Flexibility and Productivity: Towards the Understanding of Firm Heterogeneity*

Luca Macedoni — Aarhus University

Mingzhi (Jimmy) Xu — Peking University

September 2021

Abstract

This paper studies the heterogeneity in scope and sales across multi-product firms. We document four new stylized facts using Chinese firm-level data to show that the standard single-attribute model of trade, in which firms only differ in productivity, must be augmented with an additional layer of heterogeneity in order to match the facts. Combining the theoretical predictions of a rich framework with the empirical results, we show that to match the evidence, a model requires heterogeneity in flexibility, namely the ability to introduce new varieties in a destination at low costs. We calibrate our model and find suggestive evidence of a trade-off between productivity and flexibility: more productive firms are likely to be less flexible, as they pay higher fixed costs to introduce new varieties. By quantifying the role of firm flexibility on the welfare gains from trade, we show that ignoring firm flexibility causes a large underestimation of the welfare effects of trade.

Keywords: international trade; multi-product firms; flexible manufacturing; firm heterogeneity, firm-level data; China.

JEL Code: F12, F14, L11

*We are grateful to Costas Arkolakis, Dirk Krueger, and anonymous referees of the International Economic Review for their valuable suggestions and constructive comments. We thank Ina Simonovska and Robert Feenstra for their early suggestions and support. We also thank Sharat Ganapati, Michael Koch, Deborah Swenson, Frederic Warzynski, and seminar participants at Aarhus University, European Trade Study Group 2017, Midwest Spring International Trade Meeting 2018, Ljubljana Empirical Trade Conference 2018, North American Summer Meeting of the Econometric Society 2018, IFN, KU Leuven, and UC Davis. Financial support from the Carlsberg Foundation is gratefully acknowledged by Luca Macedoni. Financial support from the National Science Foundation of China (Grant No. 8200905868) and Peking University Initiative Scientific Research Program (Grant No. 7101302576) and Peking University International Strategic Cooperation Research Program (Grant No. 7101702318) is gratefully acknowledged by Mingzhi Xu. Financial support from the China Scholarship Council is gratefully acknowledged.

Please directly contact Luca Macedoni at Aarhus University, email:lmacedoni@econ.au.dk; Mingzhi Xu, INSE at Peking University, email: mingzhixu@nsd.pku.edu.cn.

All errors are ours.

1 Introduction

Models of multi-product firms usually assume that one attribute of the firm, i.e., productivity, drives both the ability to produce a good efficiently and the ability to produce many goods. More productive firms produce a larger number of varieties, which is defined as scope, and enjoy larger sales (Bernard et al., 2011; Mayer et al., 2014). The positive relationship between productivity, sales, and scope is intuitive: a rise in a firm's productivity increases the sales of each existing variety and restores the profitability of discontinued varieties. Despite the usefulness of such an approach, this study provides new evidence that shows that single-attribute models cannot replicate important features of the data, such as the presence of firms with large sales but few varieties. The goal of the paper is to show theoretically and quantitatively that to explain the empirical patterns, a model requires firms to differ in their productivity *and* in the ability to introduce new varieties at low costs. As multi-product firms dominate domestic production and international trade (Bernard et al., 2007), such a two-dimensional heterogeneity has major implications for the welfare gains from trade.

While our empirical analysis indicates that a model needs two layers of heterogeneity, to pin down the second source of heterogeneity, we combine the testable predictions of a rich theoretical framework with our novel empirical facts. Using this approach, we show that a model needs firm heterogeneity in the ability to introduce new products, which we label flexibility, and that other additional sources of heterogeneity, such as product appeal or quality, fail to rationalize the evidence. We estimate our model and apply it to study the role of flexibility in affecting the welfare consequences of trade. The quantification results suggest ignoring the second layer of heterogeneity will cause a large underestimation of the welfare loss associated with the recent US and China trade tensions.

Our main empirical contribution is to document four new stylized facts for Chinese exporters, using data from the China Customs Dataset and the Annual Surveys of Industrial Production (Brandt et al., 2014). Our first two stylized facts indicate that a model needs at least two layers of heterogeneity. First, we find that firm-destination specific shocks explain more than 50% of the variation in scope across firms and destinations. Second, we document a disconnect between total sales of firms and their scope in a given destination: there are several narrow- and wide-scope Chinese firms at any level of sales. These two facts are in stark contrast to the predictions of single-attribute models. In these models, the scope of firms only depends on the firm productivity and on destination characteristics (Bernard et al., 2011; Mayer et al., 2014; Macedoni, 2017) with no room for other shocks. Furthermore, such models predict a positive one-to-one mapping between sales and scope driven by productivity (Bernard et al., 2011).

The third and fourth stylized facts indicate particular characteristics that the second layer of heterogeneity must have. In our third stylized fact, we investigate the determinants of the scope conditional on sales, namely the variation in scope that is not explained by sales, which should be negligible according to single-attribute models. We find that more than half of its variation is driven by firm-destination shocks, as firm characteristics account for 30-40% of its variation and destination characteristics account for 15%.[1] The scope conditional on sales is also an important predictor of the response of firms to trade shocks. In our fourth stylized fact, we study the impact of WTO entry on the scope response of Chinese firms. We find that firms with larger scope conditional on sales in 2000 are less responsive to the reduction in tariffs of the subsequent years, meaning that they increase their scope by less.

To replicate the first stylized fact, the new layer of heterogeneity must be firm-destination specific. The trade literature has considered two additional types of heterogeneity: on the demand side, such as heterogeneity in quality or product appeal, and on the supply side, such as heterogeneity in the fixed cost per variety (references below). However, the second stylized fact clearly points to a firm-destination specific shock that affects the choice of scope but leaves the sales per variety unchanged. This means that the model needs a shock to the fixed cost per variety. We show theoretically that shocks that only affect the intensive margin, namely the sales per variety, such as quality or appeal, fail to rationalize this stylized fact.

We interpret the heterogeneity in the fixed cost as heterogeneity in the flexibility of firms' production processes. In fact, Eckel and Neary (2010) define flexible manufacturing as the ability of firms to introduce new varieties with minimal adaptations to production processes.[2] While in Eckel and Neary (2010) firms are fully flexible, and the fixed cost per variety is absent, in our framework a firm's flexibility is subject to shocks across destinations. This reflects the heterogeneous ability of a firm to adapt to foreign non-tariff barriers. Namely, a Chinese firm may find it less costly to adapt product standards and to package for the South Korean market than for the French market.[3] Since firms with higher flexibility have a wider scope, they are also less responsive to changes in trade costs, which is consistent with our fourth stylized fact.

Our theoretical model introduces, in a parsimonious way, a correlation between productivity and flexibility draws. Such an assumption allows us to match the third stylized fact whereby a sizeable portion of the variation in scope conditional on sales is driven by firm characteristics. As we leave the sign of such correlation unspecified, our theory nests the framework of flexible manufacturing by He (1992), in which more productive firms can invest in a more flexible technology or other models which feature, for instance, diseconomies of scope. Estimating such a correlation allows us to systematically evaluate the relationship between productivity and flexibility and shed

[1] Although the first and third stylized facts are similar, they fulfill different objectives, as we explain below. The first stylized fact indicates the need of a shock which is firm and destination specific. The third stylized fact indicates that the shock we introduce in our model is not i.i.d, as it depends on firm characteristics.

[2] Flexible manufacturing was first addressed by the IO literature (Eaton and Schmitt, 1994). A more general definition by Milgrom and Roberts (1990) states that flexible manufacturing allows for a quick response to changes in market conditions. Thesmar and Thoenig (2007) study how vertical integration or separation affect such flexibility. Other definitions of flexibility used in the literature are the ability to reduce delivery times (Tseng, 2004) and to change production scale with minor adjustment costs (Gal-Or, 2002).

[3] Although there is little evidence on the determinants of fixed costs of exporting, a documented reason for the variance of flexibility across destinations could be the exporting experience of firms' managers (Mion et al., 2016).

some light on industry-specific determinants of flexible manufacturing.

The main features of our model are as follows. On the consumer side, our model follows Bernard et al. (2011) with horizontally differentiated varieties aggregated in a CES fashion. As in Arkolakis et al. (2021), firms choose products among an infinite set of potential varieties. Varieties are introduced with increasing marginal cost, which represents the core competence assumption introduced in the international trade literature by Eckel and Neary (2010) and subsequently quantified by Arkolakis et al. (2021). Firms differ in the labor productivity of their core, or most efficient variety, thus linking the model to the standard Melitz (2003) case. As mentioned above, our model features a second source of heterogeneity in the fixed cost per variety and destination as in Eaton et al. (2011) for single product firms and Arkolakis et al. (2021) for multiproduct firms.[4]

We estimate the parameters of our model by matching moments from the distribution of sales and scope of Chinese exporters. To estimate the model, we follow the conventional assumptions on the functional forms for the distributions of productivity as well as the production technology within firms. In particular, we assume that productivity follows a Pareto distribution as Helpman et al. (2004) and Chaney (2008). Moreover, flexibility also follows a Pareto distribution to match the distribution of scope conditional on productivity that we observe in the data. The resulting distribution of sales resembles a log-normal distribution, which is in line with the data. This is in contrast to single-attribute models with CES preferences, in which the distribution of sales follows the distribution of productivities (Mrázová et al., 2021).

Our estimation results indicate that the distribution of firms' flexibility is negatively correlated with the distribution of productivity: more productive firms tend to be less able to introduce new varieties at low costs. This negative correlation allows the model to match the large mass of high-sales firms with narrow scope we observe in the data, as the high productivity of these firms generates high sales, and the low flexibility is associated with few products per firm. To the extent that flexibility and productivity are endogenous choices of firms, the estimation suggests a negative trade-off between the two sources of heterogeneity. If firms are able to invest in improving their productivity, they might do so at the expense of flexibility, and vice versa.[5]

We use our quantitative exercise to rule out alternative explanations for our stylized facts. First, we consider product-specific ex-post demand shocks, realized after firms' scope decisions, introduced by Arkolakis et al. (2021) in the context of multiproduct firms.[6] A model with ex-post

[4]Shocks to fixed costs are also present in Das et al. (2007) and Armenter and Koren (2015). The literature has also considered an additional layer of heterogeneity on the demand side (Kee and Krishna, 2008; Eaton et al., 2011; Demidova et al., 2012; Roberts et al., 2018; Cherkashin et al., 2015).

[5]We also estimate the parameters of the model by industry, finding a high degree of heterogeneity, both in the dispersion of the flexibility shocks and in the magnitude of the correlation between productivity and flexibility. We have verified that our results are robust to allowing a correlation in the fixed cost shocks across destinations, which is a possible outcome of firms learning by exporting.

[6]The model of Arkolakis et al. (2021) also features firm-destination specific shocks to the fixed costs. While Arkolakis et al. (2021) interpret the shock as an *i.i.d.* random cost of market access, our study argues that the shock affects firms' flexibility in production. In support of our assumption, the evidence shows that the shock is related to firms' characteristics, and the estimation of the model highlights a negative relationship between flexibility and

demand shocks can generate a disconnect between scope and sales, but can hardly match the presence of large firms with narrow scope.[7] Second, we calibrate the model by Nocke and Yeaple (2014), which features diseconomies of scope and thus provides a possible micro-foundation for our quantitative finding of a negative correlation between flexibility and productivity. We find that such a model cannot replicate salient features of the data, which suggests that diseconomies of scope can only be part of the explanation for our results.

To evaluate the importance of modeling flexibility heterogeneity, we study the welfare effects of increases in trade costs between the US and China to match the rise in tariffs following the recent trade tensions between the two countries. We exploit the gravity framework supported by our model and apply the exact hat algebra technique introduced by Arkolakis et al. (2012), which allows the study of changes in trade costs in a parsimonious way. In our model, a rise of the bilateral trade costs between China and the US by 13-17% causes a reduction in welfare of 0.1% in China and 0.08% in the US. These results are two times as large as those in a single-attribute model which would predict a fall in welfare by about 0.03-0.05% in both countries. The reason for this difference in results lies in the fact that single-attribute models cannot replicate the presence of large firms with narrow scope. This is crucial because we document that narrow-scope firms tend to be more responsive to trade shocks. Since our model matches the evidence that a large fraction of the largest firms have narrow scope, it predicts larger aggregate responses to changes in trade costs.

The remainder of the paper is organized as follows. Section 2 presents the data, briefly outlines a single-attribute model of multi-product firms, and shows how the model fails to replicate four new empirical regularities for Chinese exporters. Section 3 outlines our model of multi-product firms, which we estimate in section 4. Section 5 studies the welfare effects of the rise in tariffs between the US and China. Section 6 concludes the paper.

2 Stylized Facts for Multi-Product Firms

To demonstrate the need to introduce two dimensions of heterogeneity for multi-product firms in theory, in this section, we present the key features of a standard model and compare them to data facts. To do so, after having introduced the data, we briefly outline a standard model in which heterogeneity across firms is driven solely by productivity differences. Then, we show that such a model fails to replicate a number of new stylized facts for multi-product firms. These stylized facts include the decisions over the number of products that firms export, namely the scope, and the relationship between scope and sales. Finally, we show that the failures of the standard model can be potentially relevant for the analysis of the welfare effects of trade, as we document meaningful

productivity.

[7]Adding demand shocks improves the goodness of fit of our model, but the negative correlation between productivity and flexibility remains robust.

differences between the firm-level responses to changes in trade costs depending on the firm-level relationship between scope and sales.

Data Description. We use firm-level data from two sources. The first is the China Customs Dataset that provides data on export values at the product-firm-destination-year level for all international transactions from China. A product is a Harmonized System (HS) eight-digit code. To understand which firms' characteristics influence the scope decisions of firms, we use the Annual Surveys of Industrial Production (ASIP) that is conducted by the National Bureau of Statistics of China (Brandt et al., 2014). The dataset covers manufacturing firms with more than five million RMB in annual sales (≈ \$700k). For each firm, the ASIP provides data on employment, output, and elements of accounting statements.

Further we use name, location, zip code, and telephone number to match the firms in the two datasets, and we are able to match 30% to 40% of exporters involved in ordinary trade to the information provided in ASIP. [8] The distribution for scope and sales in the two samples are similar: although the matched sample features higher scope and revenues on average, the similarity in the shape of the distribution across the two samples is is confirmed by statistical tests (see Online Appendix). The China Customs Dataset ranges from 2000 to 2006 while the ASIP from 1998 to 2007. While the results of our paper hold for each year separately, for the sake of exposition we focus on 2006, which has the largest number of matched firms. Moreover, we refine our sample to the firms involved in ordinary trade only.[9]

2.1 The Standard Model with Only Productivity Heterogeneity

The purpose of this section is to derive the prediction over the scope and sales of firms of a standard model where firms only differ in terms of productivity. For this reason, we focus on the problem of firms from country i exporting to country j. We assume that there is a continuum of multi-product firms, each producing multiple varieties, indexed by ω, of a differentiated good. Firms are monopolistically competitive. As in the model of Bernard et al. (2011), we consider a CES demand function:

$$p_{ij}(\omega) = A_j x_{ij}(\omega)^{-\frac{1}{\sigma}} \tag{1}$$

where $p_{ij}(\omega)$ is the price of a variety ω, $x_{ij}(\omega)$ is the aggregate quantity demanded in j for variety ω produced in i, $\sigma > 1$ is the elasticity of substitution across varieties, and A_j is a destination-specific demand shifter. The main predictions we test also hold under alternative demand structures and

[8] We follow the conventional method to match firms from the China Customs Data to the ASIP (Feenstra et al., 2014; Yu, 2015; Manova and Yu, 2016).

[9] The code for ordinary trade mode is 18. According to Dai et al. (2016), firms involved in processing exports behave abnormally in China. The matching performance and the summary statistics are in Appendix A.

assumptions over the competition.[10]

Each firm produces a continuum of varieties. Labor is the only input and is paid a wage w_i. To produce and deliver a given variety in a destination, a firm pays a constant marginal cost and a fixed cost. Varieties within-firm differ in their marginal costs: firms have a core competence and marginal cost rise as they introduce varieties farther from the core (Eckel and Neary, 2010; Mayer et al., 2014). The marginal cost to produce a unit of variety ω from i to j is $\tau_{ij} w_i c h(\omega)$, where τ_{ij} is the iceberg trade cost, and $\tau_{ii} = 1$. Each firm produces the first variety, or *core*, with the lowest variable cost, and the variable cost of new varieties increases with their distance from the core. Hence, we assume that $h(\omega)$ is increasing in ω. Moreover, we normalize $h(0) = 1$ so that c is the marginal cost of the core variety, and it is randomly drawn by firms. Firms also pay a fixed cost per variety and destination f_j in destination labor units.

The only difference across firms is in the marginal cost c. The assumptions of monopolistic competition and CES preferences yields the standard constant markup pricing rule:

$$p_{ij}(\omega, c) = \frac{\sigma}{\sigma - 1} \tau_{ij} w_i c h(\omega) \tag{2}$$

Thus, the profits of a variety ω for firm c equal:

$$\pi_{ij}(\omega, c) = \frac{1}{\sigma - 1} \left[A_j \frac{\sigma - 1}{\sigma} \right]^{\sigma} (\tau_{ij} w_i c h(\omega))^{1-\sigma} - w_j f_j \tag{3}$$

Since $h' > 0$, firms introduce varieties until the last variety, denoted by $\delta_{ij}(c)$, generates zero profits. $\delta_{ij}(c)$ is also the measure of the mass of varieties produced by the firm, namely its scope. The scope is implicitly defined as:

$$h(\delta_{ij}(c))^{\sigma-1} = \frac{1}{w_j f_j (\sigma - 1)} \left[A_j \frac{\sigma - 1}{\sigma} \right]^{\sigma} (\tau_{ij} w_i c)^{1-\sigma} \tag{4}$$

We summarize the main features of (4) in the following proposition:

Proposition 1. *In a standard model with only productivity heterogeneity, the scope ($\delta_{ij}(c)$) of an exporter from i to j is a function of destination characteristics (A_j, w_j, f_j), origin-firm characteristics (w_i, c), and bilateral trade costs (τ_{ij}).*

Following the notation of Arkolakis et al. (2021), let $H(\delta_{ij}(c)) = \left[\int_0^{\delta_{ij}(c)} h(\omega)^{1-\sigma} d\omega \right]^{\frac{1}{1-\sigma}}$ be a measure of the productivity of the firm across its varieties. $H(\delta_{ij}(c))$ is declining in scope: as the firm introduces new varieties far from the most productive core, its average productivity falls. The

[10] See Mayer et al. (2014) for non-homothetic utility, and Macedoni (2017) for oligopoly and for nested-CES demand.

total sales of a firm are:

$$R_{ij}(c) = \frac{\sigma}{\sigma-1}\left[A_j\frac{\sigma-1}{\sigma}\right]^{\sigma}(\tau_{ij}w_i c H(\delta_{ij}(c)))^{1-\sigma} \tag{5}$$

Using our definition of scope (4), we can rewrite (5) as:

$$\left(\frac{h(\delta_{ij}(c))}{H(\delta_{ij}(c))}\right)^{\sigma-1} = \frac{R_{ij}(c)}{\sigma w_j f_j} \tag{6}$$

We summarize the main feature of (6) in the following proposition:

Proposition 2. *In a standard model with only productivity heterogeneity, there is a monotone positive relationship between the scope of a firm in a destination and its revenues.*

The proof comes from the assumption that $h'(\omega) > 0$, which also implies that $H'() < 0$. The proposition means that when two firms have the same scope, they also have the same revenues and vice versa. In the next section, we test how Propositions 1 and 2 perform in the data.

2.2 Challenge for the Standard Model: Cross-sectional Evidence

I. Exporter Scope Across Firms and Destinations. Let us consider the log of the scope defined in (4) for a firm denoted by f with marginal cost c:

$$\underbrace{\ln h(\delta_{ij}(c))}_{\ln(\#\ \text{Products}_{fj})} = \underbrace{\frac{1}{\sigma-1}\ln\left(\frac{(\sigma-1)^{\sigma-1}}{\sigma^{\sigma}}\right)}_{\text{Constant}} - \underbrace{\ln(w_i c)}_{\text{Firm FE}} + \underbrace{\frac{1}{\sigma}\ln\left(\frac{A_j^{\sigma}\tau_{ij}}{w_j f_j}\right)}_{\text{Destination FE}} \tag{7}$$

We approximate the left-hand side with the log of the number of products, noting that the approximation is exact in the case in which $d\ln h(\delta_{ij})/d\ln\delta_{ij}$ is constant.[11] The log of the scope of an exporter f, from China to a destination j, depends on two fixed effects. The first component is a firm fixed effect that captures firm's f productivity and costs of production common for all destinations reached. The second one is a destination fixed effect that captures characteristics of the destination j and bilateral trade costs. To quantify the explanatory power of the two margins, we regress the number of products that a firm f exports to a destination j on a firm fixed effect a_f and a destination fixed effect d_j.

[11] For this reason, we also considered an approach with higher order polynomial; the results are similar to the baseline ones.

Table 1: Decomposition of Product Scope Variation

Sample	Overall Fit (R^2)	Firm FE Only (R^2)	Destination FE Only (R^2)
All Exporters	0.37	0.36	0.01
Matched Exporters	0.44	0.42	0.02

R^2 from from regressing log(scope) on firm and destination fixed effects together and individually.

In Table 1, we report the R^2 of the regression and the contribution to the model fit of the firm and destination fixed effects. Our first finding is:

Stylized Fact 1. *More than half of the variations in the scope of Chinese exporters is explained by firm-destination specific shocks.*

The standard model represented by (7) only explains 37% of the scope variation for all exporters and 44% for the sample of matched exporters. This result directly contrasts with Proposition 1, which excludes the presence of firm-destination specific shocks.

II. Disconnect Between Sales and Scope. We now focus on the within-destination distribution of product scope and sales. Taking the log of (6), we obtain:

$$\underbrace{\ln\left(\frac{h(\delta_{ij}(c))}{H(\delta_{ij}(c))}\right)^{\sigma-1}}_{\#\text{ Products}_{fj}} = \underbrace{\ln R_{ij}(c)}_{\ln\text{Sales}_{fj}} - \underbrace{\ln(\sigma w_j f_j)}_{\text{Constant}} \tag{8}$$

The scope is a function of revenues and a constant. In other words, the standard model predicts that given two firms with the same sales, they also have the same scope. Choosing the US as destination, we plot the scope of Chinese exporters, normalized by the industry average, against the log of their sales, demeaned by the industry average, in Figure 1.[12] Our second finding is summarized as:

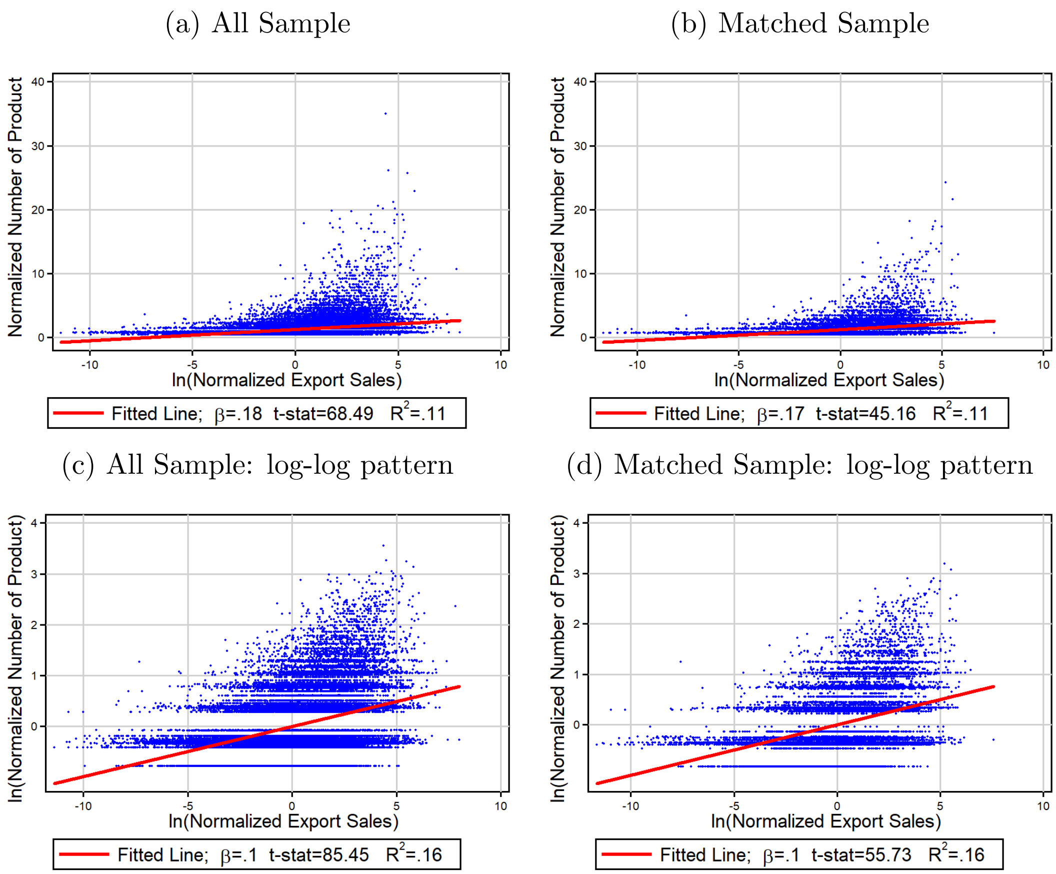
Stylized Fact 2. *At any level of sales, there are many single-product firms and multi-product firms with different scope.*

This is in contrast to Proposition 2, which stated that two firms with the same sales also have the same scope. Panel (a) displays the pattern for all Chinese exporters to the US, and Panel (b)

[12]Throughout the analysis, we only use single-industry exporters. The classification of industries is reported in Table A.3 in the appendix. For each exporter f in industry j, its demeaned variable (i.e., $\exp(\widetilde{\ln x_{fj}})$) is computed in such a way that $\widetilde{\ln x_{fj}} = \ln x_{fj} - \overline{\ln x_{fj}}$, where $\overline{\ln x_{fj}}$ denotes the arithmetic mean of $\ln x_{fj}$ across all firms within industry j. The pattern remains robust if we include multi-industry exporters in the sample, as displayed in appendix Figure C.1. In this case, we define a firm observation as a firm-industry. The normalization of firm sales and scopes remains the same as in Figure 1.

displays the matched ones. Although, on average, there is a positive relationship between sales and scope as suggested by the standard model in equation (8), the figures highlight that such a relationship is far from perfect. The scope variations explained by firm sales are very limited as the R^2 of the line of best fit is 11%. At any level of sales, there are single-product firms and multi-product firms. Particularly, contrary to the single-attribute models, there are many narrow-scope firms at the medium-high level of sales. This pattern remains robust in Panels (c) and (d), where we use the scope in logs instead of the scope in levels.

