Heterogeneous Market Structure and International Trade Patterns

Nahyeon Bak*
Amazon.com

Daisoon Kim†
North Carolina State University

Mishita Mehra‡
Grinnell College

March 1, 2022

Abstract

How do industry differences influence the home market effect? We build a model that highlights the role of scale economies generated from non-constant marginal costs in shaping trade patterns. We then provide empirical evidence that supports the model's prediction that large returns to scale industries are concentrated in large countries when non-production activities are more non-tradable input (labor) intensive than production activities. Consistent with the previous literature, high markup and/or trade cost industries are concentrated in large countries. Furthermore, our quantitative analysis reveals that heterogeneous returns to scale generate substantial variation in the bilateral export-import ratios across industries.

JEL Classification: F12, F14, L11.
Keywords: Home market effect, Markups, New trade theory, Non-constant marginal costs, Returns to scale.

*Core AI, Amazon.com, 345 Boren Avenue, Seattle, WA 98144, US. Email: nahyeonbak@gmail.com

†Department of Economics, Poole College of Management, North Carolina State University, Nelson Hall 4114, Raleigh, NC 27695, US. Email: dkim29@ncsu.edu

‡Department of Economics, Grinnell College, HSSC S1321, Park St, Grinnell, IA 50112, US. Email: mehramis@grinnell.edu

This paper started with Kim (2019); Ph.D. dissertation at University of Washington. Daisoon Kim is very grateful to Fabio Ghironi and John Geweke for their valuable advice and encouragement. The authors would like to thank Luca David Opromolla, Yu-chin Chen, Jae Wook Jung, Akbar Marvasti, Kyung In Hwang, as well as seminar participants at the University of Washington, WEAI 2019 conference, and the Korea International Economic Association Annual Meeting 2021 for useful comments. Any remaining errors are our own.
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. This paper was written prior to Nahyeon Bak joining Amazon.com.

1. Introduction

The relationship between country size and trade patterns has important implications. If trade liberalization and globalization lead to the concentration of manufacturing industries in relatively larger countries, these countries will exhibit net export gains in such industries, potentially hurting smaller countries. The home market effect (Krugman, 1980) suggests that relocation to large countries could be an essential mechanism for industries with scale economies, differentiated products, and high transport costs since large countries offer an advantageous market size and trade cost savings.[1]

Motivated by these factors, this paper aims to study how the home market effect is influenced by industrial characteristics such as returns of scale, markups, trade elasticity, and trade costs. Our main focus is on the impact of returns to scale and we attempt to shed light on a new angle on returns to scale in international trade studies. To this end, we first build a multi-country multi-industry new trade model that allows for diverse sources of supply and demand side heterogeneity. Unlike much of the previous literature on the home market effect, our model includes (a) scale economies in production emerging from non-constant marginal costs and (b) labor as well as captial (non-tradable as well as tradable inputs). Our theoretical model predicts that the impact of returns to scale on the home market effect depends on the relative labor input intensities used in production vs non-production activities. We then use narrowly-defined industries' bilateral trade flow data to estimate industrial traits' impact on the home market effect — the impact of country size on industries' net exports. Our empirical results are consistent with the model's main prediction — when the labor input is used more intensively in non-production activities, returns to scale positively impact the home market effect. We also argue that returns to scale channel leads to an additional 21 percent of cross-sectional variation in the export-import ratio across industries. These results highlight the importance of returns to scale (especially, non-constant marginal costs) and their heterogeneity in influencing industrial trade patterns.

To investigate home market effects' dependence on industrial characteristics, we extend Hanson and Xiang (2004)'s multi-industry new trade model by allowing for many countries and a more flexible preference and production structure. Our model economy's demand side characteristics are similar to Hanson and Xiang (2004). However, we use a more generalized preference structure— a nested-CES utility as in Lashkaripour (2020) — to break the tight link between trade elasticity and markups, i.e., we allow for different national-level and firm-level elasticities of substitution.[2]

The key novel feature of the model is cost-side heterogeneity across industries. Though identical within an industry, we allow firms' cost structures to differ across industries in many dimensions. Firms in each industry face non-constant marginal costs — i.e., firms' marginal costs vary with their production level, which is empirically supported (e.g., Almunia et al. 2018; Bergstrand et al. 2021; Kim 2021). This generates an additional source of scale economies in production. Additionally, our model includes two factors of production: labor and capital. The former is non-tradable, whereas the latter is tradable across countries. We allow the relative factor intensities to

[1] We interpret the home market effect as the impact of country size on industry concentration and trade surplus. A negative size impact implies a negative home market effect in an industry, which is sometimes called the inverse home market effect.

[2] Appendix A introduces a two country model with CES preferences for the rigorous mathematical proofs of the model's mechanisms.

differ (a) between production and non-production activities and also (b) across industries.[3] Then, the impact of returns to scale (arose from non-constant marginal costs) on industrial home market effects depends on the relative intensity of labor used in production and non-production activities.

If the relative labor input intensity used in firms' non-production (operation/entry) and production is identical, then an industry's returns to scale do not matter for its home market effect. In this case, the free entry condition forces the firm size to be equal across countries despite differences in country size (aggregate market size), and therefore scale economies in production do not lead to advantages for larger countries. The model predicts that returns to scale positively influences the home market effect when an industry's non-production activities are more labor input intensive than production activities. Under this condition, the industry's average firm size (i.e., employment/output) is larger in larger countries. The larger country exhibits appreciated terms of labor (non-tradable input) because of the larger market size and higher labor demand. If non-production activities in an industry require relatively more labor, firms in the industry face a more considerable increase in non-production costs compared to production costs in the larger country. These higher entry/operation costs hinder firm entry within the industry in the larger country but augment the firm size/scale. In other words, firms in this industry expand through the intensive margin rather than the extensive margin in the larger country. The large firm size causes advantages when there are economies of scale but disadvantages under diseconomies of scale. Hence, our model predicts that large countries would have a higher concentration of industries with large returns to scale. In contrast, if non-production activities require a relatively smaller amount of the labor input compared to the capital input, our model predicts that industries with large returns to scale would tend to be concentrated in relatively small countries.

The main contribution of our theoretical model is that we carefully consider the sources of scale economies as well as the cost structures of non-production and production activities. We find that these factors influence the direction of returns to scale's impact on the home market effect. There are two potential sources of scale economies — fixed costs and non-constant marginal costs. Even though fixed costs are important in trade models, the magnitude of fixed costs has a limited role in determining cross-country industrial trade patterns such as the home market effect.[4] Also, Kim (2021) documents that industry heterogeneity mainly arises from non-constant marginal costs rather than fixed costs (non-production costs) in US manufacturing industries. Moreover, conventional trade models with the home market effect such as Krugman (1980), Davis (1998), and Hanson and Xiang (2004), assume that production needs only one input, labor. In that case, the home market effect's strength is always independent of the degree of returns to scale. Our results show that if we ignore tradable inputs such as capital, we would not be fully capturing the

[3]Non-production costs include fixed, sunk, and overhead costs, which relate to operations and entry.

[4]Due to modeling tractability, conventional new trade models such as Krugman (1979, 1980), Melitz (2003), and Hanson and Xiang (2004) widely assume flat marginal cost curves with fixed costs. In these models, economies of scale are derived only from fixed costs. This assumption has several drawbacks in multi-industry models. First, fixed costs cannot generate dis-economies of scale, but in data, we see that average costs increase in some industries, as also supported by many empirical studies (e.g., Basu and Fernald 1997 and Lee 2007). Second, a flat marginal cost curve implies that an individual firm's decisions in domestic and export markets are separate, which is not supported by recent firm-level studies (e.g., Vannoorenberghe 2012, Soderbery 2014, Berman et al. 2015, and Almunia et al. 2018).

role of scale economies in international trade.[5]

We then test the main prediction of our theoretical model. Our empirical analysis uses narrowly defined industry-level data to investigate the impact of returns to scale on the home market effect. We measure industries' home market effect by the estimated export-import ratio's elasticity with respect to the relative size of origin and destination countries. This elasticity depends on industry traits such as returns to scale and markups. We use the export-import ratio, which cancels out symmetric variables between origin and destination countries in the regression equations. We estimate our specification using bilateral trade data from Schott (2008) and his updates. We use data on the five-digit North American Industry Classification System (NAICS) manufacturing industries from 1989 through 2011 for 21 advanced economies among the top 40 trade partners of the U.S.

For our empirical analysis, in addition to sectoral returns to scale, we estimate demand side characteristics like markups, trade elasticity, and effective trade costs. These demand side characteristics have been found to be important in previous studies on the home market effect and we include them as controls in our industrial traits. For returns to scale and markups, we use the NBER-CES Manufacturing Industry Database from 1989 through 2011, following Hall (1988), Basu and Fernald (1997), and De Loecker et al. (2020).

As an indicator of product differentiation, the previous literature on the home market effect includes an estimated elasticity of substitution or a proxy for product differentiation from demand-side data. However, these measures are hard to get for many industries due to data limitations. Therefore, in contrast to the previous literature, we use markups from supply-side data that are easily accessible for many narrowly defined industries. This is another contribution of our paper in the home market literature. We construct trade elasticity estimates from Fontagne et al. (2020)'s six-digit Harmonized System (HS) categories in each industry.[6]

For evaluating the impact of returns to scale, we are interested in measuring the relative labor (non-tradable) input intensities in production vs non-production activities. However, the main challenge for the empirical analysis is that our data set unfortunately does not distinguish between capital and material inputs used in non-production and production activities even though we can observe non-production and production workers' payrolls and employment. Thus, we cannot directly calculate labor intensities in non-production and production costs. To overcome this problem, we use microeconomics about a cost function, factor demand, price, and associated relations. Shephard's lemma implies that the gap between wage elasticity of non-production and production labor input is approximately equal to the gap between labor intensities of non-production and production labor input.[7] Thus, we use the wage elasticity gap as a proxy for the labor input intensity gap. Using this measure, we provide empirical evidence that the impact of returns to scale on the home market effect depends on differences in input intensities in production and non-production activities.

The empirical results related to markups show that relatively large countries tend to have an

[5]As another example, consider Grossman and Helpman (1991)'s economy with labor and capital in which firm entry (i.e., non-production) needs only labor (non-tradable), while production needs both labor and capital. In that scenario, industries with high returns to scale (derived from decreasing marginal costs) are concentrated in larger countries.

[6]See Imbs and Mejean (2017), Boehm et al. (2020), and Giri et al. (2020) for recent studies related to trade elasticity estimation at the product- or sector-level. In this paper, trade elasticity serves as a robustness check because it influences product differentiation.

[7]The equality holds under Cobb-Douglas functional form.

export-import ratio greater than one in high-markup industries.[8] This result is akin to previous findings — a positive impact of product differentiation on the home market effect — in Krugman (1980) and Hanson and Xiang (2004), since the level of product differentiation is typically tightly linked with firm markups. A low price elasticity (more differentiated products) allows firms to charge a higher markup over the marginal cost, reinforcing the gains from locating in countries with the larger market size.

Our empirical results highlight the role of industry traits in shaping the home market effect. Most importantly, we show that returns to scale positively (negatively) impact the home market effect when the non-production activity is more (less) labor input intensive than the production activity. We also show that the degree of markups, trade elasticity, and effective trade costs positively impact the home market effect coefficient. The latter result is consistent with previous papers that show that the home market effect increases with product differentiation and trade costs.[9]

Finally, we quantify the role of returns to scale from non-constant marginal costs in trade patterns using our theoretical framework and measurements of industrial traits. Our quantitative analysis computes the marginal impact of heterogeneous returns to scale and labor input intensities on cross-sectional variations in export-import ratios. We find that the returns to scale channel contributes to 21% of the observed heterogeneous trade patterns measured by the cross-sectional standard deviation of log export-import ratios across industries. Also, the channel accounts for the evolution of trade heterogeneity during the sample period (1989 – 2011).

This paper illuminates new angles concerning international trade patterns of narrowly defined industries. The theoretical and empirical findings suggest that one of the fundamental elements of the new trade model — scale economies — is crucial for understanding narrowly defined industry-level international trade patterns. Distinguishing this paper from the previous literature on the home market effect, we highlight the role of scale economies that arise from non-constant marginal costs as well as the relative intensities of labor and capital inputs used in firms' production and non-production activities.

Related Literature. Our paper contributes to a vast theoretical and empirical literature on the home market effect. Though the economic literature seems to have reached a consensus about the home market effect, its magnitude and conditional constraints are still under much debate. Country size, trade costs, elasticities of demand, elasticities of substitution, and country cost structure are all variables that determine the magnitude of the home market effect.

Our paper is closely related to Hanson and Xiang (2004), who construct a multi-industry new trade model that allows for different demand elasticities and trade costs across industries. They theoretically show that industries with high trade costs and more differentiated goods are concentrated in larger countries. Also, using a difference-in-difference approach, they find empirical evidence for their theoretical predictions. Laussel and Paul (2007) also address the link between product differentiation, transport costs, and the home market effect of industries. The authors develop a two-industry model in which the elasticity of substitution across goods differs between industries,

[8] Markups have received recent attention in many areas of economics, including international economics. For example, see Rodriguez-Lopez (2011), Edmond et al. (2015), De Loecker et al. (2016), Keller and Yeaple (2020), and many others.

[9] For instance, using a multi-industry framework, Hanson and Xiang (2004)'s difference-in-difference analyses shows that the home market effect increases with product differentiation and trade costs.

and they conclude that a large country becomes a net exporter of more differentiated goods when transport costs are sufficiently high.

Much of the earlier literature on the home market effect — for instance, Krugman (1980), Helpman and Krugman (1985), and Davis (1998) — focus on very broadly defined industries: a manufacturing sector (with monopolistic competition and economies of scale) and a relatively homogeneous outside sector (with perfect competition and constant returns to scale).[10] Feenstra et al. (1998), Davis and Weinstein (1999, 2003), and Head and Ries (2001) empirically test the home market effect based on broadly defined industries. Crozet and Trionfetti (2008) study nonlinearities in the home market effect when the outside homogeneous good assumption is eliminated. Their result highlights the importance of considering multi-industries (or controlling for other industries) to investigate home market effects.

When there is no firm entry, scale economies directly imply that large countries have comparative advantages in production. In that case, large countries host a disproportionately high number of industries with large scale economies, as shown in Panagariya (1981) and Holmes and Stevens (2005). Holmes and Stevens (2005)'s new trade model shows that a large country becomes a net exporter in industries with large economies of scale. Furthermore, Panagariya (1981)'s perfect competition model predicts that a small country is specialized in industries with low returns to scale. Their results only hold under the assumption of no firm entry. Since the home market effect's trademark is a concentration of firms in large countries (and since such a concentration implies firm entry), the authors' mechanism differs from the original spirit of the home market effect.[11]

Furthermore, Fajgelbaum et al. (2011) address the home market effect from the perspective of different product qualities as well as product differentiation. Relatedly, Lashkaripour (2020) investigates within-industry quality specialization and market power. He finds that the composition of country-level exports aligns with the prediction that high-wage and distant economies export relatively more in high market power segments of each industry — this result illuminates another channel through which market power and markups affect exports. Using firm-level data, Dingel (2016) studies home-market effect across cities derived from the relationship between quality and income in the US.

Our empirical approach is related to previous estimation approaches in this literature but differs in some ways. Rather than exports, we regress the ratio of exports to imports, which eliminates symmetric bilateral variables such as distance, trade agreements, language, and borders. Also, in contrast to much of the literature (for instance, Hanson and Xiang 2004 and Pham et al. 2014), we use cost side markup estimates developed by Hall (1988) and De Loecker et al. (2020), the measure of which is continuous. Hanson and Xiang (2004) and Pham et al. (2014) use a discrete measure for product differentiation based on Rauch (1999)'s product classification.[12] Additionally,

[10]Instead of using broadly-defined level data, Costinot et al. (2019) study home market effects from detail data within a specific industry (the pharmaceutical industry).

[11]Related to this, Head and Ries (2001) show that when the number of firms is fixed (under increasing returns to scale) or when products are differentiated by nationality, there is a 'reverse home market effect' — a country's share of output increases less than proportionately with an increase in the country's share of demand.

[12]Rauch (1999) divides the 4-digit Standard International Trade Classification (SITC) commodities into three types: an organized exchange, reference priced, and differentiated.

our robustness checks are related to Pham et al. (2014).[13]

Our paper is closely related to the literature that has examined the role of economies of scale in international economics. Antweiler and Trefler (2002) document that allowing for increasing returns to scale in production significantly heightens the ability to predict international trade flows. Anderson et al. (2016) investigate the link between increasing returns and exchange rate pass-through. Grossman and Rossi-Hansberg (2010), Lyn and Rodriguez-Clare (2013), Kucheryavyy et al. (2016), and Bartelme et al. (2019) study implications of national- or industry-level external economies of scale for trading partners and welfare implications. Recently, Almunia et al. (2018), Bergstrand et al. (2021), and Kim (2021) emphasize the importance of internal scale economies derived from non-constant marginal costs for firms' trade, gains from trade agreements, and the international business cycle, respectively. In particular, Bergstrand et al. (2021) shows that allowing increasing marginal costs to the Melitz-type trade model leads to notable different impacts of trade policy on welfare. Even if their estimated marginal cost coefficients significantly differ across industries as in Kim (2021), their model and quantitative analysis are with a single sector and focus on an aggregate economies and trade outcomes. In contrast, we investigate how non-constant marginal costs and their heterogeneity account for cross-sectional different trade patterns.

Our quantitative analysis complements the recent quantitative trade literature related to scale economies. Kucheryavyy et al. (2016) propose a quantitative multi-sector gravity framework with scale economies from differentiated products (demand-side) and external economies of scale (supply-side). Bartelme et al. (2019) provide evidence for sector-level external economies of scale and its heterogeneity at the broadly defined sectoral level and then evaluate gains from industrial policies in an open economy. Lashkaripour (2020)'s quantitative framework (monopolistic competition and non-CES preference) and results also highlight scale economies' importance for industrial policy evaluation, focusing on scale economies derived from the demand side. Similar to our paper, Bergstrand et al. (2021) focus on the role of internal economies of scale. They provide evidence for increasing marginal costs (decreasing returns to scale) and study the impacts on welfare gains from trade policies in Melitz (2003)'s type economy.

This paper is organized as follows. Section 2 presents a multi-industry, multi-country new trade framework to understand how returns to scale heterogeneity shapes the home market effect patterns across industries. Section 3 guides our empirical analysis based on the main channels presented in 2. Section 4 describes the main components of the data and variables. Section 5 empirically documents the home market effect's systemic variation with respect to industries' market characteristics. Section 6 provides a wide range of robustness check exercises. Section 7 uses the previous sections' frameworks and results for the quantification of the returns to scale channel in generating variation in the bilateral export-import ratios across industries. The last section concludes.

