Journal of International Economics 122 (2020) 103266

Contents lists available at ScienceDirect

Journal of International Economics

journal homepage: www.elsevier.com/locate/jie

Full Length Articles

Competition, markups, and gains from trade: A quantitative analysis of China between 1995 and 2004

Wen-Tai Hsu [a,*], Yi Lu [b], Guiying Laura Wu [c]

[a] School of Economics, Singapore Management University. 90 Stamford Road, 178903, Singapore
[b] School of Economics and Management, Tsinghua University, Beijing 100084, China
[c] Division of Economics, School of Humanities and Social Sciences, Nanyang Technological University, 14 Nanyang Drive, 637332, Singapore

ARTICLE INFO

Article history:
Received 2 August 2017
Received in revised form 27 September 2019
Accepted 12 October 2019
Available online 4 November 2019

Key words:
Gains from trade
Markups
Pro-competitive effects
ACR formula
Head-to-head competition
Chinese economy

ABSTRACT

This paper provides a quantitative analysis of gains from trade in a model with head-to-head competition using Chinese firm-level data from Economic Censuses in 1995 and 2004. We find a significant reduction in trade cost during this period, and total gains from such improved openness during this period is 7.1%. The gains are decomposed into a Ricardian component and two pro-competitive ones. The pro-competitive effects account for 20% of the total gains. Moreover, the total gains from trade are 13 − 31% larger than what would result from the formula provided by ACR (Arkolakis et al., 2012), which nests a class of important trade models, but without pro-competitive effects. We find that head-to-head competition is the key reason behind the larger gains, as trade flows do not reflect all of the effects via markups in an event of trade liberalization.

1. Introduction

It has been well understood that competition may affect gains from trade via changes in the distribution of markups. For example, when markups are the same across all goods, first-best allocative efficiency is attained because the condition that the price ratio equals the marginal cost ratio, for any pair of goods, holds. In other words, in an economy with variable markups, trade liberalization may improve *allocative efficiency* if the dispersion of markups is reduced.[1] Moreover, the *relative markup effect* also matters because welfare improves with trade liberalization when consumers benefit from lower markups of the goods they consume and when producers gain from higher markups (hence higher profits) in foreign markets. The effects of trade liberalization via changes in both the mean and dispersion of markups are generally termed *pro-competitive effects of trade*.

A natural question is then whether competition and markups are *quantitatively important* in gains from trade. To address this, this paper conducts quantitative analyses of the gains from trade using a model that features *head-to-head* competition to investigate the role of pro-competitive effects. We use Chinese firm-level data in Economic Censuses in 1995 and 2004 to quantify our model. China in between these two years is an important case, as this was a period when China drastically improved openness – not only was transport infrastructure rapidly expanded, but joining the World Trade Organization (WTO) in 2001 also drastically reduced trade barriers.[2] Recently, Brandt et al. (2017) and Lu and Yu (2015) have both estimated firm-level markups using Chinese manufacturing data and the approach by De Loecker and Warzynski (2012; henceforth DLW). Lu and Yu (2015) show that the larger the tariff reduction due to the WTO entry in one industry, the greater the reduction in the dispersion of markups in that industry. Brandt et al. present similar results on levels of markups. These empirical results suggest that pro-competitive effects might be present in the case of China, but a formal quantitative welfare analysis is warranted.

* Corresponding author.
E-mail addresses: wentaihsu@smu.edu.sg (W.-T. Hsu), luyi@sem.tsinghua.edu.cn (Y. Lu), guiying.wu@ntu.edu.sg (G.L. Wu).

[1] The idea of allocative efficiency dates back to Robinson (1934, Ch. 27) and Lipsey and Lancaster (1956-57). Note that allocative efficiency is determined by how the production resources are allocated across firms with different markups. Thus, both relative revenue/employment and the dispersion of markups matter. This point is made clear by Arkolakis et al. (2019) and also in the formulation of the current paper.

[2] Between 1995 and 2004, the import share increased from 0.13 to 0.22, whereas the export share increased from 0.15 to 0.25. The proportion of exporters among manufacturing firms also increased from 4.4% to 10.5%.

https://doi.org/10.1016/j.jinteco.2019.103266
0022-1996/

The two focuses of our welfare analyses are how head-to-head competition matters for the *total* gains from trade and the *decomposition* of these total gains into a standard Ricardian component and pro-competitive effects. To appreciate what we do, it is important to review recent related studies. First, Arkolakis et al. (2012; henceforth ACR) show that for a class of influential trade models, welfare gains from trade ($\hat{W} \equiv W'/W$) can be simply calculated by $(v'/v)^{1/\epsilon} = \hat{v}^{1/\epsilon}$, where v is domestic expenditure share, and $\epsilon < 0$ is the trade elasticity. As both v and ϵ depend on trade flows, trade flows provide key information regarding gains from trade. However, this class of models features no pro-competitive effects. To investigate pro-competitive effects, Edmond et al. (2015; henceforth EMX) use a model of distinct-product Cournot competition *a lá* Atkeson and Burstein (2008) and find that pro-competitive effects account for 11 − 38% of total gains from trade. Moreover, even though EMX's model deviates from the ACR class and sizable pro-competitive effects are found, it turns out their total gains from trade is well captured by the local version of the ACR formula. Similar results are also found by Feenstra and Weinstein (2017). Whereas ACR (p. 116) state, " While the introduction of these pro-competitive effects, which falls outside the scope of the present paper, would undoubtedly affect the composition of the gains from trade, our formal analysis is a careful reminder that it may not affect their total size", the present paper will revisit both the *total* and *composition* of gains from trade, and show how head-to-head competition matters.

Our quantitative framework is a variant of Bernard et al. (2003; henceforth BEJK). To understand our framework, we first note three features of BEJK. First, the productivity of firms is heterogeneous and follows a Fréchet distribution. Second, firms compete in Bertrand fashion good by good and market by market with active firms charging prices at the second lowest marginal costs. Third, although differences in markups are driven by productivity differences through limit pricing, it turns out that the resulting markup distribution is invariant to trade costs. Holmes et al. (2014) find that this invariance is due to the assumption that the productivity distribution is fat-tailed (Fréchet). If productivity draws are from a non-fat-tailed distribution, then the distribution of markups may change with the trade cost, and pro-competitive effects of trade may be observed.

Figure 1 shows the distribution of markups in China in 1995 and 2004. The distributions are highly skewed to the right, and it is clear that the distribution in 2004 is more condensed than that in 1995. Indeed, the (unweighted) mean markup decreases from 1.43 to 1.37 and almost all percentiles decrease from 1995 to 2004 (See Section 3 for more details). A two-sample Kolmogorov–Smirnov test clearly rejects the null hypothesis that the two samples (1995 and 2004) are drawn from the same distribution.[3] Under the BEJK structure, this suggests that one needs to deviate from fat-tailed distributions to account for such changes.[4]

We thus adopt the model of Holmes et al. (2014) with the productivity drawn from log-normal distributions. The log-normal distribution has been widely used in empirical applications; in particular, Head et al. (2014) argue that log-normal distribution offers a better approximation to firm sizes than Pareto. We describe the model in detail in Section 2. In Section 3, we structurally estimate the model using Simulated Method of Moments (SMM) in each data year, as if we are taking snapshots of the Chinese economy in the respective years. Thus, all parameters are allowed to change between these two years to reflect changes in the environment of the Chinese economy. In our main quantitative exercise, we vary only the trade cost. In particular, we gauge the effect of "factual improvement in openness" by examining the effect of changing trade costs from 1995 to 2004. As we focus on competition, our empirical implementation relies heavily on markups. We estimate firm-level markups following DLW and then use moments of markups to discipline model parameters, along with some macro moments.

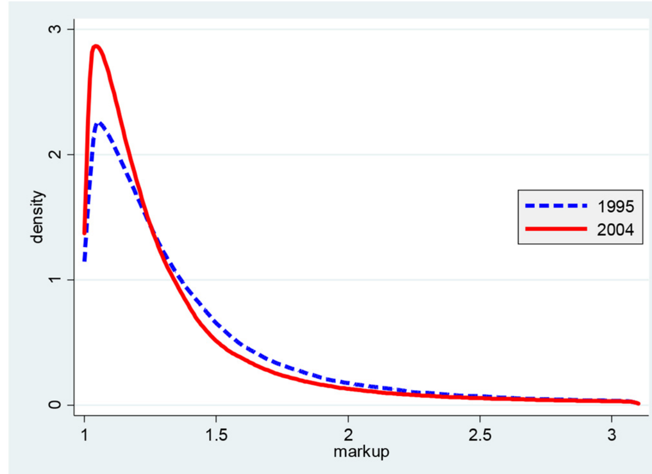

Fig. 1. Markup distributions (1995 versus 2004).

In Section 4, we conduct welfare analysis on gains from trade. Our benchmark counter-factual analysis is based on 2004 estimates and reverts the trade cost back to the level estimated using 1995 data to gauge the gains from the improved openness in this period. The gain is 7.1% of real income, and the contribution of the pro-competitive effects is 19.9%. The improvement of allocative efficiency accounts for the bulk of pro-competitive effects. The overall gains at 7.1% seems a relatively large number compared with those found in the literature. The sources of the larger gains compared with the literature can be understood as three-fold. First, there is a large reduction in trade cost (from an iceberg cost of 2.31 to 1.78) that is essentially inferred by the large increase in trade flows during 1995–2004. For a given trade elasticity ϵ, this implies large gains by the ACR formula, as $\hat{W} = \hat{v}^{1/\epsilon}$ and $\hat{v}$ here is small.

Second, Simonovska and Waugh (2014) and Melitz and Redding (2015) argue that new trade models with micro mechanisms such as firm heterogeneity, selection, variable markup, etc, imply lower estimates of absolute value of trade elasticity $|\epsilon|$. By accounting for markup dispersion in the data, our quantification also entails smaller $|\epsilon|$,[5] which hovers around 3. As the trade elasticity is a variable in our model, we calculate the gains by the ACR formula by integrating the local formula, and we obtain 5.9%. The gains by the ACR formula would be 3.3% with a standard estimate of $|\epsilon| = 5$.[6] Thus, the difference in trade elasticity is the second source of the larger gains.

The total gains from trade in our model are 20%($=7.1/5.9 - 1$) larger than the ACR formula, and we investigate the reasoning behind this third source of larger gains. We prove that the extra gains come precisely from the pro-competitive effects in the special case of Cobb-Douglas preference. Under general CES preference, pro-competitive effects may be smaller or larger than the extra gains, but they are still quite close. The intuition is that trade flows do not fully reflect changes in markups in this model with head-to-head competition among firms. For example, a domestic firm may charge a lower price in the face of fiercer foreign competition, but precisely because of the lower price, foreign competitors do not enter, and no trade flows are generated due to this change in markup (See Salvo (2010) and Schmitz (2005) for empirical examples). In contrast, in either monopolistic competition models (such as Arkolakis et al. (2019), Feenstra et al., 2017 and many others)[7]

[3] The combined K-S is 0.0829 and the p-value is 0.000.

[4] Similarly, Feenstra (2018) find that in monopolistic competition models, pro-competitive effects do not exist under Pareto productivity distribution, but they reappear when the distribution deviates from Pareto.

[5] For the intuition behind the low trade elasticities in our estimated models, see Section 4.2.

[6] See Costinot and Rodríguez-Clare (2014) for a discussion on the standard estimates.

[7] There is an extensive literature exploring properties of markups under monopolistic competition; see, for example, Dixit and Stiglitz (1977), Krugman (1979), Ottaviano et al. (2002), Melitz and Ottaviano (2008), Behrens and Murata (2012), Zhelobodko et al. (2012), Feenstra et al., 2017, Weinberger (2015) and Dhingra and Morrow (2016).

or distinct-product Bertrand or Cournot competition models (such as EMX), each firm owns a variety and hence a demand curve along which pricing is determined. A change in trade cost shifts firms' demand curves through general equilibrium effects or strategic interactions and thus affects markups and trade flows simultaneously. This is not the case here with head-to-head competition.

In Section 5, we extend the model to a multi-sector economy to account for heterogeneity across sectors. The welfare results in the multi-sector economy remain similar to the one-sector economy. Exploiting the variations in sectoral markups and trade costs, we also attempt to answer the question of whether China liberalized the " right" sectors in terms of reduction in trade cost or tariffs. The rationale is that the overall allocative efficiency would be better improved if the government were to target its trade liberalization more in the higher-markup sectors because this would reduce the dispersion of markups across sectors. We find that when a sectoral markup was higher in 1995, there was a tendency for a larger reduction in the estimated trade cost or import tariff between 1995 and 2004.

A desirable feature of our oligopolistic framework for quantitative analyses with micro-level data is that it is applicable to *countries of any size*. To illustrate this point, take the closely related work by EMX, which has a sensible feature that links markups with firms' market shares. Their model is quantified using Taiwanese firm-level data, which works well for their oligopoly environment because they can go down to a very fine product level to look at a few firms to examine their market shares. However, it could be difficult to apply their framework to a large economy (such as the US or China) where even in the finest level of industry or product, there could be hundreds of firms, to the effect that firms' market shares are typically much smaller than those for a small country. The problem here is that when firms' market shares are " diluted" by country size for a given industry or product category, so are pro-competitive effects. This is not to say that pro-competitive effects do not exist in large countries; rather, it may be that there are actually several markets in an industry or product category, but we simply do not know how to separate them. In contrast, markups in our model are driven by the difference between the active firms and their latent competitors, and thus they are not tied to any given product or industrial classification. Our approach is therefore applicable to data from countries of any size.

Besides the above-mentioned studies, earlier theoretical work on how trade may affect welfare through markups include Markusen (1981), Devereux and Lee (2001), and Epifani and Gancia (2011). In particular, Markusen (1981) shows that in an environment with *head-to-head Cournot competition* and symmetric countries, trade can reduce markup dispersion and thus enhance welfare without generating trade flows. Our work differs in that we provide quantitative analyses with a richer markup-generating mechanism and by linking to the ACR formula. Whereas our model follows that in Holmes et al. (2014), our work differs in at least three aspects: (1) we quantify pro-competitive effects with Chinese data; (2) we provide theoretical and quantitative analyses on the link to the ACR formula and show that head-to-head competition adds extra gains; (3) we use multi-sector analysis to show how cross-sector markup dispersion matters.

In a monopolistic-competition framework, Arkolakis et al. (2019) study a class of models that allow general preferences and variable markups, and find that the total welfare gains in these models are slightly lower than those with constant markups. In this sense, they conclude that the pro-competitive effects of trade are elusive. Nevertheless, this approach of comparing across models is a different exercise from our welfare decomposition within the same model and from our comparison with the ACR formula. Hence, their exercises are not directly comparable with ours.

Our work is closely related to recent studies regarding how gains from trade are related to the ACR formula. By using both data on trade flows and micro-level prices, Simonovska and Waugh (2014) show that welfare gains from trade in new models with micro-level margins exceed those in frameworks without these margins. Interestingly, even though our trade elasticity is a variable, our local trade elasticities at the estimated models are quite close to their estimates of trade elasticity under the BEJK model. Our work differs by incorporating pro-competitive effects and showing that the total gains from trade can deviate from the ACR formula in the BEJK framework once the productivity draws deviate from Fréchet. Melitz and Redding (2015) show that the trade elasticity becomes variable when the distribution of productivity deviates from untruncated Pareto in the Melitz (2003) framework, and hence the global ACR formula does not apply. Obviously, their mechanism is different from ours.[8]

Our work is also related to de Blas and Russ (2015) and De Loecker et al. (2016), who provide analyses of how trade affects the distribution of markup. But these papers do not address welfare gains from trade. By looking at allocative efficiency, our paper is also broadly related to the literature of resource misallocation, including Restuccia and Rogerson (2008) and Hsieh and Klenow (2009). Recently, Asturias et al. (2017) has studied the welfare effect of transportation infrastructure in India and examined the role of allocative efficiency in a similar fashion to Holmes et al. (2014) and the current paper.

2. Model

2.1. Consumption and Production

There are two countries, which are indexed by $i = 1, 2$.[9] In our empirical application, 1 means China, and 2 means the ROW. As is standard in the literature of trade, we assume a single factor of production, labor, that is inelastically supplied, and the labor force in each country is denoted as L_i. There is a continuum of goods with measure γ, and the utility function of a representative consumer is

$$Q = \left(\int_0^{\overline{\omega}} (q_\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}} \quad \text{for } \sigma \geq 1,$$

where q_ω is the consumption of good ω, σ is the elasticity of substitution, and $\overline{\omega} \leq \gamma$ is the measure of goods that are actually produced. We will specify how $\overline{\omega}$ is determined shortly. The standard price index is

$$P_j \equiv \left(\int_0^{\overline{\omega}} p_{j\omega}^{1-\sigma} d\omega \right)^{\frac{1}{1-\sigma}}. \tag{1}$$

Total revenue in country i is denoted as R_i, which also equals the total income. Welfare of country i's representative consumer is therefore R_i/P_i, which can also be interpreted as real GDP. The quantity demanded ($q_{j\omega}$) and expenditure ($E_{j\omega}$) for the product ω in country j are given by

$$\begin{aligned} q_{j\omega} &= Q_j \left(\frac{p_{j\omega}}{P_j} \right)^{-\sigma}, \\ E_{j\omega} &= R_j \left(\frac{p_{j\omega}}{P_j} \right)^{1-\sigma}. \end{aligned} \tag{2}$$

[8] Other recent studies on gains from trade via different angles from the ACR finding include at least Caliendo and Parro (2015) on the roles of intermediate goods and sectoral linkages; Fajgelbaum and Khandelwal (2016) on the differential effects of trade liberalization on consumers with different income; and di Giovanni et al. (2014) and Hsieh and Ossa (2016) on the global welfare impact of China's trade integration and productivity growth. Our work differs in that we focus on the pro-competitive effects.

[9] Since Eaton and Kortum (2002), quantitative analysis of trade in a multiple-country framework has become computationally tractable and widely applied. See, for examples, Alvarez and Lucas (2007) and Caliendo and Parro (2015), among many others. Nevertheless, as our study focuses on the distribution of markups and relies on firm-level data, we cannot use a multiple-country framework because we do not have access to firm-level data in multiple countries.

For each good ω, there are n_ω number of potential firms. Production technology is constant returns to scale, and for a firm k located at i, the quantity produced is given by

$$q_{\omega,ik} = \varphi_{\omega,ik} l_{\omega,ik},$$

where $\varphi_{\omega,\,ik}$ is the Hicks-neutral productivity of firm $k \in \{1,2,\ldots,n_{\omega,\,i}\}$, $n_{\omega,\,i}$ is the number of entrants in country i for good ω, and $l_{\omega,\,ik}$ is the amount of labor employed. Note the subtle and important difference between subscript $j\omega$ and ω, i. The former means that it is the purchase of ω by consumers at location j, and the latter is the sales or production characteristics of the firm located at i producing ω.