Figure 1: Product Scope and Sales of Exporters in the US (2006)



In the model section, we show that introducing heterogeneity of fixed costs per variety and destination rationalizes this stylized fact. This can be intuitively understood here: if the fixed cost f_j also varies across firms, then two firms with the same sales can have different scope. We will make this point both theoretically and quantitatively. At this stage, it is important to notice that a shock to the fixed cost can explain the result. However, to match the presence of large (productive) firms, with narrow scope, the distribution of the shock has to be such that these large firms are likely to get high fixed cost draws.

III. Disconnect across Firms and Destinations. Next, we investigate whether the deviations of the exporters' scope from the value predicted by export values depend on other characteristics of firms. We first run the following regression for Chinese exporters across destinations j:

$$\ln(\#\ \text{Products}_{fj}) = b_0 + b_1 \ln(\text{Revenue}_{fj}) + \xi_{fj} \tag{9}$$

and record the residual ξ_{fj}, which is the scope conditional on sales: the variation in firms' scope across destination that is not explained by sales. We regress ξ_{fj} on firm (α_f) and destination (d_j) fixed effect. Table 2 reports the results, which we summarize as follows:

Stylized Fact 3. *Destination and firm characteristics account for less than half of the variation in the scope conditional on sales, with firm characteristics having twice as much explanatory power as destination characteristics.*

In fact, destination characteristics alone account for 15% of the variation, while firm characteristics account for 30-40%. The largest component of the scope conditional on sales is generated by the interaction between firm and destination characteristics. This stylized fact motivates our choice of modeling the shock to the extensive margin of firms as being firm and destination specific. The scope conditional on sales ξ_{fj} can be interpreted as the additional capacity of a firm to introduce new products relative to what is predicted by its sales. The result of this section shows that this capacity, which we are going to label flexibility in the model section, depends on some firm level characteristics, i.e., some firms are better equipped at expanding their scope and destination characteristics, i.e., expanding the scope in some destinations is easier than in others. Because of the larger explanatory power of firm characteristics, in the model, we introduce a correlation between firm productivity and flexibility.

Table 2: Variations of Scope Conditional Sales

Sample	Overall Fit (R^2)	Firm FE Only (R^2)	Destination FE Only (R^2)
All Exporters	0.46	0.31	0.15
Matched Exporters	0.52	0.38	0.14

R^2 from regressing ξ_{fj} from (9) on firm and destination fixed effects together and separately.

Robustness. We perform a number of robustness checks, and all results are presented in appendix C.1. In addition to the disconnect between scope and sales, there is a robust disconnect between scope and measures of firm productivity. In Panel (a) of Figure C.2, we plot the relationship between scope and total factor productivity[13], and in Panel (b), we plot the scope and the value-added per worker.

The disconnect between scope and firm productivity becomes even starker in Figure C.3, in which we divide US exporters in quartiles by productivity and plot the distribution of product

[13] We follow Levinsohn and Petrin (2003) and estimate productivity, excluding firms whose productivity values are $\leq 1\%$ or $\geq 99\%$. Appendix B provides the detailed procedure for estimating firms' productivity.

scope conditional on the firm belonging to a certain quartile of the productivity distribution. As shown in the figure, in each quartile, the majority of firms export a single-product to the US. The distribution of scope conditional on productivity resembles a Pareto distribution. The pattern is in stark contrast to equation (7) implied by a standard model of multi-product firms, which would predict that the peak of the distribution would be shifting to higher scope as productivity increases, even allowing for some noise in the data.

The pattern uncovered for the US market in 2006 also holds across alternative years: 2000 and 2003, as shown in Figure C.4.

Another concern may arise if the classifications of products and industries are artificially driving the disconnect. First, as firms may engage in very different activities within the very aggregate industry, in Figure C.5 we repeat the analysis defining an industry as an HS 4-digit code and excluding firms with trade activities in multiple HS 4-digit industries. The disconnect between sales and scope barely changes.

In addition, as we define a product as an HS 8-digit code, each HS 8-digit code can potentially include a large number of products. Namely, two firms with the same sales may have the same number of products, but one firm's products are classified within one HS code while the other firm's products cover several HS codes. We cannot fully dismiss such a concern without access to barcode data. However, we verified in Figure C.6 that the disconnect still emerges when we consider more aggregate definitions of products (i.e., defining a product as an HS 6- or 4-digit code). Thus, the way products are classified within HS codes does not appear to be driving the result.

Last, in Figure C.7 we show that the pattern uncovered for the US market also holds for other major destinations for Chinese exporters: Japan, South Korea, Germany, and the United Kingdom.[14]

2.3 Challenge for the Standard Model: Effects of Trade on Scope

The previous section presents the cross-sectional evidence that a model in which firms differ along a single-attribute fails to replicate a large portion of the variations in scope across firms and destinations. The evidence also suggests that there are systematic deviations of scope from the value that would be predicted by the sales, and that these deviations depend on firm and destination characteristics. In addition, we further provide the over time evidence that taking these previously

[14] In the online appendix, we report the results of the following additional robustness checks. The disconnect still persists when using nonparametric methods such as the locally weighted scatter-plot smoothing. The disconnect is also robust to analyzing each industry separately. The R^2 of regressing number of products on the log of sales ranges from 4% in Plastic & Rubber to 20% in Textile. We verified that the disconnect between sales and scope is also present when considering separately consumption, intermediate, and capital goods according to the Broad Economic Category (BEC) classification, and homogeneous and differentiated goods according to the definition introduced by Rauch (1999). We also verified that the results are robust to dividing the sample of firms according to their ownership (state-owned, private, and foreign-owned).

uncovered patterns into account is important in understanding the effects of changes in trade costs on the scope of exporters. As gains from varieties constitute a large portion of the gains from trade, these results reinforce the motivation for our model, since taking these stylized facts into account can have major welfare implications.

To make this point, recall that the partial elasticity of firm scope with respect to trade costs is obtained by partially differentiating (4):

$$\frac{d\ln\delta_{ij}(c)}{d\ln\tau_{ij}} = -\frac{h(\delta_{ij}(c))}{h'(\delta_{ij}(c))\delta_{ij}(c)} \tag{10}$$

According to the functional forms commonly used in the literature (Eckel and Neary, 2010; Arkolakis et al., 2021) and in our paper, the absolute value of such an elasticity is declining in scope, namely $d\frac{h(\delta)}{h'(\delta)\delta}/d\delta < 0$; that is, trade shocks have a smaller impact on firms with wider scope.

To assess the empirical relevance of a standard model, we exploit China's entry in the WTO in 2001 to study how a reduction in tariffs differentially affects firms depending on their initial scope, controlling for sales. We base our analysis on a difference-in-difference specification:[15]

$$\begin{aligned}\ln \#\ \text{Products}_{ft} =& \beta_1 \text{Post01}_t \times \text{Tariff}_{j0}^{EX} + \\ & \beta_2 \text{Post01}_t \times \text{Tariff}_{j0}^{EX} \times \ln \#\ \text{Products}_{f0} + \\ & \beta_3 \text{Post01}_t \times \text{Tariff}_{j0}^{EX} \times \ln \text{Sales}_{f0} + \\ & \gamma \mathbb{X}_{ft} + \alpha_f + \alpha_t + u_{ft}\end{aligned} \tag{11}$$

where Post01_t is a dummy variable that equals one for years after 2001, the year China joined the WTO; Tariff_{j0}^{EX} is the initial tariff of sector j which firm f belonged to.[16] We also control for sector j's import tariffs using the interaction $\text{Post01}_t \times \text{Tariff}_{j0}^{IM}$. We investigate the heterogeneous responses of exporters that differ in the initial (log) scope ($\ln \#\ \text{Products}_{f0}$) and the initial level of sales ($\ln \text{Sales}_{f0}$) in 2000. Finally, α_f and α_t stand for the firm and year fixed effects. The industry-year specific export and import tariffs are obtained from Yu (2015), and we use the tariffs in 2000 as the initial level. As shown in Figure C.8 in the appendix, both export (the solid line) and import (the dashed line) tariffs declined over time after China joined the WTO in 2001.

Table 3 reports the regression results of equation (11). Column (1) reports the overall impact of tariff reduction on firm scope. The coefficient β_1 is significantly positive, suggesting that the removal of tariff barriers increases the number of products supplied by Chinese exporters. However, as shown in column (2), the scope expansion declines in the initial scope of the firm. This is consistent with the assumption over the marginal cost technology $h(\omega)$, which generates a decreasing

[15] We adopt the difference-in-difference specification as reductions of tariffs occurred when China joined the WTO in 2001; there were not many changes in tariffs afterward. A similar specification is employed in the analysis of Chinese cities in Erten and Leight (2019) and Liu and Ma (2020).

[16] Following Yu (2015), industry classification of Chinese firms are at the two-digit level of Chinese Industry Classification (CIC).

scope elasticity with respect to trade costs, as shown in (10).

Table 3: Tariff Reductions and Product Expansion: the Role of Flexibility

Dep var: ln(# Products_{ft})	(1)	(2)	(3)	(4)	(5)
$\text{Post01}_t \cdot \text{Tariff}_{j0}^{EX}$	0.346***	1.768***	1.023***	1.002***	1.058***
	(0.087)	(0.093)	(0.250)	(0.249)	(0.256)
$\text{Post01}_t \cdot \text{Tariff}_{j0}^{EX} \cdot \ln(\#\ \text{Products}_{f0})$		-1.223***	-1.288***	-1.273***	-1.272***
		(0.042)	(0.046)	(0.046)	(0.046)
$\text{Post01}_t \cdot \text{Tariff}_{j0}^{EX} \cdot \ln(\text{Sales}_{f0})$			0.062***	0.064***	0.061***
			(0.019)	(0.019)	(0.019)
$\text{Post01}_t \cdot \text{Tariff}_{j0}^{IM}$	0.299	2.046***	2.007***	2.163***	1.991***
	(0.428)	(0.402)	(0.403)	(0.410)	(0.416)
Observations	50,464	50,464	50,464	49,002	48,667
R-squared	0.811	0.819	0.819	0.821	0.830
Firm controls	-	-	-	Y	Y
Firm FE	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
City-Year FE	-	-	-	-	Y

Note: the dependent variable is the number of products exported by firm f in year t ($\ln(\#\ \text{Products})_{ft}$). Firm controls include firm-year specific total exports, capital share in aggregate revenues, and firm productivity measure. All variables are in log. Standard errors are clustered at the firm level and reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

The negative effect of the initial scope does not disappear when we include the interaction term with the initial sales. Therefore, the time-series evidence suggests that:

Stylized Fact 4. *Firms that have larger (smaller) scope conditional on revenues are less (more) responsive to trade shocks.*

The Stylized Fact 4 is in contrast to Proposition 2. In fact, net of model misspecification, controlling for revenues should, according to the standard model, account for all of the variation of firm scope. However, we find that the cross-sectional disconnect between sales and scope we documented earlier also plays a role in explaining firm responses to trade costs shocks over time.

The coefficient β_2 barely changes when we control for firm sales in column (3). This result remains robust in columns (4) where we control for firm characteristics, and (5) where we control for city-year fixed effects. In appendix Table C.1, we use the time-variant sector tariffs (Tariff_{jt}^{EX}) from Yu (2015), instead of $\text{Post01}_t \times \text{Tariff}_{j0}^{EX}$ in regression (11). Consistent with Table 3, we observe a similar pattern showing that firms with wider initial scope responded to tariff reduction by increasing the number of products less.

3 Going Beyond the Standard Model: A Theory

To match the stylized facts that the standard single-attribute models cannot account for, we present an alternative theory of multiproduct firms. We begin the section by analyzing the problem of a multi-product firm in a destination. The evidence of the previous section highlighted the need for a model with at least two layers of heterogeneity across firms. The literature has considered two additional shocks: to the demand at the firm or firm-product level or to the fixed cost per variety. To highlight that only two sources of firm heterogeneity, one on the intensive margin (productivity) and one on the within firm extensive margin (flexibility), are necessary to match the data, we first consider a richer model with additional heterogeneity in firm-specific and firm-product-specific demand shifters. Then, we revert to a simple model with only two layers of heterogeneity to discuss the selection of firms into export markets and the equilibrium of the model.

Consider a world of I countries: subscript i denotes an origin and j a destination. In each country j, L_j consumers, with per capita income y_j, enjoy the consumption of varieties of a differentiated good. Consumers in each country j have the same CES preferences. Firms are multiproduct and are monopolistically competitive. Labor is the only input and is paid a wage w_i.

The timing of the model is similar to Demidova et al. (2012) and Cherkashin et al. (2015). After having discovered the firm productivity, a firm has to first decide whether or not to enter a destination j. Once the destination is entered, the firm maximizes its profits and chooses prices and scope. We begin by discussing the within-destination problem of the firm to highlight how the model can rationalize the disconnect between sales and scope.

3.1 Within-Destination Sales and Scope

Let us first consider the problem of a firm from country i active in destination j. Let us consider the following CES demand for a variety ω:

$$p_{ij}(\omega) = A_j z^{\frac{\sigma-1}{\sigma}} z(\omega)^{\frac{\sigma-1}{\sigma}} x_{ij}(\omega)^{-\frac{1}{\sigma}} \tag{12}$$

where $p_{ij}(\omega)$ is the price and $x_{ij}(\omega)$ the total supply of the variety, and $\sigma > 1$ is the elasticity of substitution. There are three demand shifters: a destination-specific demand shifter A_j, a firm-specific demand shifter z, which captures the firm-level quality or appeal, and a firm-product-specific demand shifter $z(\omega)$. The firm- and product-specific demand shifters can be destination specific.

A firm is identified by a vector $\boldsymbol{\varphi}$ of characteristics, which we list hereafter. Varieties within a firm are indexed by $\omega \in [0, \delta_{ij}(\boldsymbol{\varphi})]$, where $\delta_{ij}(\boldsymbol{\varphi})$ denotes the scope of the firm from i to j. The marginal cost to produce a unit of variety ω from i to j is $\tau_{ij} w_i c \tilde{h}(\omega)$, where τ_{ij} is the iceberg trade cost, and $\tau_{ii} = 1$. $\tilde{h}(\omega)$ is the variety-specific component of the marginal costs, and c is the firm-level component. To introduce a new variety in a destination j, each firm pays a fixed cost $f_j \beta$ in destination labor units, where f_j is the deterministic component while β is subject to firm-destination specific shocks. The fixed cost represents the costs of adjusting production and distribution processes to the new variety. We interpret the fixed cost per variety as a measure of a firm's flexibility: the lower the fixed cost, the higher the flexibility.

To summarize, there are three potential dimensions along which firms differ within a destination: firm-level demand shifter z, marginal cost c, and fixed cost β. Hence, $\boldsymbol{\varphi} = [z, c, \beta]$. Furthermore, there are two firm-product specific variables $z(\omega)$ and $\tilde{h}(\omega)$. In a model extension, we show that including a firm-product specific fixed cost does not change the results. The profits of a firm $\boldsymbol{\varphi}$ from i to j are given by:

$$\Pi_{ij}(\boldsymbol{\varphi}) = \int_0^{\delta_{ij}(\boldsymbol{\varphi})} [(p_{ij}(\omega) - \tau_{ij} w_i c \tilde{h}(\omega)) x_{ij}(\omega) - w_j f_j \beta] d\omega \tag{13}$$

with $p_{ij}(\omega)$ from (12). Profit maximization yields the standard constant markup pricing rule:

$$p_{ij}(\omega, \boldsymbol{\varphi}) = \frac{\sigma}{\sigma - 1} \tau_{ij} w_i c \tilde{h}(\omega) \tag{14}$$

It is convenient to re-write the profits of a variety ω as follows:

$$\pi_{ij}(\omega, \boldsymbol{\varphi}) = \frac{1}{\sigma - 1} \left[A_j \frac{\sigma - 1}{\sigma} \right]^{\sigma} (\tau_{ij} w_i)^{1-\sigma} \left(\frac{c}{z} \right)^{1-\sigma} \left(\frac{\tilde{h}(\omega)}{z(\omega)} \right)^{1-\sigma} - w_j f_j \beta$$

Let $h(\omega) = \frac{\tilde{h}(\omega)}{z(\omega)}$, and let $h'(\omega) > 0$, so that varieties within a firm are ordered by increasing quality-adjusted marginal costs. Furthermore, for simplicity, let us normalize the term for the first, or *core*, variety $h(0) = 1$. Without loss of generality, we can also normalize z to one, which is the equivalent of having c replace c/z. This implies that the demand shifters are isomorphic to the cost shifters and, thus, variety-specific profits can be written as:

$$\pi_{ij}(\omega, \boldsymbol{\varphi}) = \frac{1}{\sigma - 1} \left[A_j \frac{\sigma - 1}{\sigma} \right]^{\sigma} (\tau_{ij} w_i c h(\omega))^{1-\sigma} - w_j f_j \beta$$

Given that $h^{'}(\omega) > 0$, a value of the fixed cost shock $\beta^*_{ij}(\boldsymbol{\varphi})$ exists such that a firm with cost draw c makes zero profits by selling its core variety $\omega = 0$. If a firm with characteristics $\boldsymbol{\varphi}$ draws $\beta < \beta^*_{ij}(\boldsymbol{\varphi})$, it sells a positive scope to consumers in j. Otherwise, if $\beta > \beta^*_{ij}(\boldsymbol{\varphi})$, the firm is inactive in j. The fixed cost cutoff declines with the marginal cost of production and delivery:

$$\beta^*_{ij}(\boldsymbol{\varphi}) = \frac{1}{w_j f_j (\sigma - 1)} \left[A_j \frac{\sigma - 1}{\sigma} \right]^{\sigma} (\tau_{ij} w_i c)^{-(\sigma-1)} \tag{15}$$

and only depends on c. Hence, we can write $\beta^*_{ij}(\boldsymbol{\varphi}) = \beta^*_{ij}(c)$. Let us rewrite per variety profits as a function of the fixed cost cutoff $\beta^*_{ij}(c)$ and β:

$$\pi_{ij}(\omega, \boldsymbol{\varphi}) = w_j f_j \beta \left[\frac{\beta^*_{ij}(c)}{\beta} h(\omega)^{1-\sigma} - 1 \right]$$

The firm introduces new varieties until the profits of the last variety $\pi_{ij}(\delta_{ij}(\boldsymbol{\varphi}), \boldsymbol{\varphi})$ equal zero. Hence, the optimal scope of the firm is implicitly defined by:

$$(h(\delta_{ij}(\boldsymbol{\varphi})))^{\sigma-1} = \left(\frac{\beta_{ij}^*(c)}{\beta}\right) \tag{16}$$

Compare (16) with the predictions from the standard model (4). In both cases, the scope is a function of firm level characteristics (c in the standard model and $\boldsymbol{\varphi}$ in our model) and destination characteristics. However, while productivity perfectly explains the scope in the standard model, there is an extra element in ours represented by β. Since β is a shock which is firm and destination specific, the model is able to replicate the first stylized fact, whereby a firm-destination specific shock is the main source of variation in scope across firms and destinations. Notice that quality differences are included in $\boldsymbol{\varphi}$ but do not otherwise affect (16).

Let $H(\delta_{ij}(\boldsymbol{\varphi})) = \left[\int_0^{\delta_{ij}(\boldsymbol{\varphi})} h(\omega)^{1-\sigma} d\omega\right]^{\frac{1}{1-\sigma}}$ be a measure of the productivity of the firm across its varieties. $H(\delta_{ij}(\boldsymbol{\varphi}))$ is declining in scope: as the firm introduces new varieties far from the most productive core, its average productivity falls. The total sales of a firm are:

$$R_{ij}(\boldsymbol{\varphi}) = \sigma w_j f_j \beta_{ij}^*(c) H(\delta_{ij}(\boldsymbol{\varphi}))^{1-\sigma} \tag{17}$$

Using our definition of scope (16), we can rewrite (17) as:

$$\left(\frac{h(\delta_{ij}(\boldsymbol{\varphi}))}{H(\delta_{ij}(\boldsymbol{\varphi}))}\right)^{\sigma-1} = \frac{R_{ij}(\boldsymbol{\varphi})}{\sigma w_j f_j \beta} \tag{18}$$

which represents the sales and scope disconnect documented in the data. With the exception of β, the equation is identical to (6). The main feature of (18) is summarized as follows:

Proposition 3. *Given two firms with the same revenues, the firm with lower β (higher flexibility) offers the larger scope.*

With Proposition 3, our model is able to explain both Stylized Facts 1 and 2. In fact, β is the firm-destination specific shock that was missing in a standard model, which is the main source of scope variation documented in Stylized Fact 1. Furthermore, if we exclude fixed cost heterogeneity and consider only heterogeneity in product or firm level quality, we are not able to match Stylized Fact 2: if two firms have different quality, but the same revenues, their scope is going to be identical. To replicate Stylized Fact 3, we are going to make the assumption that the distribution of β is not i.i.d. but is correlated with firm marginal cost c. Finally, since conditional on revenues, firms with lower β have higher scope, they will respond less to changes in trade costs, which is consistent with Stylized Fact 4

3.2 Selection and Entry

Since the demand shocks are isomorphic to productivity, a firm is only identified by the pair of β and c. Hence, for clarity we replace $\boldsymbol{\varphi} = (c, \beta)$. The profits of a firm (c, β) in a destination j are given by:

$$\Pi_{ij}(c, \beta) = w_j f_j \beta \left[\left(\frac{h(\delta_{ij}(c, \beta))}{H(\delta_{ij}(c, \beta))} \right)^{\sigma-1} - \delta_{ij}(c, \beta) \right] \tag{19}$$

which are declining in c, provided that $\frac{d}{d\delta}\left(\frac{h(\delta)}{H(\delta)}\right)^{\sigma-1} > 1$.

Firms pay a (sunk) fixed cost of entry to discover their productivity and another (sunk) fixed cost per destination to discover their flexibility.[17] The timing of firms' decisions is as follows. In each country there is a pool of potential entrants. Upon entry, a firm pays a fixed cost f_E in domestic labor unit and discovers the productivity $\frac{1}{c}$ of its core variety, where c is drawn from a distribution $G_i(c)$, with pdf $g_i(c)$ and support $[0, \bar{c}_i]$. Only a mass N_i of firms pays f_E.

To export to a destination j, a firm faces a two-stage problem similar to Demidova et al. (2012) and Cherkashin et al. (2015). In the first stage, the firm decides whether to pay a fixed cost F_j in destination labor units to discover its realization of the fixed cost β per variety in j. Only a subset of firms pay F_j. The firm draws β from a distribution $B_c(\beta)$ with pdf $b_c(\beta)$ and support $[0, \beta_{max,c}]$. The distribution of β could be firm-specific since we documented that the conditional scope is related to observable characteristics of the firm. The second stage is what we examined in the previous section, which yields profits (19).