2. Theoretical Framework

How does the home market effect vary with industry traits? We attempt to answer this question with particular emphasis on one of the primary metrics of market structure: returns to scale. To achieve

[13] Pham et al. (2014) analyze Hanson and Xiang (2004)'s empirical framework and show that the results are sensitive to the way country pairs are defined and the inclusion of zero trade flows. We consider these issues in our robustness checks.

our goal and guide our empirical analysis, we construct a multi-industry, multi-country model based on the two-country model of Hanson and Xiang (2004). The latter extends the new trade model in Krugman (1979, 1980) to include infinitely many industries that face different demand-side characteristics such as elasticity of substitution across goods, trade costs, and expenditure shares.

Our model's key novel feature is that we allow for cost-side heterogeneity across industries, for example, slopes of marginal cost functions, fixed costs, and input cost shares. As in Kim (2021), firms in each industry face a sloping marginal cost — i.e., individual firms' marginal costs vary with their production level. This detail generates an additional source of scale economies in production. Additionally, our model includes two factors of production: labor and capital. The former is non-tradable, whereas the latter is tradable across countries. Finally, we allow the relative factor intensities to differ across production and non-production activities as well as across industries. These differences are essential for understanding the role of returns to scale on the home market effect.

The demand-side characteristics in our model are similar to Hanson and Xiang (2004). However, we use more generalized preferences widely used in the recent trade literature — a nested-CES utility structure as in Lashkaripour (2020). By allowing for different elasticities of substitution at the national vs. product level, these preferences break the strict link between trade elasticity and markup as seen in the conventional CES utility structure.

2.1. Environment

We consider a world economy consisting of many countries indexed by $i, j \in \mathcal{I}$ where i and j are the origin and destination countries. In each county, there are many industries indexed by $s \in \mathcal{S}$. In county i's industry s, there is a mass $n_i(s)$ of single-product firms that compete under monopolistic competition. Within each industry, firms' are identical except for their productivity, similar to the conventional model with firm heterogeneity, entry, and exit. A free entry condition endogenously determines the mass of firms in each industry. The countries have identical industry characteristics, but we allow factor endowments to differ across them. In particular, country i has a larger factor endowment than country j and therefore is larger than country j. The model balances each country's aggregate trade, but an industry's net exports can be positive or negative.

Preferences and Demand. In each country, the representative consumer's preference is described by a three-tier Cobb-Douglas-CES utility function as follows. First, consumer's preference in the destination country j is defined over a consumption basket with many industries in the set $\mathcal{S}$

$$Y_j = \prod_{s \in \mathcal{S}} [Y_j(s)]^{\phi_j(s)}, \tag{1}$$

where the Cobb-Douglas aggregator across industries implies a constant expenditure share $\phi_j(s)$ of country j's spending on industry s goods. Second, the consumption basket of destination country j over goods produced in industry s is

$$Y_j(s) = \left\{ \sum_{i \in \mathcal{I}} [Y_{ij}(s)]^{\frac{\sigma^{\text{natl}}(s)-1}{\sigma^{\text{natl}}(s)}} \right\}^{\frac{\sigma^{\text{natl}}(s)}{\sigma^{\text{natl}}(s)-1}}, \tag{2}$$

where $Y_{ij}(s)$ is the consumption basket of country j in industry s goods produced in origin country i.

In each industry s, $\sigma^{\text{natl}}(s)$ is the elasticity of substitution across origin country goods within industries. The corresponding price index is $P_j(s) = \left\{ \sum_{i \in \mathcal{I}} [P_{ij}(s)]^{1-\sigma^{\text{natl}}(s)} \right\}^{\frac{1}{1-\sigma^{\text{natl}}(s)}}$ where $P_{ij}(s)$ is the origin-specific price index of origin country i in destination country j's industry s.

Third, country j's industry s has a mass $n_i(s)$ of infinitely many competitors $\omega \in \Omega_i(s)$ from each origin i, i.e.,

$$Y_{ij}(s) = \left\{ \int_{\Omega_i(s)} [q_{ij}(\omega; s)]^{\frac{\sigma^{\text{firm}}(s)-1}{\sigma^{\text{firm}}(s)}} d\omega \right\}^{\frac{\sigma^{\text{firm}}(s)}{\sigma^{\text{firm}}(s)-1}}, \tag{3}$$

where $\sigma^{\text{firm}}(s) \geq \sigma^{\text{natl}}(s)$ is the elasticity of substitution across products ω from country i in country j's industry s. In country j's industry s, denote the quantity and price of firm ω located in country i by $q_{ij}(\omega; s)$ and $p_{ij}(\omega; s)$, respectively. The corresponding price index is $P_{ij}(s) = \left\{ \int_{\Omega_i(s)} [p_{ij}(\omega; s)]^{1-\sigma^{\text{firm}}(s)} d\omega \right\}^{\frac{1}{1-\sigma^{\text{firm}}(s)}}$.

The above demand structure allows us to distinguish between the market power at the product-level and the national-level market. In industry s, the individual firm ω located in origin country i faces the following demand in destination country j

$$p_{ij}(\omega; s) q_{ij}(\omega; s) = \left[\frac{p_{ij}(\omega; s)}{P_{ij}(s)} \right]^{1-\sigma^{\text{firm}}(s)} \left[\frac{P_{ij}(s)}{P_j(s)} \right]^{1-\sigma^{\text{natl}}(s)} \phi_j(s) Y_j \tag{4}$$

Also, the industry-level gravity equation is given by

$$\text{EX}_{ij}(s) = \int_{\Omega_i(s)} p_{ij}(\omega; s) q_{ij}(\omega; s) d\omega = \frac{[P_{ij}(s)]^{1-\sigma^{\text{natl}}(s)}}{\sum_{i \in \mathcal{I}} [P_{ij}(s)]^{1-\sigma^{\text{natl}}(s)}} \phi_j(s) Y_j. \tag{5}$$

The national-level elasticity of substitution determines the trade elasticity as $\varepsilon^{\text{trade}}(s) = 1 - \sigma^{\text{natl}}(s) < 0$, which is the percentage change in bilateral trade flows in response to a change in national-level bilateral (iceberg) trade costs denoted by $\tau_{ij}(s) = \tau_{ji}(s)$. The firm-level elasticity of substitution shapes the degree of markups as $\mu(s) = 1 + [\sigma^{\text{firm}}(s) - 1]^{-1}$, which is the ratio of price to marginal costs.

Production and Firms. Establishing a firm is costly as each firm incurs product development and production start-up costs called non-production activities. After paying these non-production costs, all firms share the same non-production and production cost structures but have different productivity levels in production, indexed by $a(\omega; s)$. On paying the non-production costs for entry, a firm draws $a(\omega; s)$ from a known distribution $G(\cdot; s)$ with support $[\underline{a}(s), \infty)$. Since the cost is sunk, all firms survive and produce.

A firm's production and non-production activity require capital and labor inputs in each industry. We assume that capital is tradable costlessly, which implies that its price is identical across countries. The fixed (non-production) and variable (production) costs are denoted by $\text{fc}_i(s)$ and

$\mathrm{vc}_i(\omega; s)$. We assume that

$$\mathrm{fc}_i(s) = \tilde{c}(w_i, r; s) f_C(s) \quad \text{and} \quad \mathrm{vc}_i(\omega; s) = c(w_i, r; s) \left[\frac{q_i(\omega; s)}{a(\omega; s)}\right]^{\frac{1}{\alpha(s)}}, \tag{6}$$

where w_i and r are labor and capital prices, respectively. $q_i(\omega; s) = \sum_{j \in \mathcal{I}} \tau_{ij}(s) q_{ij}(\omega; s)$ denotes a firm's output. $f_C(s)$ is the non-production (fixed) cost in efficiency unit of inputs. The cost functions $\tilde{c}(w_i, r; s)$ and $c(w_i, r; s)$ must be homogeneous of degree one with respect to (w_i, r). The above cost structure implies that $a(\omega; s)$ is Hicks neutral (first degree homogeneous in the corresponding production function).

There are two sources of returns to scale: fixed costs and non-constant marginal costs. Call $\alpha(s)$ the return to scale parameter. The parameter is output elasticity of variable inputs — labor and capital in production denoted by $l_{p,i}(\omega; s)$ and $k_{p,i}(\omega; s)$ — in the corresponding production function $q = a(\omega; s)[f(l_{p,i}(\omega; s), k_{p,i}(\omega; s))]^{\alpha(s)}$ where $f(\cdot, \cdot)$ is homogeneous of degree one. The parameter represents returns to scale derived from non-constant marginal costs. The slope of logged costs with respect to logged output is a function of the returns to scale coefficient by $1/\alpha(s) - 1$.

$$\mathrm{mc}_i(\omega; s) = \left[\frac{c(w_i, r; s)}{\alpha(s) a(\omega; s)}\right] \left[\frac{q_i(\omega; s)}{a(\omega; s)}\right]^{\frac{1}{\alpha(s)} - 1}, \tag{7}$$

This detail diverges from the framework in Krugman (1979, 1980) and Melitz (2003), where only fixed costs generate scale economies in production.[14] Kim (2021) provides evidence that different marginal cost structures are the main source of heterogeneous scale economies across industries. Therefore, our model features an additional source represented by the returns to scale coefficient $\alpha(s)$.

Firms' profit maximization yields the well-known price-setting condition that the price is a markup over marginal cost as follows.

$$p_{ij}(\omega; s) = \tau_{ij}(s) \mu(s) \mathrm{mc}_i(\omega; s) = \tau_{ij}(s) \left[\frac{\mu(s)}{\alpha(s)}\right] \left[\frac{c(w_i, r; s)}{a(\omega; s)}\right] \left[\frac{q_i(\omega; s)}{a(\omega; s)}\right]^{\frac{1}{\alpha(s)} - 1} \tag{8}$$

We assume that the returns to scale coefficient is lower than the markup for a unique finite solution, i.e., $\alpha(s) < \mu(s) = 1 + [\sigma^{\text{firm}}(s) - 1]^{-1}$.

Firm Entry. To summarize all the information on the productivity distributions relevant for all aggregates, define the average productivity level by $\bar{a}_i(s) = \left[\int_{\underline{a}(s)}^{\infty} a^{1/[\mu(s) - \alpha(s)]} dG(a; s)\right]^{\mu(s) - \alpha(s)}$ and index the average firm by $(\bar{\omega}_i; s)$ that satisfies $\bar{a}_i(s) = a(\bar{\omega}_i; s)$.[15]

A free-entry condition pins down the equilibrium number of firms and average firm size in a country. In each country i's industry s, firms enter until the average profits (ex-ante profits) are

[14] More precisely, the inverse elasticity of the total cost measures the degree of economies of scale: $\text{Average Costs}/\text{Marginal Costs} = \alpha(s)[1 + \text{Fixed Costs}/\text{Variable Costs}]$ because $\alpha \times q \times \text{Marginal Costs} = \text{Variable Costs} = \text{Total Costs} - \text{Fixed Costs}$.

[15] As in Melitz (2003), constant marginal costs ($\alpha(s) = 1$) yield the conventional aggregation: $\bar{a}_i(s) = \left\{\int_{\Omega_i(s)} [a(\omega; s)]^{1/[\mu(s) - 1]} d\omega\right\}^{\mu(s) - 1} = \left\{\int_{\Omega_i(s)} [a(\omega; s)]^{\sigma^{\text{firm}} - 1} d\omega\right\}^{1/[\sigma^{\text{firm}} - 1]}$.

zero. The free entry condition is:

$$\sum_{j\in\mathcal{I}} p_{ij}(\bar{\omega}_i; s)q_{ij}(\bar{\omega}_i; s) - \mathrm{vc}_i(\bar{\omega}_i; s) = \mathrm{fc}_i(s), \tag{9}$$

where the left-hand side is the firm's total revenue minus its variable costs and the right-hand side includes the fixed costs. The cost-minimization yields the following relation from the above equation.

$$\left[\frac{\mu(s)}{\alpha(s)} - 1\right]\mathrm{vc}_i(\bar{\omega}_i; s) = \mathrm{fc}_i(s) \tag{10}$$

This connects production and non-production activities (right and left-hand side, respectively) and determines the equilibrium firm size.

Equilibrium and Market Clearing Conditions. Denote the aggregate labor and capital endowments by L_i and K_i where we assume the amount of capital guarantee zero net capital flows in equilibrium. The industry-level inputs are

$$l_i(s) == n_i(s)\left\{\left[\frac{q_i(\bar{\omega}_i; s)}{\bar{a}_i(s)}\right]^{\frac{1}{\alpha(s)}}\left[\frac{\partial c(w_i, r; s)}{\partial w_i}\right] + f_C(s)\left[\frac{\partial \tilde{c}(w_i, r; s)}{\partial w_i}\right]\right\} \tag{11}$$

$$k_i(s) = n_i(s)\left\{\left[\frac{q_i(\bar{\omega}_i; s)}{\bar{a}_i(s)}\right]^{\frac{1}{\alpha(s)}}\left[\frac{\partial c(w_i, r; s)}{\partial r}\right] + f_C(s)\left[\frac{\partial \tilde{c}(w_i, r; s)}{\partial r}\right]\right\}, \tag{12}$$

where the factor market clearing conditions imply $L_i = \sum_s n_i(s)l_i(s)$ and $K_i = \sum_s n_i(s)k_i(s)$. Then, the balanced aggregate trade and zero profits imply the following aggregate accounting in each country:

$$w_i L_i + rK_i = Y_i, \tag{13}$$

where the GDP is the sum of industrial GDP: $Y_t = \sum_s Y_t(s)$. The industrial GDP is

$$Y_t(s) = n_i(s)p_i(\bar{\omega}_i; s)q_i(\bar{\omega}_i; s) = \left[\frac{\mu(s)}{\alpha(s)}\right]n_i(s)\mathrm{vc}_i(\bar{\omega}_i; s) = \left[\frac{\mu(s)}{\mu(s) - \alpha(s)}\right]n_i(s)\mathrm{fc}_i(s), \tag{14}$$

where we use the zero profit in equation (10).

2.2. Heterogeneous Returns to Scale and Home Market Effects Across Industries

This section carefully considers returns to scale (from non-constant marginal costs) as a determinant of the home market effect. Even though economies of scale are one of the new trade models' primary building blocks, their impact on the home market effect is not straightforward. The direction in which returns to scale impacts the home market effect depends on the cost structures of non-production and production activities because the differences in the cost structure determine cross-country firm sizes (production scale) through the free entry condition. Thus, international trade studies, including those that study home market effects, cannot fully investigate the impact of scale economies without considering the differences between non-production and production cost structures. This paper investigates the cost structures and shows how these factors influence the direction of returns to scale's impact on the home market effect. In Appendix A, we provide a two-country model with CES preferences to provide rigorous theoretical predictions for the results we discuss in this section.

To show this, note that the free entry condition and the functional forms of cost functions (Equations 6 and 10) give us the relative size of firms in countries i and j:

$$\left[\frac{q_i(\bar{\omega}_i;s)/\bar{a}_i(s)}{q_j(\bar{\omega}_j;s)/\bar{a}_j(s)}\right]^{\frac{1}{\alpha(s)}} = \frac{\tilde{c}(w_i,r;s)}{c(w_i,r;s)}\frac{c(w_j,r;s)}{\tilde{c}(w_j,r;s)}, \tag{15}$$

where the size is in terms of factors in production. Substituting Equation (15) in the expression for marginal costs, the terms of trade is given by

$$\begin{aligned}\mathrm{TOT}_{ij}(s) = \frac{p_{ij}(\bar{\omega}_i;s)}{p_{ji}(\bar{\omega}_i;s)} &= \left[\frac{\bar{a}_j(s)}{\bar{a}_i(s)}\right]\left[\frac{c(w_i,r;s)}{c(w_j,r;s)}\right]\left[\frac{q_i(\bar{\omega}_i;s)/\bar{a}_i(s)}{q_j(\bar{\omega}_j;s)/\bar{a}_j(s)}\right]^{1-\frac{1}{\alpha(s)}} \\ &= \left[\frac{\bar{a}_j(s)}{\bar{a}_i(s)}\right]\left[\frac{c(w_i,r;s)}{c(w_j,r;s)}\right]\underbrace{\left[\frac{\tilde{c}(w_i,r;s)}{c(w_i,r;s)}\frac{c(w_j,r;s)}{\tilde{c}(w_j,r;s)}\right]^{1-\alpha(s)}}_{\text{Scale economies in production}},\end{aligned} \tag{16}$$

which represents cross-country competitiveness. Using the industry-level export-import ratio equation (Equation 5), we obtain

$$\frac{\mathrm{EX}_{ij}(s)}{\mathrm{IM}_{ij}(s)} = \left\{\left[\frac{n_i(s)}{n_j(s)}\right]^{\mu(s)-1}\frac{1}{\mathrm{TOT}_{ij}(s)}\right\}^{-\varepsilon^{\text{trade}}(s)}\left[\frac{P_j(s)}{P_i(s)}\right]^{-\varepsilon^{\text{trade}}(s)}\left[\frac{\phi_j(s)}{\phi_i(s)}\frac{Y_j}{Y_i}\right], \tag{17}$$

where the term $[n_i(s)/n_j(s)]^{\mu(s)-1}$ represents scale effects from the demand side, and $\mu(s)-1 = [\sigma^{\text{firm}}(s)-1]^{-1}$ is called the scale elasticity in the previous stduies.

Equations (15), (16), and (17) are crucial for understanding the impact of returns to scale on the home market effect. The first thing to note is that under the assumption of symmetric fixed costs in efficiency units across countries, fixed costs have a limited role in determining the direction of home market effects as in previous studies such as Hanson and Xiang (2004). The fixed cost in efficiency units, denoted by $f_C(s)$, disappears in the above equation that determines relative output size. Thus, returns to scale derived from fixed costs do not matter. That is why our primary interest in this paper is scale economies derived from non-constant marginal costs.[16]

Importantly, we get an identical average size of firms across countries in Equation (15) when production and non-production activities have identical cost structures within countries, $c(\cdot,\cdot;s) = \tilde{c}(\cdot,\cdot;s)$.[17] Differences in relative size arise only when production and non-production activities have different cost structures, i.e., they differ in their relative input intensity. Only in the latter case do returns to scale from non-constant marginal cost, $\alpha(s) \neq 1$, affect the relative output in the countries.

[16]The other reason is that using the US data, Kim (2021) provides evidence that the primary source of heterogeneous scale economies across industries is non-constant marginal costs, not fixed costs.

[17]That result holds even if average firm productivity differs across countries.

The role played by different cost structures of production vs. non-production activities has not been discussed previously in this context because most home market effect models assume only one input (labor) that is non-tradable. That model environment forces the production and non-production activities to be identical because the cost function should be homogeneous of degree one with respect to an input price vector. In that case, returns to scale derived from non-constant marginal costs do not matter because the free entry leads to no scale difference across countries, as we discussed.