2.2. Measure of Goods and Number of Entrants

The number of entrants for each good $\omega \in [0,\gamma]$ in each country i is a random realization from a Poisson distribution with mean λ_i. That is, the density function is given by

$$f_i(n) = \frac{e^{-\lambda_i}\lambda_i^n}{n!}.$$

Poisson parameters provide a parsimonious way to summarize the overall competitive pressure (or entry effort) in the economy.[10] The total number of entrants for good ω across the two countries is $n_\omega = n_{\omega,\,1} + n_{\omega,\,2}$. There are goods that have no firms from either countries, and the total number of goods actually produced is given by

$$\overline{\varpi} = \gamma[1 - f_1(0) f_2(0)] = \gamma\left[1 - e^{-(\lambda_1+\lambda_2)}\right]. \tag{3}$$

There is also a subset of goods produced by only one firm in the world, and in this case, this firm charges monopoly prices in both countries. For the rest, the number of entrants in the world are at least two, and firms engage in Bertrand competition. We do not model entry explicitly. By this probabilistic formulation, we let λ_i summarize the entry effort in each country. From (3), we see that the larger the mean numbers of firms λ_i, the larger the $\overline{\varpi}$.

2.3. Productivity, Trade Cost, Pricing and Markups

Let wages be denoted as w_i. If the productivity of a firm is $\varphi_{i\omega}$, then its marginal cost is $w_i/\varphi_{i\omega}$ before any delivery. Assume standard iceberg trade costs $\tau_{ij} \geq 1$ (to deliver one unit to j from i, it must ship τ_{ij} units). Let $\tau_{ii} = 1$ for all i. Hence, for input ω, the delivered marginal cost from country i's firm k to country j is therefore $\frac{\tau_{ij} w_i}{\varphi_{\omega,ik}}$. For each $i\omega$, productivity $\varphi_{\omega,\,ik}$ is drawn from log-normal distribution, i.e., $\ln\varphi_{\omega,\,ik}$ is distributed normally with mean μ_i and variance η_i^2. Let $\varphi^*_{\omega,\,i}$ and $\varphi^{**}_{\omega,\,i}$ be the first and second highest productivity draws among the $n_{i\omega}$ draws.[11]

For each ω, the marginal cost to deliver to location 1, for the two lowest cost producers at 1, and the two lowest cost producers at 2, are then

$$\left\{\frac{\tau_{1j}w_1}{\varphi^*_{\omega,1}}, \frac{\tau_{1j}w_1}{\varphi^{**}_{\omega,1}}, \frac{\tau_{2j}w_2}{\varphi^*_{\omega,2}}, \frac{\tau_{2j}w_2}{\varphi^{**}_{\omega,2}}\right\}. \tag{4}$$

If the number of entrants is 1, 2, or 3, then we can simply set the missing element in the above set to infinity. Let $a^*_{j\omega}$ and $a^{**}_{j\omega}$ be the lowest and second lowest elements of this set. The monopoly pricing for goods sold in country j is $\overline{p}_{j\omega} = \frac{\sigma}{\sigma-1} a^*_{j\omega}$. In the equilibrium outcome of Bertrand competition, price equals the minimum of the monopoly price and the marginal cost $a^{**}_{j\omega}$ of the second lowest cost firm to deliver to j, i.e.

$$p_{j\omega} = \min\left(\overline{p}_{j\omega}, a^{**}_{j\omega}\right) = \min\left\{\frac{\sigma}{\sigma-1} a^*_{j\omega}, a^{**}_{j\omega}\right\}. \tag{5}$$

The markup of good ω at j is therefore

$$m_{j\omega} = \frac{p_{j\omega}}{a^*_{j\omega}} = \min\left\{\frac{\sigma}{\sigma-1}, \frac{a^{**}_{j\omega}}{a^*_{j\omega}}\right\}. \tag{6}$$

Note that firms' markups may differ from the markups for consumers. A non-exporter's markup is the same as the markup facing consumers, but an exporter has one markup for each market. Let the markup of an exporter producing ω be denoted as m^f_ω. Then, due to constant returns to scale,

$$m^f_\omega = \left(\frac{\text{costs}}{\text{revenue}}\right)^{-1} = \left(\frac{E_{1\omega}}{E_{1\omega}+E_{2\omega}} m^{-1}_{\omega,1} + \frac{E_{2\omega}}{E_{1\omega}+E_{2\omega}} m^{-1}_{\omega,2}\right)^{-1}.$$

In other words, an exporter's markup is a harmonic mean of the markups in each market, weighted by relative revenue.

We can now define *producers' aggregate markup*, M_i^{sell}. Let $\chi_j^*(\omega) \in \{1,2\}$ denote the source country for any particular good ω at destination j, and let $\phi_{j\omega} \equiv \left(\frac{p_{j\omega}}{P_j}\right)^{1-\sigma}$ denote country j's spending share on good ω. Then, we have

$$\begin{aligned} M_i^{sell} = \frac{R_i}{w_i L_i} &= \frac{\int_{\{\omega:\chi_1^*(\omega)=i\}} \phi_{1\omega} R_1 d\omega + \int_{\{\omega:\chi_2^*(\omega)=i\}} \phi_{2\omega} R_2 d\omega}{\int_{\{\omega:\chi_1^*(\omega)=i\}} m_{1\omega}^{-1}\phi_{1\omega} R_1 d\omega + \int_{\{\omega:\chi_2^*(\omega)=i\}} m_{2\omega}^{-1}\phi_{2\omega} R_2 d\omega} \\ &= \left(\int_{\{\omega:\chi_1^*(\omega)=i\}} m_{1\omega}^{-1} \frac{\phi_{1\omega} R_1}{R_i} d\omega + \int_{\{\omega:\chi_2^*(\omega)=i\}} m_{2\omega}^{-1} \frac{\phi_{2\omega} R_2}{R_i} d\omega\right)^{-1}, \end{aligned} \tag{7}$$

which is the revenue-weighted harmonic mean of markups of all goods with *source* at location i. Similarly, *consumers' aggregate markup* M_i^{buy} is the revenue-weighted harmonic mean across goods with *destination* at i:

$$M_i^{buy} = \left(\int_0^{\overline{\varpi}} m_{i\omega}^{-1} \phi_{i\omega} d\omega\right)^{-1}.$$

2.4. Wages and General Equilibrium

Observe that the total imports of country j from country i is

$$R_{j,i} = \int_{\left\{\omega:\chi_j^*(\omega)=i\right\}} E_{j\omega} d\omega = R_j \int_{\left\{\omega:\chi_j^*(\omega)=i\right\}} \left(\frac{p_{j\omega}}{P_j}\right)^{1-\sigma} d\omega \equiv R_j \phi_{j,i}, \tag{8}$$

where $\chi_j^*(\omega) \in \{1,2\}$ denotes the source country for any particular good ω at destination j and $\phi_{j,\,i}$ denote the spending share of country j's consumers on the goods produced in i. The balanced trade condition is therefore $R_{2,\,1} = R_{1,\,2}$, or equivalently,

$$R_2 \phi_{2,1} = R_1 \phi_{1,2}. \tag{9}$$

Combine (7) and (9), and we have

$$\begin{aligned} M_1^{sell} &= \left(\int_{\{\omega:\chi_1^*(\omega)=1\}} m_{1\omega}^{-1}\phi_{1\omega} d\omega + \int_{\{\omega:\chi_2^*(\omega)=1\}} m_{2\omega}^{-1}\phi_{2\omega} \frac{R_2}{R_1} d\omega\right)^{-1} \\ &= \left(\int_{\{\omega:\chi_1^*(\omega)=1\}} m_{1\omega}^{-1}\phi_{1\omega} d\omega + \int_{\{\omega:\chi_2^*(\omega)=1\}} m_{2\omega}^{-1}\phi_{2\omega} \frac{\phi_{1,2}}{\phi_{2,1}} d\omega\right)^{-1}. \end{aligned}$$

[10] Eaton et al. (2013) also model finite number of firms as a Poisson random variable, but for a very different purpose.

[11] Another non-fat-tailed distribution that is often used is bounded Pareto, e.g. Helpman et al. (2008) and Melitz and Redding (2015).

We choose country 1's labor as numeraire, and hence $w_1 = 1$, and $w \equiv w_2$ is also the wage ratio. It is readily verified that $\phi_{j\omega}$ depends only on relative wage w, but not on R_1, R_2, L_1, or L_2 directly. Hence, M_1^{sell} becomes a function of w only. Similarly, we have

$$M_2^{sell}(w) = \left(\int_{\{\omega:\chi_1^*(\omega)=2\}} m_{\omega,1}^{-1}\phi_{\omega,1}\frac{\phi_{2,1}}{\phi_{1,2}}(w)d\omega + \int_{\{\omega:\chi_2^*(\omega)=2\}} m_{\omega,2}^{-1}\phi_{\omega,2}d\omega \right)^{-1}.$$

We can then define R_1 and R_2 as a function of w:

$$R_1(w) = M_1^{sell}(w)L_1$$
$$R_2(w) = M_2^{sell}(w)wL_2.$$

Note that the above two equations are actually the labor market clearing conditions. Combining these two equations, we thus arrive at the following one equation in one unknown w:

$$\frac{R_1(w)}{R_2(w)} = \frac{M_1^{sell}(w)}{M_2^{sell}(w)}\frac{L_1}{wL_2}. \quad (10)$$

Once w is computed, R_1, R_2, M_1^{sell}, M_2^{sell}, and the trade flows $R_{j,\,i}$ are computed according to the above procedure. Also note that from (10), what matters for an equilibrium is the ratios R_2/R_1 and L_2/L_1, rather than the levels. This is not surprising as the model features constant returns to scale.

2.5. Welfare Decomposition

In this subsection, we show the decomposition of welfare, which is exactly that provided by Holmes et al. (2014). Here, we attempt to be brief and at the same time self-contained.

Let A_j be the price index at j when all goods are priced at marginal cost:

$$A_j = \int_0^{\overline{\omega}} a_{j\omega}^* \tilde{q}_{j\omega}^a d\omega,$$

where $\tilde{\mathbf{q}}_j^a = \{\tilde{q}_{j\omega}^a : \omega \in \mathbf{0}, \overline{\omega}]\}$ is the expenditure-minimizing consumption bundle that delivers one unit of utility under marginal cost pricing. Total welfare is defined as real income R_j/P_j. As the product of producers' aggregate markup and labor income entails total revenue (7), we can write welfare at location i as

$$\begin{aligned} W_j^{Total} &= \frac{R_j}{P_j} = w_j L_j \times M_j^{sell} \times \frac{1}{P_j} \\ &= w_j L_j \times \frac{1}{A_j} \times \frac{A_j \times M_j^{buy}}{P_j} \times \frac{M_j^{sell}}{M_j^{buy}} \\ &\equiv w_j L_j \times W_j^{Prod} \times W_j^A \times W_j^R. \end{aligned}$$

Without loss of generality we focus on the welfare of country 1, and by choosing numeraire, we can let $w_1 = 1$. As the labor supply L_j is fixed in the analysis, the first term in the welfare decomposition is a constant that we henceforth ignore. The second term $1/A_j$ is the *productive efficiency index* W_j^{Prod}; this is what the welfare index would be with constant markup. The index varies when there is technical change determining the underlying levels of productivity. It also varies when trade costs decline, decreasing the cost for foreign firms to deliver goods to the domestic country. Terms-of-trade effects also show up in W_j^{Prod} because a lower wage from a source country raises the index.

The third term is the *allocative efficiency index* W_j^A

$$W_j^A \equiv \frac{A_j \times M_j^{buy}}{P_j} = \frac{\int_0^{\overline{\omega}} a_{j\omega}^* \tilde{q}_{j\omega}^a d\omega}{\int_0^{\overline{\omega}} a_{j\omega}^* \tilde{q}_{j\omega} d\omega} \leq 1. \quad (11)$$

The inequality follows from the fact that under marginal cost pricing, $\tilde{q}_{j\omega}^a$ is the optimal bundle, whereas $\tilde{q}_{j\omega}$ is the optimal bundle under actual pricing. If markups are constant, then for any pair of goods, the ratio of actual prices equals the ratio of marginal cost. In this case, the two bundles become the same and $W_j^A = 1$. Once there is any dispersion of markups, welfare deteriorates because resource allocation is distorted. Goods with higher markups are produced less than optimally (employment is also less than optimal), and those with low markups are produced more than optimally (employment is also more than optimal).

The fourth term is a "terms of trade" effect on markups that depends on the ratio of producers' aggregate markup to consumers' aggregate markup; thus we call it *relative markup effect* W_j^R. This term is intuitive because a country's welfare improves when its firms sell goods with higher markups while its consumers buy goods with lower markups. This term drops out in two special cases: under symmetric countries where the two countries are mirror images of each other; and under autarky, as there is no difference between the two aggregate markups.

Note that as Holmes et al. focus on the symmetric country case, they do not explicitly analyze the relative markup effect W_j^R. As fitting to the Chinese economy, we allow asymmetries between countries in all aspects of the model (labor force, productivity distribution, entry, and wages). Also note that the above decomposition requires only homothetic preference and is thus applicable to all market structures.[12]

2.6. The Productive Efficiency and the ACR Formula

As is well known, the ACR welfare formula captures the gains from trade globally (i.e., for arbitrary changes in trade cost) in a certain class of models with a constant trade elasticity. This class includes BEJK and features no pro-competitive effect. In our model in which pro-competitive effects may exist and trade elasticity may vary, the ACR formula does not hold for arbitrary changes in trade costs. Nevertheless, as indicated by ACR, for models with variable trade elasticity, the ACR formula may still capture the total gains from trade locally (i.e., for infinitesimal changes in trade cost).[13] Thus, we are interested in examining whether our model predicts larger/smaller or similar total gains from trade as compared with the local ACR formula.

We start the comparison by examining the similarity between the productive efficiency W_j^{Prod} and the ACR welfare formula. Note that ACR's proof of their theorems covers both perfect competition and monopolistic competition. They do not prove why the BEJK model, which features head-to-head Bertrand competition, fits their formula. As Holmes et al. (2014) highlights, the distributional assumption and the number of firms are the key. Whereas BEJK features a constant trade elasticity, the trade elasticity in our model is a variable, and thus the macro restriction R3 in ACR does not hold here.

Following ACR, the import demand system is a mapping from $(\{w_i\}, \{\tau_{ij}\}, \{N_i\})$ into $\mathbf{X} \equiv \{X_{ij}\}$, where X_{ij} is the trade flow from i to j and N_i is the measure of goods that is produced in each country i. R3 in ACR is a restriction on partial trade elasticity $\epsilon_j^{ii'} \equiv \partial \ln (X_{ij}/X_{jj})/\partial \ln \tau_{i'j}$ of this system such that for any importer j and any pair of exporters $i \neq j$ and $i' \neq j$, $\epsilon_j^{ii'} = \epsilon < 0$ if $i = i'$, and zero otherwise. Since there are only two countries in our model, we are not concerned with the country index $i' \neq i, j$ here, and thus we simply denote $\epsilon_j^{ii'}$

[12] For welfare decomposition under non-homothetic preference and monopolistic competition, see Weinberger (2015) and Dhingra and Morrow (2016).

[13] See footnote 13 and page 109 in ACR. This statement is true if the restriction R3 in their paper holds locally.

as ϵ_j^i. Let v_{ij} be the share of country j's expenditure on goods from i. Then, in our two-country model, for any $i \neq j$,

$$\epsilon_j^i = \frac{\partial \ln\left(\frac{X_{ij}}{X_{jj}}\right)}{\partial \ln \tau_{ij}} = \frac{\partial \ln\left(\frac{1-v_{jj}}{v_{jj}}\right)}{\partial \ln \tau_{ij}}. \tag{12}$$

Suppose we are in the class of models characterized in ACR with only two countries i and j. Before knowing if R3 holds, the following holds for welfare in country j, W_j,

$$\begin{aligned} d \ln W_j &= -\left(v_{ij}\frac{d \ln v_{ij} - d \ln v_{jj}}{\epsilon_j^i} + v_{jj}\frac{d \ln v_{jj} - d \ln v_{jj}}{\epsilon_j^i}\right) \\ &= \frac{1}{\epsilon_j^i} d \ln v_{jj}, \end{aligned} \tag{13}$$

where the last line uses $v_{ij} + v_{jj} = 1$, which implies that $v_{ij} d \ln v_{ij} + v_{jj} d \ln v_{jj} = 0$.[14] If R3 holds so that ϵ_j^i is a constant ϵ across i and j and across different levels of variable trade costs, then the local ACR formula can be expressed as

$$d \ln W_j^{ACR} = \frac{1}{\epsilon} d \ln v_{jj}. \tag{14}$$

Moreover, the global formula $W_j'/W_j = (v_{jj}'/v_{jj})^{\frac{1}{\epsilon}}$ holds when R3 holds. We repeat the derivation in ACR in (13) here to clarify that if R3 does not hold, the appropriate local trade elasticity should be ϵ_j^i, which by definition is the elasticity of $(1 - v_{jj})/v_{jj}$ to τ_{ij}. Thus, when numerically computing the trade elasticity in Section 4.2 for China's welfare ($j = 1$), it is done by varying τ_{21} by a small amount rather than by varying the symmetric cost $\tau_{21} = \tau_{12} = \tau$.[15]

We examine how productive efficiency in our model is related to the ACR formula. Recall that $W_j^{Prod} = 1/A_j$, where A_j is the price index under marginal cost pricing, i.e., the equilibrium price index when all goods are priced at marginal costs. Thus, ACR's proof of Proposition 1 for the perfect competition case actually applies up to Step 3 with W_j and P_j there replaced with W_j^{Prod} and A_j here. That is, letting $\tilde{v}_{ij}$ and $\tilde{\epsilon}_j^i$ be the share of country j's expenditure on goods from i and the trade elasticity under marginal cost pricing, we have

$$d \ln A_j = \sum_{i=1}^{n} \tilde{v}_{ij} \frac{d \ln \tilde{v}_{ij} - \ln \tilde{v}_{jj}}{\tilde{\epsilon}_j^i}.$$

Similar to (13), for any $i \neq j$, the above implies

$$d \ln W_j^{Prod} = -d \ln A_j = \frac{1}{\tilde{\epsilon}_j^i} d \ln \tilde{v}_{jj}. \tag{15}$$

Note that the ACR formula (14) should be applied using actual trade flow to calculate trade elasticity and domestic expenditure share (that is, actual pricing (5) should be used), whereas (15) uses those under marginal cost pricing. However, there is a special case in which $\tilde{v}_{jj} = v_{jj}$ and hence $\tilde{\epsilon}_j^i = \epsilon_j^i$. When $\sigma = 1$, the preference becomes Cobb-Douglas:

$$U = \exp\left(\int_0^{\varpi} \ln q_\omega d\omega\right),$$

and the expenditure share on each good becomes the same (not responsive to prices). As the domestic expenditure share is simply the fraction of all goods consumed in country j that originate in country j, we have $\tilde{v}_{jj} = v_{jj}$. By (12), we also have $\tilde{\epsilon}_j^i = \epsilon_j^i$. In this case, $d \ln W_j^{ACR} = d \ln W_j^{Prod}$ with the trade elasticity being ε_j^i.

