Comparative Advantage, Competition, and Firm Heterogeneity

Hanwei Huang and Gianmarco I.P. Ottaviano *

May, 2019

preliminary

Abstract

This paper examines how firm heterogeneity shapes comparative advantage. Drawing on matched customs and firm-level data from China, we find that export participation, exported product scope and product mix, and firm mix within industries vary systematically with firms' labour intensity. This is rationalized by a model in which firms from industries of comparative disadvantage face tougher competition in the export market. The competitive effect induces reallocation within and across firms and generates endogenous Ricardian comparative advantage which dampens ex ante comparative advantage. We develop a new sufficient statistics approach to measure and decompose comparative advantage and find the dampening mechanism is quantitatively important in shaping comparative advantage for a calibrated Chinese economy.

Key Words: Comparative Advantage, Competition, Multi-product Firm, Sufficient Statistics, Firm Heterogeneity

*Hanwei Huang: University College London and City University of Hong Kong. huanghanwei@gmail.com. Gianmarco I.P. Ottaviano: Bocconi University, CEP and CEPR. gianmarco.ottaviano@unibocconi.it. For helpful comments, we would like to thank Eric Bartelsman, Andy Bernard, Lorenzo Caliendo, Dave Donaldson, Elhanan Helpman, James Harrigan, Mengxiao Liu, Hong Ma, Kalina Manova, Thierry Mayer, Peter Neary, Albert Park, Ariell Reshef, Steve Redding, Gerard Roland, Pete Schott, and seminar participants at the LSE, AMES 2017, CESI 2017, XX Applied Economics Meeting, RES annual meeting 2019. All errors are our own responsibilities.

1 Introduction

Comparative advantage which was first articulated by David Ricardo in 1817, has been one of the corner stones of international trade theory in the last 200 years. In the past two decades, firm heterogeneity has taken the centre stage in this research area. Despite the growing interest on its macro implications on productivity and welfare (see, e.g. Melitz, 2003; Arkolakis et al., 2012; Melitz and Redding, 2015; Arkolakis et al., forthcoming), we know relatively little about its impact on comparative advantage. Bernard, Redding, and Schott (2007) famously demonstrate that firm heterogeneity amplifies comparative advantage which increases the welfare gains from trade. In this paper, we show that in an environment with variable mark-ups where the pro-competitive effect is essential, there is another channel through which firm heterogeneity *dampens* comparative advantage. We find this new mechanism to be quantitatively more important than the amplifying mechanism in shaping comparative advantage in a calibrated Chinese economy.

We motivate our theory by four stylized facts about intra- and inter-firm reallocations generated from matched customs and firm-level data from China. First, compared with labour intensive firms, capital intensive Chinese firms are less likely to export. Second, capital intensive exporters export fewer products on average than labour intensive exporters. Third, exports of capital intensive exporters are more skewed toward better performing products than labour intensive exporters. Finally, the skewness of domestic sales across labour intensive firms is higher than across capital intensive firms. The first two facts, which concern the extensive margin of reallocation within and across firms, can be rationalized by extending models such as Arkolakis and Muendler (2010), or Bernard, Redding, and Schott (2011) to multiple industries. However, their assumptions of CES demand and a continuum of firms impose an exogenously fixed mark-up across destinations and industries. The different market conditions therefore have no effect on the export product mix (the relative distribution of exports across products) or the variation of skewness of domestic sales across firms. The third and fourth stylized facts, which concern reallocations along the intensive margin, thus cannot be reconciled with models of this type.

Our theory explains all these facts simultaneously. We extend the analysis of Mayer, Melitz, and Ottaviano (2014) to a continuum of industries by embedding it in Dornbusch, Fischer, and Samuelson (1977). The model features heterogeneous firms and variable mark-ups as in Melitz and Ottaviano (2008). Each firm possesses a "core competency" and has access to a multi-product technology. The marginal cost of producing a product increases as it moves away from the firm's core competency. There are two countries. In industries of comparative advantage, firms are assumed to be more likely to have lower marginal costs than firms from the other country. Exporters in comparative disadvantage industries face tougher competition in the export market, which shifts the whole distribution of mark-ups downwards. The tougher the competition is, the more exporters have to cut the scope of their export product and skew exports toward the better performing products. The relative ease of competition at home in comparative disadvantage industries also induces firms to sell more at home rather than export, thereby reducing their propensity to export. However, competition is tougher in comparative advantage industries in the domestic market, which induces reallocations of domestic sales toward the better performing firms.

Our theory generates new predictions about the effect of firm heterogeneity on comparative advantage. Melitz (2003) predicts that opening up to trade reallocates resources toward more productive firms. In a Heckscher-Ohlin model with heterogeneous firms, Bernard et al. (2007) find that the reallocation effect differs systematically across industries. Due to higher expected profits, an industry with comparative advantage has more entry and stronger selection. This generates endogenous Ricardian comparative advantage, which amplifies the *ex ante* comparative advantage. In our model, there is a new mechanism working on the top of this. In industries of comparative disadvantage, tougher competition in the foreign market will induce more export sales toward the high productivity firms and the better performing products after the country has been opened up to trade. The more competitive the foreign market is, the more exporters have to toughen up. Such endogenous response reduces the relative productivity differences between the two countries and dampens comparative advantage. We also use the model to theoretically decompose Ricardian comparative advantage and find that the productivity measure matters for the decomposition. Industry productivity measures which only capture selections along the extensive margin fail to capture the dampening component. Productivity measures which take into account selections along both the extensive and intensive margins capture both the amplifying and the dampening components.

To test the mechanism of the model, we first extend the empirical analysis of Mayer et al. (2014) to incorporate the competition due to comparative advantage. They examine how French exporters vary their export product mix across markets with different sizes. We construct new variables which measure the competition faced by firms in each market due to comparative advantage. The idea is that capital intensive exporters face tougher competition when exporting to capital abundant markets; labour intensive exporters face tougher competition when exporting to labour abundant markets. Regressions using the matched customs and firm-level data confirm the model's predictions. Exporters export fewer products and skew exports more toward better performing products in markets where they face tougher competition due to comparative advantage, conditioning on the effect of market size.

We then employ a sufficient statistic approach to quantify the different components of comparative advantage. Comparative advantage is not directly observable. We show that, given the trade elasticity, iceberg trade costs, and domestic export participation (export intensities, and export propensities measured by the percent of firms that export), we can infer the home country's comparative advantage against the rest of the world (RoW). The intuition is that, conditional on trade costs and trade elasticity, firms' export participation reveals their relative competitiveness. The higher the fraction of firms that export and the more that exporters export, the stronger the country's comparative advantage. This echoes Balassa's idea of "*Revealed Comparative Advantage*" (RCA).[1] Our sufficient statistics result also allows us to decompose comparative advantage and evaluate the importance of individual components. Using this identification result, we estimate our two-country model for the case of China vs. RoW. We find that the dampening component appears to dominate the amplifying component. Ignoring the dampening component would lead to overestimations of comparative advantage.

Finally, we parametrize our model and conduct simulations on the effect of trade liberalization. We find that bilateral trade liberalization tends to strengthen the endogenous comparative advantage. Taken together, however, whether trade liberalization strengthens the overall comparative advantage or not depends on what kind of productivity measure is used, and which of the endogenous components dominates. It tends to strengthen comparative advantage if the productivity measure captures only the extensive margin. However, if the productivity measure also incorporates the intensive margin and the dampening component is more pronounced, bilateral trade liberalization weakens comparative advantage. As regards welfare, our simulated model with variable mark-ups and firm heterogeneity (Melitz and Ottaviano, 2008) generates higher welfare gains from trade than a simulated model with variable mark-ups but without firm heterogeneity (e.g., Ottaviano, Tabuchi, and Thisse, 2002).

Our paper contributes to the following strands in the literature. Our work is closely related to that of recent authors who study the macro implication of firm heterogeneity. We show that there is a new channel through which firm heterogeneity shapes comparative advantage,

[1] While Balassa (1965) measured the underlying pattern of comparative advantage by relative exports across industries, we use data on firms' export participations together with estimated trade costs and trade elasticity. As noted by Costinot et al. (2012) and French (2017), the RCA index would not necessarily coincide with the underlying ranking of relative productivities. In contrast, our measure is theoretically consistent.

namely that tougher competition in the export market induces reallocations such that *ex ante* comparative advantage is dampened. This contrasts with the amplifying mechanism found in Bernard et al. (2007).[2] Arkolakis et al. (2012) find that for a group of models which satisfy certain restrictions, the formula for the welfare gains from trade is the same.[3] Melitz and Redding (2015) show that the Melitz model with firm heterogeneity implies higher welfare gains from trade than the Krugman model with homogeneous firms. Compared with their results, our model features variable mark-ups. However, we also find trade yields higher welfare gains in the *simulated* heterogeneous firm model than it does in the homogeneous model.

We also contribute to the literature on the measurement of comparative advantage. Comparative advantage is the basis of classic trade theory. However, it has remained challenging to measure. Balassa's RCA index has in the last few decades been the key tool in measuring comparative advantage. There has been a renaissance in quantifying Ricardian comparative advantage since the seminal contribution by Eaton and Kortum (2002), which provides a tractable multi-country Ricardian model.[4] We provide sufficient statistic results, which identify comparative advantage directly and decompose it into exogenous and endogenous components. The sufficient statistic approach, as argued in Arkolakis et al. (2012), saves us from solving all the endogenous variables but still provides estimates for the object of interest. As far as we know, this paper is the first to provide sufficient statistics for comparative advantage.[5] We also show that, in measuring comparative advantage, the exact productivity measures matter. Measures that capture only the extensive margin miss an important determinant of comparative advantage and bias our estimations.

Finally, the literature both theoretical and empirical on multi-product firm has been booming.[6] Our analysis highlights how comparative advantage affects resource reallocation along the

[2]Recent contributions include Lu (2010), Huang et al. (2017), and Burstein and Vogel (2017). Gaubert and Itskhoki (2018) also study a multi-sector Ricardian model with heterogeneous firms but their focus is on the effect of the granularity force on comparative advantage. Ma et al. (2014) build on Bernard et al. (2011) and study within-firm specialization across products with different factor proportions.

[3]The restrictions include CES preferences and a constant trade elasticity. Arkolakis et al. (forthcoming) depart from these two restrictions and study welfare gains from trade in models with variable mark-ups.

[4]Costinot *et al.* (2012) estimate the importance of Ricardian comparative advantage on trade patterns and welfare using an extended Eaton-Kortum model. Relatedly, Levchenko and Zhang (2016) use the gravity equation to infer comparative advantage from trade flows and its evolution over time. Costinot et al. (2016) focus on the agriculture sector for which the parcel-level productivity of lands can be precisely estimated for different crops. Gaubert and Itskhoki (2018), Huang et al. (2017) instead use the two-country DFS framework to work out comparative advantage by structural estimation.

[5]The sufficient statistic approach has gained popularity in the field of public finance (Chetty, 2009). Arkolakis *et al.* (2012) shows that within a set of trade models which satisfy certain conditions, trade elasticity and the share of expenditure on domestic goods are sufficient statistics for welfare gains from trade.

[6]Feenstra and Ma (2009), and Eckel and Neary (2010) examine the effect of competition on the distribution of sales and the cannibalization effect for multi-product firms. Arkolakis and Muendler (2010), and Bernard et

intra-firm extensive and intensive margins for multi-product firms, and how it feeds back to comparative advantage. The mechanism is similar to that in Mayer et al. (2014). Their focus is on the competition due to market size while the present paper concentrates on comparative advantage. Our model therefore provides a finer characterization of multi-product exports in a world with many industries.

The remainder of the paper is arranged as follows. Section 2 presents four stylized facts which motivate our theory. Section 3 presents the model and provides predictions on comparative advantage. Section 4 contains two sets of empirical analyses. Section 5 conducts numerical simulations on the effect of trade liberalization. Section 6 concludes.

2 Motivating Evidence

2.1 Data

In this section, we present a few stylized facts on the way in which export participation, exporters' product scope and product mix, and firm mix vary with capital intensity. These facts are generated using matched customs and firm-level data from China for the period 2000-2006. The first dataset that we use is the Chinese Annual Industrial Survey (CAIS) which covers all State Owned Firms (SOE) and non-SOEs with sales above 5 million Chinese Yuan. These data provide rich information on firms' financial statements, and forms of identification such as name, address, ownership, and number of employees. The other dataset that we employ is Chinese Customs data, which cover all China's import and export transactions. For each transaction, we know the Chinese importer/exporter, the product (at HS-8 level), value, origin, destination, etc. There is no common firm identifier between the two datasets. We match the two datasets on the basis of firm's name, address, telephone number, and zip code.[7] The sample of matched exports represents about 37% of all Chinese exports reported in the customs data for 2000 and 52% for 2006.

We focus on the Chinese manufacturers and exclude firms from the mining and utility sectors in CAIS, and wholesalers or intermediaries in the customs data. We use capital intensity to capture comparative advantage: given the abundance of labour endowment in China, we expect

al. (2011) emphasize selection along the extensive margin, while Mayer et al. (2014) focus on selection along the intensive margin. Manova and Yu (2017) instead appraise quality differentiation and study product selection along the quality margin. Bernard et al. (2010), Iacovone and Javorcik (2010), and Mayer et al. (2016) investigate product churning over time in response to changes in market conditions.

[7]Such matching method has been used in a few number of papers, including Ma et al. (2014), Yu (2015), and Manova and Yu (2016).

the country to have comparative advantage relative to the RoW in labour intensive industries and comparative disadvantage in capital intensive industries. We follow Schott (2004) and Huang et al. (2017) to define industries as "Heckscher-Ohlin aggregates" and group Chinese firms into 100 bins according to their capital intensity. Schott (2004) argues that traditional industry classification, which defines industries according to the final use of goods, aggregates goods that are produced using different factor proportions. Similarly, Huang et al. (2017) show that such industry classification also aggregates firms which use different technologies. Capital intensity is defined as $1 - \frac{Labour\,Costs}{Value\,Added}$ for each firm. For example, firms with capital intensity between 0 and 0.1 are defined as industry 1.[8] Under this classification, which we use for the rest of the paper, the following stylized facts are found using data for the year 2006.

2.2 Stylized Facts

Stylized fact 1: *Export propensity and export intensity decline with capital intensity.*

This is captured in Figure 1. The left panel plots the export propensity of each industry, where export propensity is defined as the total number of exporters divided by the total number of firms. The right panel plots the export intensity, where export intensity is defined as total exports divided by the total sales for each industry. As the figures indicate, both measures decline with capital intensity. This is consistent with our expectation that China has comparative advantage in labour intensive industries and labour intensive firms are more likely to export.

Stylized fact 2: *Exporters' export product scope declines with capital intensity.*

A firm's export product scope is defined as the number of products it exports. We measure each exporter's export product scope by counting the number of distinctive HS-8 products exported to all destinations in the customs data. The left panel of Figure 2 plots the export product scope averaged across exporters for each industry. As we can see, it falls with capital intensity. The right panel of the figure plots the share of single-product exporters, which are firms exporting one HS-8 product only. It is obvious that single-product exporters are more prevalent in the capital intensive industries in China.

Stylized fact 3: *The export product mix is more skewed in capital intensive industries than in*

[8] We follow the traditional two-factor Heckscher-Ohlin paradigm to consider labour vs. capital. Here "capital" includes all non-labour factors, such as energies. Labour costs include payable wages, labour and employment insurance fees, and the total of employee benefits payable. We exclude firms with capital intensities which are negative or greater than 1. Their presence is very likely to be due to misreporting or errors. We also exclude firms with negative value added, employment or assets. Firms with fewer than 8 employees are also excluded since they are under different legal regime. The results using data for other years are qualitatively the same.

labor intensive industries.

This is captured by Figure 3. The left panel plots the average of the log-ratios between the exports of the core product to the second best product. The core product is defined as the product that makes up the greatest part of the total exports for each firm. As we can see, this measure tends to be higher in capital intensive industries. Exports are therefore more concentrated on the better performing products in capital intensive industries. However, this measure captures only the skewness of exports across a few products. To show the presence of such a relationship across all exported products, we use a measure which captures the skewness of the whole distribution of exports. The right panel plots the average Theil index of firm exports across products. Again, the skewness of exports across products tends to increase with capital intensity.

Stylized fact 4: *The skewness of domestic sales decreases with capital intensity.*

In the left panel of Figure 4, we plot the log-ratios of domestic sales between the 75th-percentile firm and the 25th-percentile firm. We measure a firm's domestic sales by deducting exports from its total sales. As is obvious from the figure, the skewness tends to be higher in labour intensive industries. We also use the Theil index to capture the skewness of the whole distribution of domestic sales across firms. This is shown in the right panel. Still, the skewness tends to decline with capital intensity.

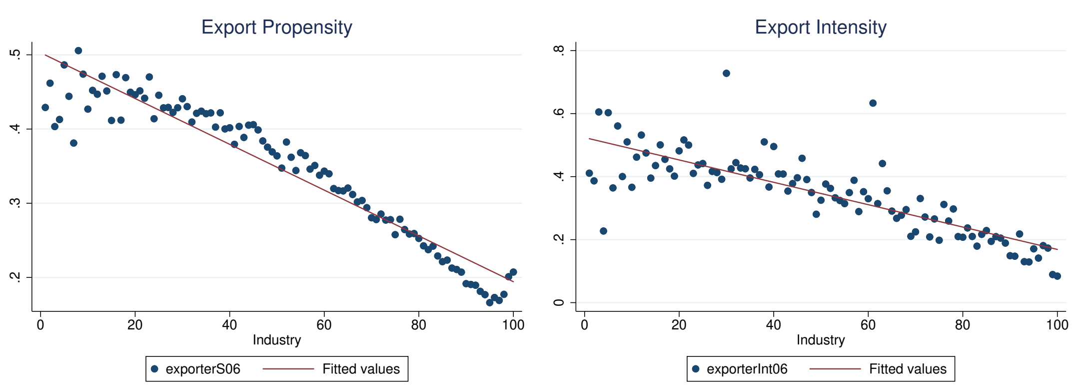


Figure 1: Export propensity and export intensity

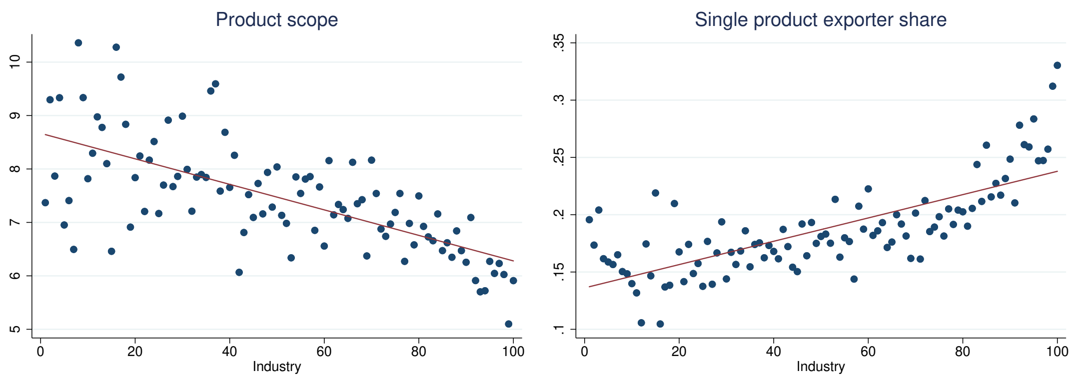


Figure 2: Number of products exported

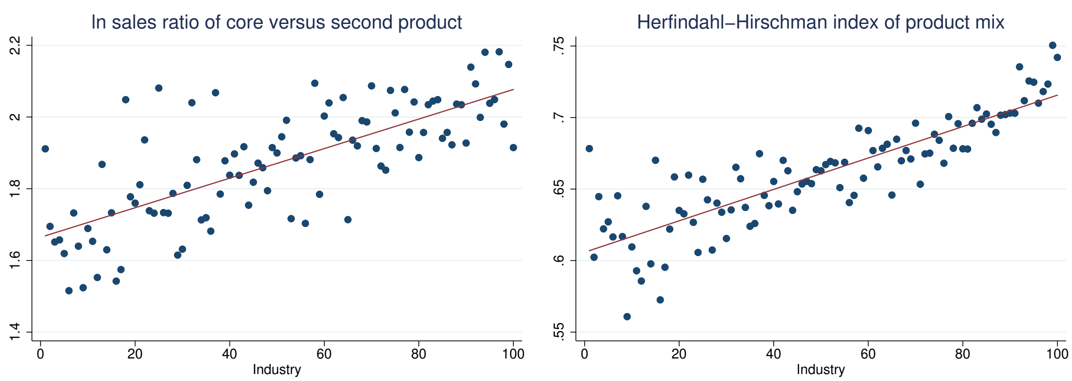


Figure 3: Skewness of export product mix

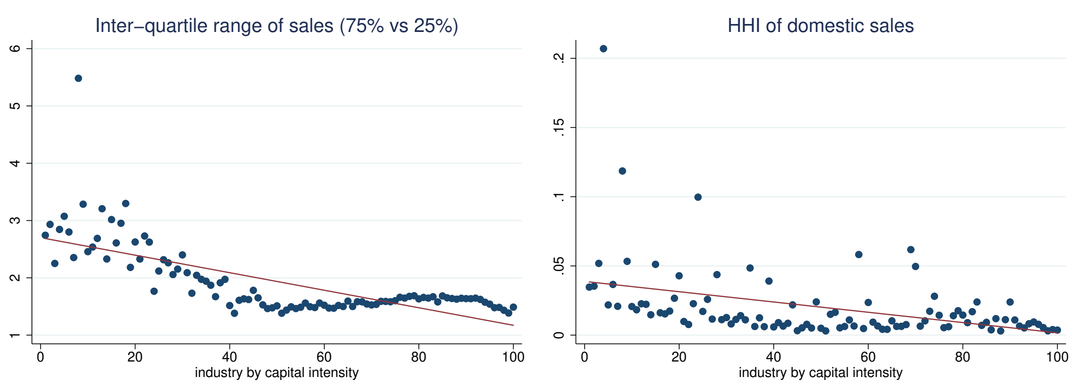


Figure 4: Skewness of domestic sales across firms

2.3 Discussion

So far, our results are only graphical evidence.[9] In Appendix 7, we provide further regression evidence on the robustness of the stylized facts. In Appendix Table C1, we confirm fact 1, that export propensity and export intensity decline strongly with capital intensity, using data from 2000-2006. To deal with concerns that many Chinese exporters were processing traders, and China went through a period of state-owned-enterprise (SOE) reform which might have affected firms' exports, we examine whether fact 1 is true or not for non-processing traders and none-SOEs by excluding them from our sample. Still, fact 1 remains highly robust. Similarly, we examine the robustness of fact 2 on export product scope in Appendix Table C2, and find that it holds for the full sample and sub-sample of exporters. In Appendix Table C3-C5, we examine the robustness of fact 3 on the product mix of exporters and use alternative measures of skewness such as the Herfindahl index. Again, fact 3 is robust to alternative measures and data samples. Similarly, we examine the robustness of fact 4 on the skewness of domestic sales in Appendix Table C6-C8. It is still the case the skewness of domestic sales is higher in labour intensive industries.

