheless, you persisted.

Acknowledgements

First of all, I would like to thank my advisor Jaume Ventura. Jaume used to say that I should go to see him more often. The truth is that meetings with Jaume were so intellectually dense and intense that afterwards I needed a full hour to let my brain rest, then a week to digest his feedback, and then another week to battle against the frustration that I simply had no idea what he meant. It is usually by the end of the third week that I began to understand the insights embedded in his comments, which always push my research towards the right direction. I thoroughly enjoyed learning from Jaume and I am extremely grateful to him for his guidance and support throughout my PhD years.

I also had the privilege to learn from many other professors at Pompeu Fabra and CREI. I would like to thank Julian di Giovanni and Gino Gancia for their continuous feedback on my research and for their generous support during the job market process. I benefited significantly from the many conversations with Dávid Nagy and Giacomo Ponzetto. In addition, Andrea Caggese, Alessandra Bonfiglioli, Fernando Broner, Jordi Galí, Manuel García-Santana, and Alberto Martin all offered me valuable advice and suggestions. Even though I did not pursue research in their fields, taking courses with Babara Rossi and Jan Eeckhout is among the most inspiring experiences for me in graduate school.

I would not be able to finish this PhD (literally) without the help of Marta Araque, whose efficiency and kindness always managed to save the day. I am also fortunate to have Julia and Marta as my "academic siblings". Going through this intellectual journey with them had been a great pleasure. I would also like to thank Eda, Francesca, Eugina, and Yining, for sharing with me the ups and downs of research and for reminding me that life holds much more than academics.

Finally, I would like to say thank you to those who had been standing behind me long before I arrived in Barcelona. To Aly, Shuang, and Jeff, for constantly being my inspirations and for believing in me when I did not even believe in myself; to Professor Mark Kuperberg, for first introducing me to the fascinating field of macroeconomics; and to my parents, for always letting me think and decide for myself, for encouraging me to go see the world and welcoming me home whenever the world seems too overwhelming. You are the best parents one could ask for and I love you very much.

Abstract

This thesis explores the macroeconomic implications of endogenous production networks, defined as the collection of input-output linkages in the economy. In the first chapter, I develop a model of international trade to study how production networks adjust to the forces of globalization. Due to the inefficiency of the market equilibrium, the welfare implications of trade liberalization is ambiguous in general. Calibrating the model to trade data between the United States and the rest of the world, I find that a significant part of the welfare gains from trade arises from the endogenous rearrangement of linkages among firms. The second chapter studies the formation of input-output linkages in the context of economic growth. I establish theoretically that, with endogenous input-output linkages, the static cross-industry difference in linkage fixed costs can lead to different productivity growth rates, which in turn give rise to structural changes. A simple calibration of the model to the U.S. economy suggests that, comparing to a model with a fixed production network, the endogenous adjustment of linkages and the resulting structural changes double the welfare gains from a technology shock that lowers the linkage fixed cost universally.

Resumen

Esta tesis explora las implicaciones macroeconómicas de la existencia de redes de producción endógenas, definidas como el conjunto de vínculos "input-output" en la economía. En el primer capítulo, desarrollo un modelo de comercio internacional para estudiar cómo las redes de producción se ajustan a las fuerzas de la globalización. Debido a que el equilibrio de mercado es ineficiente, las implicaciones para el bienestar de una liberalización comercial son ambiguas en general. Calibrando el modelo con datos de comercio entre los Estados Unidos y el resto del mundo, descubro que una parte importante de las ganancias de bienestar derivadas del comercio surge de la reorganización endógena de los vínculos entre las empresas. El segundo capítulo estudia la formación de vínculos input-output en el contexto del crecimiento económico. Establezco teóricamente que, con vínculos de entrada y salida (input-output) endógenos, la diferencia estática en el coste fijo de crear vínculos entre industrias puede conducir a diferentes tasas de crecimiento de la productividad, que a su vez dan lugar a cambios estructurales. Una simple calibración del modelo para la economía de EE. UU. sugiere que, comparado con un modelo con una red de producción fija, el ajuste endógeno de los vínculos, y los cambios estructurales que provoca este ajuste, duplican las ganancias de bienestar esperadas de un shock tecnológico que redujera el coste fijo de la formación de vínculos para todas las empresas.

Preface

The modern economy features increasingly complex goods that use intermediate inputs sourced from various suppliers located all over the world. Take an iPhone for instance, apart from being designed in California and assembled in China, it also embodies a display from Japan, a cellular modem from Germany, and processors from South Korea, all supplied by firms other than Apple. Input-output linkages connect firms with each other, forming production networks that extend beyond the boundaries of industries and countries. Gross output of the United States, which includes intermediate inputs, is 1.75 times its gross domestic product and trade in intermediate inputs account for two thirds of the global trade flows.[1]

Despite their ubiquitous presence, input-output linkages are costly to form and maintain. Suppliers and customers often go through non-trivial search processes before meeting the right business partners. The supplier may then be required to customize its product for the need of the customer and the customer may have to modify its existing production line so that the new parts can be incorporated. Finally, both parties have to communicate and coordinate to ensure the smooth handling and delivery of the intermediate inputs. All of these activities demand economic resources from firms. In fact, the worldwide market for supply chain management software alone has reached an annual turnover of more than 12 billion dollars.[2]

This thesis aims to improve the understanding of the formation of production networks and its macroeconomic implications. Specifically, I ask the following questions. What economic trade-off do firms face when establishing supplier-customer relationships with each other? How do production networks arise from the linkage decisions made by individual firms? How does the distribution of input-output linkages respond to changes in the economic environment? Is the decentralized formation of production networks socially optimal and, if not, where does the inefficiency stem from? I explore the answers to these questions in the context of international trade and economic growth.

In the first chapter of this thesis, I develop a quantitative trade model with

[1]The U.S statistics is for the first quarter of 2018 and drawn from the U.S. Bureau of Economic Analysis. The world trade statistics is for the year 2011 and drawn from the OECD database.

[2]"SAP Leading The Fast-Growing SCM Market With 26% Share." *Forbes*, July 28, 2018. https://www.forbes.com/sites/louiscolumbus/2018/07/28/sap-leading-the-fast-growing-scm-market-with-26-share

endogenous production networks. In the model, firms form linkages with each other both within and across borders, balancing the trade-off between extra revenue brought in by downstream connections and fixed costs required to establish these relationships. The structure of equilibrium production networks depends both on variable trade costs and linkage fixed costs. In particular, trade integration can lead to structural transformations of global production networks, which in turn bring about technological changes on both the firm and the aggregate level. The joint adjustments of domestic and international linkages constitute a new margin along which trade liberalization can affect welfare. I calibrate the model to trade data between the United States and the rest of the world (ROW) over 2000-2014. The model is able to replicate the actual time trend of the value added share in gross trade, as well as several cross-sectional patterns observed in the US-ROW input-output networks. Applying the model, I quantify the welfare gains of moving from autarky to the 2014 equilibrium to be 15.5%, with a quarter of these gains arising solely from the rearrangement of linkages among firms.

The second chapter of this thesis presents a model of economic growth with endogenous input-output linkages to study the interplay between structural changes and the evolution of production networks. Endogenous linkages translate the static difference among industries – the fixed cost of forming firm relationships being lower in some industries than in others – into different productivity growth rates, bringing about structural changes. The expanding industries also become more prominent intermediate input suppliers, a prediction consistent with the empirical pattern in the United States. A simple calibration of the model to the U.S. economy suggests that, comparing to a model with a fixed production network, the endogenous adjustment of linkages and the resulting structural changes double the welfare gains from a technology shock that lowers the linkage fixed cost universally.

Contents

Chapter 1

ENDOGENOUS PRODUCTION NETWORKS AND GAINS FROM TRADE

1.1 INTRODUCTION

Production has become a global process. Nowadays, supply chains extend far beyond borders, forming input-output networks that involve multiple countries. The emergence of international production sharing is evident from two aggregate trends shown in Figure 1. First, the share of intermediate inputs in world exports has been rising steadily over the past two decades. As producers source inputs from foreign suppliers, domestic factors no longer account for all the value embodied in a country's gross exports. The more production processes fragment across borders, the more gross shipments exaggerate the value added content of trade. The second time series in Figure 1 traces the declining ratio of value-added to gross exports for the world as a whole, further demonstrating the growing scope of global production sharing.

Indeed, underlying these aggregate trade patterns is the evolution of global production networks, which I define as the collection of all input-output linkages within and across borders. I provide two pieces of suggestive evidence for the changing network structure of the global economy. First, on the country level, Figure 2 plots the separate shares of domestic and imported intermediate inputs in aggregate output. In general, expenditure shares on domestic intermediate inputs are about three times as high as those on the imported ones, suggesting that input-output linkages are still concentrated within borders. However, international linkages are gaining importance, as most of the countries experienced a rise in the expenditure share on imported intermediate inputs between 2000 and 2014. Second, on the industry level, I focus on the trade between the United States and the rest of the world (ROW), using Figure 3 to visualize how the US-ROW input-output table evolves over time. Between 2000 and 2014, U.S. industries have increased their reliance on foreign suppliers, while U.S. domestic linkages have undergone mixed changes. The takeaway from Figure 2 and 3 is that domestic input-output linkages are more prominent than their international counterparts, but both evolve over time with the latter expanding in recent years. Therefore, the structure of global production networks is neither random nor static.[1]

All of the above facts lead one naturally to the following questions. What economic forces govern the formation of firm linkages within and across borders? How does the structure of global production networks respond to changes in trade frictions? Can endogenous linkages generate new insight, both theoretical and quantitative, into welfare gains from trade? Standard trade models are silent on these issues, for they impose an exogenous input-output structure on the economy. In this paper, I propose a framework capable of answering these questions. To address the first question, I let profit-maximizing firms form relationships with both their domestic and foreign peers, subject to the following tradeoff. On one hand, due to input specificity, trade of intermediate goods can occur only along established buyer-seller relationships. Therefore, firms have the incentive to form linkages in order to acquire input customers and raise sales. On the other hand, setting up business relationships is often a costly activity for firms in the real world. I thus assume that a fixed cost must be paid for every linkage established. All else equal, firms prefer connecting with their peers at home to the ones abroad, because the benefit of international linkages is lower given the cost of shipping goods across borders. The linkage decisions by all individual firms jointly determine the structure of production networks in equilibrium, which in turn determines the availability of intermediate inputs facing each firm. This micro-macro connection then allows me to answer the second question. Trade liberalization, modeled as reductions in trade

[1] The responsive nature of production networks to changes in the economic environment also finds ample support in firm level data. For example, Goldberg et al. (2010) report evidence that tariff reductions in India lead to the expansion of intermediate input variety used by firms. Gopinath and Neiman (2014) identify that, during the 2000-2002 Argentine crisis, up to 45 percent of the collapse in aggregate imports can be accounted for by firms dropping varieties from their intermediate input bundles. Using microdata of Hungarian importers, Halpern et al. (2015) attribute an annual growth of 5.9 percentage points in aggregate imports during the 1990s to firms gaining access to new foreign inputs.

costs, increases the returns to international linkages and prompts firms to expand overseas relationships. Moreover, since firms in my model simultaneously decide on both their domestic and foreign connections, the expansion of international linkages affects the returns to domestic relationships and therefore leads to structural changes also in production networks within borders.

Responding to the third question, I use my model to evaluate gains from trade both theoretically and quantitatively. The results crucially depend on the presence of firm heterogeneity. If firms are homogeneous, whether international trade in intermediate inputs is positive in equilibrium depends on both variable trade costs and linkage fixed costs. For linkage fixed costs above a threshold, intermediate input trade will remain zero despite trade liberalization, because no firm will find it profitable to acquire foreign buyers given the high costs of establishing such connections. Only when linkage fixed costs become sufficiently low will we see falling trade costs lead to the formation of international linkages and therefore the emergence of intermediate input trade. I prove that, below this threshold level of linkage fixed costs, welfare gains from trade are larger than those predicted by the benchmark Krugman (1980) model. These extra gains arise from changes in firm technology, as the expansion of international linkages enables producers to adopt a wider range of intermediate inputs and lower the cost of production. However, when firms are heterogeneous in productivity, trade shocks always induce firms to adjust customer relationships regardless of linkage fixed costs, provided that the productivity support is continuous and unbounded from above. In addition to firm level technological changes, trade liberalization also affects aggregate production technology by reshaping the firm productivity distribution. As firms rearrange downstream linkages in response to falling trade costs, some producers attract new suppliers while others lose input diversity. The differential impact of trade liberalization on heterogeneous firms, together with free entry by productivity, lead to a redistribution of firm mass over the productivity support. Therefore, the heterogeneous firm model embodies an extra channel through which gains from trade can potentially arise from the production side. Nevertheless, the overall welfare impact of trade liberalization is ambiguous in general for the following reason. As trade costs fall, firms may want to connect with high-productivity input customers abroad instead of the low-productivity ones at home. This redistribution of linkages could lead to some firms losing suppliers in absolute terms and thus having to raise prices because of higher production costs.

To settle the ambiguity in theory, I calibrate the heterogeneous firm model to trade data between the United States and the rest of the world from 2000 to 2014, reported by the World Input Output Database. Even though my calibration does not target the value added content of trade, the model successfully replicates the evolution of trade in value added, delivering a model-data correlation above 0.7 over the sample period. I then apply the model to quantify the effect of trade liberalization through two exercises. First, I decompose the actual welfare changes according to their sources. Reductions in variable trade costs and linkage fixed costs contribute respectively 19.3% and 24.2% of the 2000-2014 cumulative welfare gains. Second, I compare the observed equilibrium in 2014 with the autarky equilibrium. To assess specifically the importance of endogenous linkages, I conduct the counterfactuals with both the baseline model and an alternative scenario where linkage distribution (i.e., the matching pattern between suppliers and customers) is fixed at the 2014 equilibrium outcome. Welfare gains from autarky to the trade equilibrium vary from 25.4% to 10.7%, for a range from 4 to 8 of the elasticity of substitution among closely-related varieties. In contrast, the gains solely due to linkage rearrangement appear robust, ranging from 3.3% to 3.9% (equivalent to 13%-37% of the overall gains).

This paper contributes to three strands of literature. The first strand studies the formation of production networks either in the closed economy (Lim 2017, Oberfield 2017) or with international trade (Chaney 2012, Antràs et al. 2017, Tintelnot et al. 2017). Comparing to these existing frameworks, the strength of my model lies in its ability to generate realistic network structure, eliminating several restrictive assumptions adopted by the literature. First, I do not group firms ex-ante into buyers and sellers, nor do I restrict the number of linkages a firm can possess. In my model, any firm may become both a supplier and a customer, with the number of upstream and downstream relationships endogenously determined in equilibrium. Second, I do not assume that production is sequential, thereby allowing the network structure to be cyclic. Third, I do not impose a stochastic process on firms' linkage decision. Instead, firms always have the opportunity to adjust linkages in accordance with profit maximization. Despite such flexibility, the model remains highly tractable and permits analytic characterization of the equilibrium even when firms are heterogeneous.

Next, the paper joins the collective effort to quantify gains from trade in the presence of input-output linkages, such as Goldberg et al. (2010), Costinot and Rodríguez-Clare (2014), Melitz and Redding (2014), Caliendo and Parro (2015), Halpern et al. (2015), Bernard et al. 2017, and Blaum et al. (2017). Unlike these previous frameworks, in my model not only international but also domestic linkages are responsive to trade liberalization. Furthermore, I demonstrate that the rearrangement of firm relationships, both within and across borders, constitutes a quantitatively relevant margin of adjustment. Therefore, ignoring the structural changes of either domestic or international production networks leads one to an incomplete understanding of gains from trade. In a broader sense, this paper is also related to the macro approach of quantifying gains from trade initiated by Arkolakis, Costinot and Rodríguez-Clare (2012, henceforth ACR). The ACR framework identifies two sufficient statistics for the size of gains that are applicable to all trade models satisfying three macro-level restrictions. I show that two of the three ACR restrictions fail to hold once firm linkages become endogenous: both the number of firms and trade elasticity change with variable trade costs, instead of being constant. Consequently, the ACR formula for gains from trade does not apply to my model even in the absence of firm heterogeneity, for the sufficient statistics implied by my model also include production-side moments such as network densities and the factor share of intermediate inputs.

Finally, the paper also speaks to the empirical analysis of trade in value added, from the pioneering work of Hummels et al. (2001) to recent studies by Johnson and Noguera (2012, 2017), Koopman et al. (2014), and Kee and Tang (2016). My contribution to this strand of literature is that I provide a general equilibrium theory for the value added content of trade, where the ratio of value-added to gross trade is determined endogenously as an aggregate outcome of linkage formation by individual firms. In addition, the model succeeded in replicating the observed trend in value-added trade, lending support to the theory. The general equilibrium nature of the model makes it desirable for conducting conterfactuals to identify the key determinants of trade in value added. Calibrating the model to U.S. trade data, I find that reductions in variable trade costs account for most of the observed fall in the value added content of U.S. imports.

The rest of the paper takes the following structure. Section 2 illustrates the model mechanism in a simplified setup with homogeneous firms. Section 3 generalizes the model to accommodate firm heterogeneity. Section 4 derives testable predictions of the model and describes the calibration strategy. Section 5 assesses model fit, conducts counterfactuals, and checks robustness. Section 6 concludes. Proofs of all theoretical results are relegated to the appendix.

1.2 HOMOGENEOUS FIRM MODEL

In this section, I build the intuition of endogenous production networks through a simplified model with homogeneous firms. I first setup the model in Section 2.1, describing the problems that households and firms face. Next, I define and solve for the equilibrium in Section 2.2. To characterize the equilibrium, I begin with the limiting case of autarky in Section 2.3 and then consider the impact of trade in Section 2.4.

1.2.1 SETUP

The world consists of two symmetric countries, each hosting a continuum of industries indexed by i over the unit interval. All industries feature monopolistic competition with free entry of firms. Each firm produces a differentiated variety for two purposes: households may consume the variety as a final good, and other firms may adopt the variety as an intermediate input. Trade is costless within borders, whereas an iceberg transport cost $\tau \geqslant 1$ is imposed on the buyer for any international shipment of goods. I model trade liberalization as reductions in τ. Notation developed below takes into account the fact that I will search for symmetric equilibrium when solving the model.

1.2.1.1 HOUSEHOLDS

A mass L of identical households reside in each country. They supply labor inelastically and consume final good varieties according to nested CES preferences:

$$X = \left[\int_0^1 \left(\int_0^{2N_i} X_{ij}^{\beta} dj \right)^{\frac{\alpha}{\beta}} di \right]^{\frac{1}{\alpha}}$$

where $1/(1-\alpha)$ and $1/(1-\beta)$ are respectively the elasticities of substitution between any two varieties from different industries and from the same industry. Since consumers typically find it easier to substitute among goods within an industry than goods across industries, I assume that $0 < \alpha < \beta < 1$. Goods are indexed by j over $[0, 2N_i]$, where N_i is the total number of varieties produced in industry i per country and is to be determined in equilibrium by free entry of firms. The symmetry of firms implies the symmetry of industries. Therefore, $N_i = N$ for all $i \in [0,1]$ and firms charge the same price p in equilibrium. Household optimization yields the following demand for domestic goods X_H and imported goods X_F:

$$X_H = X \left(\frac{P}{p} \right)^{\frac{1}{1-\alpha}} \left[\left(1 + \tau^{\frac{\beta}{\beta-1}} \right) N \right]^{\frac{\beta-\alpha}{\beta(\alpha-1)}}, \quad X_F = \tau^{\frac{1}{\beta-1}} X_H,$$

where P is the consumer price index:

$$P \equiv \left\{ \left[\left(1 + \tau^{\frac{\beta}{\beta-1}} \right) N \right]^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)} p^{\frac{\alpha}{\alpha-1}} \right\}^{\frac{\alpha-1}{\alpha}}.$$

1.2.1.2 FIRMS

Firms produce differentiated varieties by combining labor with a bundle of intermediate inputs in a Cobb-Douglas fashion:

$$q = \frac{\theta}{\sigma^{\sigma} (1-\sigma)^{1-\sigma}} l^{\sigma} m^{1-\sigma}$$

where $0 < \sigma < 1$ and θ is the common level of total factor productivity (TFP).[2] The intermediate input bundle m aggregates all available varieties in a nested CES form, with separate elasticities of substitution within and across industries:

$$m = \left[\int_0^1 \left(\int_0^{2N_i} \mathbb{I}_{ij} x_{ij}^{\beta} dj \right)^{\frac{\alpha}{\beta}} di \right]^{\frac{1}{\alpha}}$$

where $\mathbb{I}_{ij}$ is an indicator variable taking on value 1 if good j of industry i is accessible via upstream linkages and 0 otherwise. I assume that it is easier for producers to substitute between intermediate inputs from the same industry than between those from different industries: $0 < \alpha < \beta < 1$. Without loss of generality, I order the varieties of any industry i such that those over $[0, N_i]$ are produced by domestic firms and those over $(N_i, 2N_i]$ are produced abroad. Let $\mu_{i,H} \equiv \int_0^{N_i} \mathbb{I}_{ij} dj / N_i$ and $\mu_{i,F} \equiv \int_{N_i}^{2N_i} \mathbb{I}_{ij} dj / N_i$ denote respectively the fraction of domestic and foreign firms from industry i that have made themselves available as input suppliers. Then, $\mu_{i,H} N_i$ and $\mu_{i,F} N_i$ correspond to the number of domestic and foreign suppliers from industry i, both to be determined in equilibrium. Since industries are symmetric, the equilibrium distribution of firm linkages is uniform across industries: $\mu_{i,H} = \mu_H$ and $\mu_{i,F} = \mu_F$ for all $i \in [0, 1]$. Taking supplier availability μ_H and μ_F as given, firms choose quantities of labor l, domestic intermediate inputs x_H, and foreign intermediate inputs x_F to minimize the variable production cost,

[2]In this simplified model, the TFP level θ is no more than a scalar. Later on in Section 3, I will allow θ to follow a non-degenerate distribution as a way of introducing firm heterogeneity.

resulting in the factor demand functions below:

$$x_H = m\left(\frac{P^S}{p}\right)^{\frac{1}{1-\alpha}}\left[\left(\mu_H + \tau^{\frac{\beta}{\beta-1}}\mu_F\right)N\right]^{\frac{\beta-\alpha}{\beta(\alpha-1)}}, \quad x_F = \tau^{\frac{1}{\beta-1}}x_H,$$
$$m = (1-\sigma)\frac{q}{\theta}\left(\frac{w}{P^S}\right)^{\sigma}, \qquad l = \sigma\frac{q}{\theta}\left(\frac{P^S}{w}\right)^{1-\sigma},$$

where P^S is the producer price index:

$$P^S \equiv \left\{\left[\left(\mu_H + \tau^{\frac{\beta}{\beta-1}}\mu_F\right)N\right]^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)} p^{\frac{\alpha}{\alpha-1}}\right\}^{\frac{\alpha-1}{\alpha}}.$$

Following the standard monopolistic pricing rule, firms set variety prices at a constant markup over the marginal cost: $p = w^{\sigma}\left(P^S\right)^{1-\sigma}/(\alpha\theta)$. In addition to fulfilling consumption demand, firms may also sell their output to other firms as intermediate inputs. In order to serve as an input supplier, the firm must first establish a bilateral relationship with the customer by paying a fixed cost of $\kappa\theta$ labor units.[3] Given the symmetric structure of the model, the firm's linkage formation problem boils down to choosing a fraction $\mu_H \in [0,1]$ of domestic firms and a fraction $\mu_F \in [0,1]$ of foreign firms to sell to.[4] Each firm then has a mass of $\mu_H N$ input customers at home and a mass of $\mu_F N$ input customers from abroad. The demand for a variety consists of consumption needs from domestic households X_H and foreign households τX_F, as well as intermediate input needs at home $x_H\mu_H N$ and from abroad $\tau x_F \mu_F N$:

$$q = X_H + \tau X_F + x_H\mu_H N + \tau x_F \mu_F N. \tag{1.2.1}$$

[3]One justification for linkage fixed costs being proportional to the TFP level θ is that, the higher the productivity, the higher the opportunity cost of removing labor from production to linkage formation.

[4]Since firms are symmetric within a country, a supplier does not care about the specific identities of its customers, as long as they exhibit the same demand for intermediate inputs. I thus assume that the specific identities of the customers are selected at random. As a result, μ_H is also the probability of any two firms from the same country having an input-output link between them, and μ_F gives such linking probability between any two firms from different countries.

The first order conditions for the linkage choice variables μ_H and μ_F are given by

$$(1-\alpha)x_H p - \kappa\theta w \begin{cases} \leqslant 0 & \text{if } \mu_H = 0 \\ = 0 & \text{if } \mu_H \in (0,1) \\ \geqslant 0 & \text{if } \mu_H = 1 \end{cases};$$

$$(1-\alpha)\,\tau x_F p - \kappa\theta w \begin{cases} \leqslant 0 & \text{if } \mu_F = 0 \\ = 0 & \text{if } \mu_F \in (0,1) \\ \geqslant 0 & \text{if } \mu_F = 1 \end{cases}.$$

These conditions highlight the tradeoff that a firm faces when forming downstream linkages. On the benefit side, acquiring an extra customer adds $(1-\alpha)x_H p$ to firm profits if the customer is domestic and $(1-\alpha)\,\tau x_F p$ if the customer is foreign. On the cost side, setting up the supplier-customer relationship requires an upfront payment of $\kappa\theta w$. In the presence of trade costs ($\tau > 1$), returns to domestic relationships always exceed those to the international ones, because clients at home place larger orders than buyers from abroad ($x_H > \tau x_F = \tau^{\frac{\beta}{\beta-1}} x_H$). Therefore, a firm never reaches to customers overseas when there are still domestic firms left unconnected. In a symmetric equilibrium, μ_H and μ_F correspond to the densities of domestic and international production networks. Hence, the pair of variables $\{\mu_H, \mu_F\}$ completely define the equilibrium structure of global production networks. In standard trade models where the input-output structure is exogenously imposed, μ_H and μ_F are merely fixed parameters irresponsive to changes in the economic environment.[5] In my model, μ_H and μ_F are equilibrium objects rooted in the firm's profit maximization problem and therefore naturally reactive to shocks. As I show in the forthcoming sections, endogenous production networks make the economy adjust to trade liberalization in new and important ways.

1.2.2 EQUILIBRIUM DEFINITION AND SOLUTION

I close the model with the assumption about firm entry: to enter any industry, a firm must pay a cost of $\nu\theta$ labor units.[6] Normalizing wage w to one, I define the symmetric equilibrium below.

---

[5] For example, the Krugman (1980) model fixes μ_H and μ_F at zero by excluding intermediate inputs from factors of production. Following Ethier (1982), models that feature roundabout production, where final goods can be used as intermediate inputs, essentially assumes fully-connected production networks with $\mu_H = \mu_F = 1$.

[6] One justification for entry costs being proportional to the TFP level θ is that, the higher the productivity, the higher the opportunity cost of removing labor from production to firm creation.

Definition (homogeneous firm equilibrium) *Given parameters α, β, σ, L, κ, ν, τ and θ, a symmetric equilibrium of the homogeneous firm economy consists of variety price p, firm revenue $r \equiv pq$, linkage densities μ_H and μ_F, and firm mass N such that the following conditions hold:*

1. *Households maximize utility; i.e. the consumption demand functions X_H and X_F are given by Section 2.1.1.*

2. *Firms maximize profit; i.e. the factor demand functions (x_H, x_F, l, and m) as well as the pricing rule p are given by Section 2.1.2, and the linkage choice variables μ_H and μ_F satisfy the first order conditions laid out in the same section.*

3. *There is free entry of firms in all industries; i.e. profits are driven down to cover solely the entry costs: $(1-\alpha)r - \kappa\theta(\mu_H+\mu_F)N = \nu\theta$.*

4. *Goods market clears according to equation (1.2.1).*

5. *Labor market clears: $L = [l + \kappa\theta(\mu_H+\mu_F)N + \nu\theta]N$.*

To solve for the equilibrium, I first express variety price p and firm revenue r in terms of the production network structure $\{\mu_H, \mu_F\}$ and aggregate firm mass N:

$$p = (\alpha\theta)^{-\frac{1}{\sigma}} \left[\left(\mu_H + \tau^{\frac{\beta}{\beta-1}}\mu_F\right)N\right]^{-\left(\frac{1-\beta}{\beta}\right)\left(\frac{1-\sigma}{\sigma}\right)}, \quad r = \frac{L}{[1-\alpha(1-\sigma)]N}.$$

Variety price decreases in the total mass of firm linkages (with international linkages discounted by a factor of $\tau^{\frac{\beta}{\beta-1}}$), a direct result of the intermediate input bundle taking nested CES form. The more varieties enter the intermediate input bundle, the lower the marginal cost of production and hence the lower the monopolistic price. Free entry and labor market clearing jointly imply the following relationship between aggregate firm mass N and aggregate linkage mass $(\mu_H+\mu_F)N$:

$$\left[\frac{1-\alpha}{1-\alpha(1-\sigma)}\right]\frac{L}{N} = \nu\theta + \kappa\theta(\mu_H+\mu_F)N \tag{1.2.2}$$

which, together with the following optimality conditions of firm linkages, allow me to solve for the equilibrium value of N, μ_H, and μ_F.

$$\frac{\alpha(1-\alpha)(1-\sigma)L}{[1-\alpha(1-\sigma)]N^2\left(\mu_H + \tau^{\frac{\beta}{\beta-1}}\mu_F\right)} - \kappa\theta \begin{cases} \leqslant 0 & \text{if } \mu_H = 0 \\ = 0 & \text{if } \mu_H \in (0,1) \\ \geqslant 0 & \text{if } \mu_H = 1 \end{cases} \tag{1.2.3}$$

$$\frac{\tau^{\frac{\beta}{\beta-1}}\alpha(1-\alpha)(1-\sigma)L}{[1-\alpha(1-\sigma)]N^2\left(\mu_H+\tau^{\frac{\beta}{\beta-1}}\mu_F\right)}-\kappa\theta\begin{cases}\leqslant 0 & \text{if } \mu_F=0\\ =0 & \text{if } \mu_F\in(0,1)\\ \geqslant 0 & \text{if } \mu_F=1\end{cases}\quad(1.2.4)$$

In the optimal linkage conditions (1.2.3) and (1.2.4), the left hand side of the equality/inequality signs gives the returns to linkages, which are decreasing functions of the equilibrium network densities μ_H and μ_F. This is because intermediate input demands (x_H and x_F) decrease in the number of available varieties, as a result of the assumption that varieties are more substitutable within an industry than across industries $(\alpha<\beta)$. Whether the firm's linkage problem yields an interior solution depends on the linkage fixed cost parameter κ. For sufficiently high κ, marginal profit of foreign relationships is negative even when international linkages are entirely missing.[7] For sufficiently low κ, the marginal profit of domestic (foreign) relationships remains positive even when the firm already supplies inputs to all its domestic (foreign) peers. I formalize these intuitions in Lemma 1.

Lemma 1 *The domestic production network of each country has density:*

$$\mu_H=\begin{cases}\frac{\alpha(1-\sigma)\theta\nu^2}{(1-\alpha)[1-\alpha(1-\sigma)]\kappa L} & \textit{if } \kappa\in(\underline{\kappa}_H,\infty)\\ 1 & \textit{if } \kappa\in[0,\underline{\kappa}_H]\end{cases}$$

where the linkage fixed cost threshold is given by

$$\underline{\kappa}_H=\frac{\alpha(1-\sigma)\theta\nu^2}{(1-\alpha)[1-\alpha(1-\sigma)]L}.$$

The international production network has density

$$\mu_F=\begin{cases}0 & \textit{if } \kappa\in[\overline{\kappa}_F,\infty)\\ \frac{4\alpha(1-\alpha)(1-\sigma)\kappa L\left(\tau^{\frac{\beta}{1-\beta}}-1\right)^2}{[1-\alpha(1-\sigma)]\left[\nu\theta-\sqrt{(\nu\theta)^2-4\left(\tau^{\frac{\beta}{1-\beta}}-1\right)(1-\alpha)\kappa L\theta}\right]^2}-\tau^{\frac{\beta}{1-\beta}} & \textit{if } \kappa\in(\underline{\kappa}_F,\overline{\kappa}_F)\\ 1 & \textit{if } \kappa\in[0,\underline{\kappa}_F]\end{cases}$$

[7]On the contrary, marginal profit of domestic relationships is never negative, for marginal revenue of domestic linkages becomes infinite when the total number of links in the economy approaches zero.

where the linkage fixed cost thresholds are given by

$$\overline{\kappa}_F = \tau^{\frac{\beta}{\beta-1}} \left[\frac{1-\alpha(1-\sigma)}{1-\tau^{\frac{\beta}{\beta-1}}\alpha(1-\sigma)} \right]^2 \underline{\kappa}_H,$$

$$\underline{\kappa}_F = \tau^{\frac{\beta}{\beta-1}} \left(1+\tau^{\frac{\beta}{\beta-1}}\right) \left[\frac{1-\alpha(1-\sigma)}{1+\tau^{\frac{\beta}{\beta-1}}-2\tau^{\frac{\beta}{\beta-1}}\alpha(1-\sigma)} \right]^2 \underline{\kappa}_H.$$

To assess the aggregate impact of trade liberalization, I measure welfare by real wage $W \equiv 1/P$ which has the following expression in equilibrium:

$$W = \left(\Lambda_c \Lambda_m^{\frac{1-\sigma}{\sigma}} \mu_H^{\frac{\sigma-1}{\sigma}} N^{-\frac{1}{\sigma}} \right)^{\frac{\beta-1}{\beta}} (\alpha\theta)^{\frac{1}{\sigma}} \tag{1.2.5}$$

where $\Lambda_c \equiv 1/\left(1+\tau^{\frac{\beta}{\beta-1}}\right)$ is the domestic expenditure share for consumption goods and $\Lambda_m \equiv \mu_H/\left(\mu_H+\tau^{\frac{\beta}{\beta-1}}\mu_F\right)$ is the domestic expenditure share for intermediate inputs. One immediately observes that, when production technology relies entirely on labor ($\sigma = 1$), the economy reduces to a version of Krugman (1980) with nested CES preferences. This should not be surprising as the novelty of my model lies in firm linkages, which are relevant only when factors of production include intermediate inputs.

Proposition 1 *The homogeneous firm model converges to a generalized version (with nested CES preferences) of Krugman (1980) in the limiting case of $\sigma \to 1$. At this benchmark, the elasticity of welfare with respect to market size $\left(1+\tau^{\frac{\beta}{\beta-1}}\right)L$ (i.e. the degree of increasing returns to scale) is $(1-\beta)/\beta$.*

1.2.3 AUTARKY

The autarky case serves a good starting point for equilibrium characterization, since entering free trade from autarky is equivalent to doubling population size in the closed economy. When trade costs are prohibitively high ($\tau \to \infty$), firms find no incentive to establish customer-supplier relationships with their foreign counterparts. The fixed cost thresholds for foreign linkages tend to zero in this limiting case:

$$\lim_{\tau\to\infty} \overline{\kappa}_F = \lim_{\tau\to\infty} \underline{\kappa}_F = 0.$$

As population L increases, not only does the number of varieties grow, but the scale of production also expands, provided that linkage fixed costs κ are low enough. Lemma 2 below summarizes the market size effects on variety mass N and (revenue-based) firm size r.

Lemma 2 *If linkage fixed costs are high such that $\kappa \in [\underline{\kappa}_H, \infty)$, market expansions raise the number of varieties ($dN/dL > 0$) but do not change the size of firms ($dr/dL = 0$). If linkage fixed costs are low such that $\kappa \in [0, \underline{\kappa}_H)$, larger markets lead to not only more varieties ($dN/dL > 0$) but also larger production scale ($dr/dL > 0$).*

The above results showcase the key difference between the Krugman model and my framework. In the former, the benefit of larger markets arises entirely from a more diverse consumption basket, as production scale is unaffected by market size. Contrasting to the Krugman benchmark, my model possesses an additional mechanism of gains from trade, operating on the production side. Specifically, larger markets enable firms to expand downstream linkages, potentially boosting the range of intermediate inputs available for adoption and thereby reducing the marginal cost of production. The relevance of this extra welfare-enhancing mechanism depends naturally on the fixed cost of forming a supplier-customer relationship, regulated by the parameter κ. High levels of κ discourage firms from taking advantage of the widened customer pool to create more connections, thus rendering this additional mechanism irrelevant. The next proposition elaborates on the above argument.

Proposition 2 *In a closed economy with endogenous firm linkages, the degree of aggregate increasing returns to scale depends on linkage fixed costs κ: $d\ln W/d\ln L = (1-\beta)/\beta$ for $\kappa \in [\underline{\kappa}_H, \infty)$, whereas $d\ln W/d\ln L > (1-\beta)/\beta$ for $\kappa \in [0, \underline{\kappa}_H)$. Therefore, when the equilibrium production network is incomplete ($0 \leqslant \mu_H < 1$), the model implies the same market size effect on welfare as in Krugman (1980). However, gains from market expansions are strictly larger than those in Krugman (1980) when the equilibrium production network is complete ($\mu_H = 1$).*

When customer-supplier relationships are relatively expensive to form ($\kappa \geqslant \underline{\kappa}_H$), the marginal profit of downstream linkages falls to zero before firms exhaust all potential clients. Since firms are satiated with the amount of input orders they already receive, enlarging the market alone would not motivate them to acquire additional customers, and the production network is left unchanged. Thus, the market size effect manifests solely in extra consumption varieties. However, when linkage fixed costs are so low that firms find it profitable to connect with every other firm ($\kappa < \underline{\kappa}_H$), the shadow value of customer-supplier relationships is positive but firms are constrained by the number of clients they can possibly obtain. In this case, increasing market size relaxes the customer availability constraint facing suppliers and fosters new firm linkages. Finally, the threshold $\underline{\kappa}_H$ itself is decreasing in market size L, suggesting that small economies enjoy more scope for the production-side gains from trade than large countries.