To the extent that flexibility is a feature of the design and production of goods by a firm, we might expect the draws of the shock to the fixed cost to be correlated across destinations within-firm. This possibility is suggested by our evidence that firm characteristics explain 30-40% of the variation in the scope conditional on sales. To reflect this channel, our model features a distribution of the fixed cost shock which is correlated with the firm's productivity. We leave the sign of the correlation unspecified and will determine it with the estimation of the model. For instance, a firm may have designed its products in order to easily adapt them to regulations and standards of any destination. Such a design may come at the cost of reducing the total volumes that the firm can produce, or it may increase them.

The firm's expected profits in j equal:

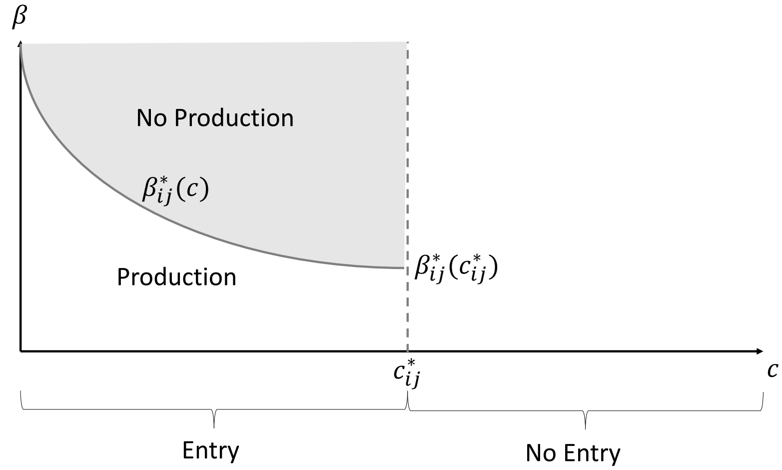
$$E[\Pi_{ij}(c)] = \int_0^{\beta^*_{ij}(c)} \Pi_{ij}(c, \beta) b_c(\beta) d\beta \tag{20}$$

A firm with cost draw c enters a destination j as long as its expected profits over the possible realizations of the fixed cost shock exceed the fixed cost of entry F_j.[18] Hence, a marginal firm with

[17] According to the estimates of Das et al. (2007), the fixed cost of entry in a foreign market is quite large.

[18] The expected profits are declining in c if $\int_0^{\beta^*_{ij}(c)} \left[\frac{\partial \Pi_{ij}(c,\beta)}{\partial c} b_c(\beta) + \Pi_{ij}(c, \beta) \frac{\partial b_c(\beta)}{\partial c} \right] d\beta < 0$.

Figure 2: Entry and Production Decisions



cost draw c^*_{ij} exists such that:

$$E[\Pi_{ij}(c^*_{ij})] = w_j F_j \tag{21}$$

Figure 2 summarizes the entry and production decisions of firms. A firm with cost draw $c > c^*_{ij}$ does not pay the fixed cost F_j and, thus, decides not to enter the destination. On the other hand, a firm with $c < c^*_{ij}$ pays the fixed cost F_j and discovers its draw of β. If $\beta > \beta^*_{ij}(c)$, the firm does not produce any variety. If β equals the cutoff, the firm is indifferent between producing the core variety and not producing, thus, having a scope of zero. For β below the threshold, the firm has a positive scope and sales. For a given c, firms with low β have a higher scope $\delta_{ij}(c, \beta)$.

Firms pay the fixed cost f_E if their expected profits across all destinations exceed the fixed cost of entry. In particular, ex-ante expected profits across all destinations are given by:

$$\pi_i = \sum_{j=1}^{I} G_i(c^*_{ij}) \int_0^{c^*_{ij}} E[\Pi_{ij}(c)]\mu_i(c)dc \tag{22}$$

where $E[\Pi_{ij}(c)]$ is defined in (20), and $\mu_i(c)$ is the distribution of cost draws conditional on c being below the threshold c^*_{ij}. In particular, $\mu_i(c) = \frac{g_i(c)}{G_i(c^*_{ij})}$ if $c < c^*_{ij}$ and zero otherwise. Free entry implies that firms' expected profits equal the fixed cost of entry:

$$\pi_i = w_i f_E \tag{23}$$

3.3 Equilibrium

We compute the mass of firms N_{ij} that sell to a destination j by integrating the probabilities over the white area (labeled "Production") of Figure 2:

$$N_{ij} = N_i \int_0^{c_{ij}^*} \int_0^{\beta_{ij}^*(c)} dB_c(\beta) dG_i(c) \tag{24}$$

The total revenues from i to j are:

$$T_{ij} = N_i \int_0^{c_{ij}^*} \int_0^{\beta_{ij}^*(c)} R_{ij}(c, \beta) b_c(\beta) g_i(c) d\beta dc \tag{25}$$

where the firm's revenues are defined in (17). Goods market clearing implies that:

$$\sum_{i=1}^{I} T_{ij} = w_j L_j \tag{26}$$

In equilibrium, free entry drives expected profits equal to the fixed cost of entry (23), goods markets clear (26), and trade is balanced, namely $\sum_{j=1}^{I} T_{ij} = \sum_{j=1}^{I} T_{ji}$.

Model Extension. In our baseline model, the fixed cost per variety of a firm in a destination is constant for all varieties. In an extension, we allow the fixed cost per variety to be variety-specific and also to depend on the total number of varieties sold, to reflect some economies or diseconomies of scope. We show that these two extensions do not change the ability of our model to match the stylized facts. In fact, to generate the disconnect between sales and scope, the model still requires a firm-destination specific shock to the fixed costs. Furthermore, under proper parametrization, we show that the scope of firms in this extension is isomorphic to that of our baseline model, which is an important result for the following quantitative section.

4 Estimation

Next we parametrize the model and use a simulated method of moments (SMM) algorithm to estimate the parameters controlling the distribution of productivity and flexibility. We choose a SMM estimator in order two deal with the two random processes, in productivity and flexibility, that determine firms' performance. Furthermore, the identification of firms' flexibility and its underlining characteristics through alternative estimators such as OLS is problematic due to possible selection effects.

4.1 Parametrization of the Model

To derive the model's predictions, we choose the following functional forms for the distributions of productivity, flexibility, and within-firm marginal costs. Following Helpman et al. (2004) and Chaney (2008), the inverse of a firm's productivity c is drawn from a Pareto distribution with CDF $G_i(c) = \left(\frac{c}{\bar{c}_i}\right)^{\kappa}$, where $c \in [0, \bar{c}_i]$, κ is the shape parameter, common across all countries, and $\bar{c}_i$ is an origin-specific location parameter.

The distribution of the fixed cost shock β follows a Pareto distribution with CDF $B_c(\beta) = \left(\frac{\beta}{\beta_m c^{\alpha}}\right)^{\gamma}$, where $\beta \in [0, \beta_m c^{\alpha}]$, γ is the shape parameter and $\beta_m c^{\alpha}$ is a firm-specific location parameter. We chose a Pareto distribution to replicate the observed distribution of scope conditional on firms' productivity (Figure C.3). The parameter α controls whether or not firms with high c are likely to get higher levels of β. The expected value of β, conditional on c, is:

$$E[\beta|c] = \int_0^{\beta_m c^{\alpha}} \frac{\gamma \beta^{\gamma}}{(\beta_m c^{\alpha})^{\gamma}} = \frac{\gamma \beta_m c^{\alpha}}{\gamma + 1} \tag{27}$$

If $\alpha = 0$, the fixed cost shock and the cost draw are unrelated, as the expected value of β is independent of c. If $\alpha > 0$, the fixed cost shock is positively correlated with the cost draw, and, thus, productivity and flexibility are positively correlated. In fact, higher levels of c are associated with higher levels of $E[\beta|c]$. If $\alpha < 0$, productivity and flexibility are negatively correlated, since firms with high cost draw c have lower expected realization for β.[19]

The firm-specific location parameter is a shorthand that captures alternative ways with which firms' characteristics affect firms' flexibility. Our model can be thought of as a reduced-form expression of He (1992), in which firms can choose between an output-specific or a flexible technology. If the cost for the flexible technology is constant, more productive firms are endogenously more flexible too. In contrast, if choosing the flexible technology reduces the efficiency of production, the opposite case would occur. The presence of diseconomies of scope á la Nocke and Yeaple (2014) can be represented by $\alpha < 0$ while the presence of economies of scope by $\alpha > 0$.[20]

Finally, our marginal cost per variety ω is $h(\omega) = \exp(\theta\omega)$, with $\theta > 0$, which has the desired properties of $h(0) = 1$ and $h'(\omega) > 0$. The implication of such functional form is that firms with wider scope respond less to trade shocks than narrow-scope firms, which is consistent with

[19] Aside from the timing, the distribution of productivity and flexibility resemble a bivariate Pareto distribution. If c and β were to be drawn from a bivariate Pareto distribution with shift parameters b_c and b_β and shape parameter κ, the pdf of c is given by $f(c) = \frac{\kappa b_c}{c^{\kappa+1}}$ and the conditional pdf of β is $f(\beta|c) = \frac{(\theta+1)\left(\frac{b_\beta}{b_c}\beta\right)^{\kappa+1}}{\left(c+\frac{b_\beta}{b_c}\beta - b_\beta\right)^{\kappa+1}}$. Hence, our formulation is more general in that we allow for the shape parameter to vary across the two distributions of productivity and flexibility. Furthermore, while there is a positive correlation in a bivariate Pareto, we can leave the sign of the correlation unspecified with the parameter α.

[20] The model by Arkolakis et al. (2021) features fixed costs per variety which are declining in the firm's scope, and, thus, features economies of scope within the destination. Although their model features shocks to the fixed costs, these shocks are assumed to be *i.i.d.*. Thus, the source of economies of scope in their model comes from the technological assumption on the fixed costs and not on a relationship between fixed costs shocks and productivity.

our evidence. We derive the closed-form expressions for scope and revenues, and the equilibrium conditions for the parametrized model in appendix D.1.

4.2 Estimation Strategy

Our goal is to estimate the set of parameters $\Theta = [\theta, \sigma, \gamma, \alpha, \beta_m, \kappa]$, which fully characterizes the technology of multi-product firms and the distribution of productivity and flexibility across firms. Given the candidate parameters Θ, we design an algorithm that simulates the revenues and scope across destinations of $1,000,000$ hypothetical Chinese firms. We focus on the eighteen most popular destinations for Chinese exporters. To maximize the number of observations we use for the simulation, we condition our sample to contain only exporters that enter a given reference destination r. We use $r = US$, namely the firms that export to the US, and $r = China$, namely firms that sell in the domestic economy (or all Chinese exporters).[21] The algorithm finds the values of Θ that minimize the weighted sum of squared difference between the simulated vector of moments to the vector of moments in the data.

We follow Arkolakis et al. (2021) and focus on 118 moments, which are divided into five sets:

1. We consider moments from the distribution of sales, within-firm and within-destination. The distribution of sales within-firm is regulated by firms' technology and by consumers' preferences. Higher elasticity of marginal costs with respect to the distance of a variety from the core (θ) causes product sales to decline faster the farther they are from the core. A similar relationship arises between product sales and the elasticity of substitution σ.

2. We match moments from the distribution of the sales of core products within-destination. As we focus on the core products, the distribution of sales is independent of the ability of firms to expand their scope. Absent any heterogeneity in flexibility, the distribution of the core products sales would reflect the distribution of productivity and the degree of substitutability between products. The presence of heterogeneity in flexibility affects the extensive margin of the distribution of core sales.

3. We include the distribution of exporters' scope within-destination. In single-attribute models, such a distribution would only depend on the distribution of firms' productivity and on the distribution of within-firm's productivity, which we targeted in the previous two sets of moments. In our framework, the scope distribution within-destination is affected both by the degree of dispersion of the shock to firms' flexibility and by the correlation between flexibility and productivity.

4. The fourth set of moments takes into account the distribution of firms' total sales within-destination. Combined with the third set of moments, the fourth set exploits the observed

[21] We focus on exporters, as our dataset does not cover the domestic scope of Chinese firms.

disconnect between sales and scope, which is explained by the presence of heterogeneity in flexibility. Without such a heterogeneity, given the distribution of productivity and scope, the distribution of sales would be easily determined. However, in our framework, there is a disconnect between sales and scope that is informative of the heterogeneity in flexibility.

5. The fifth set of moments considers the distribution of firms' scope across destinations. Absent any flexibility shocks, controlling for firms' and destinations' characteristics, the distribution of scope across destinations should be identical across firms. Any deviations from the predictions of single-attribute models are captured by the shock to firms' flexibility.

Our estimation procedure closely follows Arkolakis et al. (2021). The detailed description of the algorithm, moments, and the statistical inference are provided in Appendix E.

4.3 Estimation Results

Table 4 shows the estimated parameters for the full sample of manufacturing firms. All parameters are statistically significant. Moreover, a regression of the calibrated moments against the moments in the data yields an overall R^2 of 80%, and 45% to 97% across the five sets of moments we consider. Details are in Appendix E.5. The model performs well in matching the moments 2-5, which involve the distribution of sales and scope within and across destinations. The result is reassuring, as the two sources of heterogeneity we model are aimed at matching these distributions. The model is less successful in matching the distribution of sales within a firm: there is more variation in the data than our mechanical core competence assumption allows for. We use both the United States and China as the reference countries, and the estimates remain similar.

The shape parameter of the distribution of the fixed cost γ approximately equals 3.1, which suggests a large degree of dispersion in the shock. Without heterogeneity in the fixed cost per variety and destination, such a parameter would be assuming larger values. Moreover, the correlation between the productivity and flexibility distribution is negative: more productive firms are more likely to draw unfavorable draws of flexibility. The result suggests that, to the extent by which flexibility and productivity are endogenous choices of firms, there is a trade-off between the ability to increase volumes at low costs and the ability to expand the scope at low costs.

As the distribution of sales within a firm is controlled by the parameters $\theta(\sigma-1)$, the relatively small value of θ suggests that the skewness of within-firm sales is mainly driven by the demand rather than the marginal cost schedule of firms. In other words, although the marginal cost rises relatively slowly with the scope, the differences in demand are magnified by the demand elasticity of substitution.

Table 4: Estimation of Θ (SMM)

	Pooling Sample Estimation		
	Description \Reference Country	(1) United States	(2) China
$\hat{\kappa}$	Shape par. of productivity distr	1.011***	1.088***
		(0.032)	(0.031)
$\hat{\alpha}$	Correlation productivity fixed cost	-0.744**	-0.742***
		(0.291)	(0.200)
$\hat{\beta}_m$	Shift par. of fixed cost distr.	1.160**	1.111***
		(0.476)	(0.690)
$\hat{\gamma}$	Shape par. of fixed cost distr.	3.119**	4.899***
		(1.147)	(0.690)
$\hat{\theta}$	Elas. of m.cost with distance from core	0.205***	0.131***
		(0.068)	0.013
$\hat{\sigma}$	Elasticity of substitution	1.812***	1.826***
		(0.222)	(0.142)

Bootstrapped standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

To understand the role played by the parameters that control the distribution of flexibility, we re-estimate our model under alternative parameters restrictions (see Table E.1 for the goodness of fit). First, we consider the case of independence between flexibility and productivity ($\alpha = 0$); the details on the model fit are reported in Table E.1. In such a case, the model's ability to match the moments from the data worsens across all five sets of moments, and particularly so for the distribution of within-firm sales. The absence of a negative correlation between productivity and flexibility fails to replicate the presence of firms with large sales and narrow scope uncovered in the data.

Second, we consider the role of the shape parameter of the distribution of flexibility γ. Although in both the case where we double and halve γ, respectively, the goodness of fit of the model declines according to Table E.1, the case with higher γ is the least able to replicate our stylized facts. In fact, higher γ implies lower dispersion in the shock to the fixed cost. Such a lower dispersion causes the disconnect between sales and scope to dwindle, as the relationship between the two variables becomes less noisy and more linear. Furthermore, as in the previous case of $\alpha = 0$, the model fails to generate large firms with low scope.[22]

Robustness. We estimate the parameters of our baseline model at the industry-level. As our moments require the presence of firms selling a sufficient number of products (32 in our case), we focus on the following eight industries out of fifteen: *Vegetable Products*, *Foodstuffs*, *Metals*, *Transportation*, *Chemicals & Allied Industries*, *Machinery & Electrical*, *Textile*, and *Miscellaneous*. Results are provided in Table E.2. Given that the sample size of each industry becomes much

[22]In a previous version of the manuscript, we examined a variance decomposition of firms scope and revenues to assess how much of the variation is explained by the various parameters of the model. Details are available upon request.

more limited, and that firms exporting a sufficient number of products also become less ubiquitous within each industry, the by-industry parameters are less precisely estimated relative to our baseline specification, particularly so for Foodstuffs and Transportation.

The shape of the distributions of productivity and flexibility, and their correlation, exhibit a large degree of heterogeneity across industries. In particular, γ ranges between 1.6 and 4.5, which highlights higher dispersion in the flexibility shock at the industry-level relative to the sample of all manufacturing firms. The correlation between productivity and flexibility, captured by the parameter α, ranges between -1.4 and -0.6. Though for many industries we cannot reject the null hypothesis that $\alpha = 0$, we detect the statistical significance for *Machinery & Electrical* and *Miscellaneous* industries. The prevalent negative sign on α suggests that more productive firms are likely drawing unfavorable levels of flexibility.

In Appendix F, we show that our estimation results are robust to allowing a correlation of the fixed costs shocks across destinations with similar demand. This assumption reflects the possibility that a firm that receives a favorable fixed cost shock in one destination may learn of its flexibility in other similar destinations. The model performance remains similar to our baseline case and the correlation between productivity and flexibility remains negative and attains a larger value in absolute terms. The result suggests that, although learning may play a role in the scope decisions of firms, it is not driving the results of the paper.

4.4 Performance Across Models

To get a comprehensive overview, we compare the quantitative performance of our model relative to the standard model, as well as to others that feature multiple layers of heterogeneity across firms. In particular, we estimate the parameters of four additional models and report their goodness of fit in Table 5. First, we consider the standard model with one dimension of heterogeneity. As highlighted in section 2, such a model cannot match both the distribution of scope and of sales. In fact, the standard model is able to adequately match the scope distribution (R^2 between model and data of 87%), but it is not able to match the distribution of sales of the core products and of firms (R^2 of 35-54% against the 90% of our model). Furthermore, the standard model cannot replicate the distribution of product scope within a firm across destination: the standard model predicts that the ratio of the scope across two destinations is identical across firms, while in our baseline model, it depends on the realization of the flexibility shock.[23]

There are two additional models that, in theory, can account for the disconnect between sales and scope. As a result, we rely on the estimation of the model to verify if the two models can match

[23] A possible concern is that, by adding parameters, we are mechanically increasing the fit of the model. To address the concern, in the bottom panel of Table 5, we compare models based on the D-test as introduced by Newey and West (1987). Specifically, suppose one model can be expressed as a special (or restricted) case of the other (or unrestricted) model. One can perform a statistical comparison that looks very much like a likelihood ratio test. According to the statistics, we reject the null hypothesis that the standard model with only productivity difference performs better in fitting data than our baseline model or the model with ex-post demand shocks.

the disconnect under reasonable parameter assumptions. The first one is the presence of ex-post product-specific demand shocks as in Arkolakis et al. (2021). As we showed in the previous section, heterogeneity in demand shifters across firm products does not affect the disconnect as long as the firm internalizes these demand shifters. However, if these demand shocks are realized upon consumption, when the scope decision of the firm is sunk, two firms with the same scope can have different revenues. This can be interpreted as a short run model for the scope decisions of firms[24] and, provided enough dispersion in the demand shocks, can rationalize part of the disconnect. In fact, comparing the performance of our baseline model and a model with ex-post demand shocks shows that the latter performs better in the moments involving the sales distribution, while the former performs better in the moments regarding the scope distribution. However, since the scope predictions of the model with ex-post demand shocks are identical to the standard model, this case cannot match the distribution of within-firm scope.

Table 5: Models' Goodness of Fit

Moments	(1) Baseline Model	(2) Standard Model	(3) Ex-Post Demand Shocks	(4) NY Model
M1 - Within-Firm & Destination Sales Distr.	0.45	0.35	0.64	-
M2 - Within-Destination Core Sales Distr.	0.90	0.32	0.96	-
M3 - Within-Destination Scope Distr.	0.97	0.84	0.78	0.97
M4 - Within-Destination Sales Distr.	0.90	0.54	0.95	0.50
M5 - Within-Firm Scope Distr.	0.80	0	0	-
M1-5 - All Moments	0.80	0.15	0.74	0.51
Hypothesis Testing:				
H_0: Standard Model; H_a: Baseline Model		D-statistics = 72.35, reject H_0 at 99% level		
H_0: Standard Model; H_a: Ex-Post Demand Shocks		D-statistics = 94.63, reject H_0 at 99% level		

R^2 from regressing the moments predicted by the model against the moments from the data. The list of moments M1-5 is provided in the main text. Hypothesis testing is based on χ^2 difference statistics introduced by Newey and West (1987), who call it the "D-test". The critical value at 99% confidence for $\chi^2(1)$ is 6.64.

An alternative framework is that of Nocke and Yeaple (2014) (NY). In their model, firms differ along two dimensions: organizational efficiency, which is related to our model's productivity, and organizational capital, which is related to our model's flexibility.[25] Although the two sources of

[24] If, in the long run, firms adjust their scope according to the realized demand shocks (or if firms learn these demand shocks in the spirit of Timoshenko (2015)), the problem of the firm becomes the one previously outlined, in which only shocks to the fixed cost can explain the disconnect between sales and scope.

[25] Firms allocate their capital across varieties to reduce the costs of production, and the ability with which capital reduces costs is a function of the organizational capital. As a result, scope is proportional to capital and weakly decreasing in productivity: more efficient firms focus on fewer products.

heterogeneity can generate a disconnect between sales and scope, the model cannot fully replicate the patterns in the data. As shown in column (4) of Table 5, the NY model can match the scope distribution within a destination (M3), but it cannot successfully match the sales distribution within a destination (M4).[26] The main reason for the inability of the NY model to match the data is for firms to have large sales and low scope, they need to attain a level of productivity which is close to the maximum theoretical level. Such an event tends to be rare for reasonable distributional assumption, and it is at odds with the data in which a large set of firms produce few products with large sales.

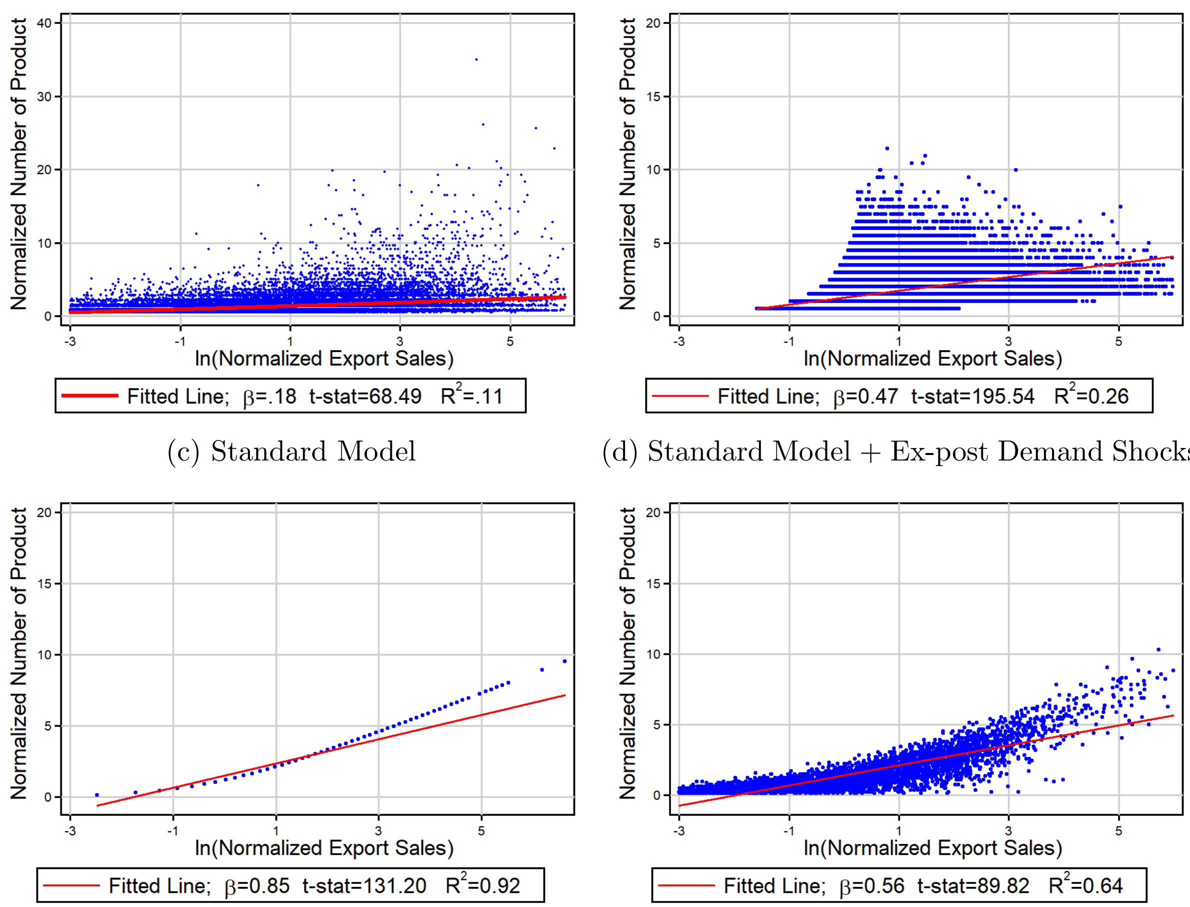
Figure 3: Sale and Scope Disconnect

(a) Data

(b) Baseline Model

(c) Standard Model

(d) Standard Model + Ex-post Demand Shocks



Disconnect Between Sales and Scope. To graphically compare the different models, in Figure 3, we present the predicted relationship between sales in our baseline model, a standard model, and a standard model with ex-post demand shocks. The figure clearly shows that a standard model

[26]Since the estimation of the NY model is restricted to two sets of moments (M3 and M4), we estimate our baseline model again limiting ourselves to these sets of moments. The R^2 of a regression of simulated moments on the data is 0.91. Details are provided in the online appendix or upon request.

cannot replicate the evidence. However, if we simply focus on this figure, it appears that both our baseline model and a model with ex-post demand shocks play a complementary role in explaining the data. In particular, our model is able to match the behavior of larger firms, which can have both a wide and a narrow number of products. However, our model is not able to match the presence of a large number of very small firms. The model with ex-post demand shocks behaves in the opposite way. In fact, such a model is able to match the smallest firms, which receive negative demand shocks, and still remain in the market despite the losses, because of the sunk nature of the scope decisions. In contrast, such a model fails to match the scope of the largest firms.