We have now proved the following proposition. Note in particular that this proposition is applicable to all distributions of productivity draws and of per-product number of firms.

Proposition 1. *For infinitesimal changes in τ, the change in the productive efficiency W_j^{Prod} can be expressed as*

$$d \ln W_j^{Prod} = \frac{1}{\tilde{\epsilon}_j^i} d \ln \tilde{v}_{jj},$$

where $\tilde{\epsilon}_j^i$ and $\tilde{v}_{jj}$ are the trade elasticity and domestic expenditure share under marginal cost pricing. When $\sigma = 1$ (Cobb-Douglas case), $\tilde{v}_{jj} = v_{jj}$, $\tilde{\epsilon}_j^i = \epsilon_j^i$, and $d \ln W_j^{Prod} = d \ln W_j^{ACR}$.

In the case of $\sigma = 1$, this proposition says that for infinitesimal changes in τ, the ACR formula captures productive efficiency but not the total gains from trade. That is, in this case,

$$d \ln W_j^{Total} - d \ln W_j^{ACR} = d \ln W^A + d \ln W_j^R.$$

The distributional assumption in BEJK entails $d \ln W^A + d \ln W_j^R = 0$ because the resulting markup distribution is invariant to trade cost. This is not the case here. In the case of $\sigma > 1$, our quantitative analysis using Chinese data in Section 4.2 reveals that $d \ln W_j^{ACR}$ is still relatively close to $d \ln W_j^{Prod}$, and that $d \ln W_j^{Total}$ are larger than $d \ln W_j^{ACR}$.

For the intuition behind the gap, we distinguish all possible six cases of pricing, markups, and trade flows in the following table. Without loss of generality, we focus on the market at country 1, i.e., $j = 1$. Denote (i, i') as the pair of locations where the first and second lowest marginal costs to deliver to country 1 are located. We use $(\bar{i})$ to denote the case when the lowest marginal cost is from country i and it charges the monopoly price in equilibrium.

	(1,1)	(1,2)	(2,1)	(2,2)	$(\bar{1})$	$(\bar{2})$
markup	$\frac{\varphi_1^*}{\varphi_1^{**}}$	$\frac{\tau w \varphi_1^*}{\varphi_2^*}$	$\frac{\varphi_2^*}{\tau w \varphi_1^*}$	$\frac{\varphi_2^*}{\varphi_2^{**}}$	$\frac{\sigma}{\sigma-1}$	$\frac{\sigma}{\sigma-1}$
price	$\frac{1}{\varphi_1^{**}}$	$\frac{\tau w}{\varphi_2^*}$	$\frac{1}{\varphi_1^*}$	$\frac{\tau w}{\varphi_2^{**}}$	$\frac{\sigma}{\sigma-1}\frac{1}{\varphi_1^*}$	$\frac{\sigma}{\sigma-1}\frac{\tau w}{\varphi_2^*}$
markup affected by τ	No	Yes	Yes	No	No	No
import affected by τ	No	No	No	Yes	No	Yes

Note that for infinitesimal changes, the effect of a good ω switching between cases can be ignored because at the border between any two cases, the markups must be the same. Thus, apart from the general equilibrium effect on macro variables, the above table provides a comprehensive anatomy of the effect of changes in τ. Thus, apart from the general equilibrium effect on R_j and P_j, import is affected by τ directly in the cases where prices are affected by τ and the suppliers are located at country 2. We ignore the effect on exports because imports are what is needed for the ACR formula. To look at pro-competitive effects, we look only at two cases where markups are affected by trade cost – (1,2) and (2,1). In Case (1,2), a lower τ decreases both the price and markup but has no effect on import because the supplier is domestic; this is similar to the entry-deterrence example mentioned in the

[14] The expression in (13) can be easily obtained in ACR's proof of Proposition 1 in the perfect competition case. In the case of monopolistic competition, the same expression can be obtained by observing (A37), $d \ln W_j = -d \ln P_j$, $d \ln \alpha_{ij}^* = d \ln \xi_{ij}/(1-\sigma) = 0$ (p. 126) and $d \ln N_j = 0$ (p.127). Since we will apply the ACR formula in our model, $d \ln \xi_{ij} = 0$ because there are no fixed exporting costs. ACR show that R1 and R2 imply $d \ln N_j = 0$.

[15] Note that in Melitz and Redding (2015), when they calculate trade elasticity in the case when it is a variable, they vary τ instead of τ_{21}. This is because they assume countries are symmetric and thus domestic expenditure shares v_{jj} are the same across countries.

introduction. In Case (2,1), a lower τ increases the markup but does not affect the price and import because the foreign supplier is constrained only by the domestic best. Thus, in cases where markups are affected by τ, imports are unaffected. If the expenditure share of each case is unaffected by small changes in τ, then the welfare impacts of τ via markups are totally independent of imports (Proposition 1). The reason why Proposition 1 need not hold under $\sigma > 1$ is that changes in trade cost τ may change the expenditure shares across goods and hence across different cases. Nevertheless, it will be seen in the quantitative analysis in Section 4.2 that the effects due to changes in expenditure share are minor, as the extra gains from trade over the ACR formula remain roughly those due to pro-competitive effects.

In sum, Proposition 1 and the above table show how head-to-head competition separates markups and trade flows, and hence make the total welfare gains from trade in our model different from the ACR formula. In contrast, the total gains from trade in EMX can be captured by the ACR formula because even with finite number of firms, each firm owns a variety and hence a demand curve along which the pricing is determined, taking into account strategic interactions among firms. A change in τ changes the foreign supplier's delivered marginal cost, and therefore changes the price, markup, and import simultaneously. Similarly, even though the ACR formula must be modified in Arkolakis et al. (2019) to account for the change from CES preference to a general preference that allows variable markup, the fact that each firm owns a variety under monopolistic competition still makes trade flows sufficient statistics for welfare gains from trade.

3. Quantifying the Model

As the markup distribution is the central focus of the paper, our approach of quantifying the model relies heavily on the distribution of markups, which is estimated from the Chinese firm-level data. Note that unlike EMX whose benchmark focuses on symmetric countries, our empirical implementation focuses on asymmetric countries, as the large wage gap between China and the ROW should not be ignored since it may have a large impact on parameter estimates, as well as potential large general equilibrium effects in counter-factuals. Despite the lack of firm level data in the ROW, we demonstrate that separating moments of exporters and nonexporters can help identify the different parameters of the two countries. In this subsection, we first describe the data and approach by which our model is quantified. We then present and discuss the estimation results, and make a comparison with the BEJK model.

3.1. Data

Our firm-level data set comes from the Economic Census data (1995 and 2004) from China's National Bureau of Statistics (NBS), which covers all manufacturing firms, including SOEs. The sample sizes for 1995 and 2004 are 458, 327 and 1, 324, 752, respectively.[16] The advantage of using this data set, instead of the commonly used firm-level survey data set, which reports all SOEs and only those private firms with revenues of at least 5 million renminbi, is that we do not have to deal with the issue of truncation. As we are concerned with potential resource misallocation between firms, it is important to have the entire distribution. We estimate the models separately for the years 1995 and 2004.

We obtain world manufacturing GDP and GDP per capita from the World Bank's World Development Indicators (WDI). The aggregate Chinese trade data is obtained from the UN COMTRADE.

3.2. Estimation of Markups

Under the constant returns to scale assumption, a natural way to estimate markups is by taking the ratio of revenue to total costs, i.e., revenue productivity, or what we call *raw markup*. However, it is important to recognize that, in general, raw markups may differ across firms, not only because of the real markup differences, but also because of differences in the technology with which they operate. To control for this potential source of heterogeneity, we use modern IO methods to purge our markup estimates of the differences in technology. In particular, we estimate markups following DLW's approach,[17] who calculate markups as

$$m_\omega = \frac{\theta_\omega^X}{\alpha_\omega^X},$$

where θ_ω^X is the input elasticity of output for input X, and α_ω^X is the share of expenditure on input X in total revenue. To map our model into firm-level data, we relax the assumptions of a single factor of production and constant returns to scale. Following DLW, we assume a translog production function.[18] The estimation of firm-level markup hinges on choosing an input X that is free of any adjustment costs and the estimation of the elasticity of output to this input, θ_ω^X. As labor is largely not freely chosen in China (particularly SOEs) and capital is often considered a dynamic input (which makes its input elasticity difficult to interpret), we choose intermediate materials as the input to estimate firm markup (see also DLW). The full details of the markup estimation are relegated to Appendix A1.

Table 1 gives summary statistics of the markup distribution,[19] with breakdowns in each year and between exporters and non-exporters. Observe that the (unweighted) mean markups all decrease between 1995 and 2004 for all firms, both exporters and non-exporters. The (unweighted) standard deviation of markups decreases for non-exporters, but increases slightly for exporters. Because there are more non-exporters than exporters and the decrease in the standard deviation of non-exporters is larger than the increase in the standard deviation of exporters, the overall standard deviation decreases. Almost all of the percentiles decreased between 1995 and 2004. This is consistent with the pattern seen in Figure 1 where the entire distribution becomes more condensed.

However, we note that the pattern described in Table 1 only hints at the existence of pro-competitive effects. The reduction of dispersion of firm markups does not necessarily mean that the allocative efficiency increases, as allocative efficiency depends on consumer markups rather than firm markups. It does show that the markets facing Chinese firms became more competitive. Also, we cannot reach a conclusion yet about the relative markup effect, as we do not observe the consumers' aggregate markup directly. We need to quantify the model and simulate both types of markups to conduct welfare analysis.

3.3. Simulated Method of Moments

3.3.1. Method

Given measures of $\{w, R_1, R_2\}$, we use moments of markups, trade flows, number of firms and fraction of exporters to estimate all parameters except L_2/L_1 by SMM. We estimate the parameters using SMM for

[16] The original data sets have larger sample sizes, but they also include some (but not all) non-manufacturing industries, as well as firms without independent accounting and village firms, which entail numerous missing values. The final sample is obtained after excluding these cases and adjusting for industrial code consistency.

[17] We also conduct estimation and counter-factual analysis under raw markups as a robustness check.

[18] In our implementation of the DLW approach using Chinese firm-level data under the translog production function, which allows variable returns to scale, it turns out that the returns to scale are quite close to constant. See Panel B of Table A1 in Appendix A1. Interestingly, EMX also found similar results using Taiwanese firm-level data.

[19] Following the literature, e.g., De Loecker et al. (2016) and Lu and Yu (2015), we trim the estimated markup distribution in the top and bottom 2.5 percentiles to alleviate the concern that the extreme outliers may drive the results. Our results are robust to alternative trims (e.g., the top and bottom 1%; results are available upon request). We also drop estimated markups that are lower than one, as our structural model does not generate such markups.

Table 1
Detailed markup distributions.

	All firms		Exporters		Non-exporters	
Year	1995	2004	1995	2004	1995	2004
Mean	1.428	1.372	1.340	1.318	1.432	1.379
Std. dev.	0.495	0.479	0.431	0.438	0.498	0.483
p1	1.005	1.004	1.003	1.004	1.005	1.004
p5	1.022	1.019	1.017	1.017	1.023	1.019
p10	1.044	1.036	1.034	1.032	1.045	1.037
p25	1.114	1.091	1.084	1.077	1.116	1.093
p50	1.262	1.207	1.120	1.168	1.266	1.213
p75	1.538	1.437	1.414	1.362	1.544	1.447
p90	2.015	1.893	1.784	1.747	2.023	1.909
p95	2.464	2.379	2.199	2.183	2.475	2.400
p99	3.528	3.509	3.299	3.364	3.537	3.523

1995 and 2004 separately. To adequately reflect the changes in the environment of the Chinese economy, it is important to allow all parameters to vary between the two years. If we instead have the change in trade cost τ in between two years explain all the changes in the observed moments, then the role of trade cost may be exaggerated.

For $i=1,2$, the parameters to be estimated are

τ : trade cost
γ : total measure of goods
λ_i : mean number of entrants per product
μ_i : mean parameter of log-normal productivity draw
η_i : standard deviation parameter of log-normal productivity draw
σ : elasticity of substitution

Note that for productivity, we normalize $\mu_2=0$ (when $\ln\varphi$ is zero, $\varphi=1$) because only the relative magnitude of μ_1 to μ_2 matters. Choosing μ_2 amounts to choosing a unit. Let the inverses of markups be called *cost shares*, as they are the shares of costs in revenues. As shown in Section 2.3, aggregate markups are harmonic means of markups, which are the inverses of the arithmetic means of cost shares weighted by revenues. However, it is unclear how a "harmonic variance" could be defined. Since the (arithmetic) variance of markup is positively related to the variance of cost shares, we choose to work with cost shares in calculating moments.

In order to use SMM to estimate these eight parameters, we need at least eight moments. We use the following 14 moments: the import and export shares; relative number of firms; fraction of exporters; revenue-weighted mean and standard deviation of cost shares for both exporters and non-exporters; and the 50th, 95th, and 99th percentiles of the markup distribution for exporters and non-exporters.[20]

Trade is highly aggregated; hence almost all parameters affect the three trade moments. Nevertheless, we expect them to be more sensitive to the trade cost τ and hence help identify this parameter. We will explain shortly how the relative number of firms help identify γ. The standard deviation and the two measures of centrality are meant to capture the overall pattern of the markup distributions. As the model is about the top firms, the 95th and 99th percentiles are important moments to have. Moreover, the model implies that the range of markup is $m\in[1,\frac{\sigma}{\sigma-1}]$, and hence the monopoly markup is the upper bound of markup distribution. Recall the economics behind this. An active firm of a product charges the second lowest marginal cost when such cost is sufficiently low. When the second marginal cost is high, the markup is bounded by the monopoly markup because the firm's profit is still subject to the substitutability between products. The higher the substitutability (σ), the lower the monopoly markup the firm will charge. Thus, another important role of the 99th percentile is to reflect this upper bound and help identify σ.

As our data shows whether a firm is an exporter or not, we use moments of exporters and non-exporters separately because the way in which parameters of countries 1 and 2 (China and the ROW) enter these moments differs between these two groups. The intuition is clear: Chinese exporters face direct competition in the ROW's markets and non-exporters face foreign competition on their home turf. As we lack firm-level data from the ROW, this approach is crucial for backing out the parameters of the ROW.[21] The parameters of the ROW is not needed in a symmetric-country estimation/calibration, which may explain why it is often adopted in the literature. We will also estimate a symmetric country version for comparison. Nevertheless, our exercise demonstrates that this approach of separating moments of exporters and non-exporters works well for asymmetric-country estimation.

The model structure implies that all of the above-mentioned moments can be simulated for a given set of all parameters besides L_2/L_1 and observed macro variables $\{w,R_1,R_2\}$. Specifically, the actual supplier and the second best supplier in each market are identified by (4). Pricings and markups are given by (5) and (6). Then, the price index and sales are given by (1) and (2). Import and export are then aggregated from sales. The general equilibrium conditions are not used in SMM since the macro variables $\{w,R_1,R_2\}$ are given, and hence L_2/L_1 is not needed. However, L_2/L_1 is still needed in counter-factuals, and it is inferred by matching the observed w with the estimated parameters using the final general equilibrium condition (10).[22] The key advantage of this approach is that it makes the multiple-sector estimation possible because it allows a convenient sector-by-sector estimation where different sectors are linked by a Cobb-Douglas utility function. There are 29 manufacturing sectors in China, and each sector carries 8 parameters. If one were to also include L_2/L_1 in the SMM procedure (and perhaps use observed w as a moment), this implies the need to estimate $29\times 8+1=233$ parameters simultaneously in an over-identified estimation with w computed as a general equilibrium object. This is not feasible. For more details of the multiple-sector estimation, see Section 5.1. Note that the model simulated R_2/R_1 is generally different from the data ones. We will check the external validity of our approach by comparing the data and simulated R_2/R_1.

Recall that the actual measure of goods is given by (3): $\overline{\varpi}=\gamma[1-e^{-(\lambda_1+\lambda_2)}]$, but this is not directly observed. What is observable is the number of active Chinese firms:

$$N_1=\gamma\left(1-e^{-\lambda_1}\right)\times\ \Pr\left[\frac{1}{\varphi^*_{1\omega}}<\frac{w\tau}{\varphi^*_{2\omega}}\right].$$

As our simulation of the number of goods must be discrete, γ must be a large number in order to match the large number of Chinese firms in the data. To calculate the relative number of Chinese firms, we divide both sides of the above equation by $\overline{N}$, a large number that is chosen for normalizing the moment:

$$\frac{N_1}{\overline{N}}=\frac{\gamma\left(1-e^{-\lambda_1}\right)}{\overline{N}}\times\ \Pr\left[\frac{1}{\varphi^*_{1\omega}}<\frac{w\tau}{\varphi^*_{2\omega}}\right]. \tag{16}$$

Here, we define $\tilde{\gamma}$ by $\gamma\equiv$ *total_goods_baseline*$^*\tilde{\gamma}$, where *total_goods_baseline* is set to be 250,000. We will report $\tilde{\gamma}$ in the place of γ to avoid reporting a very large number. The roles of

[20] The import share is the import penetration ratio, i.e. IM/(R1-EX+IM), and the export share is the total export divided by the same denominator. All the cost share moments are weighted by revenues.

[21] Whereas using such firm-level data with information on firms' exporting status gives the advantage of backing out parameters for the ROW, it also implies that one cannot use an n-country model with $n\geq 2$ unless one can gather firm-level data for all of these countries, which is a daunting task.