1.2.4 THE OPEN ECONOMY

I characterize the open economy equilibrium by focusing on the parameter space of trade costs τ and linkage fixed costs κ. Lemma 1 provides a mapping from any combination of the cost parameters $(\tau, \kappa) \in [1,\infty) \times [0,\infty)$ to a configuration $(\mu_H, \mu_F) \in [0,1] \times [0,1]$ of global production networks. To facilitate discussion, I first categorize all possible structure of global production networks into four classes, according to linkage location and density:

Definition (network configuration) *Global production networks are national-partial (NP) if $\mu_H \in (0,1)$ and $\mu_F = 0$; national-complete (NC) if $\mu_H = 1$ and $\mu_F = 0$; international-partial (IP) if $\mu_H = 1$ and $\mu_F \in (0,1)$; international-complete (IC) if $\mu_H = 1$ and $\mu_F = 1$.*

When global production networks are NP or NC, they are no more than a collection of two separate national networks, and international trade involves consumption goods only. Cross-country trade in intermediate inputs occurs only when the world economy exhibits network structure IP or IC. The next lemma partitions the parameter space of trade costs τ and linkage fixed costs κ into areas where one or more of the four aforementioned network structure emerges in equilibrium.

Lemma 3 *If $\alpha(1-\sigma) \leqslant 1/2$, the linkage fixed cost thresholds satisfy $0 \leqslant \underline{\kappa}_F \leqslant \overline{\kappa}_F \leqslant \underline{\kappa}_H < \infty$ for any $(\tau, \kappa) \in [1,\infty) \times [0,\infty)$, and the trade equilibrium is unique. If $\alpha(1-\sigma) > 1/2$, the ranking $0 \leqslant \underline{\kappa}_F \leqslant \overline{\kappa}_F \leqslant \underline{\kappa}_H < \infty$ holds only when trade costs fall below a threshold: $\tau \leqslant \widehat{\tau}$. For sufficiently high trade costs ($\tau > \widehat{\tau}$), the linkage fixed cost thresholds satisfy $0 \leqslant \overline{\kappa}_F < \underline{\kappa}_F < \underline{\kappa}_H < \infty$, and multiple equilibria can be sustained over a non-empty interval of κ. The trade cost threshold $\widehat{\tau}$ solely depends on exogenous parameters α, β, and σ.*

Figure 4 plots the linkage fixed cost thresholds $\underline{\kappa}_H$, $\overline{\kappa}_F$, and $\underline{\kappa}_F$ as functions of trade costs τ for the case of $\alpha(1-\sigma) \leqslant 1/2$. Global production networks are NP if $\underline{\kappa}_H < \kappa < \infty$, NC if $\overline{\kappa}_F \leqslant \kappa \leqslant \underline{\kappa}_H$, IP if $\underline{\kappa}_F < \kappa < \overline{\kappa}_F$, and IC if $0 \leqslant \kappa \leqslant \underline{\kappa}_F$. A key insight from Lemma 3 is that the level of linkage fixed costs κ determines whether alleviating trade barriers can affect the structure of global production networks. For countries with $\kappa \geqslant \underline{\kappa}_H$, trade integration does not trigger adjustments in firm linkages, as neither the domestic network density μ_H nor the threshold $\underline{\kappa}_H$ depends on trade costs τ. However, when κ is below the threshold $\underline{\kappa}_H$, trade cost reductions may prompt firms to form supplier-customer relationships with their peers abroad, since the thresholds $\overline{\kappa}_F$ and $\underline{\kappa}_F$ are both downward-sloping functions of trade costs τ. In this case, trade integration can induce global production networks to undergo structural transformation from NC to IP or even IC, during which trade in intermediate inputs appears endogenously. To fully gauge the welfare impact of trade, we also need to consider the response of aggregate firm mass N, which is the subject of the next lemma.

Lemma 4 *If global production networks are NP, NC, or IC, $dN/d\tau = 0$. If global production networks are IP, $dN/d\tau > 0$. Therefore, as countries connect more densely with each other through intermediate input trade, the world economy becomes more concentrated, with firm size r rising and firm mass N shrinking.*

Smaller trade costs τ make foreign markets more accessible. When international linkages are too costly to set up, only consumers can benefit from the enhanced accessibility of foreign goods, while firms face the same pool of domestic input varieties and hence maintain the same production technology. Cross-country firm relationships are a necessary condition for producers to also reap the benefits of larger markets. By adopting a wider range of intermediate inputs through connections to foreign suppliers, firms are able to lower the marginal cost of production and thereby grow in size. Even though concentration increases with trade liberalization when global production networks are IP, the gains from intermediate input diversity more than compensate for the losses of consumption variety, and therefore the market size effect still exceeds that implied by the Krugman model. In terms of welfare gains from trade, Proposition 3 below formally contrasts my model of endogenous firm linkages with the Krugman benchmark.

Proposition 3 *If global production networks are NP or NC in equilibrium, the elasticity of welfare with respect to market size is $d\ln W/d\ln\left(1+\tau^{\frac{\beta}{\beta-1}}\right) = (1-\beta)/\beta$. If global production networks are IP or IC in equilibrium, $d\ln W/d\ln\left(1+\tau^{\frac{\beta}{\beta-1}}\right) > (1-\beta)/\beta$. Therefore, the open economy equi-*

librium implies the same market size effect on welfare as in Krugman (1980) when international linkages are absent. However, the degree of aggregate increasing returns to scale is strictly higher than that in Krugman (1980) as long as trade in intermediate inputs is positive (i.e. international linkages are present).

Having shown that the model implies larger welfare gains from trade than the Krugman benchmark, I take a further step by asking if these larger gains are in fact "new gains". Arkolakis et al. (2012) point out two sufficient statistics for the welfare impact of globalization, namely the trade elasticity and the share of domestic goods in aggregate expenditure, which apply to an important class of quantitative trade models. I now examine whether endogenous production networks indeed provide an extra margin of adjustment missing in models with fixed linkages, such as those considered by Arkolakis et al. (2012). The next proposition expresses the welfare gains from trade implied by my model in terms of observable empirical moments and contrasts this expression with the ACR formula.

Proposition 4 *With endogenous production networks, the following empirical moments constitute sufficient statistics for the welfare impact of globalization (conditional on the value of σ): trade elasticity ε, the aggregate domestic expenditure share Λ, the domestic expenditure share for intermediate inputs Λ_m, the share of intermediate inputs in gross output ι, the density ratio of international networks to domestic networks $\widehat{\mu} \equiv \mu_F / \mu_H$, and total firm mass N. Specifically, the response of welfare with respect to a small change in variable trade costs τ is given by*

$$d \ln W = \left(\frac{1+\chi_1}{\varepsilon} \right) (d \ln \Lambda + d \ln \chi_2) \tag{1.2.6}$$

with

$$\chi_1 = \frac{\iota \Lambda_m (1-\Lambda_m)(1-\widehat{\mu})}{1-\iota \Lambda_m (1-\widehat{\mu})} - \frac{\iota \Lambda_m (1-\Lambda_m)\left(\widehat{\mu}^{-1}-1\right)}{1+\iota (1-\Lambda_m)(\widehat{\mu}^{-1}-1)},$$

$$\chi_2 = \frac{\Lambda_m^{\frac{1-\sigma}{\sigma}} N^{-\frac{1}{\sigma}}}{1+\iota (1-\Lambda_m)(\widehat{\mu}^{-1}-1)}.$$

In general, the above expression is different from the ACR formula ($d \ln W = d \ln \Lambda / \varepsilon$), because endogenous production networks violate two macro-level restrictions required for the ACR formula. First, as shown by Lemma 4, when trade cost reductions prompt firms to expand foreign relationships, the reallocation of labor towards linkage formation crowds out firm creation, leading to a decrease in the total number of varieties and hence violating restriction R2 in Arkolakis et al. (2012). Second, globalization affects intermediate inputs trade not only on the intensive margin but also on the extensive margin through the creation of international linkages. Therefore, trade elasticity is no longer constant, violating restriction R3 in the ACR framework. Since input-output linkages are endogenous variables in my model and therefore responsive to trade shocks, the set of sufficient statistics for welfare changes naturally expands to include production-side moments: ι captures the contribution of intermediate inputs in aggregate output, while Λ_m and $\widehat{\mu} \equiv \mu_F/\mu_H$ reflect how globally integrated the markets for intermediate inputs are. When the intermediate inputs share $\iota = \alpha(1-\sigma)$ is zero (i.e., $\sigma \to 1$), equation (1.2.6) converges to the ACR formula, which is expected as the Krugman model satisfies the ACR restrictions. In the limiting case of no international linkage ($\mu_F \to 0$ and hence $\widehat{\mu} \to 0$), only domestic inputs are used in production ($\Lambda_m = 1$) and equation (1.2.6) reduces to $d\ln W = d\ln\Lambda/[\varepsilon(1-\iota)]$. At the other extreme where production networks are globally complete ($\mu_F \to 1$ and hence $\widehat{\mu} \to 1$), the domestic expenditure share for intermediate inputs coincides with that for consumption goods ($\Lambda_m = \Lambda_c = \Lambda$) and equation (1.2.6) reduces to $d\ln W = d\ln\Lambda/(\varepsilon\sigma)$. In both limiting cases, the ACR formula understates the true welfare gains from trade.

1.3 Heterogeneous firm model

In this section, I generalize the model by introducing firm heterogeneity. Without repeating the maintained assumptions, Section 3.1 only describes how firm heterogeneity affects the choices made by households and producers. I then define and solve for the heterogeneous firm equilibrium in Section 3.2. In Section 3.3, I use stylized examples to illustrate key differences in the impact of trade between the heterogeneous and the homogeneous firm models, while referring those interested in general comparative statics to Appendix A.

1.3.1 Setup

The heterogeneous firm model shares the same setup as previously described, except that productivity now varies across industries. Specifically, industry TFP θ (common to all firms within the industry) follows an exogenous nondegenerate distribution with probability density function $g(\theta)$ over positive support $\Theta \equiv [\underline{\theta}, \overline{\theta}]$, where $\overline{\theta}$ can potentially be infinite. From now on, I switch the industry index from $i \in [0,1]$ to $\theta \in \Theta$ via the mapping implied by $g(\theta)$. Firms within an industry are still symmetric, and therefore I index firms by the productivity levels θ of their respective industries.

1.3.1.1 HOUSEHOLDS

In this generalized setup, consumption demand by households varies across firm productivity levels θ:

$$X_H(\theta) = X\left[\frac{P}{p(\theta)}\right]^{\frac{1}{1-\alpha}}\left[\left(1+\tau^{\frac{\beta}{\beta-1}}\right)N(\theta)\right]^{\frac{\beta-\alpha}{\beta(\alpha-1)}}, \qquad X_F(\theta) = \tau^{\frac{1}{\beta-1}}X_H(\theta),$$

where the consumer price index now has the following expression:

$$P \equiv \left\{\int_\Theta p(\theta)^{\frac{\alpha}{\alpha-1}}\left[\left(1+\tau^{\frac{\beta}{\beta-1}}\right)N(\theta)\right]^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)} g(\theta)\,d\theta\right\}^{\frac{\alpha-1}{\alpha}}.$$

1.3.1.2 FIRMS

With firm heterogeneity, demand for intermediate inputs depends on the productivity level of both the supplier and the customer. Hereafter, I refer to a firm (or an industry) with TFP level θ as a type-θ firm (industry). A type-θ firm sources intermediate inputs $\{x_H(\theta,\theta')\}_{\theta'\in\Theta}$ from its domestic suppliers and $\{x_F(\theta,\theta')\}_{\theta'\in\Theta}$ from its foreign suppliers. These factor demands are functions of both the customer productivity θ and the supplier productivity θ':

$$x_H(\theta,\theta') = m(\theta)\left[\frac{P^S(\theta)}{p(\theta')}\right]^{\frac{1}{1-\alpha}}\left\{\left[\mu_H(\theta,\theta')+\tau^{\frac{\beta}{\beta-1}}\mu_F(\theta,\theta')\right]N(\theta')\right\}^{\frac{1}{\beta}\left(\frac{\alpha-\beta}{1-\alpha}\right)},$$
$$x_F(\theta,\theta') = \tau^{\frac{1}{\beta-1}}x_H(\theta,\theta'),$$

where the producer price index is now specific to the customer firm:

$$P^S(\theta) \equiv \left(\int_\Theta\left\{\left[\mu_H(\theta,\theta')+\tau^{\frac{\beta}{\beta-1}}\mu_F(\theta,\theta')\right]N(\theta')\right\}^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)} p(\theta')^{\frac{\alpha}{\alpha-1}} g(\theta')\,d\theta'\right)^{\frac{\alpha-1}{\alpha}}. \tag{1.3.1}$$

The linkage densities $\mu_H(\theta,\theta') \in [0,1]$ and $\mu_F(\theta,\theta') \in [0,1]$ also become relationship-specific in the context of heterogeneous firms.[8] Specifically,

[8] Throughout the heterogeneous firm model, wherever a variable is dependent on the types of both the buyer and the seller, the first index always refers to the customer type while the second index always refers to the supplier type.

$\mu_H(\theta,\theta')N(\theta')$ and $\mu_F(\theta,\theta')N(\theta')$ give the number of domestic and foreign suppliers that firm θ has in any type-θ' industry. The linkage formation problem for firm θ is to choose, in every industry $\theta' \in \Theta$, a fraction $\mu_H(\theta',\theta) \in [0,1]$ of domestic firms and a fraction $\mu_F(\theta',\theta) \in [0,1]$ of foreign firms to sell to. As a result, in any type-θ' industry, firm θ acquires a mass of $\mu_H(\theta',\theta)N(\theta')$ customers at home and a mass of $\mu_F(\theta',\theta)N(\theta')$ customers from abroad and therefore faces the following demand function:

$$q(\theta) = X_H(\theta) + \tau X_F(\theta) + \int_\Theta x_H\left(\theta',\theta\right)\mu_H\left(\theta',\theta\right)N\left(\theta'\right)g\left(\theta'\right)d\theta' \\ + \int_\Theta \tau x_F\left(\theta',\theta\right)\mu_F\left(\theta',\theta\right)N\left(\theta'\right)g\left(\theta'\right)d\theta'. \tag{1.3.2}$$

The first order conditions for the linkage choice variables $\{\mu_H(\theta',\theta)\}_{\theta'\in\Theta}$, $\{\mu_F(\theta',\theta)\}_{\theta'\in\Theta}$ balance the same tradeoff as in the homogeneous firm model, except that now returns to linkages vary across firm pairs. For firm θ, the benefit of having an additional type-θ' customer is $(1-\alpha)x_H(\theta',\theta)p(\theta)$ if the relationship is domestic and $(1-\alpha)\tau x_F(\theta',\theta)p(\theta)$ if the relationship is international. In equilibrium, the two linkage density mappings $\mu_H,\mu_F : \Theta\times\Theta \to [0,1]$ jointly define the structure of global production networks: the mapping μ_H describes the linkage distribution within borders and the mapping μ_F characterizes the distribution of international linkages.

1.3.2 EQUILIBRIUM DEFINITION AND SOLUTION

As before, I state the firm entry assumption to close the model. Firms can choose which industry to enter, knowing that they will adopt the productivity level of the industry upon entry. To enter an industry with TFP level θ, a firm must pay a cost of $\nu\theta$ labor units. The heterogeneous firm equilibrium is defined analogously to that in Section 2.2, with equilibrium variables replaced by equilibrium mappings whose domains are the productivity space Θ.

Definition (heterogeneous firm equilibrium) *Given a set of parameters* $\{\alpha,\beta,\sigma,L,\kappa,\nu,\tau\}$ *and a probability distribution function of industry productivities* $g:\Theta\to[0,\infty)$, *a symmetric equilibrium of the heterogeneous firm economy includes mappings of variety prices* $p:\Theta\to\mathbb{R}^+$, *firm revenues* $r = p\circ q : \Theta\to\mathbb{R}^+$, *linkage densities* $\mu_H,\mu_F:\Theta\times\Theta\to[0,1]$, *and firm masses* $N:\Theta\to\mathbb{R}^+$ *such that the following conditions hold:*

1. *Households maximize utility according to Section 3.1.1.*

2. *Firms maximize profit according to Section 3.1.2.*

3. *There is free entry of firms in all industries: for all $\theta \in \Theta$,*

$$(1-\alpha)\,r(\theta) - \kappa\theta \int_{\Theta} \left[\mu_H\left(\theta',\theta\right) + \mu_F\left(\theta',\theta\right)\right] N\left(\theta'\right) g\left(\theta'\right) d\theta' = \nu\theta.$$

4. *Goods market clears according to equation (1.3.2) for all firms $\theta \in \Theta$.*

5. *Labor markets clear:*

$$\begin{aligned} L = & \int_{\Theta} l(\theta)\,N(\theta)\,g(\theta)\,d\theta \\ & + \int_{\Theta} \kappa\theta \int_{\Theta} \left[\mu_H\left(\theta',\theta\right) + \mu_F\left(\theta',\theta\right)\right] N\left(\theta'\right) g\left(\theta'\right) d\theta' N(\theta)\,g(\theta)\,d\theta \\ & + \int_{\Theta} \nu\theta N(\theta)\,g(\theta)\,d\theta. \end{aligned}$$

Even though the distribution of industry productivity $g(\theta)$ is exogenous, the distribution of firm productivity is endogenous, because the free entry condition determines the number of firms $N(\theta)$ operating in each industry. To facilitate equilibrium characterization, I define aggregate productivity A as an unweighted sum of the productivity levels of all firms in the economy:

$$A \equiv \int_{\Theta} \theta N(\theta)\,g(\theta)\,d\theta.$$

The endogenous nature of aggregate productivity A arises from $N(\theta)$, the equilibrium distribution of firms across industries. Thus, A can also be thought of as a weighted sum of industry TFP levels, with the weights endogenously given by the mass of firms.

In the presence of firm heterogeneity, the payoff to a supplier-customer relationship depends on the productivity of both parties. When a firm decides which potential customers to sell to, it naturally favors those with higher demand since the fixed cost of establishing a downstream link does not vary across buyer types. I conjecture and later verify that, for any seller type θ, intermediate input demand $x_H\left(\theta',\theta\right)$ and $x_F\left(\theta',\theta\right)$ are increasing functions of the buyer's productivity level θ'. As a result, the supplier's optimal linkage formation gives rise to a set of productivity cutoffs $b_H(\theta)$, $\underline{b}_F(\theta)$, and $\overline{b}_F(\theta)$ for selecting customers:

$$\mu_H\left(\theta',\theta\right) = \begin{cases} \in (0,1) & \text{if } \underline{\theta} \leqslant \theta' < b_H(\theta) \\ 1 & \text{if } b_H(\theta) \leqslant \theta' \leqslant \overline{\theta} \end{cases};$$

$$\mu_F\left(\theta',\theta\right) = \begin{cases} 0 & \text{if } \underline{\theta} \leqslant \theta' \leqslant \underline{b}_F(\theta) \\ \in (0,1) & \text{if } \underline{b}_F(\theta) < \theta' < \overline{b}_F(\theta) \\ 1 & \text{if } \overline{b}_F(\theta) \leqslant \theta' \leqslant \overline{\theta} \end{cases}.$$

Conditional on buyer type θ', any supplier would find domestic relationships more profitable than international ones, so long as trade remains costly ($\tau > 1$). Therefore, the customer-selection productivity cutoffs must satisfy

$$b_H(\theta) < \underline{b}_F(\theta) < \overline{b}_F(\theta)$$

which means that a firm would never sell to a foreign type-θ' firm before it has exhausted all domestic type-θ' customers. The next lemma characterizes these productivity cutoffs in equilibrium.

Lemma 5 *The cutoff productivity levels for selecting customers are the same for all suppliers. For all $\theta \in \Theta$,*

$$\begin{aligned} b_H(\theta) &= b_H \equiv \max\left\{\underline{\theta}, \kappa[1-\alpha(1-\sigma)]A^2/[\alpha(1-\alpha)(1-\sigma)L]\right\}, \\ \underline{b}_F(\theta) &= \underline{b}_F \equiv \tau^{\frac{\beta}{1-\beta}} b_H, \\ \overline{b}_F(\theta) &= \overline{b}_F \equiv \left(1+\tau^{\frac{\beta}{1-\beta}}\right) b_H. \end{aligned}$$

Lemma 5 allows me to group firms according to how their productivity levels θ compare with the customer-selection cutoffs. If $\theta \in [\underline{\theta}, b_H)$, the firm has no foreign suppliers and is reached only by a subset of domestic firms; if $\theta \in [b_H, \underline{b}_F]$, the firm has upstream links with all other producers in the country but with none abroad; if $\theta \in (\underline{b}_F, \overline{b}_F)$, the firm's suppliers include all of its domestic peers plus a subset of foreign firms; if $\theta \in [\overline{b}_F, \overline{\theta}]$, the firm sources intermediate inputs from all other firms, home and abroad. Since trade costs τ alone entirely determine the relative positions of b_H, $\underline{b}_F$, and $\overline{b}_F$, solving for one of the three cutoffs is sufficient for carrying out the partition. I thus choose the cutoff b_H as the unknown to be solved for. The next lemma establishes that all firm level outcomes depend on equilibrium aggregate variables only through the customer-selection cutoff b_H and aggregate productivity A.

Lemma 6 *Equilibrium firm size is increasing in the firm productivity level θ:*

$$r(\theta) = \frac{\theta L}{[1-\alpha(1-\sigma)]A}.$$

Equilibrium variety price is decreasing in the firm productivity level θ:

$$p(\theta) = \begin{cases} \theta^{-\left[1+\left(\frac{1-\beta}{\beta}\right)(1-\sigma)\right]} \Psi_P(b_H, A) & \text{if } \underline{\theta} \leqslant \theta < b_H \\ \theta^{-1}\left\{\frac{\kappa[1-\alpha(1-\sigma)]A^2}{\alpha(1-\alpha)(1-\sigma)L}\right\}^{\left(\frac{\beta-1}{\beta}\right)(1-\sigma)} \Psi_P(b_H, A) & \text{if } b_H \leqslant \theta \leqslant \tau^{\frac{\beta}{1-\beta}} b_H \\ \theta^{-\left[1+\left(\frac{1-\beta}{\beta}\right)(1-\sigma)\right]} \tau^{1-\sigma} \Psi_P(b_H, A) & \text{if } \tau^{\frac{\beta}{1-\beta}} b_H < \theta < \left(1+\tau^{\frac{\beta}{1-\beta}}\right) b_H \\ \theta^{-1}\left\{\left(1+\tau^{\frac{\beta}{\beta-1}}\right)\frac{\kappa[1-\alpha(1-\sigma)]A^2}{\alpha(1-\alpha)(1-\sigma)L}\right\}^{\left(\frac{\beta-1}{\beta}\right)(1-\sigma)} \Psi_P(b_H, A) & \text{if } \left(1+\tau^{\frac{\beta}{1-\beta}}\right) b_H \leqslant \theta \leqslant \overline{\theta} \end{cases}$$

The distribution of firms across industries is given by

$$N(\theta) = \begin{cases} \theta^{\frac{(2\alpha-1)\beta}{\beta-\alpha} + \frac{\alpha(1-\beta)}{\beta-\alpha}(1-\sigma)} \Psi_N(b_H, A) & \text{if } \underline{\theta} \leqslant \theta < b_H \\ \theta^{\frac{(2\alpha-1)\beta}{\beta-\alpha}} \left\{ \frac{\kappa[1-\alpha(1-\sigma)]A^2}{\alpha(1-\alpha)(1-\sigma)L} \right\}^{\frac{\alpha(1-\beta)}{\beta-\alpha}(1-\sigma)} \Psi_N(b_H, A) & \text{if } b_H \leqslant \theta \leqslant \tau^{\frac{\beta}{1-\beta}} b_H \\ \theta^{\frac{(2\alpha-1)\beta}{\beta-\alpha} + \frac{\alpha(1-\beta)}{\beta-\alpha}(1-\sigma)} \tau^{\frac{\alpha\beta}{\alpha-\beta}(1-\sigma)} \Psi_N(b_H, A) & \text{if } \tau^{\frac{\beta}{1-\beta}} b_H < \theta < \left(1+\tau^{\frac{\beta}{1-\beta}}\right) b_H \\ \theta^{\frac{(2\alpha-1)\beta}{\beta-\alpha}} \left\{ \left(1+\tau^{\frac{\beta}{\beta-1}}\right) \frac{\kappa[1-\alpha(1-\sigma)]A^2}{\alpha(1-\alpha)(1-\sigma)L} \right\}^{\frac{\alpha(1-\beta)}{\beta-\alpha}(1-\sigma)} \Psi_N(b_H, A) & \text{if } \left(1+\tau^{\frac{\beta}{1-\beta}}\right) b_H \leqslant \theta \leqslant \overline{\theta} \end{cases}$$

where $\Psi_P(b_H, A)$ and $\Psi_N(b_H, A)$ are expressions common to all productivity types, and are dependent on two endogenous variables only: the customer-selection cutoff b_H and aggregate productivity A (the full expressions are given in the proof).

Global production networks in the heterogeneous firm model are defined by $\mu_H, \mu_F : \Theta \times \Theta \to [0,1]$, two mappings from the space of productivity type pairs to the unit interval. In equilibrium, $\mu_H(\theta', \theta)$ gives the probability that a type-θ supplier is connected with a type-θ' customer of the same country, and $\mu_F(\theta', \theta)$ gives such a linking probability across countries. I show in the next lemma that the productivity cutoff b_H for customer selection is all one needs to fully characterize both domestic and international production networks.

Lemma 7 *Domestic production networks are characterized by the linking probability function:*

$$\mu_H\left(\theta', \theta\right) = \begin{cases} \frac{\alpha(1-\alpha)(1-\sigma)L\theta'}{\kappa[1-\alpha(1-\sigma)]A^2} & \text{if } \underline{\theta} \leqslant \theta' < b_H \\ 1 & \text{if } b_H \leqslant \theta' \leqslant \overline{\theta} \end{cases}.$$

International production networks are characterized by the linking probability function:

$$\mu_F\left(\theta', \theta\right) = \begin{cases} 0 & \text{if } \underline{\theta} \leqslant \theta' \leqslant \tau^{\frac{\beta}{1-\beta}} b_H \\ \frac{\alpha(1-\alpha)(1-\sigma)L\theta'}{\kappa[1-\alpha(1-\sigma)]A^2} - \tau^{\frac{\beta}{1-\beta}} & \text{if } \tau^{\frac{\beta}{1-\beta}} b_H < \theta' < \left(1+\tau^{\frac{\beta}{1-\beta}}\right) b_H \\ 1 & \text{if } \left(1+\tau^{\frac{\beta}{1-\beta}}\right) b_H \leqslant \theta' \leqslant \overline{\theta} \end{cases}$$

where θ is the productivity level of the supplier and θ' that of the customer.

Lemma 6 and 7 jointly paint a picture of the equilibrium at the firm level. More productive firms operate on a larger scale, demand more intermediate inputs, attract a wider range of suppliers, enjoy superior production technology thanks to higher input diversity, and charge lower prices accordingly. The substitutability of varieties across industries (as measured by the parameter α) governs how the concentration of firms $N(\theta)$ varies with industry TFP levels. If $0 < \alpha \leqslant \beta / [2\beta + (1-\beta)(1-\sigma)]$, products from different industries are not close substitutes, and firm concentration is weakly increasing in industry productivity. In this case, high-productivity industries host less albeit larger firms than low-productivity industries. If $1/2 \leqslant \alpha < 1$, firm concentration is weakly decreasing in industry productivity, and high-productivity industries feature not only larger firms but also more firms. For the intermediate range of cross-industry substitutability such that $\beta / [2\beta + (1-\beta)(1-\sigma)] < \alpha < 1/2$, the relationship between firm concentration and industry type is no longer monotonic.

Finally, I solve for the two remaining unknowns: the productivity cutoff b_H for selecting customers and aggregate productivity A. By the two previous lemmas, one can express a firm's total number of downstream links $\int_\Theta [\mu_H(\theta', \theta) + \mu_F(\theta', \theta)] N(\theta') g(\theta') d\theta' \equiv z(b_H, A)$ in terms of the two aggregate unknowns (the full expression is given in the appendix). Thus, the free entry (FE) condition together with the customer selection (CS) condition given by Lemma 5 provide two separate relationships between the customer-selection cutoff b_H and aggregate productivity A, whose behavior in the (b_H, A) space jointly determines equilibrium existence and uniqueness:

$$\frac{(1-\alpha)L}{[1-\alpha(1-\sigma)]A} - \kappa z(b_H, A) = \nu \qquad \text{(FE)}$$

$$b_H = \max\left\{\underline{\theta}, \frac{\kappa[1-\alpha(1-\sigma)]A^2}{\alpha(1-\alpha)(1-\sigma)L}\right\} \qquad \text{(CS)}$$

So far I have not imposed any restriction on the exogenous distribution of industry types $g(\theta)$. To sharpen the predictions about equilibrium existence and uniqueness, I adopt the standard Pareto distribution for $g(\theta)$ with support $[\underline{\theta}, \infty)$, i.e. $\overline{\theta} \to \infty$. The following proposition details the parametric condition under which the heterogeneous firm equilibrium exists and is unique.

Proposition 5 *Suppose that the exogenous distribution of industry productivity is Pareto with shape parameter ζ such that $\zeta > \alpha(2\beta - 1)/(\beta - \alpha)$ and $\zeta \neq \alpha[\beta - \sigma(1-\beta)]/(\beta - \alpha)$, then the heterogeneous firm equilibrium exists. For linkage fixed costs above a certain threshold ($\kappa > \overline{\kappa}$), the customer-selection cutoff b_H lies in the interior of the productivity support ($\underline{\theta} < b_H \leqslant \infty$) and the equilibrium is unique; for sufficiently low linkage* *fixed costs ($\kappa \leqslant \overline{\kappa}$), the customer-selection cutoff b_H hits the lower bound of the productivity support ($b_H = \underline{\theta}$) and the equilibrium is unique in the limiting case of $\tau \to \infty$. The threshold level $\overline{\kappa}$ depends only on exogenous parameters.*

Restrictions on the shape parameter ζ ensure the convergence of various integrals, so that equilibrium prices and quantities are finite. As with the homogeneous firm model, I measure welfare by real wage $W \equiv 1/P$. In the presence of firm heterogeneity, aggregate variables alone are no longer sufficient for capturing the welfare impact of trade liberalization. Instead, the response of welfare to trade depends on the entire distribution of the following firm level outcomes: the domestic expenditure share for intermediate inputs $\Lambda_m(\theta) \equiv \mu_H(\theta,\theta') / \left[\mu_H(\theta,\theta') + \tau^{\frac{\beta}{\beta-1}} \mu_F(\theta,\theta')\right]$, the ratio of foreign suppliers to the domestic ones $\widehat{\mu}(\theta) \equiv \mu_F(\theta,\theta') / \mu_H(\theta,\theta')$, and firm sales $r(\theta)$. In the next proposition, I derive a formula for welfare gains from trade in the spirit of Arkolakis et al. (2012), showing that the micro structure of the model remains relevant for welfare analysis.

Proposition 6 *In the heterogeneous firm model, the response of welfare with respect to a small change in the variable trade cost τ is given by*

$$d\ln W = \left(\frac{1+\chi_1}{\varepsilon}\right)(d\ln\Lambda + d\ln\chi_2)$$

with

$$\chi_1 = \frac{\int_\Theta \iota\Lambda_m(\theta)[1-\Lambda_m(\theta)][1-\widehat{\mu}(\theta)]\, r(\theta)\widehat{g}(\theta)\,d\theta}{\overline{r} - \int_\Theta \iota\Lambda_m(\theta)[1-\widehat{\mu}(\theta)]\, r(\theta)\widehat{g}(\theta)\,d\theta} - \frac{\int_\Theta \iota\Lambda_m(\theta)[1-\Lambda_m(\theta)]\left[\widehat{\mu}(\theta)^{-1}-1\right] r(\theta)\widehat{g}(\theta)\,d\theta}{\overline{r} + \int_\Theta \iota[1-\Lambda_m(\theta)]\left[\widehat{\mu}(\theta)^{-1}-1\right] r(\theta)\widehat{g}(\theta)\,d\theta}$$

$$\chi_2 = \frac{\overline{r}\left\{\int_\Theta \left[\theta^{\frac{\beta}{1-\beta}} \Lambda_m(\theta)^{\sigma-1} \mu_H(\theta,\theta')^{1-\sigma} N(\theta)\right]^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)} g(\theta)\,d\theta\right\}^{-\frac{1}{\sigma}\frac{\beta}{\alpha}\left(\frac{1-\alpha}{1-\beta}\right)}}{\overline{r} + \int \iota[1-\Lambda_m(\theta)]\left[\widehat{\mu}(\theta)^{-1}-1\right] r(\theta)\widehat{g}(\theta)\,d\theta}$$

where $\widehat{g}(\theta) \equiv N(\theta)g(\theta) / \int_\Theta N(\theta)g(\theta)\,d\theta$ is the endogenous distribution of firm productivity and $\overline{r}$ the average firm sales.

The formula above reduces to that in the homogeneous firm case once the productivity differential is removed. Proposition 6 suggests that trade liberalization can affect welfare through the production side, on both the firm and the aggregate level. On the firm level, trade induces linkage redistribution, thereby changing a firm's set of suppliers and ultimately its cost of production. These firm level technological changes are reflected by $\Lambda_m(\theta)$, $\widehat{\mu}(\theta)$, and $r(\theta)$. On the aggregate level, trade reshapes the distribution of firm productivity as some industries expand relative to the others. Aggregate production technology, defined by the firm productivity distribution $\widehat{g}(\theta)$, therefore is no longer fixed as in the homogeneous firm model, but responsive to changes in market size. I elaborate on how trade transforms the production side of the economy in the next section, emphasizing on the differences between the heterogeneous and the homogeneous firm models.

1.3.3 THE ROLE OF FIRM HETEROGENEITY

This section highlights the role played by firm heterogeneity in determining the impact of trade integration. I construct two examples, one with small firm heterogeneity and the other featuring large productivity dispersion, and characterize these two stylized worlds respectively in autarky ($\tau \to \infty$) and in the integrated economy ($\tau \to 1$). In both examples, I consider the simplest form of firm heterogeneity by assuming that the productivity support Θ contains only two values: $\Theta = \{\underline{\theta}, \overline{\theta}\}$. Thus, the unit mass of industries fall into two categories: those with a low productivity level $\underline{\theta}$ and those with a higher one $\overline{\theta} \in (\underline{\theta}, \infty)$. A key variable that governs the responses of production networks to trade integration is the following ratio of two uncentered moments of θ:

$$M_\theta \equiv \frac{\int_\Theta \theta^{\frac{\alpha(2\beta-1)}{\beta-\alpha}} g(\theta)\, d\theta}{\underline{\theta} \int_\Theta \theta^{\frac{\beta(2\alpha-1)}{\beta-\alpha}} g(\theta)\, d\theta}.$$

Over the parameter space where $\beta > 1/2$, the moment ratio M_θ increases in the productivity advantage $\overline{\theta}/\underline{\theta}$ enjoyed by the high type firms over their low type peers, conditional on the probability mass function $g(\theta)$. As productivity dispersion vanishes ($\overline{\theta} \to \underline{\theta}$), M_θ approaches its minimum at 1. Thus, M_θ can be considered as a measure of firm heterogeneity.

EXAMPLE 1: SMALL FIRM HETEROGENEITY

Suppose that firm heterogeneity is small such that $1 < M_\theta < \alpha(1-\sigma)\left(1-\sqrt{\Phi}\right)^{-1}$, where Φ is a constant depending on exogenous parameters only.[9] Equi-

[9] Specifically, $\Phi \equiv \alpha(1-\sigma)[1-\alpha(1-\sigma)]\underline{\theta}\nu^2/[2(1-\alpha)\kappa L]$. In order for this interval to be non-empty, population L needs to be sufficiently large: $L > \alpha(1-\sigma)\underline{\theta}\nu^2/\{2(1-\alpha)[1-\alpha(1-\sigma)]\kappa\}$. Both examples are constructed as follows: first, I conjecture the network structure in equilibrium; next, I use the free entry condition as

librium production networks in this case are complete within borders in autarky and globally complete in the free trade world. On the firm level, trade integration leads to a universal improvement of production technology, as firms of both types gain access to foreign inputs and therefore are able to lower production costs. Consequently, prices fall across the board: $p(\theta)|_{\tau=1} < p(\theta)|_{\tau=\infty}$ for all $\theta \in \{\underline{\theta}, \overline{\theta}\}$. On the aggregate level, the firm mass ratio $N(\overline{\theta})/N(\underline{\theta})$ remains unchanged from autarky to full integration, which implies that the average productivity $A/\int_{\Theta} N(\theta)\, g(\theta)\, d\theta$ of the economy also stays the same. In this case, the welfare impact of trade integration is unambiguously positive, since all firms benefit from the structural transformation of production networks.