Overall, these stylized facts reveal how comparative advantage shapes firm sales within and across firms at home and abroad. The first two stylized facts focus on the extensive margin, the third and fourth on the intensive margin. The first two facts can be easily explained by existing models, such as that of Bernard, Redding, and Schott (2007, 2011), by introducing multiple industries and multi-product firms.[10] However, the third and fourth stylized facts are not consistent with these models which impose CES demand and a continuum of firms assumptions. These two assumptions imply a fixed mark-up across markets and industries. There is therefore no variation in the intra-firm product mix or relative sale across firms in different markets or industries.[11] Mayer, Melitz, and Ottaviano (2014) provide a multi-product model built on Melitz and Ottaviano (2008), which features variable mark-ups. Their model explains how French exporters vary their sales across markets of different sizes. They find firms which export to larger markets skew their exports toward their better selling products. Such a mechanism should work at the industry level as well, and in principle can explain the third and fourth stylized facts. This motivates our theory in the following section.

[9] In Appendix 9, we provide figures for the years 2000 and 2006, and other measures to capture the skewness of distributions.

[10] For example, Huang *et al.* (2007) provide a multi-sector extension of Bernard et al. (2007). Bernard et al. (2011) discuss an extension of their benchmark multi-product model to multiple industries in the appendix.

[11] The relative sale of different products only depends on the relative firm or product productivity in these type of models.

3 Theory

We build a model which simultaneously explains the stylized facts discovered in the previous section. Our model extends Mayer, Melitz, and Ottaviano (2014) to a continuum of industries by embedding it in Dornbusch, Fischer, and Samuelson (1977). The model makes novel predictions on the effect of firm heterogeneity on comparative advantage.

3.1 Closed Economy

We first consider the closed economy. Suppose there are two countries, Home and Foreign. The consumers in each country have identical preference given by

$$U = q_0^c + \int_0^1 [\alpha \int_{i\in\Omega(z)} q_i^c(z)di - \frac{\gamma}{2}\int_{i\in\Omega(z)} (q_i^c(z))^2 di - \frac{\eta}{2}(\int_{i\in\Omega(z)} q_i^c(z)di)^2]dz,$$

where q_0^c denotes the consumption of the numeraire good and $q_i^c(z)$ the consumption of the differentiated variety i in industry z. z indexes the continuum of industries and has a support of $[0, 1]$. $\Omega(z)$ is the set of differentiated varieties in industry z. The parameters capturing the substitution pattern between the differentiated varieties and numeraire good are α and η. As is obvious in the demand function below, a higher α or smaller η will lead to a higher demand for the differentiated varieties relative to the numeraire good. The parameter capturing the substitution pattern of the differentiated varieties within each industry is given by γ. The degree of differentiation increases with γ. In the extreme case that $\gamma = 0$, the differentiated varieties become perfect substitutes.

We normalized the price of the numeraire good to be 1. The budget constraint faced by consumers is given by

$$q_0^c + \int_0^1 \int_{i\in\Omega(z)} p_i^c(z)q_i^c(z)didz = y_0^c + I,$$

where y_0^c is the endowment of the numeraire good and I the labour income. Assuming that consumers have positive demand for the numeraire good, solving the consumers' problem delivers the following demand for the differentiated varieties

$$p_i(z) = \alpha - \gamma q_i^c(z) - \eta Q^c(z).$$

Then the corresponding market demand is

$$q_i(z) = Lq_i^c(z) = \frac{L}{\gamma}(p_{\max}^z - p_i(z)),$$

where L is the number of consumers in the home country and $p_{\max}^z$ is the choke price of industry z. Then a firm with marginal cost c operating in industry z faces the following problem

$$\max_{p(z)}(p(z,c) - c)q(z).$$

Solving the firm's problem, we have

$$\begin{aligned}
p(z,c) &= \frac{1}{2}(p_{\max}^z + c), \\
\mu(z,c) &= \frac{1}{2}(p_{\max}^z - c), \\
q(z,c) &= \frac{L}{2\gamma}(p_{\max}^z - c), \\
\pi(z,c) &= \frac{L}{4\gamma}(p_{\max}^z - c)^2,
\end{aligned}$$

where $p(z,c)$, $\mu(z,c)$, $q(z,c)$, and $\pi(z,c)$ are the price, mark-up, output, and profit, respectively.

Each industry has a pool of potential entrants. Firms pay a fixed cost of f_E, and draw their marginal costs from a common distribution $G(z,c)$ defined on the support of $[0, C_M(z)]$ for industry z. Firms with marginal costs higher than the threshold $C_D(z) = p_{\max}^z$ will exit from the market. Free entry implies that

$$\int_0^{C_D(z)} \pi(z,c)dG(z,c) = f_E.$$

Under the Pareto distribution assumption that

$$G(z,c) = (\frac{c}{C_M(z)})^k,\ c \in [0,\ C_M(z)],$$

the cut-off marginal cost under autarky is given by

$$C_D(z)^A = [\frac{2(k+1)(k+2)\gamma C_M(z)^k f_E}{L}]^{1/(k+2)}. \tag{3.1}$$

Similarly, for the foreign country, we have[12]

$$C_D(z)^{*A} = [\frac{2(k+1)(k+2)\gamma C_M(z)^{*k} f_E}{L^*}]^{1/(k+2)}. \tag{3.2}$$

3.2 Open Economy with Single-product Firms

We now consider the open economy case without multi-product firms. The key purpose is to study how competition varies across industries when countries start trading with each other. To export to the foreign country, we assume that domestic firms need to pay an iceberg trade cost of τ. Foreign firms face the iceberg trade cost of τ^*.

Free entry implies that the sum of expected profits from both markets equals the fixed entry cost. The free entry condition therefore becomes

$$\int_0^{C_D(z)} \pi_d(z,c)dG(z,c) + \int_0^{C_X(z)} \pi_x(z,c)dG(z,c) = f_E,$$

where $C_X(z) = C_D^*/\tau$ is the marginal cost cut-off for exporters. Thanks to the Pareto distribution assumption, this can be simplified as

$$LC_D(z)^{k+2} + \rho L^* C_D^*(z)^{k+2} = \beta C_M(z)^k, \tag{3.3}$$

where $\rho = \tau^{-k} \in [0,1]$ is the freeness of trade and $\beta = 2\gamma(k+1)(k+2)f_E$ is a constant.

Similarly, for the foreign country, we have

$$L^* C_D^*(z)^{k+2} + \rho^* LC_D(z)^{k+2} = \beta C_M^*(z)^k, \tag{3.4}$$

where $\rho^* = \tau^{*-k}$. Combining the two equations above, we have[13]

$$C_D(z)^{k+2} = \frac{\beta[C_M(z)^k - \rho C_M^*(z)^k]}{L(1-\rho\rho^*)}, \tag{3.5}$$

$$C_D^*(z)^{k+2} = \frac{\beta[C_M^*(z)^k - \rho^* C_M(z)^k]}{L^*(1-\rho\rho^*)}. \tag{3.6}$$

Following Dornbusch et al. (1977), we rank the industries such that $\frac{\partial C_M(z)}{\partial z} > 0$ and $\frac{\partial C_M^*(z)}{\partial z} < 0$. That is, domestic firms in industries with higher z draw their marginal costs from a wider

[12] Variables with asterisk are for the foreign country.

[13] To ensure that the equations have real solutions, we assume that $\rho \le \frac{C_M(z)^k}{C_M(z)^{*k}} \le \frac{1}{\rho^*}$ for all z.

support, while the converse is true for the foreign firms. Under such assumptions, the home country will have comparative advantage in industries with lower z. There are different ways that these assumptions can be micro-founded. For example, they can be generated by the Heckscher-Ohlin force. Following Corcos et al. (2011), suppose that firms use a Cobb-Douglas production technology with z indexing the capital intensity, and $C_M(z) = w^{1-z}r^z$ and $C^*_M(z) = w^{*1-z}r^{*z}$. Then $\frac{\partial C_M(z)}{\partial z} = C_M(z)\ln\frac{r}{w}$ and $\frac{\partial C^*_M(z)}{\partial z} = C^*_M(z)\ln\frac{r^*}{w^*}$. If the home country is labour abundant relative to the foreign such that $\frac{r^*}{w^*} < 1 < \frac{r}{w}$, then we have $\frac{\partial C_M(z)}{\partial z} > 0$ and $\frac{\partial C^*_M(z)}{\partial z} < 0$. Under this interpretation, the home country has comparative advantage in labour intensive industries while the foreign country has comparative advantages in capital intensive industries.

Given the assumption that $\frac{\partial C_M(z)}{\partial z} > 0$ and $\frac{\partial C^*_M(z)}{\partial z} < 0$, it is easy to verify that

$$\frac{\partial C_D(z)}{\partial z} > 0, \text{ and } \frac{\partial C^*_D(z)}{\partial z} < 0.$$

So the cut-offs are lower in industries of comparative advantage. Exporters therefore face tougher competition to sell in the foreign market in industries where the foreign country has comparative advantage. Immediately, we have the following proposition.

Proposition 1. *Export propensity* $\chi(z) \equiv (\frac{C_X(z)}{C_D(z)})^k$ *and export intensity* $\lambda(z) \equiv \frac{Exports(z)}{Total\ Sales(z)}$ *increase with comparative advantage.*

Proof. See Appendix 10.1. □

This proposition implies that firms are more likely to export in industries of comparative advantage. This is consistent with Stylized fact 1 if we believe that China has comparative advantage in labour intensive industries.

Following Melitz and Ottaviano (2008), the number of entrants in each industry is given by

$$\begin{aligned} N_E(z) &= \frac{2C_M(z)^k(k+1)\gamma}{\eta(1-\rho\rho^*)}\left(\frac{\alpha - C_D(z)}{C_D(z)^{k+1}} - \rho^*\frac{\alpha - C^*_D(z)}{C^*_D(z)^{k+1}}\right), \\ N^*_E(z) &= \frac{2C^*_M(z)^k(k+1)\gamma}{\eta(1-\rho\rho^*)}\left(\frac{\alpha - C^*_D(z)}{C^*_D(z)^{k+1}} - \rho\frac{\alpha - C_D(z)}{C_D(z)^{k+1}}\right). \end{aligned}$$

If $\frac{\alpha - C^*_D(z)}{C^*_D(z)^{k+1}} \leq \rho\frac{\alpha_z - C_D(z)}{C_D(z)^{k+1}}$, we have $N^*_E(z) \leq 0$ so that there is no foreign firm in such industries. In this case, the home country specializes in these industries. This is more likely to happen if the freeness of trade ρ is sufficiently high, or $C^*_D(z)$ is greater than $C_D(z)$. Intuitively, in such cases, foreign firms face tough competition and get eliminated from the market. Similarly, the foreign country will specialize in industries where $\frac{\alpha - C_D(z)}{C_D(z)^{k+1}} \leq \rho^* \frac{\alpha - C_D^*(z)}{C_D^*(z)^{k+1}}$ is satisfied.[14]

3.3 Open Economy with Multi-product Firms

Now we extend the model to allow firms producing multiple products by following Mayer et al. (2014). Each firm's marginal cost of producing the core competency is given by c. Varieties are ranked in increasing order of distance from the core competency and indexed by m. The marginal cost of producing variety m is given by $v(m,c) = \varpi^{-m}c$, and $\varpi \in (0,1)$. So the marginal cost increases as we move away from the core competency.[15] Firms will keep adding products until the marginal cost is higher than the choke price. Therefore, the number of varieties produced by each firm is given by

$$M_d(z,c) = \begin{cases} 0, \; if \; c > C_D(z), \\ \max\{m | v(m,c) \leq C_D(z)\} + 1, \; if \; c \leq C_D(z). \end{cases}$$

The number of varieties exported to the foreign country by domestic firms is given by

$$M_x(z,c) = \begin{cases} 0, \; if \; c > C_X(z), \\ \max\{m | v(m,c) \leq C_X(z) = \frac{C_D^*(z)}{\tau}\} + 1, \; if \; c \leq C_X(z). \end{cases}$$

The free entry condition now becomes

$$\int_0^{C_D(z)} \Pi_d(z, v(m,c)) dG(z,c) + \int_0^{C_X(z)} \Pi_x(z, v(m,c)) dG(z,c) = f_E, \tag{3.7}$$

where firm profits from Home $\Pi_d(z,c)$, and Foreign $\Pi_x(z, v(m,c))$, are the sum of the profits made from each product sold in the respective market:

$$\begin{aligned} \Pi_d(z,c) &= \sum\nolimits_{m=0}^{M_d(z,c)-1} \pi_d(z, v(m,c)), \\ \Pi_x(z, v(m,c)) &= \sum\nolimits_{m=0}^{M_x(z,c)-1} \pi_x(z, v(m,c)). \end{aligned}$$

According to Mayer et al. (2014), the free entry condition Equation (3.7) can be simplified as

$$LC_D(z)^{k+2} + \rho L^* C_D^*(z)^{k+2} = \frac{\beta C_M(z)^k}{\Psi}, \tag{3.8}$$

[14] Given that China imports and exports in every industry, we assume for the rest of the paper that the no-specialization conditions are always satisfied .

[15] Eckel and Neary (2010) provide an alternative way to model the asymmetries between products on the cost side. Eckel et al. (2015) further allow firms to invest in quality.

where $\Psi = (1 - \varpi^k)^{-1}$ is an index of multi-product flexibility. Similarly, for the foreign country, we have

$$L^* C_D^*(z)^{k+2} + \rho^* L C_D(z)^{k+2} = \frac{\beta C_M^*(z)^k}{\Psi}.$$

We can solve the two equations above for the choke prices:

$$C_D(z)^{k+2} = \frac{\beta[C_M(z)^k - \rho C_M^*(z)^k]}{\Psi L(1 - \rho\rho^*)}, \tag{3.9}$$

$$C_D^*(z)^{k+2} = \frac{\beta[C_M^*(z)^k - \rho^* C_M(z)^k]}{\Psi L^*(1 - \rho\rho^*)}. \tag{3.10}$$

It is easy to see that we still have $\frac{\partial C_D(z)}{\partial z} > 0$ and $\frac{\partial C_D^*(z)}{\partial z} < 0$ under the assumptions that $\frac{\partial C_M(z)}{\partial z} > 0$ and $\frac{\partial C_M^*(z)}{\partial z} < 0$. Therefore, Propositions 1 still holds in an environment with multi-product firms. The following two propositions focus on the variations in the product scope and product mix across industries, which the single-product model cannot explain.

Proposition 2. *The export product scope increases weakly with comparative advantage.*

Proof. See Appendix 10.2. □

Proposition 2 implies that the export product scope tends to be lower in the industries of comparative disadvantage. For firms with the same marginal cost, those exporting in the industries of comparative disadvantage are more likely to be single-product exporters. This is consistent with Stylized Fact 2.

Proposition 3. *Exports are skewed toward better products in the industries of comparative disadvantage.*

Proof. See Appendix 10.3. □

In industries of comparative disadvantage, the export market is more competitive. The tougher competition induces exporters to reallocate more sales to the better selling products. If we agree that capital intensive industries are the industries of comparative disadvantage for China, we should expect capital intensive exporters to have a more skewed export product mix. This is consistent with Stylized Fact 3.

Proposition 4. *Domestic sales tend to skew toward more productive firms in comparative advantage industries.*

Proof. See Appendix 10.4. $\square$

In comparative advantage industries, the domestic market is more competitive. Such tougher competition would induce reallocations of sales more toward products that are produced with lower marginal costs. Since such products are more likely to be produced by firms with higher core efficiencies, outputs are reallocated toward these firms. In the end, domestic sales are skewed toward these firms and Stylized fact 4 is also rationalized.

3.4 Comparative Advantage

Our model has new implications for comparative advantage. Bernard et al. (2007) show that the different degree of selection across industries generates endogenous Ricardian comparative advantage which amplifies ex ante comparative advantage. In this subsection, we show that variable mark-ups allow for selections along the intensive margin, which generate endogenous Ricardian comparative advantage that dampens ex ante comparative advantage. We also find that the measure of sectoral productivity matters for the estimate of comparative advantage. Comparative advantage is usually measured by relative productivity (Costinot et al., 2012). If the productivity measure captures only selections along the extensive margin, we miss the dampening effect of intensive margin selections and overestimate comparative advantage.

3.4.1 Relative average marginal cost

Comparative advantage is defined as the relative productivity between home and foreign for each industry. We can measure productivity as the inverse of the simple average marginal cost across firms within each industry. The average marginal cost of industry z in the home country is given by

$$\overline{c}(z) = \int_0^{C_D(z)} c dG(z, c) = \frac{k}{k+1} C_D(z).$$

For the foreign country, it is $\overline{c}(z)^* = \frac{k}{k+1} C_D(z)^*$. Then using Equations (3.1) and (3.2), the relative average marginal cost under autarky is given by:

$$\begin{aligned} \frac{\overline{c}(z)}{\overline{c}(z)^*} &= \frac{C_D(z)^A}{C_D(z)^{*A}} \\ &= \left(\frac{L^*}{L} \frac{C_M(z)^k}{C_M^*(z)^k}\right)^{1/(k+2)}. \end{aligned}$$

If we denote the cost cut-offs in the open economy as $C_D(z)^T$ and $C_D(z)^{*T}$, according to

Equations (3.9) and (3.10) the relative marginal cost between home and foreign is[16]

$$\begin{aligned}\frac{\overline{c}(z)}{\overline{c}(z)^*} &= \frac{C_D(z)^T}{C_D(z)^{*T}} \\ &= (\frac{L^*}{L}\frac{C_M(z)^k - \rho C_M^*(z)^k}{C_M^*(z)^k - \rho^* C_M(z)^k})^{1/(k+2)}.\end{aligned}$$

Proposition 5. *Comparative advantage as measured by the relative simple average of margin costs between home and foreign* $\frac{\overline{c}(z)}{\overline{c}(z)^*}$ *is amplified after opening up to trade as*

$$\frac{C_D(z)^T}{C_D(z)^{*T}} = \underbrace{\frac{C_D(z)^A}{C_D(z)^{*A}}}_{ex\ ante}\underbrace{\left[\frac{1-\rho\frac{C_M^*(z)^k}{C_M(z)^k}}{1-\rho^*\frac{C_M(z)^k}{C_M^*(z)^k}}\right]^{\frac{1}{k+2}}}_{amplifying}. \tag{3.11}$$

Proof. See Appendix 10.5. □

As noted above, the amplifying mechanism in Bernard et al. (2007) is also present in our model. This proposition shows not only that it exists but also how to tease it out by decomposition. However, the productivity measure varies only with the cost cut-off, or the productivity of the marginal survival firm, which misses the details of allocations across the inframarginal firms which form the majority in each industry. We next show that a different result arises if the inframarginal firms also matter for the productivity measure.