[22] We also opt not to measure L_2/L_1 from data, as this is a difficult object to measure because it involves considerations of human capital and non-manufacturing sectors, etc. Moreover, it is also difficult to find a robust way to combine the population size and human capital across different countries in the rest of the world.

total_goods_baseline and $\overline{N}$ are different. The former affects how precise a simulation is; the larger the *total_goods_baseline*, the more goods and firms there are and hence the more precise a simulation will be. The constant $\overline{N}$ is only for normalization and does not affect the estimates.[23]

How the macro variables $\{w, R_1, R_2\}$ are obtained from data is as follows. To calculate $w = w_2/w_1$, we first obtain the GDP per capita of China and the ROW from WDI.[24] We then proxy w_i by multiplying GDP per capita by the labor income shares for the ROW and China, which are taken from Karabarbounis and Neiman (2014).[25] For R_1 and R_2, we first obtain the manufacturing GDPs of China and the ROW from WDI data. We then use the input-output table for China (2002) and the US (1997–2005) to obtain GDP's share of total revenue. We then use such shares and the manufacturing GDPs to impute R_1 and R_2 as total revenue. Although our model does not distinguish value added and revenue, we choose to interpret R_i as total revenue rather than GDP to be consistent with our export and import moments, which are also in terms of revenue.

We use the equal-weight weighting matrix in our SMM implementation.[26] The nature of the moments implies that some empirical moments, e.g., mean and standard deviation of cost shares, would be estimated more accurately than others, e.g., the top quantiles of markups. Thus, under the optimal weighting matrix calculated from the inverse of the variance-covariance matrix of the empirical moments, those top-quantile moments would tend to have smaller weights. However, these top-quantile moments are crucial in identifying key model parameters $(\lambda_i, \eta_i, \sigma)$ as explained in Section 3.3.2 below, and thus we choose to treat each moment equally in our estimation procedure. The standard errors are calculated by the standard approach as in Adda and Cooper (2003, p. 88). As will be seen shortly, the standard errors in our implementation tend to be rather small due to the large sample sizes.

3.3.2. SMM Result

The estimation result is shown in Table 2. The model fits the data moments reasonably well, and the small standard errors indicate that each parameter is relatively precisely estimated. The bottom row reports data and model simulated R_2/R_1, and they turn out to be reasonably close, serving as additional validation of the model.

As we estimate the models for 1995 and 2004 separately, the changes of the parameters are strikingly consistent with well-known empirical patterns about the Chinese economy during this period. From 1995 to 2004, the estimate of τ shows a dramatic decrease from 2.31 to 1.78. The measure of goods γ more than triples from 0.19 to 0.66. This basically reflects the sharp increase in the number of firms between the two Economic Censuses, from 458,327 in 1995 to 1,324,752 in 2004, which is almost triple. The mean number of entrants per product in China (λ_1) increased from 2.46 to 2.62, whereas in the ROW it decreased from 5.54 to 5.02. China's mean log productivity (μ_1) relative to the ROW increased from −2.40 to −1.76. These numbers are negative, meaning that China's productivity is lower than that of the ROW. Also, we see a slight decrease (increase) in the dispersion parameter of the productivity distribution in China (ROW). Interestingly, the productivity dispersion is larger in China than in the ROW, which is consistent with the finding by Hsieh and Klenow (2009).[27]

Table 2
SMM results.

		1995		2004	
Predetermined					
w	Relative wages (the ROW to China)	10.25		5.18	
R1	China's manufacturing sales ($b)	918,291		2,343,328	
R2	ROW's manufacturing sales ($b)	9,397,500		14,737,500	
Moments for SMM		Data	Model	Data	Model
Import share		0.130	0.148	0.222	0.252
Export share		0.153	0.176	0.249	0.273
Relative number of firms		0.210	0.193	0.596	0.605
Fraction of exporters		0.044	0.023	0.105	0.064
Mean cost share for exporters		0.845	0.802	0.801	0.789
Std of cost share for exporters		0.135	0.135	0.142	0.139
p50 markup for exporters		1.196	1.207	1.168	1.224
p95 markup for exporters		2.199	2.173	2.183	2.207
p99 markup for exporters		3.299	3.202	3.364	3.225
mean cost share for non-exporters		0.789	0.715	0.829	0.763
std of cost share for non-exporters		0.147	0.185	0.139	0.161
p50 markup for non-exporters		1.266	1.383	1.213	1.285
p95 markup for non-exporters		2.475	2.752	2.400	2.193
p99 markup for nonexporters		3.537	3.202	3.523	2.735
Parameter values		Estimates	s.e.	Estimates	s.e.
τ, trade cost		2.311	0.027	1.782	0.007
γ, measure of goods relative to total_goods_baseline		0.186	0.002	0.659	0.003
λ_1, Poisson parameter, China		2.455	0.037	2.618	0.017
λ_2, Poisson parameter, ROW		5.535	0.037	5.024	0.048
μ_1, mean of log productivity, China relative to ROW		−2.397	0.023	−1.756	0.012
η_1, std of log productivity, China		0.450	0.004	0.425	0.002
η_2, std of log productivity, ROW		0.351	0.022	0.357	0.011
σ, elasticity of substitution		1.454	0.003	1.449	0.003
Simulated R2/R1		Data	Model	Data	Model
R2/R1		10.234	9.388	6.289	5.875

Notes: All units, if any, are in billions USD, current price. The import share is the import penetration ratio, i.e. IM/(R1-EX+IM), and the export share is the total export divided by the same denominator. All the cost share moments are weighted by firms' revenues. Recall that a firm's cost share is the inverse of its markup. p# denotes the #-th percentile.

The σ estimate we obtain is approximately 1.45 in both years.[28] This estimate is quite low compared with those estimates in models that feature constant markups (often a CES preference coupled with either monopolistic competition or perfect competition), and this is driven mainly by the need to fit the two 99th percentiles in the markup distribution. Recall that $\sigma/(\sigma-1)$ in our model is the upper bound rather than the average of markups. Under a constant-markup model and using the harmonic mean of firm markups in 1995, 1.259, this implies $\sigma = 4.86$. However, in the current model, this value of σ implies that $m \in [1, 1.259]$, which cuts 50.6% off the estimated markup distribution. Then, these large markups where most distortions come from are ignored. In fact, the pro-competitive effects of trade become negligible under $m \in [1, 1.259]$ because the associated allocative efficiency is much closer to the first-best case (constant markup) without the very skewed larger half of the markups. EMX also found that the extent of pro-competitive effects depends largely on the extent to which markups can vary in the model. In fact, our estimate is strikingly similar

[23] We set $\overline{N} = 2,000,000$ so that the relative numbers of firms are 0.210 and 0.596 in 1995 and 2004, respectively. We initially set *total_goods_baseline* to 2,500,000. However, we find that the calculated moments under *total_goods_baseline*=250,000 are virtually the same as those under 2,500,000. For faster computing speed, we thus set *total_goods_baseline*=250,000. To fit the above-mentioned relative numbers of firms (the left-hand side of [16]) in the SMM procedure, we set $\overline{N} = 200,000$ so that we effectively scale both the numerator and denominator of the right-hand side of (16) by 1/10.

[24] The ROW's GDP per capita is the population-weighted average of GDP per capita across all countries other than China.

[25] The ROW's labor share is the weighted average of labor share across all countries besides China, with the weight being relative GDP.

[26] Specifically, this weighting matrix is the inverse of the diagonal matrix with each diagonal element being the square of each data moment. This is equivalent to using the identity matrix if the moment error is first normalized by the data moment. Normalization is needed because the magnitudes of the data moments vary substantially from 0.13 to 3.54.

[27] The mean of a log-normal distribution is $e^{\mu+\eta^2/2}$. According to our estimates of μ_1 and η_1 in these two years, this translates to an annual productivity growth rate of 7.25%. This impressive growth rate is actually similar to the 7.96% estimated by Brandt et al. (2012). Note that the 7.25% growth rate here is relative to the ROW. If the ROW also grows in their productivity, the actual productivity growth rate could be even higher. In fact, Brandt et al. (2017) find a 12% average TFP growth rate at industry level. The data used in both above-mentioned papers is the annual manufacturing survey data from 1998 to 2007.

[28] Note that this estimate of σ is not sensitive to sample size. In our multi-sector exercise, the unweighted means of σ_s are 1.56 and 1.53 for 1995 and 2004, respectively, and 24 of the 29 σ_s are within one standard deviation from the mean in both years. See Section 5.1.4.

Table 3
Jacobian matrix.

Moments	τ	γ	λ_1	λ_2	μ_1	η_1	η_2	σ
Import share	−0.409	0.005	−0.072	0.001	0.251	−0.019	0.554	−0.030
Export share	−0.742	0.003	0.082	−0.101	0.966	0.665	−0.204	−0.026
Relative number of firms	0.252	0.971	0.102	−0.015	0.429	0.179	−0.769	0.000
Fraction of exporters	−0.167	0.001	0.006	−0.018	0.130	0.167	0.012	0.000
Mean cost share for exporters	−0.031	−0.001	0.023	0.012	0.040	−0.140	−0.034	0.021
Std of cost share for exporters	0.022	0.003	−0.016	−0.016	−0.041	−0.152	0.062	−0.050
p50 markup for exporters	0.055	−0.006	−0.033	−0.016	−0.052	0.269	0.010	0.006
p95 markup for exporters	1.597	−0.408	−1.949	−0.556	−2.088	−13.072	1.419	−0.070
p99 markup for exporters	0.000	0.000	0.000	0.000	0.000	0.000	0.000	−5.107
Mean cost share for non-exporters	−0.088	−0.003	0.021	0.002	−0.063	−0.360	0.089	0.014
Std of cost share for non-exporters	0.054	0.003	−0.008	−0.001	0.030	0.175	−0.023	−0.010
p50 markup for non-exporters	0.154	0.011	−0.044	−0.005	0.141	0.639	−0.173	0.000
p95 markup for non-exporters	0.929	−0.004	−0.159	−0.027	0.660	2.973	−0.480	0.000
p99 markup for non-exporters	1.914	0.055	−0.205	−0.067	0.880	3.636	−0.491	−0.877

Notes: Each entry of this table gives the rate of change of a moment to a parameter. This is based on the benchmark estimation of the 2004 model. The larger the absolute value of the rate of change, the more sensitive this moment is to the parameter, and the more useful this moment is in identifying this parameter.

to the estimate of the same parameter (1.37) in Simonovska and Waugh (2014) with the optimal weighting matrix in their method of moments procedure.

Note that in BEJK, the trade elasticity is given by the tail index of the Fréchet distribution, and is independent of the elasticity of substitution σ. In our model where the productivity draws deviate from Fréchet, σ may potentially matter in determining trade elasticity, but the effect seems small, as we will see in Section 4.2 that the trade elasticities in our model are quite close to those found by Simonovska and Waugh (2014) under the BEJK model.

Based on the 2004 estimation, we calculate a Jacobian matrix in which each entry gives a rate of change of a moment to a parameter; this is shown in Table 3. The larger the absolute value of a rate of change, the more sensitive this moment is to the parameter, and hence the more useful this moment is in identifying this parameter, at least at the local area of the optimal estimates. With such Jacobian matrices, the asymptotic variance-covariance matrices of the optimal estimates can be calculated to produce the standard errors reported in Table 2.

It is obvious from Table 3 that σ is almost single-handedly determined by the 99th percentile of markups for exporters, and this moment has little influence on other parameters. Trade cost τ affects almost all moments significantly except the 99th percentile of markups for exporters. It is natural to see that the two trade moments, the relative number of Chinese firms and the fraction of exporters are particularly strong for identifying this. Interestingly, when τ increases, the 95th percentiles of markups for both exporters and non-exporters, as well as the 99th percentile of markups for exporters, increase sharply. For non-exporters, this is intuitive because a higher τ provides non-exporters more insulation from foreign competition, and the top non-exporters gain more from this. For exporters, a higher τ makes it harder for them to compete in foreign markets, but recall that an exporter's markup is a harmonic mean of the markups in both the domestic and foreign markets. It must be that the gains in markups at home outweigh the losses in markups in foreign markets.

For the identification of λ_1, the top percentiles of markups play the dominant role. The intuition is as follows. Fixing other parameters, when λ_1 increases, the number of entrants per product in China increases. Due to the non-fat-tailed nature of the productivity distribution, the ratio between the top two draws is narrowed, but since this ratio is indeed the markup and since this is particularly pronounced for the top markups, the top percentiles are particularly useful in identifying this parameter. The relative number of firms also plays some role, as (16) shows that the larger the λ_1, the larger the probability that China draws a positive number of firms from the Poisson distribution. For λ_2, the 95th percentile of markups for exporters and the export share are the key moments. A larger λ_2 implies fiercer competition on the foreign turf for exporters as it brings out better competitors from the ROW, reducing both China's export share and the 95th percentile of markups for exporters.

For the measure of goods γ, the relative number of Chinese firms is the most useful moment. An increase in mean productivity parameter μ_1 increases export share, the number of Chinese firms, and the fraction of exporters, but decreases the import share. These are all intuitive. However, an increase in μ_1 sharply increases the 95th percentile markup for non-exporters but sharply decreases the 95th percentile markup for exporters. This is because top non-exporters are actually not the most productive firms – their productivities are somewhere in the middle of the distribution and hence they gain in markup by having higher productivity. In contrast, top exporters are the most productive firms, and they lose in markup when they become even more productive, due to the compression at the upper tail of the productivity distribution.

For η_1 and η_2, first note that they are not only dispersion parameters, but their increases induce increases in means as well. Hence, the direction of changes due to a change in η_1 is similar to that of a change in μ_1, but the intensities are quite different. For example, η_1 has much larger effects on almost all moments of markups than μ_1, but it has smaller impacts on the trade moments. In particular, the 95th percentile markup for exporters is extremely sensitive to η_1 because η_1 affects the top productivities much more than μ_1. Also note the interesting pattern: η_1 and η_2 affect many moments in opposite ways. An increase in η_2 increases both the mean and dispersion of the ROW's productivity, and this increases China's import share, and decreases China's export share and number of firms. It compresses the markup distribution of non-exporters, but it increases the 95th percentile of markups for exporters.

Finally, we discuss a point that is often mentioned in studies of the Chinese economy. China underwent various reforms, including but not limited to trade reforms, in this decade. One notable reform is that of SOEs during the late 90s, which is well known to have made China's various industries more competitive. Although we do not model the source of distortion explicitly in our model and rather treat markups (and their distribution) as a reflection of distortion, the fact that we observe increases in both λ_1 and γ may be partly due to these reforms. The compression in markup distribution (Table 1 and Figure 1) and the increasing number of manufacturing firms are also consistent with the above-mentioned reforms.

3.4. Comparison with the BEJK Model

As mentioned, there are no pro-competitive effects and the ACR formula is satisfied in the BEJK model. A natural question is whether our model fits the data better than the BEJK model, at least for some aspects of data patterns. This question is important because if the BEJK model dominates our model in almost all aspects of data patterns, then one may be less interested in our welfare analysis. To examine this, we conduct two sets of comparisons. The first set is to fit the BEJK model by SMM to the above-mentioned moments that discipline our estimation and then compare with our benchmark result from Table 2. The second set is to add moments that BEJK were concerned with matching and use SMM to estimate both the BEJK and our model. We find that our model fits better than the BEJK model in both sets of comparison. The details of the comparison are given in Appendix A2.

4. Gains from Trade

In this section, we conduct a battery of counter-factual analyses to examine the welfare gains from trade.

4.1. Welfare Analysis: Between 1995 and 2004 and from Autarky

To examine gains from trade, we conduct two counter-factual analyses by fixing all parameter values at the 2004 level and changing only τ. In the first analysis, we simulate welfare and its components when τ is changed to the 1995 level, and we calculate the percentage changes of welfare and its components. In the second analysis, we take τ to a prohibitive value so that the economy goes to autarky.

The results are shown in Table 4. The welfare gains of changing τ from 1995's level to 2004's level are 7.1%, in which allocative efficiency accounts for 20.6% (1.5/7.1) and relative markup effect accounts for −0.7%. Thus, these pro-competitive effects jointly account for 19.9% of the total gains from trade. In fact, both aggregate markups M^{sell} and M^{buy} decrease during this period, which is a natural result under trade liberalization, but the percentage decrease in the consumers' aggregate markup M^{buy} is smaller. Overall, although the relative markup effect is negative, it is relatively small, whereas the combined effect can account for about one fifth of the total gains. The total gains from autarky to 2004's τ are, of course, much larger, at 28%, but the decomposition is similar to the first analysis.

To understand the intuition behind the gains due to allocative efficiency, we follow Holmes et al. (2014) to distinguish markup changes by a cost channel and a price channel. An increase in τ increases marginal costs proportionately in events (2,1), (2,2), and $(\overline{2})$ as defined in Section 2.6. This channel reduces markups when prices are held fixed. An increase in τ increases prices proportionately in events (1,2), (2,2), and $(\overline{2})$ and markups when costs are held fixed. From country 1's viewpoint, Holmes et al. show that

$$\eta^A \equiv \frac{d\ \ln W^A}{d\ \ln \tau} = \left(\eta^A_{cost} + \eta^A_{price}\right)\frac{d\ \ln(\tau w)}{d\ \ln \tau}, \tag{17}$$

where

$$\eta^A_{cost} = s_2\left(\frac{Em_2^{\sigma-1}}{Em^{\sigma-1}} - \frac{Em}{Em_2}\right), \tag{18}$$

$$\eta^A_{price} = -\sigma s_{\Omega^{p\dagger}}\left(1 - \frac{Em}{Em_{\Omega^{p\dagger}}}\right). \tag{19}$$

Here, E is the revenue-weighted *harmonic* expectation operator, and thus Em is actually the consumers' aggregate markup M^{buy} as defined earlier. Also, s denotes an expenditure share, and the subscript 2 denotes the set of imported goods (from country 2), which also represents the cost channel. The subset $\Omega^{p\dagger}$ is the set of goods affected by the price channel.

As the general equilibrium effect term $d \ln(\tau w)/d \ln \tau$ is generally positive, the sign of the allocative efficiency hinges on the signs of η^A_{cost} and η^A_{price}. Both terms are negative under the 1995 and 2004 estimates, implying positive allocative efficiency gains when trade costs reduce. More specifically, the mean markup of imported goods Em_2 is lower than that of all goods (so that $Em/Em_2 > 1 > Em_2^{\sigma-1}/Em^{\sigma-1}$). When prices are held fixed, a trade liberalization reduces costs and increases markups of imported goods, reducing the discrepancy between Em_2 and Em and mitigating distortion. Similarly, a trade liberalization decreases prices and markups through the price channel. Under our estimates, $Em_{\Omega^{p\dagger}}$ is greater than Em, and thus a trade liberalization reduces markup dispersion and distortion through the price channel, too.