Because of this result, a possible concern is that the presence of demand shocks can conflate our estimates of firm flexibility. Hence, we estimate a version of our model that includes ex-post demand shocks. Including demand shocks improves the goodness of fit of the model in regards to moments targeting the sales distribution. The model is able to better match the distribution of sales within a firm.[27] However, despite the improvement in performance, we find that the main parameter of interest, α, the correlation between productivity and flexibility remains negative and has a somewhat larger magnitude (-0.9;-1.1). The parameter capturing the dispersion of the shock to the fixed costs remains similar to the baseline case (3.1;3.5). The only difference relative to the baseline model is the shape parameter of the distribution of marginal costs κ, which assumes higher values (2.6;3.1).

The reason why there is a higher density of points in Panels (c) and (d) is that these models predict more dispersed scope distribution than our baseline model. As a result, normalizing the scope of simulated firms by the destination average generates a denser plot relative to our baseline model. Notice that, in Panel (a), the high density of points in the data is due to the normalization across industries and not to the high dispersion of scope. In fact, our model performs better in matching the scope distribution (M3 in Table 5). For instance, in the data, the share of firms that export more than 4 products in a destination is 23% on average. Our baseline model predicts a value of 19%, while models represented in Panels (c) and (d) overestimate the value to 54% and 58%.

4.5 Other Moments

We further present model fit in terms of other data moments to see whether the introduction of firm-destination specific shock to the fixed costs comes at the cost of missing important features of the data.[28] The model can replicate the disconnect observed in the data between productivity

[27]In appendix Table E.1, we report the model fit of the baseline model with demand shocks. For each model setting (i.e., models with $\alpha = 0$, high and low γ, NY model, extended model with demand shocks), we provide detailed model derivations and calibration results in the online appendix.

[28]Based on the baseline parameters as provided in column 1 of Table 4, we simulate the behavior of $1,000,000$ single-industry firms exporting to the eighteen major destinations. We normalize the scope for the active exporters by the destination average. Specifically, for each exporter f in country c, its demeaned variable (i.e., $\exp(\widetilde{\ln x_{fc}})$) is computed in such a way that $\widetilde{\ln x_{fc}} = \ln x_{fc} - \overline{\ln x_{fc}}$, where $\overline{\ln x_{fc}}$ denotes the arithmetic mean of $\ln x_{fc}$ across

and scope. In Panel (a) of Figure E.2, we measure productivity as the reciprocal of the marginal cost draw c, and in Panel (b), we measure productivity as value added per worker.[29]

Figure E.3 shows the distribution of the scope, conditional on the realization of c, which follows a distribution that is similar to the distribution observed in the data as displayed in Figure C.3: in each quartile, the majority of firms export a single-product, and the distribution of scope conditional on productivity resembles a Pareto distribution. We should note that a model with ex-post demand shocks cannot replicate this pattern.

Figure E.4 displays the distributions of scope and sales generated by our model, which are consistent with what we observe in the data. In single-attribute models with CES preferences, the distribution of sales follows the distribution of productivities (Mrázová et al., 2021). Hence, if productivity is Pareto distributed, so are the sales. In our model, the two Pareto distributions of productivity and flexibility generate a distribution of sales that resembles a log-normal distribution. To understand the result, consider the firms at the bottom of the productivity distribution. As the Pareto distribution has a large mass at the bottom, there is a large mass of less productive firms with low sales. The shock to the fixed cost per variety causes some of these firms to not produce and others to produce more varieties than a single-attribute model would predict. Both channels reduce the mass of firms at low sales and shift it towards higher levels of revenues.

5 Welfare Implications: US-China Trade Tension

As gains from varieties and products make up a large portion of the gains from trade, taking into account heterogeneity in flexibility can have major welfare implications. To demonstrate this, we study the welfare effects of the trade tension between the US and China in recent years. In particular, we consider the effects of an increase in the iceberg trade costs from the US to China by 12.9% and from China to the US by 17.0%, which corresponds to the average increase in the tariffs in the years 2017-2019 (Benguria et al., 2020). Since tariffs are part of the iceberg trade costs, it is reasonable to model the tariff war as an increase in τ, although we should note that we are abstracting from changes in government revenues.

The multi-country nature of our model is important for two reasons. First, bilateral changes in trade costs between the US and China affect the welfare of other countries as well. For instance, because of the reductions of Chinese exports to the US, Mexican firms may export more to the US. We find that both Hong Kong and Mexico benefit from the trade tension and their welfare changes are maximized in our baseline model, relative to the single-attribute model. Second, ignoring the presence of multiple countries may overestimate the welfare costs associated with higher trade costs, as it may overestimate the changes in terms of trade.

To compute the welfare changes due to the change in trade costs, we use the exact hat algebra

all firms within country c.

[29] Theoretical derivations of value added per worker across exporters are provided in Appendix D.1.2.

approach introduced in the trade literature by Arkolakis et al. (2012). This procedure allows to exactly measure changes in welfare given changes in trade costs, and it has parsimonious data and parameters requirements. Specifically, the welfare change ($\hat{W}_j$) for country j resulting from the US-China trade tension tariffs can be derived as:[30]

$$\hat{W}_j = \hat{\lambda}_{jj}^{-\frac{\alpha\gamma+\tilde{\gamma}}{\tilde{\gamma}\kappa}}$$

where λ_{jj} is the domestic expenditure share of country j and $\hat{x} = \frac{x_{new}}{x_{old}}$. Given the parameters we estimated in the previous section, we can predict the welfare effects of changes in trade costs for a wide range of counterfactual changes in trade costs, provided with data on countries' size and on the matrix of bilateral expenditure shares. We gather population, average income, GDP, export and import information across countries from World Development Indicators (WDI), and bilateral trade flows that were harmonized and generously provided by the UC Davis Center for International Data. For consistency with the firm data used for calibration, we use the year 2006.

Applying the exact hat algebra, we estimate that the increase in trade costs generates a reduction in welfare in China of 0.1% and in the US by almost 0.08%. In appendix Table F.1, we report the change in trade shares and welfare for all countries. We compare the welfare predictions of our baseline model with the calibrated standard model. As shown in Figure 4, in both cases, the standard model underestimates the welfare losses as it predicts a reduction in welfare by about 0.03-0.05% instead of 0.08-0.1%.[31] These results further motivate and confirm the important role of the second layer of heterogeneity across firms in flexibility in understanding the welfare consequence of trade.

The reason for such different results lies in the ability of our model to match the presence of large firms with narrow scope. Since the elasticity of a firm's scope and revenues with respect to trade costs is declining in the firm scope, it is not surprising that narrow-scope firms are the ones most affected by the shocks and that experience the largest reduction in scope and sales. In standard models, narrow-scope firms are also those with the smallest revenues and, thus, that have smaller aggregate implications. By contrast, in our model, narrow-scope firms can also have large revenues and, thus, have larger aggregate implications.

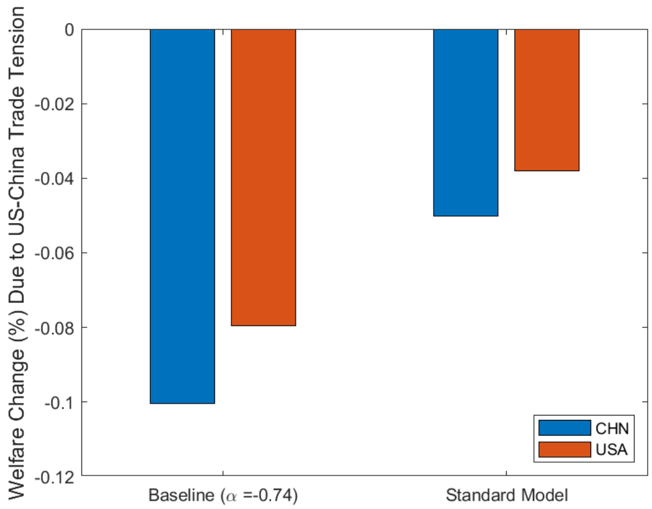
6 Conclusions

We have argued that single-attribute models, in which productivity is the only source of firms' heterogeneity, fail to explain several new stylized facts for multi-product exporters, such as the quantitative relevance of firm-destination specific shocks to the scope. Furthermore, the standard model cannot match the presence of narrow- and wide-scope firms at any level of sales. This has

[30] Detailed theoretical derivations and data information are provided in Appendix G.

[31] In appendix Table F.2, we report the change in trade share and welfare in the standard model.

Figure 4: Welfare Effects of Increase in Trade Costs



The figure shows the percentage change in welfare for the US and China after an increase in the iceberg trade costs from the US to China by 23.9% and from China to the US by 25.3%. The changes in welfare are computed by use of the hat algebra introduced by Arkolakis et al. (2012).

important implications as we find that the scope conditional on sales is an important predictor of the response of exporters to trade liberalization.

To rationalize the evidence, we have proposed a parsimonious model with an additional layer of heterogeneity across firms in flexibility - the ability with which they introduce new varieties at low costs. Our modeling choice is supported by our theoretical and quantitative work which have ruled out alternative frameworks with additional dimensions of heterogeneity, such as to product quality, as they cannot match the empirical evidence.

Failing to account for the new layer of heterogeneity in flexibility has large quantitative implications for the welfare costs of increases in trade costs. Applying our model to quantify the welfare costs of the US-China trade tensions, we find that the losses are twice as large in our model as they are in single-attribute models. This result is driven by the ability of our model to match the presence of large firms with narrow scope. These firms are quite responsive to trade shocks, and since they account for a large share of total exports, they magnify the aggregate effects of trade shocks.

References

C. Arkolakis, A. Costinot, and A. Rodriguez-Clare. New trade models, same old gains? *American Economic Review*, 102(1):94–130, 2012.

C. Arkolakis, S. Ganapati, and M.-A. Muendler. The extensive margin of exporting products: A firm-level analysis. *forthcoming in American Economic Journal: Macroeconomics*, 2021.

R. Armenter and M. Koren. Economies of Scale and the Size of Exporters. *Journal of the European Economic Association*, 13(3):482–511, 2015.

F. Benguria, J. Choi, D. L. Swenson, and M. Xu. Anxiety or pain? the impact of tariffs and uncertainty on chinese firms in the trade war. Technical report, National Bureau of Economic Research, 2020.

A. B. Bernard, J. B. Jensen, S. J. Redding, and P. K. Schott. Firms in International Trade. *Journal of Economic Perspectives*, 21(3):105–130, 2007.

A. B. Bernard, S. J. Redding, and P. K. Schott. Multiproduct Firms and Trade Liberalization. *The Quarterly Journal of Economics*, 126(3):1271–1318, 2011.

L. Brandt, J. Van Biesebroeck, and Y. Zhang. Challenges of working with the Chinese NBS firm-level data. *China Economic Review*, 30:339–352, 2014.

T. Chaney. Distorted Gravity: The Intensive and Extensive Margins of International Trade. *American Economic Review*, 98(4):1707–21, 2008.

I. Cherkashin, S. Demidova, H. L. Kee, and K. Krishna. Firm heterogeneity and costly trade: A new estimation strategy and policy experiments. *Journal of International Economics*, 96(1): 18–36, 2015.

M. Dai, M. Maitra, and M. Yu. Unexceptional exporter performance in China? The role of processing trade. *Journal of Development Economics*, 121:177–189, 2016.

S. Das, M. J. Roberts, and J. R. Tybout. Market Entry Costs, Producer Heterogeneity, and Export Dynamics. *Econometrica*, 75(3):837–873, 2007.

S. Demidova, H. L. Kee, and K. Krishna. Do trade policy differences induce sorting? Theory and evidence from Bangladeshi apparel exporters. *Journal of International Economics*, 87(2):247 – 261, 2012.

B. C. Eaton and N. Schmitt. Flexible Manufacturing and Market Structure. *American Economic Review*, 84(4):875–88, 1994.

J. Eaton, S. Kortum, and F. Kramarz. An Anatomy of International Trade: Evidence From French Firms. *Econometrica*, 79(5):1453–1498, 2011.

C. Eckel and J. P. Neary. Multi-Product Firms and Flexible Manufacturing in the Global Economy. *Review of Economic Studies*, 77(1):188–217, 2010.

B. Erten and J. Leight. Exporting out of agriculture: The impact of WTO accession on structural transformation in China. *Review of Economics and Statistics*, pages 1–46, 2019.

R. C. Feenstra, Z. Li, and M. Yu. Exports and credit constraints under incomplete information: Theory and evidence from China. *Review of Economics and Statistics*, 96(4):729–744, 2014.

A. M. Fernandes, P. J. Klenow, S. Meleshchuk, M. D. Pierola, and A. Rodrıguez-Clare. The intensive margin in trade: moving beyond pareto. Policy research working paper 8625 washington, d.c. : World bank group, 2018.

E. Gal-Or. Flexible manufacturing systems and the internal structure of the firm. *International Journal of Industrial Organization*, 20(8):1061 – 1096, 2002.

H. He. Investments in flexible production capacity. *Journal of Economic Dynamics and Control*, 16(3):575–599, 1992.

E. Helpman, M. J. Melitz, and S. R. Yeaple. Export versus fdi with heterogeneous firms. *American Economic Review*, 94(1):300–316, 2004.

H. L. Kee and K. Krishna. Firm-Level Heterogeneous Productivity and Demand Shocks: Evidence from Bangladesh. *American Economic Review*, 98(2):457–62, May 2008.

J. Levinsohn and A. Petrin. Estimating Production Functions Using Inputs to Control for Unobservables. *The Review of Economic Studies*, 70(2):317–341, 2003.

Q. Liu and H. Ma. Trade policy uncertainty and innovation: Firm level evidence from China's WTO accession. *Journal of International Economics*, 127:103387, 2020.

L. Macedoni. Large multiproduct exporters across rich and poor countries: Theory and evidence. *Working Paper*, 2017.

K. Manova and Z. Yu. How firms export: Processing vs. ordinary trade with financial frictions. *Journal of International Economics*, 100:120–137, 2016.

T. Mayer, M. J. Melitz, and G. I. P. Ottaviano. Market Size, Competition, and the Product Mix of Exporters. *American Economic Review*, 104(2):495–536, 2014.

M. J. Melitz. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6):1695–1725, 2003.

P. Milgrom and J. Roberts. The Economics of Modern Manufacturing: Technology, Strategy, and Organization. *American Economic Review*, 80(3):511–528, June 1990.

G. Mion, L. D. Opromolla, and A. Sforza. The Diffusion of Knowledge via Managers' Mobility. *CEPR Discussion Paper 11706*, 2016.

M. Mrázová, P. Neary, and M. Parenti. Sales and markup dispersion: theory and empirics. *forthcoming in Econometrica*, 2021.

W. K. Newey and K. D. West. Hypothesis testing with efficient method of moments estimation. *International Economic Review*, pages 777–787, 1987.

D. X. Nguyen. Demand uncertainty: Exporting delays and exporting failures. *Journal of International Economics*, 86(2):336–344, 2012.

V. Nocke and S. Yeaple. Globalization and multiproduct firms. *International Economic Review*, 55(4):993–1018, 2014.

J. E. Rauch. Networks versus markets in international trade. *Journal of International Economics*, 48(1):7–35, 1999.

M. J. Roberts, D. Y. Xu, X. Fan, and S. Zhang. The Role of Firm Factors in Demand, Cost, and Export Market Selection for Chinese Footwear Producers. *Review of Economic Studies*, 85(4), October 2018.

D. Thesmar and M. Thoenig. From Flexibility to Insecurity: How Vertical Separation Amplifies Firm-Level Uncertainty. *Journal of the European Economic Association*, 5(6):1161–1202, December 2007.

O. Timoshenko. Product switching in a model of learning. *Journal of International Economics*, 95(2):233–249, 2015.

M.-C. Tseng. Strategic choice of flexible manufacturing technologies. *International Journal of Production Economics*, 91(3):223 – 227, 2004.

J. Wang, F. Cheng, and L. Yang. Smooth simultaneous confidence bands for cumulative distribution functions. *Journal of Nonparametric Statistics*, 25(2):395–407, 2013.

M. Yu. Processing trade, tariff reductions and firm productivity: evidence from Chinese firms. *The Economic Journal*, 125(585):943–988, 2015.

Appendix

A Summary Statistics for ASIP and Customs Data

Table A.1: Match Statistics Between ASIP and China Custom Data

Year	# Matched Firms	% of Total Number	% of Total Export
2000	16,596	35.32%	20.91%
2001	19,597	37.44%	24.41%
2002	23,112	37.51%	27.17%
2003	27,333	35.49%	26.35%
2004	41,955	41.52%	37.95%
2005	44,212	35.62%	35.61%
2006	50,648	33.42%	36.28%

The matching criteria are in the main text of the paper.

Table A.2: Summary Statistics (2006)

Product Scope	Number of Firms		Export Value	
	N	% of total	Value	% of total
All Exporters				
1	34,879	23.02	26,624.71	6.40
2	21,728	14.34	23,346.14	5.61
3	14,860	9.81	21,976.69	5.28
4	10,870	7.17	18,912.32	4.54
5	8,496	5.61	15,538.85	3.73
6 − 10	22,380	14.77	53,065.52	12.75
11 − 30	21,391	14.12	79,534.55	19.10
31 − 50	5,290	3.49	31,152.95	7.48
51 − 70	2,982	1.97	17,115.56	4.11
> 70	8,658	5.71	129,058.06	31.00
Total	151,534	100	416,325.3	100
Matched Sample				
1	10,892	21.51	13,197.05	8.74
2	7,946	15.69	12,402.61	8.21
3	5,936	11.72	11,797.43	7.81
4	4,462	8.81	10,838.26	7.18
5	3,528	6.97	8,911.25	5.90
6 − 10	9,186	18.14	29,835.19	19.75
11 − 30	7,406	14.62	37,467.43	24.81
31 − 50	934	1.84	10,839.69	7.18
51 − 70	219	0.43	4,691.00	3.11
> 70	139	0.27	11,051.66	7.32
Total	50,648	100.00	151,031.57	100.00

A product is a HS eight-digit code. Export value is in million of U.S. dollars.

Table A.3: Classification of Industries

Industry Description	Range of HS 2-digit
Animal & Animal Products	01 - 05
Vegetable Products	06 - 15
Foodstuffs	16 - 24
Mineral Products	25 - 26
Chemicals & Allied Industries	28 - 38
Plastics & Rubbers	39 - 40
Raw Hides, Skins, Leather & Furs	41 - 43
Wood & Wood Products	44 - 49
Textile	50 - 63
Footwear & Headgear	64 - 67
Stone & Glass	68 - 71
Metals	72 - 83
Machinery & Electrical	84 - 85
Transportation	86 - 89
Miscellaneous	90 - 97

Table A.4: Share of Exporters by Destination

Destination	Share of Exporters
United States	23.90%
Japan	17.78%
Hong Kong	16.78%
South Korea	15.29%
Germany	13.99%
Italy	12.17%
United Kingdom	11.27%
India	11.14%
Netherlands	11.12%
Taiwan	9.82%
Russia	9.75%
United Arab Emirates	9.34%
Canada	9.25%
Spain	8.99%
Australia	8.71%
Singapore	8.40%
Indonesia	6.93%
France	6.54%

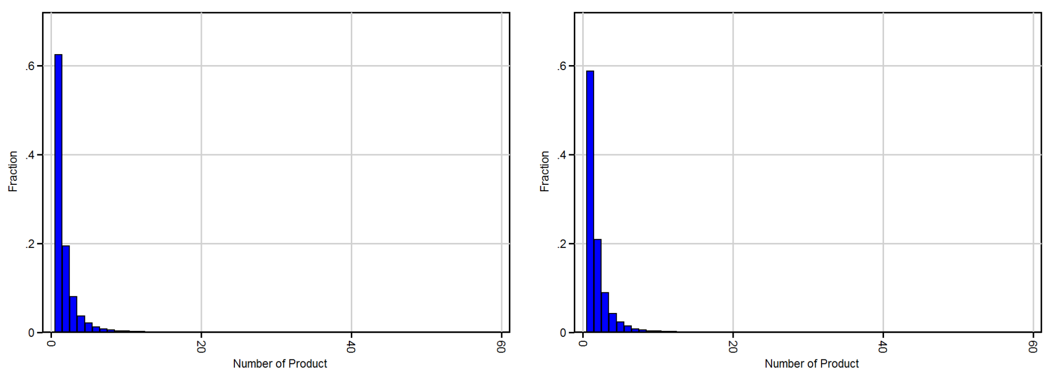
Matched sample. Probability of being exporter in a destination is the ratio of the number of exporters in that destination relative to the total number of matched domestic firms in 2006.

Chinese multi-product firms dominate the country's exports: 77% of exporters sell at least two products, that is two HS eight-digit codes, in a destination, and they account for 94% of total export value. Such results are in line with the evidence documented for several other countries by Bernard et al. (2007), Mayer et al. (2014), Arkolakis et al. (2021), and Macedoni (2017). Our sample of matched firms exhibits a similar distribution: 78% of the firms in the sample are multi-product, and they account for 96% of the sample's total exports.

As our main empirical results focus on the distribution of firms' scope, sales, and productivity within a destination, we focus on the US market, the most popular destination for Chinese exports. Figure A.1 illustrates the distribution of the number of HS eight-digit goods per firm that Chinese exporters sell to the US. The first graph uses all exporters exporting to the US while the second focuses on the sample of firms with matched characteristics from ASIP. The two distributions are remarkably similar as they both exhibit the largest mass for a scope of one. In particular, 36% of all Chinese exporters exporting to the US and 40% of the matched sample export a single HS eight-digit good to the US. However, matched firms have a wider scope, on average, than the firms in the entire sample.

Figure A.1: Distribution of Product Scope in the US (2006)

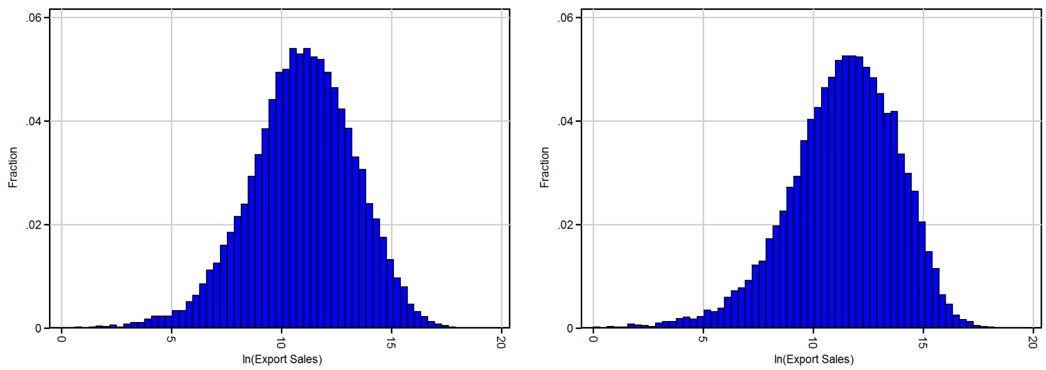
(a) All Sample (b) Matched Sample



While the scope distribution across firms resembles a Pareto distribution, sales appear to be log-normally distributed. Figure A.2 shows the distribution of log sales within the US, for the entire sample of firms and the matched sample. Although the average sale is larger for the matched sample, the two distributions look virtually identical. We test for the similarity of the CDF of the sale and scope distribution between the entire sample of exporters and the sample of matched firms. In particular, we apply the kernel Kolmogorov-Smirnov test (Wang et al., 2013) to a bootstrap sample of firms, which we repeat 50 times. In more than 90 % of tests, we fail to reject the hypothesis that the two distributions are different. Details are in the online appendix.