EXAMPLE 2: LARGE FIRM HETEROGENEITY

I now increase firm heterogeneity so that $\alpha(1-\sigma)\left(1-\sqrt{\Phi}\right)^{-1} < M_\theta < \alpha(1-\sigma)\left(1-\sqrt{2\Phi}\right)^{-1}$ and $\overline{\theta} > [1-\alpha(1-\sigma)]^2\,\underline{\theta}/\Phi$. The autarky network structure is still domestically complete just as in Example 1. Unlike the previous case, global production networks in the integrated economy are no longer complete. Specifically, firms with low productivity $\underline{\theta}$ are connected to only a subset of all suppliers in the world, even though their high-productivity peers can source inputs from all producers: $\mu_H(\underline{\theta}, \theta)|_{\tau=1} \in (0,1)$ and $\mu_H(\overline{\theta}, \theta)|_{\tau=1} = 1$ for $\theta \in \{\underline{\theta}, \overline{\theta}\}$. On the firm level, trade integration grants more technological advantage (in terms of input diversity) to high-productivity firms than the less productive ones, as reflected in the fallen price ratio $p(\overline{\theta})/p(\underline{\theta})$ of high-type goods to the low-type ones. Such biased technological changes make entry into high-productivity industries relatively more attractive, raising the firm mass ratio $N(\overline{\theta})/N(\underline{\theta})$ of high-type industries to the low-type ones. As the result of firm mass redistribution, trade integration raises the average productivity of the economy, thus affecting production technology not only on the firm level but also on the aggregate level. In this case, trade integration has ambiguous welfare consequences, because the structural transformation of production networks could result in low-productivity firms losing suppliers in absolute terms.

Comparing the above two examples leads one to the following observations. Small firm heterogeneity can give rise to homogeneous network adjustment pattern, as in Example 1. Since firms of different types have the same amount of input-output linkages, the technological consequences of trade integration are essentially identical to those in the homogeneous firm

well as the first order conditions of the firm linkage formation problem to establish the range of M_θ over which the conjectured network structure can be sustained as an equilibrium.

model: prices of individual goods decline, whereas average productivity stays constant. On the other hand, sufficiently large firm heterogeneity results in heterogeneous linkage responses, as in Example 2. Trade integration in this case generates biased technological change at the firm level, favoring the more productive ones. Free entry then leads to a redistribution of firm mass towards high-productivity industries, thereby reshaping the firm productivity distribution. Hence, trade integration also brings about technological changes on the aggregate level, which constitutes an extra margin of adjustment absent in the homogeneous firm model.

The intuition of these two examples carry over to the general setup, where the number of productivity types are infinite and the industry type distribution $g(\theta)$ is continuous. I present the general comparative statics with respect to market size in Appendix A, with $g(\theta)$ being Pareto as assumed in Proposition 5. As shown by Example 2, technological changes on the firm and the aggregate level do not necessarily affect welfare in the same direction. Thus, for a general set of parameters, it is difficult to theoretically pinpoint the overall production-side impact of trade liberalization on welfare. Instead, the rest of this paper brings the heterogeneous firm model to data and use the calibrated model to quantify welfare gains from trade.

1.4 Empirical application: trade in value added

Having explored the theoretical properties of the model, I now demonstrate how the framework can be applied in a quantitative setting. In particular, the model yields an analytic expression for the ratio of value-added to gross trade, even in the general case with firm heterogeneity. Therefore, the model provides a solid theoretical foundation for analyzing data on trade in value added. In Section 4.1, I summarize several testable model predictions, which I revisit later to check model fit. In Section 4.2, I describe the data source for empirical exercises and outline the strategy for calibrating the model.

1.4.1 Model predictions

I present two sets of model predictions that have clear empirical equivalents, which serve as the basis for checking model fit. First, I derive the analytic expression of the value added content of trade. Second, I characterize the network structure of the economy, focusing on the cross-sectional heterogeneity in an industry's role as input user and supplier. Both sets of outcome are observable in a world input-output table.

VALUE-ADDED EXPORTS

To derive the value added content of aggregate trade flows, I first define the direct requirement coefficients on the industry level:
$\omega_H^1(\theta',\theta) \equiv x_H(\theta',\theta)\mu_H(\theta',\theta)N(\theta)p(\theta)/r(\theta')$ gives the expenditure on a type-θ industry for every value unit of a type-θ' industry's output when the customer industry resides in the same country as the supplier industry; $\omega_F^1(\theta',\theta) \equiv \tau x_F(\theta',\theta)\mu_F(\theta',\theta)N(\theta)p(\theta)/r(\theta')$ then gives the correspondent expenditure share if the industry pair are located in different countries. In particular, the shares $\left\{\omega_H^1(\theta',\theta),\omega_F^1(\theta',\theta)\right\}_{\theta,\theta'\in\Theta}$ constitute the elements of the direct requirement input-output (I-O) tables. Specifically, for a world with two identical countries, the global input-output table is a two-by-two symmetric block matrix with the national I-O tables on the diagonal and the international I-O tables off the diagonal. Then, the domestic shares $\left\{\omega_H^1(\theta',\theta)\right\}_{\theta',\theta\in\Theta}$ fill the diagonal blocks of the world I-O matrix, while the foreign shares $\left\{\omega_F^1(\theta',\theta)\right\}_{\theta',\theta\in\Theta}$ are the elements of the off-diagonal blocks. As suggested by their names, the direct requirement coefficients capture the input transactions between customers and their immediate (i.e., first order) suppliers. One may also define indirect requirement coefficients to account for the contribution of higher-order suppliers. For example, the coefficients $\omega_H^2(\theta',\theta)$ and $\omega_F^2(\theta',\theta)$ defined below represent the spending per dollar of output by the downstream industries, indexed by θ', on their second-order suppliers (i.e., suppliers of suppliers) at home and abroad, indexed by θ:

$$\begin{aligned}
\omega_H^2\left(\theta',\theta\right) \equiv & \int_\Theta \omega_H^1\left(\theta',\theta''\right)\omega_H^1\left(\theta'',\theta\right)g\left(\theta''\right)d\theta'' \\
& + \int_\Theta \omega_F^1\left(\theta',\theta''\right)\omega_F^1\left(\theta'',\theta\right)g\left(\theta''\right)d\theta'' \\
\omega_F^2\left(\theta',\theta\right) \equiv & \int_\Theta \omega_H^1\left(\theta',\theta''\right)\omega_F^1\left(\theta'',\theta\right)g\left(\theta''\right)d\theta'' \\
& + \int_\Theta \omega_F^1\left(\theta',\theta''\right)\omega_H^1\left(\theta'',\theta\right)g\left(\theta''\right)d\theta''
\end{aligned}$$

In general, the n-th-order coefficients $\omega_H^n(\theta',\theta)$ and $\omega_F^n(\theta',\theta)$, whose expressions are relegated to the Appendix, reflect the indirect input requirements for industry pairs, within and across national borders, separated by n degrees along supply chains. Thus, we can compute an infinite series of indirect requirement coefficients to account for all higher-order input usage. Value-added exports at the industry level $\{VA(\theta)\}_{\theta\in\Theta}$ then follow the identity below, where $1-\alpha(1-\sigma)$ is the share of value added in gross output:

$$\frac{VA(\theta)}{1-\alpha(1-\sigma)} = \underbrace{\tau X_F(\theta)\, p(\theta) N(\theta)}_{\text{final good exports}} + \underbrace{\int_\Theta \sum_{n=1}^{\infty} \omega_H^n(\theta',\theta)\, \tau X_F(\theta')\, p(\theta') N(\theta') g(\theta')\, d\theta'}_{\text{indirect intermediate input exports}} + \underbrace{\int_\Theta \sum_{n=1}^{\infty} \omega_F^n(\theta',\theta)\, X_H(\theta')\, p(\theta') N(\theta') g(\theta')\, d\theta'}_{\text{direct intermediate input exports}}$$

This identity highlights three channels through which output of a type-θ industry can contribute to the consumption basket abroad. First, the industry exports its product as final goods directly to foreign households. Second, the industry supplies intermediate inputs to other domestic industries whose products are then exported as final goods. This second channel can be thought of as exporting intermediate inputs indirectly via domestic linkages. Third, the industry exports directly to their foreign counterparts, who then use the imported intermediate inputs to produce final goods consumed locally. Summing over all industries, I obtain the following expression for the aggregate *value-added exports* (VA):

$$\frac{VA}{L} \equiv \frac{1}{L}\int_\Theta VA(\theta)\, g(\theta)\, d\theta = 1-\Lambda_c + \frac{\iota(2\Lambda_c-1)(1-\Lambda_m)}{1-\iota[1-2(1-\Lambda_m)]} \tag{1.4.1}$$

where $\iota \equiv \alpha(1-\sigma)$ is the share of intermediate inputs in gross output, $\Lambda_c \equiv 1/\left(1+\tau^{\frac{\beta}{\beta-1}}\right)$ and $\Lambda_m \equiv \int_\Theta\int_\Theta \omega_H^1(\theta,\theta')\, g(\theta')\, d\theta' r(\theta) N(\theta) g(\theta)\, d\theta / \left[\iota \int_\Theta r(\theta) N(\theta) g(\theta)\, d\theta\right]$ are respectively the domestic expenditure shares for consumption goods and intermediate inputs (the derivation details for the expression of VA are given in the appendix). In contrast, *gross exports* (EX) simply comprise the value of all goods sold to foreign households and firms:

$$\frac{EX}{L} = 1-\Lambda_c + \frac{\iota}{1-\iota}(1-\Lambda_m) \tag{1.4.2}$$

Since gross exports trace sales rather than value added, double counting arises every time intermediate inputs cross borders. The more extensive international input-output linkages are, the larger the discrepancy between value-added exports and gross exports. Therefore, the ratio of value-added to gross exports ($VAX \equiv VA/EX$) has become a standard measure for the prominence of global supply chains (Johnson and Noguera 2017). My model predicts that the *VAX* ratio depends on three easily observable aggregate statistics: intermediate inputs share in gross output ι, domestic expenditure shares for consumption goods Λ_c and intermediate inputs Λ_m. In the absence of international linkages, value-added exports coincide with gross exports ($\lim_{\Lambda_m \to 1} VAX = 1$) as intermediate inputs never cross borders and therefore double counting never occurs. As the economy approaches complete integration ($\Lambda_c, \Lambda_m \to 1/2$), the ratio of value-added to gross exports tends to the share of value-added in gross output $1-\iota$ since production in this limiting case is fully shared between the two countries.

NETWORK CHARACTERISTICS

My model also predicts a pair of standard statistics for characterizing production networks: the weighted outdegree and the weighted indegree (Acemoglu et al., 2012). In equilibrium, firm heterogeneity gives rise to an uneven distribution of linkages among industries, which can be measured by the weighted degrees. The *weighted outdegree* of a type-θ industry is the sum of all direct requirement coefficients where the industry serves as a supplier: $d_{OUT}(\theta) \equiv \int_\Theta \left[\omega_H^1(\theta', \theta) + \omega_F^1(\theta', \theta)\right] g(\theta')\, d\theta'$. This statistic captures the strength of an industry's downstream linkages, ranging from 0 if the industry does not play the role of suppliers at all to 2 if it is the only source of intermediate inputs for every producer in the world. The model implies that the weighted outdegrees are positively correlated with industry revenue:

$$d_{OUT}(\theta) = \alpha(1-\sigma)\frac{R(\theta)}{R}$$

where $R(\theta) \equiv r(\theta) N(\theta)$ is the total sales of a type-θ industry and $R \equiv \int_\Theta r(\theta) N(\theta) g(\theta)\, d\theta$ is the national gross output. In other words, larger industries are also more important as suppliers in production networks. An industry's *weighted indegree* simply equals the share of intermediate inputs in gross output and reflects its reliance on upstream linkages: $d_{IN}(\theta) \equiv \int_\Theta \left[\omega_H^1(\theta, \theta') + \omega_F^1(\theta, \theta')\right] g(\theta')\, d\theta'$. Since the model keeps the Cobb-Douglas exponent constant across firms, all industries adopt the same share of intermediate inputs, and therefore the weighted indegrees are independent of industry characteristics:

$$d_{IN}(\theta) = \alpha(1-\sigma)$$

1.4.2 DATA AND CALIBRATION STRATEGY

I use the World Input-Output Database (WIOD) Release 2016 as the main data source for quantitative exercises (Timmer et al., 2015). This dataset covers 44 countries (28 EU members, 15 other major economies, and the rest of the world) and 56 industries corresponding roughly to the two-digit ISIC (revision 4) level, spanning 15 years from 2000 to 2014. To comply with the two-county setup of my model, I transform the dataset to feature the United States and the rest of the world (ROW) by merging the 43 non-US economies together and treating the trade within them as domestic transactions.

The heterogeneous firm model contains nine parameters to be calibrated. In the baseline calibration, I set the elasticity of substitution among varieties within the same industry $1/(1-\beta)=6$, which stands in the middle of the range reported by Anderson and van Wincoop (2004). Later on in the robustness checks, I re-calibrate the model by setting the within-industry elasticity at 4 and 8. For the between-industry elasticity of substitution $1/(1-\alpha)$ and the labor share σ in the Cobb-Douglas production function, I choose their values to strike a balance between the following two empirical facts. First, the elasticity of substitution across broadly defined industries should be close to one according to the recent estimates by Atalay (2017) and Oberfield and Raval (2014). Second, using WIOD, I find that the share of intermediate inputs in aggregate gross output $\iota \equiv \alpha(1-\sigma)$ for the United States ranges from 0.41 to 0.45 over the data period. Since both α and σ have to lie on the unit interval, these two facts create a tension on the value of α, with the first fact demanding α to be close to zero and the second fact imposing a lower bound on α at 0.41. To find a middle ground, I set $\alpha=0.5$ and $\sigma=0.1$, corresponding to a between-industry elasticity of substitution $1/(1-\alpha)=2$ and an aggregate intermediate input share $\iota=0.45$.

I then calibrate the two parameters that define the distribution of industry TFP levels: the lower bound of the productivity support $\underline{\theta}$ and the Pareto shape parameter ζ. Since varying the productivity minimum $\underline{\theta}$ has only level effects on welfare, I normalize $\underline{\theta}=1$. The Pareto shape parameter ζ determines the equilibrium distribution of firm sizes $l(\theta)$, whose right tail can be shown to follow a Pareto distribution with tail index $\beta(2\alpha-1)/(\beta-\alpha)$ $-\zeta$. Thus, I choose a value of ζ such that $\beta(2\alpha-1)/(\beta-\alpha)-\zeta=1.06$, which is consistent with the empirical evidence on U.S. firm size distribution (Axtell, 2001). Furthermore, I assume that the five aforementioned parameters associated with preferences and technology (α,β, σ, $\underline{\theta}$, and ζ) are constant over time.

I allow the four remaining parameters (various costs τ, κ, ν and population L) to be time-varying. I calibrate the iceberg transport cost τ to match the US import penetration ratio (one minus the share of aggregate expenditure on domestic goods Λ) for each year observed in WIOD. Linkage fixed costs κ and entry costs ν jointly govern the allocation of labor among three activities: goods production, linkage formation, and firm creation. Therefore, I calibrate κ and ν to match the total number of firms in the US with at least one employee and a minimum firm size $\min\{l(\theta)\}_{\theta\in\Theta} = 1$. Finally, I set population L to be the US employment, where both the employment and the firm number data are from Statistics of U.S. Businesses by the US Census Bureau. Figure 5 plots the evolution of the three calibrated cost parameters τ, κ, and ν. Except for the Great Trade Collapse during 2008-2009, variable trade costs τ display a steady downward trend. In contrast, linkage fixed costs κ and entry costs ν appear to be more volatile without an obvious time trend.

1.5 QUANTITATIVE RESULTS

This section reports the results of quantitative exercises. First, I assess model fit in Section 5.1. I then conduct a series of counterfactuals in Section 5.2 to answer two quantitative questions: (i) which exogenous force is most responsible for the observed time trend in welfare and the value added content of trade? (ii) Do endogenous linkages, as opposed to a fixed network structure, make a quantitative difference in predicting welfare gains from trade? Finally, in Section 5.3, I check whether the quantitative results are robust to alternative parameter values.

1.5.1 MODEL FIT

Guided by the theoretical predictions in Section 4.1, I compare the model-generated outcomes against data on two dimensions: (i) the value added content of gross trade flows; (ii) the relationship between an industry's weighted degrees in global production networks and its gross output.

Figure 6 plots the evolution of value-added trade, contrasting the model outcomes with the US pattern observed in WIOD.[10] *Model 1* corresponds to the full calibration of structural parameters as described in Section 4.2. *Model 2* computes the intermediate inputs share ι as well as the domestic expenditure shares Λ_c and Λ_m directly from data and then substitutes these three statistics into equations (1.4.1) and (1.4.2). Thus, comparing *Model 1* with data accesses the quality of the calibration, whereas contrasting *Model 2* against data tests the theoretical prediction that Λ_c, Λ_m, and ι are sufficient

[10]The model assumes balanced trade, whereas the US runs a nontrivial trade deficit in reality. Since I calibrate variable trade costs τ to match the pattern of gross imports, the proper benchmark for the value added content of trade should accordingly be the ratio of value-added to gross imports.

statistics for the value added content of trade. The figure shows that the model-generated time series are able to trace the actual series closely. The satisfactory fit of the model is further confirmed by high correlations between the model outcomes and data (0.71 and 0.78 for *Model 1* and *Model 2* respectively), as reported in Table 1.

In addition to replicating aggregate trade patterns, the model also succeeds in qualitatively accounting for the linkage distribution across industries. Figure 7 illustrates the relationship between an industry's weighted degrees in global production networks and its gross output. Consistent with the theoretical predictions in Section 4.1, the weighted outdegrees (the left panel) are positively correlated with industry output, whereas the weighted indegrees (the right panel) display little correlation with industry output. To corroborate the visual evidence of Figure 7, I regress the weighted degrees on industry output, focusing on the explained variation rather than causality. Table 2 reports the extent to which the cross-industry variation in output can explain the cross-industry variation in the weighted degrees. Across specifications, industry output can account for more than 40% of the observed variation in the weighted *outdegree*, but only 0.02% to 3.1% of that in the weighted *indegree*.

1.5.2 COUNTERFACTUALS

Having established the model's consistency with the data, I conduct two sets of counterfactual exercises to study the welfare impact of trade frictions. First, I decompose the changes in welfare over 2000-2014 according to four exogenous driving forces. Second, I compare the autarky equilibrium predicted by the model against the observed equilibrium in 2014 to gauge the total welfare gains from trade.

DECOMPOSING THE 2000-2014 WELFARE CHANGES

In Figure 8, the baseline in the left panel traces the changes in welfare (real wage) implied by the model. Over the sample period, cumulative welfare gains amount to 87.3% of the 2000 welfare level. This welfare improvement can be attributed to four sources, given by changes in the four time-varying parameters: trade costs τ, linkage fixed costs κ, entry costs ν, and labor endowment L. To access the relative importance of these four driving forces, I compute a series of counterfactual welfare, each time keeping one of the four parameters constant at its 2000 level. The gap between the baseline and each counterfactual then gives the contribution of the correspondent channel. Table 3 summarizes the decomposition results. Jointly, reductions in trade costs τ and linkage fixed costs κ contribute to more than 40% of the total welfare gains during 2000-2014. In Figure 8, the distance between the baseline and the "invariant τ" counterfactual visualizes the contribution of trade cost changes (19.3%), while the gap between the two plotted counterfactuals illustrates the relevance of varying linkage costs κ (24.2%). Even though linkage fixed costs κ play an equally, if not more, important role in explaining the evolution of welfare, the trend of trade in value added is almost solely accounted for by the cross-time variation in international trade frictions τ, as evidenced by the right panel of Figure 8.

How do the true welfare gains from trade liberalization compare with those predicted by the ACR formula? Since the ACR framework considers only international trade shocks, I first re-run the model holding all domestic factors (κ, ν, and L) constant at their initial levels. The baseline in Figure 9 plots the welfare pattern when trade costs τ are the only exogenous source of cross-time variation. I then compute the welfare series predicted by the standard ACR formula $d\ln W = d\ln\Lambda/\varepsilon$, where the domestic expenditure share Λ for each year is calculated directly from WIOD and the constant trade elasticity ε takes the value (8.22) implied by the model for year 2000. As shown by the dotted line in Figure 9, the ACR formula without intermediate inputs barely predicts any welfare change over the sample period despite of the sizable gains implied by the model. To give the ACR framework a fairer chance, I compute another welfare series using the extended ACR formula $d\ln W = d\ln\Lambda/(\sigma\varepsilon)$, which assumes that an intermediate input bundle aggregates all varieties in the same CES fashion as the final good basket and enters the production function in the Cobb-Douglas manner with exponent $1-\sigma$. The extended ACR formula is consistent with an exogenous production network that is globally complete ($\mu_H(\theta,\theta') = \mu_F(\theta,\theta') = 1$ for all $\theta,\theta' \in \Theta$) at zero linkage cost ($\kappa = 0$). Given by the dashed line in Figure 9, the ACR formula with the intermediate input extension predicts significantly larger welfare changes than the standard formula, as intermediate inputs amplify trade shocks. Nevertheless, compared to the baseline, the extended ACR formula still understates the welfare impact of trade cost adjustments, by nearly 1% of the 2000 real GDP in some years. Since the equilibrium network structure is in fact less than complete, the extended ACR formula enjoys an advantage over the endogenous network model with respect to the amplification mechanism provided by intermediate inputs. Thus, the fact that the model still generates larger welfare gains/losses than the extended ACR formula suggests that endogenous input-output linkages enable new adjustment margins beyond the two sufficient statistics identified by the ACR formula (the domestic expenditure share Λ and trade elasticity ε). The next set of counterfactuals shed more light on these additional margins.

RETURN TO AUTARKY

I now quantify the total gains from trade by comparing the latest observed equilibrium (year 2014) with the counterfactual autarky equilibrium where trade costs tend to infinity ($\tau \to \infty$). To gauge the quantitative importance of endogenous linkages, I let the model compete against an alternative version where the distribution of linkages $\{\mu_H(\theta,\theta'), \mu_F(\theta,\theta')\}_{\theta,\theta'\in\Theta}$ are fixed at the 2014 equilibrium outcome regardless of how trade costs τ change. I refer to this scenario as "Fixed Network" to emphasize the exogenous nature of linkage distribution, even though the total number of linkages may still respond to trade shocks through firm entry. Following the previous exercise, I also report the welfare gains implied by the standard and the extended ACR formulas, based on the observed domestic expenditure share and the model-implied trade elasticity (8.01) in 2014. The standard ACR formula is consistent with the Krugman (1980) model, whereas the extended formula applies to the Krugman model with intermediate inputs (i.e., complete global production networks at zero linkage cost). Table 4 column 1 reports the welfare gains from trade implied by the baseline model as well as the three alternative scenarios. The endogenous network model implies more welfare gains than all three alternatives with exogenous linkages. In particular, by allowing firms to rearrange their supplier-customer relationships, the endogenous network model enhances welfare gains from trade by almost a third from the "Fixed Network" scenario where the distribution of linkages is irresponsive to trade shocks.

Figure 10 illustrates the aggregate consequences of returning to autarky, under both the endogenous and fixed network assumptions. Comparing to the fixed network scenario, the model with endogenous linkages predicts a larger rise in the total number of firms N and a greater fall in average productivity A/N. However, the two models exhibit virtually no difference in the ratio of value-added to gross trade (the endogenous network model predicts a smaller *VAX* ratio by a negligible margin during the transition to autarky). To understand the causes behind these differential aggregate responses, I delve deeper into the micro-level adjustments to trade shocks.

First, I look at how global production networks undergo structural transformations as countries retreat to autarky. Figure 11 shows that surges in trade costs raise the productivity cutoffs for selecting both domestic (b_H) and foreign customers ($\underline{b}_F$ and $\overline{b}_F$). Figure 12 visualizes how the probability of forming a relationship depends on the customer and the supplier productivity in a two-country world, contrasting the trade equilibrium result against the autarky outcome. The difference between these two linkage density distributions is then highlighted in Figure 13, which reveals where the loss of firm linkages occurs. Rising international shipping costs force firms to abandon not only their customers abroad, but also those at home even though domestic trade frictions remain unchanged. Thus, international trade shocks have a spillover effect on domestic production networks, as long as there are still cross-country linkages left to function as the transmission channel. Specifically, when a firm loses international suppliers who retreat in the face of surging trade costs, the customer firm charges a higher price for its good to reflect a now inferior production technology due to less input diversity. This price increase is then passed on via downstream linkages to all of the firm's customers, including those at home. An analogous transmission mechanism also operates upward: when a firm cuts back on quantity produced due to its withdrawal from foreign markets, the resulting fall in intermediate input demand is then transmitted through upstream linkages to all of its suppliers, some of which are domestic.

Second, I examine how the price distribution and the firm productivity distribution in autarky differ from those in the trade equilibrium. In the two upper panels of Figure 14, I plot the relationship between productivity type θ and prices $p(\theta)$ as well as firm mass $N(\theta)g(\theta)$. In either the trade or the autarky equilibrium, prices decrease in firm productivity, whose equilibrium distribution is heavy-tailed. In the two lower panels, I compare the endogenous network (EN) and the fixed network (FN) models by graphing their respective predictions on the changes in prices and firm mass from the open economy to autarky. The endogenous network model implies a more prominent price hike than the fixed network alternative, because it allows trade shocks to affect production costs not only on the usual intensive margin but also through the extensive margin of linkage destruction. For firm mass, both the endogenous and the fixed network models predict that, moving from the trade to the autarky equilibrium, firm entry increases in low-productivity industries while decreasing in the high-productivity ones. This leftward shift in the firm productivity distribution has the following intuition: in the open economy, firms from high-productivity industries rely more heavily on imported intermediates than their low-productivity counterparts. Thus, the negative impact of rising trade costs on profit is relatively larger for high-type firms, driving entrants towards industries with relatively low productivity. The leftward shift of the firm productivity distribution depresses the average productivity of the economy, as shown in Figure 10. Furthermore, the endogenous network model implies a greater autarky total firm mass than the fixed network model, because the structural transformation of production networks frees up more labor from forming linkages to creating firms.

1.5.3 Robustness Check

Finally, I check whether the quantitative welfare gains from trade are sensitive to alternative parameter values. I focus on β, which determines the elasticity of substitution among varieties within the same industry $1/(1-\beta)$ and influences the magnitude of trade costs τ as well as trade elasticity ε. I re-calibrate the model by setting $1/(1-\beta)$ to 4 and 8 respectively (the baseline calibration sets this value at 6). Columns 2 and 3 of Table 4 reports the welfare gains from autarky to the 2014 equilibrium predicted by these two alternative calibration exercises. In addition, Table 5 reports the model-implied trade costs τ and trade elasticity ε for different values of β.

A low elasticity of substitution implies a low trade elasticity, and in turn larger welfare gains regardless of whether linkages are endogenous or fixed. On the other hand, the quantitative relevance of linkage redistribution (measured by the difference between the "Endogenous Network" and the "Fixed Network" predictions) increases in the value of β. Recall that the difference between the within-industry and the cross-industry substitutability of goods $\beta-\alpha$ regulates how responsive the demand for intermediate inputs is with respect to the number of similar suppliers. Therefore, the lower the value of β, the smaller this elasticity differential $\beta-\alpha$ becomes, diminishing the impact of linkage redistribution on intermediate demand, production costs, and ultimately aggregate welfare. As noted by the literature, the overall welfare gains exhibit considerable sensitivity to the substitutability across similar goods: the gains implied by the endogenous network model more than double from 10.7% to 25.4% as the within-industry elasticity of substitution rises from 4 to 8. In contrast, the welfare gains purely due to linkage redistribution is much more robust to recalibration: the welfare gains differential between the "Endogenous Network" and the "Fixed Network" scenarios ranges from 3.3% to 3.9% across the three calibration exercises.

1.6 Conclusion

This paper proposes a framework to study production networks in the context of international trade. I endogenize the input-output structure of the global economy, allowing firms to form supplier-customer relationships with their counterparts both at home and from abroad. Endogenous firm linkages generate new welfare implications of trade liberalization. In addition to increasing the variety of consumption goods, trade integration can also raise welfare from the production side, as both domestic and international linkages respond to falling trade costs. Since the structure of global production networks determines the set of intermediate inputs available to each firm, trade liberalization brings about technological changes through the joint adjustments of linkages within and across borders. Quantitative exercises using the World Input-Output Database confirm the relevance of endogenous linkages in delivering welfare gains from trade.

The equilibrium of this economy is inefficient, with inefficiency stemming from two aspects (in addition to the standard inefficiency of monopolistic competition). First, linkages amplify the distortion due to monopolistic markups, because the markup charged by a supplier affects not only its immediate customers but also firms further downstream. Second, firms do not take into account the effect of their relationship choices on the overall structure of production networks. In particular, the second type of inefficiency will persist even if firms are allowed to split linkage fixed costs through bilateral bargaining, because the bilateral surplus perceived by a supplier-customer pair does not necessarily coincide with the social returns to that relationship. Therefore, one follow-up to this paper could be to explore under which conditions does inefficiency lead to an under-connected versus over-connected economy.

I conclude this paper by suggesting several directions in which this framework can be of use. For theoretical research, this framework offers a tractable way of introducing endogenous linkages into standard macroeconomic models. Existing business cycle literature often assumes an exogenous input-output structure when studying how production networks propagate macroeconomic shocks. My model suggests one approach to make production networks themselves responsive to shocks without losing analytic tractability. For empirical work, the framework is particularly useful for policy evaluation. In the model, both linkage fixed costs and variable trade costs are crucial in shaping the structure of production networks. One may then ask to what extent can reducing linkage fixed costs, a domestic policy choice, be a substitute for international trade agreements in terms of welfare benefits. My framework offers a suitable laboratory for conducting such policy experiment.

1.7 Appendix A: Figures and Tables

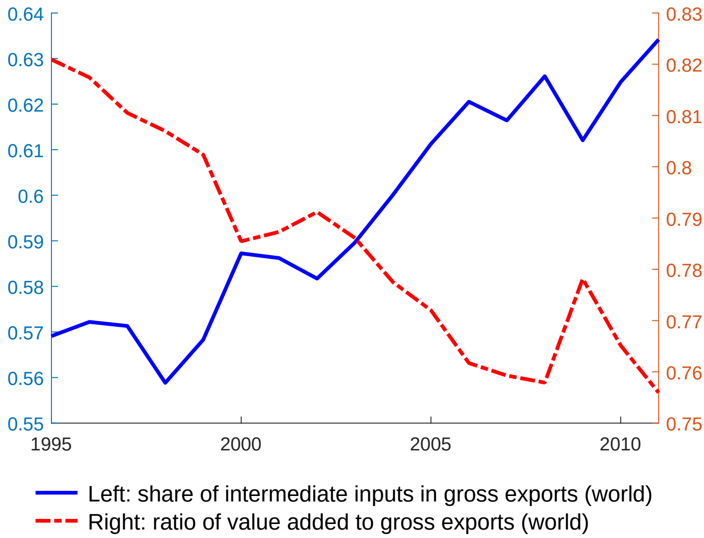

Figure 1.7.1: The share of intermediate inputs in world exports (the left axis) and the ratio of value added to gross exports for the world (the right axis). The data source is the OECD Trade in Value Added database.

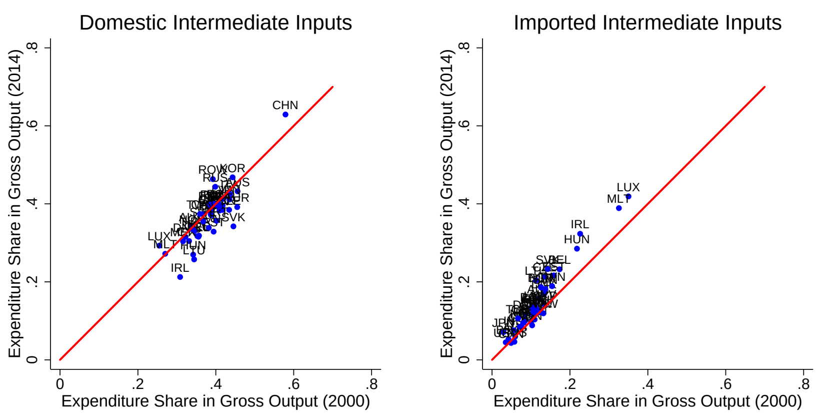

Figure 1.7.2: Country-level expenditure shares on domestic intermediate inputs (the left panel) and imported intermediate inputs (the right panel) in gross output, for year 2000 and year 2014. The data source is the World Input-Output Database.

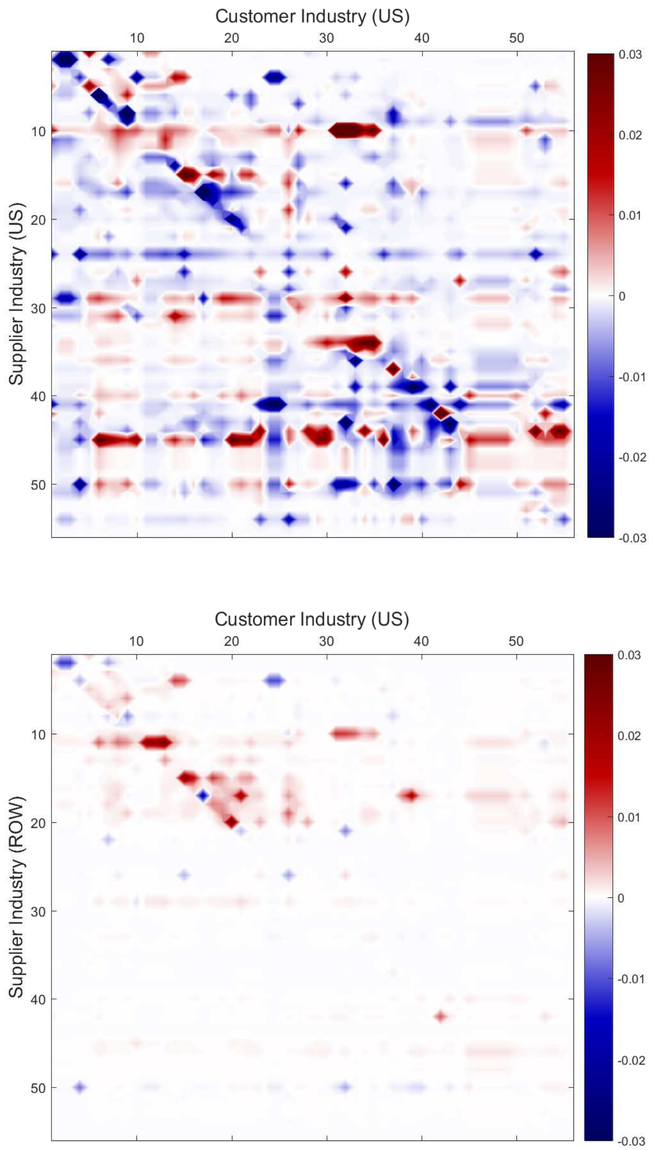

Figure 1.7.3: Changes between 2000 and 2014 in the US domestic direct requirement (the upper panel) and the US direct requirement of imported intermediate inputs (the lower panel). In each matrix, the (i, j)-th entry gives the change between 2000 and 2014 in the expenditure on industry i per value unit of industry j's gross output, i.e. the direct requirement coefficients. The axis labels correspond to industry index, with 1-4 belonging to the agriculture and mining sector, 5-23 the manufacturing sector, and 24-56 the service sector. The data source is the World Input-Output Database.

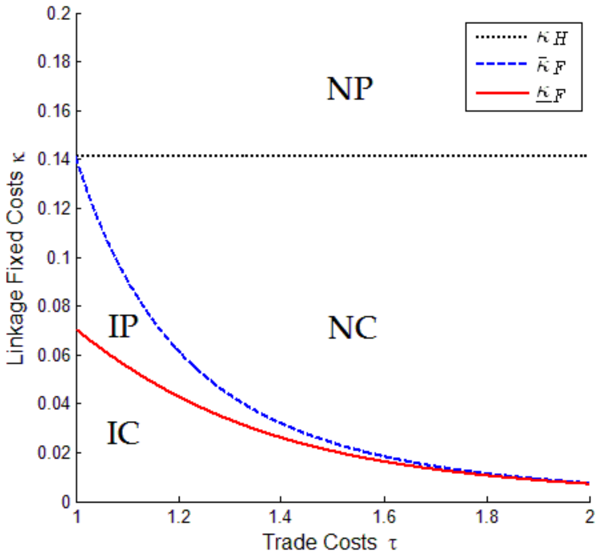

Figure 1.7.4: Linkage fixed cost thresholds in the homogeneous firm model. Simulation parameters: $\alpha = 0.3$, $\beta = 0.8$, $\sigma = 0.7$, $L = 1$, $\nu = 1$, $\theta = 1$.