The key of the above discussion is that $Em_2 < Em < Em_{\Omega^{p\dagger}}$ under our model and estimates. In the BEJK model, $Em_2 = Em$ because the markup distribution conditional on the source is actually independent of the source. Hence, the cost channel disappears ($\eta^A_{cost} = 0$). As the markup distribution in any country is invariant to trade costs, $\eta^A_{cost} + \eta^A_{price} = 0$, and thus $\eta^A_{price} = 0$. Foreign firms selling to China need to pay trade costs that domestic firms avoid. Other things being equal, this drives up the costs and lowers the markups of foreign firms so that $Em_2 < Em$. However, there is also a selection of productivities that only relatively productive foreign firms can export, and this tends to increase markups of foreign firms. This latter force critically depends on the upper tail of the productivity distribution. Under the Fréchet structure in BEJK, this force is so strong that the two forces exactly counterbalance, entailing $Em_2 = Em$. In contrast, the selection force is not as strong with log-normal productivity draws in our model, and thus $Em_2 < Em$. The intuition for why $Em < Em_{\Omega^{p\dagger}}$ is similar.[29]

Table 4
Counter-factual analysis.

Panel A: Counter-factual from 2004 estimates

	Under 2004 Estimates	τ at 1995	% change	Autarky	% change
τ, trade cost	1.782	2.311		1,000,000	
Welfare					
Total welfare	2.12E+12	1.98E+12	7.1%	1.66E+12	28.0%
W_Prod	2.21E+12	2.09E+12	5.6%	1.83E+12	20.7%
W_A	0.958	0.944	1.5%	0.904	6.0%
W_R	1.000	1.000	-0.1%	1.000	0.0%
Contribution to total welfare					
W_A and W_R			19.9%		21.5%
W_A			20.6%		21.6%

Panel B: Counter-factual from autarky

	Autarky	10% import share	% change from autarky	20% import share	% change from 10% import share
τ, trade cost	1,000,000	2.424		1.916	
Welfare					
Total welfare	1.66E+12	1.96E+12	18.4%	2.07E+12	5.6%
W_Prod	1.83E+12	2.08E+12	13.6%	2.17E+12	4.2%
W_A	0.904	0.942	4.2%	0.954	1.3%
W_R	1.000	1.000	0.0%	1.000	0.0%
Contribution to total welfare					
W_A and W_R			23.1%		23.5%
W_A			22.9%		24.2%

Notes: In Panel A, all the analysis is done under 2004 estimates, and only the trade cost (τ) changes. The reported percentage changes in this panel are under the changes from the corresponding τ to 2004's τ. Panel B reports results when τ is changed from an inhibitive level (autarky) to the level that entails 10%, and then from 10% to 20%, with other parameters fixed at the 2004 estimates.

[29] Observe that the cost and price channels have two sets in common; the only difference is that the cost channel has the case (2,1) whereas the price channel has the case (1,2). Again, the fact that domestic firms need not pay trade costs compared with foreign firms implies that the average markup in the case (1,2) is high compared with that in case (2,1) and the overall average, and thus $Em < Em_{\Omega^{p\dagger}}$. The Frechét structure in BEJK implies that the selection force introduces sufficiently productive foreign firms, and thus depresses the markups in case (1,2) so that $Em = Em_{\Omega^{p\dagger}}$. But this is not the case under log-normal productivity draws.

Next, we examine whether the result of "diminishing returns in openness" in EMX holds here. The following table summarizes the welfare gains reported in their study, as well as the breakdown in Ricardian gains and allocative efficiency. There is an obvious "diminishing returns" in allocative efficiency, as the opening up from autarky to 10% import share improves welfare by 1.2%, whereas further opening up from 10% to 20% improves welfare by only 0.3%. But such a diminishing-returns pattern does not show up in the Ricardian component. As a result, the contribution of allocative efficiency diminishes rapidly from $1.2/3.1 \approx 38\%$ to $0.3/2.8 \approx 10.7\%$.

Import share	%Δ in EMX			Importance of W^A
	Total welfare	Ricardian	W^A	
0–10%	3.1	1.9	1.2	38.7%
10%–20%	2.8	2.5	0.3	10.7%

Panel B of Table 4 reports the result from a similar exercise. Note that EMX's pro-competitive effect includes only allocative efficiency but not the relative markup effect, as their formulation focuses on symmetric countries. To compare, we ignore the relative markup effect. A similar diminishing returns pattern in allocative efficiency is obvious, dropping from 4.2% to 1.3%. But, unlike in EMX, we also see sharp diminishing returns in our counter-factuals for total welfare and the Ricardian component. As a result, we do not see a diminishing contribution in allocative efficiency. Indeed, the contribution stays around 23%, which is quite close to the results reported in Panel A.

Looking at both panels together, the contribution of pro-competitive effects ranges from 19.9% to 23.5%, and the contribution of allocative efficiency ranges from 20.6% to 24.2%. Despite the differences in model structures, our estimates turn out to be in the ballpark of EMX's estimates, which range from 11% to 38%.

4.2. Comparison with the ACR Formula

In this subsection, we compare the welfare gains in this model with the ACR formula in two ways. First, we compare with the local ACR formula for small changes in trade cost. Second, as trade elasticity is a variable, the global ACR formula does not apply, but one can integrate the local formula to examine the gains from 1995's τ to 2004's τ in a similar fashion to Panel A of Table 4.

For the first comparison, recall from Section 2.6 that for the case of $\sigma = 1$ (Cobb-Douglas), the ACR formula captures the gains in productive efficiency for small changes in trade costs, but not the total gains from trade. For general $\sigma > 1$, analytical results on the comparison with the ACR formula are not available, and here we provide a quantitative analysis based on the estimated models at 1995 and 2004. For this exercise, we investigate the effect of a small reduction h in the logarithm of trade cost so that $\ln\tau' = \ln(\tau) - h$. The results are reported in Table 5. Here, the welfare gains are expressed in terms of elasticity to trade cost, i.e., $d\ln(W)/d\ln\tau$, where W can be W^{Total}, W^{Prod}, $W^A \times W^R$, or W^{ACR}. As discussed in Section 2.6, the trade elasticity used in evaluating $d\ln W^{ACR}$ is ε_1^2.[30]

The local trade elasticities in our estimated model in 1995 and 2004 are −2.85 and −3.30, which are lower than standard estimates in the literature but surprisingly close to the estimates of trade elasticity under the BEJK model in Simonovska and Waugh (2014), which range from −2.74 to −3.32.[31] As explained by Simonovska and Waugh (2014), different micro mechanisms imply different distributions of price gaps of goods *between countries*. For example, price gaps are on average smaller in the BEJK model than those in the Eaton-Kortum model. To match the same mean price gap from the data, the productivity dispersion needs to be larger in the BEJK model than the Eaton-Kortum one. This implies a smaller tail index of the Fréchet distribution and hence smaller trade elasticity in the BEJK model. The degrees of dispersion in productivity, price gaps, and markups are all positively related. Also, the differences in the moments of markups between exporters and non-exporters may play a similar role to the price gaps in Simonovska and Waugh (2014). Thus, to fit the moments of markups, firm productivities in our model also need to be sufficiently dispersed; it turns out the local trade elasticities in this model are close to their estimates.

In 1995, the welfare elasticity of trade cost is 0.249, meaning a 1% reduction in trade cost τ induces a 0.249% increase in real income; 23.2% of this elasticity is from pro-competitive effects. The ACR formula entails a welfare elasticity of 0.190, which is quite close to the elasticity of productive efficiency 0.191. As a result, the totals gains from trade are larger than the gains predicted by the ACR formula by 31.3%; most of these extra gains are from pro-competitive effects. In the 2004 model, the contribution of pro-competitive effects and the additional gains over the ACR formula are smaller at 20.3% and 13.0%, respectively. Note that the contributions of pro-competitive effects in Table 5 are still relatively similar in magnitude to those reported in Table 4.

Note that the total gains and their components (as point elasticities) are all larger in 2004 than in 1995. To understand why this is the case, we focus on the allocative efficiency first. We compute η^A_{cost} and η^A_{price} according to (18) and (19), and find that both η^A_{cost} and η^A_{price} are negative, implying positive allocative efficiency gains from a decrease in τ. The $\eta^A_{cost} + \eta^A_{price} \equiv \dfrac{d\ \ln W^A}{d\ \ln(\tau w)}$ term is −0.049 and −0.042 in 1995 and 2004, respectively, but the general equilibrium term $d\ln(\tau w)/d\ln\tau$ is 1.18 and 1.73, respectively. It is then clear that the allocative efficiency gains are larger in 2004 mainly because of the general equilibrium term. As w is 10.25 in 1995 (compared with 5.18 in 2004), a percentage decrease in τ in 1995 can only induce a smaller percentage decrease in τw because w is large. Apart from the general equilibrium term, $\eta^A_{cost} + \eta^A_{price}$ is smaller in 2004, which is in line with the facts that the productivity distribution in China is less skewed in 2004 and that the difference in average productivity between China and the ROW has become smaller.

As the productive efficiency gains in our model approximate (roughly) the ACR gains, consider the ACR gains under symmetric countries for clearer intuition. In this situation, the ACR formula becomes $d\ln W^{ACR}/d\ln\tau = -(1-v)$. The smaller the trade cost τ, the smaller the domestic expenditure share v and the larger the ACR gains. Hence, the productive efficiency gains is larger in 2004 mainly because the trade cost estimate τ is significantly lower in 2004 than in 1995.

For the second comparison, we integrate the local ACR formula to compute what their formula would predict for a change in trade cost between 1995 and 2004's levels. That is, we compute $W_1^{ACR,\,2004}/W_1^{ACR,\,1995}$

$$\ln\frac{W_1^{ACR,2004}}{W_1^{ACR,1995}} = \int_{\tau_{1995}}^{\tau_{2004}} d\ \ln W_1^{ACR}(\tau) = \int_{\tau_{1995}}^{\tau_{2004}} d\left(\frac{\ln v_{11}(\tau)}{\epsilon(\tau)}\right). \tag{20}$$

We relegate the calculation details to Appendix A3. The result is that the gains from trade according to the ACR formula are 5.9%. The total gains from trade 7.1% (Table 4) are 20.3% higher than the ACR formula.

In the benchmark exercise, the overall gains at 7.1% seems a relatively large number compared with those found in the literature. The sources of the larger gains compared with the literature are three-fold. The first source is the large reduction in trade cost, which is essentially inferred by the large increases in trade flows and the fraction of exporter during 1995–2004. This means wider bounds of integration in (20). If the trade elasticity were a constant, (20) would reduce to $\hat{W} = \hat{v}^{1/\epsilon}$, and the large change in trade cost would imply a low value of $\hat{v}$. Second,

[30] Here, we set we set $h = 0.0001$, and thus τ' is about 0.1% off τ. To reduce the secant error in calculating trade elasticity, we use two-point formula: $f'(x) = [f(x+h) - f(x-h)]/2h$, where $x = \ln(\tau_{21})$ and $f = \ln((1 - v_{11})/v_{11})$. Note that when calculating the trade elasticity ε_1^2, wages are taken as fixed, as in ACR.

[31] See Tables 4 and 7 in their paper.

Table 5
Welfare analysis local to the estimated model and comparison with the ACR formula.

	Total welfare gains	Gains in productive efficieny	Pro-competitive gains	Contribution of pro-competitive effects	Trade elasticity	Gains by the ACR formula	Additional gains over the ACR formula
1995	0.249	0.191	0.058	23.2%	−2.85	0.190	31.3%
2004	0.362	0.289	0.073	20.3%	−3.30	0.320	13.0%

Notes: All the welfare gains here are calculated in terms of welfare elasticity to trade cost, i.e., dln(W)/dln(τ), where W could be total welfare or its components, or the one according to the ACR formula. For both the 1995 and 2004 models, we calculate the welfare gains and its components from estimated tau to the case where ln(tau')=ln(tau)-h, where h = 0.0001. To reduce secant error in calculating trade elasticity, we use two-point formula: f'(x) = (f(x+h)-f(x-h))/2h, and here x=ln(τ21) and f = ln((1-v11)/v11). As in ACR, the trade elasticity calculated here is partial, i.e., wages are fixed at the initial equilibrium.

even though trade elasticity is a variable in our model, it ranges from −2.28 to −3.58 for $\tau \in [\tau_{2004}, \tau_{1995}]$. Thus, the $W_1^{\mathrm{ACR},\ 2004}/W_1^{\mathrm{ACR},\ 1995}$ computed by (20) is bounded below by the computed number where $\epsilon(\tau)$ is forced to be a constant −3.58. If we force $\epsilon(\tau) = -5$, as is standard in the literature, then $W_1^{ACR,\ 2004}/W_1^{ACR,\ 1995} = 1.033$.[32] Thus, the second source of the larger gains is this gap between 3.3% and 5.9% that is created by the difference in trade elasticity. Then, the third source of the larger gains is the gap between 5.9% and 7.1% that is mainly due to pro-competitive effects.[33]

As explained in Section 2.6, the main reason why Proposition 1 ($d \ln W_j^{Prod} = d \ln W_j^{ACR}$ when $\sigma = 1$) need not hold under $\sigma > 1$ is that changes in trade cost τ may change the expenditure shares across goods and hence across different cases in the table in that subsection. Recall that when a good experiences a change in markup due to a change in trade cost, there is no change in trade flow (or, more precisely, imported value). These are actually cases (1,2) and (2,1) in that table. Under the 2004 estimates, the fraction of goods and expenditure share in these two cases combined are 0.48 and 0.46, respectively. When trade cost τ is changed to the 1995 level, these numbers become 0.52 and 0.49, respectively. Thus, there are significant portions of goods and expenditure where changes in markups are not associated with import. Moreover, the changes in these two magnitudes are slight, corroborating the intuition highlighted in Section 2.6 that most of the extra gains compared with the ACR formula come from pro-competitive effects.

4.3. Symmetric Countries

For the purposes of investigating the role played by the asymmetry between China and the rest of the world, especially in terms of the differences in relative wage and productivity, we also estimate a symmetric-country case. The assumption of symmetric countries is often made in the literature because it allows greater tractability and requires less data. Nevertheless, ignoring cross-country differences may obscure important gains from trade. We demonstrate this point here.

The estimation results are shown in Table 6 and the counter-factual results in Table 7. The changes in trade cost τ, measure of goods γ and number of entrants per product λ between 1995 and 2004 are all in the same direction as in the benchmark case. Note that the estimated λ is similar to a weighted average of estimated λ_1 and λ_2, with the ROW weighted more heavily, since the ROW is much larger than China. Also, observe that the fit of moments becomes significantly worse. This is because there are fewer parameters in the symmetric-country estimation, reflecting the fact that the symmetric-country estimation obscures the large discrepancy in entry and productivity distribution seen in Table 2. It may also be partly because the symmetric-country model fails to reflect the general equilibrium effect in the adjustment of relative wages, which change from 10.25 to 5.18 (See Table 2), meaning that Chinese wages relative to the ROW almost double in this decade.

Table 6
SMM results (symmetric countries).

	1995		2004	
Predetermined				
w Relative wages (the ROW to China)	1.0		1.0	
R1 China's manufacturing sales ($b)	918291		2343328	
R2 ROW's manufacturing sales ($b)	918291		2343328	
Moments	Data	Model	Data	Model
Import share	0.130	0.049	0.222	0.117
Export share	0.153	0.047	0.249	0.114
Relative number of firms	0.210	0.205	0.596	0.627
Fraction of exporters	0.044	0.058	0.105	0.138
Mean cost share for exporters	0.845	0.721	0.801	0.717
Std of cost share for exporters	0.135	0.151	0.142	0.151
p50 markup for exporters	1.196	1.398	1.168	1.394
p95 markup for exporters	2.199	2.255	2.183	2.239
p99 markup for exporters	3.299	3.604	3.364	3.018
Mean cost share for non-exporters	0.789	0.758	0.829	0.772
std of cost share for non-exporters	0.147	0.171	0.139	0.161
p50 markup for non-exporters	1.266	1.286	1.213	1.262
p95 markup for non-exporters	2.475	2.372	2.400	2.168
p99 markup for non-exporters	3.537	3.377	3.523	2.794
Parameter values	Estimate	s.e.	Estimate	s.e.
τ, trade cost	2.376	0.006	1.876	0.003
$\gamma/\overline{N}$, measure of goods relative to $\overline{N}$	0.174	0.001	0.573	0.002
λ, Poisson parameter	4.217	0.033	4.893	0.069
η, std. of log productivity	0.438	0.004	0.495	0.003
σ, elasticity of substitution	1.384	0.003	1.353	0.019

Notes: All the units, if any, are in billions USD, current price. For the detailed definition of moments, see Table 2.

For counter-factual results, first note that the relative markup effect does not show up in Table 7 because this term drops out under symmetric countries. Note that the overall welfare gains become much smaller than the benchmark case (e.g. 2.4% versus 7.1%). Both components also become much smaller. However, the contribution of the pro-competitive effect is still quite close to the benchmark case between 1995 and 2004. In comparison with autarky, the contribution of the pro-competitive effect increases to 29.4%. As the distributions of the number of entrants and productivity draws become the same between the two countries, the Ricardian gains are reduced because active firms' productivity differences between two countries are now reduced. Moreover, not only do the distribution of markups become similar, but the dispersion of markups also becomes smaller. In fact, looking at autarky, we see that the allocative efficiency is much larger in the symmetric-country case than in the benchmark case (0.938 versus 0.904). As the allocative efficiency is larger to start with, it is not surprising that the gains in allocative efficiency are smaller (0.5% versus 1.5% and 2.0% versus 6.0%). The same rationale explains why we see a pronounced diminishing-returns (dropping from 30.8% to 12.4%) pattern in Panel B that is absent in the asymmetric-country case.

[32] Under the 2004 parameters, domestic consumption share drops from 0.8855 to 0.7516 when trade cost τ falls from the 1995 level to the 2004 level. See Costinot and Rodríguez-Clare (2014) for a discussion on the standard estimates of trade elasticities.

[33] Another way to understand the larger gains is to observe that the ACR formula becomes $d \ln W = (1-\lambda) d \ln \tau$ under the case of symmetric countries and no distortion. In this case, trade elasticity does not matter. This formula does not hold in our model because (1) the countries are asymmetric so that there are changes in terms of trade (and hence trade elasticity matters) and (2) there are distortions due to variable markups. Even though $d \ln W^{Prod}$ and $d \ln W^{ACR}$ are the same under $\sigma = 1$ and close under $\sigma > 1$, $d \ln W^{ACR}$ would be higher in our model since our model implies a lower trade elasticity. Thus, reasons (1) and (2) above actually correspond to the second and third sources of larger gains discussed here.