3.4.2 Relative TFPQ

Now we consider a quantity-based TFP (TFPQ) from Mayer et al. (2014). It measures industry output per worker and captures both the intensive and the extensive margins by incorporating the fact that firms have different amounts of inputs and outputs, and only a subset of firms export. In the closed economy, the TFPQ of industry z is

$$\overline{\Phi}(z)^A = \frac{\int_0^{C_D(z)^A} Q(z,c)dG(z,c)}{\int_0^{C_D(z)^A} C(z,c)dG(z,c)} = \frac{k+2}{k}\frac{1}{C_D(z)},$$

[16]The single-product economy gives the same result. Equations (3.5) and (3.6) differ only by a constant as compared with Equations (3.9) and (3.10).

where $Q(z,c)=\sum_{m=0}^{M_d(z,c)-1}q(z,v(m,c))$ and $C(z,c)=\sum_{m=0}^{M_d(z,c)-1}v(m,c)q(z,v(m,c))$ are firm outputs and inputs, respectively. The relative TFPQ under autarky is given by

$$\frac{\overline{\Phi}(z)^A}{\overline{\Phi}(z)^{*A}}=\frac{C_D(z)^{A*}}{C_D(z)^A}=(\frac{L}{L^*}\frac{C_M^*(z)^k}{C_M(z)^k})^{1/(k+2)}.$$

which is the *ex ante* comparative advantage before countries open to trade. It coincides with Equation (3.11) which measures the relative average marginal cost under autarky. In the open economy, we need to account for exports. The TFPQ is then given by

$$\overline{\Phi}(z)^T=\frac{\int_0^{C_D(z)}Q_d(z,c)dG(z,c)+\int_0^{C_X(z)}Q_x(z,c)dG(z,c)}{\int_0^{C_D(z)}C_d(z,c)dG(z,c)+\int_0^{C_X(z)}C_x(z,c)dG(z,c)}.$$

It can be shown that the total industry outputs and inputs for each market are

$$\begin{aligned}
\int_0^{C_D(z)}Q_d(z,c)dG(z,c) &= \frac{LC_D(z)^{k+1}}{2\gamma C_M^k(k+1)}\frac{1}{1-\varpi^k},\\
\int_0^{C_X(z)}Q_x(z,c)dG(z,c) &= \frac{\rho L^*C_D^*(z)^{k+1}}{2\gamma C_M^k(k+1)}\frac{1}{1-\varpi^k},\\
\int_0^{C_D(z)}C_d(z,c)dG(z,c) &= \frac{kLC_D(z)^{k+2}}{2\gamma C_M^k(k+1)(k+2)}\frac{1}{1-\varpi^k},\\
\int_0^{C_X(z)}C_x(z,c)dG(z,c) &= \frac{\rho kL^*C_D^*(z)^{k+2}}{2\gamma C_M^k(k+1)(k+2)}\frac{1}{1-\varpi^k}.
\end{aligned}$$

substituting these results into the definition of the TFPQ, we have

$$\overline{\Phi}(z)^T=\frac{k+2}{k}\Big[\frac{LC_D(z)^{k+2}}{LC_D(z)^{k+2}+\rho L^*C_D^*(z)^{k+2}}\frac{1}{C_D(z)}+\frac{\rho L^*C_D^*(z)^{k+2}}{LC_D(z)^{k+2}+\rho L^*C_D^*(z)^{k+2}}\frac{1}{C_D^*(z)}\Big],$$

which is a weighted average of the competitiveness of each market. The weights are given by the total costs for the goods sold in each market. If $\rho=0$, we go back to the closed economy case. Using the free entry condition (3.8), it can be further simplified as

$$\overline{\Phi}(z)^T=\frac{(k+2)\Psi}{k\beta C_M(z)^k}(LC_D(z)^{k+1}+\rho L^*C_D^*(z)^{k+1}).$$

There is a similar equation for the foreign country. The relative TFPQ between home and foreign for each industry is therefore given by

$$\frac{\overline{\Phi}(z)^T}{\overline{\Phi}^*(z)^T} = \frac{C_M^*(z)^k}{C_M(z)^k}\frac{LC_D(z)^{k+1}+\rho L^*C_D^*(z)^{k+1}}{L^*C_D^*(z)^{k+1}+\rho^*LC_D(z)^{k+1}}.$$

Proposition 6. *Comparative advantage as measured by the relative quantity-based TFP between home and foreign* $\frac{\overline{\Phi}(z)}{\overline{\Phi}^*(z)}$ *after opening up to trade can be decomposed into three components: an ex ante component, an amplifying component, and a dampening component as:*

$$\frac{\overline{\Phi}(z)^T}{\overline{\Phi}^*(z)^T} = \underbrace{\frac{\overline{\Phi}(z)^A}{\overline{\Phi}(z)^{*A}}}_{ex\ ante} \underbrace{(\frac{\overline{\Phi}(z)^A}{\overline{\Phi}(z)^{*A}})^{k+1}}_{amplifying} \underbrace{\frac{L^*}{L} \frac{\frac{L}{L^*}(\frac{C_D(z)}{C_D^*(z)})^{k+1} + \rho}{1 + \rho^* \frac{L}{L^*}(\frac{C_D(z)}{C_D^*(z)})^{k+1}}}_{dampening}. \tag{3.12}$$

Proof. See Appendix 10.6. □

As pointed out by Bernard et al. (2007), given the higher expected profits of exporting, there will be more entrants and more intense selection in the comparative advantage industries. This tends to enlarge the relative productivity differences across industries and amplify comparative advantage. Such a channel is preserved in this measure and given by the term $(\frac{\overline{\Phi}(z)^A}{\overline{\Phi}(z)^{*A}})^{k+1}$ which is positively correlated with the ex ante component $\frac{\overline{\Phi}(z)^A}{\overline{\Phi}(z)^{*A}}$.[17]

However, their assumptions of CES demand and a continuum of firms impose a constant exogenous mark-up. This implies that the relative revenue in each market between firms depends only on relative productivity and has nothing to do with market conditions. So selections along the intensive margin are constant across markets and industries. Our model with variable mark-ups has different implications. Tougher competition would induce reallocations of resources toward more productive firms and better performing products, as evident from Propositions 3 and 4. In other words, in tougher markets or industries, firms toughen up. This channel tends to dampen their comparative disadvantage.

If we follow Mayer et al. (2014) to define revenue-based TFP (TFPR) as

$$\overline{P}(z) = \frac{\int_0^{C_D(z)} R_d(z,c)dG(z,c) + \int_0^{C_X(z)} R_x(z,c)dG(z,c)}{\int_0^{C_D(z)} Q_d(z,c)dG(z,c) + \int_0^{C_X(z)} Q_x(z,c)dG(z,c)},$$

where $R_d(z,c)$ and $R_x(z,c)$ are firms' domestic and foreign revenues, respectively, we get the

[17] In Appendix 11.1 we extend the model by Bernard et al. (2007) to multiple industries. To simplify the analysis, we use a quasi-CES preference and decompose comparative advantage in similar manners.

same result since

$$\begin{aligned}\overline{\Phi}_R(z) &= \frac{(\int_0^{C_D(z)} R_d(z,c)dG(z,c) + \int_0^{C_X(z)} R_x(z,c)dG(z,c))/\overline{P}(z)}{\int_0^{C_D(z)} C_d(z,c)dG(z,c) + \int_0^{C_X(z)} C_x(z,c)dG(z,c)} \\ &= \overline{\Phi}(z).\end{aligned}$$

4 Empirical Analysis

in this section, we provide two empirical tests on our theory. The first one is a reduced form analysis, which shows that exporters skew more of their exports toward the better selling products in markets where they face tougher competition due to comparative advantage. This is followed by a more structural analysis, in which we calibrate our model to the Chinese economy and quantify the different components of comparative advantage via a sufficient statistic approach.

4.1 Comparative Advantage and Export Product Mix

Exporters face different levels of competition across different markets. For example, they face tougher competition in larger markets (Mayer et al., 2014). Our theory emphasizes competition induced by comparative advantage. Capital intensive firms face tougher competition in capital abundant markets while labour intensive firms face tougher competition in labour abundant markets. To capture this channel, we need to first measure the competition faced by firms due to comparative advantage in each market. We propose the following two measures. The first is given by

$$CA1_{ij} = (z_i - \overline{z})(\ln \frac{K_j}{L_j} - \ln \overline{\frac{K}{L}}),$$

where z_i is the capital intensity of firm i and $\overline{z}$ is the average capital intensity of all Chinese manufacturing firms, $\frac{K_j}{L_j}$ is the capital to labour ratio of market j and $\overline{\frac{K}{L}}$ is the average capital to labour ratio of all markets (other than China). The larger that $CA1_{ij}$ is, the tougher the competition that exporter i will face in market j. The reason is that $CA1_{ij}$ would be higher if z_i is high above $\overline{z}$ and $\frac{K_j}{L_j}$ is also high above $\overline{\frac{K}{L}}$, or if z_i is far below $\overline{z}$ and $\frac{K_j}{L_j}$ is also far below $\overline{\frac{K}{L}}$. In both cases, firm i faces tough competition in market j since the market is abundant in the factor that firm i uses intensively. Alternatively, we can use firms' capital to labour ratio instead of the capital intensity and have the following measure

$$CA2_{ij} = (\frac{k_i}{l_i} - \overline{\frac{k}{l}})(\ln \frac{K_j}{L_j} - \ln \overline{\frac{K}{L}}),$$

where $\frac{k_i}{l_i}$ is the capital to labour ratio of firm i and $\overline{\frac{k}{l}}$ is the average capital to labour ratio of all Chinese firms.

To construct these measures, we need to estimate the capital to labour ratio for each destination market. We use the Penn World Table 9.0, which provides estimates of capital stock (at constant prices) and employment.[18] The capital to labour ratio of each country is then computed as the ratio of capital stock to employment. The world average capital to labour ratio $\overline{\frac{K}{L}}$ is computed as the average of the capital to labour ratio across all countries except China.[19] For the firm level measures, capital intensity is the same measure that we used in constructing motivating evidence. To measure the capital to labour ratio of each firm, following Brandt et al. (2012), we first estimate the capital stock for each firm using the perpetual inventory method. Labour is measured as the total number of employees. The average capital intensity $\overline{z}$ and capital to labour ratio $\overline{\frac{k}{l}}$ are computed as the simple average across all Chinese firms.

To compare our results with Mayer et al. (2014), we also use data for the year 2003. Table 1 shows our first result, which extends their basic empirical analysis on the exporters' product mix by including our new competition measures. The dependent variable is the logarithm of the ratio of exports between the core product and the second best product in each market for each firm.[20] We include the GDP of each market to capture the competition due to the effect of market size. Following Mayer et al. (2014), we also include the supply potential to capture the competition due to geography.[21] As we can see, the market size effect highlighted in their paper remains highly significant. The supply potential is positive but not precisely estimated. Our new competition measures are positive and significant. That is to say, in markets where firms face tougher competition due to comparative advantage, exports are more skewed toward the core product.

Table 2 looks at the skewness across all products that firms export to a market. The skewness is measured by the Herfindhal or the Theil index. Here, we control for the market fixed effect and firm fixed effect. The market fixed effect will capture the size and geography of the destination market. As we can see, the skewness measures tend to be higher in markets where firms have comparative disadvantage. That is to say, exports are more skewed in foreign markets where

[18]The data are available at http://www.rug.nl/ggdc/

[19]We exclude China from the sample to make the measure more exogenous, but in fact adding China makes little difference.

[20]The product rank is the rank at the local market.

[21]Markets which are closer to other markets have more potential competitors. The supply potential variable is constructed as the predicted aggregate exports to a market based on a gravity regression with the usual gravity variables and fixed effects.

Table 1: Comparative advantage and sales ratio between the core and second best product: 2003

	(1)	(2)	(3)	(4)
dependent variable				
ln GDP	0.0165***	0.0145***	0.0149***	0.0151***
	(0.00372)	(0.00385)	(0.00399)	(0.00401)
ln supply potential		0.0123*	0.0128	0.0134
		(0.00747)	(0.00869)	(0.00818)
CA1			0.0702***	
			(0.0243)	
CA2				0.00764**
				(0.00365)
Constant	-0.000704	-0.000600	0.000837	-0.000153
	(0.00762)	(0.00760)	(0.00864)	(0.00813)
Within R^2	0.000119	0.000128	0.000233	0.000166
No. of observations	85104	85104	85103	85063

Notes: The dependent variable is the logarithm of the ratio of exports between the core product (m=0) and second best product (m=1) in each market for each firm. CA1 and CA2 measure competition due to comparative advantage (higher value is associated with tougher competition). CA1 is an interaction term between firms' capital intensity (relative to all other firms) and the destination market's capital-labour ratio (relative to the world average). CA2 is another measure, which is an interaction term between firms' capital-labour ratio (relative to all other firms) and the destination market's capital to labour ratio (relative to the world average). We apply country-specific random effects on firm-demeaned data. Standard errors are reported in parentheses. The number of observations is significantly less than the following two tables since the dependent variable can be constructed only if the firms export at least two products at the destination. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01 respectively.

Table 2: Comparative advantage and the skewness of export sales: 2003

	(1)	(2)	(3)	(4)
dependent variable	Herfindal	Herfindal	Theil	Theil
CA1	0.0142***		0.0264***	
	(0.00272)		(0.00537)	
CA2		0.00206***		0.00398***
		(0.000503)		(0.00105)
country fixed effect	Y	Y	Y	Y
firm fixed effect	Y	Y	Y	Y
No. of observations	187180	187180	187180	187180

Notes: CA1 and CA2 measure competition due to comparative advantage (higher value is associated with tougher competition). CA1 is an interaction term between firms' capital intensity (relative to all other firms) and the destination market's capital-labour ratio (relative to the world average). CA2 is another measure, which is an interaction term between firms' capital-labour ratio (relative to all other firms) and the destination market's capital to labour ratio (relative to the world average). Standard errors are clustered at firm level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01 respectively.

Table 3: Comparative advantage and firms' export product scope: 2003

	(1)	(2)	(3)	(4)	(5)	(6)
dependent variable	ln product #	ln product #	product #	product #	product #	product #
CA1	-0.0343***		-0.146***		-0.114***	
	(0.00867)		(0.0227)		(0.00920)	
CA2		-0.00371**		-0.0372***		-0.0274***
		(0.00161)		(0.00451)		(0.00159)
Constant					1.548***	1.546***
					(0.00764)	(0.00765)
country fixed effect	Y	Y	Y	Y	Y	Y
firm fixed effect	Y	Y	Y	Y	Y	Y
R^2	0.451	0.451				
No. of observations	187180	187180	187180	187180	187180	187180

Notes: Columns (1) and (2) use OLS method and ln(product count) as the dependent variable. Columns (3) to (6) use product count as the dependent variable. Columns (3) and (4) use Poisson method while columns (5) and (6) use negative binomial method. CA1 and CA2 measure competition due to comparative advantage (higher value is associated with tougher competition). CA1 is an interaction term between firms' capital intensity (relative to all other firms) and the destination market's capital-labour ratio (relative to the world average). CA2 is another measure, which is an interaction term between firms' capital-labour ratio (relative to all other firms) and the destination market's capital to labour ratio (relative to the world average). Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01 respectively.

exporters face tougher competition due to comparative advantage.

Table 3 examines the effect on product scope, that is, the number of products exported by firms to each market. Again, we control for the market fixed effect and firm fixed effect. We use ln(product count) as the dependent variable in columns (1) and (2), and product count from column (3) to (6). In all cases, firms tend to export fewer products in markets where they face tougher competition due to comparative advantage.

To sum up, the evidence is consistent with Propositions 2 and 3 that firms facing tougher competition due to comparative advantage have narrower export product scope and more skewed export product mix.

4.2 Quantification of Comparative Advantage

We have shown that different measures of comparative advantage capture different margins of reallocations in action. Measures capturing only the extensive margin miss the dampening force. Does such a distinction quantitatively make a difference? How important are the different components of comparative advantage? To answer these questions, we need to quantify and decompose comparative advantage using these different measures. However, there are a few challenges in doing so. First, we do not observe the *ex ante* comparative advantage, which depends on the relative productivity differences across industries between the home country and the RoW under autarky. We observe only the open economies.[22] Second, even for the open economies which we can observe, measuring the relative productivities between the home country and the RoW remains difficult. One practical obstacle is that we do not have the firm-level data for the RoW. Even if we had the data, selection into exports would posit a significant challenge in estimating the underlying productivities (Costinot et al., 2012).[23] Finally, it is challenging to measure the endogenous components directly. They depend either on the relative cost upper-bound of the Pareto distribution or on the relative cost cut-off between the two economies, which do not have clear empirical counterparts.

Given these challenges, we provide an identification result which shows that only the export propensity $\chi(z)$ and export intensity $\lambda(z)$ for the *home country*, trade elasticity k, and trade freeness ρ and ρ^* are needed to measure and decompose comparative advantage. In other words, $\chi(z)$, $\lambda(z)$, k, ρ, and ρ^* are *sufficient statistics* for comparative advantage and its subcomponents.

Proposition 7. *We can write comparative advantage (as defined in Propositions 5 or 6) and its different subcomponents as functions of the trade elasticity k, the trade freeness ρ and ρ^*, the export propensity $\chi(z)$, and export intensity $\lambda(z)$.*

Proof. See Appendix 10.7. □

To quantify comparative advantage using this result, we calibrate the model to the Chinese economy. We first estimate the Pareto shape parameter following the method proposed by Corcos et al. (2011). The basic idea is conduct a log-log regression of $G(z, c)$ on c, while firm marginal cost is approximated by the inverse of the estimated TFP.[24] Using their method, we estimate the Pareto shape parameter to be $k = 5.51$. ρ and ρ^*, the freeness of trade, are estimated using the Head-Ries Index (Head and Ries, 2001) and the World Input Output Database. The details of the estimation are in Appendix 8. The results are presented in Appendix Table C10. As we

[22] Most modern economies are far from economic autarky. Historically, economic autarky is less unusual but only a few cases have been studied by economists. Bernhofen and Brown (2004) investigate the sudden opening-up of Japan in the 1860s to test the theory of comparative advantage. Irwin (2005) studies the welfare cost of the Jeffersonian Trade Embargo from 1807 to 1809. Kung and Ma (2014) exploit the severe trade suppression during 1550-1567 in Ming China to study the relationship between autarky and piracy.

[23] Costinot et al. (2012) argue that relative producer prices are good measures of relative productivity. However, we do not have the relative producer prices between China and the RoW.

[24] Given that we do not observe output price in the data we have, only TFPR can be estimated. Corcos et al. (2011) find that the structure of the model can be exploited to correct for such a bias. We estimate firms' revenue productivity for each CIC 2-digit sector using the method by Ackerberg et al (2015) and follow the procedures from Corcos et al. (2011) to estimate the Pareto shape parameter, conditioning year and industry fixed effects. For robustness, we have also checked the results using the median trade elasticity of 5.03 from the literature (Head and Mayer, 2014) and experimented with relative low and high elasticities of 2.5 and 7.5. The results are qualitatively the same.

can see, the freeness of trade between China and the RoW has been increasing over time. The average freeness was 0.043 in 2000 and rose to 0.058 in 2003 and 0.071 in 2006. Given the trade elasticity $k = 3.43$, the implied average iceberg trade costs dropped from 2.50 in 2000 to 2.16 in 2006. Finally, we measure the export propensity $\chi(z)$ by the percentage of firms that export, and the export intensity $\lambda(z)$ by the percent of sales exported, which are the data underlining Figure 1.

4.2.1 Validating the calibration

Before getting the result, we validate the estimation by evaluating the model prediction on moments that have not been used in the estimation. Our sufficient statics rely only on information about exports. We can evaluate the model prediction on imports. According to the model, the volume of exports from China to the RoW in industry z is given by $EXP(z) = \frac{1}{2\gamma(k+2)C_M(z)^k} N_E(z) C_D^*(z)^{k+2} L^* \rho$, and the volume of imports from the RoW to China is $IMP(z) = \frac{1}{2\gamma(k+2)C_M^*(z)^k} N_E^*(z) C_D(z)^{k+2} L \rho^*$. Therefore, the ratio of imports to exports is

$$\frac{IMP(z)}{EXP(z)} = \frac{L}{L^*} \frac{\rho^*}{\rho} \frac{N_E^*(z)}{N_E(z)} \frac{C_M(z)^k}{C_M^*(z)^k} \frac{C_D(z)^{k+2}}{C_D^*(z)^{k+2}},$$

which depends on the relative market size $\frac{L}{L^*}$, relative trade freeness $\frac{\rho^*}{\rho}$, the relative number of entrants $\frac{N_E^*(z)}{N_E(z)}$, and comparative advantage captured by $\frac{C_M(z)^k}{C_M^*(z)^k} \frac{C_D(z)^{k+2}}{C_D^*(z)^{k+2}}$. We can estimate $\frac{C_D(z)^{k+2}}{C_D^*(z)^{k+2}} \frac{C_M(z)^k}{C_M^*(z)^k}$ directly using the sufficient statistics results from Proposition 7. How well does this explain the variation of $\frac{IMP(z)}{EXP(z)}$ in the data? Answering this question helps to validate the calibrated model.