In contrast, returns to scale derived from non-constant marginal costs can matter when there are labor and capital inputs. Equations (15) and (16) indicate that input intensities play a crucial role in determining the cross-country relative firm size and the impact of returns to scale on the home market effect. Suppose that $\tilde{c}(\cdot,\cdot;s)$ and $c(\cdot,\cdot;s)$ follow Cobb-Douglas functional forms with labor share $\tilde{\vartheta}_l(s) \in [0,1]$ and $\vartheta_l(s) \in [0,1]$, respectively.[18] The degree of homogeneity represents the labor intensities in non-production and production activities. Then, Equations (15) and (16) give

$$\left[\frac{q_i(\bar{\omega}_i;s)/\bar{a}_i(s)}{q_j(\bar{\omega}_j;s)/\bar{a}_j(s)}\right]^{\frac{1}{\alpha(s)}} = \mathrm{TOL}_{ij}^{\tilde{\vartheta}_l(s)-\vartheta_l(s)} \tag{18}$$

$$\mathrm{TOT}_{ij}(s) = \left[\frac{\bar{a}_j(s)}{\bar{a}_i(s)}\right]\mathrm{TOL}_{ij}^{\vartheta_l(s)}\underbrace{\mathrm{TOL}_{ij}^{[\tilde{\vartheta}_l(s)-\vartheta_l(s)][1-\alpha(s)]}}_{\text{Scale economies in production}},, \tag{19}$$

where $\mathrm{TOL}_{ij} = w_i/w_j$ is the terms of labor (the non-tradable input). Inserting the functional forms into Equation (17), the export-import ratio equation can be written as

$$\frac{\mathrm{EX}_{ij}(s)}{\mathrm{IM}_{ij}(s)} = \left\{ \overbrace{\left[\frac{n_i(s)}{n_j(s)}\right]^{\mu(s)-1}}^{\text{Demand-side scale effect}} \left[\frac{1}{\mathrm{TOL}_{ij}^{\vartheta_l(s)}}\frac{\bar{a}_i(s)}{\bar{a}_j(s)}\right] \underbrace{\mathrm{TOL}_{ij}^{[\tilde{\vartheta}_l(s)-\vartheta_l(s)][\alpha(s)-1]}}_{\text{Returns to scale channel}} \right\}^{-\varepsilon^{\text{trade}}(s)}$$

$$\times \left[\frac{P_j(s)}{P_i(s)}\right]^{-\varepsilon^{\text{trade}}(s)} \left[\frac{\phi_j(s)}{\phi_i(s)}\frac{Y_j}{Y_i}\right]. \tag{20}$$

Here, we would obtain the conventional (export-import ratio) gravity equation in the quantitative international trade literature when there is only labor input ($\vartheta_l(s) = 1$) and constant marginal costs ($\alpha(s) = 1$). Equation (20) describes the relationships between industrial home market effect, returns to scale, and input intensities.

If country i is larger than country j, the large country i tends to face a higher labor input price, which indicates an appreciated terms of labor due to higher input demand.[19]

$$\zeta = \frac{\partial \ln \mathrm{TOL}_{ij}}{\partial \ln Y_i/Y_j} > 0 \tag{21}$$

[18] Since $\tilde{c}(\cdot,\cdot;s)$ and $c(\cdot,\cdot;s)$ are homogeneous with degree one, they are homogeneous with degree $1-\tilde{\vartheta}_l(s)$ and $1-\vartheta_l(s)$ with respect to capital, respectively.

[19] See Krugman (1980); Hanson and Xiang (2004) and also Lemma A1 in Appendix A for the details.

Consider industries for which the non-production activity uses labor inputs more intensively than production activities, i.e., $\tilde{\vartheta}_l(s) > \vartheta_l(s)$. Then, an appreciated wage (an increase in terms of labor, TOL_{ij}) means that firm entry/operations are more costly compared to production activities, which restricts firm entry in these industries, and leads to a larger equilibrium firm size/scale in country i compared to country j. In industries with increasing returns to scale, i.e., $\alpha(s) > 1$, firms in country i take advantage of scale economies due to their large production scale, which leads to a depreciation of their terms of trade, i.e., firms in these industries are more competitive in country i relative to country j. Thus, the large country i is concentrated in industries with large returns to scale. In contrast, in industries where the non-production activity uses labor inputs less intensively than production activities, the small country j is concentrated in industries with large returns to scale.

To sum up, the role of returns to scale derived from non-constant marginal costs in shaping industry trade patterns depends on labor intensities in non-production and production activities, which we denote by $\tilde{\vartheta}_l(s)$ and $\vartheta_l(s)$, respectively. From Equation (20), we obtain the following relation:

$$\left[\frac{\partial \ln \mathrm{EX}_{ij}(s)/\mathrm{IM}_{ij}(s)}{\partial \ln \mathrm{TOL}_{ij}}\right] \Big/ \left[\frac{\partial \ln Y_i/Y_j}{\partial \ln \mathrm{TOL}_{ij}}\right] \propto -\vartheta_l(s) + [\tilde{\vartheta}_l(s) - \vartheta_l(s)][\alpha(s) - 1]. \tag{22}$$

Ex-ante, we can expect the impact of returns to scale on the home market effect to be either positive or negative, depending on whether non-production activities are more or less labor-intensive compared to production activities; i.e., it depends on whether $\tilde{\vartheta}_l(s) - \vartheta_l(s) > 0$ or < 0, respectively. This result implies that our empirical analysis needs careful consideration of the cost structures of non-production and production activities when studying the impact of returns to scale on the home market effect.

3. Empirical Strategy

To estimate how the home market effect varies with industries' traits, including returns to scale, markups, trade costs and elasticity, this section introduces a regression model by linearizing our theoretical framework's predictions because it is hard to get the explicit functional form for the solution to Equation (20).

To do that, define the home market effect coefficient by the derivative of the export-import ratio with respect to the relative GDP.

$$\beta(s) = \frac{\partial \ln \mathrm{EX}_{ij,t}(s)/\mathrm{IM}_{ij,t}(s)}{\partial \ln Y_{ij,t}}, \tag{23}$$

where $Y_{ij,t} = Y_{i,t}/Y_{j,t}$ is the relative GDP. There are home market effects in industry s when $\beta(s) > 0$. In contrast, $\beta(s) < 0$ indicates inverse home market effects.

To estimate each industry's home market effect coefficient $\beta(s)$, we begin with the ratio of exports to imports from country i to country j as follows.

$$\ln \frac{\mathrm{EX}_{ij,t}(s)}{\mathrm{IM}_{ij,t}(s)} = \beta(s) \ln Y_{ij,t} + \nu_i(s) - \nu_j(s) + \eta_{i,t} - \eta_{j,t} + \epsilon_{ij,t}(s) \tag{24}$$

The above equation expresses the cross-country differences of the exporter's relative attributes by $\beta(s)\ln Y_{ij}$. The origin and destination's country-time unobserved components, $\eta_{i,t}$ and $\eta_{j,t}$, are any time-varying differences across industries within a country such as macroeconomic conditions. The origin and destination unobserved components of each industry are denoted by $\nu_i(s)$ and $\nu_j(s)$, which relate to differences in sectoral cross-country tastes as well as comparative advantages of country i over country j in industry s. The systemic origin-destination components across bilateral trade partners such as distance, trade agreements, language difference, borders, and so on disappear, but $\epsilon_{ij,t}(s)$ contains the unsystematic origin-destination components.

Section 2 predicts that the home market coefficients are a function of industry traits. To test the prediction, we introduce an interaction term between the relative GDP and industry characteristics.

$$\ln \frac{\mathrm{EX}_{ij,t}(s)}{\mathrm{IM}_{ij,t}(s)} = [\mathbf{trait}'_{ij,t}(s)\mathbf{b}(s)] \times \ln Y_{ij,t} + \nu_i(s) - \nu_j(s) + \eta_t(s) - \eta_t(s) + \epsilon_{ij}(s), \tag{25}$$

where the vector $\mathbf{trait}_{ij}(s)$ includes a set of control variables to account for confounding industry-level factors, and $\mathbf{b}(s)$ is the corresponding vector of coefficients. Most importantly, Equation (22) yields that the marginal impact of returns to scale on the home market effect depends on labor intensities in non-production and production activities:

$$\frac{\partial\beta(s)}{\partial\alpha(s)} \propto \tilde{\vartheta}_l(s) - \vartheta_l(s). \tag{26}$$

Thus, $\mathbf{trait}_{ij,t}(s)$ includes $[\tilde{\vartheta}_l(s) - \vartheta_l(s)]\alpha(s)$ instead of $\alpha(s)$. Our main hypothesis is that its coefficient is positive.

Along the lines of the home market literature, there are three pillars of home market effects; economies of scale, product differentiation, and trade costs. Thus, our regression should control for product differentiation and trade costs as determinants of the home market effect coefficient $\beta(s)$.

According to the conventional home market effect theory, firms prefer locating in large countries due to cost advantages, which is much stronger when the market is more differentiated. In other words, large countries tend to be concentrated in industries with more differentiated products compared to industries with less differentiated products. Thus, large countries are more concentrated in industries with smaller elasticities of substitution at the product and national levels.

The home market effect in an industry increases with its degree of markups due to the one-to-one negative relationship between the markup and elasticity of substitution across products (product-level product differentiation): $\mu(s) = \sigma^{\text{firm}}(s)/[\sigma^{\text{firm}}(s) - 1]$.

In a similar way, the industrial home market effect decreases with trade elasticity because of the relationship between trade elasticity and the national-level elasticity of substitution: $\varepsilon^{\text{trade}}(s) = 1-\sigma^{\text{natl}}(s) < 0$. Industries with high trade elasticities (small absolute value) witness higher product differentiation and a larger home market effect.

In addition to the above channel (national-level product differentiation), industrial heterogeneity in trade elasticity potentially plays a role in influencing industry trade patterns in another way: effective trade costs. This channel generates an opposite impact of trade elasticity on the home market effect. The effective trade cost in industry s is defined by $x_{ij}(s) = [\tau_{ij}(s)]^{\sigma^{\text{natl}}(s)-1} = [\tau_{ij}(s)]^{-\varepsilon^{\text{trade}}(s)}$, and is decreasing with the trade elasticity. The conventional home market effect theory predicts that industries with high effective trade costs tend to be more concentrated in large economies compared to industries with low effective trade costs because firms prefer locations where they can save more on trade costs. Therefore, we would expect a negative relationship between the trade elasticity and the home market effect from this channel, for a fixed degree of product differentiation.

To sum up, the industrial home market effect — the impact of country size on trade surplus and the location of industries — is also a function of markups, trade elasticity, effective trade costs, and their interactions in addition to returns to scale and the difference between non-production and production cost structures. Hence, $\texttt{trait}_{ij,t}(s)$ should include these variables.

4. Data and Variables

4.1. International Trade Flows

We aim to estimate the impact of market structure on the home market effect by comparing the country size impacts of U.S. trading partners. We fix the destination country j to be the U.S. economy. The U.S. manufacturing industries' bilateral export and import flows are from the U.S. Census Bureau, which is constructed by Schott (2008) using the concordances from Bartelsman and Doms (2000) and Pierce and Schott (2009). The real GDP data are from the Penn World Table 9.1.

Home market effects are the central prediction of the new trade theory that economists use to explain intra-industry trade between advanced countries. Therefore, our empirical investigation focuses on advanced economies. We choose all 21 non-OPEC advanced economies among the top 40 U.S. trading partners.[20] Our benchmark analysis focuses on the five-digit NAICS classification that contains 166 industries. The panel data set contains data for 166 manufacturing industries in 21 countries from 1989 through 2011, which yields a sample of 21 $\times$ 166 $\times$ 23 = 80,178 observations. See Appendix C for data-related details.

Some industries have zero bilateral trade flows. To account for these, we modify the dependent variables in the regression Equation (25) with fixed destination j = US as follows.

$$v_{i,t}(s) = \ln \frac{\text{ex}_{i\text{US},t}(s) + \$1}{\text{im}_{i\text{US},t}(s) + \$1} \tag{27}$$

Figure 1 plots summary statistics of the logarithmic export-import ratio $v_{i,t}(s)$. The distribution is skewed to the left. The mean and median values tend to rise in our sample. The standard deviations are 3.4 on average and falls over time from 3.8 to 3.2. In Appendix Figures D2 and D3, the kernel densities of each year do not drastically change over time. Within-industry standard deviations are around 2.7 on average (over-time).[21] These statistics imply that within- and between-industry variations lead to approximately 80% and 20% of overall cross-sectional variations in trade patterns measured by log export-import ratio's standard deviation.

[20]The sample countries' ISO codes are AUS, AUT, BEL, CAN, CHE, DEU, DNK, ESP, FIN, FRA, GBR, HKG, IRL, ISR, ITA, JPN, KOR, NLD, SGP, SWE, and TWN. The U.S. trade volume (i.e., exports plus c.i.f. imports) with this group declines during the sample period: from 77% (1989) to 51% (2011) of the total manufacturing trade volume. Appendix Figure D1 presents the share of trade volume to the entire manufacturing sector in the U.S. over time.

[21]Their maximum and minimum values are 0.7 and 10.0, respectively.

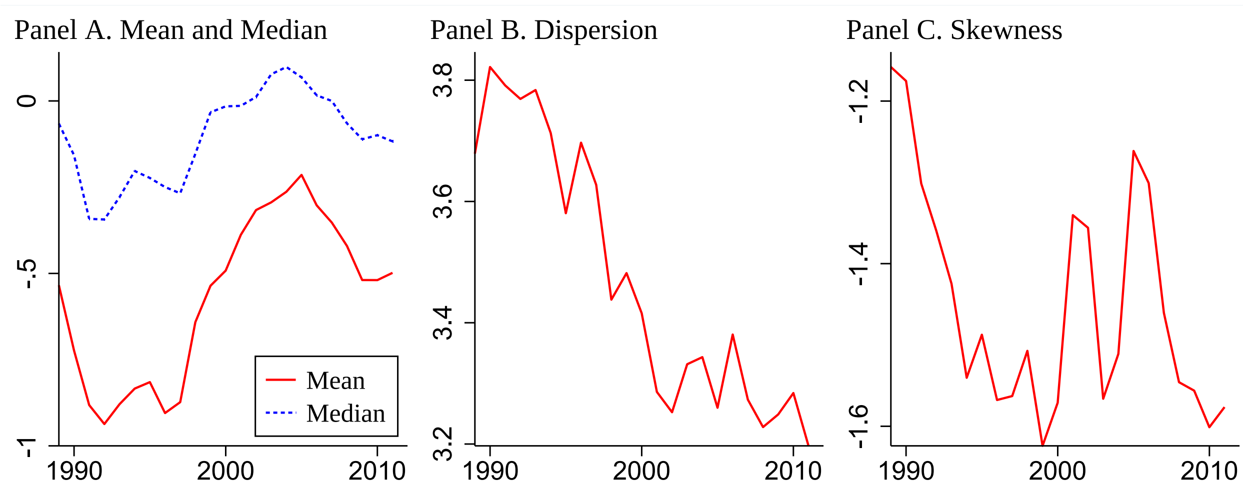

Figure 1: Summary Statistics of Logarithmic Export-Import Ratio to the US

Notes: The logarithmic export-import ratio of trading partners to the U.S. denoted by $v_{i,t}(s)$ is defined in Equation (27). The dispersion and skewness are measured by $\mathrm{m}_2^{1/2}$ and $\mathrm{m}_3\mathrm{m}_2^{-3/2}$, respectively, where m_n is the sample n-th central moment.

.

4.2. Industry Traits

We estimate industry market structure characteristics (returns to scale and markups) using industry-level data. As an indicator of product differentiation, the previous literature on the home market effect includes an estimated elasticity of substitution or a proxy for product differentiation using demand-side data. However, these measures are hard to get for many industries due to data limitations. Therefore, in contrast to the previous literature, we use markups from supply-side data that are easily accessible for many narrowly defined industries. In addition, as discussed above, we use trade elasticity as a variable affecting product differentiation and effective trade costs. Therefore, we include effective trade costs and trade elasticity separately, as well as their interaction. The data used to measure the various industry traits are detailed below.

Returns to Scale and Markups (α and μ). To measure the degree of returns to scale $\alpha(s)$, we estimate the output elasticity of cost-shared inputs (i.e., cost-weighted growth of inputs). For industry s, the degree of returns to scale $\alpha(s)$ is

$$\Delta y_t(s) = \alpha(s)\Delta x_t(s) + \varpi_0 + \varpi_1\Delta e_t(s) + \varpi_2 t + \text{errors}_t(s), \tag{28}$$

where $\Delta y_t(s)$ and $\Delta x_t(s)$ are the industry s growth rates (i.e., log-differences between t and $t-1$) of real output and cost-shared inputs, respectively. The growth rate of real energy spending $\Delta e_t(s)$

Table 1: Industry Traits

Industry Characteristics	Sample Period	N	Mean	Standard Deviation	Percentiles 10%	50%	90%
Returns to scale	–	166	0.922	0.364	0.419	0.987	1.293
Wage elasticity of non-production and production labor							
Non-production	–	166	-0.507	0.354	-1.000	-0.433	0.000
Production	–	166	-0.279	0.266	-0.761	-0.220	0.000
Gap: $\varepsilon_{np}^{\text{wage}}(s) - \varepsilon_p^{\text{wage}}(s)$	–	166	-0.228	0.421	-0.774	-0.212	0.228
Interaction b/w Returns to scale and gap: $\alpha(s)[\varepsilon_{np}^{\text{wage}}(s) - \varepsilon_p^{\text{wage}}(s)]$							
	–	166	-0.217	0.412	-0.687	-0.162	0.181
Markup	89–00	166×12	1.215	0.574	0.494	1.223	1.831
	01–11	166×11	1.536	0.866	0.601	1.465	2.350
Cost to revenue ratio	89–00	166×12	0.786	0.130	0.635	0.793	0.932
	01–11	166×11	0.648	0.183	0.460	0.642	0.825
Trade elasticity	89–00	166×12	-8.931	5.408	-15.923	-7.292	-4.300
	01–11	166×11	-8.849	5.217	-15.962	-7.366	-4.404
Labor cost share	89–00	166×12	0.226	0.091	0.114	0.223	0.337
	01–11	166×11	0.209	0.089	0.102	0.201	0.313
Material cost share	89–00	166×12	0.681	0.097	0.560	0.683	0.797
	01–11	166×11	0.703	0.098	0.585	0.705	0.826
Effective trade costs	89–00	166×12×21	1.062	0.057	1.013	1.049	1.120
	01–11	166×11×21	1.059	0.054	1.013	1.046	1.112
Industry relative size	89–00	166×12×21	0.575	0.755	0.103	0.341	1.107
(percent)	01–11	166×11×21	0.576	0.975	0.072	0.315	1.095

Notes: Returns to scale and wage elasticities are constant over time across countries. Markups, cost to revenue ratio, trade elasticity, and cost shares are identical across origin countries.

is a proxy for factor utilization.[22] Second, as shown in Hall (1988), cost minimization implies

$$\mu_t(s) = \alpha_t(s)/\lambda_t(s), \tag{29}$$

where $\lambda_t(s)$ is the cost to revenue ratio, which equals the returns to scale divided by markups in equilibrium.