Figure A.2: Distribution of ln(Export Sales) in the US (2006)

(a) All Sample (b) Matched Sample



B Productivity Estimation

We use Levinsohn and Petrin (2003) method to estimate firm productivity. This method is applied to our dataset of Chinese firms in ASIP as follows. Let y_{it} denotes the value added of firm i in year t, and the production function is

$$\begin{aligned} y_t =& \beta_0 + \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \omega_{it} + \epsilon_{it} \\ =& \beta_l l_{it} + \phi(k_{it}, m_{it}) + \epsilon_{it} \end{aligned}$$

where $\phi(k_{it}, m_{it}) \equiv \beta_0 + \beta_k k_{it} + \omega(k_{it}, m_{it})$. We substitute $\phi(k_{it}, m_{it})$ in equation of y_{it} with a third-order polynomial approximation in k_{it} and m_{it}, after which we are able to obtain the consistent estimation of the value-added equation using OLS as

$$y_{it} = \delta_0 + \beta l_{it} + \sum_{s=0}^{3} \sum_{j=0}^{3-s} \delta_{sj} k_{it}^s m_{it}^j + \epsilon_{it}$$

This provides the first stage of estimation, from which we obtain $\hat{\beta}_l$ and an estimate of $\hat{\phi}_{it}$.

Next, we estimate $\hat{\beta}_k$, which begins by computing the estimated value of $\hat{\phi}_{it}$:

$$\hat{\phi}_{it} = \hat{y}_{it} - \hat{\beta}_l l_{it} = \hat{\delta}_0 + \sum_{s=0}^{3} \sum_{j=0}^{3-s} \hat{\delta}_{sj} k_{it}^s m_{it}^j - \hat{\beta}_l l_{it}$$

For any candidate value of β_k^*, we compute the prediction for ω_{it} as $\hat{\omega}_{it} - \hat{\phi}_{it} - \beta_k^* k_{it}$. We assume a consistent (nonparametric) approximation to $E(\widehat{\omega_{it}|\omega_{it-1}})$ as

$$\hat{\omega}_{it} = \gamma_0 + \gamma_1 \omega_{t-1} + \gamma_2 \omega_{t-1}^2 + \gamma_3 \omega_{t-1}^3 + e_{it}$$

Given $\hat{\beta}$, β_k^* and $E(\widehat{\omega_{it}|\omega_{it-1}})$, we rewrite the sample residuals as

$$\widehat{\epsilon_{it} + e_{it}} = y_{it} - \hat{\beta}_l l_{it} - \beta_k^* k_{it} - E(\widehat{\omega_{it}|\omega_{it-1}})$$

Finally, we estimate $\hat{\beta}_k$ as the solution to:

$$\min_{\beta_k^*} \sum_i \sum_t \left(y_{it} - \hat{\beta}_l l_{it} - \beta_k^* k_{it} - E(\widehat{\omega_{it}|\omega_{it-1}}) \right)^2$$

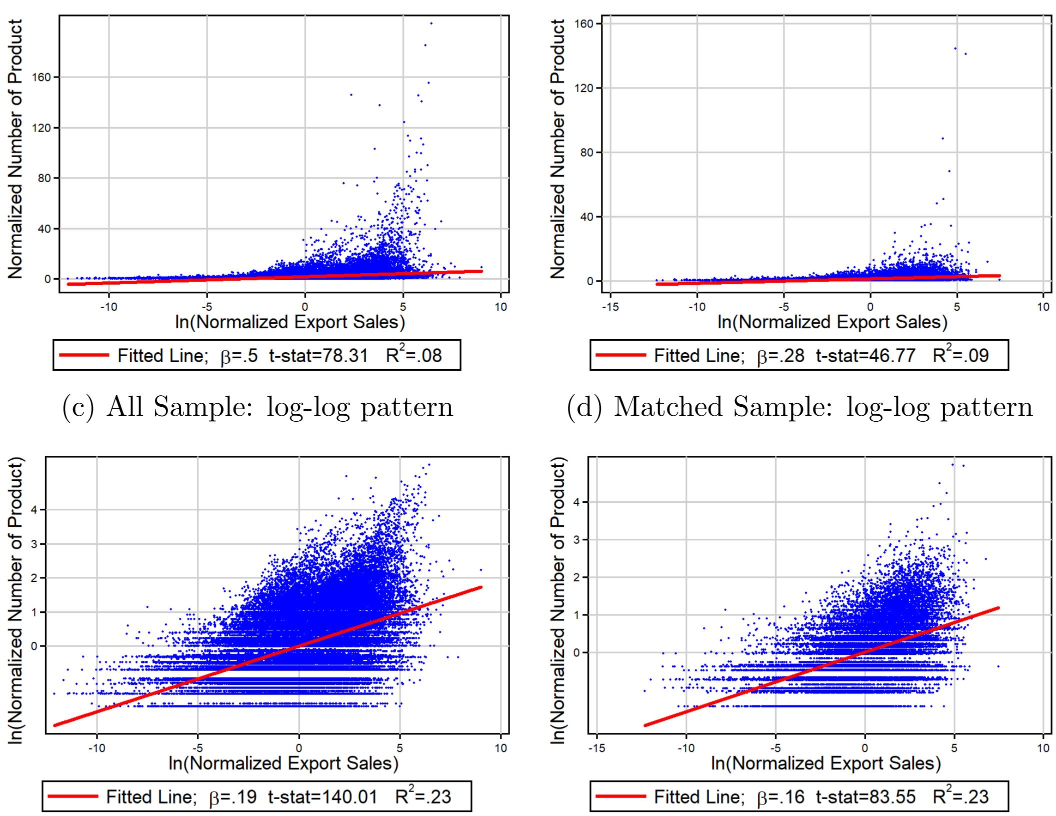
In practice, we use data of ASIP for 2005 and 2006. We measure value added as outcome variable y_{it}, employment as the freely chosen variable (l_{it}), total fixed asset as capital (k_{it}), and intermediate input and the sales cost as the proxy variables.[32]

C Empirical Robustness

C.1 Stylized Facts: Sales and Scope Disconnect

Figure C.1: Product Scope and Sales of Exporters in the US with Multi-industry Exporters (2006)

(a) All Sample (b) Matched Sample



[32] The sales cost for a firm is the cost of merchandise in its beginning inventory plus the net cost of merchandise purchased minus the cost of merchandise in its ending inventory. We try different specifications using different variables to measure the m_{it}, and the estimated productivity are highly positively correlated.

Figure C.2: Product Scope and Characteristics of Exporters in the US (2006)

(a) Productivity Measure (b) Value-added Per Worker

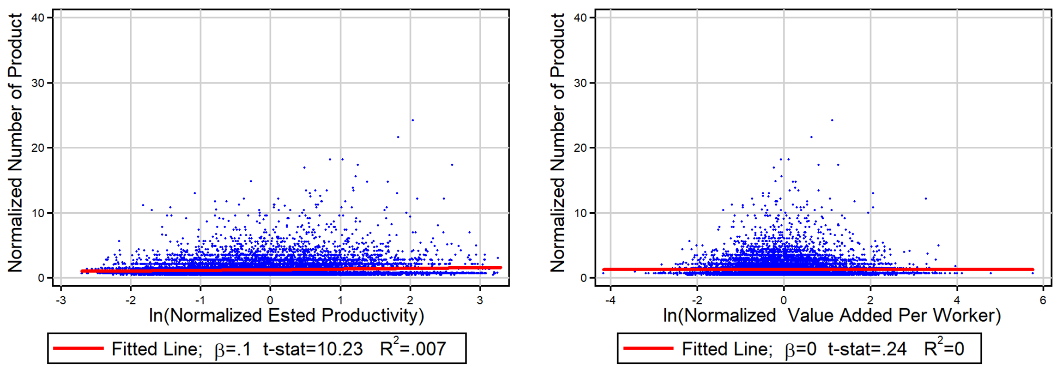


Figure C.3: Scope Distribution by Productivity Quartiles (2006)

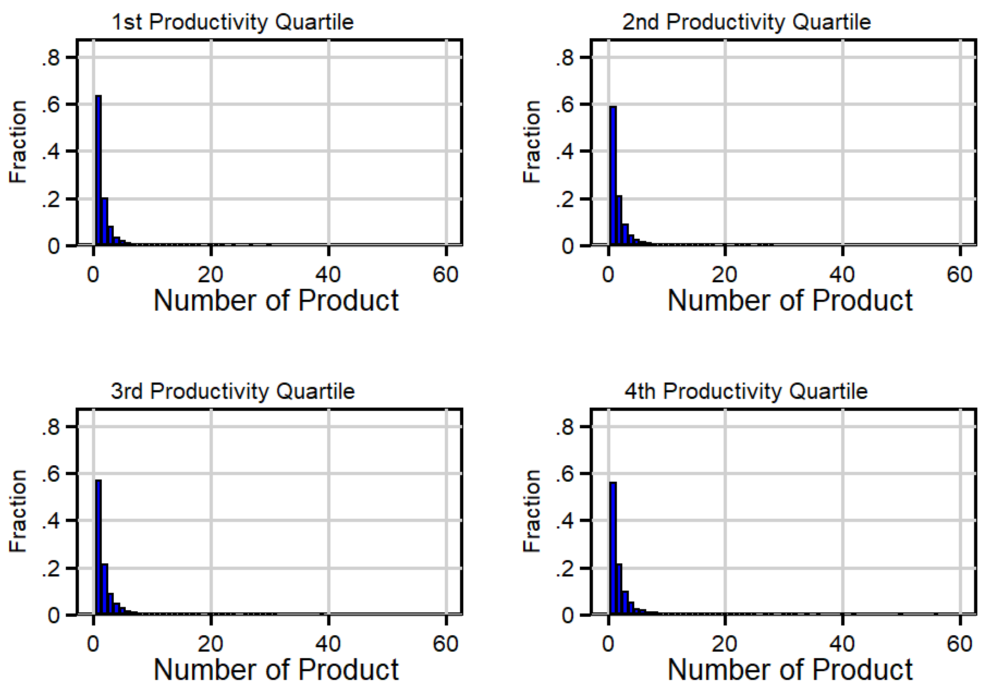
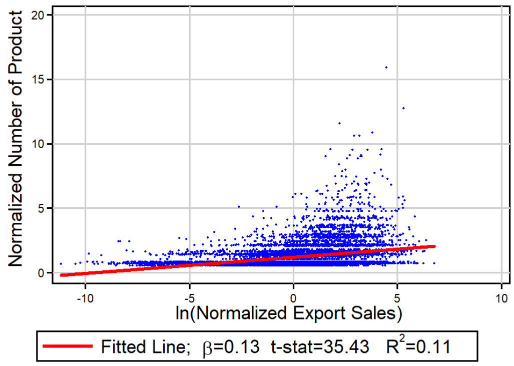
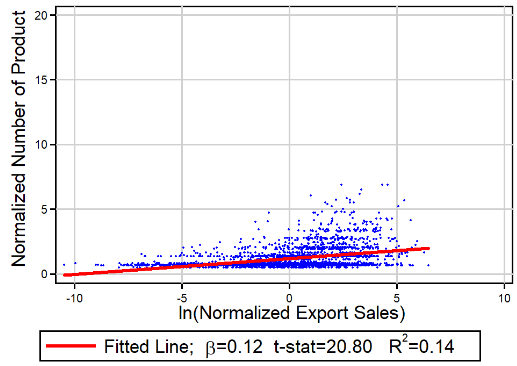


Figure C.4: Product Scope and Sales of Exporters in the US: *Alternative Years*

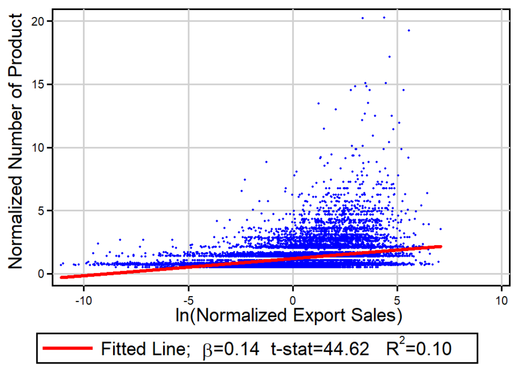
(a) All Sample: 2000



(b) Matched Sample: 2000



(c) All Sample: 2003



(d) Matched Sample: 2003

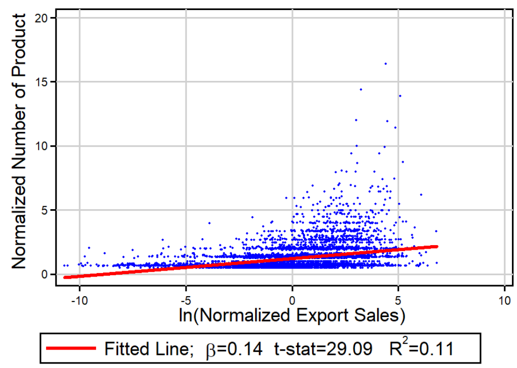
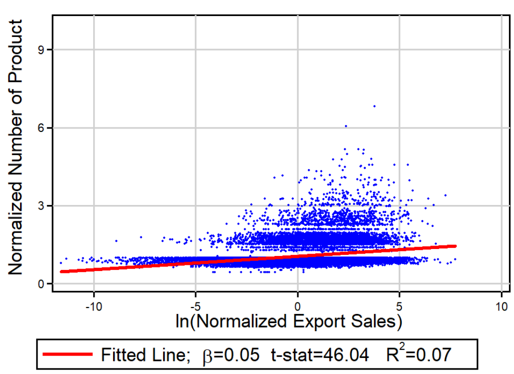


Figure C.5: Product Scope and Sales of Exporters in the US (2006): *Define an Industry by HS 4-digit Code*

(a) All Sample



(b) Matched Sample

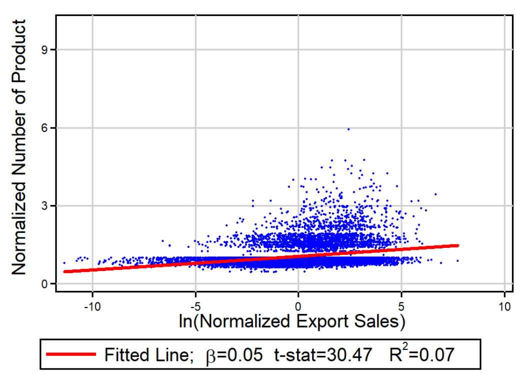
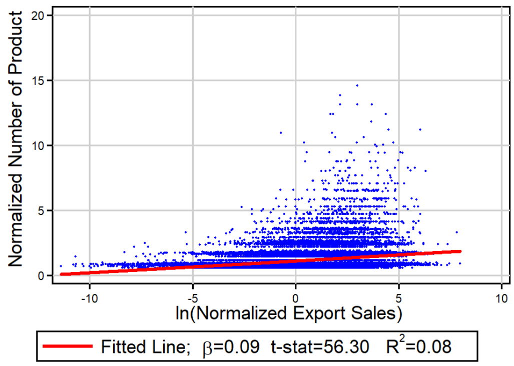
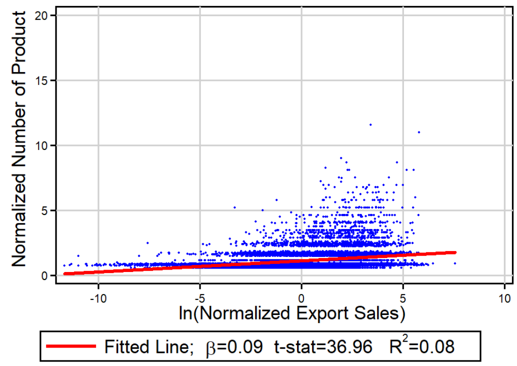


Figure C.6: Product Scope and Sales of Exporters in the US (2006): *Define a Product by HS 4- and 6-digit Codes*

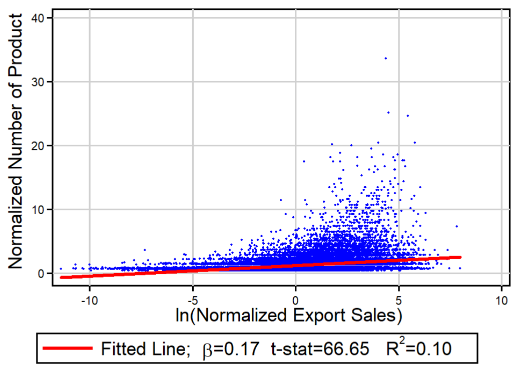
(a) All Sample: HS 4-digit Product



(b) Matched Sample: HS 4-digit Product



(c) All Sample: HS 6-digit Product



(d) Matched Sample: HS 6-digit Product

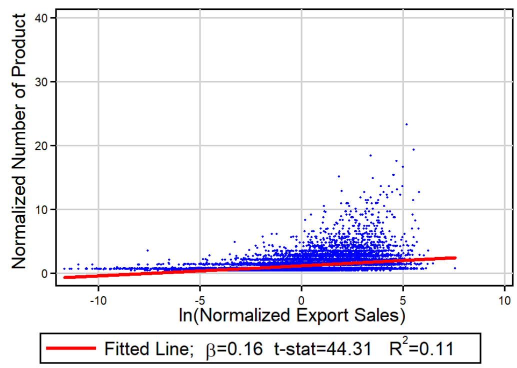
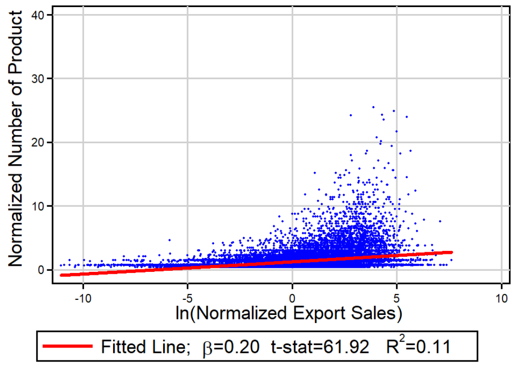
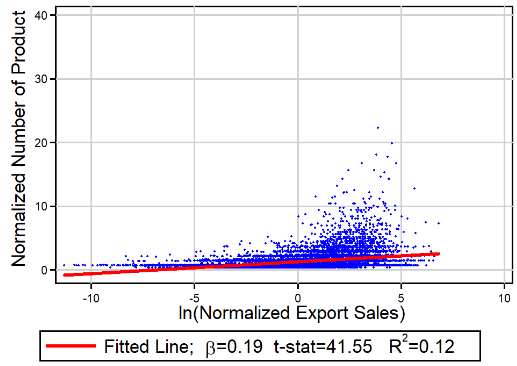


Figure C.7: Product Scope and Sales of Exporters in *Other Countries* (2006)

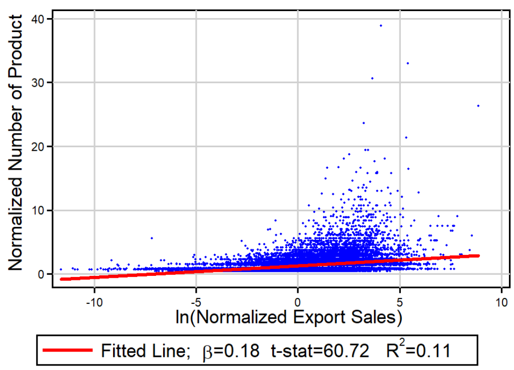
(a) All Sample: Japan



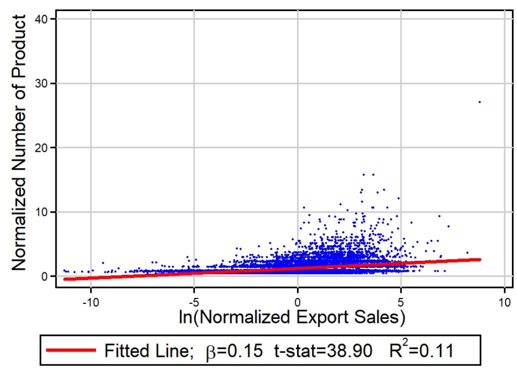
(b) Matched Sample: Japan



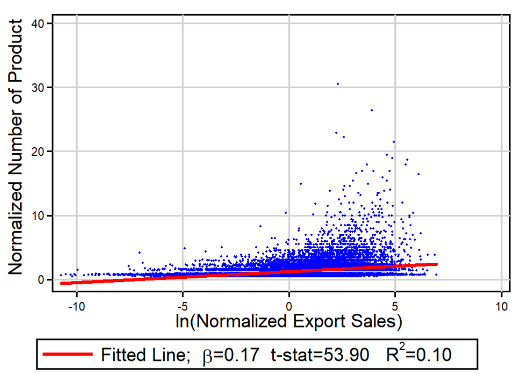
(c) All Sample: South Korea



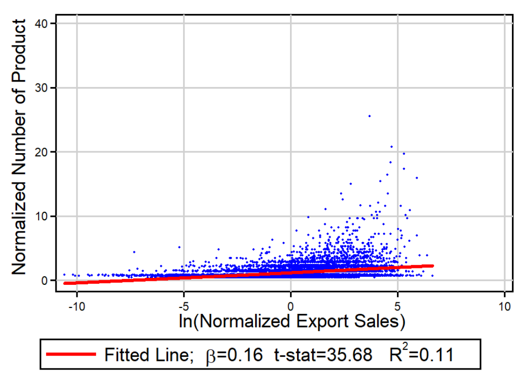
(d) Matched Sample: South Korea



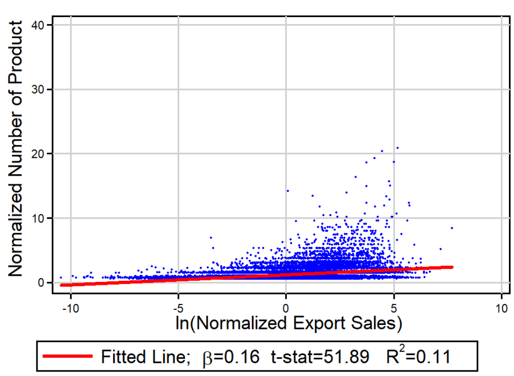
(e) All Sample: Germany



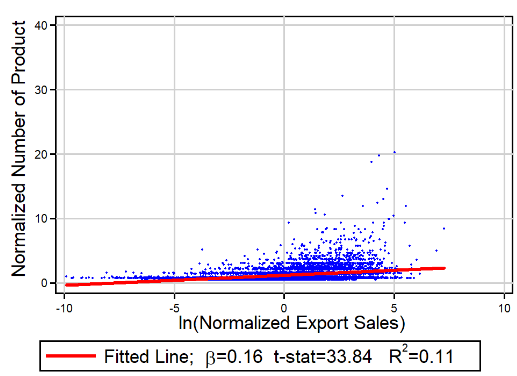
(f) Matched Sample: Germany



(g) All Sample: United Kingdom



(h) Matched Sample: United Kingdom



C.2 Effects of Trade on Scope

Figure C.8: Tariff Reduction Over Time

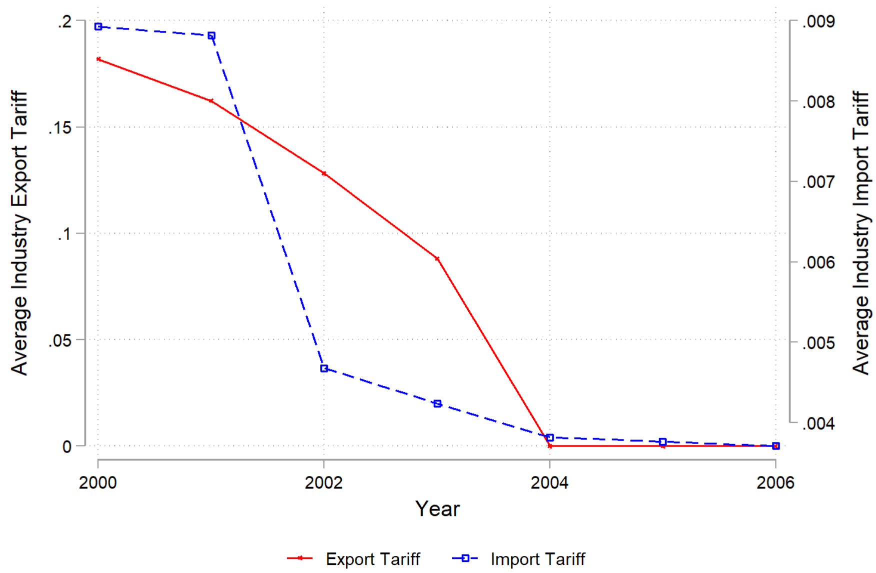


Table C.1: Robustness: Sector-Year Specific Tariff

Dep var: ln(# Products$_{ft}$)	(1)	(2)	(3)	(4)	(5)
Tariff^{EX}_{jt}	-0.537***	-1.878***	-0.891***	-0.809**	-1.000***
	(0.096)	(0.108)	(0.345)	(0.349)	(0.359)
$\text{Tariff}^{EX}_{jt} \cdot \ln(\#\ \text{Products}_{f0})$		1.107***	1.191***	1.186***	1.184***
		(0.051)	(0.057)	(0.058)	(0.058)
$\text{Tariff}^{EX}_{jt} \cdot \ln(\text{Sales}_{f0})$			-0.080***	-0.087***	-0.070***
			(0.026)	(0.026)	(0.027)
Tariff^{IM}_{jt}	0.251	-0.810**	-0.802*	-1.015**	-1.199***
	(0.426)	(0.409)	(0.409)	(0.414)	(0.422)
Observations	50,717	50,717	50,717	49,252	48,909
R-squared	0.812	0.818	0.818	0.820	0.829
Firm controls	-	-	-	Y	Y
Firm FE	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
City-Year FE	-	-	-	-	Y

Note: for detailed notes, see Table 3. Because we use a different way of measuring the trade shock (Tariff^{EX}_{jt}) from that ($\text{Post01}_t \times \text{Tariff}^{EX}_{j0}$) in (11), we obtain the opposite sign patterns from those expressed in the Table 3. However, the interpretations of the sign remain the same and robust.