Table 7
Counter-factual analysis (symmetric countries).

Panel A: Counter-factual from 2004 estimates

	Under 2004 Estimates	τ at 1995	% change	Autarky	% change
τ, trade cost	1.876	2.376		1,000,000	
Welfare					
Total welfare	3.50E+16	3.42E+16	2.4%	3.28E+16	6.7%
W_Prod	3.66E+16	3.59E+16	1.9%	3.50E+16	4.7%
W_A	0.957	0.952	0.5%	0.938	2.0%
Contribution to total welfare					
W_A			19.4%		29.4%

Panel B: Counter-factual from autarky

	Autarky	10% import share	% change from autarky	20% import share	% change from 10% import share
τ, trade cost	1,000,000	1.970		1.550	
Welfare					
Total welfare	3.28E+16	3.48E+16	6.1%	3.62E+16	4.0%
W_Prod	3.50E+16	3.64E+16	4.1%	3.77E+16	3.5%
W_A	0.938	0.956	1.9%	0.960	0.5%
Contribution to total welfare					
W_A			30.8%		12.4%

Notes: Under symmetric countries, W_R = 1. In Panel A, all the analysis is done under 2004 estimates, and only the trade cost (τ) changes. The reported percentage changes in this panel are under the changes from the corresponding τ to 2004's τ. Panel B reports results when τ is changed from an inhibitive level (autarky) to the level that entails 10%, and then from 10% to 20%, with other parameters fixed at the 2004 estimates.

Under symmetric countries, the results in EMX rely on the cross-country productivity differences across different sectors to generate pro-competitive effects. However, our exercise indicates that asymmetries between countries could also be important sources of gains, both in the Ricardian component and the pro-competitive effects. Not finding these gains in the symmetric-country implementation indicates the importance of asymmetric-country quantification, especially when the country of concern is a developing one, such as China. Our approach of using moments from both exporters and non-exporters proves to be instrumental in such an implementation.

4.4. Comparative Statics and Robustness Checks

As the entry parameter λ and the productivity-dispersion parameter η play important roles in our model, we provide the following comparative statics in Appendix A4. In the first set of comparative statics, we study the welfare gains and their components in a closed economy when population L and entry parameter λ are simultaneously doubled and when the productivity-dispersion parameter η is doubled. In both cases, we find that welfare and its components increase except that the allocative efficiency decreases when η doubles. The second set of comparative statics studies gains from trade under different levels of η and λ in a symmetric-country open economy. When η increases, both the total gains from trade and the productive-efficiency component increase, but the allocative-efficiency component remains roughly the same. The total gains from trade and their two components all decrease in λ. See more discussions in Appendix A4.

We conduct four robustness checks. Recall that in the benchmark case, the counter-factual analyses are based on 2004 estimates and change τ back to the 1995 level. The first robustness check is to conduct a counter-factual analysis based on 1995 estimates and change τ to the 2004 level. In our second check, we use an alternative measure of markups to estimate the model and run counter-factuals. That is, by invoking the constant-returns-to-scale assumption, we calculate *raw markups* by taking the ratio of revenue to total costs. There were substantial trade surpluses in China in both 1995 and 2004. They account for 2.25% of China's manufacturing sales in 1995 and 2.63% in 2004. Our third check is to accommodate trade imbalance in the model. Another potential concern on our results is that a substantial fraction of the Chinese trade is intermediate goods and " processing". In the benchmark, processing trade is included in the total import and export when calculating the import and export shares. Our fourth check is based on the export and import figures that exclude "processing trade". The details of these robustness checks are given in Appendix A5.

We find that the total gains from trade between 1995 and 2004 range from 5.0% to 9.2%, and the contribution of pro-competitive effects ranges from 13.1% to 32.1%, and that of allocative efficiency ranges from 15.6% to 30.7%. These indicate that the magnitude of pro-competitive effects remains similar, and the allocative efficiency still accounts for the bulk of gains from trade. See more discussions in Appendix A5.

5. Multiple-Sector Economy

The framework in this paper can be easily extended to a multiple-sector economy, which we do for three reasons. First, the model is more realistically matched to data, taking into account the cross-sector heterogeneity in trade costs, as well as in productivity distribution, entry effort and preference parameters. Second, we conduct similar welfare analyses to gauge the robustness of our previous results for this multiple-sector extension. Third, exploiting the variations in sectoral markups and trade costs, we attempt to answer the question of whether China liberalized the " right" sectors by examining whether there was larger trade liberalization in sectors with higher initial markups in 1995.

5.1. Model and Estimation

5.1.1. Model Modification

There are S sectors, which are indexed by $s = 1, 2, \ldots, S$. The utility function of a representative consumer is

$$U = \Pi_{s=1}^{S} (Q_s)^{\alpha_s},$$

where $\alpha_s \in (0,1)$, $\sum_{s=1}^{S} \alpha_s = 1$, and Q_s is the consumption of the composite good of sector s given by a CES aggregator:

$$Q_s = \left(\int_0^{\overline{\omega}_s} (q_{s,\omega})^{\frac{\sigma_s - 1}{\sigma_s}} d\omega \right)^{\frac{\sigma_s}{\sigma_s - 1}}, \quad \text{for } \sigma_s > 1,$$

where σ_s is the elasticity of substitution of sector s. The aggregate and sectoral price indices are therefore

$$P_j = \Pi_{s=1}^{S} \left(\frac{P_{js}}{\alpha_s} \right)^{\alpha_s}$$

$$P_{js} \equiv \left(\int_0^{\overline{\omega}_s} p_{js\omega}^{1-\sigma_s} d\omega \right)^{\frac{1}{1-\sigma_s}}.$$

The Cobb-Douglas structure implies that $P_{js} Q_{js} = \alpha_s R_j$, and country j's total expenditure of good $s\omega$ is given by

$$E_{js\omega} = \alpha_s R_j \left(\frac{p_{js\omega}}{P_{js}} \right)^{1-\sigma_s} \equiv \alpha_s R_j \phi_{js\omega}, \tag{21}$$

and the total revenue of all firms at i in sector s is

$$R_{s,i} = \int_{\{s\omega:\chi_1^*(s\omega)=i\}} \alpha_s R_1 \phi_{1s\omega} d\omega + \int_{\{s\omega:\chi_2^*(s\omega)=i\}} \alpha_s R_2 \phi_{2s\omega} d\omega.$$

For each sector s, all the parameters in the one-sector economy now become sector-specific. That is, for each sector s there is a τ_s and a γ_s, and for sector s and country i, there is a set $\{\lambda_{is}, \mu_{is}, \eta_{is}\}$. For each sector, pricing and markups follow the previous formulation.

5.1.2. Wages and General Equilibrium

For trade flows, observe that country j's total import from country i is

$$R_{j,i} = \sum_{s=1}^{S} \int_{\left\{s\omega:\chi_j^*(s\omega)=i\right\}} E_{js\omega} d\omega = R_j \phi_{j,i}$$

where $\chi_j^*(s\omega) \in \{1,2\}$ denotes the source country for any particular good $s\omega$ at destination j and $\phi_{j,\,i}$ is the total spending share of j on i's goods:

$$\phi_{j,i} = \sum_{s=1}^{S} \alpha_s \int_{\left\{s\omega:\chi_j^*(s\omega)=i\right\}} \phi_{js\omega} d\omega. \tag{22}$$

The balanced trade condition $R_2\phi_{2,\,1} = R_1\phi_{1,\,2}$ holds in equilibrium. The algorithm for calculating an equilibrium in a multiple-sector economy is similar to the one-sector case. From (23) and (24), we can derive the following formula for M_1^{sell} and M_2^{sell}:

$$M_1^{sell}(w) = \left[\sum_{s=1}^{S} \alpha_s \left(\int_{\{\omega:\chi_{s1}^*(\omega)=1\}} m_{1s\omega}^{-1} \phi_{1s\omega} d\omega + \int_{\{\omega:\chi_{s2}^*(\omega)=1\}} m_{2s\omega}^{-1} \phi_{2s\omega} \frac{\phi_{1,2}}{\phi_{2,1}} d\omega\right)\right]^{-1}$$

$$M_2^{sell}(w) = \left[\sum_{s=1}^{S} \alpha_s \left(\int_{\{\omega:\chi_{s1}^*(\omega)=2\}} m_{1s\omega}^{-1} \phi_{1s\omega} \frac{\phi_{2,1}}{\phi_{1,2}}(w) d\omega + \int_{\{\omega:\chi_{s2}^*(\omega)=2\}} m_{2s\omega}^{-1} \phi_{2s\omega} d\omega\right)\right]^{-1},$$

in which $\phi_{j,\,i}$ is the total spending share of j on i's goods given in (22). Then, we still calculate $R_1(w) = M_1^{sell}(w)L_1$, $R_2(w) = M_2^{sell}(w)wL_2$, and

$$\frac{R_1(w)}{R_2(w)} = \frac{M_1^{sell}(w)}{M_2^{sell}(w)} \frac{L_1}{wL_2}$$

to pin down the relative wage w.

5.1.3. Welfare

The welfare of country i is decomposed in the same way as before:

$$W_i^{Total} = w_i L_i \times \frac{1}{A_i} \times \frac{M_i^{sell}}{M_i^{buy}} \times \frac{A_i \times M_i^{buy}}{P_i},$$

where

$$A_i = \Pi_{s=1}^{S} \left(\frac{A_{is}}{\alpha_s}\right)^{\alpha_s}, \qquad P_i = \Pi_{s=1}^{S} \left(\frac{P_{is}}{\alpha_s}\right)^{\alpha_s},$$

$$M_i^{buy} = \left(\sum_{s=1}^{S} \alpha_s \left(M_{is}^{buy}\right)^{-1}\right)^{-1}, \qquad M_i^{sell} = \frac{R_i}{w_i L_i} = \left(\sum_{s=1}^{S} \frac{R_{s,i}}{R_i} \left(M_{is}^{sell}\right)^{-1}\right)^{-1}, \tag{23}$$

and A_{is}, P_{is}, and M_{is}^{buy} are defined in the same way as before, and M_{is}^{sell} is

$$M_{is}^{sell} = \left(\int_{\{\omega:\chi_{s1}^*(\omega)=i\}} m_{1s\omega}^{-1} \frac{\alpha_s R_1 \phi_{1s\omega}}{R_{s,i}} d\omega + \int_{\{\omega:\chi_{s2}^*(\omega)=i\}} m_{2s\omega}^{-1} \frac{\alpha_s R_2 \phi_{2s\omega}}{R_{s,i}} d\omega\right)^{-1}. \tag{24}$$

The sectoral welfare cannot be further decomposed into the three components as in the one-sector model. This breaks down because there is no simple analogue of $R_i = w_i L_i \times M_i^{sell}$ at the sectoral level. Indeed, $w_i L_i = \sum_s \frac{R_{is}}{M_{is}^{sell}}$.

5.1.4. Quantifying the Model

To quantify the model, we focus on 29 two-digit manufacturing sectors in the Chinese Industrial Classifications (CIC).[34] We first calibrate $\{\alpha_s\}_{s=1}^{S}$. Recall that $P_{1s}Q_{1s} = \alpha_s R_1$. We proxy α_s by the sectoral share of aggregate revenue of all firms in the Economic Census in each data year. Given $\{R_1, R_2, w\}$ and $\{\alpha_s\}_{s=1}^{S}$, all of the moments used in Section 3.3 for each sector can be computed using a similar procedure and with (21). Thus, we can estimate the parameters sector by sector; this largely simplifies the estimation.

The parameter estimates are shown in Tables 8A and 8B. In both tables, we also report the (unweighted) mean, standard deviation, maximum, and minimum of the estimates and percentage changes across sectors. There are substantial variations across industries in their moments. The model performs well in accommodating these variations with corresponding variations in the estimates. The changes in the unweighted means of parameters between 1995 and 2004 are consistent with the pattern observed in the one-sector case for τ, γ, λ_1, μ_1, and σ. In particular, all estimated trade costs decrease except for Tobacco Processing and Food Processing.[35] Also observe that the means of σ_s are 1.56 and 1.53, which are close to our benchmark in the one-sector economy, and σ_s in most industries (24 out of 29) are within one standard deviation from the means in both years.

To further evaluate the performance of the multi-sector model, we examine variation in the cross-sector entry parameter in China (λ_1) and that of standard indicators for entry frictions such as the SOE's share of capital in the industry, the R&D expenditure (as a proxy to sunk cost of entry), capital intensity (measured by capital labor ratio), and capital distortion (measured by the standard deviation of $y_{s\omega}/k_{s\omega}$, where $y_{s\omega}$ and $k_{s\omega}$ are value added and capital for an active firm $s\omega$ in sector s).[36] As these four variables capture different aspects of entry friction in the Chinese economy, we regress λ_1 on these entry-friction indicators jointly to examine the conditional correlation of each indicator with λ_1. The result is given in Table 9. We pool the two years of estimates, and thus the number of observations is 58. We report both cases where year fixed effects are controlled and where both year and sector fixed effects are controlled. All entry-friction indicators but the SOE share appear to be significantly correlated with λ_1, and carry the expected signs. That is, the larger the entry friction indicator, the smaller the λ_1. This lends confidence that λ_1 does reflect cross-sector variation in competition environment in China. The SOE share does not carry the expected sign but is insignificant.

From the one-sector analysis, welfare is closely (and negatively) related to the dispersion of markups. Even though there is no

[34] We include all 2-digit CIC manufacturing sectors except Sector 43 because we do not have the necessary data to calculate markups for this industry.

[35] This is mainly because the import and export shares decrease in these two sectors.

[36] We use this measure for capital distortion following the literature of factor misallocation *à la* Hsieh and Klenow (2009). The R&D expenditure is not available in the Economic Census data, and we take it from the Annual Surveys of Industrial Firms conducted by the National Bureau of Statistics of China during 1998–2007. The earliest year that the R&D expenditure data is available is 2001: we use this year's data to proxy 1995. The remaining three indicators are computed from the Economic Census data.

Table 8A
Estimation result in multi-sector model (Part A).

		α (pre-determined)		σ			γ			τ			Tariff			Non-tariff t		
cic2d	Industry definition	1995	2004	1995	2004	% change	1995	2004	% change	1995	2004	% change	1995	2004	% change	1995	2004	% change
13	Food processing	0.063	0.049	1.693	1.814	7.2	0.016	0.033	111.5	1.974	2.289	15.9	25.6	16.6	−35.2	1.57	1.96	24.9
14	Food manufacturing	0.020	0.017	1.577	1.366	−13.4	0.007	0.011	63.5	3.994	2.534	−36.6	17.5	9.9	−43.6	3.40	2.31	−32.1
15	Beverage manufacturing	0.024	0.014	1.453	1.416	−2.6	0.005	0.009	84.8	5.374	3.992	−25.7	25.3	7.7	−69.8	4.29	3.71	−13.5
16	Tobacco processing	0.022	0.014	1.466	1.374	−6.3	0.000	0.000	−10.8	5.365	5.394	0.5	37.9	9.8	−74.3	3.89	4.91	26.3
17	Textile industry	0.089	0.059	1.688	1.443	−14.5	0.012	0.036	199.9	1.797	1.601	−10.9	19.7	7.6	−61.4	1.50	1.49	−0.9
18	Garments & other fiber products	0.028	0.023	1.566	1.482	−5.4	0.007	0.019	166.8	3.098	2.869	−7.4	10.8	9.2	−15.0	2.80	2.63	−6.0
19	Leather, furs, down & related products	0.019	0.016	1.541	1.416	−8.1	0.004	0.010	150.0	1.852	1.694	−8.5	9.9	5.5	−44.3	1.69	1.61	−4.7
20	Timber processing, bamboo, cane, palm fiber & straw products	0.009	0.011	1.453	1.477	1.6	0.008	0.019	126.2	2.098	1.643	−21.7	7.8	2.6	−67.1	1.95	1.60	−17.7
21	Furniture manufacturing	0.004	0.008	1.446	1.283	−11.3	0.004	0.010	146.6	2.535	1.928	−23.9	8.3	1.0	−88.0	2.34	1.91	−18.4
22	Papermaking & paper products	0.021	0.020	1.780	1.512	−15.1	0.006	0.020	237.7	2.545	2.034	−20.1	23.7	4.0	−83.0	2.06	1.95	−5.0
23	Printing industry	0.009	0.009	1.417	1.293	−8.7	0.008	0.020	144.3	2.603	2.373	−8.9	5.3	0.9	−83.5	2.47	2.35	−4.9
24	Cultural, educational & sports goods	0.007	0.007	1.361	1.395	2.5	0.002	0.006	275.4	2.051	1.794	−12.5	4.1	1.5	−64.3	1.97	1.77	−10.3
25	Petroleum processing & coking	0.046	0.050	1.338	1.744	30.3	0.001	0.004	300.1	1.740	1.513	−13.0	8.6	5.0	−42.2	1.60	1.44	−10.0
26	Raw chemical materials & chemical products	0.077	0.072	1.649	1.489	−9.7	0.012	0.038	218.3	2.108	1.786	−15.3	14.6	7.2	−51.0	1.84	1.67	−9.4
27	Medical & pharmaceutical products	0.020	0.017	1.584	1.277	−19.4	0.003	0.004	42.0	3.553	2.916	−17.9	6.9	3.8	−44.9	3.32	2.81	−15.5
28	Chemical fiber	0.017	0.010	2.414	2.412	−0.1	0.000	0.002	242.6	2.920	2.186	−25.1	22.0	4.9	−77.7	2.39	2.08	−12.9
29	Rubber products	0.013	0.010	1.284	1.523	18.6	0.002	0.007	238.6	1.771	1.707	−3.6	20.2	11.0	−45.6	1.47	1.54	4.4
30	Plastic products	0.023	0.027	1.795	1.494	−16.7	0.008	0.030	276.2	1.714	1.713	−0.1	13.9	5.4	−61.0	1.50	1.62	8.0
31	Nonmetal mineral products	0.060	0.050	1.595	1.429	−10.4	0.024	0.063	160.3	4.588	2.388	−48.0	12.8	5.9	−54.0	4.07	2.25	−44.6
32	Smelting & pressing of ferrous metals	0.083	0.092	1.522	2.137	40.4	0.003	0.009	180.0	2.028	2.326	14.7	10.9	4.9	−55.2	1.83	2.22	21.3
33	Smelting & pressing of nonferrous metals	0.027	0.031	1.999	1.735	−13.2	0.002	0.008	337.8	2.079	1.690	−18.7	7.7	3.9	−49.4	1.93	1.63	−15.8
34	Metal products	0.030	0.032	1.495	1.466	−1.9	0.011	0.031	182.0	1.947	1.844	−5.3	13.2	4.0	−69.9	1.72	1.77	3.1
35	Ordinary machinery	0.049	0.052	1.712	1.713	0.0	0.015	0.065	321.5	2.601	1.655	−36.4	17.5	5.1	−71.0	2.21	1.58	−28.8
36	Special purpose equipment	0.041	0.030	1.548	1.430	−7.6	0.008	0.033	298.9	2.491	1.668	−33.0	16.6	5.3	−68.2	2.14	1.58	−25.8
37	Transport equipment	0.072	0.076	1.391	1.478	6.2	0.011	0.029	171.1	2.358	2.023	−14.2	43.5	12.7	−70.8	1.64	1.79	9.2
39	Electric equipment & machinery	0.055	0.061	1.569	1.656	5.5	0.009	0.030	253.1	1.718	1.598	−7.0	11.3	3.0	−73.0	1.54	1.55	0.4
40	Electronic & telecommunications equipment	0.052	0.121	1.390	1.362	−2.0	0.004	0.013	225.0	2.087	1.628	−22.0	13.5	1.3	−90.5	1.84	1.61	−12.6
41	Instruments, meters, cultural & office equipment	0.009	0.013	1.279	1.416	10.7	0.003	0.009	206.5	2.035	1.677	−17.6	15.7	4.3	−72.6	1.76	1.61	−8.6
42	Other manufacturing	0.011	0.009	1.223	1.459	19.3	0.004	0.012	178.7	2.303	1.811	−21.4	8.8	2.8	−67.8	2.12	1.76	−16.8
Mean		0.034	0.034	1.56	1.53	−0.83	0.01	0.02	190.65	2.60	2.15	−15.30	15.64	5.74	−61.87	2.24	2.02	−7.47
Standard deviation		0.025	0.029	0.24	0.25	14.08	0.01	0.02	84.26	1.03	0.82	14.11	9.12	3.66	17.25	0.80	0.74	16.09
Max		0.089	0.121	2.41	2.41	40.41	0.02	0.06	337.82	5.37	5.39	15.91	43.46	16.57	−14.99	4.29	4.91	26.33
Min		0.004	0.007	1.22	1.28	−19.41	0.00	0.00	−10.84	1.71	1.51	−47.96	4.08	0.87	−90.54	1.47	1.44	−44.55