Taking the U.S. economy as destination country j, we estimate industry traits including the market structure of country j by collecting data on industry output, inputs, costs, and their deflators from the NBER-CES Manufacturing Industry Database for the years 1958 through 2011. We use a production function with three factors — capital, labor, and materials — as in previous empirical research related to returns to scale estimation such as Basu and Fernald (1997), Lee (2007), and many others.[23] The uninstrumented estimator is biased due to the relationship between productivity and input demand. To control for potential endogeneity, we use demand-side instruments such as oil price shocks, the president's party, government defense spending, and monetary policy shocks; these instruments are widely used in the literature. We restrict an industry's estimated degree of returns to scale to be non-negative, which does not impact our results. See Appendix C for details

[22]Removing $\Delta e_t(s)$ does not change any main result related to heterogeneity across industries in the empirical estimation. This, however, causes an upward bias in the estimated returns to scale on average.

[23]See Basu and Fernald (1997) for the biased returns to scale estimator with a value-added production function (without materials).

on data and variable construction.

Table 1 reports the U.S. manufacturing industries' estimated characteristics. On average, the estimated results show slightly decreasing returns to scale. This result is due to the presence of some industries with very low estimates, i.e., close to zero. (See Appendix Figure D4's Panel A for the histogram with kernel density.) Estimated markups increase from 1.22 to 1.54 due to the sharp decline in the cost to revenue ratio. Also, a markup increase in the top percentile leads to an increase in the cross-sectional standard deviation (See Appendix Figure D5 for the histogram and kernel density).

Equation (29) implies a high correlation between returns to scale and markups by construction. In our sample, their correlation coefficient is very highly positive, approximately 70%, which could generate a problem in identifying their marginal impacts in the regression results.

Gap Between Wage Elasticity of Non-production and Production Labor ($\varepsilon_{np}^{\textbf{wage}} - \varepsilon_{p}^{\textbf{wage}}$). We use the gap between wage elasticities of non-production and production labor as a proxy for the gap between labor intensities in non-production and production activities.

Equation (15) shows the vital role played by the labor input intensities of non-production and production activities in determining the direction of the impact of returns to scale across countries. We cannot directly measure the labor input intensity in non-production and production activities due to data limitations. Therefore, Shepard's lemma guides our estimation of the gap between the input intensity of non-production and production labor.

We use labor as a non-tradable input. In Equation (6), the first derivative of non-production and production costs with respect to wage yields that the non-production and production labor demands are $l_{np} = \text{fc}[\partial\tilde{c}(w,r)/\partial w]$ and $l_p = q^{1/\alpha}[\partial c(w,r)/\partial w]$, respectively. As an approximation, consider the Cobb-Douglas functional form. This gives us: $\ln\tilde{c}(w,r) = \tilde{\vartheta}_l \ln w + (1-\tilde{\vartheta}_l)\ln r + \text{constant}$ and $\ln c(w,r) = \vartheta_l \ln w + (1-\vartheta_l)\ln r + \text{constant}$, where $\tilde{\vartheta}_l$ and ϑ_l are the labor intensities (or labor cost shares) in non-production and production activities, respectively. Then, the wage elasticities of non-production and production labor are $\varepsilon_{np}^{\text{wage}} = \tilde{\vartheta}_l - 1$ and $\varepsilon_{p}^{\text{wage}} = \vartheta_l - 1$, respectively. Therefore, the gap between wage elasticities measures the gap between labor intensities in non-production and production activities.

Within each industry s, denote $\tilde{s}$ as the sub-industry (the six-digit NAICS level). We estimate the wage elasticities of non-production and production labor from the following panel regressions.

$$\Delta \ln l_{np,t}(\tilde{s}) = \varepsilon_{np}^{\text{wage}}(s)\Delta \ln w_t(\tilde{s}) + [\mathbf{w}_{-1}(\tilde{s});\ \Delta \ln y_t(\tilde{s})]'\mathbf{c}_{np}(s) + \delta_t(\tilde{s}) + \text{errors}_t(\tilde{s}) \tag{30}$$

$$\Delta \ln l_{p,t}(\tilde{s}) = \varepsilon_{p}^{\text{wage}}(s)\Delta \ln w_t(\tilde{s}) + \mathbf{w}_{-1}'(\tilde{s})\mathbf{c}_{p}(s) + \delta_t(\tilde{s}) + \text{errors}_t(\tilde{s}), \tag{31}$$

where $\mathbf{w}_{-1}(\tilde{s})$ is the price vector excluding the wage $w_t(\tilde{s})$ in the sub-industry $\tilde{s}$. The regressions include the sub-industry and time fixed effects: $\delta_t(\tilde{s})$. Since the cost function is homogeneous in degree one with respect to input prices, the wage elasticities are located in $[-1,\ 0]$. Thus, we replace the estimates with negative one or zero when they are smaller than negative one or larger than zero, respectively.

Table 1 reports the estimated wage elasticities and their gaps. The results imply that the mean and median labor input intensity gaps are around –0.22 and –0.21. Many industries use labor inputs more intensively in production activities compared to non-production activities. The gap is right-skewed (i.e., positively skewness). The wage elasticity gap is weakly correlated with returns to scale (the correlation coefficient is -4.8%). Thus, the interaction between returns to scale and the gap is negative on average. Appendix Figure D4 plots their histograms and kernel densities.

Trade Elasticity ($\varepsilon^{\textbf{trade}}$). We use Fontagne et al. (2020)'s trade elasticities at the product-level. The trade elasticity should be negative but not bounded below (not bounded above in terms of absolute values). Using the concordances from Bartelsman and Doms (2000) and Pierce and Schott (2009), we calculate the industry-level trade elasticity by averaging the six-digit HS product-level estimates in each five-digit NAICS industry. Due to compositional effects across HS product classification, our trade elasticity measures are time-varying. Such time-variations are not sizable. The statistics of trade elasticity do not change over time significantly in Table 1. (See Appendix Figure D6 for the histograms and kernel densities.) To minimize the impact of outliers, we winsorize the measurements at the top and bottom 1% levels.

Effective Trade Costs (x). We measure the effective trade cost from an origin country to the US as the ratio of cost, insurance, and freight (c.i.f.) value of imports to the free on board (f.o.b). value of imports. Thus, the value is not less than one but is not bounded above. To minimize the impact of outliers, we winsorize the measurements at the top and bottom 1% levels. See Appendix Figure D7 for the histograms and kernel densities.

Other Traits. Table 1 also reports other industry traits: durability, industry relative size, labor, and material cost share. In our sample, the dispersion of industry size relative to the total manufacturing industries rises over time. The labor cost share falls, the material cost-share rises, and the capital cost-share is small at around 9% on average. These results are consistent with previous empirical studies, for examples, Karabarbounis and Neiman (2013) and De Loecker et al. (2020). See Appendix C for details on their construction.

5. Estimation Results

To estimate the impact of market structure on the home market effect, we consider the following regression based on Equations (25) and (27) with fixed destination j = US as our benchmark specification:

$$v_{it}(s) = [\mathbf{trait}'_{i\text{US},t}(s)\mathbf{b}(s)] \times \ln Y_{i\text{US},t} + \tilde{\nu}_i(s) + \tilde{\eta}_t(s) + \epsilon_{i\text{US},t}(s), \tag{32}$$

where our regression includes the origin country-industry and country-time fixed effects denoted by $\tilde{\nu}_{i,t}(s) = \nu_{i,t}(s) - \nu_{\text{US},t}(s)$ and $\tilde{\eta}_{i,t}(s) = \eta_{i,t}(s) - \eta_{\text{US},t}(s)$. These fixed effects allow us to control for any time-invariant cross-industry and time-varying country differences that might influence intra-industry bilateral trade patterns. Since $\ln Y_{i\text{US},t}$ is identical across industries in origin country i over time, the country-time fixed effects control for heterogeneous home market effect coefficients (slopes of $\ln Y_{i\text{US},t}$) across origin countries derived from unobserved county-specific characteristics. Based on our discussion in previous sections, the industry characteristics that we focus on are returns to scale, markups, trade costs, wage elasticity gap, trade elasticity, and their interactions. We also consider other traits such as durability of goods, industry size, labor and material cost share.

Table 2: Home Market Effect, Product Differentiation, and Trade Costs

Dependent variable: $v_{it}(s)$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Returns to scale								
$\alpha(s) \times \ln Y_{i\text{US},t}$	5.03	4.89	4.29	3.91	4.21	3.74	4.05	4.20
	(2.98)	(2.88)	(3.08)	(2.97)	(3.10)	(2.63)	(2.76)	(3.11)
Markups								
$\mu_t(s) \times \ln Y_{i\text{US},t}$	2.63^{***}		2.61^{***}		2.64^{***}		2.70^{***}	
	(0.82)		(0.78)		(0.78)		(0.89)	
Effective trade costs								
$x_{i\text{US},t}(s) \times \ln Y_{i\text{US},t}$	66.18^{**}	64.84^{**}	66.63^{**}	65.50^{**}	66.51^{**}	35.30	35.70	60.66^{**}
	(23.33)	(22.91)	(23.64)	(23.24)	(23.57)	(21.33)	(20.94)	(21.41)
Trade elasticity								
$\varepsilon_t^{\text{trade}}(s) \times \ln Y_{i\text{US},t}$		0.04		-0.05	-0.12	2.51^{**}	2.50^{**}	
		(0.17)		(0.19)	(0.20)	(0.98)	(0.90)	
Interaction between trade costs and elasticity								
$x_{i\text{US},t}(s) \times \varepsilon_t^{\text{trade}}(s) \times \ln Y_{i\text{US},t}$						-2.34^{**}	-2.39^{**}	-0.46^{*}
						(0.90)	(0.85)	(0.24)
Overall product differentiation								
$[-\varepsilon_t^{\text{trade}}(s)] \times [\mu_t(s) - 1] \times \ln Y_{i\text{US},t}$								0.17^{**}
								(0.07)
Other controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Country-industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	80178	80178	80178	80178	80178	80178	80178	80178
R-squared	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67

Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors (three-way) clustered within country (21), industry (166), and year (23) are in parentheses. Regressions with other controls include interactions of log relative GDP with durability dummy, industry relative size, labor and material cost share.

Tables 2 and 3 report the regression results obtained from estimating Equation (32). In Table 2, regressions include the degree of returns to scale without the consideration of non-production and production cost structure differences. Then, all coefficients of returns to scale are insignificant in our specifications at the 10% level. Section 3 discussed that the impact of returns to scale on the home market effect potentially depends on labor intensities in production and non-production activities. Our theoretical model in Section 2 supports this argument. Thus, we directly test this hypothesis using a regression specification that includes the interaction term $\alpha(s) \times [\varepsilon_{np}^{\text{wage}}(s) - \varepsilon_{p}^{\text{wage}}(s)] \times \ln Y_{i\text{US},t}$ in Table 3.

In Table 3, we consider the gap between the wage elasticities of non-production and production labor inputs ($\varepsilon_{np}^{\text{wage}} - \varepsilon_{p}^{\text{wage}}$) in order to account for differences in labor input intensities in non-production and production activities ($\tilde{\vartheta}_l - \vartheta_l$), following our discussion in Section 4. The estimated coefficients of the interaction between returns to scale and the wage elasticity gap are significantly positive at the 1% level in all columns. This means that when there is a positive wage elasticity gap, i.e., the labor is used more intensively in non-production activities, returns to scale positively impact the home market effect. When the non-production activity is more labor-intensive than the production activity as in the model of Grossman and Helpman (1991), larger countries tend to be

Table 3: Home Market Effect, Returns to Scale, and Non-Tradable Inputs

Dependent variable: $v_{it}(s)$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Interaction between returns to scale and wage elasticity gap								
$\alpha(s) \times [\varepsilon_{np}^{\text{wage}}(s) - \varepsilon_{p}^{\text{wage}}(s)] \times \ln Y_{i\text{US},t}$	4.93***	5.22***	5.21***	5.61***	5.16***	5.58***	5.12***	4.86***
	(1.58)	(1.54)	(1.58)	(1.63)	(1.57)	(1.71)	(1.65)	(1.47)
Markups								
$\mu_t(s) \times \ln Y_{i\text{US},t}$	2.60***		2.56***		2.60***		2.66***	
	(0.84)		(0.79)		(0.80)		(0.91)	
Effective trade costs								
$x_{i\text{US},t}(s) \times \ln Y_{i\text{US},t}$	65.91**	64.58**	66.42**	65.30**	66.29**	35.06	35.45	60.47**
	(23.30)	(22.86)	(23.59)	(23.19)	(23.52)	(21.07)	(20.70)	(21.31)
Trade elasticity								
$\varepsilon_t^{\text{trade}}(s) \times \ln Y_{i\text{US},t}$		0.04		-0.05	-0.12	2.52**	2.50**	
		(0.18)		(0.20)	(0.21)	(0.97)	(0.90)	
Interaction between trade costs and elasticity								
$x_{i\text{US},t}(s) \times \varepsilon_t^{\text{trade}}(s) \times \ln Y_{i\text{US},t}$						-2.35**	-2.40**	-0.45*
						(0.90)	(0.85)	(0.25)
Overall product differentiation								
$[-\varepsilon_t^{\text{trade}}(s)] \times [\mu_t(s) - 1] \times \ln Y_{i\text{US},t}$								0.17**
								(0.07)
Other controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Country-industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	80178	80178	80178	80178	80178	80178	80178	80178
R-squared	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67

Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors (three-way) clustered within country (21), industry (166), and year (23) are in parentheses. Regressions with other controls include interactions of log relative GDP with durability dummy, industry relative size, labor and material cost share.

more concentrated in industries with higher returns to scale compared to industries with smaller returns to scale. The opposite holds for industries in which the non-production activity is less labor-intensive than the production activity. This is consistent with our discussion summarized by Equation (22). This result highlights the importance of carefully considering the production structure and the relative labor input intensity use in non-production vs. production activities, when studying the home market effect.

As discussed in Section 3, markups represent product differentiation. In the regressions of Columns (1) and (3), the industry characteristics vector $\texttt{trait}_{i\text{US},t}$ includes markups but excludes trade elasticity and related terms. In contrast, Columns (2), (4), and (6) include the trade elasticity and its interaction term but not markups because markups are tightly related to returns to scale measure by construction as Section 4 discussed. Columns (5) and (7) include markups as well as trade elasticity. In regressions of Columns (3) – (7), we control for labor cost shares, material cost shares, size of the industry, and durability of products. All regressions control for the returns to scale, effective trade costs, and fixed effects. In Column (8), we control for a degree of overall product differentiation by including a combination of national- and product-level degree of substitutions, $[-\varepsilon_t^{\text{trade}}(s)] \times [\mu_t(s) - 1]$, instead of controlling for the markup and trade elasticity separately.

According to the second row of Tables 2 and 3, the average impact of markups on the home market effect is positive at the 1% significance level. In Columns (1) – (5), the coefficients of effective trade costs are positive at the 1% significance level. In Columns (6) and (7), the coefficients are insignificant at the 10% level, but the estimates of the interaction between the trade costs and elasticity are negative at the 5% significance level in these specifications. Since the upper bound of trade elasticity is -1, we conclude that effective trade costs' net impact is significantly positive at the 5% level in all specifications. These results are consistent with the previous literature, such as Hanson and Xiang (2004). In international trade markets, a larger country has an advantage in industries with more differentiated products and more considerable trade costs.

In Columns (2), (4), and (5) of Tables 2 and 3, the impact of trade elasticity on the home market effect is statistically insignificant at the 10% level. This result can be rationalized by the fact that a lower trade elasticity leads to higher product differentiation, but also lowers effective trade costs ($x = \tau^{-\varepsilon^{\text{trade}}}$), which would have an opposing effect on the home market effect. To control for the second channel, Columns (6) and (7) include the interaction between trade costs and elasticity. For these specifications, we obtain significantly positive coefficients for the trade elasticity excluding the trade cost channel (in the last row), which is consistent with the theory.

6. Robustness

This section performs robustness check exercises. First, we consider cross-sectional regressions instead of panel regressions. Second, we include/exclude a wide range of different groups of countries, for example, developing countries (e.g., China, India, Mexico), European countries (e.g., Germany, France, Italy), and small–specialized countries (e.g., Israel, Hong Kong, Singapore, Ireland). Third, we discuss measurement issues of market structure estimates (exclusively from the US industry-level data) and introduce random slope models. Lastly, we consider (a) the subsample after 1995, (b) the subsample excluding zero trade flows, and (c) the uninstrumented estimated returns to scale.