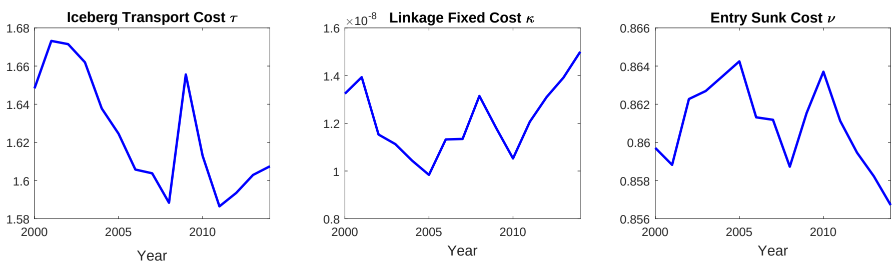

Figure 1.7.5: Calibrated value of the cost parameters

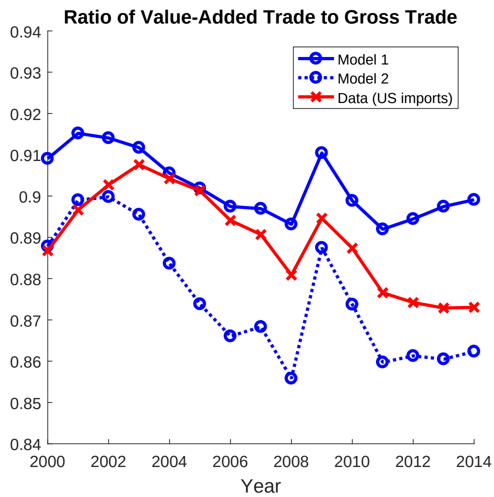

Figure 1.7.6: Ratio of value-added trade to gross trade, model predictions versus data. The data source is WIOD, which is transformed into a two-country setup (US and the rest of the world) consistent with the theoretical framework. The ratio of value-added imports to gross imports is then calculated for the US from the transformed two-country input-output table, following the method described by Johnson and Noguera (2017).

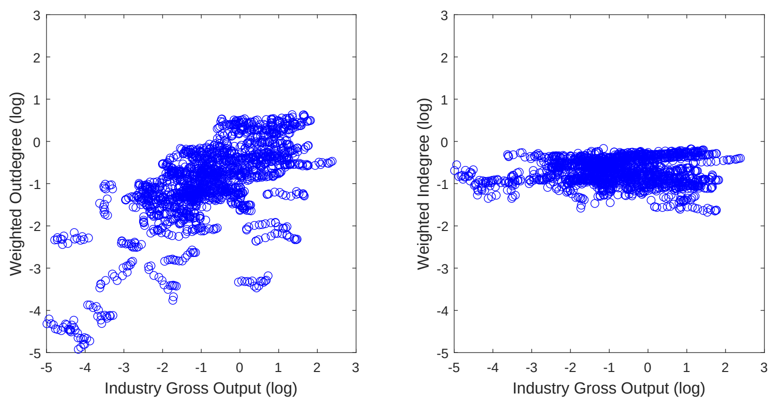

Figure 1.7.7: Industry output and weighted degrees calculated from WIOD. Each dot corresponds to an industry-year observation. The sample covers two countries (US and ROW), 55 industries (The industry defined as "Activities of extraterritorial organizations and bodies" is excluded as an outlier, because it produce a trivial amount of output but relies heavily on service inputs due to the special nature of international organizations.), and a time span from 2000 to 2014.

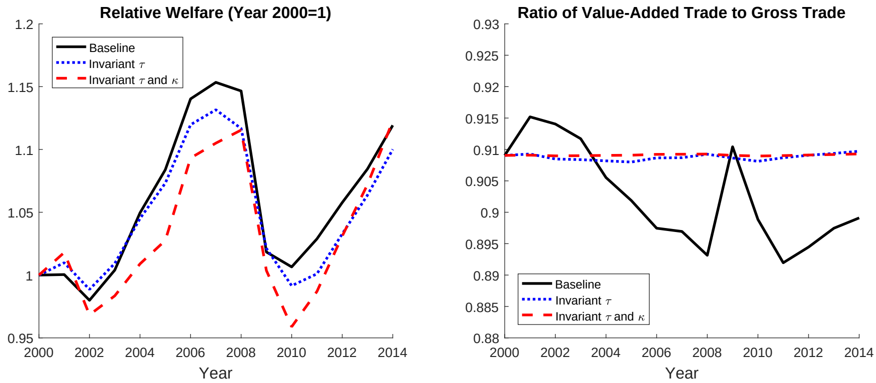

Figure 1.7.8: Contribution to the evolution of welfare (left panel) and value-added trade (right panel) by changes in trade costs τ and linkage fixed costs κ

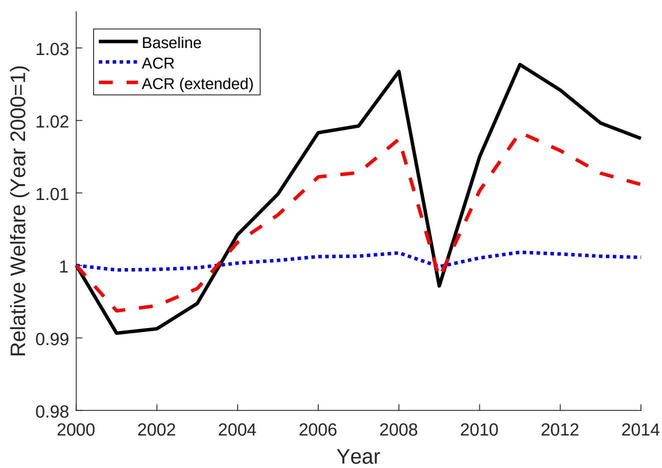

Figure 1.7.9: Welfare gains/losses solely due to changes in variable trade costs τ: model outcome versus predictions by the ACR formula

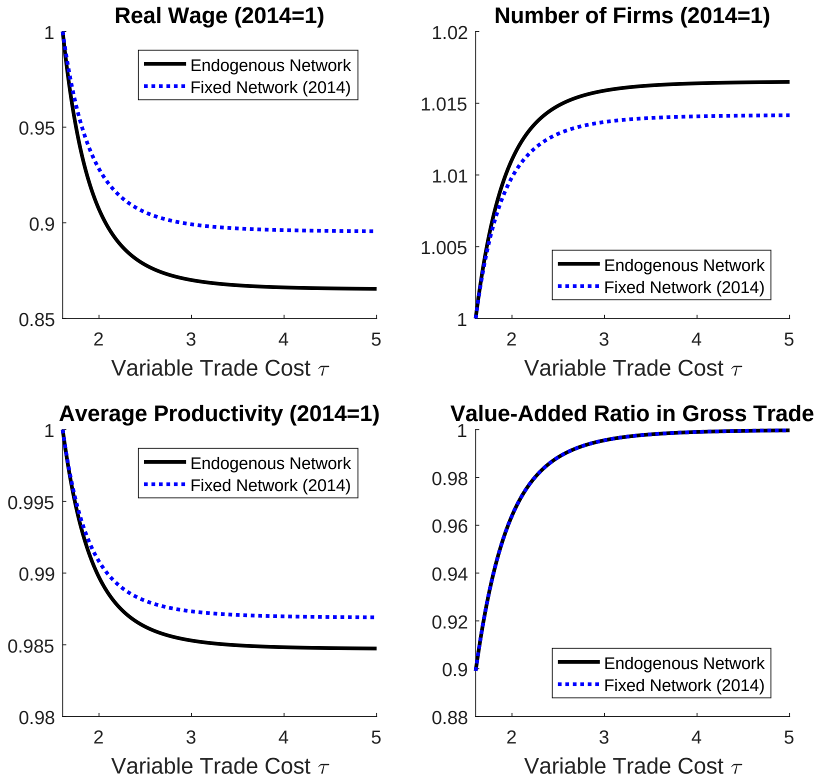

Figure 1.7.10: Aggregate consequences of moving into autarky from the 2014 equilibrium

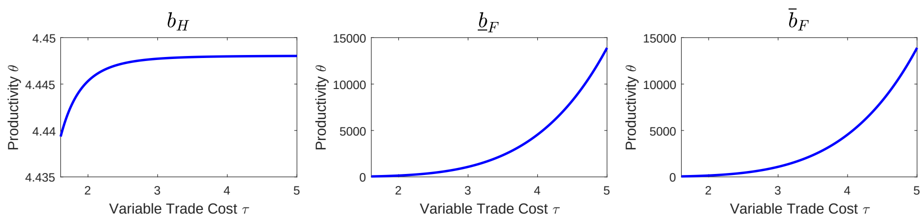

Figure 1.7.11: Responses of the customer-selection cutoffs to trade cost surges

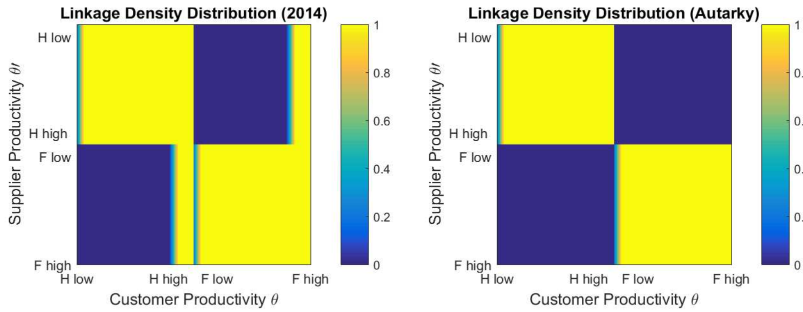

Figure 1.7.12: Linkage density distribution implied by the model. From left to right (top to bottom), customer productivity θ (supplier productivity θ') is ranked from low to high for Home (H) and Foreign (F) separately.

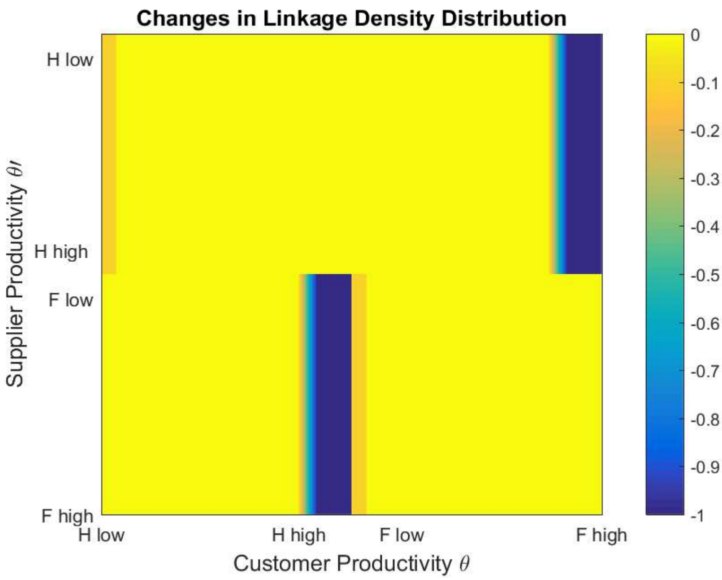

Figure 1.7.13: Changes in linkage density distribution from the 2014 equilibrium to autarky.

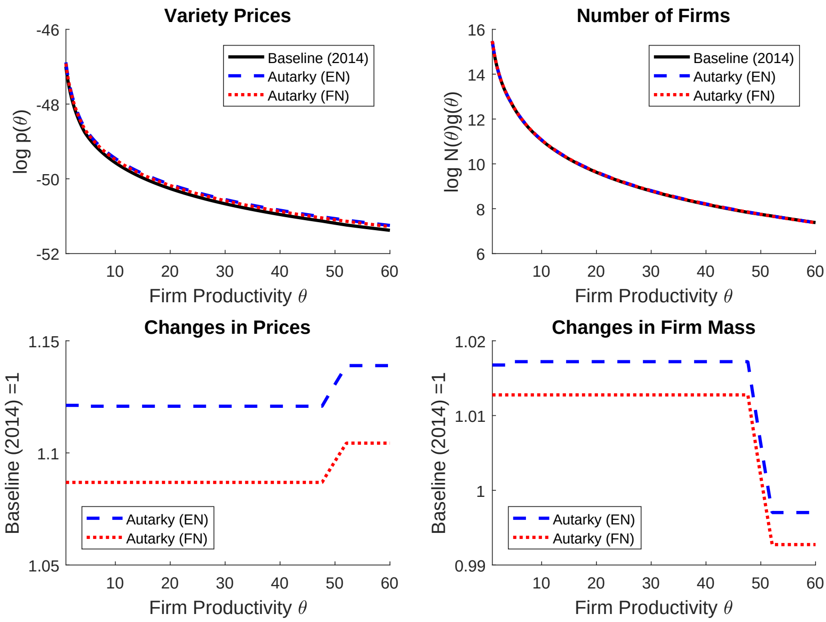

Figure 1.7.14: Firm level consequences of moving into autarky from the 2014 equilibrium, under the assumption of endogenous networks (EN) and fixed networks (FN).

Table 1.1: The ratio of value-added trade to gross trade, data and model predictions

	Data (US imports)	Model 1	Model 2
2000-2014 Mean	0.890	0.902	0.876
Correlation (model, data)		0.707	0.777

NOTE: see the legend of Figure 6 for the data source.

Table 1.2: Explaining the cross-industry variation in the weighted degrees

	DEPENDENT VARIABLE: $\ln d_{OUT}(\theta)$			DEPENDENT VARIABLE: $\ln d_{IN}(\theta)$		
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln R(\theta)$	0.477^{***}	0.497^{***}	0.563^{***}	0.003	0.002	-0.023^{***}
	(0.019)	(0.019)	(0.023)	(0.006)	(0.006)	(0.007)
Year dummies	*No*	*Yes*	*Yes*	*No*	*Yes*	*Yes*
Country dummy	*No*	*No*	*Yes*	*No*	*No*	*Yes*
Observations	1650	1650	1650	1650	1650	1650
R^2	0.419	0.438	0.458	0.0002	0.002	0.031

NOTE: Robust standard errors are in the parentheses. * Significant at 10%; ** Significant at 5%; *** Significant at 1%. The sample covers two countries (US and ROW), 55 industries (The industry defined as "Activities of Extraterritorial Organizations and Bodies" is excluded as an outlier, because it produces a trivial amount of output but relies heavily on service inputs due to the special nature of international organizations.), and a time span from 2000 to 2014.

Table 1.3: Contribution to the 2000-2014 cumulative welfare gains

	Source of gains			
	Trade costs (τ)	*Linkage costs* (κ)	*Entry costs* (ν)	*Labor* (L)
Percentage contribution	19.3	24.2	-4.16	60.6

Table 1.4: Changes in welfare from autarky to the 2014 trade equilibrium

	Baseline	Robustness Checks	
	$\beta = 5/6$	$\beta = 3/4$	$\beta = 7/8$
	(1)	(2)	(3)
Endogenous Network	15.5%	25.4%	10.7%
Fixed Network	11.7%	22.1%	6.8%
Krugman	1.0%	1.2%	0.8%
Krugman with Intermediates	10.6%	12.8%	8.7%

NOTE: the baseline calibration assumes a within-industry elasticity of substitution $1/(1-\beta) = 6$; the alternative calibration for robustness checks sets this elasticity at 4 and 8 respectively.

Table 1.5: Model-implied trade costs and trade elasticities (range during 2000-2014)

	Baseline	Robustness Checks	
	$\beta = 5/6$	$\beta = 3/4$	$\beta = 7/8$
	(1)	(2)	(3)
Trade cost τ	1.59-1.67	2.17-2.37	1.39-1.44
Trade elasticity ε	7.91-8.35	6.57-7.18	9.58-9.96

NOTE: the baseline calibration assumes a within-industry elasticity of substitution $1/(1-\beta) = 6$; the alternative calibration for robustness checks sets this elasticity at 4 and 8 respectively.

1.8 APPENDIX B: CHARACTERIZATION OF THE HETEROGENEOUS FIRM EQUILIBRIUM

This appendix characterizes the heteogeneous firm equilibrium in a generalized setup with a continuum of productivity types θ following a Pareto distribution $g(\theta)$ as described in Proposition 5. I first study the autarky equilibrium in Section A.1 and then the open economy equililbrium in Section A.2.

1.8.1 AUTARKY

I first consider the autarky equilibrium ($\tau \to \infty$) and study the market size effect by raising population L. In this case, the productivity cutoffs for foreign customers become infinitely high ($\lim_{\tau\to\infty} \underline{b}_F = \lim_{\tau\to\infty} \overline{b}_F = \infty$), and firms invest in domestic downstream links according to the productivity cutoff b_H. Specifically, a firm will supply inputs to all firms with TFP level θ above or equal to the cutoff b_H, but to only a subset of the firms with productivity below this cutoff. In the next lemma, I describe how changes in market size L affect the customer-selection cutoff b_H, aggregate productivity A, as well as the firm level outcomes.

Lemma 8 *If linkage fixed costs are high ($\kappa > \overline{\kappa}$), expanding market size results in a higher customer-selection cutoff ($db_H/dL > 0$) , higher aggregate productivity ($dA/dL > 0$), and larger production scale for all firms ($dr(\theta)/dL > 0$ for all $\theta \in \Theta$). However, the market size effect on the price levels $p(\theta)$ and firm distribution $N(\theta)$ is ambiguous in this case. If linkage fixed costs are low ($\kappa \leqslant \overline{\kappa}$), expanding market size does not affect the customer-selection cutoff ($db_H/dL = 0$) but improves aggregate productivity ($dA/dL > 0$). For all firms, the scale of production grows and prices fall ($dr(\theta)/dL > 0$, $dp(\theta)/dL < 0$ for all $\theta \in \Theta$). Moreover, the number of firms increases in all industries ($dN(\theta)/dL > 0$, for all $\theta \in \Theta$).*

The initial structure of production networks determines how the customer-selection cutoff b_H responds to market size expansions. For sufficiently low linkage fixed costs ($\kappa \leqslant \overline{\kappa}$), production networks are complete because firms find it profitable to supply inputs to all other firms, regardless of customer productivity. In this case, a marginal increase in population L leaves the customer-selection cutoff b_H unchanged at the lower bound $\underline{\theta}$, since it is still desirable for firms to connect with all its peers. Prompted by the market expansion, more firms enter business across industries, which improves the production technology at the firm level because a wider range of input varieties become available for adoption. As the marginal cost of production declines, firms are able to charge lower prices and grow in size. Even though larger markets lead to better firm level technology, the aggregate production technology remains the same because firm entry is proportional to industry size, which preserves the shape of the firm productivity distribution.

If linkage fixed costs exceed the threshold level ($\kappa > \overline{\kappa}$), production networks are incomplete and only firms with productivity levels above or equal to the customer-selection cutoff ($\theta \geqslant b_H$) have access to all input varieties. In this case, a marginal expansion of market size L elevates the customer-selection cutoff b_H. Furthermore, for any two industries with TFP levels θ_{low} and θ_{high} such that $\theta_{low} < b_H < \theta_{high}$, a marginal increase in market size L raises the price ratio $p(\theta_{low})/p(\theta_{high})$ and lowers the firm mass ratio $N(\theta_{low})/N(\theta_{high})$. A higher cutoff b_H means that some firms used to source intermediate inputs from all producers in the country, but now have access to only a subset of the input varieties being produced. As firms become more selective in the productivity of their customers, the distribution of input-output linkages shifts away from low-type customers towards the high-type ones. Such redistribution of suppliers reduces the input diversity of low-productivity firms relative to their high-productivity peers, raising the price ratio of low-type goods to high-type ones. To restore equilibrium, firm entry shifts towards high-productivity industries where market expansions bring about relatively larger benefits, thus depressing the firm mass ratio of a low-productivity industry to a high-productivity one. Contrasting the homogeneous firm model, market expansions have differential consequences on the production technology of firms, biasing towards high-type producers. In addition, the aggregate production technology benefits from market expansions due to the rightward shift of the firm productivity distribution.

Therefore, the technological consequences of larger markets, at both the firm and the aggregate level, depend crucially on the structure of production networks. It then follows that the market size effect on welfare is also dependent on the network structure and fundamentally on linkage fixed costs κ. In the following proposition, I summarize the degree of aggregate increasing returns to scale in the heterogeneous firm equilibrium, still comparing against the Krugman benchmark.

Proposition 7 *For $\kappa \in [0, \overline{\kappa}]$, the market size effect on welfare is larger than that in Krugman (1980): $d\ln W/d\ln L > (1-\beta)/\beta$. For $\kappa \in (\overline{\kappa}, \infty)$, the market size effect on welfare is ambiguous.*

When production networks are complete ($\kappa \leqslant \overline{\kappa}$, and $b_H = \underline{\theta}$), increasing market size improves the firm level production technology universally. The intuition for more aggregate increasing returns to scale than in Krugman (1980) is the same as in the homogeneous firm model: higher input diversity brings down production costs and in turn lowers the price of goods, generating welfare gains in addition to the standard benefit of expanding consumption variety. When production networks are incomplete ($\kappa > \overline{\kappa}$, and $\underline{\theta} < b_H < \infty$), the market size effect becomes more nuanced. As supplier connections and the mass of firms both redistribute towards more productive industries, firms in less productive industries may lose suppliers in absolute terms, and therefore may have to raise prices to reflect higher production costs. Such increases in price would be transmitted to all downstream firms through supplier-customer linkages, further dampening the benefit of a larger market size.

1.8.2 The open economy

Next, I consider the heterogeneous firm equilibrium in an open economy ($1 < \tau < \infty$), where the relevant market size is $\left(1 + \tau^{\frac{\beta}{\beta-1}}\right) L$. To guarantee equilibrium uniqueness, I focus on the case where the customer-selection cutoff b_H lies in the interior of the productivity support ($\underline{\theta} < b_H < \infty$), which requires that linkage fixed costs are sufficiently high ($\kappa > \overline{\kappa}$). Substituting the CS condition $b_H = \kappa [1 - \alpha (1 - \sigma)] A^2 / [\alpha (1 - \alpha)(1 - \sigma) L]$ into the FE condition, I obtain the equilibrium customer-selection cutoff b_H as the unique root to the following equation:

$$\sqrt{\frac{[1-\alpha(1-\sigma)]\,\nu^2}{\kappa\alpha(1-\alpha)(1-\sigma)L} b_H} + \frac{T_1}{T_2 - \left(\frac{b_H}{\underline{\theta}}\right)^{\zeta - \frac{\alpha\beta}{\beta-\alpha} + \sigma\frac{\alpha(1-\beta)}{\beta-\alpha}}} - \frac{1-\alpha(1-\sigma)}{\alpha(1-\sigma)} = 0$$

where constants T_1 and T_2 are expressions of the exogenous parameters α, β, σ, ζ, τ only and are given in Appendix B. Trade liberalization is modeled as the decline in trade costs τ, which enter the equilibrium condition of b_H solely through constants T_1 and T_2. For a general set of parameters, the effect of trade cost reductions on the customer-selection cutoff b_H is ambiguous. Nevertheless, for a given value of the derivative $db_H/d\tau$, comparative statics of other aggregate and firm level variables remain tractable. The next lemma presents comparative statics of aggregate productivity A and firm size $r(\theta)$ with respect to trade liberalization, conditional on the response of b_H.

Lemma 9 *If trade liberalization lowers the productivity cutoff for selecting domestic customers ($db_H/d\tau > 0$), aggregate productivity falls ($dA/d\tau >$*

0*) and the scale of production increases for all firms (*$dr(\theta)/d\tau < 0$ *for all* $\theta \in \Theta$*). If trade liberalization raises the productivity cutoff for selecting domestic customers (*$db_H/d\tau < 0$*), aggregate productivity grows (*$dA/d\tau < 0$*) and the scale of production shrinks for all firms (*$dr(\theta)/d\tau > 0$ *for all* $\theta \in \Theta$*).*

In addition to the domestic cutoff b_H, trade integration also moves the cutoffs $\underline{b}_F$ and $\overline{b}_F$ for selecting foreign customers, thus bringing structural changes to production networks both within and across borders. As production networks evolve with globalization, not only does a firm's set of suppliers adjust, but its import status may also change depending on whether it attracts foreign sellers or not. According to Lemma 5, a falling τ always shortens the intervals $[b_H, \underline{b}_F]$ and $\left[b_H, \overline{b}_F\right]$, while lengthening the interval $\left[\underline{b}_F, \overline{b}_F\right]$. However, changes in the absolute position of $\underline{b}_F$ and $\overline{b}_F$ depend on the elasticity of the domestic cutoff b_H with respect to the degree of integration $\mathscr{E}_{b_H} \equiv d\ln b_H / d\ln\left(1+\tau^{\frac{\beta}{\beta-1}}\right)$. Specifically, four cases cover all the possible adjustment patterns of global production networks in response to a small decline in trade costs τ: (i) if $\mathscr{E}_{b_H} \in \left(1+\tau^{\frac{\beta}{1-\beta}}, \infty\right)$, then $d\underline{b}_F/d\tau < 0$ and $d\overline{b}_F/d\tau < 0$; raising all three cutoffs, globalization makes suppliers more selective in both their domestic and foreign customers. (ii) if $\mathscr{E}_{b_H} \in \left(\tau^{\frac{\beta}{1-\beta}}, 1+\tau^{\frac{\beta}{1-\beta}}\right)$, then $d\underline{b}_F/d\tau > 0$ and $d\overline{b}_F/d\tau < 0$; raising b_H and $\overline{b}_F$ while lowering $\underline{b}_F$, globalization makes suppliers more selective in their domestic customers, whereas neither more nor less so in foreign customers . (iii) if $\mathscr{E}_{b_H} \in \left(0, \tau^{\frac{\beta}{1-\beta}}\right)$, then $d\underline{b}_F/d\tau > 0$ and $d\overline{b}_F/d\tau > 0$; raising b_H while lowering both $\underline{b}_F$ and $\overline{b}_F$, globalization makes suppliers more selective in their domestic customers but less so in foreign customers. (iv) if $\mathscr{E}_{b_H} < 0$, then $d\underline{b}_F/d\tau > 0$ and $d\overline{b}_F/d\tau > 0$; lowering all three cutoffs, globalization makes suppliers less selective in both their domestic and foreign customers.

1.9 APPENDIX C: PROOFS AND DERIVATION

Proof of Lemma 1

If μ_H has an interior solution in equilibrium, it must be that $\underline{\kappa}_H < \kappa < \infty$ and $\mu_F = 0$. In this case, conditions (1.2.2) and (1.2.3) imply the equilibrium value of μ_H as

$$\mu_H = \frac{\alpha(1-\alpha)(1-\sigma)L}{[1-\alpha(1-\sigma)]\kappa\theta N^2} \qquad \text{where } N = \frac{(1-\alpha)L}{\theta\nu}$$

The value of the linkage fixed cost threshold $\underline{\kappa}_H$ can be found by substituting $\mu_H = 1$ and $\kappa = \underline{\kappa}_H$ into the above equation. If μ_F has an interior solution in equilibrium, it must be that $\underline{\kappa}_F < \kappa < \overline{\kappa}_F$ and $\mu_H = 1$. In this case, conditions (1.2.2) and (1.2.4) imply the equilibrium value of μ_F as

$$\mu_F = \frac{\alpha(1-\alpha)(1-\sigma)L}{[1-\alpha(1-\sigma)]\kappa\theta N^2} - \tau^{\frac{\beta}{1-\beta}} \tag{1.9.1}$$

where N is given by the following quadratic equation

$$\left(\tau^{\frac{\beta}{1-\beta}} - 1\right)\kappa\theta N^2 - \nu\theta N + (1-\alpha)L = 0 \tag{1.9.2}$$

The larger root of the above equation is discarded, because it implies $d\mu_F/d\kappa > 0$ and hence $\mu_F \in (0,1)$ cannot be an equilibrium over $\underline{\kappa}_F < \kappa < \overline{\kappa}_F$. The value of the linkage fixed cost threshold $\overline{\kappa}_F$ can be found by substituting $\mu_F = 0$ and $\kappa = \overline{\kappa}_F$ into the system given by (1.9.1) and (1.9.2). Similarly, the value of the linkage fixed cost threshold $\underline{\kappa}_F$ can be found by substituting $\mu_F = 1$ and $\kappa = \underline{\kappa}_F$ into the aforementioned system.

Proof of Proposition 1

By Lemma 1, $\lim_{\sigma\to1}\mu_H = \lim_{\sigma\to1}\mu_F = 0$. Thus, in the limit of $\sigma = 1$, condition (4) implies that $N = (1-\alpha)L/(\nu\theta)$. Welfare as measured by real wage in this limiting case is

$$\lim_{\sigma\to1} W = \left[\left(1+\tau^{\frac{\beta}{\beta-1}}\right)\lim_{\sigma\to1} N\right]^{\frac{1-\beta}{\beta}} \alpha\theta = \left[\left(1+\tau^{\frac{\beta}{\beta-1}}\right)\frac{(1-\alpha)L}{\nu\theta}\right]^{\frac{1-\beta}{\beta}} \alpha\theta$$

which implies an elasticity with respect to market size $d\ln W/d\ln\left[\left(1+\tau^{\frac{\beta}{\beta-1}}\right)L\right] = (1-\beta)/\beta$.

Proof of Lemma 2 and Proposition 2

This lemma considers the limiting case of autarky: $\tau\to\infty$. If $\kappa\in[\underline{\kappa}_H,\infty)$, we have $N = (1-\alpha)L/(\nu\theta)$ and $r = \nu\theta/\{(1-\alpha)[1-\alpha(1-\sigma)]\}$, which implies $dN/dL > 0$ and $dr/dL = 0$. In this case, the welfare expression becomes

$$W = \left(\mu_H^{\frac{\sigma-1}{\sigma}} N^{-\frac{1}{\sigma}}\right)^{\frac{\beta-1}{\beta}} (\alpha\theta)^{\frac{1}{\sigma}}$$
$$= L^{\frac{1-\beta}{\beta}} \left\{\frac{\alpha(1-\sigma)\nu}{[1-\alpha(1-\sigma)]\kappa}\right\}^{\left(\frac{1-\beta}{\beta}\right)\frac{1-\sigma}{\sigma}} \left(\frac{1-\alpha}{\nu\theta}\right)^{\frac{1-\beta}{\beta}} (\alpha\theta)^{\frac{1}{\sigma}}$$

which implies $d\ln W/d\ln L = (1-\beta)/\beta$. If $\kappa \in [0, \underline{\kappa}_H)$, equilibrium admits an corner solution $\mu_H = 1$, and the total number of firms is given by the following quadratic equation

$$\kappa\theta N^2 + \nu\theta N - \left[\frac{1-\alpha}{1-\alpha(1-\sigma)}\right] L = 0$$

Applying the implicit function theorem yields $dN/dL > 0$. The last equation also implies

$$r = \frac{L}{[1-\alpha(1-\sigma)]N} = \frac{\kappa\theta N + \nu\theta}{1-\alpha}$$

and therefore $dr/dL \propto dN/dL > 0$. In this case, the welfare expression becomes$W = N^{\frac{1-\beta}{\sigma\beta}} (\alpha\theta)^{\frac{1}{\sigma}}$, which implies

$$\frac{d\ln W}{d\ln L} = \frac{1-\beta}{\sigma\beta}\frac{d\ln N}{d\ln L} = \frac{1-\beta}{\sigma\beta}\left(\frac{\kappa N + \nu}{2\kappa N + \nu}\right)$$

The condition for $d\ln W/d\ln L > (1-\beta)/\beta$ is $(\kappa N + \nu)/(2\kappa N + \nu) > \sigma$, which always holds if $\sigma \leqslant 1/2$. If $\sigma > 1/2$, the condition $(\kappa N + \nu)/(2\kappa N + \nu) > \sigma$ requires that

$$\kappa < \frac{[1-\alpha(1-\sigma)]\left[1-(2\sigma-1)^2\right]\theta\nu^2}{4(1-\alpha)(2\sigma-1)^2 L}$$

The above inequality holds for all $\kappa \in [0, \underline{\kappa}_H)$ if the right hand side is larger than the fixed cost cutoff $\underline{\kappa}_H$:

$$\underline{\kappa}_H = \frac{\alpha(1-\sigma)\theta\nu^2}{(1-\alpha)[1-\alpha(1-\sigma)]L} < \frac{[1-\alpha(1-\sigma)]\left[1-(2\sigma-1)^2\right]\theta\nu^2}{4(1-\alpha)(2\sigma-1)^2 L}$$

The last inequality implies $\alpha\sigma < 1$, which is true by assumption. Therefore, for all $\kappa \in [0, \underline{\kappa}_H)$, we have $d\ln W/d\ln L > (1-\beta)/\beta$.

Proof of Lemma 3

By Lemma 1, the condition for $\underline{\kappa}_F \leqslant \overline{\kappa}_F$ is

$$\left(1+\tau^{\frac{\beta}{\beta-1}}\right)\left[1+\tau^{\frac{\beta}{\beta-1}}-2\tau^{\frac{\beta}{\beta-1}}\alpha(1-\sigma)\right]^{-2} \leqslant \left[1-\tau^{\frac{\beta}{\beta-1}}\alpha(1-\sigma)\right]^{-2}$$

which is equivalent to $1-2\alpha(1-\sigma)+\left\{[1-\alpha(1-\sigma)]^2+2\alpha^2(1-\sigma)^2\right\}\tau^{\frac{\beta}{\beta-1}} \geqslant \alpha^2(1-\sigma)^2\tau^{\frac{2\beta}{\beta-1}}$. If $\alpha(1-\sigma) \leqslant 1/2$, the condition for $\underline{\kappa}_F \leqslant \overline{\kappa}_F$ holds for all $\tau \geqslant 1$. If $\alpha(1-\sigma) > 1/2$, the condition for $\underline{\kappa}_F \leqslant \overline{\kappa}_F$ is satisfied only for sufficiently small τ such that

$$\tau \leqslant \widehat{\tau} \equiv \left\{\frac{1}{2}\left[\frac{1-\alpha(1-\sigma)}{\alpha(1-\sigma)}\right]^2+1-\sqrt{\left\{\frac{1}{2}\left[\frac{1-\alpha(1-\sigma)}{\alpha(1-\sigma)}\right]^2+1\right\}^2+\frac{1-2\alpha(1-\sigma)}{\alpha^2(1-\sigma)^2}}\right\}^{\frac{\beta-1}{\beta}}$$

In this case, global production networks are NP if $\underline{\kappa}_H < \kappa < \infty$, NC if $\max\{\overline{\kappa}_F, \underline{\kappa}_F\} \leqslant \kappa \leqslant \underline{\kappa}_H$, IP if $\underline{\kappa}_F < \kappa < \overline{\kappa}_F$, and IC if $0 \leqslant \kappa \leqslant \min\{\overline{\kappa}_F, \underline{\kappa}_F\}$. For $(\tau, \kappa) \in (\widehat{\tau}, \infty) \times [\overline{\kappa}_F, \underline{\kappa}_F]$, there are two possible trade equilibria, with NC and IC as the respective equilibrium network structure.

Proof of Lemma 4

Global production networks are NP if $\underline{\kappa}_H < \kappa < \infty$; in this case we have $N = (1-\alpha)L/(\nu\theta)$ by the proof of Lemma 1 and therefore $dN/d\tau = 0$. Global production networks are NC if $\overline{\kappa}_F \leqslant \kappa \leqslant \underline{\kappa}_H$; in this case the total number of firms is given by the equilibrium condition (1.2.2) which becomes $\kappa\theta N^2 + \nu\theta N - (1-\alpha)L/[1-\alpha(1-\sigma)] = 0$ given $\mu_H = 1$ and $\mu_F = 0$. Since this equation does not involve τ, we have $dN/d\tau = 0$. Global production networks are IP if $\underline{\kappa}_F < \kappa < \overline{\kappa}_F$; in this case, the proof of Lemma 1 establishes that total firm mass N is given by the smaller root of the quadratic equation (1.9.2). Applying the implicit function theorem to this quadratic equation, we have $dN/d\tau^{\frac{\beta}{1-\beta}} = \kappa\theta N^2\left[(\nu\theta)^2-4\kappa\theta(1-\alpha)L\left(\tau^{\frac{\beta}{1-\beta}}-1\right)\right]^{-\frac{1}{2}} > 0$, which implies $dN/d\tau > 0$. Global production networks are IC if $0 \leqslant \kappa \leqslant \underline{\kappa}_F$; in this case the total number of firms is given by the equilibrium condition (1.2.2) which becomes $2\kappa\theta N^2 + \nu\theta N - (1-\alpha)L/[1-\alpha(1-\sigma)] = 0$ given $\mu_H = \mu_F = 1$. Since this equation does not involve τ, we have $dN/d\tau = 0$.