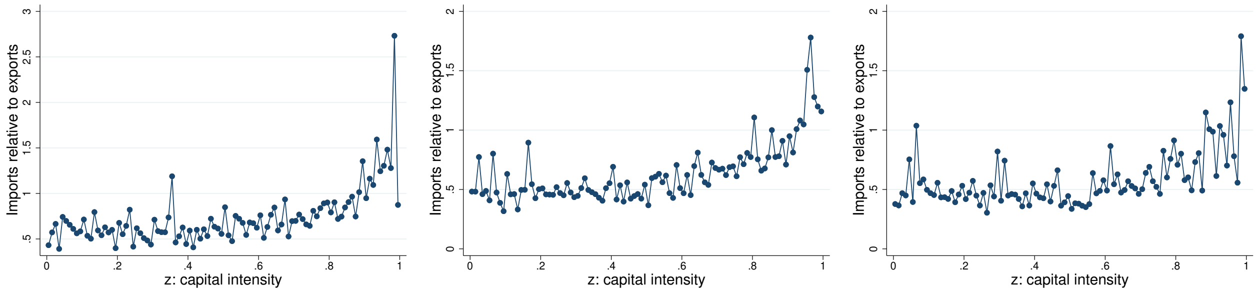
Our matched firm and customs data contain imports for importers. We assume that the imports of industry z from the RoW are the total imports of importers from industry z in China.[25] Under this assumption, we find $\frac{IMP(z)}{EXP(z)}$ tends to increase with capital intensity z, as shown in Figure 5. For the most capital intensive industries, China ran trade deficits since $\frac{IMP(z)}{EXP(z)} > 1$.

On Figure 6, we plot $\ln(\frac{IMP(z)}{EXP(z)})$ against $\ln(\frac{C_M(z)^k}{C_M^*(z)^k} \frac{C_D(z)^{k+2}}{C_D^*(z)^{k+2}})$. As can be seen, there is a very strong positive correlation. China tends to run trade deficits in industries of comparative disadvantage. We confirm this by regressing $\ln(\frac{IMP(z)}{EXP(z)})$ on $\ln(\frac{C_M(z)^k}{C_M^*(z)^k} \frac{C_D(z)^{k+2}}{C_D^*(z)^{k+2}})$.[26] The results are

[25] Ideally, we would like the firm-level data for the RoW to get exports to China by capital intensity.

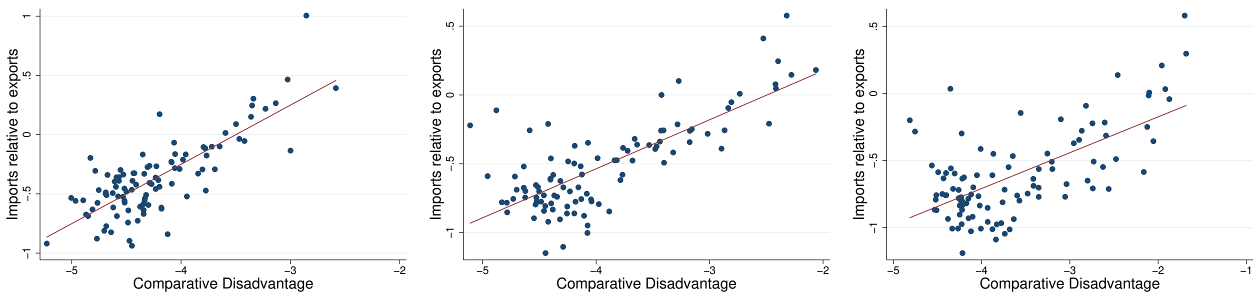
[26] Ideally, we would like to run this regression: $ln(\frac{IMP(z)}{EXP(z)}) = a_0 + a_1 \ln(\frac{C_M(z)^k}{C_M^*(z)^k} \frac{C_D(z)^{k+2}}{C_D^*(z)^{k+2}}) + a_2 \ln \frac{N_E^*(z)}{N_E(z)} + \varepsilon_z$.

shown in Appendix Table C9. The coefficients for $\ln(\frac{C_M(z)^k}{C_M^*(z)^k}\frac{C_D(z)^{k+2}}{C_D^*(z)^{k+2}})$ are positive and highly significant. Comparative advantage explains around half of the variation in $\ln(\frac{IMP(z)}{EXP(z)})$ and remains robust after controlling for capital intensity.



Year 2000 Year 2003 Year 2006

Figure 5: Chinese imports relative to exports by industries



Year 2000 Year 2003 Year 2006

Notes: The horizontal axis is $ln(\frac{C_M(z)^k}{C_M^*(z)^k}\frac{C_D(z)^{k+2}}{C_D^*(z)^{k+2}})$. Higher values indicate greater comparative disadvantage of China relative to the RoW. The vertical axis plots the logarithm of Chinese imports from the RoW relative to exports to the RoW.

Figure 6: Imports relative to exports and comparative disadvantage

4.2.2 Results

Armed with the calibrated parameters and the data, we quantify and decompose the comparative advantage of China relative to the RoW in 2000, 2003, and 2006, using Proposition 7. First, the *ex ante* component is the same for the two measures of comparative advantage as shown in the proof of Proposition 7 (we invert relative cost so that it is comparable with relative TFPQ). This is plotted in Figure 7. To filter out the noise in the data, we use local polynomials to

Our theory predicts that $a_0 = \frac{L}{L^*}\frac{\rho^*}{\rho}$, $a_1 = a_2 = 1$. However, $\frac{N_E^*(z)}{N_E(z)}$ is not observable.

fit the data, with confidence intervals indicated. According to the result, China had *ex ante* comparative advantage in the labour intensive industries. Over time, the *ex ante* component appears to favour the labour intensive industries.[27]

To single out the endogenous components of comparative advantage, we divide the overall comparative advantage by the *ex ante* component and get

$$\frac{C_D(z)^T}{C_D(z)^{*T}} / \frac{C_D(z)^A}{C_D(z)^{*A}} = \left[\frac{1-\rho C_M^*(z)^k / C_M(z)^k}{1-\rho^* C_M(z)^k / C_M^*(z)^k}\right]^{\frac{1}{k+2}}, \tag{4.13}$$

$$\frac{\overline{\Phi}(z)^T}{\overline{\Phi}^*(z)^T} / \frac{\overline{\Phi}(z)^A}{\overline{\Phi}(z)^{*A}} = \left(\frac{\overline{\Phi}(z)^A}{\overline{\Phi}(z)^{*A}}\right)^{k+1} \frac{L^*}{L} \frac{\frac{L}{L^*}\left(\frac{C_D(z)}{C_D^*(z)}\right)^{k+1} + \rho}{1+\rho^* \frac{L}{L^*}\left(\frac{C_D(z)}{C_D^*(z)}\right)^{k+1}}. \tag{4.14}$$

The right hand side of the equations above are left with the endogenous components. They are plotted in Figures 8 and 9. As can be seen from Figure 8, the endogenous component of relative cost cut-off tends to favour labour intensive industries (relative cost is inverted to be comparable with relative TFPQ. Here, it is the RoW relative to China). This is not surprising given that the theory predicts that it is positively correlated with the *ex ante* component which favours the labour intensive industries.

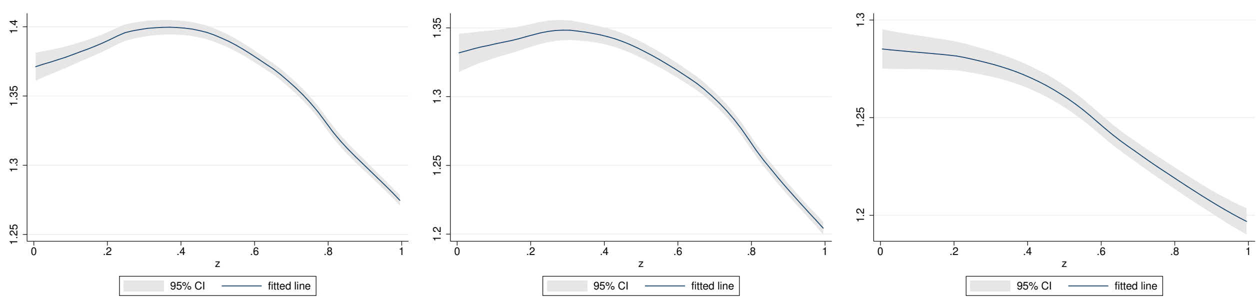
However, as is evident from Figure 9, the endogenous component of relative TFPQ tends to favour capital intensive industries. Given that our theory predicts that the dampening component is negatively correlated with the *ex ante* component, the dampening component would favour the capital intensive industries. This implies that the dampening component must have dominated the amplifying component such that it determines how the endogenous component will vary with capital intensity.

To examine the effect of the endogenous components on comparative advantage, we plot the inferred overall comparative advantage in Figures 10 and 11. Figure 10 plots the overall comparative advantage which captures only the extensive margin, i.e., $\frac{C_D(z)^{*T}}{C_D(z)^T}$. Figure 11 plots the overall comparative advantage which captures both the extensive and intensive margins, i.e., $\frac{\overline{\Phi}(z)^T}{\overline{\Phi}^*(z)^T}$. Both measures tend to favour the labour intensive industries. Given that the endogenous component of the measure capturing only the extensive margin amplifies the *ex ante* component as we saw in Figure 8, it is more variant than the *ex ante* component since all the lines are

[27] Huang et al. (2017) also find that the exogenous Ricardian comparative advantage increasingly favoured labour intensive industries in China during the period 1999-2007.

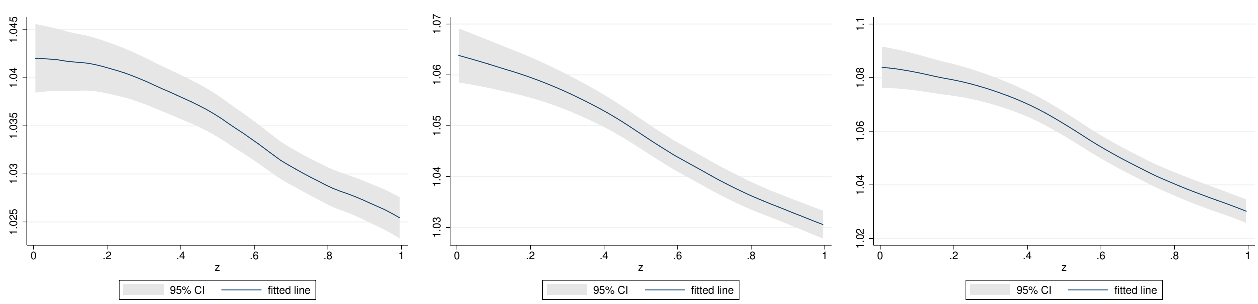
steeper in Figure 10 than in Figure 7. However, due to the dampening effect of the endogenous component, the measure that captures both margins is less variant than the *ex ante* component since all the lines in Figure 11 are flatter than in Figure 7.

Table 4 confirms a similar message. Column (1) reports the regression coefficients of capital intensity out of an OLS regression which regresses the *ex ante* comparative advantage on capital intensity. Indeed, given the negative coefficients, the home country tends to be less productive in the capital intensive industries *ex ante*. These coefficients become even more negative in Column (2) when we replace the dependent variable by the measure of comparative advantage which captures only the extensive margin. This shows the effect of the amplifying endogenous component. However, the coefficients become less negative in Column (3) when we replace the dependent variable by relative TFPQ which captures both margins. This again shows that the dampening component dominates the amplifying component.



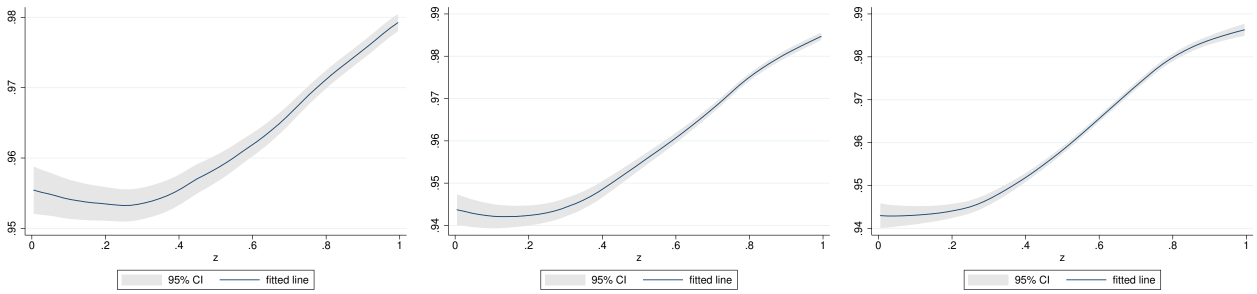
year 2000 year 2003 year 2006

Figure 7: The ex ante component of comparative advantage



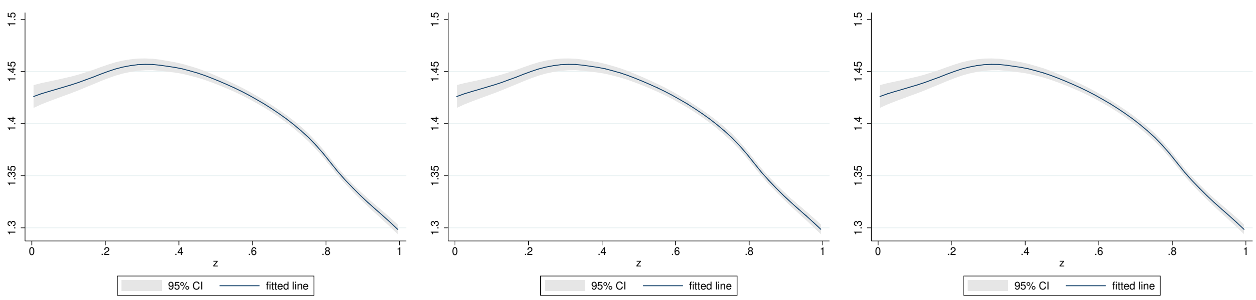
year 2000 year 2003 year 2006

Figure 8: The endogenous component of relative marginal cost cut-off



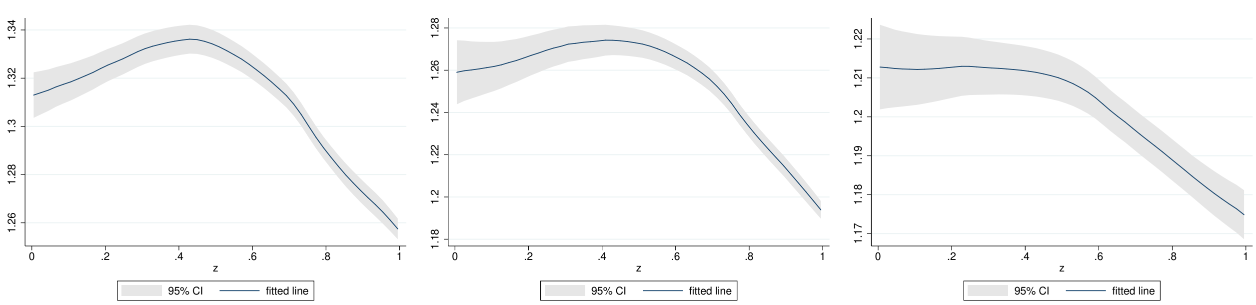
year 2000 year 2003 year 2006

Figure 9: The endogenous component of relative TFPQ



year 2000 year 2003 year 2006

Figure 10: Comparative advantage measured by relative marginal cost cut-off



year 2000 year 2003 year 2006

Figure 11: Comparative advantage measured by relative TFPQ

Table 4: Comparing different measures of comparative advantage

Regressor capital intensity	Dependent variables		
	(1)	(2)	(3)
year	*ex ante* comparative advantage	relative cost cut-off	relative TFPQ
2000	-.109	-.143	-.066
2003	-.142	-.218	-.069
2006	-.127	-.232	-.053

Notes: The table reports the coefficients of capital intensity out of regressions which regress the different measures of comparative advantage on capital intensity. The dependent variable of column (1) is the ex ante comparative advantage implied by the sufficient statistics. For column (2), it is the relative cost cut-off estimated using the sufficient statistics. For column (3), it is the relative TFPQ estimated from the sufficient statistics. All coefficients are significant at the 0.1% level.

5 Numerical Simulations

In this section, we parametrize and simulate the single-product model.[28] We are particularly interested in the way that trade liberalization (lower variable trade costs) affects comparative advantage. We are also interested in comparing the associated welfare change with the homogeneous firm model.

5.1 Parameters

We assume that $C_M(z)$ and $C_M(z)^*$, the cost upper bounds for the home country and the RoW, which determine the ex ante comparative advantage, can be parametrized as

$$\begin{aligned} C_M(z) &= az + b, \\ C_M(z)^* &= a^*z + b^*. \end{aligned}$$

We also assume that $a > 0$ and $a^* < 0$. Therefore, the home country has comparative advantage in low z industries and the RoW has comparative advantage in high z industries. The key parameters of the model are given by Table 5. We assume that $C_M(z) = 1.3 + 0.3z$, and $C_M(z)^* = 1.6 - 0.3z$. As can be seen in panel (a) of Figure 12, the cost upper bounds are symmetric around $z = 0.5$ for the two economies.

Moreover, we set the size of the two economies to be the same $L = L^* = 10$. This is to neutralize the market size effect. Consumers' endowments of the homogeneous good and incomes

[28]The purpose of the simulations is to demonstrate the model channel and provide numerical comparative statics. For future work, we will structurally estimate a multi-product model.

Table 5: Model parameters

Parameter	Definition	Value
a	the slope of $C_M(z)$	0.3
b	the intercept of $C_M(z)$	1.3
a^*	the slope of $C_M(z)^*$	-0.3
b^*	the intercept of $C_M(z)^*$	1.6
f_E	fixed cost of firm entry	0.5
k	Pareto Shape	2.5
α	consumer preference	5
γ	consumer preference	1
η	consumer preference	0.5
τ	iceberg trade cost, home to foreign	[1.3, 1.4]
τ^*	iceberg trade cost, foreign to home	[1.3, 1.4]
L	number of consumers in the home country	10
L^*	number of consumers in the foreign country	10
y_0^c	home country consumers' endowments of homogeneous good	50
y_0^{c*}	foreign country consumers' endowments of homogeneous good	50
I	home country consumers' labour income	50
I^*	foreign country consumers' labour income	50

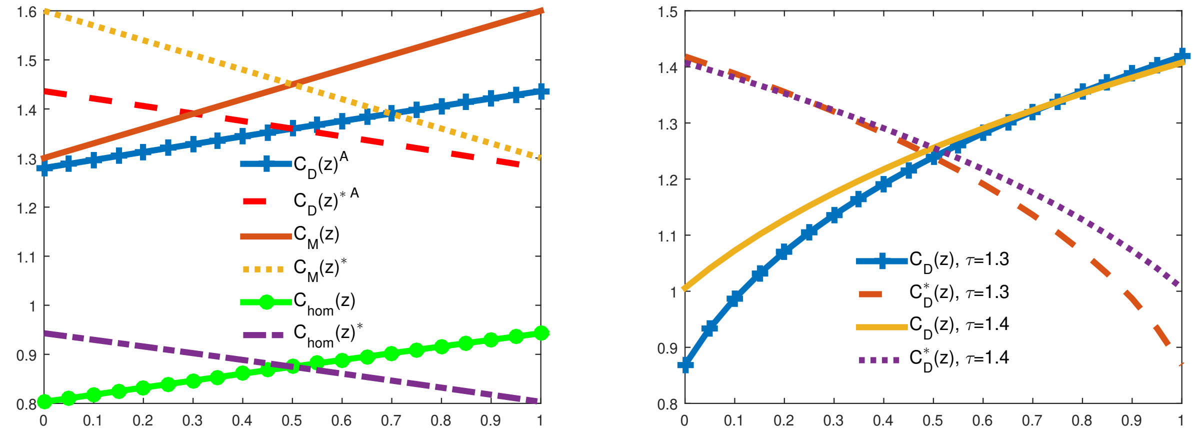
are set to be $y_0^c = y_0^{c*} = 50$ and $I = I^* = 50$, respectively. These values are high enough to ensure that the demand for the homogeneous good is positive.