6.1. Cross-Sectional Analysis

We use panel regression analyses to consider both within- and between-industry variation. However, cross-sectional regressions would be more appropriate when within-industry changes over time are not random, since in that case, time variation introduces additional endogeneity concerns. Thus, we consider the following regression equation in each year.

$$v_i(s) = [\mathbf{trait}'_{i\text{US}}(s)\mathbf{b}(s)] \times \ln Y_{i\text{US}} + \tilde{\nu}_{i\text{US}}(s) + \epsilon_{i\text{US}}(s) \tag{33}$$

Figure 2 plots the year-by-year cross-sectional estimated coefficients and their 90% confidence intervals for our main coefficient of interest; impact of scale economies in production on the home market effect, i.e., the coefficient of $[\varepsilon_{np}^{\text{wage}}(s) - \varepsilon_{p}^{\text{wage}}(s)] \times \ln Y_{i\text{US},t}$. In Panels A and B, we consider industry fixed effects. Panels B and C plot the results with both industry and country fixed effects. Panels A and C regressions (specifications I and III) include interaction between returns to scale and wage elasticity gap, markups, effective trade costs, and other controls, corresponding to Column (3) in Table 3. Panels B and D regressions (specifications II and IV) include the interaction between returns to scale and wage elasticity gap, effective trade costs, trade elasticity, interaction

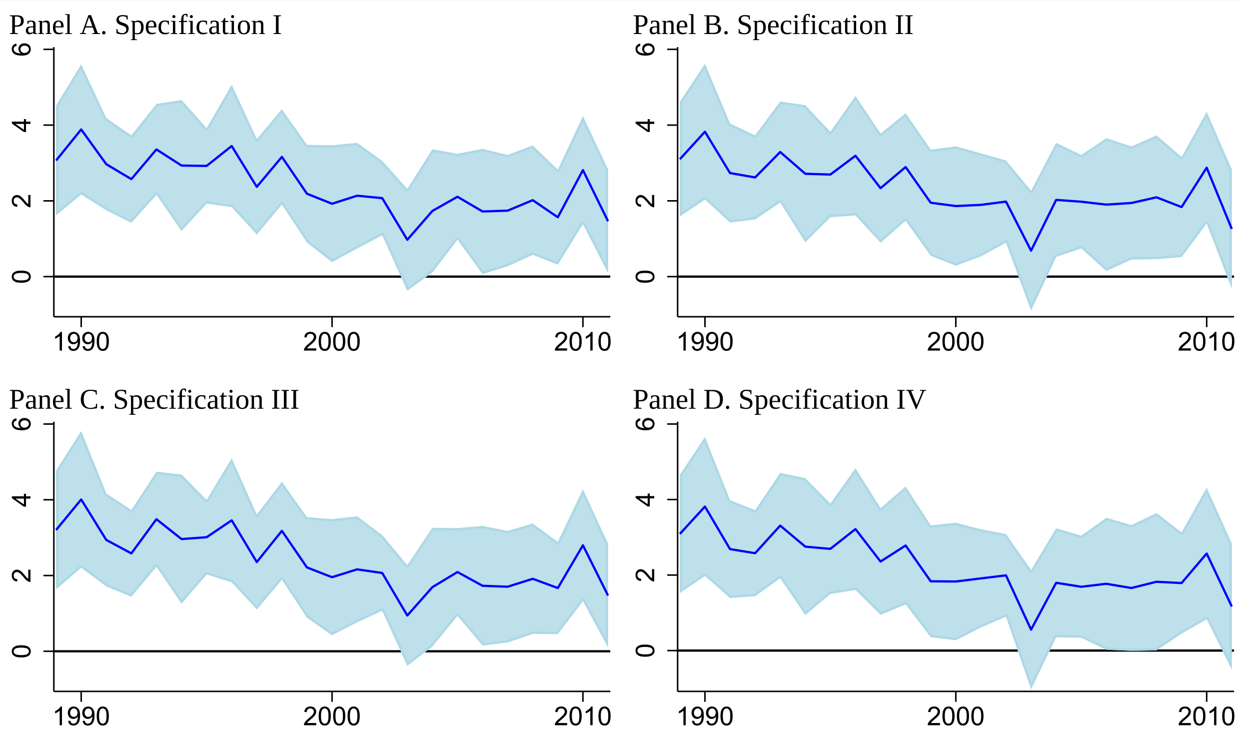

Figure 2: Cross-Sectional Estimates: Returns to Scale Channel

Notes: The figures plot the year-by-year estimated coefficients of the interaction between returns to scale and wage elasticity gap, denoted by $\alpha(s)[\varepsilon_{np}^{\text{wage}}(s) - \varepsilon_{p}^{\text{wage}}(s)] \times \ln Y_{i\text{US},t}$, and their 90% confidence intervals with clustered standard errors at the country level, where panels A – B regressions include industry fixed costs, and Panels C – D regressions include country fixed effects and industry fixed effects. Panels A and C regressors are corresponding to Column (3) in Table 3. Panels B and D regressors are corresponding to Column (6) in Table 3.

.

between trade costs and elasticity, and other controls, corresponding to Column (6) in Table 3. We find evidence for positive impacts of returns to scale on home market effects when non-production activities use labor more intensively than production activities. In Figure 2, specifications I – IV yield positive estimated coefficients 2.40, 2.33, 2.42, and 2.25, respectively, . All estimated coefficients are positive in the range of [0.55, 4.01]. They are statistically significant at the 10% level, except for 2003 (Panels A – D), 2007, and 2011 (Panels B and D).

6.2. Different Groups of Countries

In Table 4, we consider three different sample groups of countries to understand Section 5's results better and check the robustness.

Developing Countries. Our framework is based on the new trade theory that economists use primarily to explain intra-industry trade between advanced countries. Therefore, our empirical investigation focuses on advanced countries in the previous sections. In manufacturing sectors, some developing countries have however became primary trade partners with the US economy during

Table 4: Robustness Checks: Different Groups of Countries

Dependent variable: $v_{it}(s)$	(1) Including Developing Countries	(2)	(3) Non-European Countries Only	(4)	(5) Excluding Small & Specialized Countries	(6)
Interaction between returns to scale and wage elasticity gap						
$\alpha(s) \times [\varepsilon_{np}^{\text{wage}}(s) - \varepsilon_{p}^{\text{wage}}(s)] \times \ln Y_{i\text{US},t}$	1.98	1.75	3.79***	3.75***	5.24***	5.58***
	(1.23)	(1.15)	(0.91)	(0.99)	(1.54)	(1.66)
Markups						
$\mu_t(s) \times \ln Y_{i\text{US},t}$	1.25***		2.52		2.31***	
	(0.41)		(1.47)		(0.70)	
Effective trade costs						
$x_{i\text{US},t}(s) \times \ln Y_{i\text{US},t}$	49.55***	19.81	54.08	26.16	59.99**	29.17
	(15.16)	(14.25)	(34.13)	(25.47)	(21.55)	(19.49)
Trade elasticity						
$\varepsilon_t^{\text{trade}}(s) \times \ln Y_{i\text{US},t}$		2.32***		2.20**		2.50**
		(0.53)		(0.65)		(1.01)
Interaction between trade costs and elasticity						
$x_{i\text{US},t}(s) \times \varepsilon_t^{\text{trade}}(s) \times \ln Y_{i\text{US},t}$		-2.38***		-2.17**		-2.31**
		(0.39)		(0.81)		(0.95)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Country-industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	103086	103086	30544	30544	61088	61088
R-squared	0.68	0.68	0.74	0.74	0.66	0.66

Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors (three-way) clustered within country (27 in Columns (1) – (2), 8 in Columns (3) – (4), and 16 in Columns (5) – (6)), industry (166), and year (23) are in parentheses. This table repeats the analysis in Columns (4) and (5) in Table 3 using alternative specifications to check the robustness of our main results. Regressions with other controls include interactions of log relative GDP with durability dummy, industry relative size, labor and material cost share.

our sample period (1989 – 2011). Thus, we include Brazil, China, India, Malaysia, Mexico, and Thailand, leading to a sample size increase from 80,178 to 103,086 observations. According to Appendix Figure D1, we observe that the six developing countries' trade share in the US manufacturing sector sharply increases over time. Our benchmark sample and the six countries cover almost 90% of the US manufacturing sector trade.

According to Columns (1) and (2) in Table 4, adding developing countries does not affect the demand-side variables. The estimated coefficients related to markups, trade costs, trade elasticity, and their interaction terms conclude the same results in Tables 2 and 3. The estimated coefficients of interaction terms between returns to scale and wage elasticity gap are positive in Columns (1) and (2), but their p-values are 10.7% and 12.8%, which are higher than 10%. The R-squared values are lower than our benchmark sample regressions: 68% vs. 76%. These results are not surprising because our theoretical predictions are more appropriate for accounting for advanced economies' intra-industry trade patterns rather than inter-industry trade derived from comparative advantages.

Non-EU Countries. European Union members comprise more than half of the original sample of countries (13 out of 21). Trade between EU members is more important than their trade with the US, which may lead to systemic differences between EU and non-EU countries in determining production location and home market effects in the US market.

In line with these concerns, Columns (3) and (4) in Table 4 report regression results with only the 8 non-EU countries in our sample.[24] Column (4)'s results are consistent with the estimation results in Table 3. The coefficient associated with our main variable of interest — returns to scale effect — is positive at the 1% significance level. The coefficients related to markups and effective trade costs are positive but insignificant in Column (3).

Small Countries and Industrial Policies. We exclude some small countries as these could be specialized in certain industries due to specific industrial polices. This would cause exogenous impacts on location decisions and home market effects. Even though we consider county-industry fixed effects, this channel would not be accounted for if industrial policies of individual countries changed during the sample period 1989 – 2011.

To ensure that our results are not driven by these factors, regressions in Columns (5) and (6) in Table 4 exclude small countries.[25] The estimated coefficients are consistent with the estimation results in Table 3.

6.3. Measurement Issues

Many industry traits are measured using the destination country (US) data because it is difficult to collect narrowly defined industry-level production data for other countries. However, industry traits could differ across countries. Denote the difference by $\text{me}_{i\text{US}}(s)$ and assume that it is linearly separable and time-invariant. Then, we obtain the following regression equation from Equation (32).

$$\begin{aligned} v_{it}(s) &= [\mathbf{trait}'_{i\text{US},t}(s)\mathbf{b}(s) + \text{me}_{i\text{US}}(s)] \times \ln Y_{i\text{US},t} + \tilde{\nu}_i(s) + \tilde{\eta}_{i,t} + \epsilon_{i\text{US},t}(s) \\ &= [\mathbf{trait}'_{i\text{US},t}(s)\mathbf{b}(s)] \times \ln Y_{i\text{US},t} + \text{me}_{i\text{US}}(s) \times \ln Y_{i\text{US},t} + \tilde{\nu}_i(s) + \tilde{\eta}_{i,t} + \epsilon_{i\text{US},t}(s), \end{aligned} \tag{34}$$

where $\text{me}_{i\text{US}}(s) \times \ln Y_{i\text{US},t}$ represents a random slope. Due to computational costs, we consider demeaned variables within an industry-country level instead of industry-country fixed effects.

Table 5 reports the estimated coefficients with multi-level random coefficient models for two specifications. First, in Columns (3) – (4), we assume that measurement errors, $\text{me}_{i\text{US}}(s)$, are different for each country-industry pair, generating different slopes for each country-industry pair. Alternatively, we assume $\text{me}_{i\text{US}}(s) = \text{me}_{i\text{US}} + \text{me}(s)$ in Columns (1) – (2). This implies two sources of random slopes: one due to the country and another due to the industry. In both cases, the estimated coefficients are consistent with the estimation results in Table 3.

6.4. Additional Robustness Exercises

After the World Trade Organization. We account for structural changes following the establishment of the World Trade Organization (WTO) that officially commenced on 1 January 1995. The WTO prohibits discrimination between trading partners. Therefore, as a robustness check,

[24] The ISO code of non-EU members are AUS, CAN, CHE, HKG, ISR, JPN, KOR, SGP, and TWN.

[25] The ISO codes of small countries excluded in the benchmark sample are CHE, ISR, IRL, HKG, and SGP. The results are robust on excluding DNK, FIN, SWE, and TWN in addition to the five countries.

Table 5: Robustness Checks: Random-Coefficient Models

Dependent variable: de-meaned $v_{it}(s)$	(1)	(2)	(3)	(4)
Interaction between returns to scale and wage elasticity gap				
de-meaned $\alpha(s) \times [\varepsilon_{np}^{\text{wage}}(s) - \varepsilon_{p}^{\text{wage}}(s)] \times \ln Y_{i\text{US},t}$	4.95***	5.18***	4.95***	5.18***
	(1.38)	(1.38)	(1.38)	(1.38)
Markups				
de-meaned $\mu_t(s) \times \ln Y_{i\text{US},t}$	2.60***		2.60***	
	(0.25)		(0.25)	
Effective trade costs				
de-meaned $x_{i\text{US},t}(s) \times \ln Y_{i\text{US},t}$	65.91***	35.40***	65.91***	35.40***
	(1.98)	(3.59)	(1.98)	(3.59)
Trade elasticity				
de-meaned $\varepsilon_t^{\text{trade}}(s) \times \ln Y_{i\text{US},t}$		2.53***		2.53***
		(0.27)		(0.27)
de-meaned $x_{i\text{US},t}(s) \times \varepsilon_t^{\text{trade}}(s) \times \ln Y_{i\text{US},t}$		-2.26***		-2.26***
		(0.23)		(0.23)
Other controls	Yes	Yes	Yes	Yes
Country-time FE	Yes	Yes	Yes	Yes
Random slope of log relative GDP	$\text{me}_i + \text{me}(s)$	$\text{me}_i + \text{me}(s)$	$\text{me}_i(s)$	$\text{me}_i(s)$
Observations	80178	80178	80178	80178

Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Instead of adding country-industry fixed effects, we demean the dependent and independent variables within a country-industry level to reduce computational burdens. Regressions with other controls include interactions of log relative GDP with durability dummy, industry relative size, labor and material cost share.

we consider a subsample after 1995 rather than in 1989. In Table 6's Columns (1) and (2), the estimated coefficients are insignificantly different from the estimation results in Table 3.

Non-zero Trade Flow. We check the impact of zero trade flows by removing observations with zero trade flows. Eaton and Tamura (1994) highlight the importance of controlling for zero trade values in a gravity equation. Since zero trade flows could be an exogenous outcome of huge comparative advantage differences, the robustness check without zero trade flows is reasonable in our empirical framework to control for such differences.

Hanson and Xiang (2004) document that their home market effect results are robust to including or excluding zero trade flows. However, Pham et al. (2014) conclude that Hanson and Xiang (2004)'s results are sensitive. In Columns (3) and (4) of Table 6, the estimated impacts of returns to scale are consistent with the results in Table 3. However, without zero trade flows, the effective trade cost coefficients are negative at the 10% significance level. Also, the estimated trade elasticity coefficient is positive but insignificant at the 10% level. These are in contrast to the results in Table 3. These findings show that treatment of zero trade flows is essential in empirical works related to the home market effects.

Alternative Estimation of Returns to Scale and Markups. We also estimate the relationship between market structure characteristics and trade patterns using instrumented estimates of returns to scale and corresponding markup estimates. As discussed in Basu and Fernald (1997), instru-

Table 6: Additional Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: $v_{it}(s)$	After WTO (95)		Non-Zero Trade Flow		OLS Mkt Structure	
Interaction between returns to scale and wage elasticity gap						
$\alpha(s) \times [\varepsilon_{np}^{\text{wage}}(s) - \varepsilon_{p}^{\text{wage}}(s)] \times \ln Y_{i\text{US},t}$	6.48***	6.57***	4.34***	4.48***	6.29***	6.47***
	(1.63)	(1.65)	(1.43)	(1.40)	(1.49)	(1.42)
Markups						
$\mu_t(s) \times \ln Y_{i\text{US},t}$	1.80**		1.42**		2.86**	
	(0.83)		(0.61)		(1.06)	
Effective trade costs						
$x_{i\text{US},t}(s) \times \ln Y_{i\text{US},t}$	66.54**	28.87	-15.27*	-24.36*	66.29**	34.71
	(24.80)	(19.65)	(7.57)	(12.02)	(23.68)	(21.18)
Trade elasticity						
$\varepsilon_t^{\text{trade}}(s) \times \ln Y_{i\text{US},t}$		2.76***		0.65		2.55**
		(0.56)		(0.45)		(0.98)
Interaction between trade costs and elasticity						
$x_{i\text{US},t}(s) \times \varepsilon_t^{\text{trade}}(s) \times \ln Y_{i\text{US},t}$		-2.87***		-0.70		-2.36**
		(0.57)		(0.50)		(0.90)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Country-industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	59262	59262	75981	75981	80178	80178
R-squared	0.71	0.71	0.78	0.78	0.67	0.67

Notes: $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors are (three-way) clustered within country (21), industry (166), and year (17 in Columns (1) – (2) and 23 in Columns (3) – (6)) are in parentheses. This table repeats the evaluation Columns (4) and 5 in Table 3 using alternative specifications to check the robustness of our main results. Regressions with other controls include interactions of log relative GDP with durability dummy, industry relative size, labor and material cost share.

mented estimates can be more biased compared to uninstrumented estimates when the instruments are not completely exogenous and are only weakly correlated with regressors. According to Tables 3 and 6, the results are robust to the choice of instrumented and uninstrumented estimates of market structure.

7. Quantitative Analysis

As a final step, we quantify the contribution of returns to scale in generating heterogeneous trade patterns in the US manufacturing sector. To this end, we combine Section 2's theoretical framework and Section 4's measurements of traits of narrowly defined industries such as $\alpha(s)$, $\tilde{\vartheta}_l(s) - \vartheta_l(s)$, and $\varepsilon^{\text{trade}}(s)$. Our goal is to calculate the marginal contribution of the returns to scale channel, i.e., how much the returns to scale channel contributes to the additional cross-sectional variation in the export-import ratio. In particular, we ask how much would the trade variation increase if we allow for both returns to scale, i.e., $\alpha(s) \neq 1$, and different input intensities in non-production and production activities, i.e., $\tilde{\vartheta}_l(s) \neq \vartheta_l(s)$.

From Equation (20), the impact of the returns to scale channel on the industrial export-import

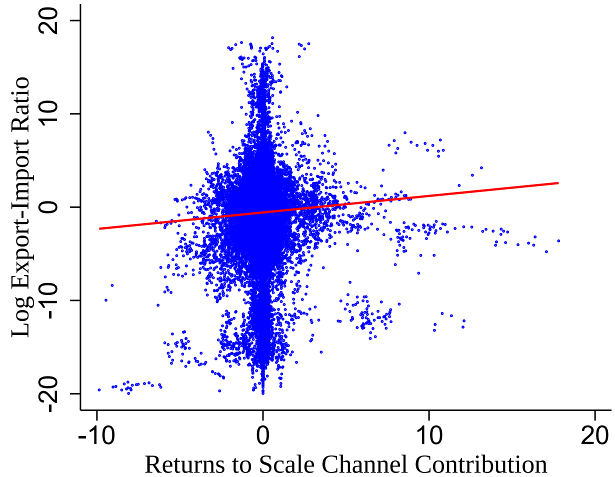

Figure 3: Export-Import Ratio and Returns to Scale Channel

Notes: In Panel A, x- and y-axis are the log export-import ratio and its returns to scale channel that are the left- and right-hand sides of Equation (35), respectively. (i.e., $\ln \mathrm{EX}_{ij}(s)/\mathrm{IM}_{ij}(s)$ and $[-\varepsilon^{\text{trade}}(s)][\alpha(s)-1][\tilde{\vartheta}_l(s)-\vartheta_l(s)] \times \ln \mathrm{TOL}_{ij}$.)

.

ratio is given by

$$\ln \frac{\mathrm{EX}_{ij}(s)}{\mathrm{IM}_{ij}(s)} = -\varepsilon^{\text{trade}}(s)[\alpha(s)-1][\tilde{\vartheta}_l(s)-\vartheta_l(s)] \ln \mathrm{TOL}_{ij} + \text{other terms}. \tag{35}$$

The above equation tells us that returns to scale leads to an additional source of variation in the export-import ratio. Therefore,

$$\mathrm{sd}\left(\ln \frac{\mathrm{EX}_{ij}(s)}{\mathrm{IM}_{ij}(s)} \right) \approx \mathrm{sd}\big([-\varepsilon^{\text{trade}}(s)][\alpha(s)-1][\tilde{\vartheta}_l(s)-\vartheta_l(s)] \ln \mathrm{TOL}_{ij}\big) + \text{other terms} \tag{36}$$

where the marginal contribution of returns to scale increases with $\alpha(s)$ and equates to zero if labor input intensities between non-production and production activities are identical. i.e., $\tilde{\vartheta}_l(s) = \vartheta_l(s)$.