Proof of Proposition 3

When global production networks are NP or NC, $\Lambda_m = 1$ since there is no intermediate input trade. Lemma 1 implies $d\mu_H/d\tau = 0$ and Lemma 4 shows $dN/d\tau = 0$. Using the welfare expression (1.2.5), we have $d\ln W/d\ln\left(1+\tau^{\frac{\beta}{\beta-1}}\right) = (\beta-1)/\beta d\ln\Lambda_c/d\ln\left(1+\tau^{\frac{\beta}{\beta-1}}\right) = (1-\beta)/\beta$. When global production networks are IP,
$d\ln\Lambda_m/d\ln\left(1+\tau^{\frac{\beta}{\beta-1}}\right) = 2\ln N/d\ln\left(1+\tau^{\frac{\beta}{\beta-1}}\right) - d\ln\tau^{\frac{\beta}{\beta-1}}/d\ln\left(1+\tau^{\frac{\beta}{\beta-1}}\right)$.
Changes in welfare then follow from (1.2.5) as

$$\begin{aligned}\frac{d\ln W}{d\ln\left(1+\tau^{\frac{\beta}{\beta-1}}\right)} &= \frac{\beta-1}{\beta}\frac{d\ln\Lambda_c}{d\ln\left(1+\tau^{\frac{\beta}{\beta-1}}\right)} + \left(\frac{1-\sigma}{\sigma}\right)\left(\frac{\beta-1}{\beta}\right)\frac{d\ln\Lambda_m}{d\ln\left(1+\tau^{\frac{\beta}{\beta-1}}\right)} \\ &\quad + \frac{1-\beta}{\sigma\beta}\frac{d\ln N}{d\ln\left(1+\tau^{\frac{\beta}{\beta-1}}\right)} \\ &= \frac{1-\beta}{\beta} + \left(\frac{1-\sigma}{\sigma}\right)\left(\frac{1-\beta}{\beta}\right)\frac{1+\tau^{\frac{\beta}{\beta-1}}}{\tau^{\frac{\beta}{\beta-1}}}\left[1-\left(\frac{1-2\sigma}{1-\sigma}\right)\frac{d\ln N}{d\ln\tau^{\frac{\beta}{\beta-1}}}\right]\end{aligned}$$

where $d\ln N/d\ln\tau^{\frac{\beta}{\beta-1}} = -\kappa N\tau^{\frac{\beta}{1-\beta}}/\left[2\kappa N\left(\tau^{\frac{\beta}{1-\beta}}-1\right)-\nu\right]$ by applying the implicit function theorem to (1.9.2). Thus $d\ln W/d\ln\left(1+\tau^{\frac{\beta}{\beta-1}}\right) > (1-\beta)/\beta$ provided that $(1-2\sigma)\,d\ln N/d\ln\tau^{\frac{\beta}{\beta-1}} < 1-\sigma$, which is equivalent to $\left[\tau^{\frac{\beta}{1-\beta}} - 2(1-\sigma)\right]\kappa N < (1-\sigma)\nu$. Furthermore, applying the implicit function theorem to equation (1.9.2) yields $dN/d\kappa = \theta N^2\left(\tau^{\frac{\beta}{1-\beta}}-1\right)\times \left[(\nu\theta)^2 - 4\kappa\theta(1-\alpha)L\left(\tau^{\frac{\beta}{1-\beta}}-1\right)\right]^{-\frac{1}{2}} > 0$. Therefore, a sufficient condition for $d\ln W/d\ln\left(1+\tau^{\frac{\beta}{\beta-1}}\right) > (1-\beta)/\beta$ is that the inequality $\left[\tau^{\frac{\beta}{1-\beta}} - 2(1-\sigma)\right]\kappa N < (1-\sigma)\nu$ holds when $\kappa = \overline{\kappa}_F$. At $\kappa = \overline{\kappa}_F$, total firm mass $N_{|\overline{\kappa}_F} = \left[1-\tau^{\frac{\beta}{\beta-1}}\alpha(1-\sigma)\right](1-\alpha)L/\{\nu\theta[1-\alpha(1-\sigma)]\}$. Rearranging $\left[\tau^{\frac{\beta}{1-\beta}} - 2(1-\sigma)\right]\overline{\kappa}_F N_{|\overline{\kappa}_F} < (1-\sigma)\nu$ yields $\alpha < 1+\alpha(1-\sigma)\tau^{\frac{\beta}{\beta-1}}$, which holds by assumption. Finally, when global production networks are

IC, domestic expenditure shares for final goods and intermediate inputs are the same: $\Lambda_m = \Lambda_c$. By Lemma 4 $dN/d\tau = 0$ in this case, and changes in welfare are given by $d\ln W/d\ln\left(1+\tau^{\frac{\beta}{\beta-1}}\right) = (1-\beta)/(\sigma\beta) > (1-\beta)/\beta$ for all $\sigma < 1$.

Proof of Proposition 4

The aggregate domestic expenditure share Λ is defined as the proportion of aggregate expenditure that goes to domestically produced goods, including both final goods and intermediate inputs: $\Lambda = (1-\iota)\Lambda_c + \iota\Lambda_m$. We can rewrite Λ as below

$$\begin{aligned}
\Lambda &= \frac{1-\iota}{1+\tau^{\frac{\beta}{\beta-1}}} + \frac{\iota\mu_H}{\mu_H + \tau^{\frac{\beta}{\beta-1}}\mu_F} \\
&= \frac{1}{1+\tau^{\frac{\beta}{\beta-1}}}\left[1+\iota\frac{\tau^{\frac{\beta}{\beta-1}}\mu_F}{\mu_H+\tau^{\frac{\beta}{\beta-1}}\mu_F}\left(\frac{\mu_H}{\mu_F}-1\right)\right] \\
&= \Lambda_c\left[1+\iota(1-\Lambda_m)\left(\frac{\mu_H}{\mu_F}-1\right)\right]
\end{aligned}$$

which implies

$$\frac{1-\Lambda}{\Lambda} = \frac{1-\Lambda_c}{\Lambda_c}\left[\frac{1-\iota\Lambda_m\left(1-\frac{\mu_F}{\mu_H}\right)}{1+\iota(1-\Lambda_m)\left(\frac{\mu_H}{\mu_F}-1\right)}\right]$$

Thus, (partial) trade elasticity is given by

$$\begin{aligned}
\varepsilon &\equiv \frac{\partial\ln\left(\frac{1-\Lambda}{\Lambda}\right)}{\partial\ln\tau}|_{\mu_F} \\
&= \frac{\beta}{\beta-1} + \frac{\partial\ln\left[\frac{1-\iota\Lambda_m\left(1-\frac{\mu_F}{\mu_H}\right)}{1+\iota(1-\Lambda_m)\left(\frac{\mu_H}{\mu_F}-1\right)}\right]}{\partial\ln\tau}|_{\mu_F} \\
&= \frac{\beta}{\beta-1}\left[1+\frac{\iota\Lambda_m(1-\Lambda_m)\left(1-\frac{\mu_F}{\mu_H}\right)}{1-\iota\Lambda_m\left(1-\frac{\mu_F}{\mu_H}\right)} - \frac{\iota\Lambda_m(1-\Lambda_m)\left(\frac{\mu_H}{\mu_F}-1\right)}{1+\iota(1-\Lambda_m)\left(\frac{\mu_H}{\mu_F}-1\right)}\right]
\end{aligned}$$

Substituting the expressions of Λ and ε into the welfare expression (1.2.5), we have

$$
\begin{aligned}
d\ln W &= \left(\frac{\beta-1}{\beta}\right)\left[d\ln\Lambda_c + d\ln\left(\Lambda_m^{\frac{1-\sigma}{\sigma}} N^{-\frac{1}{\sigma}}\right)\right] \\
&= \left(\frac{1+\chi_1}{\varepsilon}\right)\left\{d\ln\left[\frac{\Lambda}{1+\iota\left(1-\Lambda_m\right)\left(\frac{\mu_H}{\mu_F}-1\right)}\right] + d\ln\left(\Lambda_m^{\frac{1-\sigma}{\sigma}} N^{-\frac{1}{\sigma}}\right)\right\} \\
&= \left(\frac{1+\chi_1}{\varepsilon}\right)\left(d\ln\Lambda + d\ln\chi_2\right)
\end{aligned}
$$

As $\mu_F \to 0$ (hence $\widehat{\mu} \to 0$), we have $\Lambda_m = 1$, $\chi_1 = -\left[\iota/\left(1-\iota\right)\right]\left(1-\Lambda\right)/\Lambda$, $d\ln N = 0$ (by Lemma 4), and $d\ln\chi_2 = \left[\iota/\left(\Lambda-\iota\right)\right] d\ln\Lambda$. Substituting these results into equation (1.2.6) yields $d\ln W = d\ln\Lambda/\left[\varepsilon\left(1-\iota\right)\right]$. As $\mu_F \to 1$ (hence $\widehat{\mu} \to 1$), we have $\Lambda_m = \Lambda$, $\chi_1 = 0$, $d\ln N = 0$ (by Lemma 4), and $d\ln\chi_2 = \left[\left(1-\sigma\right)/\sigma\right] d\ln\Lambda$. Substituting these results into equation (1.2.6) yields $d\ln W = d\ln\Lambda/\left(\varepsilon\sigma\right)$.

Proof of Lemma 5

If in equilibrium $b_H\left(\theta\right) > \underline{\theta}$, by the first order conditions for linkage formation, the marginal domestic and foreign customers (whose productivity levels coincide with the three cutoffs) for a type-θ supplier should satisfy

$$
\begin{aligned}
r\left(b_H\left(\theta\right)\right) P^S\left(b_H\left(\theta\right)\right)^{\frac{\alpha}{1-\alpha}} p\left(\theta\right)^{\frac{\alpha}{\alpha-1}} N\left(\theta\right)^{\frac{1}{\beta}\left(\frac{\alpha-\beta}{1-\alpha}\right)} \theta^{-1} &= \frac{\kappa}{\alpha\left(1-\alpha\right)\left(1-\sigma\right)} \\
r\left(\underline{b}_F\left(\theta\right)\right) P^S\left(\underline{b}_F\left(\theta\right)\right)^{\frac{\alpha}{1-\alpha}} p\left(\theta\right)^{\frac{\alpha}{\alpha-1}} N\left(\theta\right)^{\frac{1}{\beta}\left(\frac{\alpha-\beta}{1-\alpha}\right)} \theta^{-1} &= \tau^{\frac{\beta}{1-\beta}} \frac{\kappa}{\alpha\left(1-\alpha\right)\left(1-\sigma\right)} \\
r\left(\overline{b}_F\left(\theta\right)\right) P^S\left(\overline{b}_F\left(\theta\right)\right)^{\frac{\alpha}{1-\alpha}} p\left(\theta\right)^{\frac{\alpha}{\alpha-1}} N\left(\theta\right)^{\frac{1}{\beta}\left(\frac{\alpha-\beta}{1-\alpha}\right)} \theta^{-1} &= \tau^{\frac{\beta}{1-\beta}} \left(1+\tau^{\frac{\beta}{\beta-1}}\right)^{\frac{1}{\beta}\left(\frac{\beta-\alpha}{1-\alpha}\right)} \times \\
&\quad \frac{\kappa}{\alpha\left(1-\alpha\right)\left(1-\sigma\right)}
\end{aligned}
$$

which implies that

$$
\begin{aligned}
&\frac{r\left(\underline{b}_F\left(\theta\right)\right) P^S\left(\underline{b}_F\left(\theta\right)\right)^{\frac{\alpha}{1-\alpha}}}{r\left(b_H\left(\theta\right)\right) P^S\left(b_H\left(\theta\right)\right)^{\frac{\alpha}{1-\alpha}}} = \tau^{\frac{\beta}{1-\beta}} \quad \text{and} \\
&\frac{r\left(\overline{b}_F\left(\theta\right)\right) P^S\left(\overline{b}_F\left(\theta\right)\right)^{\frac{\alpha}{1-\alpha}}}{r\left(b_H\left(\theta\right)\right) P^S\left(b_H\left(\theta\right)\right)^{\frac{\alpha}{1-\alpha}}} = \tau^{\frac{\beta}{1-\beta}} \left(1+\tau^{\frac{\beta}{\beta-1}}\right)^{\frac{1}{\beta}\left(\frac{\beta-\alpha}{1-\alpha}\right)} \qquad (1.9.3)
\end{aligned}
$$

Therefore, the productivity cutoffs for foreign customers $\underline{b}_F\left(\theta\right)$ and $\overline{b}_F\left(\theta\right)$ depend on the supplier type θ only through their dependence on the productivity cutoff for domestic customers $b_H(\theta)$, as given by the above two relationships. Rewriting the free entry condition using the variety market clearing condition and the first order conditions for linkage formation, we have

$$(1-\alpha)\,p(\theta)^{\frac{\alpha}{\alpha-1}}\,N(\theta)^{\frac{\beta-\alpha}{\beta(\alpha-1)}}\,\theta^{-1}\left[\left(1+\tau^{\frac{\beta}{\beta-1}}\right)^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)} XP^{\frac{1}{1-\alpha}}+\mathscr{I}_1\left(b_H(\theta)\right)\right] -\kappa\mathscr{I}_2\left(b_H(\theta)\right)=\nu$$

where we save notation by defining the following sums of integrals:

$$\begin{aligned}
\mathscr{I}_1\left(b_H(\theta)\right) \equiv & \int_{b_H(\theta)}^{\underline{b}_F(\theta)} \alpha(1-\sigma)\,r(\theta')\,P^S(\theta')^{\frac{\alpha}{1-\alpha}}\,N(\theta')\,g(\theta')\,d\theta' \\
& +\left(1+\tau^{\frac{\beta}{\beta-1}}\right)^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)} \int_{\overline{b}_F(\theta)}^{\overline{\theta}} \alpha(1-\sigma)\,r(\theta')\,P^S(\theta')^{\frac{\alpha}{1-\alpha}}\,N(\theta')\,g(\theta')\,d\theta' \\
\mathscr{I}_2\left(b_H(\theta)\right) \equiv & \int_{b_H(\theta)}^{\underline{b}_F(\theta)} N(\theta')\,g(\theta')\,d\theta' + \left(1-\tau^{\frac{\beta}{1-\beta}}\right)\int_{\underline{b}_F(\theta)}^{\overline{b}_F(\theta)} N(\theta')\,g(\theta')\,d\theta' \\
& +2\int_{\overline{b}_F(\theta)}^{\overline{\theta}} N(\theta')\,g(\theta')\,d\theta'
\end{aligned}$$

By the previous argument, these two sums of integrals above depend on the supplier type θ only via their dependence on $b_H(\theta)$. Since the above equation must hold for all values of θ, it must be that $b_H(\theta)$ and $p(\theta)^{\frac{\alpha}{\alpha-1}}N(\theta)^{\frac{\beta-\alpha}{\beta(\alpha-1)}}\theta^{-1}$ are both constant across firm types: for all $\theta\in\Theta$, $b_H(\theta)=b_H$ and $p(\theta)^{\frac{\alpha}{\alpha-1}}N(\theta)^{\frac{\beta-\alpha}{\beta(\alpha-1)}}\theta^{-1}=D$ where b_H and D are some constants to be solved for in equilibrium. Accordingly, $\underline{b}_F(\theta)$ and $\overline{b}_F(\theta)$ must also be independent of the supplier type: for all $\theta\in\Theta$, $\underline{b}_F(\theta)=\underline{b}_F$ and $\overline{b}_F(\theta)=\overline{b}_F$. Substituting these results back into the variety market clearing condition shows that firm revenue is proportional to firm productivity: $r(\theta)=B\theta$, where B is a constant depending on aggregate variables only and is to be solved for in equilibrium. Furthermore, the constant cutoffs for customer selection imply that $\mu_H(b_H,\theta)=\mu_H(\underline{b}_F,\theta)=\mu_H\left(\overline{b}_F,\theta\right)=1$, $\mu_F(b_H,\theta)=\mu_F(\underline{b}_F,\theta)=0$, and $\mu_F\left(\overline{b}_F,\theta\right)=1$ for all supplier type $\theta\in\Theta$. Therefore, the producer price index for the marginal customer firms with productivity level $\theta=b_H$ is given by

$$\begin{aligned}
P^S(b_H)^{\frac{\alpha}{\alpha-1}} &= \int_\Theta p(\theta')^{\frac{\alpha}{\alpha-1}}\left\{\left[\mu_H(b_H,\theta')+\tau^{\frac{\beta}{\beta-1}}\mu_F(b_H,\theta')\right]N(\theta')\right\}^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)} g(\theta')\,d\theta' \\
&= DA
\end{aligned}$$

where $A\equiv\int_\Theta \theta N(\theta)\,g(\theta)\,d\theta$ is aggregate productivity.

Similarly, we can derive $P^S(\underline{b}_F)^{\frac{\alpha}{\alpha-1}} = DA$ and $P^S(\overline{b}_F)^{\frac{\alpha}{\alpha-1}} = \left(1+\tau^{\frac{\beta}{\beta-1}}\right)^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)} DA$. Substituting the expressions of firm revenue and producer price index into the equations (1.9.3) yields $\underline{b}_F = \tau^{\frac{\beta}{1-\beta}} b_H$, *and* $\overline{b}_F = \left(1+\tau^{\frac{\beta}{1-\beta}}\right) b_H$. Finally, substituting $r(\theta) = B\theta$ into the variety market clearing condition allows us to solve for the constant term $B = L/\{[1-\alpha(1-\sigma)]A\}$, which then allows us to deduce from the equation at the beginning of this proof that $b_H = \kappa[1-\alpha(1-\sigma)]A^2/[\alpha(1-\alpha)(1-\sigma)L]$.

Proof of Lemma 6

Firm revenue $r(\theta) = \theta L/\{[1-\alpha(1-\sigma)]A\}$ follows directly from the proof of Lemma 5. Variety prices $p(\theta)$ can be derived by substituting the linkage first order conditions, the monopolistic pricing condition $p(\theta) = P^S(\theta)^{1-\sigma}/(\alpha\theta)$, and $p(\theta)^{\frac{\alpha}{\alpha-1}} N(\theta)^{\frac{\beta-\alpha}{\beta(\alpha-1)}} \theta^{-1} = D$ (see the proof of Lemma 5) into the definition of the producer price index (1.3.1). The number of firms in each industry $N(\theta)$ then follows from $N(\theta) = D^{\frac{\beta(1-\alpha)}{\alpha-\beta}} p(\theta)^{\frac{\alpha\beta}{\alpha-\beta}} \theta^{-\frac{\beta(1-\alpha)}{\beta-\alpha}}$. The constant term D can be solved from the definition of aggregate productivity $A \equiv \int_\Theta \theta N(\theta) g(\theta) d\theta = \int_{\underline{\theta}}^{b_H} \theta N(\theta) g(\theta) d\theta + \int_{b_H}^{\underline{b}_F} \theta N(\theta) g(\theta) d\theta + \int_{\underline{b}_F}^{\overline{b}_F} \theta N(\theta) g(\theta) d\theta + \int_{\overline{b}_F}^{\overline{\theta}} \theta N(\theta) g(\theta) d\theta$, which implies that

$$D = \left\{\frac{\kappa[1-\alpha(1-\sigma)]A^2}{\alpha(1-\alpha)(1-\sigma)L}\right\}^{-\frac{\alpha(1-\beta)}{\beta(1-\alpha)}\left(\frac{1-\sigma}{\sigma}\right)} \left[\frac{A^{1-(1-\sigma)\frac{\beta(1-\alpha)}{\beta-\alpha}}}{\mathscr{I}_3(b_H)}\right]^{-\frac{\beta-\alpha}{\sigma\beta(1-\alpha)}} \alpha^{-\frac{1}{\sigma}}$$

where we save notation by defining the following sum of integrals:

$$\begin{aligned}
\mathscr{I}_3(b_H) \equiv & \int_{\underline{\theta}}^{b_H} \theta^{\frac{\alpha\beta}{\beta-\alpha}-\sigma\frac{\alpha(1-\beta)}{\beta-\alpha}} g(\theta) d\theta \\
& + \left\{\frac{\kappa[1-\alpha(1-\sigma)]A^2}{\alpha(1-\alpha)(1-\sigma)L}\right\}^{\frac{\alpha(1-\beta)}{\beta-\alpha}(1-\sigma)} \int_{b_H}^{\underline{b}_F} \theta^{\frac{\alpha\beta}{\beta-\alpha}-\frac{\alpha(1-\beta)}{\beta-\alpha}} g(\theta) d\theta \\
& + \left(\tau^{\frac{\beta}{\beta-1}}\right)^{\frac{\alpha(1-\beta)}{\beta-\alpha}(1-\sigma)} \int_{\underline{b}_F}^{\overline{b}_F} \theta^{\frac{\alpha\beta}{\beta-\alpha}-\sigma\frac{\alpha(1-\beta)}{\beta-\alpha}} g(\theta) d\theta \\
& + \left(1+\tau^{\frac{\beta}{\beta-1}}\right)^{\frac{\alpha(1-\beta)}{\beta-\alpha}(1-\sigma)} \left\{\frac{\kappa[1-\alpha(1-\sigma)]A^2}{\alpha(1-\alpha)(1-\sigma)L}\right\}^{\frac{\alpha(1-\beta)}{\beta-\alpha}(1-\sigma)} \times \\
& \int_{\overline{b}_F}^{\overline{\theta}} \theta^{\frac{\alpha\beta}{\beta-\alpha}-\frac{\alpha(1-\beta)}{\beta-\alpha}} g(\theta) d\theta
\end{aligned}$$

The above integral sum (and hence the constant term D) only depends on two equilibrium unknowns: b_H and A, because the relationship between the other two cutoffs ($\underline{b}_F$ and $\overline{b}_F$) and b_H is readily given by Lemma 5. Henceforth I use the notation $D(b_H,A)$ to make clear its dependence on the two equilibrium unknowns. Finally, the two common terms $\Psi_P(b_H,A)$ and $\Psi_N(b_H,A)$ from the equilibrium variety prices and firm mass respectively are given as below:

$$\Psi_P(b_H,A) \equiv \left\{\frac{\kappa[1-\alpha(1-\sigma)]A^2}{\alpha(1-\alpha)(1-\sigma)L}\right\}^{\left(\frac{1-\beta}{\beta}\right)(1-\sigma)} [D(b_H,A)A]^{-\left(\frac{1-\alpha}{\alpha}\right)(1-\sigma)} \alpha^{-1}$$

$$\Psi_N(b_H,A) \equiv \left\{\frac{\kappa[1-\alpha(1-\sigma)]A^2}{\alpha(1-\alpha)(1-\sigma)L}\right\}^{-\frac{\alpha(1-\beta)}{\beta-\alpha}(1-\sigma)} \left[D(b_H,A)A^{-\left(\frac{1-\sigma}{\sigma}\right)}\right]^{-\frac{\beta(1-\alpha)}{\beta-\alpha}\sigma} \alpha^{\frac{\alpha\beta}{\beta-\alpha}}$$

Proof of Lemma 7

The interior solutions of $\mu_H(\theta',\theta)$ and $\mu_F(\theta',\theta)$ can be derived by substituting the definition of producer price index (1.3.1), equilibrium firm revenue $r(\theta) = \theta L/\{[1-\alpha(1-\sigma)]A\}$, and $p(\theta)^{\frac{\alpha}{\alpha-1}} N(\theta)^{\frac{\beta-\alpha}{\beta(\alpha-1)}} \theta^{-1} = D$ (see the proof of Lemma 5) into the linkage first order conditions. The partition of the productivity support $\Theta \equiv \left[\underline{\theta},\overline{\theta}\right]$ into intervals of interior/corner solutions follows directly from Lemma 5.

Proof of Proposition 5

First, we express a type-θ firm's total number of downstream links in terms of the two equilibrium unknowns b_H and A using the results about equilibrium firm mass $N(\theta)$ and linkage densities from Lemma 6 and 7:

$$\begin{aligned} z(b_H,A) &\equiv \int_\Theta \left[\mu_H(\theta',\theta) + \mu_F(\theta',\theta)\right] N(\theta') g(\theta') d\theta' \\ &= \int_{\underline{\theta}}^{b_H} \frac{\alpha(1-\alpha)(1-\sigma)L\theta'}{\kappa[1-\alpha(1-\sigma)]A^2} N(\theta') g(\theta') d\theta' + \int_{b_H}^{\underline{b}_F} N(\theta') g(\theta') d\theta' \\ &+ \int_{\underline{b}_F}^{\overline{b}_F} \left\{1 + \frac{\alpha(1-\alpha)(1-\sigma)L\theta'}{\kappa[1-\alpha(1-\sigma)]A^2} - \tau^{\frac{\beta}{1-\beta}}\right\} N(\theta') g(\theta') d\theta' \\ &+ 2\int_{\overline{b}_F}^{\overline{\theta}} N(\theta') g(\theta') d\theta' \end{aligned}$$

$$\Leftrightarrow z(b_H,A) = \Psi_N(b_H,A)\frac{\alpha(1-\alpha)(1-\sigma)L}{\kappa[1-\alpha(1-\sigma)]A^2}\times$$
$$\int_{\underline{\theta}}^{b_H}(\theta')^{\frac{\alpha(2\beta-1)}{\beta-\alpha}+\frac{\alpha(1-\beta)}{\beta-\alpha}(1-\sigma)}g(\theta')\,d\theta'$$
$$+\Psi_N(b_H,A)\left\{\frac{\kappa[1-\alpha(1-\sigma)]A^2}{\alpha(1-\alpha)(1-\sigma)L}\right\}^{\frac{\alpha(1-\beta)}{\beta-\alpha}(1-\sigma)}\times$$
$$\int_{b_H}^{\underline{b}_F}(\theta')^{\frac{(2\alpha-1)\beta}{\beta-\alpha}}g(\theta')\,d\theta'$$
$$+\Psi_N(b_H,A)\,\tau^{\frac{\alpha\beta}{\alpha-\beta}(1-\sigma)}\times$$
$$\int_{\underline{b}_F}^{\overline{b}_F}\left\{1+\frac{\alpha(1-\alpha)(1-\sigma)L\theta'}{\kappa[1-\alpha(1-\sigma)]A^2}-\tau^{\frac{\beta}{1-\beta}}\right\}(\theta')^{\frac{(2\alpha-1)\beta}{\beta-\alpha}+\frac{\alpha(1-\beta)}{\beta-\alpha}(1-\sigma)}g(\theta')\,d\theta'$$
$$+\Psi_N(b_H,A)\,2\left[\left(1+\tau^{\frac{\beta}{\beta-1}}\right)\frac{\kappa[1-\alpha(1-\sigma)]A^2}{\alpha(1-\alpha)(1-\sigma)L}\right]^{\frac{\alpha(1-\beta)}{\beta-\alpha}(1-\sigma)}\times$$
$$\int_{\overline{b}_F}^{\overline{\theta}}(\theta')^{\frac{(2\alpha-1)\beta}{\beta-\alpha}}g(\theta')\,d\theta'$$

If linkage fixed costs are sufficiently high ($\kappa > \overline{\kappa}$), b_H has an interior solution ($\underline{\theta} < b_H < \infty$). In this case, the CS condition is given by $b_H = \kappa[1-\alpha(1-\sigma)]A^2/[\alpha(1-\alpha)(1-\sigma)L]$, which can be substituted into the FE condition to obtain the following equation with b_H as the only equilibrium unknown:

$$\sqrt{\frac{[1-\alpha(1-\sigma)]\,\nu^2}{\kappa\alpha(1-\alpha)(1-\sigma)L}b_H}+\frac{T_1}{T_2-\left(\frac{b_H}{\underline{\theta}}\right)^{\zeta-\frac{\alpha\beta}{\beta-\alpha}+\sigma\frac{\alpha(1-\beta)}{\beta-\alpha}}}-\frac{1-\alpha(1-\sigma)}{\alpha(1-\sigma)}=0 \tag{1.9.4}$$

where we save notation by defining below two expressions T_1 and T_2 that summarize the impact of variable trade costs τ:

$$\frac{T_1}{\frac{\alpha\beta}{\beta-\alpha}-\sigma\frac{\alpha(1-\beta)}{\beta-\alpha}-\zeta}\equiv\left(\tau^{\frac{\beta}{\beta-1}}\right)^{-\left[\frac{\alpha\beta}{\beta-\alpha}-\frac{\alpha(1-\beta)}{\beta-\alpha}-\zeta\right]}\left(1+\tau^{\frac{\beta}{\beta-1}}\right)^{\frac{\alpha\beta}{\beta-\alpha}-\sigma\frac{\alpha(1-\beta)}{\beta-\alpha}-\zeta}\times$$
$$\left[\frac{1}{\frac{\alpha\beta}{\beta-\alpha}-\frac{\alpha(1-\beta)}{\beta-\alpha}-\zeta}-\frac{2}{\frac{\alpha\beta}{\beta-\alpha}-\frac{\beta(1-\alpha)}{\beta-\alpha}-\zeta}\left(\frac{\tau^{\frac{\beta}{\beta-1}}}{1+\tau^{\frac{\beta}{\beta-1}}}\right)\right]$$
$$+\frac{\left(\tau^{\frac{\beta}{\beta-1}}\right)^{-\left[\frac{\alpha\beta}{\beta-\alpha}-\frac{\beta(1-\alpha)}{\beta-\alpha}-\zeta\right]}-1}{\frac{\alpha\beta}{\beta-\alpha}-\frac{\beta(1-\alpha)}{\beta-\alpha}-\zeta}-\frac{\left(\tau^{\frac{\beta}{\beta-1}}\right)^{-\left[\frac{\alpha\beta}{\beta-\alpha}-\frac{\alpha(1-\beta)}{\beta-\alpha}-\zeta\right]}-1}{\frac{\alpha\beta}{\beta-\alpha}-\frac{\alpha(1-\beta)}{\beta-\alpha}-\zeta}$$

$$-\left(\tau^{\frac{\beta}{\beta-1}}\right)^{-\left[\frac{\alpha\beta}{\beta-\alpha}-\frac{\alpha(1-\beta)}{\beta-\alpha}-\zeta\right]}\left(1-\tau^{\frac{\beta}{\beta-1}}\right)\times$$
$$\left[\frac{\left(1+\tau^{\frac{\beta}{\beta-1}}\right)^{\frac{\alpha\beta}{\beta-\alpha}-\sigma\frac{\alpha(1-\beta)}{\beta-\alpha}-\zeta-1}-1}{\frac{\alpha\beta}{\beta-\alpha}-\sigma\frac{\alpha(1-\beta)}{\beta-\alpha}-\zeta-1}\right]$$

$$\frac{T_2}{\frac{\alpha\beta}{\beta-\alpha}-\sigma\frac{\alpha(1-\beta)}{\beta-\alpha}-\zeta}\equiv\left\{\left(\tau^{\frac{\beta}{\beta-1}}\right)^{-\left[\frac{\alpha\beta}{\beta-\alpha}-\frac{\alpha(1-\beta)}{\beta-\alpha}-\zeta\right]}\left[\left(1+\tau^{\frac{\beta}{\beta-1}}\right)^{\frac{\alpha\beta}{\beta-\alpha}-\sigma\frac{\alpha(1-\beta)}{\beta-\alpha}-\zeta}-1\right]+1\right\}\times$$
$$\left[\frac{1}{\frac{\alpha\beta}{\beta-\alpha}-\sigma\frac{\alpha(1-\beta)}{\beta-\alpha}-\zeta}-\frac{1}{\frac{\alpha\beta}{\beta-\alpha}-\frac{\alpha(1-\beta)}{\beta-\alpha}-\zeta}\right]$$

Let $F(b_H)$ denote the left hand side of the equilibrium condition (1.9.4) as a function of b_H. $F(b_H)$ is increasing in b_H ($F'(b_H) > 0$) because $\left[\zeta-\frac{\alpha\beta}{\beta-\alpha}+\sigma\frac{\alpha(1-\beta)}{\beta-\alpha}\right]T_1$ can be shown to always be positive. In addition, we can establish the following limits:

$$\lim_{b_H\to\underline{\theta}} F(b_H) = \sqrt{\frac{[1-\alpha(1-\sigma)]\,v^2}{\kappa\alpha(1-\alpha)(1-\sigma)L}\underline{\theta}}+\frac{T_1}{T_2-1}-\frac{1-\alpha(1-\sigma)}{\alpha(1-\sigma)}$$
$$\lim_{b_H\to\infty} F(b_H) = \infty$$

Therefore, an unique solution exists for $b_H\in(\underline{\theta},\infty)$, provided that

$$\sqrt{\frac{[1-\alpha(1-\sigma)]\,v^2}{\kappa\alpha(1-\alpha)(1-\sigma)L}\underline{\theta}}+\frac{T_1}{T_2-1}-\frac{1-\alpha(1-\sigma)}{\alpha(1-\sigma)}<0$$

which is equivalent to

$$\kappa > \frac{\underline{\theta}[1-\alpha(1-\sigma)]\,v^2}{\alpha(1-\alpha)(1-\sigma)L}\left[\frac{1-\alpha(1-\sigma)}{\alpha(1-\sigma)}-\frac{T_1}{T_2-1}\right]^{-2}=\overline{\kappa}$$

If linkage fixed costs are sufficiently low ($\kappa\leqslant\overline{\kappa}$), b_H has a corner solution ($b_H=\underline{\theta}$). In this case, substituting the CS condition $b_H=\underline{\theta}$ into the FE condition yields the following equation with aggregate productivity A as the only equilibrium unknown:

$$\frac{(1-\alpha)L}{[1-\alpha(1-\sigma)]A}-\kappa z(\underline{\theta},A) = v$$

In the autarky limit ($\tau \to \infty$), the above equilibrium condition for aggregate productivity A reduces to

$$\frac{\kappa}{\underline{\theta}}\left[\frac{\frac{\alpha\beta}{\beta-\alpha}-\frac{\alpha(1-\beta)}{\beta-\alpha}-\zeta}{\frac{\alpha\beta}{\beta-\alpha}-\frac{\beta(1-\alpha)}{\beta-\alpha}-\zeta}\right]A^2+\nu A-\frac{(1-\alpha)L}{[1-\alpha(1-\sigma)]}=0 \qquad (1.9.5)$$

which admits an unique solution for $A \in (0,\infty)$.

Proof of Proposition 6

Since $\mu_H(\theta,\theta')$ and $\mu_F(\theta,\theta')$ are independent of the supplier type θ', using the monopolistic pricing rule $p(\theta) = w^{\sigma}P^S(\theta)^{1-\sigma}/(\alpha\theta)$ we can rewrite the definition of producer price index (1.3.1) as follows:

$$[\alpha\theta p(\theta)]^{\frac{1}{1-\sigma}\left(\frac{\alpha}{\alpha-1}\right)} = \left[\mu_H(\theta,\theta')+\tau^{\frac{\beta}{\beta-1}}\mu_F(\theta,\theta')\right]^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)} \times \int_{\Theta} p(\theta')^{\frac{\alpha}{\alpha-1}} N(\theta')^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)} g(\theta')\,d\theta'$$

$$\Leftrightarrow \int_{\Theta} p(\theta)^{\frac{\alpha}{\alpha-1}} N(\theta)^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)} g(\theta)\,d\theta =$$

$$\left[\int_{\Theta}(\alpha\theta)^{\frac{\alpha}{1-\alpha}}\left\{\left[\mu_H(\theta,\theta')+\tau^{\frac{\beta}{\beta-1}}\mu_F(\theta,\theta')\right]^{(1-\sigma)} N(\theta)\right\}^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)} g(\theta)\,d\theta\right]^{\frac{1}{\sigma}}$$

Substituting the last equation into the definition of consumer price index P yields the expression of welfare changes:

$$d\ln W = \frac{\beta-1}{\beta}d\ln\Lambda_c + \frac{\beta-1}{\beta}d\ln\left\{\int_{\Theta}\theta^{\frac{\alpha}{1-\alpha}}\left[\Lambda_m(\theta)^{-(1-\sigma)}\mu_H(\theta,\theta')^{(1-\sigma)}N(\theta)\right]^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)} g(\theta)\,d\theta\right\}^{-\frac{\beta}{\sigma\alpha}\left(\frac{1-\alpha}{1-\beta}\right)} \qquad (1.9.6)$$

Substituting the expression of $\Lambda_m(\theta) = \mu_H(\theta,\theta')/\left[\mu_H(\theta,\theta')+\tau^{\frac{\beta}{\beta-1}}\mu_F(\theta,\theta')\right]$ into the identity of the aggregate domestic expenditure share $\Lambda = (1-\iota)\Lambda_c + \iota\Lambda_m$, we have

$$\Lambda = \Lambda_c\int_{\Theta}\left\{1+\iota\left[1-\Lambda_m(\theta)\right]\left[\frac{\mu_H(\theta,\theta')}{\mu_F(\theta,\theta')}-1\right]\right\}\frac{r(\theta)}{\bar{r}}\widehat{g}(\theta)\,d\theta \qquad (1.9.7)$$

Thus, (partial) trade elasticity is given by

$$\begin{aligned}
\varepsilon \equiv & \frac{d\ln\left(\frac{1-\Lambda}{\Lambda}\right)}{d\ln\tau}\Big|_{\mu_H,\mu_F} \\
= & \frac{\beta}{\beta-1} + \frac{d\ln\left\{\frac{\bar{r}-\int_\Theta \iota\Lambda_m(\theta)\left[1-\frac{\mu_F(\theta,\theta')}{\mu_H(\theta,\theta')}\right]r(\theta)\widehat{g}(\theta)d\theta}{\bar{r}+\int_\Theta \iota[1-\Lambda_p(\theta)]\left[\frac{\mu_H(\theta,\theta')}{\mu_F(\theta,\theta')}-1\right]r(\theta)\widehat{g}(\theta)d\theta}\right\}}{d\ln\tau}\Big|_{\mu_H,\mu_F} \\
= & \frac{\beta}{\beta-1}(1+\chi_1) \qquad (1.9.8)
\end{aligned}$$

Substituting (1.9.7) and (1.9.8) back into the expression of welfare changes (1.9.6), we obtain the formula presented in the proposition.