decomposition of welfare into three components at the sectoral level, it is worthwhile examining how markup dispersion is correlated with trade cost. In particular, if the reduction in sectoral trade cost is positively correlated with the reduction of sectoral markup dispersion, then trade liberalization is a possible cause for improving allocative efficiency. We examine the correlation between the percentage change in estimated trade cost τ_s with changes in the following three measures of markup dispersion: standard deviation, coefficient of variation, and the Theil index of markups. The respective correlations are 0.264, 0.242, and 0.241. The correlations are indeed positive. While the correlations are not particularly strong, notice that the identification of τ_s are mainly driven by sectoral trade moments rather than the markup moments.[37]

[37] The general message here is similar to Lu and Yu (2015), who find that tariff reduction leads to reduced markup dispersion. Their exercise is different as they look for causality, they have controls, and they examine manufacturing industries at a finer level.

Table 8B
Estimation result in multi-sector model (Part B).

cic2d	Industry definition	λ_1 1995	λ_1 2004	λ_1 % change	λ_2 1995	λ_2 2004	λ_2 % change	μ_1 1995	μ_1 2004	μ_1 % change	η_1 1995	η_1 2004	η_1 % change	η_2 1995	η_2 2004	η_2 % change
13	Food processing	3.141	3.450	9.8	6.276	7.372	17.5	−2.367	−1.644	30.5	0.348	0.458	31.4	0.308	0.407	32.2
14	Food manufacturing	3.413	3.416	0.1	4.082	4.620	13.2	−1.808	−1.499	17.1	0.142	0.347	143.6	0.356	0.402	12.8
15	Beverage manufacturing	3.583	3.679	2.7	5.802	4.366	−24.7	−0.961	−1.254	−30.5	0.093	0.260	180.3	0.070	0.189	171.5
16	T obacco processing	2.504	3.467	38.5	6.143	6.243	1.6	−1.902	−1.339	29.6	0.355	0.454	27.9	0.217	0.196	−9.3
17	Textile industry	3.009	3.248	8.0	5.205	5.323	2.3	−2.280	−1.762	22.7	0.282	0.335	19.1	0.342	0.138	−59.5
18	Garments & other fiber products	3.633	3.510	−3.4	4.820	6.704	39.1	−2.168	−0.920	57.6	0.510	0.437	−14.3	0.091	0.450	393.4
19	Leather, furs, down & related products	2.829	3.503	23.8	5.178	4.548	−12.2	−2.179	−1.724	20.9	0.316	0.395	25.0	0.387	0.315	−18.6
20	Timber processing, bamboo, cane, palm fiber & straw products	2.846	3.028	6.4	5.246	5.543	5.7	−2.425	−1.751	27.8	0.427	0.342	−19.9	0.379	0.231	−39.2
21	Furniture manufacturing	2.793	3.086	10.5	5.130	5.111	−0.4	−1.881	−1.593	15.3	0.257	0.352	36.7	0.291	0.062	−78.5
22	Papermaking & paper products	3.158	2.761	−12.6	5.533	6.092	10.1	−2.211	−1.765	20.2	0.266	0.388	45.6	0.623	0.420	−32.6
23	Printing industry	2.969	2.760	−7.0	6.198	5.600	−9.7	−2.380	−1.769	25.7	0.359	0.435	21.2	0.103	0.179	73.8
24	Cultural, educational & sports goods	2.616	3.479	33.0	5.246	4.629	−11.7	−2.123	−1.707	19.6	0.370	0.451	21.8	0.267	0.176	−34.1
25	Petroleum processing & coking	2.703	2.958	9.4	5.368	6.417	19.5	−2.286	−1.843	19.4	0.168	0.327	94.4	0.291	0.289	−0.8
26	Raw chemical materials & chemical products	3.045	2.161	−29.0	4.775	6.112	28.0	−2.507	−1.829	27.0	0.347	0.477	37.4	0.453	0.432	−4.6
27	Medical & pharmaceutical products	2.728	2.658	−2.6	6.473	4.998	−22.8	−2.160	−1.611	25.4	0.474	0.514	8.4	0.136	0.398	192.1
28	Chemical fiber	2.444	2.982	22.0	5.474	3.985	−27.2	−2.570	−1.730	32.7	0.403	0.209	−48.0	0.527	0.446	−15.3
29	Rubber products	3.431	2.607	−24.0	5.798	4.493	−22.5	−2.362	−1.713	27.5	0.269	0.329	22.4	0.118	0.124	5.2
30	Plastic products	3.144	3.142	−0.1	4.860	5.113	5.2	−2.360	−1.643	30.4	0.289	0.315	8.7	0.255	0.260	1.8
31	Nonmetal mineral products	2.858	3.035	6.2	4.919	5.539	12.6	−1.633	−1.388	15.0	0.283	0.293	3.6	0.245	0.357	45.9
32	Smelting & pressing of ferrous metals	3.021	3.536	17.0	7.705	4.388	−43.0	−2.306	−1.512	34.5	0.125	0.124	−1.3	0.291	0.485	66.5
33	Smelting & pressing of nonferrous metals	2.742	2.700	−1.5	5.169	5.512	6.6	−2.472	−1.740	29.6	0.356	0.340	−4.5	0.280	0.297	5.9
34	Metal products	3.099	2.884	−6.9	5.164	5.426	5.1	−2.418	−1.807	25.3	0.361	0.415	15.1	0.232	0.113	−51.4
35	Ordinary machinery	2.238	2.738	22.3	4.639	6.954	49.9	−2.403	−1.744	27.4	0.333	0.421	26.4	0.524	0.378	−27.9
36	Special purpose equipment	2.065	2.326	12.6	5.165	5.608	8.6	−2.512	−1.801	28.3	0.338	0.383	13.5	0.627	0.434	−30.8
37	Transport equipment	2.786	2.863	2.8	6.788	6.257	−7.8	−2.356	−1.687	28.4	0.369	0.387	4.9	0.305	0.313	2.8
39	Electric equipment & machinery	2.803	2.961	5.6	5.796	5.601	−3.4	−2.468	−1.694	31.4	0.338	0.399	18.0	0.189	0.405	114.8
40	Electronic & telecommunications equipment	2.414	2.468	2.2	5.545	5.907	6.5	−2.354	−1.738	26.2	0.490	0.570	16.3	0.595	0.455	−23.5
41	Instruments, meters, cultural & office equipment	2.389	2.145	−10.2	4.960	5.930	19.6	−2.441	−1.697	30.5	0.538	0.554	2.9	0.600	0.525	−12.5
42	Other manufacturing	2.740	3.544	29.3	5.052	5.829	15.4	−2.251	−1.766	21.6	0.498	0.502	0.9	0.282	0.192	−31.8
Mean		2.87	3.00	5.69	5.47	5.52	2.79	−2.23	−1.64	24.72	0.33	0.39	25.43	0.32	0.31	22.35
Standard deviation		0.38	0.43	15.56	0.73	0.83	20.05	0.33	0.20	13.21	0.11	0.10	45.17	0.16	0.13	95.11
Max		3.63	3.68	38.48	7.70	7.37	49.91	−0.96	−0.92	57.56	0.54	0.57	180.33	0.63	0.53	393.41
Min		2.06	2.15	−29.02	4.08	3.98	−43.05	−2.57	−1.84	−30.51	0.09	0.12	−48.03	0.07	0.06	−78.54

5.2. Gains from Trade

When examining the welfare analysis in the multi-sector economy, we focus on the two key counter-factuals shown in Table 10. Whereas we changed τ in the one-sector economy, we now change $\{\tau_s\}$ for all sectors s from the 2004 values to the 1995 values (or to inhibitive values). The total gains from trade are 4.5% between 1995 and 2004 and 21.7% from autarky. The contribution of pro-competitive effects here is around 21%, which is quite close to the numbers in Table 4. Similarly, allocative efficiency accounts for almost all of the pro-competitive effects.

Table 9
Entry frictions and **λ_1** estimates.

Dependent variable	Indicators of entry frictions: SOE share	R&D expenditure	Capital intensity	Capital distortion	Fixed effect	R-Squared
λ_1	0.483	−0.132***	−0.160*	−0.000525**	Year	0.258
	(0.301)	(0.0325)	(0.0813)	(0.000251)		
λ_1	0.657	−0.529*	−0.210**	−0.000655*	Year and Sector	0.799
	(0.437)	(0.275)	(0.0750)	(0.000361)		

Notes: As there are two years and 29 manufacturing sectors, the number of observations is 58. Robust standard errors are indicated in parentheses. *** p<0.01, ** p<0.05, *p<0.1. R&D expenditure and the capital intensity (measured by capital labor ratio) are both in logarithms.

5.3. Did China Liberalize the Right Sectors?

In this subsection, we try to answer the question of whether China liberalized the right sectors. We examine the relationship between trade liberalization and sectoral consumers' aggregate markup (M_{1s}^{buy}) under the 1995 model. That is, if a sector has a higher M_{1s}^{buy} in 1995, do we also actually see a larger degree of trade liberalization between 1995 and 2004? The rationale is as follows. Recall from (23) that aggregate markup M_1^{buy} is a harmonic mean of sectoral markups (M_{1s}^{buy}). From both one-sector and multi-sector welfare analysis, we observe that most pro-competitive gains from trade are due to allocative efficiency. As the overall allocative efficiency depends on the dispersion of markups across sectors, if a sector s has higher M_{1s}^{buy} initially, then allocative efficiency will improve more if the government targets its trade liberalization more in these higher markup sectors.

Table 10
Counter-factual analysis in multiple-sector economy.

	Welfare				Contribution to total welfare	
	Total welfare	W_Prod	W_A	W_R	W_A and W_R	W_A
% change from τ at 1995 to τ at 2004	4.5%	3.5%	0.8%	0.1%	21.0%	18.5%
% change from autarky to τ at 2004	21.7%	16.5%	4.6%	−0.1%	20.9%	21.3%

Notes: Similar to Table 4, all the analyses in Panel A are done under 2004 estimates, and only the trade costs change. The reported percentage changes in this panel are under the changes from the corresponding $\boldsymbol{\tau}$ to 2004's.

Table 11
Did China liberalize the right sectors?

Dependent variable	Changes in trade costs between 1995 and 2004				Changes in import tariffs between 1995 and 2004			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sectoral markup at 1995	−1.315**	−1.065	−1.629*	−1.670*	−0.475***	−0.491***	−0.700***	−0.700***
	(0.590)	(0.663)	(0.903)	(0.927)	(0.147)	(0.161)	(0.238)	(0.241)
SOE share		0.244*		−0.709		−0.0202		−0.00429
		(0.127)		(0.752)		(0.0261)		(0.149)
Log wage at 1995			−0.106	−0.102			−0.0903	−0.0905
			(0.201)	(0.202)			(0.0617)	(0.0646)
Log employment at 1995			−0.144	−0.289			−0.0388	−0.0398
			(0.124)	(0.234)			(0.0307)	(0.0538)
Log export at 1995			0.160***	0.174**			−0.0429***	−0.0427**
			(0.0560)	(0.0697)			(0.0142)	(0.0153)
Log import at 1995			−0.0122	0.0559			0.0490***	0.0494**
			(0.0531)	(0.111)			(0.0174)	(0.0235)
R^2	0.076	0.101	0.289	0.334	0.203	0.206	0.463	0.463

Notes: The regression is weighted by sectoral trade volume and sectoral imports when the dependent variable is the change in trade cost and import tariff, respectively. Note that the sample size is small (29), and hence one should use caution when interpreting the significance levels.* Significant at the 10 percent level.** Significant at the 5 percent level. *** Significant at the 1 percent level.

A quick examination is to rank the 29 sectors by their values of M_{1s}^{buy} at 1995 and divide them into two groups – the first being 15 sectors with the smaller values of M_{1s}^{buy} and the second being those with the larger values. The revenue-weighted harmonic mean of the M_{1s}^{buy} are then 1.22 and 1.35, respectively. The revenue-weighted mean of the changes in trade costs τ_s are −0.162 and −0.690, respectively. An alternative measure of trade liberalization is the changes in sectoral import tariffs,[38] which directly relate to the WTO entry but do not account for other factors of trade liberalization. In this case, the corresponding revenue-weighted mean of the changes in import tariff are −8.36 and −13.65 percentage points, respectively. These simple statistics show a tendency where the higher the initial level of sectoral markups, the larger the reduction in trade costs (or import tariffs).

Columns 1 and 5 of Table 11 show similar results by regressing the changes in sectoral trade costs and in sectoral import tariffs on sectoral markups M_{1s}^{buy} at 1995.[39] Note that these descriptive results suffice for our purpose, as we seek only to examine whether China on average liberalized the right sectors, smoothing the dispersion of markups across sectors, even if this happened by chance. In other words, we do not attempt to establish causality. Nevertheless, we also examine conditional correlations by following Trefler (2004) in accounting for factors that may affect the changes in tariffs. Columns 3 and 7 show the results when we add controls for log of wage rates, employment, exports, and imports, all at 1995. The rationale of these controls is that they are highly correlated with various kinds of protectionism.[40] As the share of SOEs is presumably a good indicator of protectionism in China, we also add this as a control (see columns 2, 4, 6, and 8). The above-mentioned tendency still remains.[41]

One often-mentioned merit of trade liberalization (or tariff reduction) is that it is an easier route to reducing domestic protectionism compared with using domestic industrial policies. Before joining the WTO, import tariffs varied greatly in China, but the WTO conditions generally require larger tariff reductions in those industries with higher initial tariffs (see Lu and Yu, 2015). We do not know whether the Chinese government had benevolent motives and sought to enhance welfare; it could simply be a mechanical result of China wanting to enter the WTO. In any case, our structural approach allows a welfare assessment in the context of sectoral reallocation both in terms of improved overall allocative efficiency (Table 10) and the results in this subsection.

6. Conclusion

Using Chinese firm-level data at 1995 and 2004, this paper studies pro-competitive effects of trade quantitatively under head-to-head competition. The benchmark counter-factual shows that total gains from such improved openness during this period is 7.1%. The pro-competitive effects account for 19.9% of the total gains from trade from 1995 to 2004 and 21.5% from autarky to 2004. Allocative efficiency plays a much more important role than the relative markup effect.

For small changes in trade costs in each estimated model at 1995 and 2004, the total gains from trade are larger than the gains predicted by the ACR formula by 31% and 13%, respectively. The total gain from the change in trade cost between 1995's and 2004's levels is 20.3% larger than the ACR formula. These additional gains are mostly from pro-

[38] The tariff data is obtained from World Integrated Trade Solution (WITS), which was developed by the World Bank and incorporates trade data from various sources. In particular, we use TRAINS as it covers more countries and more years. An observation of tariff is an average tariff at the HS 6-digit product level. We use " effectively applied rates" (AHS). As WITS does not report China's import tariffs in 1995, we take averages of the 1994 and 1996 tariffs as proxies. In calculating sectoral import tariffs, we use the mapping of HS 6-digit to CIC 2-digit manufacturing sectors using the concordance table from the National Bureau of Statistics of China. For each sector, we then use imports in the corresponding product or industry from the previous year (1994 and 2003) as weights to calculate average import tariffs.

[39] As sector-level data is grouped data from either firms or products, we weight the regressions by trade volume and imports when the dependent variables are changes in trade cost and import tariffs, respectively.

[40] For a detailed explanation, see Trefler (2004), p. 878.

[41] All the coefficients on sectoral markup at 1995 are significant except in Column 2. As the sample size is small (29), one should use caution when interpreting the significance levels.

competitive effects. This is a result that is absent in models when a firm monopolizes a variety, such as in Arkolakis et al. (2019), EMX, Feenstra and Weinstein (2017), and other monopolistic competitive models. Head-to-head competition is the main reason driving this difference, as explained in Sections 2.6 and 4.2.

The total gains from trade are relatively large compared with other estimates in the literature. Beside the fact that there is a large reduction in trade cost during this period, the two channels for the larger gains are the above-mentioned finding that pro-competitive effects increase the total gains and the lower trade elasticities in our estimated models as discussed in Section 4.2. Both channels are important.

We find that the gains from trade and its components are substantially smaller in the symmetric-country case compared with the benchmark case, indicating the important role played by the differences in productivities and markups across countries. The fact that the symmetric-country implementation may obscure sizable gains from trade indicates the importance of implementing asymmetric-country estimation, especially when the country of concern is a developing one, such as China. Our approach of separating moments from exporters and non-exporters proves to be instrumental in such an implementation.