Then, from the Penn World Table 9.1., we calculate country i's terms of labor relative to the US by

$$\mathrm{TOL}_{i\mathrm{US}} = \frac{\mathrm{LaborIncome}_i}{\mathrm{LaborIncome}_{\mathrm{US}}} \frac{\mathrm{AVH}_{\mathrm{US}} \times \mathrm{EMP}_{\mathrm{US}}}{\mathrm{AVH}_i \times \mathrm{EMP}_i}, \tag{37}$$

where $\mathrm{LaborIncome}_i$, AVH_i, and EMP_i are the real labor compensation, average annual working hours, and employment, respectively. Finally, we measure the returns to scale channel's contribution to the export-import ratio by $[-\varepsilon^{\text{trade}}(s)][\alpha(s)-1][\tilde{\vartheta}_l(s)-\vartheta_l(s)] \ln \mathrm{TOL}_{ij}$. Figure 3 plots the calculated values.

Figure 4 highlights the importance of the returns to scale channel in generating cross-sectional heterogeneity in industrial trade patterns. Panel A shows that the cross-sectional standard deviation

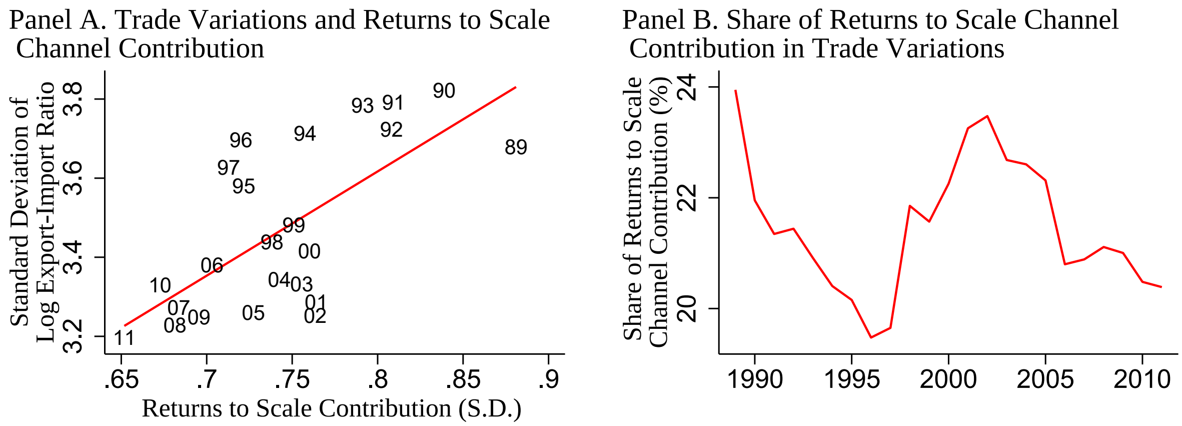

Figure 4: Industrial Trade's Cross-Sectional Variations and Contributions of Returns to Scale Channel

Notes: In Panel A, y- and x-axis are the cross-sectional standard deviations of $\ln \mathrm{EX}_{ij}(s)/\mathrm{IM}_{ij}(s)$ and $[-\varepsilon^{\text{trade}}(s)][\alpha(s)-1][\tilde{\vartheta}_l(s)-\vartheta_l(s)]\ln \mathrm{TOL}_{ij}$. Panel B plots their ratio (%): $\mathrm{sd}\big(\ln \mathrm{EX}_{ij}(s)/\mathrm{IM}_{ij}(s)\big)$ divided by $\mathrm{sd}\big([-\varepsilon^{\text{trade}}(s)][\alpha(s)-1][\tilde{\vartheta}_l(s)-\vartheta_l(s)]\ln \mathrm{TOL}_{ij}\big)$.

.

of the export-import ratio co-moves with the cross-sectional standard deviation of the returns to scale channel, as seen in the right hand side of Equation (36). In Panel B, the returns to scale channel contributes to around 21% of the trade variation on average. During our sample period, the maximum and minimum contributions are 19% and 24%, respectively.

8. Concluding Remarks

This paper explores the impact of market structure characteristics on intra-industry trade patterns, both theoretically and empirically. Our work focuses on important determinants of specialization and industry location — returns to scale in addition to product differentiation and trade costs.

The theoretical framework delivers some novel insights about the impact of returns to scale on international trade patterns. The role of scale economies depends on (a) the source (i.e., non-constant marginal costs vs. fixed costs) and (b) the relative use of the labor input (non-tradable) in production and non-production activities. The results highlight the importance of considering the role of production cost structures in multi-industry models with scale economies.

From bilateral trade flow data, we find empirical evidence that the difference between non-production and production cost structures influence the impact of returns to scale on the home market effect. There is no impact of returns to scale when the labor input shares (intensities) in production and non-production activities are identical. These results are consistent with our theoretical model predictions. Also, larger countries tend to have a higher concentration of industries with higher markups (i.e., markups positively impact the home market effect), which is consistent with the prior literature.

Finally, we use the previous results to quantify the impact of returns to scale on international trade patterns. The quantitative framework indicates that returns to scale lead to a sizable amount of cross-sectional variation in the export-import ratio.

References

Almunia, Miguel, Pol Antràs, David Lopez-Rodriguez, and Eduardo Morales, "Venting Out: Exports During a Domestic Slump," Working Paper 25372, National Bureau of Economic Research December 2018.

Anderson, James E., Mykyta Vesselovsky, and Yoto V. Yotov, "Gravity with Scale Effects," *Journal of International Economics*, 2016, *100*, 174 – 193.

Antweiler, Werner and Daniel Trefler, "Increasing Returns and All That: A View from Trade," *The American Economic Review*, 2002, *92* (1), pp. 93–119.

Baily, Martin Neil, Charles Hulten, David Campbell, Timothy Bresnahan, and Richard E. Caves, "Productivity Dynamics in Manufacturing Plants," *Brookings Papers on Economic Activity. Microeconomics*, 1992, *1992*, 187–267.

Bartelme, Dominick G, Arnaud Costinot, Dave Donaldson, and Andres Rodriguez-Clare, "The Textbook Case for Industrial Policy: Theory Meets Data," Working Paper 26193, National Bureau of Economic Research August 2019.

Bartelsman, Eric J. and Mark Doms, "Understanding Productivity: Lessons from Longitudinal Microdata," *Journal of Economic Literature*, 2000, *38* (3), pp. 569–594.

_ **and Wayne Gray**, "The NBER Manufacturing Productivity Database," Working Paper 205, National Bureau of Economic Research October 1996.

Basu, Susanto and John G. Fernald, "Returns to Scale in U.S. Production: Estimates and Implications," *Journal of Political Economy*, 1997, *105* (2), pp. 249–283.

Bergstrand, Jeffrey H., Stephen R. Cray, and Antoine Gervais, "Increasing Marginal Costs, Firm Heterogeneity, and the Gains from "Deep" International Trade Agreements," Working Paper, University of Notre Dame, July 2021.

Berman, Nicolas, Antoine Berthou, and Jerome Hericourt, "Export dynamics and sales at home," *Journal of International Economics*, 2015, *96* (2), 298 – 310.

Boehm, Christoph E, Andrei A Levchenko, and Nitya Pandalai-Nayar, "The Long and Short (Run) of Trade Elasticities," Working Paper 27064, National Bureau of Economic Research April 2020.

Costinot, Arnaud, Dave Donaldson, Margaret Kyle, and Heidi Williams, "The More We Die, The More We Sell? A Simple Test of the Home-Market Effect," *The Quarterly Journal of Economics*, 01 2019, *134* (2), 843–894.

Crozet, Matthieu and Federico Trionfetti, "Trade Costs and the Home Market Effect," *Journal of International Economics*, 2008, *76* (2), 309 – 321.

Davis, Donald R., "The Home Market, Trade, and Industrial Structure," *The American Economic Review*, 1998, *88* (5), 1264–1276.

_ and David E. Weinstein, "Economic Geography and Regional Production Structure: An Empirical Investigation," *European Economic Review*, 1999, *43* (2), 379 – 407.

_ and _ , "Market access, economic geography and comparative advantage: an empirical test," *Journal of International Economics*, 2003, *59* (1), 1 – 23.

Dingel, Jonathan I., "The Determinants of Quality Specialization," *The Review of Economic Studies*, 10 2016, *84* (4), 1551–1582.

Eaton, Jonathan and Akiko Tamura, "Bilateralism and Regionalism in Japanese and U.S. Trade and Direct Foreign Investment Patterns," *Journal of the Japanese and International Economies*, 1994, *8* (4), 478 – 510.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu, "Competition, Markups, and the Gains from International Trade," *American Economic Review*, October 2015, *105* (10), 3183–3221.

Fajgelbaum, Pablo, Gene M. Grossman, and Elhanan Helpman, "Income Distribution, Product Quality, and International Trade," *Journal of Political Economy*, 2011, *119* (4), 721–765.

Feenstra, Robert C., James A. Markusen, and Andrew K. Rose, "Understanding the Home Market Effect and the Gravity Equation: The Role of Differentiating Goods," Working Paper 6804, National Bureau of Economic Research November 1998.

_ , Robert Inklaar, and Marcel P. Timmer, "The Next Generation of the Penn World Table," *American Economic Review*, October 2015, *105* (10), 3150–82.

Fontagne, Lionel, Houssein Guimbard, and Gianluca Orefice, "Product-Level Trade Elasticities: Worth Weighting For," CESifo Working Paper Series 8491, CESifo 2020.

Giri, Rahul, Kei-Mu Yi, and Hakan Yilmazkuday, "Gains from Trade: Does Sectoral Heterogeneity Matter?," Working Paper 26741, National Bureau of Economic Research February 2020.

Grossman, Gene and Elhanan Helpman, *Innovation and Growth in the Global Economy*, 1 ed., Vol. 1, The MIT Press, 1991.

Grossman, Gene M. and Esteban Rossi-Hansberg, "External Economies and International Trade Redux," *The Quarterly Journal of Economics*, 2010, *125* (2), 829–858.

Hall, Robert E., “The Relation between Price and Marginal Cost in U.S. Industry,” *Journal of Political Economy*, 1988, *96* (5), pp. 921–947.

Hamilton, James D., “What is an oil shock?,” *Journal of Econometrics*, 2003, *113* (2), 363 – 398.

Hanson, Gordon H. and Chong Xiang, “The Home-Market Effect and Bilateral Trade Patterns,” *The American Economic Review*, 2004, *94* (4), 1108–1129.

Head, Keith and John Ries, “Increasing Returns versus National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade,” *American Economic Review*, September 2001, *91* (4), 858–876.

Helpman, Elhanan E. and Paul R. Krugman, *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy*, 1 ed., Vol. 1 of *MIT Press Books*, Cambridge MA: The MIT Press, April 1985.

Holmes, Thomas J. and John J. Stevens, “Does home market size matter for the pattern of trade?,” *Journal of International Economics*, 2005, *65* (2), 489 – 505.

Imbs, Jean and Isabelle Mejean, “Trade Elasticities,” *Review of International Economics*, 2017, *25* (2), 383–402.

Karabarbounis, Loukas and Brent Neiman, “The Global Decline of the Labor Share,” *The Quarterly Journal of Economics*, 10 2013, *129* (1), 61–103.

Keller, Wolfgang and Stephen R. Yeaple, “Multinationals, Markets, and Markups,” Working Paper, Pennsylvania State University June 2020.

Kim, Daisoon, “Essays on International Trade and Business Cycles,” Ph.D. Thesis, University of Washington 2019.

_ , “Economies of scale and international business cycles,” *Journal of International Economics*, 2021, *131*, 103459.

Krugman, Paul R., “Increasing returns, monopolistic competition, and international trade,” *Journal of International Economics*, 1979, *9* (4), 469 – 479.

_ , “Scale Economies, Product Differentiation, and the Pattern of Trade,” *The American Economic Review*, 1980, *70* (5), pp. 950–959.

Kucheryavyy, Konstantin, Gary Lyn, and Andres Rodriguez-Clare, “Grounded by Gravity: A Well-Behaved Trade Model with Industry-Level Economies of Scale,” Working Paper 22484, National Bureau of Economic Research August 2016.

Lashkaripour, Ahmad, “Within-Industry Specialization and Global Market Power,” *American Economic Journal: Microeconomics*, February 2020, *12* (1), 75–124.

Laussel, Didier and Thierry Paul, “Trade and the location of industries: Some new results,” *Journal of International Economics*, 2007, *71* (1), 148 – 166.

Lee, Yoonsoo, "The Importance of Reallocations in Cyclical Productivity and Returns to Scale: Evidence from Plant-Level Data," Working Papers, U.S. Census Bureau, Center for Economic Studies 2007.

Loecker, Jan De, Jan Eeckhout, and Gabriel Unger, "The Rise of Market Power and the Macroeconomic Implications," *The Quarterly Journal of Economics*, 01 2020, *135* (2), 561–644.

_ , Pinelopi K. Goldberg, Amit K. Khandelwal, and Nina Pavcnik, "Prices, Markups, and Trade Reform," *Econometrica*, 2016, *84* (2), 445–510.

Lyn, Gary and Andres Rodriguez-Clare, "External Economies and International Trade Redux: Comment," *The Quarterly Journal of Economics*, 2013, *128* (4), 1895–1905.

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

Panagariya, Arvind, "Variable Returns to Scale in Production and Patterns of Specialization," *The American Economic Review*, 1981, *71* (1), 221–230.

Pham, Cong S., Mary E. Lovely, and Devashish Mitra, "The home-market effect and bilateral trade patterns: A reexamination of the evidence," *International Review of Economics* & *Finance*, 2014, *30*, 120 – 137.

Pierce, Justin and Peter K. Schott, "A Concordance Between Ten-Digit U.S. Harmonized System Codes and SIC/NAICS Product Classes and Industries," NBER Working Papers 15548, National Bureau of Economic Research, Inc 2009.

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

Rodriguez-Lopez, Jose Antonio, "Prices and Exchange Rates: A Theory of Disconnect," *The Review of Economic Studies*, 2011.

Schott, Peter K., "The Relative Sophistication of Chinese Exports," *Economic Policy*, 2008, *23* (53), 5–49.

Soderbery, Anson, "Market size, structure, and access: Trade with capacity constraints," *European Economic Review*, 2014, *70*, 276 – 298.

Vannoorenberghe, G., "Firm-level volatility and exports," *Journal of International Economics*, 2012, *86* (1), 57 – 67.

Appendix

A. A Two-Country Model with CES preference

This section constructs a multi-industry, two-country new trade model with CES preference where the product-level and national level elasticities of substitution are identical: $\theta(s) = \sigma^{\text{firm}}(s) = \sigma^{\text{natl}}(s)$

The model features two countries — Home and Foreign. The two countries have identical industry characteristics, but we allow factor endowments to differ across them. For example, Home has greater factor endowments than Foreign, and therefore is larger than Foreign. A free entry condition endogenously determines the number of firms in each country. We denote Foreign variables with an asterisk. The model balances aggregate trade between the two countries, but an industry's net exports can be positive or negative.

A.1. Endowments

Table A1: Model Environments

Market	Environment
Aggregate market	Foreign GDP is normalized by one: $Y^* = 1$.
Labor market	Labor is immobile: no immigration.
	Home is larger than Foreign: $L > L^*$.
International capital market	Capital is internationally tradable without costs: $r = r^*$.
	K and K^* satisfy the zero net-capital flow.

There are two factors of production in each country: labor and capital. All factor markets are perfectly competitive. Without loss of generality, we assume that the Home country is larger than the Foreign country (i.e., $L > L^* > 0$, where L and L^* are the labor endowments at Home and Foreign, respectively). Since there is no migration in this model, the labor endowments are not tradable. In contrast, the two countries can trade capital costlessly and without any frictions. The initial capital endowments — denoted K and K^* in Home and Foreign, respectively — guarantee zero net capital flows in equilibrium. We denote the price of capital at Home and Foreign with r and $r*$, respectively, and costless trade implies that $r = r^*$. Zero profits imposed by the free entry condition of firms means that Home and Foreign income is given by $Y = wL + rK$ and $Y^* = w^*L^* + r^*K^*$, respectively, where w and w^* denote wages in Home and Foreign, respectively. Without loss of generality, we normalize $Y^* = 1$. Since all Home and Foreign features are identical with the exception of factor endowments, all cross-country differences are derived from endowment heterogeneity. Hence, the Home country is larger than the Foreign country in terms of income (GDP) (i.e., $Y > 1$).

A.2. Heterogeneous Industry Structure

There is a continuum of industries indexed by $s \in [0, 1]$, with monopolistically competitive firms in each industry. Consumer expenditure shares in industry s are constant and denoted by $\phi(s) \in$

$(0, 1)$, such that $\int_0^1 \phi(s)\,\mathrm{d}s = 1$. Therefore, the total consumer expenditure for goods produced in industry s is given by $\phi(s)Y$ at Home and $\phi(s)Y^*$ at Foreign.

The industry has a continuum of firms that produce differentiated goods with constant elasticity of substitution $\theta(s) > 1$ at the both product and national levels. The mass of firms in each industry is given by $n(s)$ and will be endogenously determined by the free entry condition. To focus on industry-level analysis, we assume that all firms are identical in industry s. Therefore, we do not include a firm index. To sell goods in an export market, firms in each industry face an iceberg trade cost denoted by $\tau(s)$. Therefore, to export one unit of a good, a firm in industry s must ship $\tau(s) > 1$ units of good. Then, the effective trade cost, denoted by $x(s) > 1$, satisfies $x(s) = [\tau(s)]^{\theta(s)-1}$. As in Hanson and Xiang (2004), we assume that $x(s)$ is given rather than $\tau(s)$.

Each firm's total and marginal cost functions, denoted by $\mathrm{tc}(s)$ and $\mathrm{mc}(s)$, in industry s are given by

$$\mathrm{tc}(s) = \left[w^{\vartheta_l(s)} r^{\vartheta_k(s)}\right]\left[q(s)\right]^{1/\alpha(s)} + \left[w^{\tilde{\vartheta}_l(s)} r^{\tilde{\vartheta}_k(s)}\right] f_C(s) \tag{A1}$$

$$\text{and} \quad \mathrm{mc}(s) = \frac{1}{\alpha(s)}\left[w^{\vartheta_l(s)} r^{\vartheta_k(s)}\right]\left[q(s)\right]^{1/\alpha(s)-1}, \tag{A2}$$

where $q(s)$ is the quantity produced and $f_C(s) > 0$ is the non-production cost (or fixed cost) in efficiency units of inputs. The returns to scale coefficient is denoted by $\alpha(s)$, which represents the output elasticity with respect to production inputs. Let $\mu(s)$ denote the markup of firm prices over marginal costs. We assume that the returns to scale coefficient is lower than the markup for a unique finite solution (i.e., $\alpha(s) < \mu(s)$).