5.2 Simulation Results

Given that the two economies have the same size, we expect the cost cut-offs under autarky $C_D(z)^A$ and $C_D(z)^{A*}$ to be symmetric around $z = 0.5$ as well. This is indeed the case in Figure 12 (a).[29] Since we are interested in comparing the welfare change with the homogeneous firm model, we follow Melitz and Redding (2015) to find marginal cost profiles $C_{hom}(z)$ and $C_{hom}(z)^*$ for the homogeneous firm model such that it has the same welfare level as the heterogeneous firm model under autarky. The detailed procedure in finding $C_{hom}(z)$ and $C_{hom}(z)^*$ are in Appendix 11.2. For the model that we have parametrized, the associated $C_{hom}(z)$ and $C_{hom}(z)^*$ are plotted on Figure 12 (a) as well. They turn out to follow a pattern similar to $C_M(z)$ and $C_M(z)^*$ as well and are symmetric around $z = 0.5$.

Figure 12 (b) plots the equilibrium cost cut-offs in the open economy $C_D(z)$ and $C_D(z)^*$ for two scenarios, one in which the iceberg trade costs are $\tau = \tau^* = 1.4$, and the other with

[29] Our choice of parameters guarantees that $C_D(z)^A < C_M(z)$ and $C_D(z)^{*A} < C_M(z)^*$ to ensure that there is export selection. The conditions are $C_M(z) > \sqrt{2(k+1)(k+2)\gamma/L}$ and $C_M(z)^* > \sqrt{2(k+1)(k+2)\gamma/L^*}$.

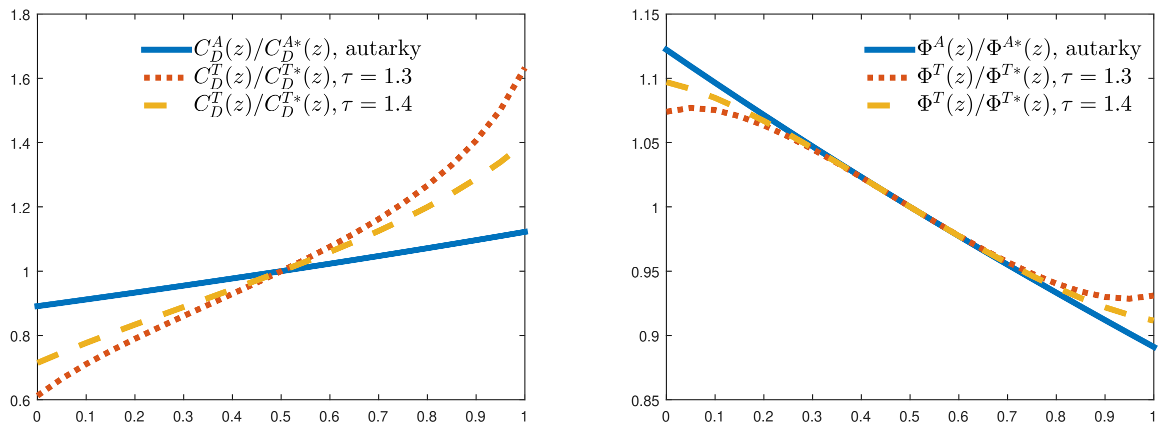


Notes: The solid lines indicate the home country while the dashed or dotted lines indicate the RoW. $C_{hom}(z)$ and $C_{hom}(z)^*$ are the marginal costs inferred for the homogeneous firm model such that it has the same welfare level as the heterogeneous firm model under autarky (see Appendix 11.2 for more details).

Figure 12: Cost upper bounds and cut-off costs

bilateral trade liberalization such that the iceberg trade costs are reduced to $\tau = \tau^* = 1.3$. Two observations result from comparing these two scenarios. First, bilateral trade liberalization appears to widen the gap between $C_D(z)$ and $C_D(z)^*$. Both $C_D(z)$ and $C_D(z)^*$ become steeper in the case with lower trade costs. It suggests that if we use industry productivity measures which capture only the extensive margin, we will find comparative advantage strengthened by bilateral trade liberalization. Second, bilateral trade liberalization can change the cut-offs of different industries in different directions. In Figure 12 (b), when trade costs are reduced, the home country cost cut-offs of industries closer to $z = 1$ rise, while those closer to $z = 0$ fall. This result can be explained as follows. In industries where the home country has comparative advantage, trade liberalization reduces the cut-offs by increasing the accessibility of the foreign market. Furthermore, there will be more entrants in these industries, given the rising profits of exporting. This drives the cut-offs further down. Trade liberalization also increases market accessibility for the industries where the home country has comparative disadvantage. However, it also makes the very efficient foreign competitors even more efficient in serving the home country, which deters home entrants and lessens competition in the home country. If the entry channel is more pronounced, the cut-offs will rise after trade liberalization.[30]

[30] The mechanism is similar to the discussion of preferential trade liberalization in Melitz and Ottaviano (2008).

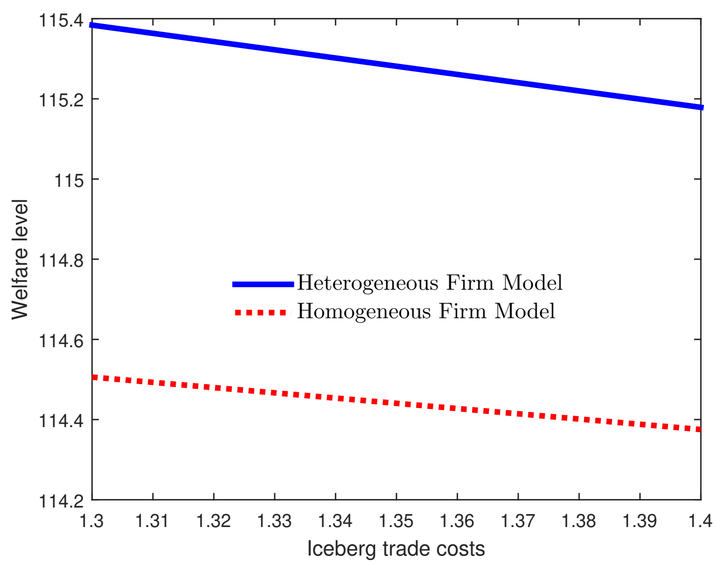


Relative average marginal cost

Relative TFPQ

Notes: Panel (a) plots the relative average firm marginal cost between the home country and the foreign country under autarky and open economy, which are simply the relative cost cut-off. Panel (b) plots the relative quantity-based TFP (TFPQ) between the home country and the foreign under economic autarky and open economy.

Figure 13: Bilateral trade liberalization and comparative advantage



Notes: This figure plots the welfare level of the homogeneous firm and heterogeneous firm model with respect to iceberg trade costs. The model parameters are chosen such that the two models have the same welfare level under autarky.

Figure 14: Trade liberalization and welfare

We now examine how bilateral trade liberalization affects comparative advantage. As suggested above, if the industry productivity measure incorporates only the extensive margin, we expect comparative advantage to be strengthened. This is indeed the case in Figure 13 (a). As trade costs go down, the relative cost cut-off get steeper. In contrast, if we use TFPQ to measure industry productivity, the relative TFPQ becomes flatter when we reduce trade costs. Therefore, comparative advantage is weakened by bilateral trade liberalization. This suggests that the endogenous dampening force becomes stronger and dominates when trade costs fall.

Finally, we compare the welfare level of our heterogeneous firm model with a homogeneous model which has the same welfare level under autarky. The details on the welfare formula are given in Appendix 11.2. Melitz and Redding (2015) show that the Melitz model with heterogeneous firms predicts higher welfare gains from trade than the Krugman model with homogeneous firms. While the mark-up is constant in the models that they consider, it is variable in the models that we examine. As Figure 14 indicates, at least in the parameter space that we specify, the heterogeneous firm model still predicts higher welfare gains from trade than the homogeneous firm model.

6 Conclusion

We uncover new stylized facts on the way in which comparative advantage shapes intra- and inter-firm reallocations. Not all of the facts can be reconciled with existing models with constant mark-ups. We construct a model which interacts firm heterogeneity with comparative advantage, featuring variable mark-ups. The model simultaneously explains the facts and generates novel insights on the way in which firm heterogeneity affects comparative advantage. We find that exporters face tougher competition in comparative disadvantage industries. Such an effect from competition induces exporters to cut their product scope and skew their product mix in the comparative disadvantage industries. We also find that export selections along the intensive margin generate endogenous Ricardian comparative advantage, which is *negatively* correlated with the *ex ante* comparative advantage. This contrasts with the amplifying mechanism found by Bernard et al. (2007). In both our calibrated Chinese economy and the simulated model, we find that the dampening force can dominate the amplifying force.

To conclude, while comparative advantage has important implications for the micro be-

In our case, formally, we have $\frac{\partial C_D(z)^{*k+2}}{\partial \rho} = \frac{\beta}{L^2}\frac{2\rho C_M(z)^k-(1+\rho^2)C_M^*(z)^k}{(1-\rho^2)^2}$. Therefore, the effect of bilateral trade liberalization on cost cut-offs is positive if $2\rho C_M(z)^k \geq (1+\rho^2)C_M^*(z)^k$ is satisfied. This is more likely to be the case for high z industries in the home country.

haviour of individual firms, the micro level responses from firms have profound macro implications. Some of the macro implications, such as welfare gains from trade, appear robust to the model specification. Other macro implications, such as comparative advantage, appear to depend on whether we allow for variable mark-ups or not.

References

[1] Ackerberg, D.A., Caves, K. and Frazer, G., 2015. Identification properties of recent production function estimators. Econometrica, 83(6), pp.2411-2451.

[2] Arkolakis, C., Costinot, A. and Rodríguez-Clare, A., 2012. New trade models, same old gains?. The American Economic Review, 102(1), pp.94-130.

[3] Arkolakis, C., Costinot, A., Donaldson, D. and Rodríguez-Clare, A., 2015. The elusive pro-competitive effects of trade. The Review of Economic Studies. Forthcoming.

[4] Arkolakis, C. and Muendler, M.A., 2010. The extensive margin of exporting products: A firm-level analysis (No. w16641). National Bureau of Economic Research.

[5] Balassa, B., 1965. Trade liberalisation and "revealed" comparative advantage. The Manchester School, 33(2), pp.99-123.

[6] Bernard, A.B., Jensen, J.B., Redding, S.J. and Schott, P.K., 2007. Firms in international trade. The Journal of Economic Perspectives, 21(3), pp.105-130.

[7] Bernard, A.B., Redding, S.J. and Schott, P.K., 2007. Comparative advantage and heterogeneous firms. The Review of Economic Studies, 74(1), pp.31-66.

[8] Bernard, A.B., Redding, S.J. and Schott, P.K., 2010. Multiple-product firms and product switching. The American Economic Review, 100(1), pp.70-97.

[9] Bernard, A.B., Redding, S.J. and Schott, P.K., 2011. Multiproduct firms and trade liberalization. The Quarterly Journal of Economics, 126(3), pp.1271-1318.

[10] Bernhofen, D.M. and Brown, J.C., 2004. A direct test of the theory of comparative advantage: the case of Japan. Journal of Political Economy, 112(1), pp.48-67.

[11] Brandt, L., Van Biesebroeck, J. and Zhang, Y., 2012. Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing. Journal of Development Economics, 97(2), pp.339-351.

[12] Burstein, A. and Vogel, J. International trade, technology, and the skill premium. Journal of Political Economy. Forthcoming.

[13] Corcos, G., Del Gatto, M., Mion, G. and Ottaviano, G.I., 2011. Productivity and firm selection: quantifying the 'new'gains from trade. The Economic Journal, 122(561), pp.754-798.

[14] Costinot, A., Donaldson, D. and Komunjer, I., 2012. What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas. The Review of Economic Studies, pp.581-608.

[15] Costinot, A., Donaldson, D. and Smith, C., 2016. Evolving comparative advantage and the impact of climate change in agricultural markets: Evidence from 1.7 million fields around the world. Journal of Political Economy, 124(1), pp.205-248.

[16] Dornbusch, R., Fischer, S. and Samuelson, P.A., 1977. Comparative advantage, trade, and payments in a Ricardian model with a continuum of goods. The American Economic Review, 67(5), pp.823-839.

[17] Eaton, J. and Kortum, S., 2002. Technology, geography, and trade. Econometrica, 70(5), pp.1741-1779.

[18] Eckel, Carsten, and J. Peter Neary. Multi-product firms and flexible manufacturing in the global economy. The Review of Economic Studies 77.1 (2010): 188-217.

[19] Eckel, C., Iacovone, L., Javorcik, B. and Neary, J.P., 2015. Multi-product firms at home and away: Cost-versus quality-based competence. Journal of International Economics, 95(2), pp.216-232.

[20] Feenstra, R. and Ma, H., 2009. Optimal Choice of Product Scope for Multiproduct Firms under Monopolistic Competition," E. Helpman, D. Marin and T. Verdier, eds., The Organization of Firms in a Global Economy , Harvard University Press, 173-199.

[21] French, S., 2017. Revealed comparative advantage: What is it good for?. Journal of International Economics, 106, pp.83-103.

[22] Gaubert, C. and Itskhoki, O., 2018. Granular comparative advantage (No. w24807). National Bureau of Economic Research.

[23] Head, K. and Mayer, T., 2014. Gravity equations: Workhorse, toolkit, and cookbook. Handbook of International Economics, 4.

[24] Head, K. and Ries, J., 2001. Increasing returns versus national product differentiation as an explanation for the pattern of US-Canada trade. American Economic Review, pp.858-876.

[25] Huang, H., Ju, J. and Yue, V.Z., 2017. Structural adjustments and international trade: Theory and evidence from China. Working Paper.

[26] Iacovone, L. and Javorcik, B.S., 2010. Multi-Product Exporters: Product Churning, Uncertainty and Export Discoveries. The Economic Journal, 120(544), pp.481-499.

[27] Kung, J.K.S. and Ma, C., 2014. Autarky and the rise and fall of piracy in Ming China. The Journal of Economic History, 74(2), pp.509-534.

[28] Levchenko, A.A. and Zhang, J., 2016. The evolution of comparative advantage: Measurement and welfare implications. Journal of Monetary Economics, 78, pp.96-111.

[29] Lu, D., 2010. Exceptional exporter performance? Evidence from Chinese manufacturing firms. manuscript, University of Chicago.

[30] Ma, Y., Tang, H. and Zhang, Y., 2014. Factor intensity, product switching, and productivity: Evidence from Chinese exporters. Journal of International Economics, 92(2), pp.349-362.

[31] Manova, K. and Yu, Z., 2016. How firms export: Processing vs. ordinary trade with financial frictions. Journal of International Economics, 100, pp.120-137.

[32] Manova, K. and Yu, Z., 2017. Multi-product firms and product quality. CEP Discussion Paper No 1469.

[33] Mayer, T., Melitz, M.J. and Ottaviano, G.I., 2014. Market size, competition, and the product mix of exporters. The American Economic Review, 104(2), pp.495-536.

[34] Mayer, T., Melitz, M.J. and Ottaviano, G.I., 2016. Product mix and firm productivity responses to trade competition (No. w22433). National Bureau of Economic Research.

[35] Melitz, M.J., 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. Econometrica, 71(6), pp.1695-1725.

[36] Melitz, M.J. and Ottaviano, G.I., 2008. Market size, trade, and productivity. The Review of Economic Studies, 75(1), pp.295-316.

[37] Melitz, M.J. and Redding, S.J., 2015. New trade models, new welfare implications. American Economic Review, 105(3), pp.1105-46.

[38] Ottaviano, G., Tabuchi, T. and Thisse, J.F., 2002. Agglomeration and trade revisited. International Economic Review, pp.409-435.

[39] Schott, P.K., 2004. Across-product versus within-product specialization in international trade. The Quarterly Journal of Economics, 119(2), pp.647-678.

[40] Timmer, M.P., Dietzenbacher, E., Los, B., Stehrer, R. and Vries, G.J., 2015. An illustrated user guide to the world input–output database: the case of global automotive production. Review of International Economics, 23(3), pp.575-605.

[41] Yu, M., 2015. Processing trade, tariff reductions and firm productivity: evidence from Chinese firms. The Economic Journal, 125(585), pp.943-988.

7 Complementary Tables

Table C1: Export propensity and intensity: 2000-2006

dependent variable:	Export Propensity			Export Intensity		
Sample:	All Exporters	Non-SOEs	Non-Processing Firms	All Exporters	Non-SOEs	Non-Processing Firms
capital intensity	-0.247***	-0.314***	-0.169***	-0.247***	-0.334***	-0.150***
	(0.0140)	(0.0126)	(0.0146)	(0.0103)	(0.0121)	(0.00905)
year FE	Y	Y	Y	Y	Y	Y
R^2	0.793	0.880	0.654	0.648	0.708	0.608
Observations	700	700	700	700	700	700

Notes: Export propensity is the percent of firms that are exporters. Export intensity is defined as the share of goods exported. Each observation is a year-industry while industry is defined as "HO aggregates". Year fixed effect is included in each regression. OLS is used. Standard errors clustered at each industries are reported in parentheses. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01 respectively.

Table C2: Export product scope: all exporters 2000-2006

dependent variable:	Mean Product Scope			Share of Single Product Firms		
Sample:	All Exporters	Non-SOEs	Non-Processing Firms	All Exporters	Non-SOEs	Non-Processing Firms
capital intensity	-1.921***	-1.989***	-0.719***	0.0950***	0.0988***	0.0564***
	(0.174)	(0.162)	(0.166)	(0.00594)	(0.00608)	(0.00912)
year FE	Y	Y	Y	Y	Y	Y
R^2	0.236	0.293	0.133	0.495	0.497	0.292
Observations	700	700	700	700	700	700

Notes: Mean product scope is the average number of products exported within each industry. Industry is defined as "HO aggregates". Year fixed effect is included in each regression. OLS is used. Standard errors clustered at each industries are reported in parentheses. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01 respectively.

Table C3: Export product mix: all exporters 2000-2006

	(1)	(2)	(3)	(4)	(5)
dependent variable	core product share	m0/m1	m0/m2	mean Herfindhal	mean Theil
capital intensity	0.0742***	0.311***	0.452***	0.0920***	0.238***
	(0.00373)	(0.0271)	(0.0318)	(0.00432)	(0.00987)
year FE	Y	Y	Y	Y	Y
R^2	0.553	0.260	0.307	0.581	0.613
Observations	700	700	700	700	700

Notes: The table contains results using the full sample of exporters. Industry is defined as "HO aggregates". The regressand of column (1) is the average sales share of the core product across firms within each industry, and the log sales ratio of the core product to the second best product in column (2), and the log sales ratio of the core product to the third best product in column (3). Column (4) and (5) regress the average Herfindhal index and Theil Index of exports on capital intensity. Standard errors clustered at each industries are reported in parentheses. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01 respectively.

8 The Head-Ries Index

We estimate the trade freeness between China and the Rest of World using the Head-Reis Index (Head and Ries, 2001). If we assume symmetric trade costs $\rho_{ij} = \rho_{ji}$ and zero domestic trade costs, then

$$\rho_{ij} = \sqrt{\frac{X_{ij} X_{ji}}{X_{ii} X_{jj}}},$$

where X_{ij} is the aggregate exports from region i to region j which follows the gravity equation.[31] So if let $i = China$ and $j = RoW$, we can infer the trade freeness between China and the RoW. However, to implement this equation, we need data on local absorption X_{ii} and X_{jj}. These are not available from our firm survey or customs data but available from the World Input Output

[31] Our model generates gravity equation for the sectoral trade flow which satisfies the general gravity equations classified by Head and Mayer (2014) even if firms produce multiple products.

Table C4: Export product mix: all non-SOE exporters 2000-2006

	(1) core product share	(2) m0/m1	(3) m0/m2	(4) mean Herfindhal	(5) mean Theil
capital intensity	0.0760***	0.311***	0.456***	0.0944***	0.243***
	(0.00369)	(0.0264)	(0.0303)	(0.00433)	(0.00976)
year FE	Y	Y	Y	Y	Y
R^2	0.558	0.262	0.312	0.586	0.619
Observations	700	700	700	700	700

Notes: The table contains results using the sample of none-state-owned-enterprises (SOE) exporters. Industry is defined as "HO aggregates". The regressand of column (1) is the average sales share of the core product across firms within each industry, and the log sales ratio of the core product to the second best product in column (2), and the log sales ratio of the core product to the third best product in column (3). Column (4) and (5) regress the average Herfindhal index and Theil Index of exports on capital intensity. Standard errors clustered at each industries are reported in parentheses. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01 respectively.