D Theory Appendix

D.1 Model's Derivation

This section shows how to derive the closed-form expression of our parametrized model. Let $\tilde{\theta} = \theta(\sigma - 1)$. Using $h(\omega) = \exp(\theta\omega)$ into (16) yields:

$$\begin{aligned} h(\delta_{ij}(c,\beta))^{\sigma-1} &= \left(\frac{\beta^*_{ij}(c)}{\beta}\right) \\ \exp(\tilde{\theta}\delta_{ij}(c,\beta)) &= \left(\frac{\beta^*_{ij}(c)}{\beta}\right) \qquad (28) \\ \delta_{ij}(c,\beta) &= \frac{1}{\tilde{\theta}}\left[\ln \beta^*_{ij}(c) - \ln \beta\right] \qquad (29) \end{aligned}$$

Using (28), the firm's cost aggregator $H(\delta_{ij}(c,\beta))$ is given by:

$$H(\delta_{ij}(c,\beta)) = \left[\int_0^{\delta_{ij}(c,\beta)} h(\omega)^{1-\sigma} d\omega\right]^{\frac{1}{1-\sigma}} = \left[\frac{1}{\tilde{\theta}} - \frac{\exp(-\tilde{\theta}\delta_{ij}(c,\beta))}{\tilde{\theta}}\right]^{\frac{1}{1-\sigma}} = \left[\frac{1}{\tilde{\theta}} - \frac{\beta}{\beta^*_{ij}(c)\tilde{\theta}}\right]^{\frac{1}{1-\sigma}} \qquad (30)$$

Our sales and scope disconnect (18) becomes:

$$\begin{aligned} \left(\frac{h(\delta_{ij}(c,\beta))}{H(\delta_{ij}(c,\beta))}\right)^{\sigma-1} &= \frac{R_{ij}(c,\beta)}{\sigma w_j f_j \beta} \\ \frac{R_{ij}(c,\beta)}{\exp(\tilde{\theta}\delta_{ij}(c,\beta)) - 1} &= \frac{\sigma w_j f_j \beta}{\tilde{\theta}} \qquad (31) \end{aligned}$$

The revenues of a variety ω of firm with cost draw c and fixed cost draw β is given by:

$$r(\omega, c, \beta) = \sigma w_j f_j \beta^*_{ij}(c) h(\omega)^{1-\sigma} = \sigma w_j f_j \beta^*_{ij}(c) \exp(-\tilde{\theta}\omega) \qquad (32)$$

Integrating (32) across the varieties exported to a destination, and using (30) yields firm's aggregate revenues in destination j:

$$R_{ij}(c,\beta) = \sigma w_j f_j \beta^*_{ij}(c) H(\delta_{ij}(c,\beta))^{1-\sigma} = \frac{\sigma w_j f_j}{\tilde{\theta}}\left[\beta^*_{ij}(c) - \beta\right] \qquad (33)$$

Using our parametrized distribution for the fixed cost shock, the expected revenues over the possible realizations of β, conditional on c, are given by:

$$E[R_{ij}(c)] = \left(\frac{\beta^*_{ij}(c)}{\beta_m c^\alpha}\right)^\gamma \int_0^{\beta^*_{ij}(c)} R_{ij}(c,\beta)\frac{\gamma\beta^{\gamma-1}}{(\beta^*_{ij}(c))^\gamma} = \frac{\sigma w_j f_j}{\tilde{\theta}(\gamma+1)}\frac{(\beta^*_{ij}(c))^{\gamma+1}}{(\beta_m c^\alpha)^\gamma} \qquad (34)$$

Using (31), (33) and (29), profits of a firm with cost draw c and fixed cost shock β are:

$$\Pi_{ij}(c,\beta) = w_j f_j \beta\left[\left(\frac{h(\delta_{ij}(c,\beta))}{H(\delta_{ij}(c,\beta))}\right)^{\sigma-1} - \delta_{ij}(c,\beta)\right] = w_j f_j \beta\left[\frac{R_{ij}(c,\beta)}{\sigma w_j f_j \beta} - \delta_{ij}(c,\beta)\right] =$$

$$= w_j f_j \beta \left[\frac{1}{\beta\tilde{\theta}} \left[\beta^*_{ij}(c) - \beta\right] - \delta_{ij}(c,\beta) \right] = w_j f_j \beta \left[\frac{1}{\beta\tilde{\theta}} \left[\beta^*_{ij}(c) - \beta\right] - \frac{1}{\tilde{\theta}} \left[\ln \beta^*_{ij}(c) - \ln \beta\right] \right] =$$
$$= \frac{w_j f_j}{\tilde{\theta}} \left[\beta^*_{ij}(c) - \beta - \beta \ln \beta^*_{ij}(c) + \beta \ln \beta\right]$$

Expected profits over the possible realizations of β are given by:

$$E[\Pi_{ij}(c)] = \frac{w_j f_j}{\tilde{\theta}(\gamma+1)^2} \frac{(\beta^*_{ij}(c))^{\gamma+1}}{(\beta_m c^\alpha)^\gamma} \tag{35}$$

Let $\tilde{\gamma} = (\sigma-1)(\gamma+1)$. Setting the expected profits (35) equal to the fixed cost of entry to a destination F_j, and using the definition of $\beta^*_{ij}(c)$ yield the cost cutoff c^*_{ij}:

$$\frac{w_j f_j}{\tilde{\theta}(\gamma+1)^2} \frac{(\beta^*_{ij}(c^*_{ij}))^{\gamma+1}}{(\beta_m (c^*_{ij})^\alpha)^\gamma} = w_j F_j \tag{36}$$

$$\left[\frac{1}{f_j(\sigma-1)} \left[A_j \frac{\sigma-1}{\sigma}\right]^\sigma (\tau_{ij} w_i c^*_{ij})^{-(\sigma-1)} \right]^{\gamma+1} \frac{f_j}{\tilde{\theta}(\gamma+1)^2} \frac{1}{(\beta_m (c^*_{ij})^\alpha)^\gamma} = F_j$$
$$\left[\frac{1}{\sigma-1} \left[A_j \frac{\sigma-1}{\sigma}\right]^\sigma \right]^{\gamma+1} \frac{(\tau_{ij} w_i c^*_{ij})^{-\tilde{\gamma}} f_j^{-\gamma} F_j^{-1}}{\tilde{\theta}(\gamma+1)^2} \frac{1}{(\beta_m (c^*_{ij})^\alpha)^\gamma} = 1$$
$$\left[\frac{1}{\sigma-1} \left[A_j \frac{\sigma-1}{\sigma}\right]^\sigma \right]^{\gamma+1} \frac{(\tau_{ij} w_i)^{-\tilde{\gamma}} f_j^{-\gamma} F_j^{-1}}{\beta_m^\gamma \tilde{\theta}(\gamma+1)^2} = (c^*_{ij})^{\alpha\gamma+\tilde{\gamma}}$$

Thus, the cost cutoff c^*_{ij} equals:

$$c^*_{ij} = (\tau_{ij} w_i)^{-\frac{\tilde{\gamma}}{\alpha\gamma+\tilde{\gamma}}} f_j^{-\frac{\gamma}{\alpha\gamma+\tilde{\gamma}}} F_j^{-\frac{1}{\alpha\gamma+\tilde{\gamma}}} \left[\left[\frac{1}{\sigma-1} \left[A_j \frac{\sigma-1}{\sigma}\right]^\sigma \right]^{\gamma+1} \frac{1}{\tilde{\theta}(\gamma+1)^2 \beta_m^\gamma} \right]^{\frac{1}{\alpha\gamma+\tilde{\gamma}}} \tag{37}$$

The higher the variable or fixed costs of exporting, the lower the cutoff: a firm must have a low draw of c to decide to pay F_j to reach a destination with high trade costs. Conveniently, c^*_{ij} can be written as a function of the domestic cost cutoff c^*_{jj}:

$$c^*_{ij} = c^*_{jj} \left(\frac{\tau_{ij} w_i}{\tau_{jj} w_j} \right)^{-\frac{\tilde{\gamma}}{\alpha\gamma+\tilde{\gamma}}} \tag{38}$$

Moreover, taking the ratio between $\beta^*_{ij}(c)$ and $\beta^*_{ij}(c^*_{ij})$ yields:

$$\beta^*_{ij}(c) = \beta^*_{ij}(c^*_{ij}) \left(\frac{c}{c^*_{ij}} \right)^{-(\sigma-1)} \tag{39}$$

By (36),

$$\beta^*_{ij}(c^*_{ij}) = \left[\frac{\beta_m^\gamma F_j \tilde{\theta}(\gamma+1)^2}{f_j} \right]^{\frac{1}{\gamma+1}} (c^*_{ij})^{\frac{\alpha\gamma}{\gamma+1}} \tag{40}$$

Using (39) and (40) in (35) and (34), the expected profits and revenues can be writte as a function of c and c^*_{ij}:

$$E[\Pi_{ij}(c)] = \frac{w_j f_j}{\tilde{\theta}(\gamma+1)^2} \frac{(\beta^*_{ij}(c^*_{ij}))^{\gamma+1}}{(\beta_m c^\alpha)^\gamma} \left(\frac{c}{c^*_{ij}}\right)^{-\tilde{\gamma}} - w_j F_j = w_j F_j \left[\left(\frac{c}{c^*_{ij}}\right)^{-\tilde{\gamma}-\alpha\gamma} - 1\right] \tag{41}$$

$$E[R_{ij}(c)] = \sigma(\gamma+1) w_j F_j \left(\frac{c}{c^*_{ij}}\right)^{-\tilde{\gamma}-\alpha\gamma} \tag{42}$$

Conditional on entry in a destination j, average profits and revenues are given by:

$$\bar{\pi}_{ij} = \int_0^{c^*_{ij}} E[\Pi_{ij}(c)] \kappa \frac{c^{\kappa-1}}{(c^*_{ij})^\kappa} dc = \frac{w_j F_j(\alpha\gamma + \tilde{\gamma})}{\kappa - \alpha\gamma - \tilde{\gamma}} \tag{43}$$

$$\bar{R}_{ij} = \frac{w_j F_j \kappa \sigma(\gamma+1)}{\kappa - \alpha\gamma - \tilde{\gamma}} \tag{44}$$

Expected profits are then given by:

$$\pi^e_i = \sum_j \left(\frac{c^*_{ij}}{\bar{c}_i}\right)^\kappa \bar{\pi}_{ij} = \frac{(\alpha\gamma + \tilde{\gamma})}{\kappa - \alpha\gamma - \tilde{\gamma}} \sum_j w_j F_j \left(\frac{c^*_{ij}}{\bar{c}_i}\right)^\kappa$$

Setting the expected profits equal to the fixed cost of entry $w_i f_E$ yields:

$$\frac{(\alpha\gamma + \tilde{\gamma})}{\kappa - \alpha\gamma - \tilde{\gamma}} \sum_j w_j F_j \left(\frac{c^*_{ij}}{\bar{c}_i}\right)^\kappa = w_i f_E \tag{45}$$

Total revenues from i to j are given by:

$$T_{ij} = N_i \left(\frac{c^*_{ij}}{\bar{c}_i}\right)^\kappa \bar{R}_{ij} = N_i \left(\frac{c^*_{ij}}{\bar{c}_i}\right)^\kappa \frac{w_j F_j \kappa\sigma(\gamma+1)}{\kappa - \alpha\gamma - \tilde{\gamma}} \tag{46}$$

Thus, market clearing and trade balance imply that:

$$\sum_j T_{ij} = w_i L_i$$

$$N_i \frac{\kappa\sigma(\gamma+1)}{\kappa - \alpha\gamma + \tilde{\gamma}} \sum_j w_j F_j \left(\frac{c^*_{ij}}{\bar{c}_i}\right)^\kappa = w_i L_i \tag{47}$$

Dividing (47) by (45) yields the mass of entrants N_i:

$$N_i = \frac{L_i(\alpha\gamma + \tilde{\gamma})}{f_E \kappa\sigma(\gamma+1)} \tag{48}$$

Revenues from i to j are then given by:

$$T_{ij} = \frac{\alpha\gamma + \tilde{\gamma}}{f_E(\kappa - \alpha\gamma - \tilde{\gamma})} L_i w_j F_j \bar{c}_i^{-\kappa} c^{*\kappa}_{ij} \tag{49}$$

Using (38), the trade share becomes:

$$\lambda_{ij} = \frac{T_{ij}}{\sum_v T_{vj}} = \frac{L_i \bar{c}_i^{-\kappa} (\tau_{ij} w_i)^{-\frac{\kappa\tilde{\gamma}}{\alpha\gamma+\tilde{\gamma}}}}{\sum_v L_v \bar{c}_v^{-\kappa} (\tau_{vj} w_v)^{-\frac{\kappa\tilde{\gamma}}{\alpha\gamma+\tilde{\gamma}}}} \tag{50}$$

By market clearing, $T_{ij} = \lambda_{ij} w_j L_j$. Thus, using (49), the cost cutoff c^*_{ij} is given by:

$$c^*_{ij} = \bar{c}_i \left(\frac{\lambda_{ij} L_j w_j}{L_i w_j F_j} \right)^{\frac{1}{\kappa}} \left[\frac{f_E(\kappa - \alpha\gamma - \tilde{\gamma})}{\alpha\gamma - \tilde{\gamma}} \right]^{\frac{1}{\kappa}} \tag{51}$$

Hence, using (39), (40) and (51), the fixed cost cutoff for a firm with cost draw c is given by:

$$\begin{aligned}
\beta^*_{ij}(c) = \beta^*_{ij}(c^*_{ij}) \left(\frac{c}{c^*_{ij}} \right)^{-(\sigma-1)} &= \left[\frac{\beta_m^\gamma F_j \tilde{\theta}(\gamma+1)^2}{f_j} \right]^{\frac{1}{\gamma+1}} (c^*_{ij})^{\frac{\alpha\gamma}{\gamma+1}} \left(\frac{c}{c^*_{ij}} \right)^{-(\sigma-1)} = \\
&= \left[\frac{\beta_m^\gamma F_j \tilde{\theta}(\gamma+1)^2}{f_j} \right]^{\frac{1}{\gamma+1}} (c^*_{ij})^{\frac{\alpha\gamma+\tilde{\gamma}}{\gamma+1}} c^{-(\sigma-1)} = \\
&= \left[\frac{\beta_m^\gamma F_j \tilde{\theta}(\gamma+1)^2}{f_j} \right]^{\frac{1}{\gamma+1}} \left[\bar{c}_i \left(\frac{\lambda_{ij} L_j w_j}{L_i w_j F_j} \right)^{\frac{1}{\kappa}} \left[\frac{f_E(\kappa - \alpha\gamma - \tilde{\gamma})}{\alpha\gamma - \tilde{\gamma}} \right]^{\frac{1}{\kappa}} \right]^{\frac{\alpha\gamma+\tilde{\gamma}}{\gamma+1}} c^{-(\sigma-1)} = \\
&= \underbrace{\left[\beta_m^\gamma \tilde{\theta}(\gamma+1)^2 \right]^{\frac{1}{\gamma+1}} \left[\frac{f_E(\kappa - \alpha\gamma - \tilde{\gamma})}{\alpha\gamma - \tilde{\gamma}} \right]^{\frac{\alpha\gamma+\tilde{\gamma}}{\kappa(\gamma+1)}}}_{=\bar{B}} \left(\frac{F_j}{f_j} \right)^{\frac{1}{\gamma+1}} \bar{c}_i^{\frac{\alpha\gamma+\tilde{\gamma}}{\gamma+1}} \left(\frac{\lambda_{ij} L_j}{L_i F_j} \right)^{\frac{\alpha\gamma+\tilde{\gamma}}{\kappa(\gamma+1)}} c^{-(\sigma-1)} \\
&= \bar{B} \left(\frac{F_j}{f_j} \right)^{\frac{1}{\gamma+1}} \bar{c}_i^{\frac{\alpha\gamma+\tilde{\gamma}}{\gamma+1}} \left(\frac{\lambda_{ij} L_j}{L_i F_j} \right)^{\frac{\alpha\gamma+\tilde{\gamma}}{\kappa(\gamma+1)}} c^{-(\sigma-1)}
\end{aligned} \tag{52}$$

Using (52) into (29) yields:

$$\begin{aligned}
\delta_{ij}(c,\beta) &= \frac{1}{\tilde{\theta}} \left[\ln \beta^*_{ij}(c) - \ln \beta \right] = \\
&= \frac{\ln \bar{B}}{\tilde{\theta}} + \frac{\alpha\gamma + \tilde{\gamma}}{\tilde{\theta}(\gamma+1)} \ln \bar{c}_i + \frac{1}{\tilde{\theta}(\gamma+1)} \left[\ln \left(\frac{F_j}{f_j} \right) + \frac{\alpha\gamma+\tilde{\gamma}}{\kappa} \ln \left(\frac{\lambda_{ij} L_j}{L_i F_j} \right) \right] - \frac{1}{\theta} \ln c - \frac{1}{\tilde{\theta}} \ln \beta
\end{aligned} \tag{53}$$

D.1.1 Effects of a Reduction in Trade Costs

Let us consider the effects of a change in either variable trade costs τ_{ij} or fixed cost per variety f_j on total firm-level exports from i to j. Let $\epsilon_H(\delta) = -\frac{d \ln H(\delta)}{\delta} > 0$ and $\epsilon_h(\delta) = \frac{d \ln h(\delta)}{\delta} > 0$ be the elasticity of the firm's average productivity, and marginal costs with respect to the scope. The within-firm trade margins can be written as:

$$-\frac{d \ln R_{ij}(c,\beta)}{d \ln f_j} = \underbrace{0}_{\text{Intensive}} + \underbrace{\frac{\epsilon_H(\delta(c,\beta))}{\epsilon_h(\delta(c,\beta))}}_{\text{Extensive}}$$

$$-\frac{d\ln R_{ij}(c,\beta)}{d\ln\tau_{ij}} = \underbrace{\sigma-1}_{\text{Intensive}} + \underbrace{(\sigma-1)\frac{\epsilon_H(\delta(c,\beta))}{\epsilon_h(\delta(c,\beta))}}_{\text{Extensive}}$$

A reduction in the fixed cost per variety and destination f_j only affects the extensive margin of trade while a variation in τ_{ij} both affects the intensive and the extensive margin. Using our functional forms, $\frac{\epsilon_H(\delta(c,\beta))}{\epsilon_h(\delta(c,\beta))} = \Big(\exp(\tilde{\theta}\delta_{ij}(c,\beta))-1\Big)^{-1}$ and, thus, it decreases with firm's scope. The result arises because the scope of firms with wide scope responds less to trade shock. Moreover, the contribution to total sales of additional varieties is smaller the farther away they are from the core. Hence, the wider the current scope the smaller their contribution to the firm's sales.

D.1.2 Value Added Per Worker

In this section, we derive an expression for the value added per worker. Since there are no intermediate inputs, the value added of a firm (c,β) equals its total revenues:

$$VA(c,\beta) = \sum_{j=1}^{I} \mathbf{1}_{x_{ij}(c,\beta)>0} R_{ij}(c,\beta) = \sum_{j=1}^{I} \mathbf{1}_{x_{ij}(c,\beta)>0} \frac{\sigma w_j f_j}{\tilde{\theta}}(\beta^*_{ij}(c)-\beta) \tag{54}$$

To evaluate the employment of the firm, we need to make an assumption about the labor units used for the fixed costs of production $f_j\beta$. If these costs are paid in destination labor units, then the employment of the firm only requires the number of workers used for the payment of the fixed cost in the domestic economy. If, instead, these costs are paid in origin labor units, then the firm's employment also takes into account the fixed cost paid for each product in each destination. Similar to the paper, we consider the case of destination labor units.

This means that the employment of a firm equals:

$$\begin{aligned}
emp(c,\beta) &= \sum_{j=1}^{I} \mathbf{1}_{x_{ij}(c,\beta)>0} \int_0^{\delta_{ij}(c,\beta)} x_{ij}(\omega,c,\beta)c\tau_{ij}h(\omega)d\omega + \int_0^{\delta_{ij}(c,\beta)} w_i f_i \beta d\omega \\
&= \sum_{j=1}^{I} \mathbf{1}_{x_{ij}(c,\beta)>0} \frac{(\sigma-1)w_j f_j \beta^*_{ij}(c)}{\tilde{\theta}\beta w_i}(\beta^*_{ij}(c)-\beta) + w_i f_i \beta \delta_{ij}(c,\beta) = \\
&= \sum_{j=1}^{I} \mathbf{1}_{x_{ij}(c,\beta)>0} \frac{(\sigma-1)w_j f_j}{\tilde{\theta} w_i}(\beta^*_{ij}(c)-\beta) + \frac{w_i f_i \beta}{\tilde{\theta}}(\ln\beta^*_{ij}(c)-\ln\beta)
\end{aligned}$$

Thus, the value added per worker equals:

$$\frac{VA(c,\beta)}{emp(c,\beta)} = \frac{\sum_{j=1}^{I} \mathbf{1}_{x_{ij}(c,\beta)>0} \frac{\sigma w_j f_j}{\tilde{\theta}}(\beta^*_{ij}(c)-\beta)}{\sum_{j=1}^{I} \mathbf{1}_{x_{ij}(c,\beta)>0} \frac{(\sigma-1)w_j f_j}{\tilde{\theta} w_i}(\beta^*_{ij}(c)-\beta) + \frac{w_i f_i \beta}{\tilde{\theta}}(\ln\beta^*_{ij}(c)-\ln\beta)} \tag{55}$$

Our calibration exercise does not target the absolute value of the fixed costs $w_j f_j$. Furthermore, their derivation is not trivial since, typically, procedures can find bounds for the fixed costs, by examining the value of the revenues of the smallest exporter in a destination. Hence, to simplify the analysis without losing generality, we make the normalization that all common components of the fixed costs are equal to one ($w_j f_j = 1$). For robustness, we experiment the case where the fixed cost $w_j f_j$ is randomly drawn from uniform distribution at interval $(0, 1)$, finding similar results. We normalize the wage of workers in China to one.

D.2 Model Extension

In this appendix section, we consider an extension to the baseline model whereby the fixed cost per firm-product also depends on the number of varieties sold and it can be specific to the product. In particular, to introduce a new variety in a destination j, each firm pays a fixed cost $w_j m(\omega) f_j(\delta_{ij}(\boldsymbol{\varphi}))\beta$ where $m(\omega)$ is a component of the fixed cost which is variety specific, and $w_j f_j(\delta_{ij}(\boldsymbol{\varphi}))$ is a function of the scope. All other assumptions of the baseline model still hold. The purpose of this section is to show that this extension does not change the results and that, under proper functional forms, this specification generates a scope of firms which is isomorphic that of the baseline model.

As in the baseline model, we consider the problem of a firm from country i active in destination j. A firm is identified by a vector $\boldsymbol{\varphi} = [z, c, \beta]$, where z is the firm-level demand shifter, c the marginal cost, and β the fixed cost shock. As in the baseline model, there are two firm-product specific variables: $z(\omega)$ and $\tilde{h}(\omega)$. In addition, there is the firm-product specific component of the fixed cost $m(\omega)$.