The variable cost function is given by $\mathrm{vc}(s) = [w^{\vartheta_l(s)} r^{\vartheta_k(s)}]\,[q(s)]^{1/\alpha(s)}$, where $\vartheta_l(s)$ and $\vartheta_k(s)$ denote the labor and capital input cost shares in production, respectively. By duality, we get $q(s) = \{[l_p(s)]^{\vartheta_l(s)}[k_p(s)]^{\vartheta_k(s)}\}^{\alpha(s)}$. The fixed cost function is given by $\mathrm{fc}(s) = [w^{\tilde{\vartheta}_l(s)} r^{\tilde{\vartheta}_k(s)}] f_C(s)$, where $\tilde{\vartheta}_l(s)$ and $\tilde{\vartheta}_k(s)$ denote the labor and capital input cost shares in non-production activities (operations), respectively. Cost minimization implies that $f_C(s) = [l_{np}(s)]^{\tilde{\vartheta}_l(s)}[k_{np}(s)]^{\tilde{\vartheta}_k(s)}$, where the subscript np indicates inputs used in the non-production activities of each industry.

Importantly, we allow for differences in the input cost shares in production and non-production activities.[26] For instance, $\tilde{\vartheta}_l(s)$ and $\vartheta_l(s)$ do not necessarily have to be equal in each industry. If a firm's non-production activities only requires overhead labor, then $\tilde{\vartheta}_l(s) = 1$. In contrast, $\tilde{\vartheta}_x(s) = \vartheta_x(s)$ for $x = k$ and l if a firm's fixed costs are in unit of output.

A.3. Firm's Optimal Decision and Equilibrium

Firms' profit maximization yields the well-known price-setting condition that the price is a markup over marginal cost.

$$p(s) = \left[\frac{\mu(s)}{\alpha(s)}\right]\left[w^{\vartheta_l(s)} r^{\vartheta_k(s)}\right]\left[q(s)\right]^{1/\alpha(s)-1}, \tag{A3}$$

where $p(s)$ denotes the domestic real price of each variety produced in industry s. The markup equates to $\mu(s) = \theta(s)/\left[\theta(s) - 1\right]$, where $\theta(s)$ is the elasticity of substitution. Under the assumptions in Table A2, an individual firm's profit maximization problem has a unique interior solution

[26] Hanson and Xiang (2004) is the special case of ours with $\alpha(s) = 1$ and $\vartheta_l(s) = \tilde{\vartheta}_l(s) = 1$.

Table A2: Parameter Assumptions

Parameter	Restriction	Implication
Elasticity of substitution across products	$\theta(s) > 1$	(i) Markup: $\mu(s) > 1$ (ii) Effective trade cost: $x(s) > 1$
Returns to scale coefficient	$\alpha(s) < \mu(s)$	No corner solution
Labor cost share	$\vartheta_l(s) > 1 - 1/\alpha(s)$	$\alpha(s)\vartheta_l(s) > [\alpha(s) - 1]\tilde{\vartheta}_l(s)$
Others	$f_C(s), \alpha(s) > 0$ and $\tau(s) > 1$ $\vartheta_k(s), \vartheta_l(s), \tilde{\vartheta}_k(s), \tilde{\vartheta}_l(s) \in [0, 1]$ $\vartheta_k(s) + \vartheta_l(s) = 1 = \tilde{\vartheta}_k(s) + \tilde{\vartheta}_l(s)$ $\phi(s) \in (0, 1)$ and $\int_0^1 \phi(s)\,\mathrm{d}s = 1$	Trivial or by definition

for given aggregate variables.

The export price is given by $p_x(s) = [x(s)]^{\frac{1}{\theta(s)-1}}\, p(s)$, as the marginal cost in the export market is $[x(s)]^{\frac{1}{\theta(s)-1}}\, \mathrm{mc}(s)$. In each industry s, firms enter until the profit of each firm becomes zero. The free entry condition is:

$$\left[1 - \frac{\alpha(s)}{\mu(s)}\right] p(s)q(s) = \left[w^{\tilde{\vartheta}_l(s)} r^{\tilde{\vartheta}_k(s)}\right] f_C(s), \tag{A4}$$

where the left-hand side is the firm's revenue minus its variable costs and the right-hand side includes the fixed costs.

We define the industry's terms of trade by $\mathrm{TOT}(s) = p_x(s)/p_x^*(s)$. Given symmetric industries across countries, the terms of trade equate to the relative marginal cost of production in Home and Foreign: $\mathrm{TOT}(s) = \mathrm{mc}(s)/\mathrm{mc}^*(s)$. Equations (A3) and (A4) imply that industry market structure characteristics and the terms of labor determine the terms of trade:

$$\mathrm{TOT}(s) = \mathrm{TOL}^{\alpha(s)\vartheta_l(s) + [1-\alpha(s)]\tilde{\vartheta}_l(s)}. \tag{A5}$$

where the terms of labor determine the relative marginal costs across countries (the terms of trade) because costless capital mobility across countries implies equalization in the capital rental rates ($r = r^*$),. Notably, the impact of terms of labor on the relative marginal costs is different across industries and depends on the cost structures (i.e., the convex combination between $\vartheta_l(s)$ and $\tilde{\vartheta}_l(s)$ with $\alpha(s)$). The price competitiveness between Home and Foreign goods can be represented by $d(s) = n(s) \,/\, \{n(s) + n^*(s)[p(s)/p^*(s)]^{\theta(s)-1}/x(s)\}$ and $d^*(s) = n(s) \,/\, \{n(s) + n^*(s)[p(s)/p^*(s)]^{\theta(s)-1}x(s)\}$, respectively.

The goods market clearing condition in the Home country is given by

$$n(s)p(s)q(s) = \phi(s)Yd(s) + \phi(s)d^*(s), \tag{A6}$$

where $\phi(s)Yd(s)$ and $\phi(s)d^*(s)$ are Home and Foreign demand for Home goods in industry s, respectively. Similarly, the Foreign goods market clearing condition is given by $n^*(s)p^*(s)q^*(s) = \phi(s)Y[1 - d(s)] + \phi(s)[1 - d^*(s)]$. Balanced aggregate trade implies that aggregate accounting at Home is given by

$$Y = \int_0^1 n(s)p(s)q(s)\,\mathrm{d}s. \tag{A7}$$

Lemma A1 *Suppose that the assumptions in Table A2 hold. Then, the model environment described in Table A1 implies that*

$$\exists!\mathrm{TOL} \quad \textit{such that } 1 < \frac{w}{w^*} < \overline{\mathrm{TOL}},$$

where $\overline{\mathrm{TOL}} = \min[x(s)]^{\frac{1}{\{\alpha(s)\vartheta_l(s)+[1-\alpha(s)]\tilde{\vartheta}_l(s)\}[\theta(s)-1]+\tilde{\vartheta}_l(s)}}$.

Proof. See the appendix B. ■

The above lemma shows that there exists a unique solution of the terms of labor such that the Home country terms of labor are appreciated in equilibrium. Therefore, the larger country faces higher labor costs than the relatively smaller country. The intuition for this result is related to the number of firms in Home vs. Foreign. Since the larger economy (i.e., Home) has the larger market size, firms that locate at Home save on trade costs. Thus, Home is a more attractive location for firms than Foreign, and this expansion leads to higher input demand at Home. Since labor inputs are not tradable, the increase in labor demand puts pressure on Home wages to appreciate relative to Foreign wages. Therefore, labor market clearing implies that Home wages are higher than Foreign wages, despite a larger endowment of labor at Home.

A.4. Market Structure and Trade Patterns

In industry s, Home exports and imports are $\phi(s)d^*(s)$ and $\phi(s)Y[1-d(s)]$, respectively. Therefore, the ratio of exports to imports in the Home country can be expressed as:

$$\frac{\mathrm{ex}(s)}{\mathrm{im}(s)} = h(s)\frac{1}{Y}, \quad \text{where } h(s) = \frac{1 + x(s)[n(s)/n^*(s)][p(s)/p^*(s)]^{1-\theta(s)}}{1 + x(s)[n^*(s)/n(s)][p(s)/p^*(s)]^{\theta(s)-1}}. \tag{A8}$$

$h(s)$ determines the trade surplus patterns across industries, which can be represented by a function of the terms of labor and the relative GDP.[27] Since industry characteristics (which differ across industries) interact with terms of labor, appreciated terms of labor at Home lead to different impacts on trade surplus and location across industries.

Proposition 1 *In the unique equilibrium of Lemma A1, the Home country's ratio of exports to imports is increasing in the markup* $\mu(s)$. *Also, it is increasing, constant, or decreasing in the returns to scale coefficient* $\alpha(s)$ *if and only if* $\vartheta_l(s) < \tilde{\vartheta}_l(s)$, $\vartheta_l = \tilde{\vartheta}_l(s)$, *or* $\vartheta_l(s) > \tilde{\vartheta}_l(s)$, *respectively.*

Proof. See the appendix B. ■

In our framework, we say that industry s exhibits a home market effect if the ratio of exports to imports increases as the relative country size increases. The impact of markups on the home market

[27] For details, see the derivations and the proof of Proposition 1 in Appendix B.

effect is unambiguously positive. Since markups are negatively related to the elasticity of substitution across products, industries with lower elasticity of substitution (more differentiated products) would be concentrated in relatively larger countries. This result leads to a higher export/import ratio for industries with a lower elasticity of substitution.

Scale economies can potentially impact the home market effect through two channels — fixed costs and non-constant marginal costs. In our framework with symmetric fixed costs across countries in each industry, scale economies from fixed costs have no impact on the direction of industry-level home market effect. However, the slope of the marginal cost curve represented by the returns to scale coefficient impacts the home market effect across industries. The degree of returns to scale positively impacts the home market effect when non-production activities are more labor (or, more generally, non-tradable input) intensive than production activities. If the relative input intensities between firms' operation (or entry) and production are identical, an industry's returns to scale coefficient does not matter for its home market effect.

B. Derivations and Proofs

All derivations and proofs are similar to Hanson and Xiang (2004). See Appendix in Hanson and Xiang (2004) for the more details.

For convenience, define $\psi(s) = \mathrm{TOL}^{-\{\alpha(s)\vartheta_l(s)+[1-\alpha(s)]\tilde{\vartheta}_l(s)\}[\theta(s)-1]-\tilde{\vartheta}_l(s)}$. Then, the function is decreasing in TOL for positive TOL. For given $\mathrm{TOL} > 1$, the function is decreasing in $\theta(s)$. It is increasing, constant, or decreasing in $\alpha(s)$ if $\vartheta_l(s) < \tilde{\vartheta}_l(s)$, $\vartheta_l(s) = \tilde{\vartheta}_l(s)$, or $\vartheta_l(s) > \tilde{\vartheta}_l(s)$, respectively.

Derivation of the Number of Firms. First, rewrite $n(s)$ as a function of Y and w/w^*. Equations (A4) and (A6) imply

$$\begin{aligned}\frac{f_C(s)\, r^{\tilde{\vartheta}_r(s)}\tilde{n}(s)}{1-\alpha(s)/\mu(s)} &= n(s)p(s)q(s) = \phi(s)Yd + \phi(s)d^* \\ &= \frac{\phi(s)\, Y n(s)}{n(s)+n^*(s)\,[\mathrm{tot}(s)]^{\theta(s)-1}/x(s)} + \frac{\phi(s)\, n(s)}{n(s)+n^*(s)\,[\mathrm{tot}(s)]^{\theta(s)-1}x(s)},\end{aligned} \tag{B9}$$

where $\tilde{n}(s) = w^{\tilde{\vartheta}_l(s)} n(s)$ and $\tilde{n}^*(s) = (w^*)^{\tilde{\vartheta}_l(s)} n^*(s)$. Inserting Equation (A5) into the above, we obtain

$$\frac{f_C(s)\, r^{\tilde{\vartheta}_r(s)}\tilde{n}(s)}{1-\alpha(s)/\mu(s)} = \frac{\phi(s)\, Y x(s)\, \tilde{n}(s)}{x(s)\,\tilde{n}(s)+\tilde{n}^*(s)/\psi(s)} + \frac{\phi(s)\,\tilde{n}(s)}{\tilde{n}(s)+x(s)\,\tilde{n}^*(s)/\psi(s)} \tag{B10}$$

Also, the world goods market clearing condition is

$$\frac{f_C(s)\, r^{\tilde{\vartheta}_r(s)}\left[\tilde{n}(s)+\tilde{n}^*(s)\right]}{1-\alpha(s)/\mu(s)} = \phi(s)\,(Y+1)\,, \tag{B11}$$

where we use $r = r^*$. Thus, the number of firms is

$$n(s) = \frac{\tilde{n}(s)}{w^{\tilde{\vartheta}_l(s)}} = \frac{Y[x(s)]^2 - (Y+1)x(s)/\psi(s) + 1}{[x(s)]^2 - x(s)[1/\psi(s) + \psi(s)] + 1}\left[\frac{1-\alpha(s)/\mu(s)}{f_C(s)\, r^{\tilde{\vartheta}_r(s)} w^{\tilde{\vartheta}_l(s)}}\right]. \tag{B12}$$

Lemma B2 *The home aggregate accounting, Equation (A7), can be expressed by*

$$G(\mathrm{TOL}) = \int_0^1 \phi(s)\, g(s)\, \mathrm{d}s = 0, \tag{B13}$$

where $g(s) = Y\{x(s)[\mathrm{tot}(s)]^{\theta(s)-1}\mathrm{TOL}^{\tilde{\vartheta}_l(s)} - 1\}^{-1} - \{x(s)[\mathrm{tot}(s)]^{1-\theta(s)}\mathrm{TOL}^{-\tilde{\vartheta}_l(s)} - 1\}^{-1}$.

Proof. Insert Equation (B12) into Equation (A7).

$$Y = \int_0^1 \frac{Y[x(s)]^2 - (Y+1)x(s)/\psi(s) + 1}{[x(s)]^2 - x(s)[1/\psi(s) + \psi(s)] + 1}\phi(s)\, \mathrm{d}s \tag{B14}$$

Since $Y = \int_0^1 \phi(s)\, Y\, \mathrm{d}s$, we obtain the result.

$$0 = \int_0^1 \frac{Y[x(s) - 1/\psi(s)] - [1/\psi(s)][x(s)/\psi(s) - 1]}{[x(s) - 1/\psi(s)][x(s) - \psi(s)]}\phi(s)\, \mathrm{d}s \tag{B15}$$

$$= \int_0^1 \left\{\frac{Y}{x(s)/\psi(s) - 1} - \frac{1}{x(s)\psi(s) - 1}\right\}\phi(s)\, \mathrm{d}s = \int_0^1 \phi(s)\, g(s)\, \mathrm{d}s \tag{B16}$$

Using Equation (A5) and the definition of $\psi(s)$, we obtain the result. ■

The function $g(s)$ represents competitiveness of home industry relative to the foreign. The first part shows the competitiveness of home country that is increasing in the relative country size Y but decreasing in the entry costs (extensive margin) and the price (intensive margin) relative to foreign. Similarly, the second part represents the competitiveness of the foreign country. The above lemma implies that the sum of relative competitiveness of industries becomes zero. Even if the home country has advantages because of its large market size, the total trade balance is zero due to changes in terms of labor (relative price of non-tradable inputs).

Proof of Lemma A1. Consider $\underline{\mathrm{TOL}} \leq \mathrm{TOL} \leq \overline{\mathrm{TOL}}$ where $\underline{\mathrm{TOL}} = \min[x(s)]^{\frac{-1}{\{\alpha(s)\vartheta_l(s) + [1-\alpha(s)]\tilde{\vartheta}_l(s)\}[\theta(s)-1] + \tilde{\vartheta}_l(s)}}$ and $\overline{\mathrm{TOL}} = \min[x(s)]^{\frac{1}{\{\alpha(s)\vartheta_l(s) + [1-\alpha(s)]\tilde{\vartheta}_l(s)\}[\theta(s)-1] + \tilde{\vartheta}_l(s)}}$. Then, $\lim_{\mathrm{TOL} \searrow \underline{\mathrm{TOL}}} G(\mathrm{TOL}) = \infty$ and $\lim_{\mathrm{TOL} \nearrow \overline{\mathrm{TOL}}} G(\mathrm{TOL}) = -\infty$. Also, $G(\cdot)$ is a strictly decreasing function because $g(\cdot)$ is a strictly decreasing function for all s. Since $G(1) = (Y-1)\int_0^1 [x(s) - 1]^{-1}\, \mathrm{d}s > 0$, there exists an unique solution of the terms of labor between $1 < \mathrm{TOL} \geq \overline{\mathrm{TOL}}$.

Now, consider $\mathrm{TOL} > \overline{\mathrm{TOL}}$. If $\mathrm{TOL} \leq \max[x(s)]^{\frac{1}{\{\alpha(s)\vartheta_l(s) + [1-\alpha(s)]\tilde{\vartheta}_l(s)\}[\theta(s)-1] + \tilde{\vartheta}_l(s)}}$, there exist an industry s' such that $x(s')\psi(s') = 1$ that implies $g(s')$ has the infinite value. When $\mathrm{TOL} > \max[x(s)]^{\frac{1}{\{\alpha(s)\vartheta_l(s) + [1-\alpha(s)]\tilde{\vartheta}_l(s)\}[\theta(s)-1] + \tilde{\vartheta}_l(s)}}$, $G(\mathrm{TOL}) > 0$ due to $g(s) > 0$ for all s. Thus, there is no solution at $\mathrm{TOL} > \overline{\mathrm{TOL}}$. Similarly, it is easy to verify that there is no solution at $\mathrm{TOL} < \underline{\mathrm{TOL}}$.