Proof of Lemma 8

If $\kappa > \overline{\kappa}$, b_H has an interior solution given by the unique root to the equation (1.9.4) with constants T_1 and T_2 replaced by their values in autarky: $\lim_{\tau\to\infty} T_1$ and $\lim_{\tau\to\infty} T_2$. Applying the implicit function theorem to this equilibrium condition (1.9.4), we can establish that $0 < d\ln b_H/d\ln L < 1$. The comparative static $dA/dL > 0$ follows directly from the CS condition. Furthermore, for all $\theta \in \Theta$, we have $d\ln r(\theta)/d\ln L = (1 - d\ln b_H/d\ln L)/2 > 0$. If $\kappa \leqslant \overline{\kappa}$, b_H has a corner solution $b_H = \underline{\theta}$, which implies that $db_H/dL = 0$. In this case, aggregate productivity A is given by the unique solution to the equilibrium condition (1.9.5). Applying the implicit function theorem to this equilibrium condition (1.9.5), we can establish that $0 < d\ln A/d\ln L < 1$. Furthermore, for all $\theta \in \Theta$, we have $d\ln r(\theta)/d\ln L = 1 - d\ln A/d\ln L > 0$, $d\ln p(\theta)/d\ln L = (\beta-1)(1-\sigma)/(\beta\sigma)\, d\ln A/d\ln L < 0$, and $d\ln N(\theta)/d\ln L = d\ln A/d\ln L > 0$.

Proof of Proposition 7

If $\kappa \leqslant \overline{\kappa}$, b_H has a corner solution $b_H = \underline{\theta}$. In this case, changes in welfare with respect to market size is given by $d\ln W/d\ln L = (1-\beta)/(\sigma\beta)\, d\ln A/d\ln L$, where the responses of aggregate productivity $d\ln A/d\ln L$ can be obtained by applying the implicit function theorem to the equilibrium condition (1.9.5). The condition for $d\ln W/d\ln L > (1-\beta)/\beta$ is

$$A > \frac{(2\sigma-1)(1-\alpha)L}{[1-\alpha(1-\sigma)]\sigma\nu}$$

which holds if $2\sigma - 1 \leqslant 0$. If $2\sigma - 1 > 0$, the last inequality holds provided that

$$\kappa < \left[\frac{\frac{\alpha\beta}{\beta-\alpha} - \frac{\beta(1-\alpha)}{\beta-\alpha} - \zeta}{\frac{\alpha\beta}{\beta-\alpha} - \frac{\alpha(1-\beta)}{\beta-\alpha} - \zeta} \right] \frac{\sigma(1-\sigma)\left[1-\alpha(1-\sigma)\right]v^2\underline{\theta}}{(1-\alpha)(2\sigma-1)^2 L}$$

which is true since the right hand side is larger than the threshold $\overline{\kappa}$ (whose expression is given in the proof of Proposition 5) in the limit of $\tau \to \infty$.

Proof of Lemma 9

By the CS condition, if b_H has an interior solution in equilibrium ($\underline{\theta} < b_H < \infty$), then $db_H/d\tau$ and $dA/d\tau$ must have the same sign. By the expression of equilibrium firm sales $r(\theta)$ established in Lemma 6, $dr(\theta)/d\tau$ must have the opposite sign of $dA/d\tau$ (and hence of $db_H/d\tau$) for all $\theta \in \Theta$.

Derivation of the value added content of trade (Section 4.1)

Substituting the definition of the direct requirement coefficients $\omega_H^1(\theta',\theta)$ and $\omega_F^1(\theta',\theta)$ into the expressions of the second order requirements, we have

$$\begin{aligned}
\omega_H^2(\theta',\theta) &= \int_\Theta \omega_H^1(\theta',\theta'')\,\omega_H^1(\theta'',\theta)\,g(\theta'')\,d\theta'' \\
&\quad + \int_\Theta \omega_F^1(\theta',\theta'')\,\omega_F^1(\theta'',\theta)\,g(\theta'')\,d\theta'' \\
&= \iota\Lambda_m\omega_H^1(\theta',\theta) + \iota(1-\Lambda_m)\,\omega_F^1(\theta',\theta) \\
\omega_F^2(\theta',\theta) &= \int_\Theta \omega_H^1(\theta',\theta'')\,\omega_F^1(\theta'',\theta)\,g(\theta'')\,d\theta'' \\
&\quad + \int_\Theta \omega_F^1(\theta',\theta'')\,\omega_H^1(\theta'',\theta)\,g(\theta'')\,d\theta'' \\
&= \iota\Lambda_m\omega_F^1(\theta',\theta) + \iota(1-\Lambda_m)\,\omega_H^1(\theta',\theta)
\end{aligned}$$

where

$$\Lambda_m \equiv \frac{\int_\Theta\int_\Theta \omega_H^1(\theta,\theta')\,g(\theta')\,d\theta' r(\theta)N(\theta)\,g(\theta)\,d\theta}{\iota\int_\Theta r(\theta)N(\theta)\,g(\theta)\,d\theta}$$

is the domestic expenditure shares for intermediate inputs. Similarly, we can rewrite the expressions of the third order requirements as below:

$$\begin{aligned}
\omega_H^3(\theta',\theta) \equiv & \int_\Theta\int_\Theta \omega_H^1(\theta',\theta'')\,\omega_H^1(\theta'',\theta''')\,\omega_H^1(\theta''',\theta)\,g(\theta''')\,d\theta'''g(\theta'')\,d\theta'' \\
& + \int_\Theta\int_\Theta \omega_F^1(\theta',\theta'')\,\omega_F^1(\theta'',\theta''')\,\omega_H^1(\theta''',\theta)\,g(\theta''')\,d\theta'''g(\theta'')\,d\theta'' \\
& + \int_\Theta\int_\Theta \omega_H^1(\theta',\theta'')\,\omega_F^1(\theta'',\theta''')\,\omega_F^1(\theta''',\theta)\,g(\theta''')\,d\theta'''g(\theta'')\,d\theta'' \\
& + \int_\Theta\int_\Theta \omega_F^1(\theta',\theta'')\,\omega_H^1(\theta'',\theta''')\,\omega_F^1(\theta''',\theta)\,g(\theta''')\,d\theta'''g(\theta'')\,d\theta'' \\
& = (\iota\Lambda_m)^2\,\omega_H^1(\theta',\theta) + 2\iota^2\Lambda_m(1-\Lambda_m)\,\omega_F^1(\theta',\theta) + [\iota(1-\Lambda_m)]^2\,\omega_H^1(\theta',\theta)
\end{aligned}$$

$$\begin{aligned}
\omega_F^3(\theta',\theta) \equiv & \int_\Theta\int_\Theta \omega_F^1(\theta',\theta'')\,\omega_H^1(\theta'',\theta''')\,\omega_H^1(\theta''',\theta)\,g(\theta''')\,d\theta'''g(\theta'')\,d\theta'' \\
& + \int_\Theta\int_\Theta \omega_H^1(\theta',\theta'')\,\omega_F^1(\theta'',\theta''')\,\omega_H^1(\theta''',\theta)\,g(\theta''')\,d\theta'''g(\theta'')\,d\theta'' \\
& + \int_\Theta\int_\Theta \omega_H^1(\theta',\theta'')\,\omega_H^1(\theta'',\theta''')\,\omega_F^1(\theta''',\theta)\,g(\theta''')\,d\theta'''g(\theta'')\,d\theta'' \\
& + \int_\Theta\int_\Theta \omega_F^1(\theta',\theta'')\,\omega_F^1(\theta'',\theta''')\,\omega_F^1(\theta''',\theta)\,g(\theta''')\,d\theta'''g(\theta'')\,d\theta'' \\
& = (\iota\Lambda_m)^2\,\omega_F^1(\theta',\theta) + 2\iota^2\Lambda_m(1-\Lambda_m)\,\omega_H^1(\theta',\theta) + [\iota(1-\Lambda_m)]^2\,\omega_F^1(\theta',\theta)
\end{aligned}$$

In general, we can write the n-th order requirement coefficients as:

$$\omega_H^n(\theta',\theta) = \sum_{k=0}^{n-1}\binom{n-1}{k}(\iota\Lambda_m)^{n-1-k}[\iota(1-\Lambda_m)]^k\left\{\omega_H^1(\theta',\theta)[1-\mathbb{I}(k\text{ is odd})]+\omega_F^1(\theta',\theta)\,\mathbb{I}(k\text{ is odd})\right\}$$

$$\omega_F^n(\theta',\theta) = \sum_{k=0}^{n-1}\binom{n-1}{k}(\iota\Lambda_m)^{n-1-k}[\iota(1-\Lambda_m)]^k\left\{\omega_H^1(\theta',\theta)\,\mathbb{I}(k\text{ is odd})+\omega_F^1(\theta',\theta)[1-\mathbb{I}(k\text{ is odd})]\right\}$$

Summing up all orders of requirements, we have

$$\begin{aligned}
\sum_{n=1}^{\infty}\omega_H^n(\theta',\theta) &= \\
\sum_{n=1}^{\infty}\sum_{k=0}^{n-1}\binom{n-1}{k}(\iota\Lambda_m)^{n-1-k}[\iota(1-\Lambda_m)]^k\left\{\omega_H^1(\theta',\theta)[1-\mathbb{I}(k\text{ is odd})]+\omega_F^1(\theta',\theta)\,\mathbb{I}(k\text{ is odd})\right\} &= \\
\sum_{k=0}^{\infty}\left\{\omega_H^1(\theta',\theta)[1-\mathbb{I}(k\text{ is odd})]+\omega_F^1(\theta',\theta)\,\mathbb{I}(k\text{ is odd})\right\}[\iota(1-\Lambda_m)]^k\sum_{n=k+1}^{\infty}\binom{n-1}{k}(\iota\Lambda_m)^{n-1-\kappa} &= \\
\frac{1}{1-\iota\Lambda_m}\sum_{k=0}^{\infty}\left\{\omega_H^1(\theta',\theta)[1-\mathbb{I}(k\text{ is odd})]+\omega_F^1(\theta',\theta)\,\mathbb{I}(k\text{ is odd})\right\}\left[\frac{\iota(1-\Lambda_m)}{1-\iota\Lambda_m}\right]^{\kappa} &= \\
\frac{1}{1-\iota\Lambda_m}\left\{\omega_H^1(\theta',\theta)+\left[\frac{\iota(1-\Lambda_m)}{1-\iota\Lambda_m}\right]\omega_F^1(\theta',\theta)\right\}\sum_{k=1}^{\infty}\left[\frac{\iota(1-\Lambda_m)}{1-\iota\Lambda_m}\right]^{2(k-1)} &= \\
\frac{(1-\iota\Lambda_m)\,\omega_H^1(\theta',\theta)+\iota(1-\Lambda_m)\,\omega_F^1(\theta',\theta)}{(1-\iota)\{1-\iota[1-2(1-\Lambda_m)]\}} &
\end{aligned}$$

$$\sum_{n=1}^{\infty} \omega_F^n \left(\theta', \theta\right) = \frac{\left(1 - \iota \Lambda_m\right) \omega_F^1 \left(\theta', \theta\right) + \iota \left(1 - \Lambda_m\right) \omega_H^1 \left(\theta', \theta\right)}{\left(1 - \iota\right) \left\{1 - \iota \left[1 - 2\left(1 - \Lambda_m\right)\right]\right\}}$$

Substituting the above expressions of $\sum_{n=1}^{\infty} \omega_H^n \left(\theta', \theta\right)$ and $\sum_{n=1}^{\infty} \omega_F^n \left(\theta', \theta\right)$ into the definition of the value-added exports from a type-θ industry $VA\left(\theta\right)$ yields the expression of aggregate value-added exports VA given by (1.4.1).

Chapter 2

Endogenous Input-Output Linkages and Structural Changes

2.1 Introduction

Economic growth is often accompanied by structural changes, defined as the shifts of economic resources and activities across sectors or industries. Common measures of structural changes – sectoral shares in consumption expenditure, value-added, and employment – highlight the decline of agriculture, the hump-shaped evolution of manufacturing, and the rise of services over time.[1] However, another prominent feature of the modern economy is the extensive input-output linkages weaving firms of various industries into a production network. It is then natural to doubt the premise that production networks are fixed structures when industries experience such different patterns of growth. One may even ask: could the endogenous adjustment of linkages be driving structural changes in the first place?

Figure 1 plots the sectoral shares in the domestic intermediate input expenditure of the United States, providing a first glimpse of the changing input-output structure. Over the past two decades, the U.S. economy has decreased its use of manufacturing intermediates while relying more on the service sector for intermediate inputs. Since firms constitute not only the units of economic activities but also the "nodes" in production networks,

[1] These stylized facts are established and strengthened by a long strand of literature dating back to Kuznets (1957). Recent empirical works by Buera and Kaboski (2012) as well as Adler, Boppart, and Müller (2018) confirm these patterns for multiple countries over a long period of time.

sectoral firm dynamics are relevant to both structural changes and the evolution of the input-output structure. Figure 2 contrasts the trend in the number of manufacturing establishments with that of the service sector, suggesting a steady shift of establishment mass from manufacturing to services. Moreover, Figure 3 reveals that this reallocation of establishment mass is driven mostly by the difference in establishment entry rate between the two sectors. In summary, evolution of the production networks occurs along side structural changes: the trend in sectoral relative sizes echos the evolution of their relative importance as intermediate input suppliers.

In this paper, I propose a growth model with many industries and endogenous input-output linkages to study the interplay between structural changes and the evolution of production networks. In the model, industries have only one intrinsic difference: they vary in the efficiency of adopting upstream linkages.[2] Each firm produces a differentiated variety for two purposes: to meet final consumption demand and to satisfy the intermediate input demand of other firms. The differentiated varieties are produced from labor and bundles of intermediate inputs (i.e., other varieties) so that labor is the only factor of production in this economy. Both the consumption basket and the intermediate input bundles aggregate varieties in a nested CES fashion with the elasticity of substitution higher within than across industries. While households can access all varieties in the economy, which varieties a firm is able to source as intermediate inputs depends on the production network structure – a variety is accessible only if the linkage exists. In reality, establishing firm-to-firm relationships requires economic resources and I therefore introduce a per-linkage fixed cost that is decreasing in the linkage efficiency of the downstream firm.[3] Firms operate as long as they are not hit by the exogenous exit shock and new firms can be created subject to the sunk costs of entry. Thus, growth in this model is achieved through expanding product variety à la Romer (1990).

[2] To focus on reallocation across industries, I abstract away from asymmetry within industries.

[3] There are many reasons for why linkage fixed costs could differ across industries. For example, introducing an additional intermediate input may require significant changes to the existing production lines in some industries but little modification in others. Additionally, products differ in the extent to which their attributes can be communicated in a systematic way, that is, their codifiability. As a result, product information may be specified via the phone or the internet for some industries, while in-person product inspection may be necessary for others. Empirical works by Fort (2017) and Juhász and Steinwender (2018) suggest that product codifiability affects the degree to which firms or countries engage in domestic or international intermediate input trade.

I solve the model by looking at the problem of a social planner.[4] In each period, the planner decides how to allocate labor among firms and intermediates along input-output linkages, how to distribute firm mass across industries, and how to form linkages among firms. The dynamic problem facing the planner is then to allocate output minus the linkage fixed costs between consumption and firm creation. To focus on the role of endogenous linkages, there is no exogenous productivity growth in the model so that the economy converges to a steady state. During the transition, linkages are redistributed towards industries with relatively high linkage efficiency. The non-stationary distribution of linkages is the result of the following trade off. On one hand, the planner prefers to connect suppliers to firms with relatively high linkage efficiency because the associated fixed costs are lower. On the other hand, the planner would like to maintain certain product diversity in all industries, even the ones very inefficient in adopting linkages, due to the lower elasticity of substitution across industries than within. Consequently, the planner allows the most efficient firms to access all intermediate input varieties while connecting the less efficient ones only partially to the production network. As the economy grows through the accumulation of firm mass, the number of possible bilateral firm relationships expands exponentially, forcing the planner to be more selective in which firms to grant complete upstream linkages.

Due to the love for variety embedded in the CES production function, the redistribution of supplier linkages alters firm productivity through the changes in input diversity. Therefore, endogenous linkages translate the static difference among industries (linkage efficiency) into different productivity growth rates. The direction of structural changes then depends on the cross-industry elasticity of substitution. If varieties of different industries are substitutes, resources are reallocated towards better-connected industries; if there is complementarity among varieties across industries, then it is the less-connected industries that expand in relative sizes. In either case, industries with intermediate levels of linkage efficiency undergo a non-monotonic growth pattern, as they experience the transition from enjoying full access to all suppliers in the economy to having only an incomplete set of upstream linkages.

This paper belongs to the literature trying to understand the economic forces behind structural changes. Broadly speaking, this literature takes two approaches. One approach focuses on the demand side, emphasizing the role

[4] I then show that the planner's solution can be decentralized in a market equilibrium with monopolistically competitive firms, once the production of varieties is optimally subsidized through lump-sum taxation.

of non-homothetic preferences.[5] The other approach highlights a supply-side explanation: productivity growth rates differ across industries.[6] By adopting CES preferences, this paper falls into the second approach. However, structural changes in my model are not contingent on exogenous differences in industry productivity growth rates. In fact, industry productivity in my model grows endogenously at different speed, because the rearrangement of input-output linkages alters the production technology of firms in a biased manner, favoring those most efficient in adopting linkages.

This paper also speaks to the literature on the macroeconomic relevance of production networks, from the earlier work by Long and Plossor (1983) to recent contributions such as Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012) and Baqaee (2018). In particular, this paper joins the budding literature that introduces the endogenous formation of input-output linkages to macroeconomic models, such as Lim (2018) and Oberfield (2018). While these two papers keep the mass of firms fixed, my model accommodates firm entry which is essential in driving both the evolution of the production network and structural changes.

The rest of the paper is organized as follows. Section 2 sets up the model and solves the planner's static problem. Section 3 turns to the dynamic problem of the planner and characterizes the condition under which structural changes can occur. Section 4 presents a simple calibration of the model to the U.S. economy and studies the welfare impact of a technology shock that leads to a universal reduction in the linkage fixed cost. Section 5 concludes. Proofs and derivation details are relegated to the appendix.

2.2 THE MODEL

This section sets up a growth model with many industries and endogenous input-output linkages. I derive the equilibrium as the solution to a social planner problem. In the appendix, I show that the planner's allocation can be

[5] Examples from this strand of literature include Matsuyama (1992, 2002), Echevarria (1997), Laitner (2000), Caselli and Coleman (2001), and Buera and Kaboski (2006) among others.

[6] This technology-based explanation can be traced back to Baumol (1967) and is recently explored by Ngai and Pissarides (2007) as well as Acemoglu and Guerrieri (2008). In addition, Boppart (2014) combines non-homothetic preferences with differential productivity growth in a single framework and shows that the income effect emphasized by the demand-side approach and the relative price effect emphasized by the supply-side approach are of similar quantitative importance in accounting for the observed structural changes in the United States.

sustained in a market equilibrium with monopolistically competitive firms, once the appropriate policy intervention is in place.

2.2.1 PREFERENCES AND TECHNOLOGY

The economy hosts a constant mass L of identical households, which supply labor inelastically and have preferences

$$\int_0^\infty e^{-\rho t} \frac{c_t^{1-\gamma} - 1}{1-\gamma} dt$$

where c_t is consumption per capita at time t.

There is a unit-mass continuum of industries indexed by $i \in [0,1]$. Each industry hosts an endogenous mass of firms, each producing a differentiated variety indexed by $j \in [0, N_i]$.[7] The differentiated varieties can be used to produce the final good and be adopted as intermediate inputs by other firms. The distribution of input-output linkages determines whether a variety is accessible to intermediate customers (i.e., other firms).

PRODUCTION The unique final good is assembled from the differentiated varieties via a nested CES aggregator:

$$Y = \left\{ \int_0^1 \left[\int_0^{N_i} X_i^\beta (j)\, dj \right]^{\frac{\alpha}{\beta}} di \right\}^{\frac{1}{\alpha}}$$

where $\alpha < \beta < 1$ so that the elasticity of substitution is higher within than across industries.

The differentiated varieties are produced from labor and a bundle of intermediate inputs:

$$q_i(j) = l_i(j)^\sigma m_i(j)^{1-\sigma}$$

where $0 < \sigma < 1$ and $m_i(j)$ is the firm's intermediate input bundle which aggregates the differentiated varieties accessible via input-output linkages:

$$m_i(j) = \left\{ \int_0^1 \left[\int_0^{N_{i'}} x_{i,i'}(j,j')^\beta\, \mathbb{I}_{i,i'}(j,j')\, dj' \right]^{\frac{\alpha}{\beta}} di' \right\}^{\frac{1}{\alpha}}$$

where $\mathbb{I}_{i,i'}(j,j')$ is an indicator variable taking on value 1 if firm j' from industry i' is an intermediate input supplier and 0 otherwise.

[7] Throughout this paper, "firms" refer specifically to the producers of the differentiated varieties, and therefore I use "firms" interchangeably with "varieties" .

To focus on cross-industry differences, I assume that all varieties in any given industry are symmetric. Under this assumption of within-industry symmetry, the aggregate production function can be written as

$$Y = \left(\int_0^1 X_i^{\alpha} N_i^{\frac{\alpha}{\beta}} di \right)^{\frac{1}{\alpha}}$$

Accordingly, the intermediate input bundle used by a firm reduces to

$$m_i = \left[\int_0^1 x_{i,i'}^{\alpha} \left(\mu_{i,i'} N_{i'} \right)^{\frac{\alpha}{\beta}} di' \right]^{\frac{1}{\alpha}}$$

where $\mu_{i,i'} \equiv \int_0^{N_{i'}} \mathbb{I}_{i,i'}(j, j')\, dj'/N_{i'}$ gives the fraction of industry-i' firms supplying to a firm industry-i.[8]

LINKAGE AND FIRM CREATION Firms operate as long as they are not hit by the exogenous exit shock, which occurs with probability δ every period. Given the total stock of firms $N \equiv \int_0^1 N_i di$ in the economy, the social planner is free to redistribute firm mass across industries by choosing $\{N_i\}_{i\in[0,1]}$ period by period. The social planner also chooses input-output linkages $\left\{\mu_{i,i'}\right\}_{i,i'\in[0,1]}$, subject to a per-linkage fixed cost in units of the final good that has to be paid every period. Specifically, the fixed cost of establishing a firm-to-firm relationship between an industry-i buyer and an industry-i' seller is κ/ϕ_i, where ϕ_i is the efficiency of an industry-i firm in incorporating an additional supplier into its intermediate input bundle and is the only source of cross-industry asymmetry in this model. Consequently, the economy admits a single state variable, the aggregate firm mass N, with law of motion:

$$\dot{N}_t = N_t^e - \delta N_t$$

where N_t^e is the mass of new firms. The social planner can create firms at a cost of ν units of the final good per new firm. Therefore, the aggregate resource constraint of the economy is

$$C + \int_0^1 \int_0^1 \frac{\kappa}{\phi_i} \mu_{i,i'} N_i N_{i'} di di' + \nu N^e = Y$$

where $C \equiv cL$ is aggregate consumption and $\int_0^1 \int_0^1 (\kappa/\phi_i)\, \mu_{i,i'} N_i N_{i'} di di'$ is the total fixed costs for creating all the firm-to-firm linkages in the economy.

[8] Throughout the paper, whenever a variable has two subscripts, the former index always refers to the industry of the customer firm and the latter index always refers to the industry of the supplier firm.

2.2.2 The static problem

The static problem of the social planner is to maximize aggregate output Y net of all the linkage fixed costs by choosing: (1) the allocation of labor across industries $\{l_i\}_{i\in[0,1]}$; (2) the allocation of variety output among intermediate users $\{x_{i,i'}\}_{i,i'\in[0,1]}$; (3) the distribution of firm mass across industries $\{N_i\}_{i\in[0,1]}$; and (4) the distribution of input-output linkages $\{\mu_{i,i'}\}_{i,i'\in[0,1]}$ subject to the constraint of aggregate labor supply:

$$\int_0^1 l_i N_i di = L$$

the identity of total firm mass:

$$\int_0^1 N_i di = N$$

and the constraints of variety quantities:

$$q_i = X_i + \int_0^1 x_{i',i}\mu_{i',i}N_{i'}di' \qquad \text{for all } i \in [0,1]$$

which states that variety output q_i must meet both the final demand X_i and all the intermediate demand $\int_0^1 x_{i',i}\mu_{i',i}N_{i'}di'$.

The solution to the static problem is characterized by four sets of optimality conditions. First, the efficient allocation of labor requires the equalization of marginal product of labor across industries:

$$\frac{\partial Y}{\partial X_i}\frac{\partial X_i}{\partial l_i}\frac{1}{N_i} = \frac{\partial Y}{\partial X_{i'}}\frac{\partial X_{i'}}{\partial l_{i'}}\frac{1}{N_{i'}} \qquad \text{for all } i, i'$$

Second, when choosing the intermediate input quantity $x_{i,i'}$ that a firm in industry i sources from a supplier in industry i', the planner faces the following trade-off: increasing $x_{i,i'}$ leads to more output of industry-i variety while diverting industry-i' output from consumption to fulfilling intermediate input demand. Therefore, the efficient allocation of intermediate input varieties requires balancing this trade-off:

$$\frac{\partial Y}{\partial X_i}\frac{\partial X_i}{\partial x_{i,i'}} = -\frac{\partial Y}{\partial X_{i'}}\frac{\partial X_{i'}}{\partial x_{i,i'}} \qquad \text{for all } i, i' \tag{2.2.1}$$

Third, the planner distributes total firm mass across industries so as to equalize the marginal returns to varieties:

$$\frac{\partial Y}{\partial N_i} - \int_0^1 \frac{\kappa}{\phi_{i''}}\mu_{i'',i}N_{i''}di'' = \frac{\partial Y}{\partial N_{i'}} - \int_0^1 \frac{\kappa}{\phi_{i''}}\mu_{i'',i'}N_{i''}di'' \qquad \text{for all } i, i' \tag{2.2.2}$$

Finally, the planner sets up input-output linkages according to the following first order condition:

$$\frac{\partial Y}{\partial \mu_{i,i'}} - \frac{\kappa}{\phi_i} N_i N_{i'} \begin{cases} \leqslant 0 & \text{if } \mu_{i,i'} = 0 \\ = 0 & \text{if } \mu_{i,i'} \in [0,1] \\ \geqslant 0 & \text{if } \mu_{i,i'} = 1 \end{cases} \tag{2.2.3}$$

where the left hand side corresponds to the return to an additional firm-to-firm relationship between an industry-i buyer and an industry-i' seller. Using these four sets of optimality conditions, I characterize the solution to the planner's static problem in the proposition below.

Proposition 1 *Given the state variable, total firm mass $N \equiv \int_0^1 N_i di$, the planner's static problem yields the following solution: labor is allocated according to $l_i = L/N$ for all i; input-output linkages are formed according to*

$$\mu_{i,i'} \equiv \tilde{\mu}_i = \begin{cases} \frac{\phi_i}{\underline{\phi}} & \text{if } \phi_i < \underline{\phi} \\ 1 & \text{if } \phi_i \geqslant \underline{\phi} \end{cases} \tag{2.2.4}$$

where the efficiency cutoff is

$$\underline{\phi} = \kappa \left(\frac{\sigma}{1-\sigma} \right) \left(\frac{\beta}{1-\beta} \right) \frac{N^2}{Y}; \tag{2.2.5}$$

the distribution of firm mass across industries is given by

$$N_i = \tilde{\mu}_i^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)} \frac{N}{A} \quad \text{where } A \equiv \int_0^1 \tilde{\mu}_i^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)} di \tag{2.2.6}$$

and the allocation of intermediate inputs is given by

$$x_{i,i'} = \left(\frac{1-\sigma}{\sigma} \right) N^{-\left(\frac{1+\alpha}{\alpha}\right)} Y \tilde{\mu}_i^{-1} N_{i'}^{\frac{\beta-\alpha}{\alpha\beta}}. \tag{2.2.7}$$

Consequently, aggregate output in the social optimum is

$$Y = \sigma (1-\sigma)^{\frac{1-\sigma}{\sigma}} L A^{\frac{\beta-\alpha}{\sigma\alpha\beta}} N^{\frac{1-\beta}{\sigma\beta}} \tag{2.2.8}$$

Since all firms share the same labor productivity, the planner equalizes employment per firm across the economy. In contrast, the returns to input-output linkages depend on the efficiency ϕ_i of downstream firms in adopting additional intermediate input varieties. This leads to an efficiency cutoff $\underline{\phi}$ partitioning all firms into two groups: for firms with linkage efficiency $\phi_i \geqslant \underline{\phi}$, the planner allows them to source intermediate inputs from all other firms in the economy by setting $\tilde{\mu}_i = 1$; for firms with $\phi_i < \underline{\phi}$, the planner connects them only partially to the production networks by setting $\tilde{\mu}_i \in (0,1)$ where $\tilde{\mu}_i$ is increasing in ϕ_i. Therefore, the number of suppliers that a firm has weakly increases in its linkage efficiency ϕ_i. The distribution of firm mass depends crucially on the cross-industry elasticity of substitution $1/(1-\alpha)$. If varieties of different industries are substitutes ($\alpha > 0$), N_i increases in $\tilde{\mu}_i$, suggesting that the planner allocates more firms to better-connected industries. If there is complementarity among varieties across industries ($\alpha < 0$), N_i is then decreasing in $\tilde{\mu}_i$, implying that larger industries have fewer upstream linkages. By Proposition 1, total fixed costs of establishing all the input-output linkages in the economy is

$$\int_0^1 \int_0^1 \frac{\kappa}{\phi_i} \mu_{i,i'} N_i N_{i'} di di' = \kappa \frac{Z}{A} N^2$$

where

$$Z \equiv \int_0^1 \phi_i^{-1} \tilde{\mu}_i^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)+1} di$$

2.2.3 The dynamic problem

The dynamic problem of the planner is to allocate output minus linkage costs between consumption and firm creation:

$$\max_{\{C_t,N_t\}_{t\geqslant 0}} \int_0^\infty e^{-\rho t} \frac{C_t^{1-\gamma}-1}{1-\gamma} dt$$

subject to

$$\dot{N}_t = \frac{1}{\nu}\left(Y_t - \kappa \frac{Z_t}{A_t} N_t^2 - C_t\right) - \delta N_t \tag{2.2.9}$$

and the initial condition $N_0 > 0$. Taking relevant partial derivatives of the Hamiltonian yields the Euler equation for consumption growth:

$$\frac{\dot{C}_t}{C_t} = \frac{1}{\gamma\nu}\left[\frac{\partial}{\partial N_t}\left(Y_t - \kappa \frac{Z_t}{A_t} N_t^2\right) - \nu(\delta+\rho)\right] \tag{2.2.10}$$

with transversality condition:

$$\lim_{t\to\infty} e^{-\rho t} C_t^{-\gamma} N_t = 0$$

2.3 Structural changes and aggregate growth

This section characterizes the dynamic equilibrium of the economy and examines under what conditions structural changes can take place. To facilitate the discussion about structural changes, I define the firm mass shares of industries:

$$n_i \equiv \frac{N_i}{N} = \begin{cases} \left(\frac{\phi_i}{\underline{\phi}}\right)^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)} A^{-1} & \text{if } \phi_i < \underline{\phi} \\ A^{-1} & \text{if } \phi_i \geqslant \underline{\phi} \end{cases} \tag{2.3.1}$$

where the linkage efficiency cutoff $\underline{\phi}$ and the term A are defined as in (2.2.5) and (2.2.6). Since all firms in the economy have the same employment size by Proposition 1, $\{n_i\}_{i\in[0,1]}$ coincide with the employment shares of industries and therefore can represent industry relative sizes. The next proposition relates the growth rate of relative industry sizes n_i to the dynamics of the linkage efficiency cutoff $\underline{\phi}$, an aggregate variable.

Proposition 2 *Suppose that linkage efficiency ϕ_i follows a continuous distribution $\mathscr{F}$ with support $[\phi_{min}, \phi_{max}]$. The dynamics of relative industry sizes satisfy*

$$\frac{\dot{n}_i}{n_i} = \begin{cases} -\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)\Phi\frac{\dot{\underline{\phi}}}{\underline{\phi}} & \text{if } \phi_i < \underline{\phi} \\ \alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)(1-\Phi)\frac{\dot{\underline{\phi}}}{\underline{\phi}} & \text{if } \phi_i \geqslant \underline{\phi} \end{cases} \tag{2.3.2}$$

where

$$\Phi \equiv \frac{\int_{\underline{\phi}}^{\phi_{max}} d\mathscr{F}(\phi)}{\int_{\phi_{min}}^{\underline{\phi}} \left(\frac{\phi}{\underline{\phi}}\right)^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)} d\mathscr{F}(\phi) + \int_{\underline{\phi}}^{\phi_{max}} d\mathscr{F}(\phi)}$$

Furthermore, for two industries i and i' with $\phi_i > \phi_{i'}$, the growth rate of their relative sizes satisfies

$$\frac{\dot{n}_i}{n_i} - \frac{\dot{n}_{i'}}{n_{i'}} = \begin{cases} 0 & \text{if } \phi_{i'} < \phi_i < \underline{\phi} \text{ or } \underline{\phi} < \phi_{i'} < \phi_i \\ \alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)\frac{\dot{\underline{\phi}}}{\underline{\phi}} & \text{if } \phi_{i'} < \underline{\phi} < \phi_i \end{cases} \tag{2.3.3}$$

Proposition 2 suggests that the distribution of firm mass across industries is non-stationary as long as $\dot{\underline{\phi}} \neq 0$ and $\phi_{min} < \underline{\phi} < \phi_{max}$. Furthermore, if we divide industries into two groups according to whether the linkage efficiency is above or below the cutoff $\underline{\phi}$, then firm mass shares evolve in the same way for all industries within a group. If $\alpha > 0$, firm mass (hence employment) distribution shifts towards better-connected industries (those with $\phi_i \geqslant \underline{\phi}$) when the linkage efficiency cutoff rises over time ($\dot{\underline{\phi}} > 0$) and in the opposite direction when the cutoff falls ($\dot{\underline{\phi}} < 0$). If $\alpha < 0$, firm mass is redistributed towards industries with fewer upstream linkages (those with $\phi_i < \underline{\phi}$) when $\dot{\underline{\phi}} > 0$ and in the opposite direction when $\dot{\underline{\phi}} < 0$.

2.3.1 GROWTH WITHOUT STRUCTURAL CHANGES

I define the absence of structural changes as the state where $\dot{n}_i/n_i = \dot{n}_{i'}/n_{i'}$ for all i and i'. One immediately observes that structural changes cannot take place if the distribution of input-output linkages $\{\tilde{\mu}_i\}_{i\in[0,1]}$ is held fixed, because $\dot{\underline{\phi}} = 0$ in this case of exogenous production networks. In the case of endogenous linkages, structural changes can still be absent if all industries lie above or below the efficiency cutoff $\underline{\phi}$ (i.e., $\underline{\phi} \leqslant \phi_{min}$ or $\underline{\phi} \geqslant \phi_{max}$). Since the parameter κ regulates the fixed cost of establishing linkages, the planner's allocation features a solution for the linkage efficiency cutoff $\underline{\phi}$ outside the support $[\phi_{min}, \phi_{max}]$ when κ is sufficiently low or high. The next proposition identifies the threshold levels of κ below or above which the economy grows without structural changes and characterizes the aggregate dynamics in each case.

Proposition 3 *If $\kappa \geqslant \overline{\kappa}$, all industries lie below the linkage efficiency cutoff ($\underline{\phi} \geqslant \phi_{max}$). In this case, the dynamic equilibrium is given by the following system of differential equations:*

$$\frac{\dot{C}_t}{C_t} = \frac{1}{\gamma}\left\{\frac{1}{\nu}\left[1-2(1-\sigma)\right]\left(\frac{1-\beta}{\sigma\beta}\right)\Lambda N_t^{\frac{[1-2(1-\sigma)]\frac{1-\beta}{\sigma\beta}}{1-\left(\frac{1-\sigma}{\sigma}\right)\left(\frac{1-\beta}{\beta}\right)}-1} - \delta - \rho\right\} \tag{2.3.4}$$

$$\frac{\dot{N}_t}{N_t} = \frac{1}{\nu}\left\{\left[1-\left(\frac{1-\sigma}{\sigma}\right)\left(\frac{1-\beta}{\beta}\right)\right]\Lambda N_t^{\frac{[1-2(1-\sigma)]\frac{1-\beta}{\sigma\beta}}{1-\left(\frac{1-\sigma}{\sigma}\right)\left(\frac{1-\beta}{\beta}\right)}-1} - \frac{C_t}{N_t}\right\} - \delta \tag{2.3.5}$$

where Λ is a constant term. If $\kappa \leqslant \underline{\kappa}$, all industries lie above the linkage efficiency cutoff ($\underline{\phi} \leqslant \phi_{min}$). In this case, the dynamic equilibrium is given by the following system of differential equations:

$$\frac{\dot{C}_t}{C_t} = \frac{1}{\gamma}\left\{\frac{1}{\nu}\left[\left(\frac{1-\beta}{\beta}\right)(1-\sigma)^{\frac{1-\sigma}{\sigma}} L N_t^{\frac{1-\beta}{\sigma\beta}-1} - 2\kappa\int_{\phi_{min}}^{\phi_{max}}\frac{1}{\phi}d\mathscr{F}(\phi)N_t\right] - \delta - \rho\right\} \tag{2.3.6}$$

$$\frac{\dot{N}_t}{N_t} = \frac{1}{\nu}\left[\sigma(1-\sigma)^{\frac{1-\sigma}{\sigma}} L N_t^{\frac{1-\beta}{\sigma\beta}-1} - \kappa\int_{\phi_{min}}^{\phi_{max}}\frac{1}{\phi}d\mathscr{F}(\phi)N_t - \frac{C_t}{N_t}\right] - \delta \tag{2.3.7}$$

The full expressions of the threshold levels $\overline{\kappa}$ and $\underline{\kappa}$ as well as the constant Λ are given in the appendix. In both cases, there exists a unique steady state that is locally saddle-path stable, provided that $2(1-\sigma) < 1$ and $(1-\beta)/\beta < \sigma$. Furthermore, the steady-state aggregate firm mass N^{SS} satisfies $dN^{SS}/d\kappa < 0$ and $dN^{SS}/dL > 0$. Finally, $\underline{\kappa} < \overline{\kappa}$ provided that labor *endowment L is sufficiently large.*

For sufficiently small linkage fixed costs ($\kappa \leqslant \underline{\kappa}$), all firms in the economy are fully connected with each other via input-output linkages. Contrastingly, when linkage fixed costs are too high ($\kappa \geqslant \overline{\kappa}$), all firms in the economy adopt intermediate inputs from only a subset of their peers. In both cases, a small technology shock that reduces κ or a market size shock that raises L to the economy already at the steady state leads it to a new steady state with a higher number of firms. During the transition to the new steady state, all industries expand at the same rate $\dot{N}_t/N_t$ and the distribution of linkages $\{\tilde{\mu}_i\}_{i\in[0,1]}$ remain unchanged.