Table C5: Export product mix: all non-processing exporters 2000-2006

	(1) core product share	(2) m0/m1	(3) m0/m2	(4) mean Herfindhal	(5) mean Theil
capital intensity	0.0314***	0.123***	0.114*	0.0401***	0.100***
	(0.00494)	(0.0414)	(0.0684)	(0.00559)	(0.0120)
year FE	Y	Y	Y	Y	Y
R^2	0.178	0.0283	0.0146	0.210	0.225
Observations	700	700	700	700	700

Notes: The table contains results using the sample of non-processing exporters. Industry is defined as "HO aggregates". The regressand of column (1) is the average sales share of the core product across firms within each industry, and the log sales ratio of the core product to the second best product in column (2), and the log sales ratio of the core product to the third best product in column (3). Column (4) and (5) regress the average Herfindhal index and Theil Index of exports on capital intensity. Standard errors clustered at each industries are reported in parentheses. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01 respectively.

Table C6: Skewness of domestic sales within industry 2000-2006: all firms

	(1)	(2)	(3)
dependent variable	Herfindahl index	Theil Index	Inter quartile range of log sales
capital intensity	-0.0358***	-2.270***	-1.633***
	(0.00621)	(0.0958)	(0.0997)
year FE	Y	Y	Y
R^2	0.0905	0.712	0.601
Observations	700	700	700

Notes: The skewness of sales is measured across all firms within each industry, while industry is defined as "HO aggregates". Robust standard errors clustered at industry level are reported in parentheses. The constants are absorbed by the year fixed effects. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01 respectively.

Table C7: Skewness of domestic sales within industry 2000-2006: all non-SOEs

	(1)	(2)	(3)
dependent variable	Herfindahl index	Theil Index	Inter quartile range of log sales
capital intensity	-0.0413***	-2.529***	-2.357***
	(0.00684)	(0.128)	(0.174)
year FE	Y	Y	Y
R^2	0.0943	0.734	0.504
Observations	700	700	695

Notes: The skewness of sales is measured across all none-state-owned-firms within each industry, while industry is defined as "HO aggregates". Column (3) has less observations because the 25th percentile of non-SOE firm has zero domestic sales in certain industries. Robust standard errors clustered at industry level are reported in parentheses. The constants are absorbed by the year fixed effects. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01 respectively.

Table C8: Skewness of domestic sales within industry 2000-2006: all non-processing firms

	(1)	(2)	(3)
dependent variable	Herfindahl index	Theil Index	Inter quartile range of log sales
capital intensity	-0.0269***	-2.269***	-0.468***
	(0.00414)	(0.0826)	(0.0506)
year FE	Y	Y	Y
R^2	0.0680	0.716	0.388
Observations	700	700	700

Notes: The skewness of sales is measured across all none-processing firms (firms not engaged in processing exports) within each industry, while industry is defined as "HO aggregates". Robust standard errors clustered at industry level are reported in parentheses. The constants are absorbed by the year fixed effects. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01 respectively.

Table C9: Imports relative to exports and comparative advantage

dependent variable: imports	year 2000		year 2003		year 2006	
relative to exports $\ln(\frac{IMP(z)}{EXP(z)})$	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(\frac{C_M(z)^k}{C_M^*(z)^k}\frac{C_D(z)^{k+2}}{C_D^*(z)^{k+2}})$	0.500***	0.432***	0.356***	0.266***	.268***	0.330***
	(.050)	(.076)	(.036)	(.073)	(.0398)	(.093)
capital intensity z		0.157		.247		-.192
		(.114)		(.168)		(.24)
constant	1.75***	1.38***	0.887***	.411	.362**	0.685
	(.217)	(.377)	(.139)	(.411)	(.147)	(.453)
Adjusted R^2	0.626	0.634	0.578	0.588	0.411	0.416
N	100	100	100	100	100	100

Notes: The dependent variable is log total Chinese imports relative to exports within each industry. Robust standard errors are in parentheses. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01 respectively.

Database (WIOD).[32] Local absorption is computed as the total outputs minus total exports. We estimate trade freeness using the formula above for each sector. The summary statistics for the manufacturing sectors are displayed in Table C10.[33] The estimated average trade freeness between China and the RoW increased from 0.043 in 2000 to 0.071 in 2006. If the trade elasticity is $k = 5.51$, which is our estimated Pareto shape, then the implied iceberg trade costs $\tau = \rho^{-\frac{1}{k}}$ dropped from around 1.77 in 2000 to 1.62 in 2006. If we use the median trade elasticity 5.03 from the literature (Head and Mayer, 2014), the implied iceberg trade costs are slightly higher.

Table C10: Trade costs between China and the RoW: the Head-Ries index

year	trade freeness ρ			implied iceberg trade costs τ	
	average	min	max	k=5.51	k=5.03
2000	0.043	0.012	0.116	1.77	1.87
2001	0.045	0.012	0.129	1.75	1.85
2002	0.051	0.012	0.165	1.72	1.81
2003	0.058	0.013	0.218	1.67	1.76
2004	0.070	0.015	0.282	1.62	1.70
2005	0.073	0.015	0.323	1.61	1.68
2006	0.071	0.015	0.313	1.62	1.69

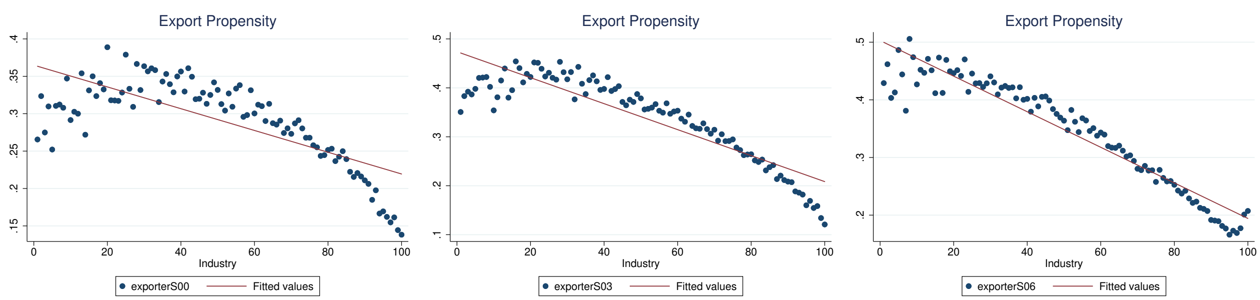
Notes: Trade freeness ρ is estimated using the Head and Ries (2001) method and the World Input Output Data for manufacturing industries. The columns titled "average", "min", and "max" are the average, minimum and maximum of the Head-Ries Index across 13 manufacturing sectors. The iceberg trade costs τ are inferred using the average trade freeness according to $\rho = \tau^{-k}$ where k is the trade elasticity.

[32] We use the 2013 release at http://www.wiod.org/database/wiots13. The details of the data can be found in Timmer, Dietzenbacher, Los and Vries (2015).

[33] There are 15 sectors of goods and 20 sectors of services. Manufacturing sectors include all the 15 goods sector except the sector of "Agriculture, hunting, forestry and fishing" and the sector of " Mining and quarrying". For brevity, we do not report the trade freeness for the service sectors. The trade freeness for services between China and the RoW is lower but increases over time.

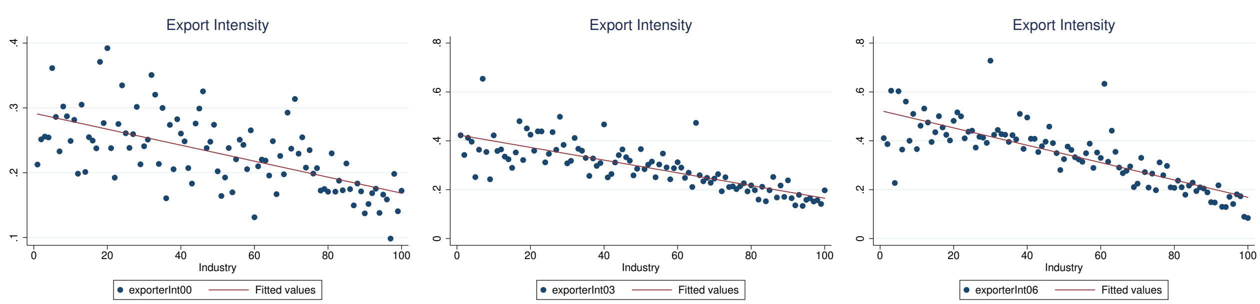
9 Complementary Figures

Our benchmark results only use data from year 2003. Now we include results using data for 2000 and 2006. Our stylized fact 3 states that export product mix is more skewed in capital intensive industries. Other than the measures of skewness used in the main text, we present results using other measures. Figure C5 plots the average sales share of the core product. The core product is defined as the product that makes up most of the total exports. As evident from the figures, the average share of sales from the core product is higher for the capital intensive industries. Figure C7 plots the average of the log-ratios between the sales of the core product to the third best product. Figure C8 plots the average Herfindhal Index of exporters for each industry. Similarly, we also include additional evidence for stylized fact 4 using alternative measures, including Figure C10 which plots the Herfindahl Index of domestic sales across firms.



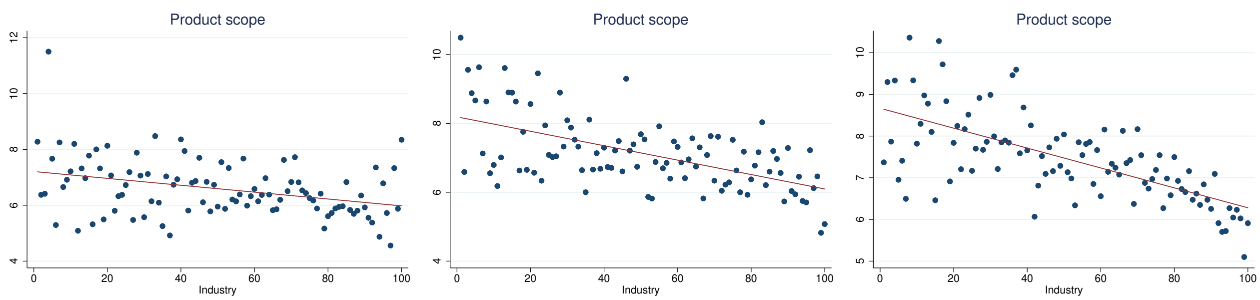
year 2000 year 2003 year 2006

Figure C1: Export propensity by industry



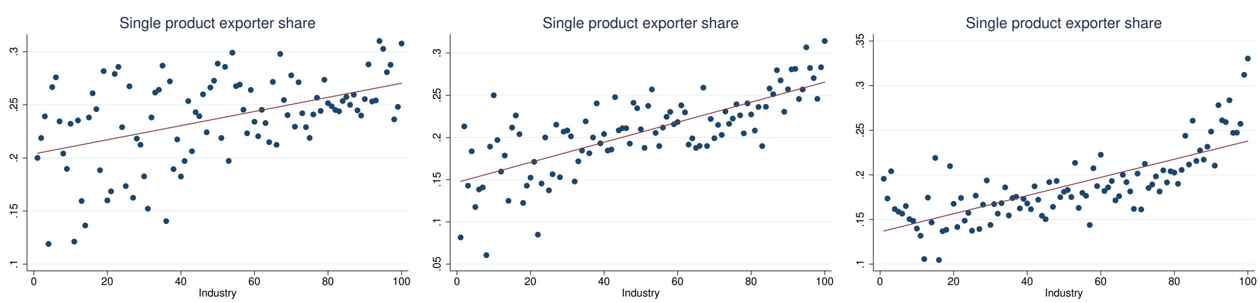
year 2000 year 2003 year 2006

Figure C2: Export intensity by industry



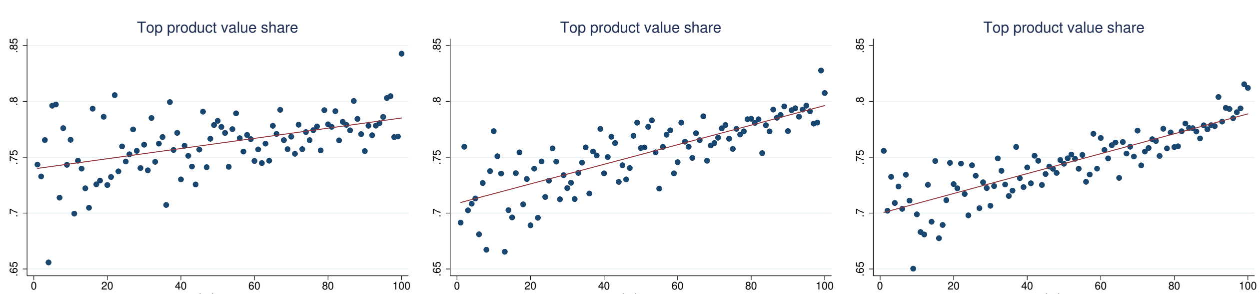
year 2000 year 2003 year 2006

Figure C3: Number of products exported



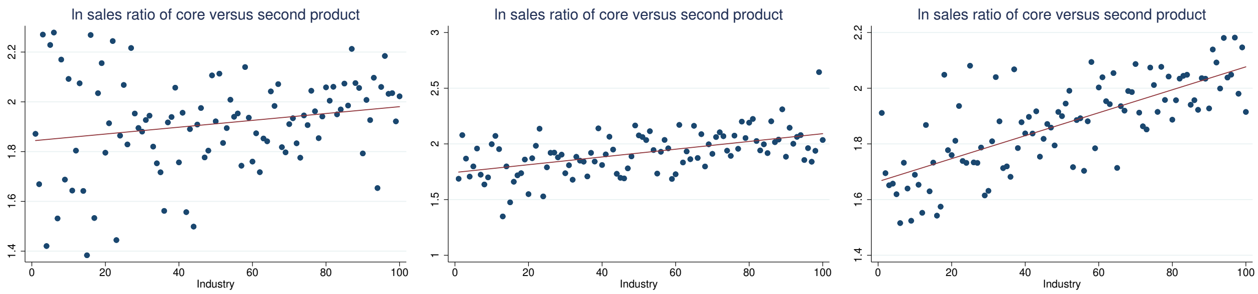
year 2000 year 2003 year 2006

Figure C4: Share of single product exporters



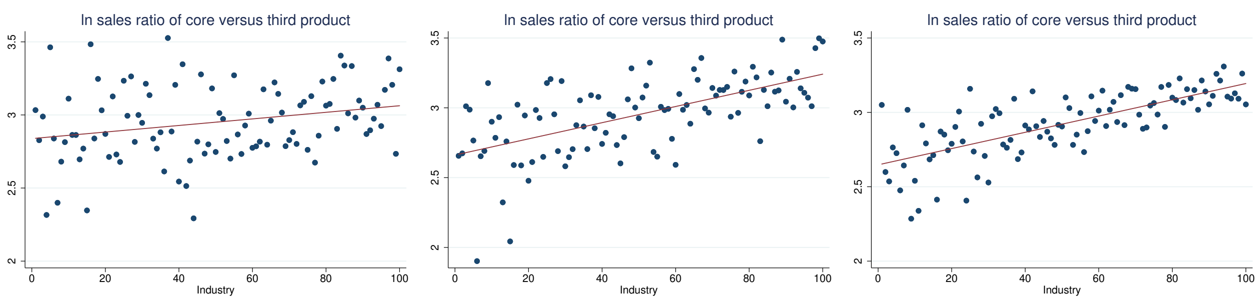
year 2000 year 2003 year 2006

Figure C5: Average value share of the core product for exporters



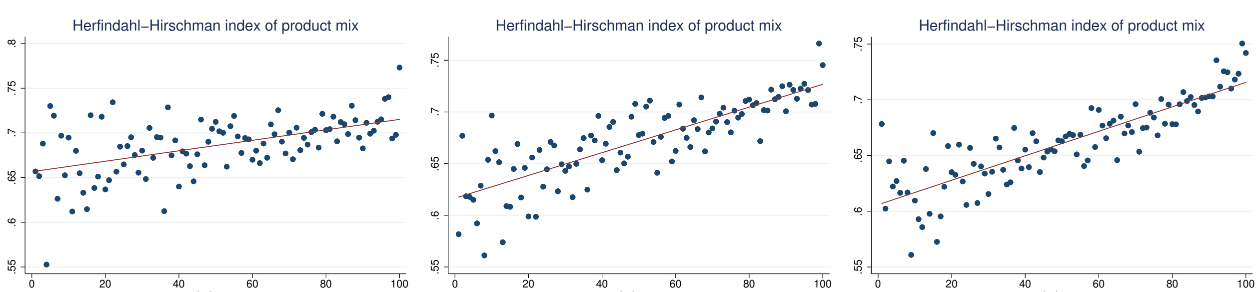
year 2000 year 2003 year 2006

Figure C6: Exports of the core product relative to the second best product



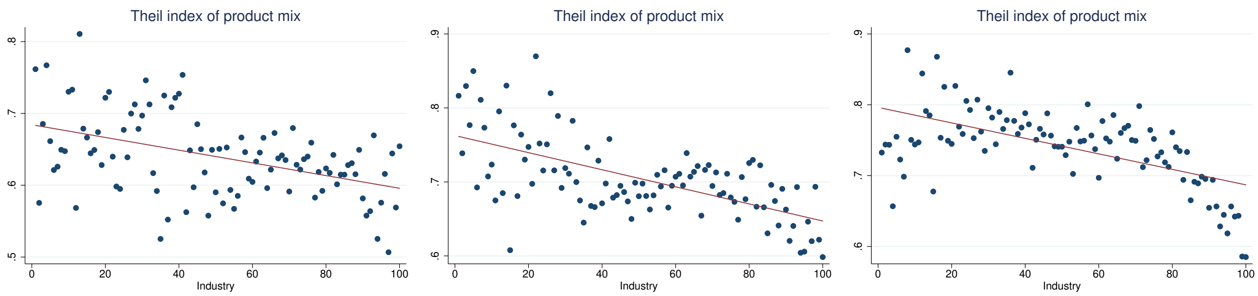
year 2000 year 2003 year 2006

Figure C7: Exports of the core product relative to the third best product



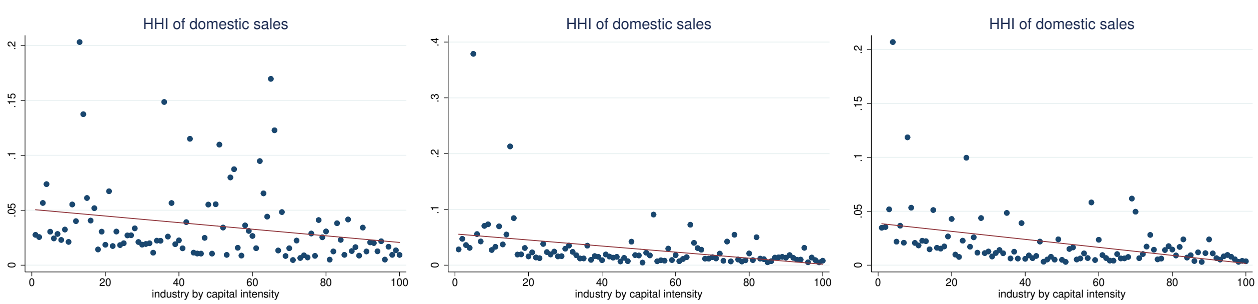
year 2000 year 2003 year 2006

Figure C8: Average Herfindhal index of exports across products



year 2000 year 2003 year 2006

Figure C9: Average Theil index of exports across products



year 2000 year 2003 year 2006

Figure C10: Herfidhal index of domestic sales across firms

10 Proofs

10.1 Proof of Proposition 1

The export propensity from the home country to the foreign in industry z is

$$\chi(z) = (\frac{C_X(z)}{C_D(z)})^k,$$

where $C_X(z)$ is the cut-off cost of export which satisfies $\tau C_X(z) = C_D^*(z)$. So we have

$$\chi(z) = \rho(\frac{C_D^*(z)}{C_D(z)})^k.$$

Given that $\frac{\partial C_D(z)}{\partial z} > 0$ and $\frac{\partial C_D^*(z)}{\partial z} < 0$, it is easy to see that $\frac{\partial \chi(z)}{\partial z} < 0$. Similarly, we can prove that $\frac{\partial \chi^*(z)}{\partial z} > 0$.

The model predicts that exports from the home country to the foreign in industry z is

$$EXP(z) = \frac{1}{2\gamma(k+2)C_M(z)^k} N_E(z) C_D^*(z)^{k+2} L^* \rho.$$

On the other hand, the sales of industry z at home is

$$S(z) = \frac{1}{2\gamma(k+2)C_M(z)^k} N_E(z) C_D(z)^{k+2} L.$$

The export intensity of industry z is thus given by

$$\begin{aligned} \lambda(z) &\equiv \frac{EXP(z)}{EXP(z)+S(z)} \\ &= \frac{L^*\rho}{L^*\rho + L(\frac{C_D(z)}{C_D^*(z)})^{k+2}} \\ &= \frac{L^*\rho}{L^*\rho + L\rho^{\frac{k+2}{k}}\chi(z)^{-\frac{k+2}{k}}}. \end{aligned}$$

Since we have $\frac{\partial \chi(z)}{\partial z} < 0$, it is easy to see that $\frac{\partial \lambda(z)}{\partial \chi(z)} > 0$, thus $\frac{\partial \lambda(z)}{\partial z} < 0$. Similar results hold for the foreign.