Profit maximization yields the standard constant markup pricing rule of the baseline model. Profits can be written as:

$$\pi_{ij}(\omega, \boldsymbol{\varphi}) = \frac{1}{\sigma - 1}\left[A_j \frac{\sigma - 1}{\sigma}\right]^{\sigma} (\tau_{ij} w_i)^{1-\sigma} \left(\frac{c}{z}\right)^{1-\sigma} \left(\frac{\tilde{h}(\omega)}{z(\omega)}\right)^{1-\sigma} - w_j m(\omega) f_j(\delta_{ij}(\boldsymbol{\varphi}))\beta$$

Let $h(\omega) = \frac{\tilde{h}(\omega)}{z(\omega)}$. We do not make an assumption on $h(\omega)$ and $m(\omega)$, but require varieties to be ordered by decreasing profit, namely $\frac{\partial \pi_{ij}(\omega, \boldsymbol{\varphi})}{\partial \omega} < 0$. We normalize the variables as follows: $h(0) = m(0) = f_j(0) = 1$. Without loss of generality, we can also normalize z to one, which is equivalent of having c replace c/z. Thus, variety-specific profits can be written as:

$$\pi_{ij}(\omega, \boldsymbol{\varphi}) = \frac{1}{\sigma - 1}\left[A_j \frac{\sigma - 1}{\sigma}\right]^{\sigma} (\tau_{ij} w_i c h(\omega))^{1-\sigma} - w_j m(\omega) f_j(\delta_{ij}(\boldsymbol{\varphi}))\beta$$

We can compute the fixed cost cutoff by setting the profits from the core variety to zero:

$$\beta^*_{ij}(\boldsymbol{\varphi}) = \frac{1}{w_j(\sigma - 1)}\left[A_j \frac{\sigma - 1}{\sigma}\right]^{\sigma} (\tau_{ij} w_i c)^{-(\sigma-1)} \tag{56}$$

Let us rewrite per variety profits as a function of the fixed cost cutoff $\beta^*_{ij}(\boldsymbol{\varphi})$ and β:

$$\pi_{ij}(\omega, \boldsymbol{\varphi}) = w_j \beta^*_{ij}(\boldsymbol{\varphi}) h(\omega)^{1-\sigma} - w_j m(\omega) f_j(\delta_{ij}(\boldsymbol{\varphi}))\beta$$

The optimal scope of the firm is implicitly defined by taking the derivative of total profits $\pi_{ij}(\boldsymbol{\varphi}) = \int_0^{\delta_{ij}(\boldsymbol{\varphi})} \pi_{ij}(\omega, \boldsymbol{\varphi}) d\omega$ with respect to the scope $\delta_{ij}(\boldsymbol{\varphi})$:

$$(h(\delta_{ij}(\boldsymbol{\varphi})))^{\sigma-1}\left(m(\delta) f_j(\delta_{ij}(\boldsymbol{\varphi})) + f'_j(\delta_{ij}(\boldsymbol{\varphi})) \int_0^{\delta_{ij}(\boldsymbol{\varphi})} m(\omega) d\omega\right) = \left(\frac{\beta^*_{ij}(\boldsymbol{\varphi})}{\beta}\right) \tag{57}$$

The right-hand side is increasing in firm's productivity ($1/c$). Imposing restrictions on the fixed cost function, we can assume that the left-hand side is increasing in scope – which is consistent with the general finding that more productive firms produce more varieties. However, because of the shock β, a given scope can be attained by a high productivity firm with low flexibility and by a low productivity firm with high flexibility. Absent any shock to β there would be a one-to-one mapping of productivity into scope.

The total sales of a firm have the same expression of the baseline model:

$$R_{ij}(\boldsymbol{\varphi}) = \sigma w_j \beta^*_{ij}(\boldsymbol{\varphi}) H(\delta_{ij}(\boldsymbol{\varphi}))^{1-\sigma} \tag{58}$$

Using our definition of scope (57), we can rewrite (58) as:

$$\left(\frac{h(\delta_{ij}(\boldsymbol{\varphi}))}{H(\delta_{ij}(\boldsymbol{\varphi}))}\right)^{\sigma-1} \left(m(\delta) f_j(\delta_{ij}(\boldsymbol{\varphi})) + f'_j(\delta_{ij}(\boldsymbol{\varphi})) \int_0^{\delta_{ij}(\boldsymbol{\varphi})} m(\omega) d\omega\right) = \frac{R_{ij}(\boldsymbol{\varphi})}{\sigma w_j \beta} \tag{59}$$

which represents the sales and scope disconnect illustrated in the baseline model. The additional assumptions on the fixed costs do not alter the fact that only a shock to the extensive margin β can generate that two firms with the same sales can have different scope. With additional assumptions on the functional forms, we can show that in two cases the scope is isomorphic that of the baseline model. Recall that we assumed the following functional form for the marginal labor requirement for variety:

$$h(\omega) = \exp(\theta\omega)$$

The scope was defined as:

$$\exp(\tilde{\theta}\delta_{ij}(\boldsymbol{\varphi})) = \left(\frac{\beta^*_{ij}(\boldsymbol{\varphi})}{\beta}\right) \tag{60}$$

where $\tilde{\theta} = \theta(\sigma - 1)$.

Case 1: Heterogeneity in Fixed Costs Across Varieties. Let us assume that the fixed cost is independent of the scope, i.e. $f'() = 0$. Furthermore, let $m(\omega) = \exp(\xi\omega)$. The scope is given by:

$$\exp((\tilde{\theta} + \xi)\delta_{ij}(\boldsymbol{\varphi})) = \left(\frac{\beta^*_{ij}(\boldsymbol{\varphi})}{\beta}\right) \tag{61}$$

which is isomorphic to (60).

Case 2: Fixed Costs as a Function of the Scope. Let $m(\omega) = 1 \ \forall\omega$ and $f_j(\delta_{ij}(\boldsymbol{\varphi})) = \exp(\psi\delta_{ij}(\boldsymbol{\varphi}))$. The scope is given by:

$$(1 + \psi)\exp((\tilde{\theta} + \psi)\delta_{ij}(\boldsymbol{\varphi})) = \left(\frac{\beta^*_{ij}(\boldsymbol{\varphi})}{\beta}\right) \tag{62}$$

which is also isomorphic to (60).

The fact that both under Case 1 and Case 2 the scope of the firms is isomorphic to the baseline one indicates that including these functional forms would mainly affect the estimation of θ and not the distribution of the fixed cost shock.

E Simulated Method of Moments Algorithm and Moments

This section describes in detail the Simulated Method of Moments (SMM) procedure we follow to estimate the set of parameters $\Theta = [\theta, \sigma, \gamma, \alpha, \beta_m, \kappa]$. First, we describe the simulation algorithm, which, for a given set of candidate parameters, generates the behavior of Chinese firms. Second, we list the moments we target. Third, we minimize the deviations of the simulated moments from the moments from the data to estimate the set of parameters Θ.

E.1 Derivations

Before we describe the algorithm, let us derive a few equations that will be used. First, let us consider the mass of firms M_{ij} exporting to a destination j :

$$\begin{aligned}
M_{ij} &= N_i \int_0^{c^*_{ij}} \int_0^{\beta^*_{ij}(c)} g_c(\beta) g(c) d\beta dc = N_i \int_0^{c^*_{ij}} \int_0^{\beta^*_{ij}(c)} \frac{\kappa\gamma}{\beta_m^\gamma \bar{c}^\kappa} \beta^{\gamma-1} c^{\kappa-\alpha\gamma-1} d\beta dc = \\
&= N_i \int_0^{c^*_{ij}} \frac{\kappa}{\beta_m^\gamma \bar{c}^\kappa} (\beta^*_{ij}(c))^\gamma c^{\kappa-\alpha\gamma-1} dc = N_i \int_0^{c^*_{ij}} \frac{\kappa}{\beta_m^\gamma \bar{c}^\kappa} (\beta^*_{ij}(c^*_{ij}))^\gamma c^{\kappa-\alpha\gamma-\gamma(\sigma-1)-1} (c^*_{ij})^{\gamma(\sigma-1)} dc = \\
&= N_i \frac{\kappa}{(\kappa - \alpha\gamma - \gamma(\sigma-1))\beta_m^\gamma \bar{c}^\kappa} (\beta^*_{ij}(c^*_{ij}))^\gamma (c^*_{ij})^{\kappa-\alpha\gamma} = \\
&= N_i \frac{\kappa}{(\kappa - \alpha\gamma - \gamma(\sigma-1))\beta_m^\gamma \bar{c}^\kappa} \left[\frac{\beta_m^\gamma F_j \tilde{\theta}(\gamma+1)^2}{f_j}\right]^{\frac{\gamma}{\gamma+1}} (c^*_{ij})^{\kappa - \frac{\alpha\gamma}{\gamma+1}}
\end{aligned}$$

We can compute the ratio of firms exporting to a destination j from i relative the mass of firms exporting to a reference country r:

$$\frac{M_{ij}}{M_{ir}} = \left[\frac{c^*_{ij}}{c^*_{ir}}\right]^{\kappa - \frac{\alpha\gamma}{\gamma+1}} \tag{63}$$

Let $D_{ij} = (\beta^*_{ij}(c^*_{ij}))(c^*_{ij})^{\sigma-1}$. Then, we can re-write the expression for the firm's scope as:

$$\exp(\tilde{\theta}\delta_{ij}(c,\beta)) = D_{ij} c^{-(\sigma-1)} \beta^{-1}$$

Let $\phi = c^{-(\sigma-1)}\beta^{-1}$ denote the entry-adjusted measure of productivity, a variable that describes the combination of c and β of a firm in a destination. The entry-adjusted productivity cutoff, which is the combination of c and β such that a firm will not produce in a destination, equals:

$$\phi^*_{ij} = D_{ij}^{-1}$$

Hence, the scope of a firm with entry-adjusted productivity ϕ equals:

$$\exp(\tilde{\theta}\delta_{ij}(\phi)) = \frac{\phi}{\phi^*_{ij}}$$

The relationship between entry-adjusted productivity cutoff and marginal cost cutoff is:

$$\frac{\phi^*_{ij}}{\phi^*_{ir}} = \frac{D_{ir}}{D_{ij}} = \frac{(\beta^*_{ir}(c^*_{ir}))(c^*_{ir})^{\sigma-1}}{(\beta^*_{ij}(c^*_{ij}))(c^*_{ij})^{\sigma-1}} = \left[\frac{c^*_{ir}}{c^*_{ij}}\right]^{\frac{\alpha\gamma}{\gamma+1}+\sigma-1} \tag{64}$$

E.2 Simulation Algorithm

Given the candidate parameters Θ, the algorithm simulates the revenues and scope of $J^{sim} = 1,000,000$ hypothetical Chinese firms, for each destination they serve $j = 1, ..., D$. Given that our model features a continuum of firms, we choose a large value for J^{sim} to approximate such a continuum. We set $D = 18$ and consider the most popular destinations for Chinese exporters - which constitute 70% of total exports. The list of eighteen destinations is provided in Table A.4. Table A.3 provides the industry coverage for the whole sample. We refine our samples to firms whose exporting scope is less than 100. The refined sample constitutes 99.5% of the total firm number and 92.75% of the total exports to the 18 countries.

We independently draw J^{sim} realizations of the firms' marginal cost percentiles u^c from the standard uniform distribution. To maximize the number of observations we use for the simulation, we condition our sample to contain only firms that enter a given reference destination r. We use $r = China$, namely firms that sell in the domestic economy and $r = US$, namely the firms that export to the US. Any iteration of the algorithm begins with the computation of the marginal cost draws of firms. Given the marginal cost percentile u^c, we recover the marginal cost of firms as:

$$c = (u^c)^{\frac{1}{\kappa}} c^*_{ir}$$

In other words, we condition our sample only to contain the firms that are productive enough to enter the domestic economy. In practice, we normalize c^*_{ir} to one. We back out the marginal cost cutoff for all other destinations using the ratio of the number of firms selling to a destination j relative to the number of firms selling to r using (63):

$$c^*_{ij} = \left[\frac{M_{ij}}{M_{ir}}\right]^{\frac{1}{\kappa - \frac{\alpha\gamma}{\gamma+1}}} c^*_{ir}$$

For each destination j and each firm c, such that $c \leq c^*_{ij}$ we draw the firm-destination specific flexibility shock percentile $u^\beta(c)_j$. The fixed-cost draw of a firm is then computed as:

$$\beta(c)_j = \beta_m c^\alpha (u^\beta(c)_j)^{\frac{1}{\gamma}}$$

We obtain the entry-adjusted productivity measure (64) as:

$$\phi^*_{ij} = \left[\frac{c^*_{ir}}{c^*_{ij}}\right]^{\frac{\alpha\gamma}{\gamma+1} + \sigma - 1} \phi^*_{ir}$$

We normalize ϕ^*_{ir} to one, thus, simulating the behavior of the firms that, with the same entry-adjusted productivity, would be active in the reference destination. We compute $\phi = c^{-(\sigma-1)}\beta^{-1}$, which is the entry-adjusted measure of productivity. A firm is active in a destination, if $\phi \geq \phi^*_{ij}$. The exporter scope of a firm in a destination is then given by:

$$\delta_{ij}(\phi) = \frac{1}{\tilde{\theta}} \left[\ln \phi - \ln \phi^*_{ij}\right]$$

For some moments we derive later, it is useful to discretize the scope of firms as follows:

$$\delta_{ij}(\phi)^{int} = 1 + int(\delta_{ij}(\phi))$$

where *int* denotes the integer part of the continuous form of the scope. We generate this firm's ω-th ranked product in destination j (with scope $\delta_{ij}(\phi)^{int}$) as[33]

$$r_{\omega ij} = \exp\left[\tilde{\theta}\left(\delta_{ij}(\phi)^{int} - \underbrace{(int(\omega)+1)}_{\text{discretize the product index}}\right)\right] \times \beta$$

where $\tilde{\theta} = \theta(\sigma - 1)$. The above expression on product revenue $r_{\omega ij}$ omits the destination-specific shifters. Both shifters are common across exporters at the same destination and firm invariant in our simulation as we normalize the relevant moments by the corresponding destination-specific median or extreme. The revenues of the firm are given by:

$$R_{ij}(\phi, \beta) = \beta[\exp(\tilde{\theta}\delta_{ij}(\phi)) - 1]$$

or

$$R_{ij}(\phi, \beta) = \sum_{\omega=1}^{\delta_{ij}(\phi)^{int}} r_{\omega ij}$$

E.3 Moments

In this appendix section, we define the moments used in the simulated method of moments algorithm. To separately identify the parameters relevant for the stochastic components (productivity and fixed cost distribution) and the parameters that govern firm's production function, we use the moments so that they are comparable across destinations by neutralizing the destination-specific shifters.

M1: Within-destination and within-firm product sales concentration

The first set of moments compares the sales of a firm's top-selling product and other products' sales of the same firm. Consider a firm with marginal and fixed cost draw (c, β) in a destination j. Let $r_{1ij}(c, \beta)$ denote the revenues of the top selling product of the firm (in our model, $\omega = 0$) in a destination j, and $r_{\tilde{\omega}ij}(c, \beta)$ denote the total revenues generated by the products in group $\tilde{\omega}$ for a firm in a destination j. The log sales ratio between the sales of products in $\tilde{\omega}$ and the core product equals:

$$\ln(r_{1\tilde{\omega}ij}(c, \beta)) = \ln\left(\frac{r_{\tilde{\omega}ij}(c, \beta)}{r_{1ij}(c, \beta)}\right) = -\tilde{\theta}\tilde{\omega} \tag{65}$$

We focus on the multi-product firms who export more than one product. We compute $\ln(r_{1\tilde{\omega}ij}(c, \beta))$ for these firms, and pool the log sales ratio $\ln(r_{1\tilde{\omega}ij}(c, \beta))$ by product group $\tilde{\omega} \in \{\{2, ..., 7\}, \{8, ...31\}, \{32, ...127\}, \{128, ...\}\}$ across firms and destination. For instance, consider $\tilde{\omega} \in= \{2, ..., 7\}$. This set contains the log sales ratio $\ln(r_{1\tilde{\omega}ij}(c, \beta))$ for $\tilde{\omega} \in \{2, ..., 7\}$ across all destinations j and all firms (c, β).

For each product group we rank order the sales ratio $\ln(r_{1\tilde{\omega}ij}(c, \beta))$ from the highest to the lowest. This gives us a collection of ordered sequence for each product group $\{\ln(r_{1\tilde{\omega}(100)}), ..., \ln(r_{1\tilde{\omega}(0)})\}$.

[33] The full expression should be $r_{\omega ij} = \sigma f_{China,j} \exp\left[\tilde{\theta}\big(\delta_{ij}(\phi)^{int} - \omega\big)\right] \times \beta$. Since we will neutralize the destination specific term by normalizations, it does not matter if we include the term $\sigma f_{China,j}$ or not.

Then for any selected percentiles p, we obtain our moment conditions:

$$M1_{\tilde{\omega}p} = \ln(r_{1\tilde{\omega}(p)}), \qquad \tilde{\omega} \in \{\{2, ..., 3\}, \{4, ...7\}, \{8, ...15\}, \{16, ...\}\}$$
$$p \in \{90, 80, 70, 60, 50, 40, 30, 20\}$$

The above procedures generate 4×8 moment conditions which is our first set of moments used in SMM algorithm.

M2: Within-destination sales of the core products

Our second set of moments compares the sales of core products across firms within the same destination. We divide firms into five groups by their (discrete) scope in a destination $\delta_{ij}(c, \beta)^{int} \in \{\{1, 2, ..., 7\}, \{8, ...15\}, \{16, ...31\}, \{32, ...\}\}$. We label the product subgroups with superscript $\tilde{\delta}$. The revenues of the core product ($\omega = 0$) of a firm with marginal and fixed cost draws (c, β), and scope group $\tilde{\delta}$ in a destination j are given by:

$$r(0, (c, \beta))^{\tilde{\delta}}_{ij} = \sigma w_j f_j \beta^*_{ij}(c^*_{ij}) \left(\frac{c}{c^*_{ij}}\right)^{-(\sigma-1)} \tag{66}$$

Let (c^m, β^m) denote the firm with the median level of core-product sales in a destination j and product group $\tilde{\delta}$. To neutralize the destination specific shifter, we normalize the sales (66), by the median sales:

$$\tilde{r}(0, (c, \beta))^{\tilde{\delta}}_{ij} = \frac{r(0, (c, \beta))^{\tilde{\delta}}_{ij}}{r(0, (c^m, \beta^m))^{\tilde{\delta}}_{ij}} = \left(\frac{c}{c^m}\right)^{-(\sigma-1)} \tag{67}$$

Next, we pool over these ratios across destinations and firms by product group $\tilde{\delta}$, and rank order the ratios from the highest to the lowest. This gives us a collection of ordered sequence $\{\tilde{r}(0)^{\tilde{\delta}}_{(100)}, ..., \tilde{r}(0)^{\tilde{\delta}}_{(0)}\}$. Consider, for instance, $\tilde{\delta} = 1$. The sequence contains the sales of the single-product exporter in a destination, normalized by the median sales of single-product firms in the same destination, pooled over destinations and firms. From the sequence we pick the observations at the select percentile $P(\omega) = p$.

$$M2_{\delta p} = \ln(\tilde{r}(0)^{\tilde{\delta}}_{(p)}), \qquad \tilde{\delta} \in \{\{1, 2, 3\}, \{4, ..., 7\}, \{8, ...15\}, \{16, ...\}\}$$
$$p \in \{95, 90, 85, 80, 70, 60, 40, 30, 20\}$$

The above procedures generate 4×9 moment conditions.

M3: Within-destination exporter scope distribution

The third set of moments captures the exporter scope distribution by destinations. We compute the share $S^{\bar{\delta}}_{ij}$ of firms with scope at least as large as $\bar{\delta}$ in a destination j as the percentage of the total number of exporters in the same destination:

$$S^{\bar{\delta}}_{ij} = \frac{\sum_{\phi > \phi^*_{ij}} \mathbb{I}(\delta_{ij}(\phi) \geq \bar{\delta})}{\sum_{\phi > \phi^*_{ij}} \mathbb{I}(\delta_{ij}(\phi) \geq 0)} \tag{68}$$

Next, we take the arithmetic means across destinations and this gives us the third set of moment conditions:

$$M3_{\bar{\delta}} = \frac{1}{D}\sum_{j=1}^{D} S_{ij}^{\bar{\delta}}, \qquad \bar{\delta} \in \{2, 4, 8, 16, 32\}$$

This procedure would provide us with 5 moment conditions.

M4: Within-destination firm sales distribution

Our fourth set of moment conditions reflects sales distribution by destinations and scope groups. We label the product subgroups with superscript $\tilde{\delta}$. For each destination j and scope group $\delta_{ij}(\phi) \in \{\{1, 2, ..., 7\}, \{8, ...15\}, \{16, ...31\}, \{32, ...\}\}$, we normalize the sales with the median sales within each destination product scope group, i.e., $R_{ij}^{\tilde{\delta}}(\phi, \beta)/R_{ij}^{m\tilde{\delta}}(\phi^m, \beta^m)$:

$$\frac{R_{ij}^{\tilde{\delta}}(\phi, \beta)}{R_{ij}^{m\tilde{\delta}}(\phi^m, \beta^m)} = \frac{\beta[\exp(\tilde{\theta}\delta_{ij}(\phi)) - 1]}{\beta^m[\exp(\tilde{\theta}\delta_{ij}(\phi^m)) - 1]}$$

For each scope group $\tilde{\delta}$, we pool over all the destination countries and firms and rank order normalized sales from the highest to the lowest. This gives us a collection of ordered sequence $\left\{\left(\frac{R^{\tilde{\delta}}}{R^{m\tilde{\delta}}}\right)_{(100)}, ..., \left(\frac{R^{\tilde{\delta}}}{R^{m\tilde{\delta}}}\right)_{(0)}\right\}$. For any select percentiles $P(\omega) = p$, we obtain the moment conditions:

$$M4_{\delta p} = \ln\left(\frac{R^{\tilde{\delta}}}{R^{m\tilde{\delta}}}\right)_{(p)}, \qquad \tilde{\delta} \in \{\{1, 2, 3\}, \{4, ..., 7\}, \{8, ...15\}, \{16, ...\}\}$$
$$p \in \{95, 90, 85, 80, 70, 60, 40, 30, 20\}$$

The above procedures generate 4×9 moment conditions.

M5: Within-firm exporter scope distribution

Our fifth set of moment conditions compares the export scopes within firms across destinations by double normalizing the export scope. To begin with, we only focus on the subset of firms that at least export to the US market. For each firm, we first calculate the firm-destination specific relative[34] exporter scope (relative to the scope in the U.S.) and denote it as $b_{ij}(c)$:

$$b_{ij}(c) = \delta_{ij}(c, \beta(c)_j) - \delta_{iUS}(c, \beta(c)_{US}) = \frac{1}{\tilde{\theta}}[\ln\beta(c)_{US} - \ln\beta(c)_j + \ln\phi_{iUS}^* - \ln\phi_{ij}^*] \qquad (69)$$

For each destination, we rank order the relative scope $b_{ij}(c)$ of all firms in that destination and further normalize $b_{ij}(c)$ with the median level within the destination[35] to obtain $\tilde{b}_{ij}(c) =$

[34]Normalizing a firm's scope with its US scope eliminates the firm specific shifter.
[35]The second normalization removes the destination specific shifter.

$b_{ij}(c) - med(b_{ij}(c))$:

$$\tilde{b}_{ij}(c) = b_{ij}(c) - med(b_{ij}(c)) = \frac{1}{\tilde{\theta}}[\ln\beta(c)_{US} - \ln\beta(c)_j - \ln\beta(c^m)_{US} + \ln\beta(c^m)_j] \tag{70}$$

Next, we pool over all firms selling in the different destinations, and rank over the double-normalized scope difference from the highest to the lowest, and we denote this new sequence as $\{\tilde{b}_{(100)}, ..., \tilde{b}_{(0)}\}$. For any select percentiles p, we obtain the moment conditions:

$$M5_p = \tilde{b}_{(p)}, \quad p \in \{90, 85, 80, 75, 60, 55, 45, 40, 25\}$$

The above procedures generate 9 moment conditions.