2.3.2 Growth with structural changes

I define structural changes as the state where $\dot{n}_i/n_i \neq \dot{n}_{i'}/n_{i'}$ for at least some i and i'. Proposition 2 establishes that structural changes can occur only if the planner's choice of the linkage efficiency cutoff $\underline{\phi}$ partitions firms into two groups by whether they have complete or incomplete upstream linkages. This case requires intermediate levels of the linkage fixed cost parameter κ. The next proposition characterizes the aggregate dynamics when growth is accompanied by structural changes.

Proposition 4 *The economy undergoes structural changes as long as $\phi_{min} < \underline{\phi} < \phi_{max}$ and $\dot{\underline{\phi}} \neq 0$. In this case, the dynamic equilibrium is given by the following system of differential equations:*

$$\frac{\dot{C}_t}{C_t} = \frac{1}{\gamma}\left\{\frac{1}{\nu}\left[\left(\frac{1-\beta}{\beta}\right)(1-\sigma)^{\frac{1-\sigma}{\sigma}} LA_t^{\frac{\beta-\alpha}{\sigma\alpha\beta}} N_t^{\frac{1-\beta}{\sigma\beta}-1} - 2\frac{Z_t}{A_t}\kappa N_t\right] - (\delta+\rho)\right\} \tag{2.3.8}$$

$$\frac{\dot{N}_t}{N_t} = \frac{1}{\nu}\left[\sigma(1-\sigma)^{\frac{1-\sigma}{\sigma}} LA_t^{\frac{\beta-\alpha}{\sigma\alpha\beta}} N_t^{\frac{1-\beta}{\sigma\beta}-1} - \frac{Z_t}{A_t}\kappa N_t - \frac{C_t}{N_t}\right] - \delta \tag{2.3.9}$$

$$\frac{\dot{\underline{\phi}}_t}{\underline{\phi}_t} = \left(2-\frac{1-\beta}{\sigma\beta}\right)A_t\left[A_t - \left(\frac{1-\sigma}{\sigma}\right)\left(\frac{1-\beta}{\beta}\right)\int_{\phi_{min}}^{\underline{\phi}_t}\left(\frac{\phi}{\underline{\phi}_t}\right)^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)} d\mathscr{F}(\phi)\right]^{-1}\frac{\dot{N}_t}{N_t} \tag{2.3.10}$$

where A_t and Z_t are functions of $\underline{\phi}_t$ whose full expressions are given in the appendix. A sufficient condition for this system to admit a unique locally stable steady state is $\phi_{min}/\phi_{max} > \alpha(1-\beta)(1-\sigma)/[(1-\alpha)\beta - \alpha(1-\beta)\sigma]$. Furthermore, the steady-state aggregate firm mass N^{SS} and linkage efficiency cutoff $\underline{\phi}^{SS}$ satisfies $dN^{SS}/d\kappa < 0$, $dN^{SS}/dL > 0$, $d\underline{\phi}^{SS}/d\kappa > 0$, and

$d\underline{\phi}^{SS}/dL > 0$. Given initial condition $N_0 < N^{SS}$, the economy experiences structural changes throughout the entire transition to the steady state if the linkage fixed cost parameter κ satisfies $\underline{\kappa}' \leqslant \kappa \leqslant \overline{\kappa}'$, where the threshold levels $\underline{\kappa}'$ and $\overline{\kappa}'$ are given in the appendix. If $\underline{\kappa} < \kappa < \underline{\kappa}'$ or/and $\overline{\kappa}' < \kappa < \overline{\kappa}$, the economy experiences structural changes only during part of the transition to the steady state. Finally, $\underline{\kappa} < \underline{\kappa}'$ and $\overline{\kappa}' < \overline{\kappa}$ provided that labor endowment L is sufficiently large.

In this model, structural changes are driven by the cross-industry differences in linkage efficiency ϕ_i, as well as the assumption that varieties are less substitutable across industries than within industries. Since the fixed costs of creating linkages decrease in the linkage efficiency of the downstream industries (i.e., the intermediate input customers), the planner prefers to concentrate linkages in industries with high ϕ_i, which is why the optimal linkage density $\tilde{\mu}_i$ weakly increases in industry linkage efficiency ϕ_i. However, the planner also faces the trade-off between allocating resources to industries most efficient in adopting linkages and maintaining sufficient product diversity in all industries, because households care about not only the total number of varieties but also how the varieties are distributed across industries (a direct consequence of the elasticity of substitution being lower across than within industries). Due to this trade-off, the constraint $\tilde{\mu}_i \leqslant 1$ becomes binding for industries with $\phi_i \geqslant \underline{\phi}$. As the economy grows, the linkage efficiency cutoff $\underline{\phi}$ rises, because the number of total possible firm relationships N^2 increases twice as fast as that of total firm mass N, prompting the planner to be more "selective" in which industries to concentrate linkages in. Consequently, $\tilde{\mu}_i$ falls in industries with $\phi_i < \underline{\phi}$ by Proposition 1, which implies that upstream linkages become sparser for these industries below the cutoff. Furthermore, the endogenous adjustment of linkages leads to differential productivity growth on the industry level:

$$\frac{\dot{(q_i/l_i)}}{q_i/l_i} = (1-\sigma)\frac{\dot{(Y/L)}}{Y/L} + (1-\sigma)\left(\frac{1-\beta}{\beta}\right)\frac{\dot{\tilde{\mu}}_i}{\tilde{\mu}_i}$$

The above equation decomposes the growth rate of industry productivity into a common component driven by aggregate productivity growth and an idiosyncratic component driven by the rearrangement of input-output linkages.[9] Specifically, industries low in linkage efficiency (those with $\phi_i < \underline{\phi}$) have relatively low productivity growth rate because linkages are being reallocated away from them ($\dot{\tilde{\mu}}_i/\tilde{\mu}_i < 0$) as the economy grows. Since the CES production function entails a "love of variety", the relative loss of input diversity due to linkage rearrangement ultimately results in productivity

[9] To derive this decomposition, we start from (2.8.3), then substitute in $l_i = L/N$, (2.2.6), and (2.2.8), and finally time-differentiate both side of the equation.

disadvantage. To summarize, endogenous linkages translate the static difference among industries (linkage efficiency) into uneven productivity growth.

To an economy already at the steady state (and satisfies $\underline{\kappa}' \leqslant \kappa \leqslant \overline{\kappa}'$), a small technology shock that reduces κ or a market size shock that raises L leads the economy to a new steady state with a larger aggregate firm mass ($N_1^{SS} > N_0^{SS}$) and a lower linkage efficiency cutoff ($\underline{\phi}_1^{SS} < \underline{\phi}_0^{SS}$). Since the linkage efficiency cutoff $\underline{\phi}$ rises as the economy accumulates firm mass during the transition, it must be that on impact $\underline{\phi}$ "overshoots" the new steady-state cutoff $\underline{\phi}_1^{SS}$, falling to $\underline{\phi}_1 < \underline{\phi}_1^{SS}$. During the transition to the new steady state, industry dynamics follow different patterns depending on the linkage efficiency ϕ_i relative to the cutoff. If $\alpha > 0$, industries with $\phi_i > \underline{\phi}_1^{SS}$ ($\phi_i < \underline{\phi}_1$) see their firm mass shares n_i grow (shrink) monotonically throughout the entire transition, whereas those with $\underline{\phi}_1 < \phi_i < \underline{\phi}_1^{SS}$ first expand then shrink in their relative sizes. If $\alpha < 0$, industries with $\phi_i < \underline{\phi}_1$ ($\phi_i > \underline{\phi}_1^{SS}$) see their firm mass shares n_i grow (shrink) monotonically throughout the entire transition, whereas those with $\underline{\phi}_1 < \phi_i < \underline{\phi}_1^{SS}$ first shrink then expand in their relative sizes.

Finally, I study the relationship between structural changes and the evolution of intermediate input expenditure shares. In a competitive equilibrium that decentralizes the planner's allocation, the price of an industry-i variety is given by $(\partial Y/\partial X_i)/N_i$. Therefore, the expenditure share of intermediate inputs supplied by industry i in the economy's total spending on intermediates is

$$\Lambda_i \equiv \frac{\int_0^1 x_{i',i}\mu_{i',i}N_iN_{i'}di' \, (\partial Y/\partial X_i)/N_i}{\int_0^1 \int_0^1 x_{i',i}\mu_{i',i}N_iN_{i'}di' \, (\partial Y/\partial X_i)/N_i di}$$

The next proposition relates the relative changes in intermediate expenditure shares to the relative changes in firm mass shares:

Proposition 5 *For any two industries i and i', their relative importance as intermediate input suppliers is related to their relative sizes as given by the following equation:*

$$\frac{\dot{\Lambda}_i}{\Lambda_i} - \frac{\dot{\Lambda}_{i'}}{\Lambda_{i'}} = \frac{\dot{n}_i}{n_i} - \frac{\dot{n}_{i'}}{n_{i'}} \tag{2.3.11}$$

Proposition 5 implies that structural changes ($\dot{n}_i/n_i \neq \dot{n}_{i'}/n_{i'}$) are necessary for there to be redistribution of intermediate input expenditure shares across industries. Specifically, industries that gain more firm mass also become more prominent intermediate input suppliers, a prediction consistent with the patterns presented in Figure 1 and Figure 3.

2.4 Quantitative Example

In this section, I perform a simple calibration of the model. I then use the calibrated model to study the impacts of a technology shock that lowers the linkage fixed cost parameter κ universally. The motivation for such a shock is the information and communication technologies (ICT) revolution during the 1990s, when the surging usage of internet and mobile phones (as illustrated in Figure 4) arguably made it easier for businesses to establish supplier-customer relationships with each other.

2.4.1 Calibration

To calibrate the model, I first impose a parametric assumption on the distribution of linkage efficiency ϕ across industries: ϕ follows a Pareto distribution with shape parameter ζ and support $[\phi_{min}, \infty)$.[10] Accordingly, this model is characterized by 11 parameters: α, β, σ, κ, ν, δ, ρ, γ, ζ, ϕ_{min}, L and one initial condition N_0. I choose values for these parameters as follows. I consider two cases of α: $\alpha = -1/3$ and $\alpha = 1/5$. These values correspond respectively to the cross-industry elasticity of substitution $1/(1-\alpha)$ being 0.75 and 1.25, consistent with the empirical finding that this high-level elasticity of substitution is close to 1 (Atalay 2017, Oberfield and Raval 2014). For the elasticity of substitution within an industry $1/(1-\beta)$, I set it at 4 (implying $\beta = 3/4$), which lies within the range estimated by the empirical literature (e.g., Broda and Weinstein 2006) and adopted by the quantitative trade literature (e.g. Costinot and Rodríguez-Clare 2014). For the inter-temporal preference parameters, I choose $\rho = 0.02$ and $\gamma = 1$ as benchmarks. Since it is the ratio $\underline{\phi}/\phi_{min}$ that matters for the aggregate variables, I thus normalize ϕ_{min} to 1.

I calibrate the rest of the parameters to relevant statistics of the U.S. economy. The exogenous firm exit rate $\delta = 0.096$ corresponds to the average of U.S. establishment exit rate during 1997-2015.[11] Labor endowment $L = 105.58$ (in millions) is set to the U.S. employment size averaged over the same time period. In the model, σ gives the common proportion of variety output that goes into final consumption: $\sigma = X_i/q_i$. I show in the appendix that σ is also the ratio of GDP to gross output and set $\sigma = 0.56$ corresponding to the period average of this ratio. Finally, I assume that the

[10] In order for output to be finite, the Pareto shape parameter must satisfy $\zeta > (1-\sigma)\alpha(1-\beta)/(\beta-\alpha)$. This condition also guarantees the local stability of the steady state when $\phi_{max} = \infty$.

[11] I choose 1997 as the initial point because expenditure data on domestic intermediates is not available for the earlier years.

U.S. economy was at the steady state in 1997 and calibrate the remaining three parameters (κ, ν, and ζ) to jointly match aggregate output $Y = 8.61$ (in trillion USD), aggregate consumption $C = 5.56$ (in trillion USD), and total firm mass $N = 5.37$ (in millions) as observed in 1997. Additional calibration details are given in the appendix.

2.4.2 THE IMPACTS OF A TECHNOLOGY SHOCK

Using the calibrated model, I now study the impacts of a 10% permanent reduction in κ, which governs the level of linkage fixed costs. Figure 5 shows the aggregate dynamics in response to this technology shock for two cases: varieties of different industries are substitutes ($\alpha = 1/5$) or complements ($\alpha = -1/3$). The left panel illustrates the overshooting behavior of the linkage efficiency cutoff $\underline{\phi}$. On impact, this cutoff drops since linkages have become cheaper to form universally. During the transition to the new steady state, the cutoff rises as the economy gains firm mass (the right panel) because the exponentially-growing linkage possibilities force the planner to be more selective in which firm pairs to connect.

Quantitatively, the cumulative welfare gains from this technology shock amount to 74.0% and 86.5% of the initial welfare level for the cases of $\alpha = 1/5$ and $\alpha = -1/3$ respectively. In an alternative model where the production network is fixed (i.e., neither the firm mass distribution $\{N_i\}_{i\in[0,1]}$ nor the linkage distribution $\{\tilde{\mu}_i\}_{i\in[0,1]}$ is responsive to shocks), the cumulative welfare gains, which is equivalent to the discounted sum of a series of static gains, are 35.9% and 40.6% for the two cases of α respectively. Therefore, more than half of the gains from this technology shock can be attributed to the endogenous adjustment of input-output linkages and the resulting structural changes.

Figure 6 shows the dynamic responses on the industry level. To highlight the patterns of structural changes, I plot the evolution of relative industry size n_i for three levels of the linkage efficiency ϕ corresponding to 0.5, 0.95, and 2 times the new steady-state linkage efficiency cutoff $\underline{\phi}$. As discussed in Section 3.2, the elasticity of substitution across industries determines the pattern of structural changes. If $\alpha > 0$, firm mass shifts towards industries with high linkage efficiency and therefore more linkages. If $\alpha < 0$, firm mass is redistributed in the opposite direction. In both cases, industries with intermediate levels of linkage efficiency that are surpassed by the rising cutoff during the transition experience non-monotonic changes in relative size.

2.5 CONCLUSION

This paper presented a growth model with endogenous input-output linkages to study the interplay between structural changes and the evolution of the production networks. The model is consistent with the stylized empirical fact that the expanding sector also becomes more important intermediate input supplier, as measured by the sectoral shares in aggregate intermediate input expenditure. The model is also able to generate non-monotonic industry growth patterns, another empirical regularity. Nevertheless, the model relies on several strong assumptions which I hope to relax in future research. First, one may introduce exogenous aggregate productivity growth to explore the existence of a constant growth path, as in Acemoglu and Guerrieri (2008), and also to bring the model closer to the actual growth experience of the U.S. both in a qualitative and a quantitative sense. Second, one may also relax the assumption of a fixed firm boundary to study the interaction among structural changes, endogenous linkages, and the organization of the firm. Third, a significant portion of input-output linkages in the real world are cross-country relationships. It would be interesting to see how the expansion of global value chains would affect structural changes in different countries.

2.6 Appendix A: Figures

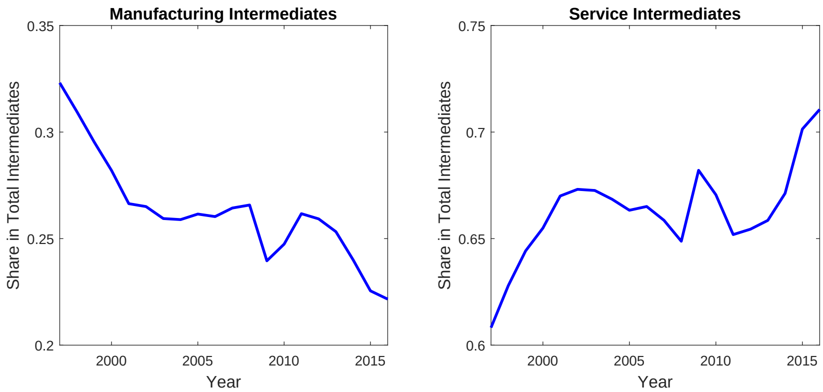

Figure 2.6.1: Shares of manufacturing and service intermediates in total U.S. domestic intermediate input expenditure. The data source is the Input-Output Accounts Data of the Bureau of Economic Analysis. Domestic intermediate input expenditure is calculated from the after-redefinition Use Tables (producer value) and Import Matrices. In this figure, "Services" include all the NAICS sectors other than "Agriculture, Forestry, Fishing, and Hunting", "Mining", "Construction", "Manufacturing", and "Public Administration".

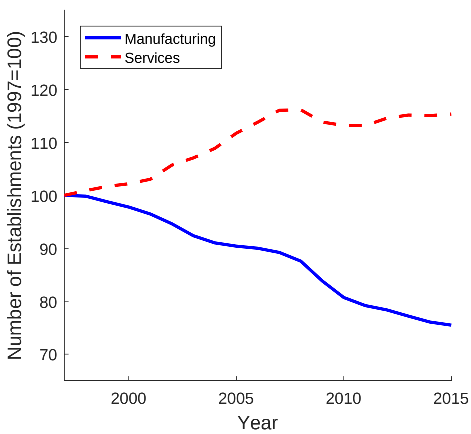

Figure 2.6.2: Number of establishments in the U.S. manufacturing and services industries (1997=100). The data source is the Longitudinal Business Database of the U.S. Census Bureau. In this figure, "Services" include the following SIC 87 major divisions: "Transportation and Public Utilities", "Wholesale Trade", "Retail Trade", "Finance, Insurance, and Real Estate", and "Services" .

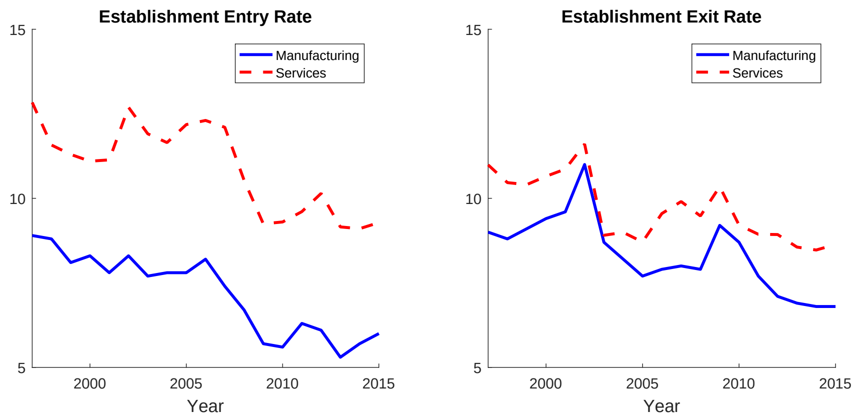

Figure 2.6.3: Establishment dynamics in the U.S. manufacturing and services industries. The data source is the Longitudinal Business Database of the U.S. Census Bureau. In this figure, "Services" include the following SIC 87 major divisions: "Transportation and Public Utilities", "Wholesale Trade", "Retail Trade", "Finance, Insurance, and Real Estate", and "Services" .

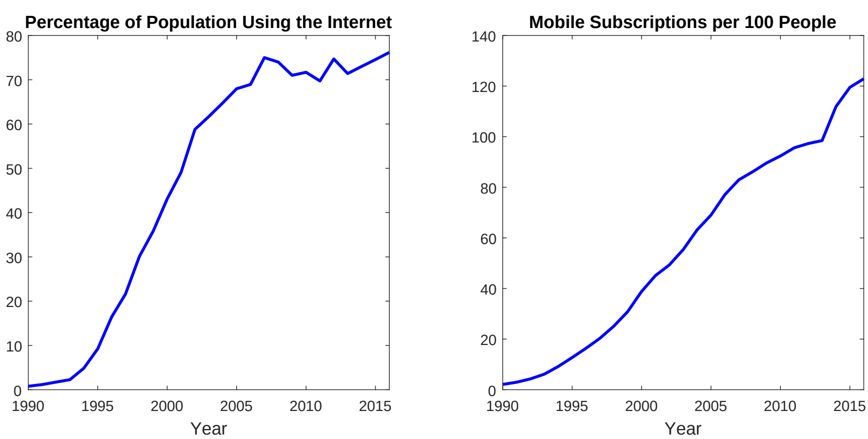

Figure 2.6.4: Time trend of the usage of information and communication technologies (ICT) in the United States. The data source is the World Bank World Development Indicators.

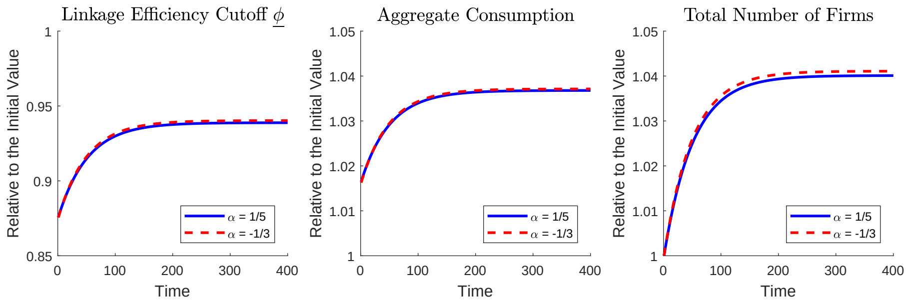

Figure 2.6.5: Aggregate dynamics in response to a 10% permanent decrease in κ, which implies a 10% universal reduction in the fixed cost per linkage.

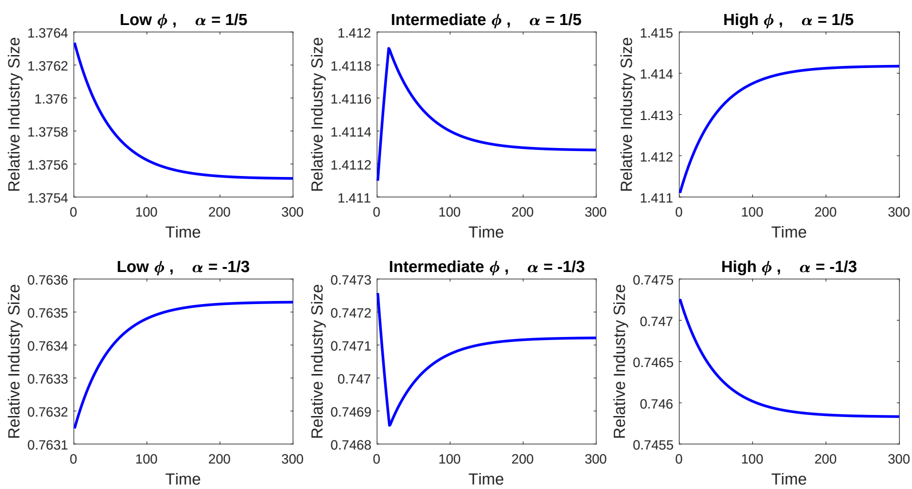

Figure 2.6.6: Patterns of structural changes in response to a 10% permanent reduction in linkage fixed cost κ. The low, intermediate, and high levels of the linkage efficiency ϕ correspond respectively to 0.5, 0.95, and 2 times the new steady-state linkage efficiency cutoff $\underline{\phi}$.

2.7 Appendix B: Decentralization

This section replicates the planner's allocation in a decentralized economy with perfectly competitive markets for labor and the final good, as well as monopolistically competitive markets with free entry for the differentiated varieties. Since monopolistic market power is the only source of distortion, I show that a production subsidy for the differentiated varieties to correct the markup is sufficient for aligning the market equilibrium with the social optimum.

Households

Households supply labor inelastically and earn competitive wage w. Static utility maximization by households implies the following final demand for the differentiated varieties:

$$X_i = YP^{\frac{1}{1-\alpha}} p_i^{\frac{1}{\alpha-1}} N_i^{\frac{\alpha-\beta}{\beta(1-\alpha)}}$$

where $P \equiv \left[\int_0^1 p_i^{\frac{\alpha}{\alpha-1}} N_i^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)} di\right]^{\frac{\alpha}{\alpha-1}}$ is the consumer price index.

Firms

The markets for varieties are characterized by monopolistic competition. In each period, firms are free to choose which industry to operate in and moving across industries incurs no cost. Firms take their upstream linkages (i.e., their suppliers) as given, but can choose downstream linkages (i.e., customers) in every period as long as they pay a fixed cost for every linkage they form. Specifically, if a firm in industry i wishes to sell its variety as intermediate inputs to a firm in industry i', it first needs to pay a fixed cost of $\kappa/\phi_{i'}$ units of the final good to establish a supply-customer relationship, where $\phi_{i'}$ is the linkage efficiency of the buyer firm. Profit maximization by firms implies the following input demand:

$$x_{i,i'} = m_i \left(P_i^S\right)^{\frac{1}{1-\alpha}} p_{i'}^{\frac{1}{\alpha-1}} \left(\mu_{i,i'} N_{i'}\right)^{\frac{\alpha-\beta}{\beta(1-\alpha)}}$$

$$m_i = q_i \left(\frac{w}{\sigma}\right)^{\sigma} \left(\frac{P_i^S}{1-\sigma}\right)^{-\sigma}$$

$$l_i = q_i \left(\frac{w}{\sigma}\right)^{\sigma-1} \left(\frac{P_i^S}{1-\sigma}\right)^{1-\sigma}$$

where $P_i^S \equiv \left[\int_0^1 p_{i'}^{\frac{\alpha}{\alpha-1}} \left(\mu_{i,i'} N_{i'}\right)^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)} di'\right]^{\frac{\alpha-1}{\alpha}}$ is the producer price index facing an industry-i firm. The firm sets monopolistic price to maximize operating profit:

$$\max_{p_i} \left(1+\psi\right) q_i p_i - q_i \left(\frac{w}{\sigma}\right)^{\sigma} \left(\frac{P_i^S}{1-\sigma}\right)^{1-\sigma}$$

where $\psi \geqslant 0$ is a subsidy financed by lump-sum taxation. The case of $\psi = 0$ corresponds to the laissez-faire market equilibrium. Optimal pricing implies

$$p_i = \frac{1}{\left(1+\psi\right)\beta} \left(\frac{w}{\sigma}\right)^{\sigma} \left(\frac{P_i^S}{1-\sigma}\right)^{1-\sigma} \tag{2.7.1}$$

The firm chooses downstream linkages to maximize flow profit, subject to the demand function for its variety:

$$\max_{\left\{\mu_{i',i}\right\}_{i'\in[0,1]}} \quad \pi_i = \left(1+\psi\right)\left(1-\beta\right) q_i p_i - \int_0^1 \frac{\kappa}{\phi_{i'}} \mu_{i',i} N_{i'} di' \tag{2.7.2}$$

$$\text{s.t.} \quad q_i = X_i + \int_0^1 x_{i',i} \mu_{i',i} N_{i'} di'$$

The first order condition for linkage formation is

$$\left(1+\psi\right)\left(1-\beta\right) x_{i',i} p_i - \frac{\kappa}{\phi_{i'}} \begin{cases} \geqslant 0 & \text{if } \mu_{i',i} = 1 \\ = 0 & \text{if } \mu_{i',i} \in [0,1] \\ \leqslant 0 & \text{if } \mu_{i',i} = 0 \end{cases}$$

Characterization of the Static Equilibrium

Using the final good as the numeraire, I normalize $P = 1$. Substituting the final demand and the intermediate demand functions into the market clearing condition for a variety, we have

$$q_i = X_i + \int_0^1 x_{i',i} \mu_{i',i} N_{i'} di'$$

$$\Leftrightarrow \quad q_i p_i^{\frac{1}{1-\alpha}} N_i^{\frac{\beta-\alpha}{\beta(1-\alpha)}} = Y + \int_0^1 \left(1+\psi\right)\beta\left(1-\sigma\right) q_{i'} p_{i'} \left(P_{i'}^S\right)^{\frac{\alpha}{1-\alpha}} \mu_{i',i}^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)} N_{i'} di' \tag{2.7.3}$$

In order for the last equality to hold for all i, we must have that $\mu_{i',i}$ does not depend on the supplier industry i (so that $\mu_{i',i} = \tilde{\mu}_i$ which is to be determined) and $q_i p_i^{\frac{1}{1-\alpha}} N_i^{\frac{\beta-\alpha}{\beta(1-\alpha)}} \equiv D$ is a constant to be determined. Furthermore, the free mobility of firms across industries implies that flow profit

(2.7.2) must be equalized throughout the economy, which in turn implies that gross output of a firm $q_i p_i$ is constant. Substituting the labor demand $l_i = (1+\psi)\beta\sigma q_i p_i / w$ into the labor market clearing condition $\int_0^1 l_i N_i di = L$, we have $q_i p_i = Lw/[(1+\psi)\sigma\beta N]$, which implies that $p_i^{\frac{\alpha}{\alpha-1}} N_i^{\frac{\alpha-\beta}{\beta(1-\alpha)}} = q_i p_i / D = Lw/[(1+\psi)\sigma\beta ND]$. Substituting the last result into the price normalization $\int_0^1 p_i^{\frac{\alpha}{\alpha-1}} N_i^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)} di = 1$, we have $D = Lw/[(1+\psi)\sigma\beta]$. Using the above results to rewrite the optimal pricing condition (2.7.1), we have

$$p_i = \frac{1}{(1+\psi)\beta}\left(\frac{w}{\sigma}\right)^{\sigma}\left\{\frac{1}{1-\sigma}\left[\int_0^1 p_{i'}^{\frac{\alpha}{\alpha-1}}\left(\tilde{\mu}_i N_{i'}\right)^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)} di'\right]^{\frac{\alpha-1}{\alpha}}\right\}^{1-\sigma}$$

$$\Leftrightarrow \quad p_i = \frac{1}{(1+\psi)\beta}\left(\frac{w}{\sigma}\right)^{\sigma}\left(\frac{1}{1-\sigma}\tilde{\mu}_i^{\frac{\beta-1}{\beta}}\right)^{1-\sigma}$$

Substituting the above results into the first order condition for linkage formation, we have the following equilibrium distribution of linkages:

$$\tilde{\mu}_i = \begin{cases} \frac{\phi_i}{\underline{\phi}} & \text{if } \phi_i < \underline{\phi} \\ 1 & \text{if } \phi_i \geqslant \underline{\phi} \end{cases}$$

where the linkage efficiency cutoff is

$$\underline{\phi} = \frac{\kappa\left[1-(1+\psi)\beta(1-\sigma)\right]N^2}{(1+\psi)^2\beta(1-\beta)(1-\sigma)Y}$$

Accordingly, the equilibrium distribution of firm mass across industries is given by $N_i/N = \tilde{\mu}_i^{(1-\sigma)\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)}/A$ where $A \equiv \int_0^1 \tilde{\mu}_i^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)} di$. Finally, substituting all of the above results into (2.7.3) and the identity $\int_0^1 N_i di = N$, we obtain the expressions of the two aggregate variables in the static equilibrium:

$$w = \left[\frac{(1+\psi)\sigma\beta}{1-(1+\psi)\beta(1-\sigma)}\right]\frac{Y}{L}$$

$$Y = \left[1-(1+\psi)\beta(1-\sigma)\right]\left[(1+\psi)\beta(1-\sigma)\right]^{\frac{1-\sigma}{\sigma}} LA^{\frac{\beta-\alpha}{\sigma\alpha\beta}} N^{\frac{1-\beta}{\sigma\beta}}$$

The Dynamic Problem of the Households

Households allocate disposable income (wage and interest earnings minus the lump-sum tax) between consumption and the holding of assets M, implying the dynamic budget constraint as follows:

$$\dot{M}_t = w_t L + r_t M_t - T_t - C_t$$

Assets in this economy take the form of a mutual fund aggregating the value of all firms:

$$M_t = \int_0^1 V_{i,t} N_{i,t} di \quad \text{where} V_{i,t} = \int_t^\infty e^{-\int_t^\tau (r_s+\delta)ds} \pi_{i,\tau} d\tau$$

Entrepreneurs borrow from the households at interest rate r_t to create firms, upon paying the entry sunk cost ν in units of the final good. The free entry condition thus implies

$$\int_0^1 V_{i,t} di = \nu$$

The lump-sum tax for financing the variety production subsidy amounts to

$$\begin{aligned} T &\equiv \int_0^1 \psi q_i p_i N_i di \\ &= \left[\frac{\psi}{1-(1+\psi)\beta(1-\sigma)} \right] Y \end{aligned}$$

Since flow profits π are equalized across industries due to the free mobility of firms, firm value V is also the same in all industries. Substituting the Bellman equation of firm value $(r_t+\delta) V_t = \pi_t + \dot{V}_t$, the free entry condition $V_t = \nu$ and the expression of T_t into the dynamic budget constraint of the households, we have

$$\begin{aligned} \dot{M}_t &= w_t L + r_t M_t - T_t - C_t \\ \Leftrightarrow \quad \nu N_t^e &= \left[\frac{(1+\psi)\sigma\beta - \psi}{1-(1+\psi)\beta(1-\sigma)} \right] Y_t + \pi_t N_t - C_t \\ \Leftrightarrow \quad C_t &= Y_t - \kappa \frac{Z_t}{A_t} N_t^2 - \nu N_t^e \end{aligned} \tag{2.7.4}$$

where the last equality follows from the expression of the flow profit:

$$\begin{aligned} \pi_i &= (1+\psi)(1-\beta) q_i p_i - \int_0^1 \frac{\kappa}{\phi_{i'}} \mu_{i',i} N_{i'} di' \\ &= \left[\frac{(1+\psi)(1-\beta)}{1-(1+\psi)\beta(1-\sigma)} \right] \frac{Y_t}{N_t} - \kappa \frac{Z_t}{A_t} N_t \end{aligned}$$

where $Z \equiv \int_0^1 \phi_i^{-1} \tilde{\mu}_i^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)+1} di.$

Implement the Social Optimum

To correct the monopolistic markup, the production subsidy ψ should be set such that $(1+\psi)\beta = 1$, which implies the optimal subsidy

$$\psi = \frac{1-\beta}{\beta}$$

It is straightforward to check that, once the optimal subsidy ψ is in place, the linkage efficiency cutoff $\underline{\phi}$ in the decentralized economy is aligned with the planner's solution, so are the distribution of linkages $\{\tilde{\mu}_i\}_{i\in[0,1]}$ and the derived expressions A and Z. Conditional on the same firm stock N, aggregate output Y in the market equilibrium under the optimal subsidy is the same as in the planner's allocation. Furthermore, once the static inefficiency is corrected, the household dynamic budget constraint (2.7.4) agrees with the aggregate resource constraint in the planner's problem, and therefore the market equilibrium replicates the social optimum.