10.2 Proof of Proposition 2

The export product scope is given by $M_x(z, c)$ for a firm with core competency c in industry z. For firms that do export, i.e., the marginal cost of their core competency satisfies $c \leq C_X(z)$.

Then $M_x(z,c) = \max\{m|v(m,c) \leq \frac{C_D^*(z)}{\tau}\} + 1$. Since $v(m,c) = \varpi^{-m}c$ and $\varpi \in (0,1)$, we have

$$M_x(z,c) = \max\{m| \ln \tau + \ln c + m \ln(\frac{1}{\varpi}) \leq \ln C_D^*(z)\} + 1.$$

Since $\frac{\partial C_D^*(z)}{\partial z} < 0$, for two industries $z\prime > z$, we have $C_D^*(z') < C_D^*(z)$, we should have

$$M_x(z',c) \leq M_x(z,c).$$

That is the export product scope is non-increasing with comparative disadvantage.

10.3 Proof of Proposition 3

The sales ratio of product m and m' for an exporter to the foreign country in industry z can be written as

$$\frac{r(z,v(m,c))}{r(z,v(m',c))} = \frac{C_D^*(z)^2 - (\tau\varpi^{-m}c)^2}{C_D^*(z)^2 - (\tau\varpi^{-m'}c)^2}.$$

Suppose m'$>$ m, so product m is closer to core: $\tau\varpi^{-m}c < \tau\varpi^{-m'}c$. Since $\frac{\partial C_D^*(z)}{\partial z} < 0$, it can be shown that

$$\frac{\partial \frac{r(z,v(m,c))}{r(z,v(m',c))}}{\partial z} > 0.$$

Export sales from the home country to the foreign country therefore become more skewed in the more comparative disadvantage industries.

10.4 Proof of Proposition 4

Consider two single-product firms in industry z such that $c_1 < c_2$, the ratio of their sales in the domestic market is given by

$$\frac{r_d(z,c_1)}{r_d(z,c_2)} = \frac{C_D^2(z) - c_1^2}{C_D^2(z) - c_2^2}.$$

Taking partial derivatives of the sales ratio with respect to $C_D(z)$, we have

$$\frac{\partial \frac{r_d(z,c_1)}{r_d(z,c_2)}}{\partial C_D(z)} = 2C_D(z)\frac{c_1^2 - c_2^2}{(C_D^2(z) - c_2^2)^2} < 0.$$

Tougher competition therefore skews more sales toward the better performing firm.

The multi-product firm case is less straightforward. Consider two firms with $c_1 < c_2$, their sales ratio in the domestic market is given by

$$\frac{R_d(z,c_1)}{R_d(z,c_2)} = \frac{\sum\limits_{m=0}^{M_1-1} r_d(z, v(m,c_1))}{\sum\limits_{m=0}^{M_2-1} r_d(z, v(m,c_2))}$$
$$= \frac{C_D^2(z)M_1 - c_1^2 \frac{w^2}{w^{2M_1}} \frac{1-w^{2M_1}}{1-w^2}}{C_D^2(z)M_2 - c_2^2 \frac{w^2}{w^{2M_2}} \frac{1-w^{2M_2}}{1-w^2}},$$

where M_1 and M_2 are the product scope of the two firms, respectively. Since $c_1 < c_2$, we have $M_2 \le M_1$. If $M_1 = M_2$, we have $\frac{\partial \frac{R_d(z,c_1)}{R_d(z,c_2)}}{\partial C_D(z)} = \frac{2C_D(z)M_1w^2}{w^{2M_1}} \frac{1-w^{2M_1}}{1-w^2} \frac{c_1^2-c_2^2}{(C_D^2(z)-c_2^2)^2} < 0$. If $M_1 > M_2$, we cannot sign the partial derivative. However, we claim that if the intra-firm reallocation is dominated by inter-firm reallocation, our result is still true. To see that, we first note that

$$\frac{R_d(z,c_1)}{R_d(z,c_2)} = \frac{r_0 + r_1 + r_2 + ... + r_{M_1-1}}{R_d(z,c_2)},$$

where $r_0 = r_d(z, v(0,c_1)), r_1 = r_d(z, v(1,c_1)), ..., r_{M_1-1} = r_d(z, v(M_1-1, c_1))$. It can be further rearranged as

$$\frac{R_d(z,c_1)}{R_d(z,c_2)} = \frac{\sum\limits_{i=0}^{M_2-1} r_i}{R_d(z,c_2)} + \frac{R_d(z,c_1) - \sum\limits_{i=0}^{M_2-1} r_i}{R_d(z,c_1)} \frac{R_d(z,c_1)}{R_d(z,c_2)}.$$

If we move the second term of the right hand side to the left, we have

$$(1 - \frac{R_d(z,c_1) - \sum\limits_{i=0}^{M_2-1} r_i}{R_d(z,c_1)}) \frac{R_d(z,c_1)}{R_d(z,c_2)} = \frac{\sum\limits_{i=0}^{M_2-1} r_i}{R_d(z,c_2)},$$

or

$$\frac{R_d(z,c_1)}{R_d(z,c_2)} \frac{\sum\limits_{i=0}^{M_2-1} r_i}{R_d(z,c_1)} = \frac{\sum\limits_{i=0}^{M_2-1} r_i}{R_d(z,c_2)}. \tag{E1}$$

Now we make two claims. First, the term on the right hand side of Equation (E1), which captures inter-firm reallocations, decreases with $C_D(z)$. This term looks at the ratio of total sales for the first M_2 products between the two firms, which is

$$\frac{\sum\limits_{i=0}^{M_2-1} r_i}{R_d(z,c_2)} = \frac{C_D^2(z)M_2 - c_1^2 \frac{w^2}{w^{2M_2}} \frac{1-w^{2M_2}}{1-w^2}}{C_D^2(z)M_2 - c_2^2 \frac{w^2}{w^{2M_2}} \frac{1-w^{2M_2}}{1-w^2}}.$$

Therefore, we have

$$\frac{\partial \frac{\sum\limits_{i=0}^{M_2-1} r_i}{R_d(z,c_2)}}{\partial C_D(z)} = 2C_D(z)M_2 \frac{w^2}{w^{2M_2}} \frac{1-w^{2M_2}}{1-w^2} \frac{c_1^2 - c_2^2}{(C_D^2(z)M_2 - c_2^2 \frac{w^2}{w^{2M_2}} \frac{1-w^{2M_2}}{1-w^2})^2} < 0.$$

Second, the intra-firm reallocation component $\frac{\sum\limits_{i=0}^{M_2-1} r_i}{R_d(z,c_1)}$ decreases with $C_D(z)$. To show that is the case, first we note that for any product i, $0 \le i \le M_1 - 1$, its share in the firms' total sales in the home market is

$$\frac{r_i}{R_d(z,c_1)} = \frac{C_D^2(z) - (c_1 w^{-i})^2}{C_D^2(z)M_1 - c_1^2 \frac{w^2}{w^{2M_1}} \frac{1-w^{2M_1}}{1-w^2}}.$$

Therefore, we have

$$\frac{\partial \frac{r_i}{R_d(z,c_2)}}{\partial C_D(z)} = \frac{2C_D c_1^2 (M_1 w^{-2i} - \frac{w^2}{w^{2M_1}} \frac{1-w^{2M_1}}{1-w^2})}{(C_D^2(z)M_1 - c_1^2 \frac{w^2}{w^{2M_1}} \frac{1-w^{2M_1}}{1-w^2})^2}. \tag{E2}$$

For $i = 0$, we have $M_1 w^{-2i} = M_1 < \frac{w^2}{w^{2M_1}} \frac{1-w^{2M_1}}{1-w^2} = \sum_{i=0}^{M_1-1} w^{-2i}$ given that $0 < w < 1$. Therefore, $\frac{\partial \frac{r_0}{R_d(z,c_1)}}{\partial C_D(z)} < 0$, which means the share of the core product must always increase when competition intensifies. For $i = M_1 - 1$, we have $M_1 w^{-2i} = M_1 w^{-2(M_1-1)} > \frac{w^2}{w^{2M_1}} \frac{1-w^{2M_1}}{1-w^2}$ since it is equivalent to $M_1 > \frac{1-w^{2M_1}}{1-w^2} = \sum_{i=0}^{M_1-1} w^{2i}$. Thus we have $\frac{\partial \frac{r_{M_1-1}}{R_d(z,c_1)}}{\partial C_D(z)} > 0$, which means that the share of the worst product must always decline when competition intensifies. Since $M_1 w^{-2i}$ is decreasing with i, there exists a product m^* such that for $i \le m^*$, we have $\frac{\partial \frac{r_i}{R_d(z,c_1)}}{\partial C_D(z)} \le 0$, and for $i \ge m^*$, we have $\frac{\partial \frac{r_i}{R_d(z,c_1)}}{\partial C_D(z)} \ge 0$. Consequently, when $C_D(z)$ increases, the market becomes less competitive and the share of the total sales of the firm's first M_2 products declines.

Given these two claims, going back to Equation (E1), if we let $f(C_D(z)) = \frac{R_d(z,c_1)}{R_d(z,c_2)}$, $h(C_D(z)) = \frac{\sum\limits_{i=0}^{M_2-1} r_i}{R_d(z,c_1)}$, and $g(C_D(z)) = \frac{\sum\limits_{i=0}^{M_2-1} r_i}{R_d(z,c_2)}$, we have

$$f(C_D(z))h(C_D(z)) = g(C_D(z)).$$

If we take the partial derivatives with respect to $C_D(z)$ for the equation above, we have

$$\frac{f'}{f} = \frac{g'}{g} - \frac{h'}{h},$$

where $f' = \partial f(C_D)/\partial C_D$, $h' = \partial h(C_D)/\partial C_D$, and $g' = \partial g(C_D)/\partial C_D$. Given the two claims that we have made above, we have $g' < 0$ and $h' < 0$. The sign of f' is therefore undetermined. It is negative if $\frac{g'}{g} < \frac{h'}{h}$, which means that inter-firm reallocations (captured by $\frac{g'}{g}$) dominates intra-firm reallocations (captured by $\frac{h'}{h}$). In the case of single-product firms, there is no intra-firm reallocation, therefore, this condition holds naturally.

10.5 Proof of Proposition 5

Comparing the relative average marginal costs between the home country and the foreign country under autarky and open economy, we have:

$$\frac{C_D(z)^T}{C_D(z)^{*T}} = \underbrace{\frac{C_D(z)^A}{C_D(z)^{*A}}}_{ex\ ante} \underbrace{\left[\frac{1-\rho\frac{C_M^*(z)^k}{C_M(z)^k}}{1-\rho^*\frac{C_M(z)^k}{C_M^*(z)^k}}\right]^{\frac{1}{k+2}}}_{amplifying}. \tag{E3}$$

where the first term is the *ex ante* comparative advantage and the second term is only present when countries open to trade. It is easy to verify that the second term increases with capital intensity z.

Depending on the relative size of $C_M(z)$ and $C_M(z)^*$ and the trade freeness, the relationship between $\frac{C_D(z)^T}{C_D(z)^{*T}}$ and $\frac{C_D(z)^A}{C_D(z)^{*A}}$ is illustrated by Figure C11. Panel (a) is when $\rho^* C_M(z)^{2k}$ is always larger than $\rho C_M^*(z)^{2k}$ so that $\frac{1-\rho\frac{C_M^*(z)^k}{C_M(z)^k}}{1-\rho^*\frac{C_M(z)^k}{C_M^*(z)^k}} > 1$, *vice versa* for panel (c). Panel (b) is when there exists an industry such that $\rho^* C_M(z)^{2k} = \rho C_M^*(z)^{2k}$. In all 3 cases, the differences in the relative average marginal costs across industries enlarge under the trade equilibrium. Hence comparative advantage is amplified by the second component.

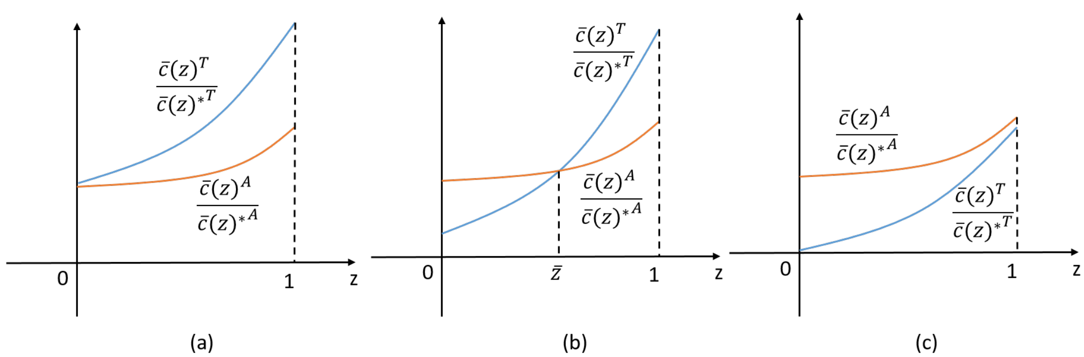


Figure C11: Relative average marginal costs: autarky vs. trade

10.6 Proof of Proposition 6

The relative quantity-based TFP between the home country and the foreign country under open economy can be rewritten as

$$
\begin{aligned}
\frac{\overline{\Phi}(z)^T}{\overline{\Phi}^*(z)^T} &= \left(\frac{L}{L^*}\frac{C_M^*(z)^k}{C_M(z)^k}\right)^{\frac{1}{k+2}}\left(\frac{L}{L^*}\frac{C_M^*(z)^k}{C_M(z)^k}\right)^{\frac{k+1}{k+2}}\frac{L^*}{L}\frac{\frac{L}{L^*}(\frac{C_D(z)}{C_D^*(z)})^{k+1}+\rho}{1+\rho^*\frac{L}{L^*}(\frac{C_D(z)}{C_D^*(z)})^{k+1}} \\
&= \underbrace{\frac{\overline{\Phi}(z)^A}{\overline{\Phi}(z)^{*A}}}_{ex\ ante}\underbrace{\left(\frac{\overline{\Phi}(z)^A}{\overline{\Phi}(z)^{*A}}\right)^{k+1}}_{amplifying}\underbrace{\frac{L^*}{L}\frac{\frac{L}{L^*}(\frac{C_D(z)}{C_D^*(z)})^{k+1}+\rho}{1+\rho^*\frac{L}{L^*}(\frac{C_D(z)}{C_D^*(z)})^{k+1}}}_{dampening}.
\end{aligned}
\tag{E4}
$$

Given that $k > 0$, it is obvious that the second component is amplifying the effect of the first component, the ex ante comparative advantage measured by relative TFPQ under autarky. For the third component, if we define as $f(z) \equiv \frac{L^*}{L}\frac{\frac{L}{L^*}(\frac{C_D(z)}{C_D^*(z)})^{k+1}+\rho}{1+\rho^*\frac{L}{L^*}(\frac{C_D(z)}{C_D^*(z)})^{k+1}}$, we have

$$
\frac{\partial f(z)}{\partial z} = \frac{(1-\rho\rho^*)(k+1)(\frac{C_D(z)}{C_D^*(z)})^k}{(1+\rho^*\frac{L}{L^*}(\frac{C_D(z)}{C_D^*(z)})^{k+1})^2}\frac{\partial(\frac{C_D(z)}{C_D^*(z)})}{\partial z} > 0.
$$

Given our assumptions that $\frac{\partial C_M(z)}{\partial z} > 0$ and $\frac{\partial C_M^*(z)}{\partial z} < 0$, we have $\frac{\partial(\frac{\overline{\Phi}(z)^A}{\overline{\Phi}(z)^{*A}})}{\partial z} < 0$. That is to say the third component is negatively correlated with the first two components. Hence, it is dampening the ex ante comparative advantage.

10.7 Proof of Proposition 7

According to Proposition 1, the export intensity is

$$
\begin{aligned}
\lambda(z) &= \frac{L^*\rho}{L^*\rho + L\rho^{\frac{k+2}{k}}\chi(z)^{-\frac{k+2}{k}}} \\
&= \frac{1}{1+\frac{L}{L^*}\rho^{\frac{2}{k}}\chi(z)^{-\frac{k+2}{k}}}.
\end{aligned}
$$

As a result, we can infer the relative market size $\frac{L}{L^*}$ as

$$
\frac{L}{L^*} = \frac{1-\lambda(z)}{\lambda(z)}\frac{\chi(z)^{\frac{k+2}{k}}}{\rho^{\frac{2}{k}}}.
\tag{E5}
$$

Again, according to Proposition 1, the export propensity in each industry is given by

$$\begin{aligned}\chi(z) &= \rho(\frac{C_D^*(z)}{C_D(z)})^k \\ &= \rho(\frac{L}{L^*}\frac{C_M^*(z)^k - \rho^* C_M(z)^k}{C_M(z)^k - \rho C_M^*(z)^k})^{\frac{k}{k+2}}.\end{aligned} \tag{E6}$$

Immediately, ratio of average costs between the home country and the foreign is given by

$$\frac{C_D^*(z)}{C_D(z)} = (\frac{\chi(z)}{\rho})^{1/k}, \tag{E7}$$

Moreover the relative cost upper bounds can be solved out of Equation (E6) as

$$\frac{C_M^*(z)^k}{C_M(z)^k} = \frac{\rho^* + \frac{L^*}{L}(\frac{\chi_z}{\rho})^{\frac{k+2}{k}}}{1 + \frac{L^*}{L}\rho^{-\frac{2}{k}}\chi(z)^{\frac{k+2}{k}}}, \tag{E8}$$

substituting the relative size of $\frac{L}{L^*}$ using Equation (E5), it can be written as a function of the observables. Then the endogenous component of the comparative advantage given by $\frac{1-\rho\frac{C_M^{*}{}^k}{C_M{}^k}}{1-\rho^*\frac{C_M{}^k}{C_M^{*}{}^k}}$ is also known. Finally, the *ex ante* comparative advantage $\frac{C_D(z)^{*A}}{C_D(z)^A} = (\frac{L}{L^*}\frac{C_M^*(z)^k}{C_M(z)^k})^{1/(k+2)}$ can also be inferred.

The *ex ante* component of comparative advantage are the same for the two measures of comparative advantage since

$$\begin{aligned}\frac{\overline{\Phi}(z)^A}{\overline{\Phi}(z)^{*A}} &= (\frac{L}{L^*}\frac{C_M^*(z)^k}{C_M(z)^k})^{1/(k+2)} \\ &= \frac{C_D(z)^{*A}}{C_D(z)^A}.\end{aligned} \tag{E9}$$

The way to quantify $\frac{\overline{\Phi}(z)^A}{\overline{\Phi}(z)^{*A}}$ therefore is the same as quantifying $\frac{C_D(z)^{*A}}{C_D(z)^A}$. Then the amplifying component $(\frac{\overline{\Phi}(z)^A}{\overline{\Phi}(z)^{*A}})^{k+1}$ is also known.

Finally, the dampening component is given by

$$\begin{aligned}\frac{L^*}{L}\frac{\frac{L}{L^*}(\frac{C_D(z)}{C_D^*(z)})^{k+1} + \rho}{1 + \rho^*\frac{L}{L^*}(\frac{C_D(z)}{C_D^*(z)})^{k+1}} &= \frac{(\frac{\chi(z)}{\rho})^{-\frac{k+1}{k}} + \rho^{1+\frac{2}{k}}\frac{\lambda(z)}{1-\lambda(z)}\chi(z)^{-\frac{k+2}{k}}}{1 + \rho^*\frac{1-\lambda(z)}{\lambda(z)}\frac{\chi(z)^{\frac{k+2}{k}}}{\rho^{\frac{2}{k}}}(\frac{\chi(z)}{\rho})^{-\frac{k+1}{k}}} \\ &= (\frac{\chi(z)}{\rho})^{-\frac{k+1}{k}}\frac{1 + \rho^{\frac{1}{k}}\frac{\lambda(z)}{1-\lambda(z)}\chi(z)^{-\frac{1}{k}}}{1 + \rho^*\frac{1-\lambda(z)}{\lambda(z)}\chi(z)^{\frac{1}{k}}\rho^{\frac{k-1}{k}}},\end{aligned}$$

which can also be inferred as long as we know $\{\rho, \rho^*, k\ \}$ and observe $\{\chi(z), \lambda(z)\ \}$.

11 Complementary Theoretical Results

11.1 A Model with Constant Mark-ups

This appendix section shows how to decompose comparative advantage in the constant mark-up heterogeneous firm model *à la* Bernard, Redding, and Schott (2007). Suppose the demand is given by the following quasi-CES preference[34]

$$U = q_0^c - \gamma \int_0^1 \ln Q(z) dz.$$

Under such a preference, solving the consumer's problem we have

$$q_i^c = -\gamma \frac{p_i^{-\sigma}}{P(z)^{1-\sigma}},$$

where $P(z) = (\int_{i \in \Omega(z)} p_i(z)^{1-\sigma} di)^{\frac{1}{1-\sigma}}$ is the price index, and $P(z)Q(z) = \int p_i(z) q_i^c(z) di = -\gamma$.