E.4 Statistical Inference

To estimate the key parameters Θ, we first take the differences between observed and simulated moments $\Delta\mathbf{M}(\Theta) = \mathbf{M}^{data} - \mathbf{M}^{sim}$. There are 118 moment conditions as implied by above. The true parameter Θ_0 will satisfy $\mathbb{E}[\Delta\mathbf{M}(\Theta_0)] = 0$. So we search for the optimal $\widehat{\Theta}$ to minimize the weighted sum of squares, i.e., $\Delta\mathbf{M}(\Theta)'\mathbf{W}\Delta\mathbf{M}(\Theta)$. Ideally, $\mathbf{W}$ is the positive semi-definite weighting matrix which is chosen at $\mathbf{W} = \mathbf{V}^{-1}$, where $\mathbf{V}$ is the variance-covariance matrix of the moments $\Delta\mathbf{M}(\Theta)$. However, the true variance-covariance matrix is unobserved, and we use its empirical estimates:

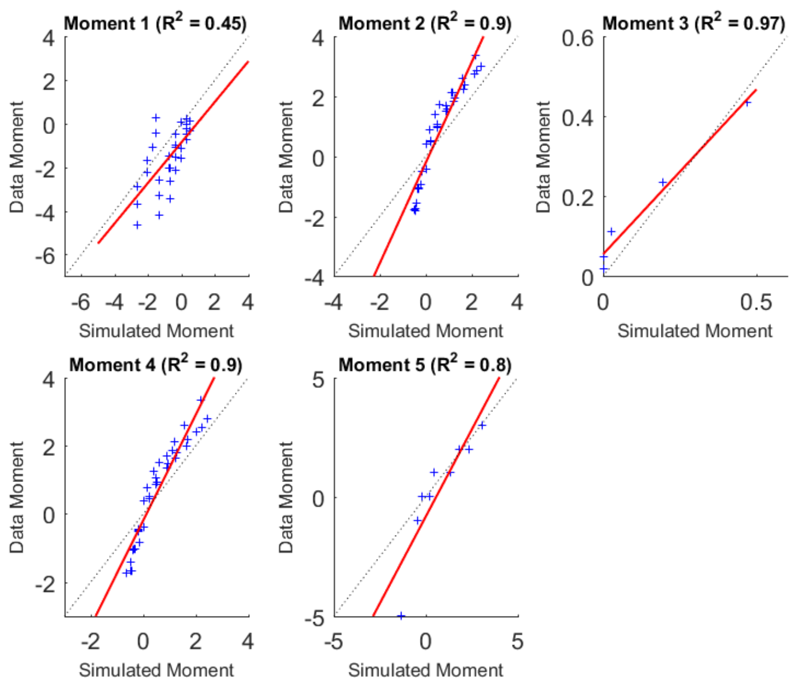
$$\widehat{\mathbf{V}} = \frac{1}{N^{Bstrp}} \sum_{i=1}^{N^{Bstrp}} (\mathbf{M}^{data} - \mathbf{M}_i^{Bstrp})\,(\mathbf{M}^{data} - \mathbf{M}_i^{Bstrp})' \tag{71}$$

where $\mathbf{M}^{data}$ is moment conditions from the data, and $\mathbf{M}_i^{Bstrp}$ denotes the moments generated from a random sample (sample i) that is drawn with replacement from the original firms by destination in the dataset. N^{Bstrp} is the number of Bootstrap samples. Due to the large dimensions, we cannot invert $\widehat{\mathbf{V}}$. Instead, we calculate the Moore-Penrose pseudo-inverse. Next, we calculate the standard deviation of $\widehat{\Theta}$ via bootstrapping. Specifically, we repeat the estimation process 30 times, using $\Delta\mathbf{M}^{Bstrp}(\Theta) = \mathbf{M}^{Bstrp} - \mathbf{M}^{sim}$ where we replace $\mathbf{M}^{data}$ in $\Delta\mathbf{M}(\Theta)$ with $\mathbf{M}^{Bstrp}$, to generate standard errors. This method is also used in Arkolakis et al. (2021), which allows for sampling and simulation errors.

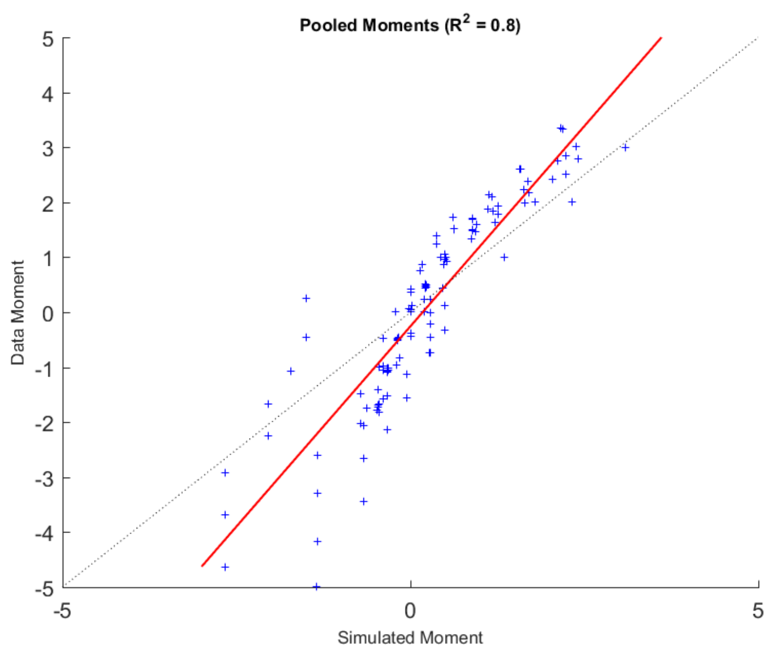
E.5 Model Fit

Figure E.1: Model Fit: Baseline

(a) Moments Comparison by Moment



(b) Moments Comparison Overall



Panel (a) of Figure E.1 provides the estimation performance by each set of moments. Panel (b) displayed the model fit for the simulated moments and the counterpart in data. The overall estimation is quite accurate according to the high R^2. In each set of moments, the scatter plots also indicate that there is no systematic difference between the simulated moments and their counterpart in data.

In Table E.1, we report the goodness of fit for a number of alternative specification. In column (1), we consider our baseline model and fix $\alpha = 0$, in column (2) and (3) we fix a value of gamma which is twice or half the value we estimated in our baseline model. In column (5), we augment our baseline model with ex-post demand shocks. In column (6), we consider firm learning (which we discuss in detail in Appendix F).

Table E.1: Models' Goodness of Fit

Moments	(1) Baseline	(2) $\alpha = 0$	(3) High γ	(4) Low γ	(5) Demand Shocks	(6) Firm Learning
M1	0.45	0.29	0.43	0.49	0.88	0.42
M2	0.9	0.89	0.87	0.93	0.96	0.93
M3	0.97	0.93	0.97	0.97	0.99	0.91
M4	0.9	0.87	0.93	0.92	0.98	0.94
M5	0.8	0.8	0.79	0.79	0.8	0.82
M1-5	0.8	0.48	0.72	0.63	0.91	0.67

R^2 from regressing the moments predicted by the model against the moments from the data. The list of moments M1-5 is provided in the main text.

E.6 Robustness

Table E.2: Estimation of Θ by Subset of Industries

Industry Category	HS-Code	$\hat{\kappa}$	$\hat{\alpha}$	$\hat{\beta}_m$	$\hat{\gamma}$	$\hat{\theta}$	$\hat{\sigma}$
Vegetable Products	6 - 15	1.157***	-0.678	1.067***	2.966*	0.258***	1.777***
		(0.395)	(0.614)	(0.365)	(1.639)	(0.052)	(0.315)
Foodstuffs	16 - 24	1.003	-0.807	1.531***	2.372	0.258***	1.844**
		(1.729)	(1.162)	(0.389)	(2.015)	(0.057)	(0.665)
Chemicals & Allied Industries	28 - 38	2.071***	-1.183	1.019***	4.458***	0.103***	2.729**
		(0.431)	(1.320)	(0.253)	(1.562)	(0.036)	(1.015)
Textile	50 - 63	1.270***	-0.596	1.053**	3.481***	0.254***	1.685**
		(0.059)	(0.900)	(0.429)	(0.897)	(0.068)	(0.627)
Metals	72 - 83	1.024***	-0.858	1.356***	1.836*	0.237***	1.913***
		(0.281)	(0.595)	(0.279)	(1.083)	(0.069)	(0.422)
Machinery & Electrical	84 - 85	1.003***	-1.456***	0.534*	1.595**	0.195***	2.252***
		(0.209)	(0.307)	(0.280)	(0.658)	(0.033)	(0.195)
Transportation	86 - 89	1.491	-0.810	0.996**	2.142*	0.258***	2.000***
		(1.284)	(0.827)	(0.436)	(1.146)	(0.014)	(0.521)
Miscellaneous	$\geq$90	1.012***	-0.719***	1.151**	2.826***	0.261***	1.746***
		(0.175)	(0.227)	(0.451)	(0.638)	(0.043)	(0.163)

United States as the reference country. The estimates using China as reference country are similar. Bootstrapped standard errors are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

E.7 Other Moments

Figure E.2: Simulated Scope and Productivity Measure (U.S.)

(a) Productivity Measure (b) Value-added Per Worker

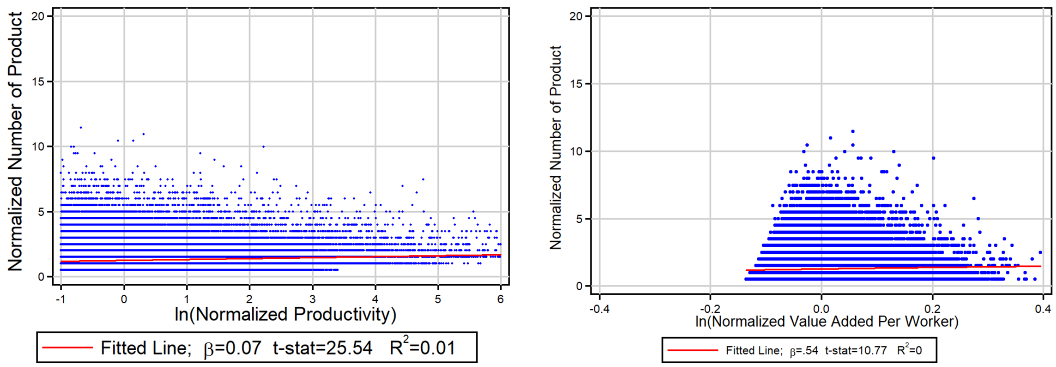


Figure E.3: Simulated Distribution of Scope Conditional on Productivity

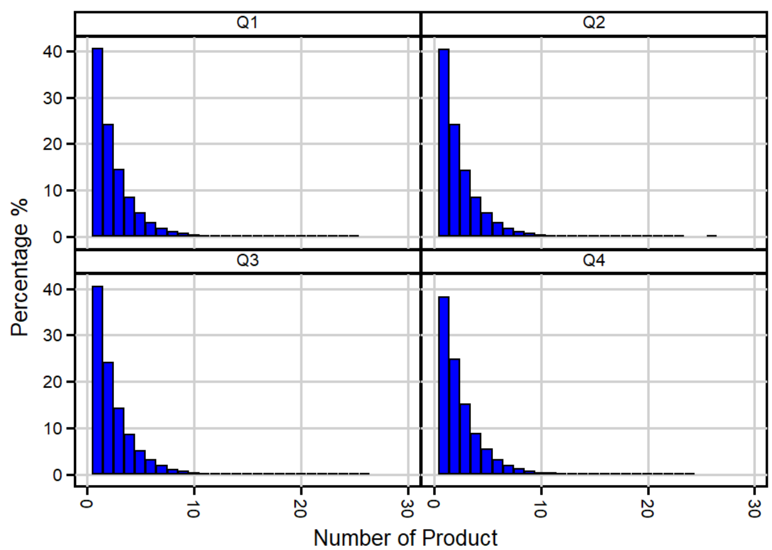
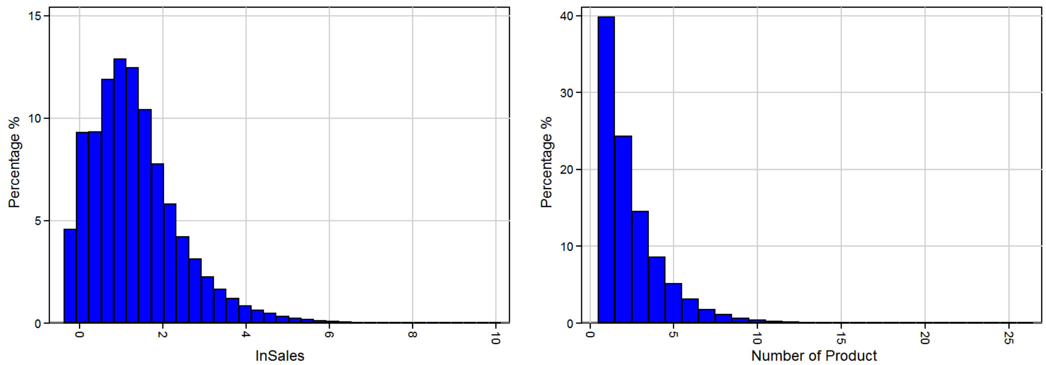


Figure E.4: Simulated Distribution of Scope and Sales



F Model with Firm Learning

We model learning as the realization of uncertainty in spirit of Nguyen (2012), where our uncertainty is on the side of fixed cost shocks. Specifically, before a firm enters a new market, we allow the firm to have some information on his flexibility by learning his experience in exporting to markets that share similar characteristic. As a simplification, we assume countries from the same geographic division have similar realizations of fixed cost shocks. To capture this idea, we first divide country into four main categories according to geographic locations, and the details are reported in Table G.1.

To capture the effects of learning in our framework, we allow for an additional shock to the fixed cost. In addition to β, a firm draws an additional shock l_r which is common across all destinations j in group r. l_r is lognormally distributed with mean zero and standard deviation σ_l. We group the 18 most popular destinations of Chinese exporters in four geographical areas. The fixed costs a firm pays becomes $f_j \beta l_r$. In the presence of learning, a Chinese firm that exports to the US can learn about its flexibility in a similar destination, i.e. Canada. We capture this effect as the shock l_r is the same for the two countries.

The estimation is similar to the baseline model. First, we assume that the there is an additional layer of fixed cost shocks denoted as l_{fr} that are log normally distributed with standard deviation σ_l and zero mean. The realized fixed cost shock for firm f in destination j is $\beta_{fj} \times l_{fr}$. Now exporters are faced with fixed cost shocks consisting of two types of uncertainty. The first term β_{fj} is firm-destination specific which is the same as what we have in the baseline. The new source of uncertainty l_{fr} is firm-region specific, which will be homogeneous for countries in the same region r for firm f.[36] Thus, we have a new parameter to estimate σ_l. Table G.2 report the calibrated parameters for the extended model with learning.

[36]For instance, firm f may learn the flexibility in Canada from the experience of exporting to US because Canada and US are in the same region. In the model, we will capture this feature by assuming $l_{f,Canada} = l_{f,USA}$. Nonetheless, firm f will still be faced with Canada-specific fixed shocks ($\beta_{f,Canada}$) which are independent of the US shocks ($\beta_{f,USA}$).

Table G.1: Division of Countries

Rank	Country Code	Export Share	Region Division (r)
1	USA	0.239	North & South America
2	JPN	0.178	Asia
3	HKG	0.168	Asia
4	KOR	0.153	Asia
5	DEU	0.140	Europe
6	GBR	0.122	Europe
7	AUS	0.113	Ocean Area
8	ITA	0.111	Europe
9	CAN	0.111	North & South America
10	TWN	0.098	Asia
11	ESP	0.097	Europe
12	SGP	0.093	Ocean Area
13	NLD	0.092	Europe
14	FRA	0.090	Europe
15	ARE	0.087	North & South America
16	IND	0.084	Asia
17	IDN	0.069	Ocean Area
18	RUS	0.065	Europe

Table G.2: Estimation of Θ (SMM): the Model Learning Effect

	Pooling Sample Estimation Description \Reference Country	(1) United States	(2) China
$\hat{\kappa}$	Shape par. of productivity distr	1.005*** (0.035)	1.003*** (0.043)
$\hat{\alpha}$	Correlation productivity fixed cost	-1.456** (0.562)	-1.804*** (0.543)
$\hat{\beta}_m$	Shift par. of fixed cost distr.	0.789** (0.370)	0.965*** (0.340)
$\hat{\gamma}$	Shape par. of fixed cost distr.	3.248*** (0.493)	4.480*** (1.184)
$\hat{\theta}$	Elas. of m.cost with distance from core	0.094*** (0.019)	0.086*** (0.018)
$\hat{\sigma}$	Elasticity of substitution	2.342*** (0.441)	2.654*** (0.427)
$\widehat{Std(l)}$	Standard deviation of the logarithmic fixed cost shocks	0.566*** (0.114)	0.340* (0.189)

United States as the reference country. Bootstrapped standard errors are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

G Welfare Implications: US-China Trade Tension

G.1 Theoretical Derivations

We leverage the hat-algebra approach of ACR, and derive the general equilibrium welfare changes due to any change in trade costs. The advantage of the approach is that it requires a parsimonious set of variables from the data to compute the counterfactuals. The first step is to describe analytically the equilibrium of the model as a function of export trade shares and of wages. The first system of equilibrium equations is the gravity equation:

$$\lambda_{ij} = \frac{L_i \bar{c}_i^{-\kappa} (\tau_{ij} w_i)^{-\frac{\kappa\tilde{\gamma}}{\alpha\gamma+\tilde{\gamma}}}}{\sum_v L_v \bar{c}_v^{-\kappa} (\tau_{vj} w_v)^{-\frac{\kappa\tilde{\gamma}}{\alpha\gamma+\tilde{\gamma}}}} \qquad \forall i,j = 1,...,I \tag{72}$$

The second system of equations is the market clearing condition:

$$\sum_j \lambda_{ij} w_j L_j = w_i L_i \quad \forall i = 1,...,I \tag{73}$$

We denote with a hat symbol the relative change in a variable: $\hat{x} = x_{new}/x_{old}$. Considering any change in iceberg trade cost and fixed costs per variety and destination, we obtain the following system of equations:

$$\hat{\lambda}_{ij} = \frac{(\hat{\tau}_{ij} \hat{w}_i)^{-\frac{\kappa\tilde{\gamma}}{\alpha\gamma+\tilde{\gamma}}}}{\sum_v \lambda_{vj} (\hat{\tau}_{vj} \hat{w}_v)^{-\frac{\kappa\tilde{\gamma}}{\alpha\gamma+\tilde{\gamma}}}} \qquad \forall i,j = 1,...,I \tag{74}$$

$$\hat{w}_i = \frac{\sum_j \lambda_{ij} w_j L_j \hat{\lambda}_{ij} \hat{w}_j}{\sum_j \lambda_{ij} w_j L_j} \quad \forall i = 1,...,I \tag{75}$$

which are obtain by applying the hat-algebra to (72) and (73). Hence, given data on λ_{ij}, L_j, and w_j, and any change in τ_{ij}, we can characterize the changes in trade shares $\hat{\lambda}_{ij}$ and wages $\hat{w}_v$ by solving the system of equations defined in (74) and (75).

The ACR conditions apply to our model, so that the equivalent variation in income due to a change in trade costs can be represented as:

$$\hat{W}_j = \hat{\lambda}_{jj}^{-\frac{\alpha\gamma+\tilde{\gamma}}{\tilde{\gamma}\kappa}} \tag{76}$$

It is straightforward to obtain this result, by applying the hat algebra changes to the definition of demand shifter A_j:

$$\hat{A}_j = \hat{U}_j^{-\frac{\sigma-1}{\sigma}}$$

Then, by the definition of the domestic cost cutoff, holding constat the domestic iceberg and fixed trade costs, as well as normalizing the wage of country j to one, without loss of generality, we obtain:

$$\hat{c}^*_{jj} = \hat{A}_j^{-\frac{\sigma\gamma}{\alpha\gamma+\tilde{\gamma}}} = \hat{U}_j^{-\frac{\tilde{\gamma}}{\alpha\gamma+\tilde{\gamma}}}$$

Furthermore, there is the following relationship between the domestic cost cutoff and the domestic trade share:

$$\hat{c}^*_{jj} = \hat{\lambda}_{jj}^{\frac{1}{\kappa}}$$

Thus, combining the last two equations, we obtain:

$$\hat{U}_j = \hat{\lambda}_{jj}^{-\frac{\alpha\gamma+\tilde{\gamma}}{\tilde{\gamma}\kappa}}$$

Since our CES preferences are homothetic, the equivalent variation in income is $\hat{W}_j = \hat{U}_j$.

G.2 Quantification

For welfare analysis, we use countries that account for 95% of the world GDP in 2006, which leaves us the total number of 49 countries. For importer i and exporter j, we compute the trade share λ_{ij} as follow:

$$\lambda_{ij} = \frac{X_{ij}}{Y_j + IM_j - EX_j}$$

where X_{ij} denotes the trade flows from country i to country j. The denominator is country j's domestic absorption, computed as gross output (Y_j) plus imports (IM_j) and minus exports (EX_j). Following Fernandes et al. (2018), we use the simplification that gross output is four times a country's GDP. The domestic expenditure of a country is computed as total domestic absorption minus total imports from foreign countries, i.e., $X_{jj} = Y_j + IM_j - EX_j - \sum_{i \neq j} X_{ij}$.

Table F.1: Change in Trade Share and Welfare (Baseline Model)

Country	$\Delta\lambda_{jj}$ (%)	Δ Welfare (%)	Country	$\Delta\lambda_{jj}$ (%)	Δ Welfare (%)
ARE	0.0010	-0.0003	IRN	-0.0023	0.0007
ARG	0.0006	-0.0002	ISR	0.0037	-0.0011
AUS	0.0004	-0.0001	ITA	-0.0007	0.0002
AUT	-0.0001	0.0000	JPN	-0.0007	0.0002
BEL	0.0022	-0.0007	KOR	0.0005	-0.0001
BRA	0.0017	-0.0005	MEX	-0.0010	0.0003
CAN	0.0043	-0.0013	MYS	0.0029	-0.0009
CHE	0.0022	-0.0007	NGA	-0.0026	0.0008
CHL	0.0026	-0.0008	NLD	-0.0001	0.0000
CHN	0.3325	-0.1005	NOR	-0.0002	0.0001
COL	0.0035	-0.0011	PAK	-0.0022	0.0007
CZE	-0.002	0.0006	PHL	0.0094	-0.0028
DEU	-0.0005	0.0001	POL	-0.0020	0.0006
DNK	-0.0003	0.0001	PRT	0.0004	-0.0001
DZA	-0.0037	0.0011	ROU	-0.0035	0.0011
ESP	-0.0005	0.0002	RUS	-0.0025	0.0008
FIN	-0.0008	0.0003	SAU	0.0022	-0.0007
FRA	0.0001	0.0000	SGP	0.0139	-0.0042
GBR	0.0005	-0.0001	SWE	0.0001	0.0000
GRC	-0.0001	0.0000	THA	-0.0014	0.0004
HKG	-0.0370	0.0112	TUR	-0.0013	0.0004
HUN	-0.0023	0.0007	USA	0.2633	-0.0796
IDN	-0.0029	0.0009	VEN	-0.0236	0.0071
IND	0.0003	-0.0001	ZAF	-0.0174	0.0053
IRL	0.0003	-0.0001			

Table F.2: Change in Trade Share and Welfare (Standard Model)

Country	$\Delta\lambda_{jj}$ (%)	Δ Welfare (%)	Country	$\Delta\lambda_{jj}$ (%)	Δ Welfare (%)
ARE	0.0019	-0.0002	IRN	-0.0049	0.0004
ARG	0.0012	-0.0001	ISR	0.0074	-0.0006
AUS	0.0008	-0.0001	ITA	-0.0016	0.0001
AUT	-0.0002	0.0000	JPN	-0.0015	0.0001
BEL	0.0044	-0.0004	KOR	0.0012	-0.0001
BRA	0.0036	-0.0003	MEX	-0.0021	0.0002
CAN	0.0085	-0.0007	MYS	0.0057	-0.0005
CHE	0.0044	-0.0004	NGA	-0.0057	0.0005
CHL	0.0054	-0.0004	NLD	-0.0002	0.0000
CHN	0.6208	-0.0501	NOR	-0.0003	0.0000
COL	0.0073	-0.0006	PAK	-0.0045	0.0004
CZE	-0.0038	0.0003	PHL	0.0203	-0.0016
DEU	-0.0011	0.0001	POL	-0.0042	0.0003
DNK	-0.0006	0.0000	PRT	0.0007	-0.0001
DZA	-0.0078	0.0006	ROU	-0.0070	0.0006
ESP	-0.0010	0.0001	RUS	-0.0055	0.0004
FIN	-0.0016	0.0001	SAU	0.0045	-0.0004
FRA	0.0001	0.0000	SGP	0.0279	-0.0023
GBR	0.0010	-0.0001	SWE	0.0003	0.0000
GRC	-0.0003	0.0000	THA	-0.0029	0.0002
HKG	-0.0707	0.0057	TUR	-0.0027	0.0002
HUN	-0.0046	0.0004	USA	0.4694	-0.0379
IDN	-0.0065	0.0005	VEN	-0.0527	0.0043
IND	0.0006	-0.0001	ZAF	-0.0390	0.0032
IRL	0.0005	0.0000			