2.8 Appendix C: Proofs and derivation

Proof of Proposition 1

First we observe that, in order for the optimality condition (2.2.2) to hold, we must have $\partial Y/\partial N_i$ be constant across industries and $\mu_{i,i'}$ be the same across supplier industries indexed by i', which allows us to write

$$\mu_{i,i'} \equiv \tilde{\mu}_i \quad \text{for all } i,i'$$

The first order condition with respect to l_i states

$$\sigma Y^{1-\alpha} X_i^{\alpha-1} N_i^{\frac{\alpha-\beta}{\beta}} q_i l_i^{-1} = \lambda_L \quad \text{for all } i \tag{2.8.1}$$

where λ_L is the Lagrange multiplier corresponding to the labor supply constraint $\int_0^1 l_i N_i di = L$. The optimality condition (2.2.1) implies

$$x_{i,i'} = \left[(1-\sigma) X_i^{\alpha-1} N_i^{\frac{\alpha-\beta}{\beta}} l_i^{\frac{\alpha\sigma}{1-\sigma}} q_i^{\frac{1-\sigma-\alpha}{1-\sigma}} \tilde{\mu}_i^{\frac{\alpha-\beta}{\beta}} X_{i'}^{1-\alpha}\right]^{\frac{1}{1-\alpha}} \quad \text{for all } i,i' \tag{2.8.2}$$

which can be substituted into the production function of differentiated varieties to yield

$$q_i = l_i \left[(1-\sigma) X_i^{\alpha-1} N_i^{\frac{\alpha-\beta}{\beta}} \tilde{\mu}_i^{\frac{1-\beta}{\beta}} Y^{1-\alpha}\right]^{\frac{1-\sigma}{\sigma}} \quad \text{for all } i \tag{2.8.3}$$

Substituting (2.8.1), (2.8.2), and (2.8.3) into the identity of variety quantity, we have

$$X_i = q_i - \int_0^1 x_{i',i} \mu_{i',i} N_{i'} di'$$

$$
\begin{aligned}
X_i &= l_i\left[(1-\sigma)X_i^{\alpha-1}N_i^{\frac{\alpha-\beta}{\beta}}\tilde{\mu}_i^{\frac{1-\beta}{\beta}}Y^{1-\alpha}\right]^{\frac{1-\sigma}{\sigma}} \\
&\quad -\int_0^1 Y^{\frac{1-\sigma-\alpha}{\sigma}}\left[(1-\sigma)X_{i'}^{\alpha-1}N_{i'}^{\frac{\alpha-\beta}{\beta}}\right]^{\frac{1}{\sigma}} l_{i'}\tilde{\mu}_{i'}^{\frac{\alpha-\beta}{\beta(1-\alpha)}+\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)\frac{1-\sigma-\alpha}{\alpha\sigma}} X_i\tilde{\mu}_{i'}N_{i'}di' \\
X_i &= l_i\lambda_L\left(\sigma Y^{1-\alpha}X_i^{\alpha-1}N_i^{\frac{\alpha-\beta}{\beta}}\right)^{-1} \\
&\quad -\int_0^1 Y^{\frac{1-\sigma-\alpha}{\sigma}}\left[(1-\sigma)X_{i'}^{\alpha-1}N_{i'}^{\frac{\alpha-\beta}{\beta}}\right]^{\frac{1}{\sigma}} l_{i'}\tilde{\mu}_{i'}^{\left(\frac{1-\beta}{\beta}\right)\left(\frac{1-\sigma}{\sigma}\right)} X_iN_{i'}di'
\end{aligned}
$$

$$
\begin{aligned}
\Leftrightarrow \int_0^1 X_i^{\alpha}N_i^{\frac{\alpha}{\beta}}di &= \lambda_L\left(\sigma Y^{1-\alpha}\right)^{-1}\int_0^1 l_iN_idi \\
&\quad -\int_0^1\int_0^1 X_i^{\alpha-1}N_i^{\frac{\alpha}{\beta}}Y^{\frac{1-\sigma-\alpha}{\sigma}}\left[(1-\sigma)X_{i'}^{\alpha-1}N_{i'}^{\frac{\alpha-\beta}{\beta}}\right]^{\frac{1}{\sigma}} l_{i'}\tilde{\mu}_{i'}^{\left(\frac{1-\beta}{\beta}\right)\left(\frac{1-\sigma}{\sigma}\right)} X_iN_{i'}di'di \\
\int_0^1 X_i^{\alpha}N_i^{\frac{\alpha}{\beta}}di &= \frac{1}{\sigma}\lambda_LY^{\alpha-1}\int_0^1 l_iN_idi - \left(\frac{1-\sigma}{\sigma}\right)\lambda_LY^{-1}\int_0^1 l_{i'}N_{i'}di'\int_0^1 X_i^{\alpha}N_i^{\frac{\alpha}{\beta}}di \\
Y^{\alpha} &= \lambda_LY^{\alpha-1}L \\
\lambda_L &= \frac{Y}{L}
\end{aligned}
$$

From (2.8.1), (2.8.3) and $\lambda_L = Y/L$, we have

$$
X_i = \left[\sigma(1-\sigma)^{\frac{1-\sigma}{\sigma}}L\right]^{\frac{\sigma}{1-\alpha}} Y^{1-\frac{\sigma}{1-\alpha}} N_i^{\frac{\alpha-\beta}{\beta(1-\alpha)}} \tilde{\mu}_i^{\left(\frac{1-\beta}{\beta}\right)\left(\frac{1-\sigma}{1-\alpha}\right)} \tag{2.8.4}
$$

which we then substitute into the aggregate production function, yielding

$$
Y = \sigma(1-\sigma)^{\frac{1-\sigma}{\sigma}}L\left[\int_0^1 \tilde{\mu}_i^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)(1-\sigma)} N_i^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)}di\right]^{\frac{1-\alpha}{\sigma\alpha}} \tag{2.8.5}
$$

The observation that $\partial Y/\partial N_i$ must be constant across industries then implies (2.2.6). By the first order condition (2.2.3), the interior solution of $\mu_{i,i'} \equiv \tilde{\mu}_i$ is given by

$$
\begin{aligned}
\frac{\partial Y}{\partial \tilde{\mu}_i} &= \frac{\kappa}{\phi_i}N_iN \\
\left(\frac{1-\beta}{\beta}\right)\left(\frac{1-\sigma}{\sigma}\right)Y\frac{\tilde{\mu}_i^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)(1-\sigma)-1}N_i^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)}}{\int_0^1 \tilde{\mu}_i^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)(1-\sigma)}N_i^{\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)}di} &= \frac{\kappa}{\phi_i}N_iN
\end{aligned}
$$

$$\Leftrightarrow \left(\frac{1-\beta}{\beta}\right)\left(\frac{1-\sigma}{\sigma}\right)Y = \frac{\kappa}{\phi_i}\tilde{\mu}_i N^2$$

where the last line follows from (2.2.6). Therefore, the solution of $\mu_{i,i'} \equiv \tilde{\mu}_i$ is given by

$$\tilde{\mu}_i = \begin{cases} \frac{\phi_i}{\kappa}\left(\frac{1-\sigma}{\sigma}\right)\left(\frac{1-\beta}{\beta}\right)\frac{Y}{N^2} & \text{if } \phi_i < \kappa\left(\frac{\sigma}{1-\sigma}\right)\left(\frac{\beta}{1-\beta}\right)\frac{N^2}{Y} \\ 1 & \text{if } \phi_i \geqslant \kappa\left(\frac{\sigma}{1-\sigma}\right)\left(\frac{\beta}{1-\beta}\right)\frac{N^2}{Y} \end{cases}$$

From (2.8.2), (2.8.3), and the identity of variety quantity, we also have

$$q_i = X_i + \sigma(1-\sigma)^{\frac{1}{\sigma}} Y^{-\alpha+\frac{1-\sigma-\alpha}{\sigma}} LX_i \int_0^1 q_{i'}\left(X_{i'}^{\alpha-1} N_{i'}^{\frac{\alpha-\beta}{\beta}}\right)^{\frac{1+\sigma}{\sigma}} \mu_{i'}^{\left(\frac{1-\beta}{\beta}\right)\left(\frac{1-\sigma}{\sigma}\right)} N_{i'} di'$$

$$q_i = X_i + \frac{\sigma(1-\sigma)^{\frac{1}{\sigma}} Y^{-\alpha+\frac{1-\sigma-\alpha}{\sigma}} LX_i \int_0^1 X_i\left(X_i^{\alpha-1} N_i^{\frac{\alpha-\beta}{\beta}}\right)^{\frac{1+\sigma}{\sigma}} \mu_i^{\left(\frac{1-\beta}{\beta}\right)\left(\frac{1-\sigma}{\sigma}\right)} N_i di}{1-\sigma(1-\sigma)^{\frac{1}{\sigma}} Y^{-\alpha+\frac{1-\sigma-\alpha}{\sigma}} L\int_0^1 X_i\left(X_i^{\alpha-1} N_i^{\frac{\alpha-\beta}{\beta}}\right)^{\frac{1+\sigma}{\sigma}} \mu_i^{\left(\frac{1-\beta}{\beta}\right)\left(\frac{1-\sigma}{\sigma}\right)} N_i di}$$

$$\frac{q_i}{X_i} = \frac{1}{\sigma}$$

Finally, substituting $q_i = X_i/\sigma$ and (2.2.6) back into (2.8.2) and (2.8.3) leads to $l_i = L/N$ and (2.2.7). Substituting (2.2.6) back into (2.2.8) yields the expression of aggregate output (2.2.8).

Proof of Proposition 2

Using the definition of A and (2.2.4), we have

$$\begin{aligned} A &\equiv \int_0^1 \mu_i^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)} di \\ &= \int_{\phi_{min}}^{\underline{\phi}} \mu(\phi)^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)} d\mathscr{F}(\phi) + \int_{\underline{\phi}}^{\phi_{max}} \mu(\phi)^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)} d\mathscr{F}(\phi) \\ &= \int_{\phi_{min}}^{\underline{\phi}} \left(\frac{\phi}{\underline{\phi}}\right)^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)} d\mathscr{F}(\phi) + \int_{\underline{\phi}}^{\phi_{max}} d\mathscr{F}(\phi) \end{aligned}$$

Time-differentiating A yields

$$\begin{aligned}\dot{A} &= \frac{dA}{d\underline{\phi}}\dot{\underline{\phi}} \\ &= \left[-\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)\underline{\phi}^{-1}\int_{\phi_{min}}^{\underline{\phi}}\left(\frac{\phi}{\underline{\phi}}\right)^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)} d\mathscr{F}(\phi) + \mathscr{F}'\left(\underline{\phi}\right) - \mathscr{F}'\left(\underline{\phi}\right)\right]\dot{\underline{\phi}} \\ &= -\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)\int_{\phi_{min}}^{\underline{\phi}}\left(\frac{\phi}{\underline{\phi}}\right)^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)} d\mathscr{F}(\phi)\frac{\dot{\underline{\phi}}}{\underline{\phi}}\end{aligned}$$

Time-differentiating both sides of (2.3.1) and using the expression of $\dot{A}$ derived above, we have (2.3.2), and (2.3.3) follows directly from (2.3.2).

Proof of Proposition 3

If $\underline{\phi} \geqslant \phi_{max}$, from Proposition 2 we have $\tilde{\mu}_i = \phi_i/\underline{\phi}$ for all i, which then implies

$$A = \int_{\phi_{min}}^{\phi_{max}}\left(\frac{\phi}{\underline{\phi}}\right)^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)} d\mathscr{F}(\phi) \quad \text{and} \quad Z = \frac{1}{\underline{\phi}}\int_{\phi_{min}}^{\phi_{max}}\left(\frac{\phi}{\underline{\phi}}\right)^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)} d\mathscr{F}(\phi)$$

In this case, aggregate output and total linkage fixed costs are respectively given by

$$Y_t = \Lambda N_t^{\frac{[1-2(1-\sigma)]\frac{1-\beta}{\sigma\beta}}{1-\left(\frac{1-\sigma}{\sigma}\right)\left(\frac{1-\beta}{\beta}\right)}} \quad \text{and} \quad \kappa\frac{Z_t}{A_t}N_t^2 = \left(\frac{1-\sigma}{\sigma}\right)\left(\frac{1-\beta}{\beta}\right)Y_t$$

where Λ is a constant given by:

$$\Lambda^{1-\left(\frac{1-\sigma}{\sigma}\right)\left(\frac{1-\beta}{\beta}\right)} \equiv \sigma(1-\sigma)^{\frac{1-\sigma}{\sigma}} L\left[\frac{1}{\kappa}\left(\frac{1-\sigma}{\sigma}\right)\left(\frac{1-\beta}{\beta}\right)\right]^{\left(\frac{1-\sigma}{\sigma}\right)\left(\frac{1-\beta}{\beta}\right)} \times \left[\int_{\phi_{min}}^{\phi_{max}}\phi^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)}d\mathscr{F}(\phi)\right]^{\frac{\beta-\alpha}{\sigma\alpha\beta}}$$

Substituting the above results into (2.2.9) and (2.2.10) yields the dynamic system (2.3.4) and (2.3.5). In the steady state, total firm mass is:

$$N^{SS} = \left\{[1-2(1-\sigma)]\left(\frac{1-\beta}{\sigma\beta}\right)\frac{\Lambda}{\nu(\delta+\rho)}\right\}^{\frac{1-\left(\frac{1-\sigma}{\sigma}\right)\left(\frac{1-\beta}{\beta}\right)}{1-\frac{1-\beta}{\sigma\beta}+\left(\frac{1-\sigma}{\sigma}\right)\left(\frac{1-\beta}{\beta}\right)}}$$

which implies $dN/d\kappa < 0$ and $dN/dL > 0$. The steady state is locally saddle-path stable provided that

$$\frac{\partial \dot{C}_t}{\partial N_t}|_{N^{SS}} < 0 \quad \Leftrightarrow \quad \frac{[1-2(1-\sigma)]^{\frac{1-\beta}{\sigma\beta}}}{1-\left(\frac{1-\sigma}{\sigma}\right)\left(\frac{1-\beta}{\beta}\right)} - 1 < 0$$

which holds under the parameter restrictions $2(1-\sigma) < 1$ and $(1-\beta)/\beta < \sigma$. Provided that the initial condition $N_0 < N^{SS}$, the linkage efficiency cutoff $\underline{\phi}$ rises over time ($\dot{\underline{\phi}} > 0$) since $d\underline{\phi}/dN > 0$. Therefore, the condition on κ for this case to prevail is

$$\underline{\phi}_0 \geqslant \phi_{max}$$
$$\Leftrightarrow \kappa \geqslant \overline{\kappa} \equiv \phi_{max}\left(\frac{1-\beta}{\beta}\right)(1-\sigma)^{\frac{1}{\sigma}} L \left[\int_{\phi_{min}}^{\phi_{max}} \left(\frac{\phi}{\phi_{max}}\right)^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)} d\mathscr{F}(\phi)\right]^{\frac{\beta-\alpha}{\sigma\alpha\beta}} N_0^{\frac{1-\beta}{\sigma\beta}-2}$$

In anticipation for the proof of Proposition 4, we also derive another threshold of κ defined by

$$\underline{\phi}^{SS} \geqslant \phi_{max}$$
$$\Leftrightarrow \kappa \geqslant \overline{\kappa}' \equiv \phi_{max}(1-\sigma)\left[\left(\frac{1-\beta}{\beta}\right)(1-\sigma)^{\frac{1-\sigma}{\sigma}} L\right]^{\frac{1}{\frac{1-\beta}{\sigma\beta}-1}} \left[\frac{\nu(\delta+\rho)}{1-2(1-\sigma)}\right]^{\frac{2-\frac{1-\beta}{\sigma\beta}}{1-\frac{1-\beta}{\sigma\beta}}} \times$$
$$\left[\int_{\phi_{min}}^{\phi_{max}} \left(\frac{\phi}{\phi_{max}}\right)^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)} d\mathscr{F}(\phi)\right]^{\frac{\frac{\beta-\alpha}{\sigma\alpha\beta}}{\frac{1-\beta}{\sigma\beta}-1}} \tag{2.8.6}$$

If $\underline{\phi} \leqslant \phi_{min}$, from Proposition 2 we have$\tilde{\mu}_i = 1$ for all i, which then implies

$$A = 1 \quad \text{and} \quad Z = \int_{\phi_{min}}^{\phi_{max}} \frac{1}{\phi} d\mathscr{F}(\phi)$$

In this case, aggregate output and total linkage fixed costs are respectively given by

$$Y_t = \sigma(1-\sigma)^{\frac{1-\sigma}{\sigma}} L N_t^{\frac{1-\beta}{\sigma\beta}} \quad \text{and} \quad \kappa\frac{Z_t}{A_t}N_t^2 = \int_{\phi_{min}}^{\phi_{max}} \frac{\kappa}{\phi} d\mathscr{F}(\phi) N_t^2$$

Substituting the above results into (2.2.9) and (2.2.10) yields the dynamic system (2.3.6) and (2.3.7). In the steady state, total firm mass N^{SS} is given implicitly by the following equation:

$$\frac{1}{\nu}\left(\frac{1-\beta}{\beta}\right)(1-\sigma)^{\frac{1-\sigma}{\sigma}}L\left(N^{SS}\right)^{\frac{1-\beta}{\sigma\beta}-1}-\frac{2}{\nu}\int_{\phi_{min}}^{\phi_{max}}\frac{\kappa}{\phi}d\mathscr{F}(\phi)N^{SS}-\delta-\rho=0$$

which implies $dN^{SS}/d\kappa<0$ and $dN^{SS}/dL>0$ by the Implicit Function Theorem. The steady state is locally saddle-path stable provided that

$$\frac{\partial\dot{C}_t}{\partial N_t}|_{N^{SS}}<0$$

which holds under the parameter restrictions $(1-\beta)/\beta<\sigma$. Since $\underline{\dot{\phi}}>0$ as argued above, the condition on κ for this case to prevail is

$$\begin{aligned}&\underline{\phi}^{SS}\leqslant\phi_{min}\\ &\Leftrightarrow\kappa\leqslant\underline{\kappa}\equiv\phi_{min}(1-\sigma)\left[\left(\frac{1-\beta}{\beta}\right)(1-\sigma)^{\frac{1-\sigma}{\sigma}}L\right]^{\frac{1}{\frac{1-\beta}{\sigma\beta}-1}}\times\\ &\left[\frac{\nu(\delta+\rho)}{1-2(1-\sigma)\int_{\phi_{min}}^{\phi_{max}}\frac{\phi_{min}}{\phi}d\mathscr{F}(\phi)}\right]^{\frac{2-\frac{1-\beta}{\sigma\beta}}{1-\frac{1-\beta}{\sigma\beta}}}\end{aligned}$$

In anticipation for the proof of Proposition 4, we also derive another threshold of κ defined by

$$\begin{aligned}\underline{\phi}_0&\leqslant\phi_{min}\\ \Leftrightarrow\quad\kappa&\leqslant\underline{\kappa}'\equiv\phi_{min}L\left(\frac{1-\beta}{\beta}\right)(1-\sigma)^{\frac{1}{\sigma}}N_0^{\frac{1-\beta}{\sigma\beta}-2}\end{aligned}\tag{2.8.7}$$

Finally, since $d\underline{\kappa}/dL<0$ while $d\overline{\kappa}/dL>0$, $\underline{\kappa}<\overline{\kappa}$ is satisfied when L is sufficiently large.

Proof of Proposition 4

If $\phi_{min}<\underline{\phi}<\phi_{max}$, Proposition 2 implies

$$A=\int_{\phi_{min}}^{\underline{\phi}}\left(\frac{\phi}{\underline{\phi}}\right)^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)}d\mathscr{F}(\phi)+\int_{\underline{\phi}}^{\phi_{max}}d\mathscr{F}(\phi)$$

$$Z=\frac{1}{\underline{\phi}}\int_{\phi_{min}}^{\underline{\phi}}\left(\frac{\phi}{\underline{\phi}}\right)^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)}d\mathscr{F}(\phi)+\int_{\underline{\phi}}^{\phi_{max}}\frac{1}{\phi}d\mathscr{F}(\phi)$$

Substituting the above results into (2.2.9) and (2.2.10) yields the differential equations (2.3.8) and (2.3.9). Both of these two differential equations depend on $\underline{\phi}_t$. To derive the law of motion of $\underline{\phi}_t$, we first substitute the expression of aggregate output (2.2.8) into that of $\underline{\phi}_t$ (2.2.5), obtaining an equation relating total firm mass N_t to $\underline{\phi}_t$:

$$N_t = \left[\left(\frac{1-\beta}{\beta}\right)(1-\sigma)^{\frac{1}{\sigma}}\frac{L}{\kappa}\underline{\phi}_t A_t^{\frac{\beta-\alpha}{\sigma\alpha\beta}}\right]^{\frac{1}{2-\frac{1-\beta}{\sigma\beta}}} \tag{2.8.8}$$

Time-differentiating both sides of the above equation yields the third differential equation (2.3.10). To study the local stability of this system, we substituting (2.3.10) into (2.3.9) to obtain

$$\frac{\dot{\underline{\phi}}_t}{\underline{\phi}_t} = \frac{\left(2-\frac{1-\beta}{\sigma\beta}\right)A_t\left\{\frac{1}{v}\left[\sigma(1-\sigma)^{\frac{1-\sigma}{\sigma}}LA_t^{\frac{\beta-\alpha}{\sigma\alpha\beta}}N_t^{\frac{1-\beta}{\sigma\beta}-1}-\kappa\frac{Z_t}{A_t}N_t-\frac{C_t}{N_t}\right]-\delta\right\}}{A_t-\left(\frac{1-\sigma}{\sigma}\right)\left(\frac{1-\beta}{\beta}\right)\int_{\phi_{min}}^{\underline{\phi}_t}\left(\frac{\phi}{\underline{\phi}_t}\right)^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)}d\mathscr{F}(\phi)}$$

Using (2.8.8), the above differential equation and (2.3.8) constitute a dynamic system of C_t and $\underline{\phi}_t$ only, where the partial derivatives evaluated at the steady state satisfy: $\partial\dot{C}_t/\partial C_t|_{SS}=0$, $\partial\dot{\underline{\phi}}_t/\partial C_t|_{SS}<0$, and $\partial\dot{\underline{\phi}}_t/\partial\underline{\phi}_t|_{SS}>0$. Thus, the system is locally saddle-path stable provided that $\partial\dot{C}_t/\partial\underline{\phi}_t|_{SS}<0$. A sufficient parameter restriction for $\partial\dot{C}_t/\partial\underline{\phi}_t|_{SS}<0$ is

$$\frac{\underline{\phi}_{min}}{\phi_{max}} > \frac{\alpha(1-\beta)(1-\sigma)}{(1-\alpha)\beta-\alpha(1-\beta)\sigma}$$

which guarantees that

$$\frac{1-\alpha}{\sigma\alpha}\left[1-\frac{\alpha}{\beta}\left(\frac{1-\beta}{1-\alpha}\right)\right]\mathscr{E}_A^{\underline{\phi}}|_{SS}-\mathscr{E}_{Z/A}^{\underline{\phi}}|_{SS}-\left[2-\frac{1}{\sigma}\left(\frac{1-\beta}{\beta}\right)\right]\mathscr{E}_N^{\underline{\phi}}|_{SS}<0$$

$$\Leftrightarrow\left[1+\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)\right]\left[1-\frac{\int_{\underline{\phi}}^{\phi_{max}}\frac{\phi}{\underline{\phi}}d\mathscr{F}(\phi)}{\int_{\underline{\phi}}^{\phi_{max}}d\mathscr{F}(\phi)}\right]<$$

$$1+\frac{\int_{\underline{\phi}}^{\phi_{max}}\frac{\phi}{\underline{\phi}}d\mathscr{F}(\phi)}{\int_{\phi_{min}}^{\underline{\phi}}\left(\frac{\phi}{\underline{\phi}}\right)^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)}d\mathscr{F}(\phi)} \tag{2.8.9}$$

where $\mathscr{E}_A^{\underline{\phi}}|_{SS}$, $\mathscr{E}_{Z/A}^{\underline{\phi}}|_{SS}$, and $\mathscr{E}_N^{\underline{\phi}}|_{SS}$ are respectively the elasticity of A, Z/A, and N with respect to $\underline{\phi}$ evaluated at the steady state. The steady-state values

N^{SS} and $\underline{\phi}^{SS}$ are jointly defined implicitly by the following two equations:

$$\nu(\delta+\rho) = \left(\frac{1-\beta}{\beta}\right)(1-\sigma)^{\frac{1-\sigma}{\sigma}} L \left(A^{SS}\right)^{\frac{\beta-\alpha}{\sigma\alpha\beta}} \left(N^{SS}\right)^{\frac{1-\beta}{\sigma\beta}-1} - 2\kappa \frac{Z^{SS}}{A^{SS}} N^{SS} \tag{2.8.10}$$

$$N^{SS} = \left[\left(\frac{1-\beta}{\beta}\right)(1-\sigma)^{\frac{1}{\sigma}} \frac{L}{\kappa} \underline{\phi}^{SS} \left(A^{SS}\right)^{\frac{\beta-\alpha}{\sigma\alpha\beta}}\right]^{\frac{1}{2-\frac{1-\beta}{\sigma\beta}}} \tag{2.8.11}$$

where A^{SS} and Z^{SS} are A and Z (both are functions of $\underline{\phi}$) evaluated at the steady state. Substituting (2.8.11) into (2.8.11), applying the Implicit Function Theorem and using $\partial \dot{C}_t / \partial \underline{\phi}_t|_{SS} < 0$, we have $d\underline{\phi}^{SS}/d\kappa > 0$ and $d\underline{\phi}^{SS}/dL > 0$. We can also show that the product $\underline{\phi}^{SS}\left(A^{SS}\right)^{\frac{\beta-\alpha}{\sigma\alpha\beta}}$ is increasing in $\underline{\phi}^{SS}$, which implies that $dN^{SS}/dL > 0$. Furthermore, under the aforementioned parameter restriction that guarantee $\partial \dot{C}_t / \partial \underline{\phi}_t|_{SS} < 0$, the elasticity of the product $\underline{\phi}^{SS}\left(A^{SS}\right)^{\frac{\beta-\alpha}{\sigma\alpha\beta}}$ with respect to κ is between 0 and 1, and therefore $dN^{SS}/d\kappa < 0$. Finally, given initial condition $N_0 < N^{SS}$, the dynamic system given by (2.3.8), (2.3.9), and (2.3.10) is applicable to the entire transition to the steady state if the linkage fixed cost parameter κ satisfies $\underline{\kappa}' \leqslant \kappa \leqslant \overline{\kappa}'$, where the threshold levels $\underline{\kappa}'$ and $\overline{\kappa}'$ are given by (2.8.6) and (2.8.7) in the proof of Proposition 3. The condition for $\underline{\kappa}' > \underline{\kappa}$ and $\overline{\kappa}' < \overline{\kappa}$ to both be satisfied is

$$L > \left(\frac{\beta}{1-\beta}\right)(1-\sigma)^{\frac{\sigma-1}{\sigma}} \left[\frac{\nu(\delta+\rho)}{1-2(1-\sigma)}\right] N_0^{1-\frac{1-\beta}{\sigma\beta}} \times$$
$$\left[\int_{\phi_{min}}^{\phi_{max}} \left(\frac{\phi}{\phi_{max}}\right)^{\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)} d\mathscr{F}(\phi)\right]^{\frac{\alpha-\beta}{\sigma\alpha\beta}}$$

which also ensures that $\underline{\kappa} < \overline{\kappa}$.

Proof of Proposition 5

In the proof of Proposition 1, we established that $q_i = X_i/\sigma$, which can be substituted into (2.8.3) to yield

$$X_i = N^{-\frac{1}{\alpha}} Y N_i^{\frac{\beta-\alpha}{\alpha\beta}}$$

Substituting the above equation and (2.2.7) into the definition of intermediate expenditure shares Λ_i, we have

$$\begin{aligned}\Lambda_i &= \frac{\int_0^1 \left(\frac{1-\sigma}{\sigma}\right)\frac{1}{N}X_i\mu_{i'}^{-1}\mu_{i'}N_iN_{i'}di'Y^{1-\alpha}X_i^{\alpha-1}N_i^{\frac{\alpha-\beta}{\beta}}}{\int_0^1\int_0^1 \left(\frac{1-\sigma}{\sigma}\right)\frac{1}{N}X_i\mu_{i'}^{-1}\mu_{i'}N_iN_{i'}di'Y^{1-\alpha}X_i^{\alpha-1}N_i^{\frac{\alpha-\beta}{\beta}}di} \\ &= \frac{N_i}{\int_0^1 N_i di} \\ &= n_i\end{aligned}$$

which implies

$$\frac{\Lambda_i}{\Lambda_{i'}} = \frac{n_i}{n_{i'}} \quad \text{for all } i, i'$$

Time-differentiating the last equation yields (2.3.11).

Derivation details of Section 4.1

In a competitive equilibrium that decentralizes the planner's allocation, the price of an industry-i variety is given by $(\partial Y/\partial X_i)/N_i$. Therefore, gross output of the economy is

$$\begin{aligned}Q &\equiv \int_0^1 q_i\left(\frac{\partial Y}{\partial X_i}N_i^{-1}\right)N_i di \\ &= \int_0^1 \frac{1}{\sigma}X_iY^{1-\alpha}X_i^{\alpha-1}N_i^{\frac{\alpha}{\beta}}di \\ &= \frac{1}{\sigma}Y^{1-\alpha}\int_0^1 X_i^{\alpha}N_i^{\frac{\alpha}{\beta}}di \\ &= \frac{1}{\sigma}Y\end{aligned}$$

Therefore, $\sigma = Y/Q$ corresponds to the ratio of GDP to gross output. To calibrate κ, ν, and ζ, I assume that the U.S. economy was at the steady state in 1997 with the initial linkage efficiency cutoff $\underline{\phi} > \phi_{min}$. Specifically, the steady state is characterized jointly by the following system of four equations:

$$\left(\frac{1-\beta}{\beta\sigma}\right)Y - 2\frac{Z}{A}\kappa N^2 = \nu(\delta+\rho)N$$

$$Y - \frac{Z}{A}\kappa N^2 - C = \nu\delta N$$

$$\left[\frac{L}{\kappa}\left(\frac{1-\beta}{\beta}\right)(1-\sigma)^{\frac{1}{\sigma}}\underline{\phi}A^{\frac{\beta-\alpha}{\sigma\alpha\beta}}\right]^{\frac{1}{2-\frac{1-\beta}{\sigma\beta}}} = N$$

$$\sigma(1-\sigma)^{\frac{1-\sigma}{\sigma}}LA^{\frac{\beta-\alpha}{\sigma\alpha\beta}}N^{\frac{1-\beta}{\sigma\beta}} = Y$$

where

$$A = \zeta\underline{\phi}^{-\zeta}\left[\frac{\underline{\phi}^{\zeta-\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)}-1}{\zeta-\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)}+\frac{1}{\zeta}\right];$$

$$Z = \zeta\underline{\phi}^{-\zeta-1}\left[\frac{\underline{\phi}^{\zeta-\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)}-1}{\zeta-\alpha\left(\frac{1-\beta}{\beta-\alpha}\right)(1-\sigma)}+\frac{1}{\zeta+1}\right].$$

Setting aggregate output Y, aggregate consumption C, and the total number of firms N at their observed values in 1997 (respectively 8.61 trillion USD, 5.56 trillion USD, and 5.37 millions) the above system gives four equations with four unknowns (κ, ν, ζ, and $\underline{\phi}$), which allows us to back out the values of κ, ν, and ζ.

Bibliography

[1] Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. 2012. “The Network Origins of Aggregate Fluctuations.” *Econometrica* 80 (5): 1977–2016.

[2] Adler, Simon, Timo Boppart, and Andreas Müller. 2018. “A Theory of Structural Change That Can Fit the Data.” mimeo.

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

[4] Antràs, Pol, Teresa C. Fort, and Felix Tintelnot. 2017. “The Margins of Global Sourcing: Theory and Evidence from U.S. Firms.” *American Economic Review* 107 (9): 2514–2564.

[5] Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare. 2012. “New Trade Models, Same Old Gains?” *American Economic Review* 102 (1): 94-130.

[6] Atalay, Enghin. 2017. “How Important Are Sectoral Shocks?” *American Economic Journal: Macroeconomics* 9 (4): 254–280.

[7] Axtell, Robert L. 2001. “Zipf Distribution of US Firm Sizes.” *Science* 293 (5536): 1818–1820.

[8] Baqaee, David. 2018. “Cascading Failures in Production Networks.” *Econometrica* forthcoming.

[9] Baumol, William. 1967. “Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis.” *American Economic Review* 57 (3): 415–426.

[10] Bernard, Andrew B., Andreas Moxnes, and Karen Helene Ulltveit-Moe. 2017. “Two-Sided Heterogeneity and Trade.” *Review of Economics and Statistics*, forthcoming.

[11] Blaum, Joaquin, Claire Lelarge, and Michael Peters. 2018. "The Gains from Input Trade with Heterogeneous Importers." *American Economic Journal: Macroeconomics*, forthcoming.

[12] Boppart, Timo. 2014. "Structural Change and the Kaldor Facts in a Growth Model with Relative Price Effects and Non-Gorman Preferences." *Econometrica* 82 (6): 2167–2196.

[13] Broda, Christian, and David Weinstein. 2006. "Globalization and the Gains from Variety." *Quarterly Journal of Economics* 121(2): 541–585.

[14] Buera, Francisco and Joseph Kaboski. 2012. "Scale and the Origins of Structural Change." *Journal of Economic Theory* 147 (2): 684–712.

[15] Caliendo, Lorenzo and Fernando Parro. 2015. "Estimates of the Trade and Welfare Effects of NAFTA." *Review of Economic Studies* 82: 1–44.

[16] Caselli, Francesco, and Wilbur John Coleman II. 2001. "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation" *Journal of Political Economy* 109 (3): 584–617.

[17] Chaney, Thomas. 2014. "The Network Structure of International Trade." *American Economic Review* 104 (11): 3600–3634.

[18] Costinot, Arnaud, and Andrés Rodríguez-Clare. 2014. "Trade Theory with Numbers: Quantifying the Consequences of Globalization." In *Handbook of International Economics*. Vol. 4, edited by Elhanan Helpman, Kenneth Rogoff, and Gita Gopinath, 197–262. Amsterdam: Elsevier North Holland.

[19] Echevarria, Cristina. 1997. "Changes in Sectoral Composition Associated With Economic Growth." *International Economic Review* 38 (2: 431–452.

[20] Ethier, Wilfred J. 1982. "National and International Returns to Scale in the Modern Theory of International Trade." *American Economic Review* 72 (3): 389–405.

[21] Fort, Teresa. 2017. "Technology and Production Fragmentation: Domestic versus Foreign Sourcing." *Review of Economic Studies* 84 (2): 650–687.

[22] Goldberg, Pinelopi K., Amit K. Khandelwal, Nina Pavcnik, and Petia Topalova. 2010. "Imported Intermediate Inputs and Domestic Product

Growth: Evidence from India." *Quarterly Journal of Economics* 125 (4): 1727–1767.

[23] Gopinath, Gita and Brent Neiman. 2014. "Trade Adjustment and Productivity in Large Crises." *American Economic Review* 104 (3): 793–831.

[24] Halpern, László, Miklós Koren, and Adam Szeidl. 2015. "Imported Inputs and Productivity. " *American Economic Review* 105 (12): 3660–3703.

[25] Hummels, David, Jun Ishii, Kei-Mu Yi. 2001. "The Nature and Growth of Vertical Specialization in World Trade." *Journal of International Economics* 54: 75–96.

[26] Johnson, Robert C. and Guillermo Noguera. 2012. "Accounting for Intermediates: Production Sharing and Trade in Value Added." *Journal of International Economics* 86: 224–236.

[27] Johnson, Robert C. and Guillermo Noguera. 2017. "A Portrait of Trade in Value Added over Four Decades." *Review of Economics and Statistics*, forthcoming.

[28] Juhász, Réka and Claudia Steinwender. 2018. "Spinning the Web: The Impact of ICT on Trade in Intermediates and Technology Diffusion." mimeo.

[29] Kee, Hiau Looi and Heiwai Tang. 2016. "Domestic Value Added in Exports: Theory and Firm Evidence from China." *American Economic Review* 106 (6): 1402–1436.

[30] Koopman, Robert, Zhi Wang, and Shang-Jin Wei. 2014. "Tracing Value-Added and Double Counting in Gross Exports." *American Economic Review* 104 (2): 459–494.

[31] Krugman, Paul. 1980. "Scale Economies, Product Differentiation, and the Pattern of Trade." *American Economic Review* 70 (5): 950–959.

[32] Kuznets, Simon. 1957. "Quantitative Aspects of the Economic Growth of Nations II: Industrial Distribution of National Product and Labor Force." *Economic Development and Cultural Change* 5 (4): 1–111.

[33] Laitner, John. 2000. "Structural Change and Economic Growth." *Review of Economic Studies* 67 (3): 545–561.

[34] Lim, Kevin. 2017. "Firm-to-Firm Trade in Sticky Production Networks." mimeo.

[35] Long, John, and Charles Plosser. 1983. “Real business cycles.” *The Journal of Political Economy* 91(1): 39–69.

[36] Matsuyama, Kiminori. 1992. “Agricultural Productivity, Comparative Advantage and Economic Growth.” *Journal of Economic Theory* 58 (2): 317–334.

[37] ———. 2002. “The Rise of Mass Consumption Societies,” *Journal of Political Economy* 110 (5): 1035–1070.

[38] Melitz, Marc J. and Stephen J. Redding. 2014. “Missing Gains from Trade?” *American Economic Review* 104 (5): 317–21.

[39] Ngai, L. Rachel, and Christopher Pissarides. 2007. “Structural Change in a Multisector Model of Growth.” *American Economic Review* 97 (1): 429–443.

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

[41] Oberfield, Ezra and Devesh Raval. 2014. “Micro Data and Macro Technology. ” mimeo.

[42] Romer, Paul. 1990. “Endogenous Technological Change.” *Journal of Political Economy* 98 (5): 71–102.

[43] Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R. and G. J. de Vries. 2015. “An Illustrated User Guide to the World Input–Output Database: the Case of Global Automotive Production.” *Review of International Economics* 23: 575–605.

[44] Tintelnot, Felix, Ken Kikkawa, Magne Mogstad, and Emmanuel Dhyne. 2017. “Trade and Domestic Production Networks.” mimeo.Acemoglu, Daron, and Veronica Guerrieri. 2008. “Capital Deepening and Nonbalanced Economic Growth.” *Journal of Political Economy* 116 (3): 467–498.