For the supply side, we follow the standard Melitz (2003) set up in the case of open economy: the entry cost is f_E and fixed cost of serving the domestic market and foreign market is f_d and f_x, respectively. On top of that, we assume that firms draw their marginal costs from the Pareto distribution $G(z, c) = (\frac{c}{C_M(z)})^k$, where $C_M(z)$ is the upper bound of the marginal cost at home. Given the market demand faced by firm at home and foreign and the iceberg cost assumption, we have

$$\begin{aligned} q_i(z) &= -\gamma L \frac{p_i^{-\sigma}}{P(z)^{1-\sigma}}, \\ q_i^*(z) &= -\gamma L^* \frac{p_i^{-\sigma}}{P^*(z)^{1-\sigma}}, \end{aligned}$$

and the optimal pricing for each market is given by

$$\begin{aligned} p_d(z, c) &= \frac{\sigma}{\sigma - 1} c, \\ p_x(z, c) &= \frac{\sigma}{\sigma - 1} \tau c. \end{aligned}$$

[34] We get rid of the income effect to simplify the algebra. Huang et al. (2017) include the income effect and arrive at similar results.

Then firm's profit functions for each market are given by

$$\pi_d(z,c) = \frac{r_d(z,c)}{\sigma} - f_d = \frac{-\gamma L}{\sigma}(\frac{p_d(z,c)}{P(z)})^{1-\sigma} - f_d,$$
$$\pi_x(z,c) = \frac{r_x(z,c)}{\sigma} - f_d = \frac{-\gamma L^*}{\sigma}(\frac{p_x(z,c)}{P^*(z)})^{1-\sigma} - f_x.$$

The zero-profit conditions are

$$\frac{-\gamma L}{\sigma}(\frac{\frac{\sigma}{\sigma-1}c_D(z)}{P(z)})^{1-\sigma} = f_d,$$
$$\frac{-\gamma L^*}{\sigma}(\frac{\frac{\sigma}{\sigma-1}\tau c_X(z)}{P^*(z)})^{1-\sigma} = f_x,$$

where $c_D(z)$ and $c_X(z)$ are the cost cut-offs. Dividing the two equations above, we have

$$\frac{c_X(z)}{c_D(z)} = \frac{P^*(z)}{\tau P(z)}(\frac{f_d L^*}{f_x L})^{\frac{1}{\sigma-1}}. \tag{E10}$$

To determine how $\frac{c_X(z)}{c_D(z)}$ varies across industries, we need to know how $\frac{P^*(z)}{P(z)}$ varies with z. To do that, we follow Bernard et al. (2007) to consider two extreme cases: free trade and autarky. Then the costly trade case would then fall between.

In the case of free trade, every surviving firm from every country exports. The number of varieties and the price charged by each firm in each market is the same. As a result, the price indexes satisfy $P(z) = (\int_{i\in\Omega(z)} p_i(z)^{1-\sigma}di)^{\frac{1}{1-\sigma}} = P(z)^* = (\int_{i\in\Omega^*(z)} p_i(z)^{1-\sigma}di)^{\frac{1}{1-\sigma}}$ under free trade. Moreover, the relative price index $\frac{P^*(z)}{P(z)}$ is constant.

Under autarky, $P(z) = (\int_{i\in\Omega(z)} p_i(z)^{1-\sigma}di)^{\frac{1}{1-\sigma}} = M_z^{\frac{1}{1-\sigma}} p_d(\widehat{c}_d(z))$ where M_z is the domestic firm mass, and $\widehat{c}_d(z)^{-1} = (\frac{1}{G(c_D(z))}\int_0^{c_D(z)} c^{1-\sigma}g(c)dc)^{\frac{1}{\sigma-1}}$ is the average marginal cost. Similarly, for the foreign country, we have $P(z)^* = M_z^{*\frac{1}{1-\sigma}} p_d^*(\widehat{c}_d^*(z))$. For the firm mass, using the market clearing condition, we have $M_z = \frac{P(z)Q(z)}{r(\widehat{c}_d(z))} = \frac{-\gamma}{r(\widehat{c}_d(z))}$. Given the CES demand, we have $\frac{r(\widehat{c}_d(z))}{r(c_D(z))} = (\frac{c_D(z)}{\widehat{c}_d(z)})^{\sigma-1}$. Combining this result with the zero profit condition, we have $r(\widehat{c}_d(z)) = r(c_D(z))(\frac{c_D(z)}{\widehat{c}_d(z)})^{\sigma-1} = \sigma f_d(\frac{c_D(z)}{\widehat{c}_d(z)})^{\sigma-1}$, which implies that the firm mass is

$$M_z = \frac{-\gamma}{\sigma f_d}(\frac{\widehat{c}_d(z)}{c_D(z)})^{\sigma-1}.$$

So the autarky price index in home country is given by

$$P(z) = [\frac{-\gamma}{\sigma f_d}(\frac{\widehat{c_d}(z)}{c_D(z)})^{\sigma-1}]^{\frac{1}{1-\sigma}} \frac{\sigma}{\sigma-1}\widehat{c_d}(z).$$

If we impose the Pareto distribution assumption, we have $\frac{\widehat{c_d}(z)}{c_D(z)} = (\frac{k-\sigma+1}{k})^{\frac{1}{\sigma-1}}$. Then the price index is

$$P(z) = [\frac{-\gamma}{\sigma f_d}\frac{k-\sigma+1}{k}]^{\frac{1}{1-\sigma}} \frac{\sigma}{\sigma-1}(\frac{k-\sigma+1}{k})^{\frac{1}{\sigma-1}} c_D(z),$$

which varies one-to-one with $c_D(z)$. To determine $c_D(z)$, we use the free entry condition under autarky which says the probability of survival times the expected profit equals to the fixed cost of entry:

$$G(c_D(z))\pi(\widehat{c_d}(z)) = f_e,$$

where $G(c_D(z)) = (\frac{c_D(z)}{C_M(z)})^k$. Since $\pi(\widehat{c_d}(z)) = \frac{r(\widehat{c_d}(z))}{\sigma} = \frac{r(c_D(z))}{\sigma}(\frac{c_D(z)}{\widehat{c_d}(z)})^{\sigma-1} = f_d(\frac{k}{k-\sigma+1})^{\frac{1}{\sigma-1}}$, it is easy to find that

$$c_D(z) = (\frac{f_e}{f_d}\frac{k-\sigma+1}{k})^{1/k} C_M(z),$$

which varies one-to-one with the cost upper bound. Therefore, under autarky, we have

$$\frac{P^*(z)}{P(z)} = \frac{C_M^*(z)}{C_M(z)},$$

which declines with z given the assumption that $\frac{\partial C_M(z)}{\partial z} > 0$ and $\frac{\partial C_M^*(z)}{\partial z} < 0$. That is to say, if we have $z' > z$, then we have

$$\begin{aligned} \frac{c_X(z)}{c_D(z)} &= \frac{P^*(z)}{\tau P(z)}(\frac{f_d L^*}{f_x L})^{\frac{1}{\sigma-1}} \\ &= \frac{c_X(z')}{c_D(z')}, \end{aligned}$$

under free trade, and

$$\frac{c_X(z)}{c_D(z)} > \frac{c_X(z')}{c_D(z')},$$

under autarky. Given the continuity of trade costs, it must be the case that under costly trade, we have

$$\frac{\partial \chi(z)}{\partial z} < 0,$$

where $\chi(z) = (\frac{c_X(z)}{c_D(z)})^k$ is the probability of export. Similarly, we can prove that $\frac{\partial \chi(z)^*}{\partial z} > 0$ holds for foreign.

Combining the zero profit condition and free entry condition under costly trade, we have

$$f_d \int_0^{C_D(z)} (\frac{\pi_x(z,c)}{\pi_x(z, C_D(z))} - 1) dG(z,c) + f_x \int_0^{C_X(z)} (\frac{\pi_x(z,c)}{\pi_x(z, C_X(z))} - 1) dG(z,c) = f_e,$$

for the home country. It can be simplified as

$$f_d C_D(z)^k + f_x C_X(z)^k = \frac{k-\sigma+1}{\sigma-1} f_e C_M(z)^k.$$

Similarly, for the foreign country, we have

$$f_d C_D^*(z)^k + f_x C_X^*(z)^k = \frac{k-\sigma+1}{\sigma-1} f_E C_M^*(z)^k.$$

These two equations imply

$$\frac{f_D C_D^*(z)^k + f_X C_X^*(z)^k}{f_D C_D(z)^k + f_X C_X(z)^k} = \frac{C_M^*(z)^k}{C_M(z)^k},$$

or

$$(\frac{C_D^*(z)}{C_D(z)})^k = \underbrace{\frac{C_M^*(z)^k}{C_M(z)^k}}_{exogenous} \underbrace{\frac{1 + \frac{f_X}{f_D}\chi(z)}{1 + \frac{f_X}{f_D}\chi(z)^*}}_{endogenous},$$

where the exogenous and endogenous components are positively correlated.

11.2 Welfare in the Homogeneous and Hetergeneous Firm Models

11.2.1 Welfare in the Heterogeneous firm model

Substituting the demand function and consumers' budget constraint into the utility function, we have

$$U = y_0^c + I + \int_0^1 \left[\frac{\gamma}{2} \int_{i \in \Omega(z)} (q_i^c(z))^2 di + \frac{\eta}{2} Q^c(z)^2 \right] dz. \tag{E11}$$

If we define average price of industry z as $\overline{p}(z) = \frac{1}{N(z)} \int_{i \in \Omega(z)} p_i^c(z) di$, and the variance of price within each industry $\sigma_p^2(z) = \frac{1}{N(z)} \int_{i \in \Omega(z)} (p_i^c(z) - \overline{p}(z)) di$, we have

$$U = y_0^c + I + \frac{1}{2} \int_0^1 \left[(\eta + \frac{\gamma}{N(z)})^{-1} (\alpha - \overline{p}(z))^2 + \frac{N(z)}{\gamma} \sigma_p^2(z) \right] dz.$$

If firm productivities are Pareto distributed, we have

$$U^{het} = y_0^c + I + \frac{1}{2\eta} \int_0^1 \left[(\alpha - C_D(z))(\alpha - \frac{k+1}{k+2} C_D(z)) \right] dz. \tag{E12}$$

11.2.2 Welfare in the homogeneous firm model

If firms are homogeneous, their profits are all given by

$$\pi_i(z) = \frac{L}{4\gamma} (p_{\max}(z) - c(z))^2.$$

Due to free entry, firms earn zero profit, and we have

$$\frac{L}{4\gamma} (p_{\max}(z) - c(z))^2 - f_E = 0,$$

which implies that the choke price is given by

$$p_{\max}(z) = \sqrt{\frac{4\gamma f_E}{L}} + c(z).$$

Then immediately, we have

$$\begin{aligned} q(z) &= \frac{L}{2\gamma} (p_{\max}(z) - c(z)) \\ &= \sqrt{\frac{f_E L}{\gamma}}. \end{aligned}$$

Therefore, the demand by each consumer is $q^c(z) = \frac{q(z)}{L} = \sqrt{\frac{f_E}{\gamma L}}$. Given the demand function, the choke price can be rewritten as

$$\begin{aligned} p_{\max}(z) &= \alpha - \eta Q^c(z) \\ &= \alpha - \eta q^c(z) N(z), \end{aligned}$$

which implies that the number of varieties is given by

$$
\begin{aligned}
N(z) &= \frac{\alpha - p_{\max}(z)}{\eta q^c(z)} \\
&= \frac{(\alpha - c(z))\sqrt{\frac{\gamma L}{f_E}} - 2\gamma}{\eta},
\end{aligned}
$$

and the overall consumption of the differentiated varieties is

$$
\begin{aligned}
Q^c(z) &= \frac{\alpha - p_{\max}(z)}{\eta} \\
&= \frac{\alpha - c(z) - \sqrt{\frac{4\gamma f_E}{L}}}{\eta}.
\end{aligned}
$$

Then using Equation (E11), we know that the welfare:

$$
\begin{aligned}
U^{\text{hom}} &= y_0^c + I + \int_0^1 \left[\frac{\gamma}{2} \int_{i \in \Omega(z)} (q_i^c(z))^2 di + \frac{\eta}{2} Q^c(z)^2 \right] dz \\
&= y_0^c + I + \int_0^1 \left[\frac{N(z) f_E}{2L} + \frac{\eta}{2} \Big(\frac{\alpha - c(z) - \sqrt{\frac{4\gamma f_E}{L}}}{\eta} \Big)^2 \right] dz \\
&= y_0^c + I + \frac{1}{2\eta} \int_0^1 \left[((\alpha - c(z))\sqrt{\frac{\gamma f_E}{L}} - 2\frac{\gamma f_E}{L}) + (\alpha - c(z) - \sqrt{\frac{4\gamma f_E}{L}})^2 \right] dz.
\end{aligned}
$$

To ensure that the welfare are the same under autarky for the models of homogeneous and heterogeneous firms, i.e., $U^{\text{hom}} = U^{het}$, we can let

$$
(\alpha - C_D(z)^A)(\alpha - \frac{k+1}{k+2} C_D(z)^A) = (\alpha - c(z))\sqrt{\frac{\gamma f_E}{L}} - 2\frac{\gamma f_E}{L} + (\alpha - c(z) - \sqrt{\frac{4\gamma f_E}{L}})^2. \quad \text{(E13)}
$$

This is a sufficient condition for $U^{\text{hom}} = U^{het}$. Let $\widetilde{U}(z)^A = (\alpha - C_D(z))(\alpha - \frac{k+1}{k+2} C_D(z))$, the equation above can be rewritten as

$$
(\alpha - c(z))^2 - 3\sqrt{\frac{\gamma f_E}{L}}\,(\alpha - c(z)) + 3\left(\sqrt{\frac{\gamma f_E}{L}}\right)^2 - \widetilde{U}(z)^A = 0.
$$

It is, however, difficult to identify which of the two roots of this quadratic equation in $(\alpha - c(z))$ is the sensible solution. Alternatively, we can write the welfare for the homogeneous firm case as the function of varieties $N(z)$:

$$\begin{aligned} U^{\hom} &= y_0^c + I + \int_0^1 \left[\frac{\gamma}{2} \int_{i\in\Omega(z)} (q_i^c(z))^2 di + \frac{\eta}{2} Q^c(z)^2 \right] dz \\ &= y_0^c + I + \int_0^1 \left[\frac{\gamma}{2} N(z) q^c(z)^2 + \frac{\eta}{2} (N(z) q^c(z))^2 \right] dz. \end{aligned}$$

Again, let $U^{\hom} = U^{het}$, we have

$$\frac{\gamma}{2} N(z)(q^c(z))^2 + \frac{\eta}{2}(N(z)q^c(z))^2 = \frac{1}{2\eta}\widetilde{U}(z)^A.$$

Since $q^c(z) = \sqrt{\frac{f_E}{\gamma L}}$, the left hand side can be rewritten as

$$\frac{\gamma}{2} N(z)(q^c(z))^2 + \frac{\eta}{2}(N(z)q^c(z))^2 = \frac{f_E}{2L} N(z) + \frac{\eta}{2}\frac{f_E}{\gamma L} N(z)^2.$$

Then we have

$$\frac{f_E}{\gamma L}\eta^2 N(z)^2 + \frac{\eta f_E}{L} N(z) - \widetilde{U}(z)^A = 0, \tag{E14}$$

which is a simple quadratic equation of $N(z)$. Given that $N(z) \geq 0$, the permissible solution is therefore given by

$$N(z)^A = \frac{\sqrt{\eta^2 f_E^2/L^2 + 4\frac{f_E}{\gamma L}\eta^2 \widetilde{U}(z)^A} - \frac{\eta f_E}{L}}{2\frac{f_E}{\gamma L}\eta^2}.$$

Then we also know the corresponding $C_{\hom}(z)$ which satisfies $N(z)^A = \frac{(\alpha - C_{\hom}(z))\sqrt{\frac{\gamma L}{f_E}} - 2\gamma}{\eta}$. Substituting the expression for $N^A(z)$ then gives

$$\begin{aligned} C_{\hom}(z) &= \alpha - \left[2\gamma + \eta \frac{1}{2}\frac{\gamma}{\eta} \left(\sqrt{1 + 4\frac{\widetilde{U}(z)^A L}{\gamma f_E}} - 1 \right) \right] \sqrt{\frac{f_E}{\gamma L}} \\ &= \alpha - \frac{1}{2}\sqrt{\frac{\gamma f_E}{L}} \left(3 + \sqrt{1 + 4\frac{\widetilde{U}(z)^A L}{\gamma f_E}} \right) \\ &= \alpha - \frac{1}{2}\left(3\sqrt{\frac{\gamma f_E}{L}} + \sqrt{\frac{\gamma f_E}{L} + 4\widetilde{U}(z)^A} \right). \end{aligned} \tag{E15}$$

It is easy to verify that this is a solution to Equation (E13). On the other hand, the other root of Equation (E13) leads to $N(z)^A = \frac{-\sqrt{\frac{\gamma f_E}{L}(\widetilde{U}(z)^A + 4\frac{\gamma f_E}{L})} - \frac{\gamma}{2}}{\eta} < 0$.

In the open economy, the free entry condition is given by

$$\begin{aligned}
\frac{L}{4\gamma}(p_{\max}(z) - c^A(z))^2 + \frac{L^*}{4\gamma}(p_{\max}(z)^* - \tau c^A(z))^2 &= f_E, \\
\frac{L^*}{4\gamma}(p_{\max}(z)^* - c^A(z)^*)^2 + \frac{L}{4\gamma}(p_{\max}(z) - \tau^* c^A(z)^*)^2 &= f_E.
\end{aligned}$$

There are two equations and two unknowns $p_{\max}(z)$ and $p_{\max}(z)^*$. In principle, we can solve for $p_{\max}(z)$ and $p_{\max}(z)^*$ for given parameters. Once the choke prices are known, we can solve for $Q^c(z)$ and $Q(z)^{c*}$ using:

$$\begin{aligned}
p_{\max}(z) &= \alpha - \eta Q^c(z), \\
p_{\max}(z)^* &= \alpha - \eta Q(z)^{c*}.
\end{aligned}$$

Moreover, firm outputs are known given that

$$\begin{aligned}
q^{HH}(z) &= \frac{L}{2}(p_{\max}(z) - c(z)), \\
q^{HF}(z) &= \frac{L^*}{2}(p_{\max}(z)^* - \tau c(z)),
\end{aligned}$$

$$\begin{aligned}
q^{FF}(z) &= \frac{L^*}{2}(p_{\max}(z)^* - c(z)^*), \\
q^{FH}(z) &= \frac{L}{2}(p_{\max}(z) - \tau^* c(z)^*).
\end{aligned}$$

Then we can solve for the number of varieties $\{n^H(z), n^F(z)\}$ using

$$\begin{aligned}
Q^c(z)L &= n^H(z)q^{HH}(z) + n^F(z)q^{FH}(z), \\
Q(z)^{c*}L^* &= n^F(z)q^{FF}(z) + n^H(z)q^{HF}(z).
\end{aligned}$$

The solution is

$$\begin{aligned}
n^H(z) &= \frac{Q(z)^{c*}L^*q^{FH}(z) - Q^c(z)Lq^{FF}(z)}{q^{FH}(z)q^{HF}(z) - q^{FF}(z)q^{HH}(z)}, \\
n^F(z) &= \frac{Q^c(z)Lq^{HF}(z) - Q(z)^{c*}L^*q^{HH}(z)}{q^{FH}(z)q^{HF}(z) - q^{FF}(z)q^{HH}(z)}.
\end{aligned}$$

The welfare for the home country is then given by

$$\begin{aligned} U &= y_0^c + I + \int_0^1 \left[\frac{\gamma}{2} \int_{i\in\Omega(z)} (q_i^c(z))^2 di + \frac{\eta}{2} (\int_{i\in\Omega(z)} q_i^c(z) di)^2 \right] dz, \\ &= y_0^c + I + \int_0^1 \left[\frac{\gamma}{2} (n^H(z)(\frac{q^{HH}(z)}{L})^2 + n^F(z)(\frac{q^{FH}(z)}{L})^2) + \frac{\eta}{2} Q^c(z)^2 \right] dz. \end{aligned}$$

For the foreign country, it is given by

$$U^* = y_0^c + I + \int_0^1 \left[\frac{\gamma}{2} (n^F(z)(\frac{q^{FF}(z)}{L^*})^2 + n^H(z)(\frac{q^{HF}(z)}{L^*})^2) + \frac{\eta}{2} Q(z)^{c*2} \right] dz.$$