Proof of Proposition 1. The ratio of export to import is $\mathrm{ex}(s)/\mathrm{im}(s) = (1/Y)\,[1 + x(s)\Psi(s)]\,/\,[1 + x(s)/\Psi(s)]$ where $\Psi(s) = [\tilde{n}(s)/\tilde{n}^*(s)]\,\psi(s)$. By using Equation (B12), $\Psi\,(s)$ can be rewritten as a function of $\psi\,(s)$.

$$\Psi\,(s) = \frac{\left\{[x\,(s)]^2\,Y + 1\right\}\psi\,(s) - (Y+1)\,x\,(s)}{[x\,(s)]^2 + Y - (Y+1)\,x\,(s)\,\psi\,(s)} \tag{B17}$$

Then, $\Psi\,(s)$ is increasing in $\psi\,(s)$ because $x\,(s) > 1$:

$$\frac{\partial\Psi\,(s)}{\partial\psi\,(s)} = \frac{\left\{[x\,(s) - 1]^2\right\}Y}{\left\{[x\,(s)]^2 + Y - (Y+1)\,x\,(s)\,\psi\,(s)\right\}^2} > 0. \tag{B18}$$

For given $\mathrm{TOL} > 1$, $\psi\,(s)$ is decreasing in $\theta(s)$, which leads that it is increasing in $\mu(s) = \theta(s)/[\theta(s) - 1]$ It is increasing, constant, or decreasing in $\alpha(s)$ if $\vartheta_l(s) < \tilde{\vartheta}_l(s)$, $\vartheta_l(s) = \tilde{\vartheta}_l(s)$, or $\vartheta_l(s) > \tilde{\vartheta}_l(s)$, respectively. Hence, the home terms of labor appreciation ($\mathrm{TOL} > 1$) implies that $\Psi\,(s)$ and $ex\,(s)/im\,(s)$ are increasing in $\mu(s)$. Also, they are increasing, constant, or decreasing in $\alpha(s)$ if $\vartheta_l(s) < \tilde{\vartheta}_l(s)$, $\vartheta_l(s) = \tilde{\vartheta}_l(s)$, or $\vartheta_l(s) > \tilde{\vartheta}_l(s)$, respectively.

C. Data and Measurement

We drop industries when they are newly created or deleted during the sample period.

C.1. Bilateral Trade and GDP

We use four, five, and six digit NAICS manufacturing sector bilateral trade flow data of U.S. from the U.S. Census Bureau from 1989 through 2011, which is constructed by Schott (2008) by using the concordances from Bartelsman and Doms (2000) and Pierce and Schott (2009). Data on country size (real GDP) and terms of labor (real GDP, labor share, working hours, and employment) is obtained from Penn World Table 9.1. (See Feenstra et al., 2015, for the details of database.)

- GDP: Real GDP at constant 2011 national prices (2011 US dollars)
- Export: Value of export (nominal: US dollars)
- Import: c.i.f. value of import (nominal: US dollars)
- Effective trade cost: c.i.f. value of import / f.o.b. value of import
- Labor share: Share of labor compensation in GDP at current national prices
- Labor income: GDP $\times$ Labor share
- Average working hours: Average annual hours worked by persons engaged
- Employment: Number of persons engaged (million)

We choose 21 economies, which are advanced economies among the 40 major trade partner economies of the U.S. The EU member countries' ISO codes are AUT, BEL, DEU, DNK, ESP, FIN, FRA, GBR, IRL, ITA, NLD, and SWE. The codes of the rest of them are AUS, CAN, CHE, HKG, ISR, JPN, KOR, SGP, and TWN. Also, we considered six developing and major trade countries: BRA, CHN, IND,MEX, MYS, and THA.

C.2. Industry Characteristics: NBER-CES Manufacturing Industry Database

We collect industry-level macro data using NBER-CES Manufacturing Industry Database from 1959 to 2011. See Bartelsman and Gray (1996) for the details.

- Revenue: Value of shipments deflated by the shipments deflator calculated from the BEA.
- Capital Input: The real capital stock. (millions of 1987 dollars)
- Capital Cost: The capital cost is not actually collected. We follow De Loecker et al. (2020) where they use the federal funds rate plus an exogenous depreciation rate and risk premium jointly (12%).
- Labor Input: The production workers' hours (production labor input) are reported in the data, but non-production workers' hours (non-production labor input) are not collected. We calculate total and non-production labor inputs as in Baily et al. (1992), where they assume that the wages of production and nonproduction workers in efficiency units are identical.
- Labor Cost: Total payroll (nominal) deflated by the shipments deflator calculated from the BEA.
- Material Input: The cost of materials deflated by the material cost deflator calculated using data from the benchmark use-make (input-output) tables and the GDP-by-Industry data from the BEA
- Material Cost: Cost of materials (nominal) deflated by the materials deflator calculated from the BEA.
- Cost shared input growth: average growth rate of labor, capital, and material inputs weighed by the previous year cost shares.
- Energy Spending: The cost of fuels and electricity deflated by the energy deflator calculated using the MECS and BLS database
- Relative size: The average of industry's value of shipments between t and $t-1$ divided by the average total value of shipments between t and $t-1$.

Durable and non-durable industries are as follows.

- Durables: 3 digit NAICS 321 and 327 – 339
- Nondurables: 3 digit NAICS 311 – 316 and 322 – 326

C.3. Instruments: Production Function Estimation

We use the following variables and their one-year lags.

- Oil price shocks: We collect monthly spot crude oil price: West Texas Intermediate (WTI) from FRED. As in Hamilton (2003), we construct the proxy of oil shocks by using the value of the oil price at time t relative to its largest value over the preceding 12 months: $\max\{0, \ln \text{Oil price}_t - \ln \text{Oil price}^{max}_{t-12,t-1}\}$ where Oil $\text{price}^{max}_{t-12,t-1}$ is the highest price of oil from $t-12$ and $t-1$. I use the real price of WTI (based on CPI). The annual oil price shocks are the sum of the monthly shocks.

- Growth rate of government defense spending (A489RA3A086NBEA from FRED): Real federal government consumption expenditures: Defense consumption expenditures: Gross output of general government: Intermediate goods and services purchased: Services (chain-type quantity index), Index 2009=100, annual

- Monetary policy shocks: The measure of monetary shocks is based on a monthly VAR model including the following log variables and 12 lags: the industrial production, the unemployment rate, the log of the CPI, and the log of a commodity price index, the federal funds rate, and M1. All data are from FRED. The error term from the fitted policy rule is the measure of the monetary shocks. The annual shocks are the sum of the monthly shocks. Exogenous time dummies, excluding the unemployment, and using T-bill interest rate instead of the federal fund rate have no impact on the results.

- President's party

D. Additional Figures

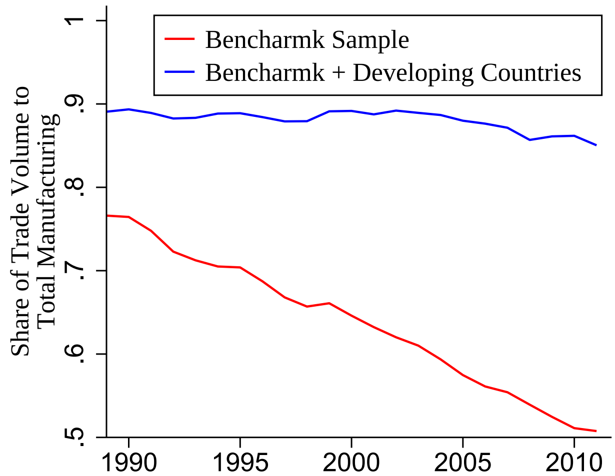

Figure D1: Sample Coverage

Notes: The trade volume is exports plus imports. The benchmark sample contains AUS, AUT, BEL, CAN, CHE, DEU, DNK, ESP, FIN, FRA, GBR, HKG, IRL, ISR, ITA, JPN, KOR, NLD, SGP, SWE, and TWN. The developing countries are BRA, CHN, IND, and MEX, MYS, and THA.

.

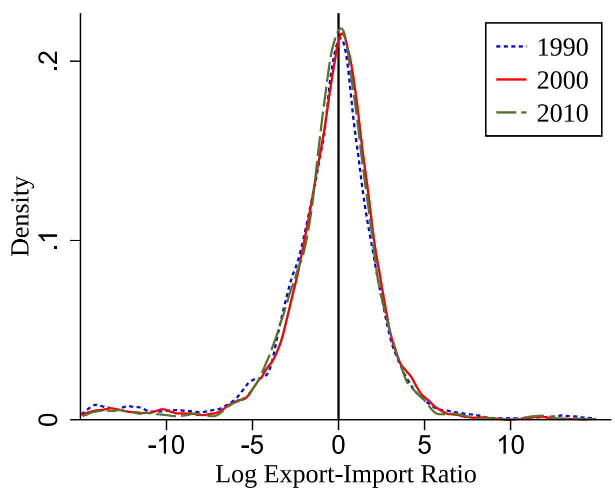

Figure D2: Kernel Densities of Logarithmic Export-Import Ratio to the US

Notes: The logarithmic export-import ratio to the US denoted $v_{i,t}(s)$ is defined in Equation (27).

.

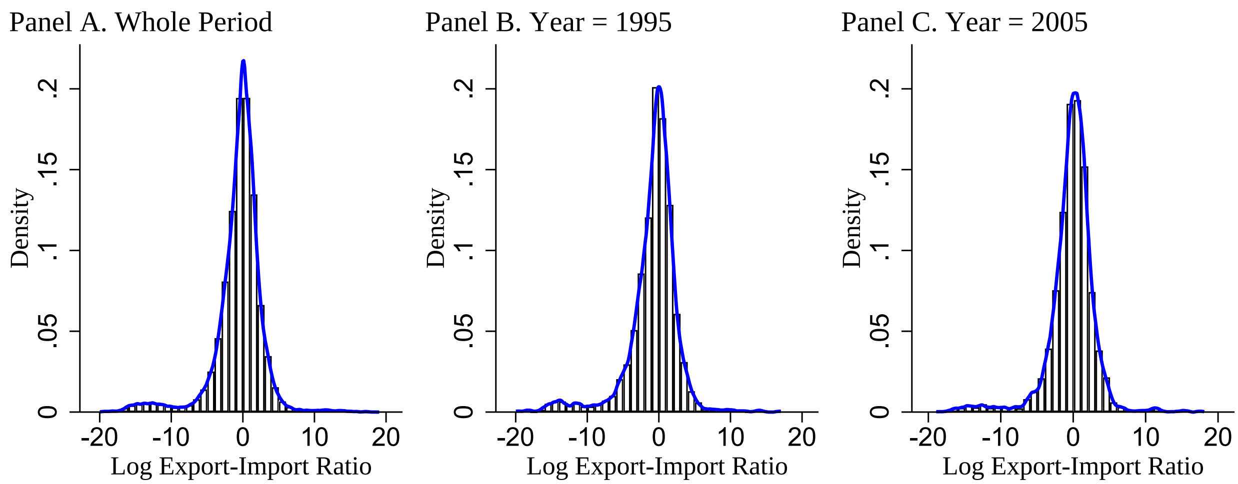

Figure D3: Histograms and Kernel Densities of Logarithmic Export-Import Ratio to the US

Notes: The logarithmic export-import ratio to the US denoted $v_{i,t}(s)$ is defined in Equation (27).

.

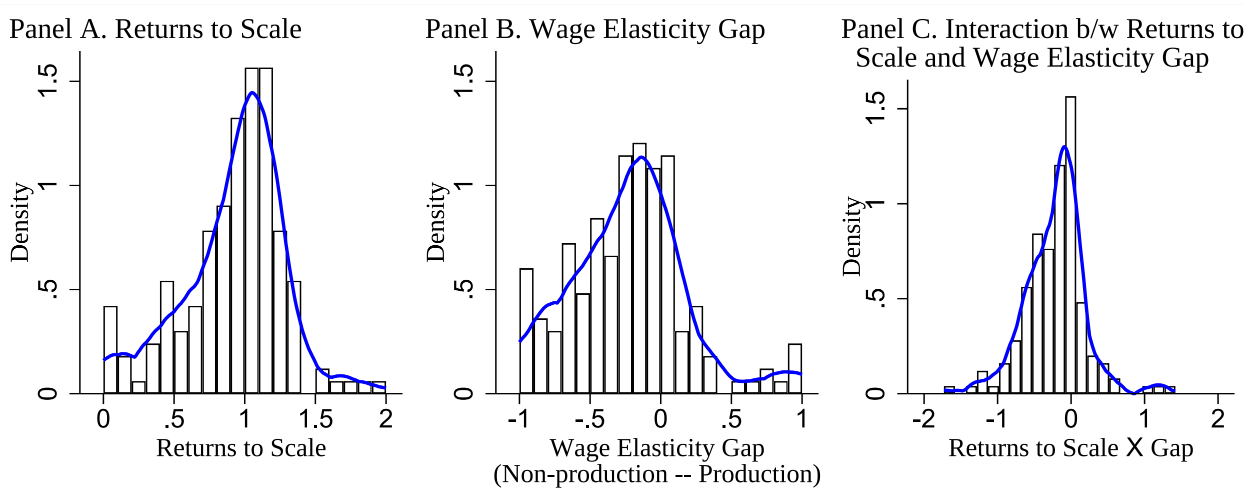

Figure D4: Histograms and Kernel Densities of Returns to Scale, Wage Elasticity Gap, and Their Interaction

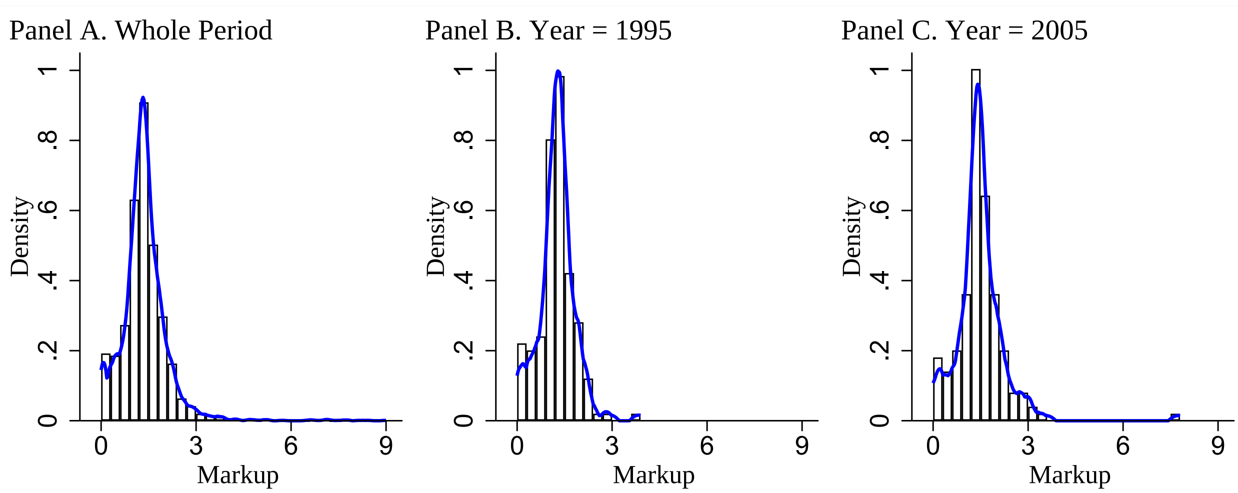

Figure D5: Histograms and Kernel Densities of Markups

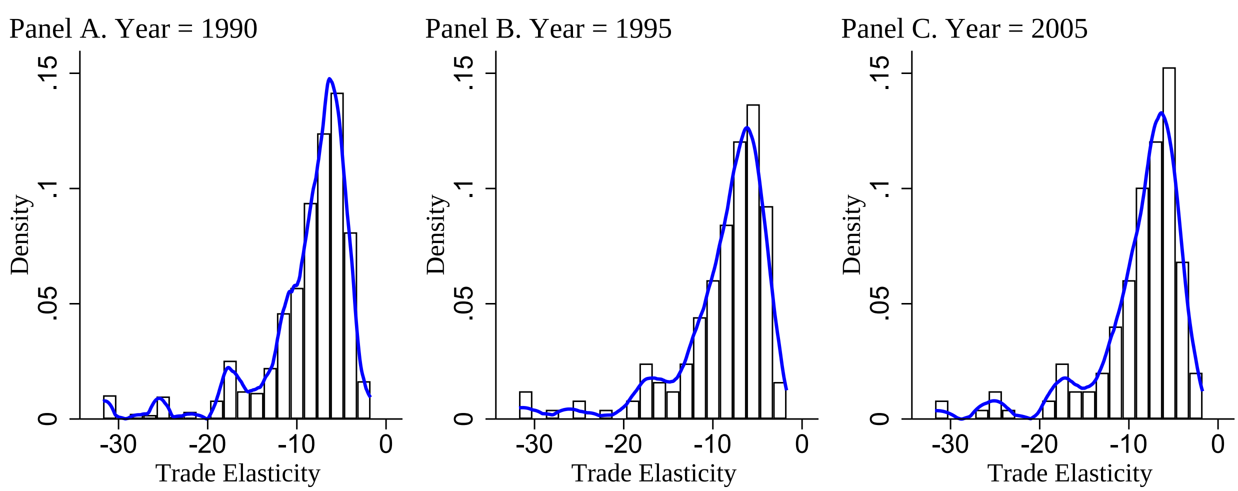

Figure D6: Histograms and Kernel Densities of Trade Elasticity

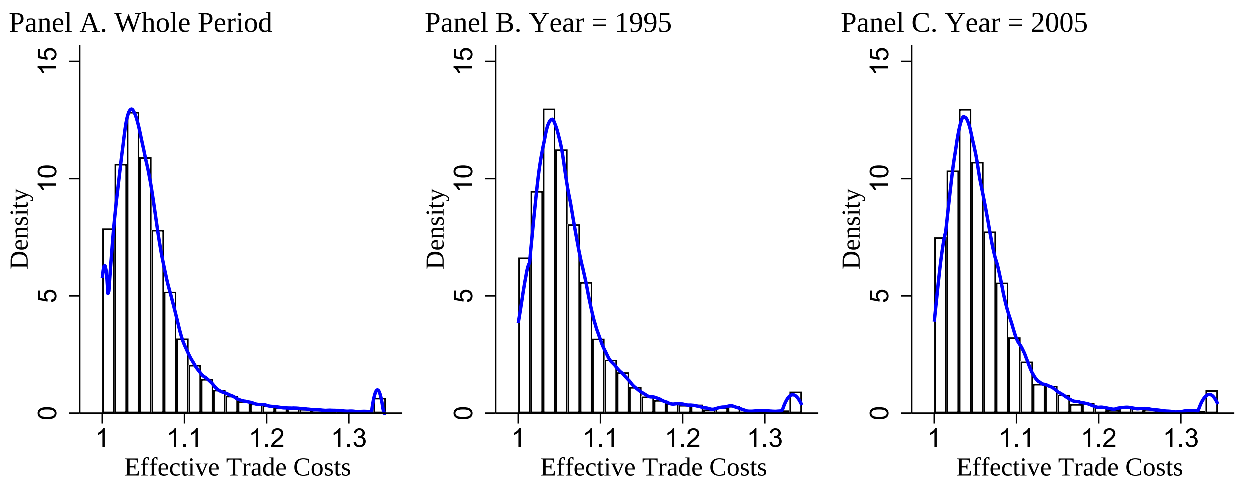

Figure D7: Histograms and Kernel Densities of Effective Trade Costs