How can one think about policy in this model? In our model, λ_1 (mean number of draws) reflects domestic industrial/competition policy, but from a welfare point of view, decreasing trade cost τ is similar to increasing λ_1. In particular, from autarky to a fully integrated world, λ_1 increases without changing any domestic industrial/competition policy. If both trade policy and domestic industrial policy are tools that a government can use, which one to use depends on the relative benefits and costs of implementing these policies.

Exploiting the variations in sectoral markups and trade costs, we find that China on average liberalized the " right" sectors in the sense that the dispersion of markups is reduced because there tended to be larger trade liberalization in sectors with higher initial markups. Even though we do not know exactly how this happened, to target trade liberalization in sectors with higher markups is a useful take-away. This is particularly so when it is difficult to eliminate distortions in some industries via domestic measures.

Acknowledgements

For their helpful comments, we thank the Editor and two anonymous referees, Costas Arkolakis, Issac Baley, Pao-Li Chang, Davin Chor, Jonathan Eaton, Tom Holmes, Chang-Tai Hsieh, Yu-Chin Hsu, Nicolas Jacquet, Sam Kortum, James Markusen, Sanghoon Lee, Hong Ma, Andres Rodríguez-Clare, Tom Sargent, Michael Zheng Song, Bob Staiger, Ping Wang, Yong Wang, Michael Waugh, and Daniel Yi Xu. We also thank Yunong Li and Xuan Luo for their excellent research assistance. Hsu and Lu gratefully acknowledge the financial support provided by the Sing Lun Fellowship of Singapore Management University and the National Natural Science Foundation of China, General Program 71773103, respectively. Part of the research work is done when Hsu visited the Institute of Economics at Academia Sinica, and their hospitality is gratefully acknowledged.

Appendix

A.1. Estimation of Markups

In this subsection, we provide the details for calculating firm markups using DLW's method. Specifically, we assume that firm i at time t has the following production technology[42]

$$Q_{it} = F_{it}(L_{it}, K_{it}, M_{it}, \omega_{it}), \tag{25}$$

where L_{it}, K_{it}, and M_{it} are the inputs of labor, capital, and intermediate materials, respectively; ω_{it} denotes firm-specific productivity. The production function $F(.)$ is assumed to be continuous and twice-differentiable with respect to all of its arguments.

Consider the following cost minimization problem firm i faces at time t:

$$\begin{aligned} &\min_{\{L_{it},K_{it},M_{it}\}} w_{it}L_{it} + r_{it}K_{it} + p_{it}^m M_{it} \\ &s.t. F_{it}(L_{it}, K_{it}, M_{it}, \omega_{it}) \geq Q_{it}, \end{aligned} \tag{26}$$

where w_{it}, r_{it}, and p_{it}^m denote the wage rate, rental price of capital and the price of intermediate inputs, respectively; Q_{it} is a given number of output.

The estimation of firm-level markup hinges on choosing an input that is free of any adjustment costs and the estimation of the elasticity of output to this input. As labor is largely not freely chosen in China (particularly state-owned enterprises) and capital is often considered a dynamic input (which makes its input elasticity difficult to interpret), we choose intermediate materials as the input to estimate firm markup (see also DLW). Specifically, the Lagrangian function associated with the optimization problem (26) can be written as

$$\begin{aligned} \mathscr{L}(L_{it}, K_{it}, M_{it}, \lambda_{it}, \eta_{it}) = &\, w_{it}L_{it} + r_{it}K_{it} + p_{it}^m M_{it} \\ &+ \lambda_{it}[Q_{it} - F_{it}(L_{it}, K_{it}, M_{it}, \omega_{it})]. \end{aligned}$$

Hence, the first-order condition for intermediate materials is

$$\frac{\partial \mathscr{L}}{\partial M_{it}} = p_{it}^m - \lambda_{it}\frac{\partial F_{it}}{\partial M_{it}} = 0. \tag{27}$$

Rearranging equation (27) and multiplying both sides by $\frac{M_{it}}{Q_{it}}$ yield

$$\begin{aligned} \frac{\partial F_{it}}{\partial M_{it}}\frac{M_{it}}{Q_{it}} &= \frac{1}{\lambda_{it}} \\ &= \frac{P_{it}p_{it}^m M_{it}}{\lambda_{it} P_{it} Q_{it}}, \end{aligned} \tag{28}$$

where P_{it} is the price of the final good.

Note that $\lambda_{it} = \frac{\partial \mathscr{L}}{\partial Q_{it}} = mc_{it}$ represents the marginal cost of production at a given level of output. Define firm markup μ_{it} as the ratio of price over marginal cost, i.e. $\mu_{it} \equiv \frac{P_{it}}{mc_{it}} = \frac{P_{it}}{\lambda_{it}}$. Hence, equation (28) leads to the following estimation expression of firm markup:[43]

$$\mu_{it} = \theta_{it}^m \left(\alpha_{it}^m\right)^{-1}, \tag{29}$$

where $\theta_{it}^m \equiv \frac{\partial F_{it}}{\partial M_{it}}\frac{M_{it}}{Q_{it}}$ is the output elasticity of intermediate materials and $\alpha_{it}^m \equiv \frac{p_{it}^m M_{it}}{P_{it}Q_{it}}$ is the share of the expenditure of intermediate materials in total revenue.

As the information about the expenditure on intermediate materials and total revenue is available in the data, α_{it}^m can be readily calculated. However, the output elasticity of intermediate materials, θ_{it}^m, must be obtained by estimating the production function (25). There is a large literature on the estimation of the production function focusing on how to control for unobserved productivity shocks (for a review, see Ackerberg et al., 2007). Solutions include instrumental variable estimation, GMM estimation, and the control function

[42] Note that the framework is robust to any arbitrary number of inputs. As we observe only three inputs (i.e., labor, capital and intermediate materials) in our data, here we focus on production technology involving only these three inputs.

[43] Note that this expression holds under any form of market competition and demand function. Specifically, DLW discuss some alternative market structures, which lead to a similar estimation expression for firm markup. These alternative market structures include Cournot competition, Bertrand competition, and monopolistic competition.

approach proposed by Olley and Pakes (1996). We adopt the control function approach developed by Ackerberg et al. (2006), which comprises a two-step estimation.

Similar to DLW, we assume a translog production function when estimating markups. Specifically, the production function to be estimated is expressed as

$$\begin{aligned} q_{it} = {} & \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \beta_{ll} l_{it}^2 + \beta_{kk} k_{it}^2 + \beta_{mm} m_{it}^2 \\ & + \beta_{lk} l_{it} k_{it} + \beta_{km} k_{it} m_{it} + \beta_{lm} l_{it} m_{it} \\ & + \beta_{lkm} l_{it} k_{it} m_{it} + \omega_{it} + \varepsilon_{it}, \end{aligned} \tag{30}$$

where the lowercase letters represent the logarithm of the uppercase letters, ω_{it} is firm-specific productivity, and ϵ_{it} is an *i.i.d.* error term. $\beta = (\beta_l, \beta_k, \beta_m, \beta_{ll}, \beta_{kk}, \beta_{mm}, \beta_{lk}, \beta_{km}, \beta_{lm}, \beta_{lkm})$ is the vector of production function coefficients.

To proxy ω_{it}, Levinsohn and Petrin (2003) assume that

$$m_{it} = m_t(k_{it}, \omega_{it}, ex_{it}),$$

where ex_{it} denotes the exporter status (i.e. taking value 1 if exporters and 0 otherwise). Given the monotonicity of $m_t(.)$, we have

$$\omega_{it} = h_t(m_{it}, k_{it}, ex_{it}).$$

In the first stage, we estimate the following equation

$$q_{it} = \phi_{it} + \epsilon_{it},$$

where

$$\begin{aligned} \phi_{it} = {} & \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \beta_{ll} l_{it}^2 + \beta_{kk} k_{it}^2 + \beta_{mm} m_{it}^2 \\ & + \beta_{lk} l_{it} k_{it} + \beta_{km} k_{it} m_{it} + \beta_{lm} l_{it} m_{it} + \beta_{lkm} l_{it} k_{it} m_{it} + h_t(m_{it}, k_{it}, ex_{it}), \end{aligned}$$

and obtain the estimates of the expected output ($\hat{\phi}_{it}$) and the error term ($\hat{\varepsilon}_{it}$).

Meanwhile, to recover all the production function coefficients β in the second stage, we model firm productivity as following a first-order Markov movement, i.e.

$$\omega_{it} = g_t(\omega_{it-1}) + \xi_{it},$$

where ξ_{it} is an idiosyncratic shock.

From the first stage, the productivity for any given value of β can be computed as

$$\omega_{it}(\boldsymbol{\beta}) = \hat{\phi}_{it} - \left(\begin{array}{l} \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \beta_{ll} l_{it}^2 + \beta_{kk} k_{it}^2 + \beta_{mm} m_{it}^2 \\ + \beta_{lk} l_{it} k_{it} + \beta_{km} k_{it} m_{it} + \beta_{lm} l_{it} m_{it} + \beta_{lkm} l_{it} k_{it} m_{it} \end{array} \right).$$

The idiosyncratic shock to productivity given $\boldsymbol{\beta}$, $\xi_{it}(\boldsymbol{\beta})$, can then be obtained through a non-parametric regression of $\omega_{it}(\boldsymbol{\beta})$ on $\omega_{it-1}(\boldsymbol{\beta})$.

To identify the coefficients of the production function, Ackerberg et al. (2006) assume that capital is determined one period beforehand and hence is not correlated with $\xi_{it}(\boldsymbol{\beta})$. Meanwhile, wage rates and prices of intermediate materials are assumed to vary across firms and be serially correlated.

Therefore, the moment conditions used to estimate the coefficients of the production function are

$$E(\xi_{it}(\boldsymbol{\beta}) \mathbf{Y}'_{it}) = 0,$$

where $\mathbf{Y}_{it} = \{l_{it-1}, l_{it-1}^2, m_{it-1}, m_{it-1}^2, k_{it}, k_{it}^2, l_{it-1} m_{it-1}, l_{it-1} k_{it}, m_{it-1} k_{it}, l_{it-1} m_{it-1} k_{it}\}$.

We estimate the translog production function (30) separately for each 2-digit industry using the Annual Survey of Manufacturing Firms conducted by the NBS from 1998 to 2005. Specifically, we use the logarithm of sales deflated by 2-digit ex-factory price indices to measure q_{it}, the logarithm of employment to measure l_{it}, and the logarithm of intermediate materials[44] deflated by input price indices to measure m_{it}; to compute the logarithm of capital k_{it}, we use the perpetual inventory method as in Brandt, Van Biesebroeck, and Zhang (2012; Online Appendix A3) to calculate real capital with yearly investments deflated by investment price indices. All price indices are provided by Brandt et al. (2012).

Once $\hat{\boldsymbol{\beta}} = (\hat{\beta}_l, \hat{\beta}_k, \hat{\beta}_m, \hat{\beta}_{ll}, \hat{\beta}_{kk}, \hat{\beta}_{mm}, \hat{\beta}_{lk}, \hat{\beta}_{km}, \hat{\beta}_{lm}, \hat{\beta}_{lkm})$ is obtained, we can readily calculate the firm markup using equation (29), i.e.

$$\hat{\mu}_{it} = \hat{\theta}_{it}^m (\alpha_{it}^m)^{-1},$$

where $\hat{\theta}_{it}^m = \hat{\beta}_m + 2\hat{\beta}_{mm} m_{it} + \hat{\beta}_{lm} l_{it} + \hat{\beta}_{km} k_{it} + \hat{\beta}_{lmk} l_{it} k_{it}$. Panel A of Table A1 reports the medians and inter-quartile ranges of input elasticities of output.

A.2. Comparison with the BEJK Model

We conduct two sets of comparisons with the BEJK model. The first set is to fit the BEJK model by SMM to the moments that discipline our benchmark estimation and then compare with our benchmark result from Table 2. The second set is to add moments that BEJK were concerned with matching and use SMM to estimate both the BEJK and our model. To save space, we conduct estimations for 2004 only.

We highlight a few features in BEJK quantification; readers are referred to their paper for details. By some change of variables, BEJK show that the Fréchet scaling parameters $\{T_i\}$, trade costs $\{\tau_{ni}\}$, and wages (or input prices) $\{w_i\}$ can be absorbed into trade shares $\{\pi_{ni}\}$, which is observable. By making assumptions on non-manufacturing sectors, BEJK treat wages $\{w_i\}$ as exogenous, and hence the observed trade shares can also be treated as parameters. Total expenditure $\{x_n\}$, which corresponds to our $\{R_n\}$ are also treated as exogenous.[45] Thus, given trade shares $\{\pi_{ni}\}$ and total expenditure $\{x_n\}$, the remaining parameters to estimate are only the Fréchet shape parameter θ and elasticity of substitution σ. BEJK use productivity and size advantages of exporters (relative to nonexporters) in their 1992 US plant-level data to back out θ and σ, where productivity is measured by value-added per worker, and size is measured by domestic sales.[46] They also examine external validity by checking the fit to the following moments: the fraction of exporters among all firms, the fraction of revenues from exports (which is also called export intensity), variability in log productivity and that in log size (measured by standard deviation).

One important difference of our model from BEJK is that $\{w_i, R_i\}$ are both endogenous because our model does not assume an outside sector and loses the feature of a constant share of total profits in total revenue. Thus, our model is harder to quantify than BEJK, with the BEJK enjoying the advantage of easy application for a multi-country quantitative analysis. Nevertheless, our focus is different as we focus on the distribution of markups. Our first set of comparisons is thus to examine the fit of the relatively simple BEJK model to the moments that concern us. The second set is then to also take into account the moments that concern BEJK. As the fraction of exporters is already included in our set of moments, the second set adds six moments (for export intensity we report both weighted mean and weighted standard deviation).[47] We compute

[44] The value of intermediate materials is calculated as (production costs)−(total wages)−(total welfare benefits)−(current-year depreciation)×(production costs)/(production costs+selling costs+administrative costs+financial costs).

[45] BEJK assumes that production functions take intermediate inputs, and so there is an involved discussion related to total expenditure. However, even without intermediates, the total expenditure (which equals total revenue) can also be taken as exogenous because their model features a constant share of total profit out of total revenue. The total revenue is thus proportional to total labor income, $w_i L_i$, which is exogenous in their model.

[46] More precisely, the productivity advantage is the ratio of the average productivity of exporters to that of non-exporters. Size advantage is defined similarly.

[47] The weight is an exporter's revenues.

Table A1
Production function estimates.

Industry	Panel A: Output elasticity with respect to …						Panel B: Returns to scale				
	Labor		Capital		Materials		Double		Triple		
	Median	IQR	Median	IQR	Median	IQR	Median	IQR	Median	IQR	Obs.
Food processing	0.09	[0.07,0.13]	0.03	[0.01,0.05]	0.86	[0.81,0.90]	0.99	[0.98,1.00]	0.99	[0.98,1.00]	104,518
Food manufacturing	0.14	[0.11,0.18]	0.05	[0.02,0.08]	0.82	[0.76,0.87]	1.02	[1.00,1.04]	1.03	[1.00,1.04]	48,295
Beverage manufacturing	0.19	[0.14,0.25]	0.02	[-0.01,0.05]	0.78	[0.71,0.84]	1.01	[0.97,1.04]	1.01	[0.98,1.04]	41,894
Tobacco processing	0.17	[0.03,0.33]	0.24	[0.10,0.35]	0.73	[0.64,0.82]	1.14	[1.05,1.23]	1.14	[1.04,1.22]	731
Textile industry	0.16	[0.11,0.22]	0.04	[0.03,0.05]	0.84	[0.77,0.89]	1.03	[0.99,1.06]	1.02	[0.99,1.05]	113,001
Garments and other fiber products	0.23	[0.15,0.35]	0.05	[0.04,0.07]	0.75	[0.64,0.84]	1.02	[1.00,1.05]	1.02	[1.00,1.05]	72,381
Leather, furs, down and related products	0.20	[0.12,0.28]	0.01	[0.00,0.02]	0.81	[0.73,0.88]	1.01	[1.00,1.03]	1.01	[1.00,1.03]	34,655
Timber processing, bamboo, cane, palm fiber and straw products	0.15	[0.10,0.21]	0.03	[0.03,0.04]	0.83	[0.76,0.88]	1.01	[0.99,1.02]	1.00	[0.99,1.02]	57,283
Furniture manufacturing	0.38	[0.33,0.44]	-0.02	[-0.03,0.00]	0.99	[0.90,1.07]	1.37	[1.30,1.44]	1.38	[1.32,1.46]	34,126
Papermaking and paper products	0.26	[0.23,0.29]	0.05	[0.04,0.06]	0.85	[0.80,0.89]	1.15	[1.13,1.19]	1.16	[1.13,1.20]	55,606
Printing industry	0.24	[0.21,0.26]	0.11	[0.08,0.15]	0.86	[0.77,0.94]	1.24	[1.17,1.29]	1.25	[1.18,1.30]	57,993
Cultural, educational and sports goods	0.23	[0.15,0.34]	0.06	[0.05,0.08]	0.79	[0.70,0.86]	1.07	[1.04,1.11]	1.06	[1.04,1.10]	20,987
Petroleum processing and coking	0.10	[0.07,0.14]	0.06	[0.05,0.07]	0.83	[0.78,0.87]	0.99	[0.98,1.00]	0.99	[0.98,1.00]	10,430
Raw chemical materials and chemical products	0.22	[0.18,0.25]	0.04	[0.03,0.05]	0.72	[0.67,0.76]	0.97	[0.96,0.97]	0.96	[0.96,0.97]	108,197
Medical and pharmaceutical products	0.25	[0.18,0.32]	0.19	[0.13,0.26]	0.65	[0.55,0.74]	1.08	[1.04,1.12]	1.08	[1.04,1.11]	17,595
Chemical fiber	0.05	[0.01,0.09]	0.16	[0.15,0.18]	0.73	[0.69,0.76]	0.94	[0.92,0.95]	0.94	[0.92,0.95]	4,925
Rubber products	0.23	[0.19,0.27]	0.06	[0.06,0.07]	0.79	[0.73,0.83]	1.08	[1.06,1.09]	1.07	[1.06,1.09]	20,664
Plastic products	0.14	[0.09,0.19]	0.06	[0.05,0.07]	0.83	[0.77,0.88]	1.01	[1.00,1.03]	1.01	[1.00,1.03]	92,509
Nonmetal mineral products	0.15	[0.09,0.22]	0.05	[0.04,0.06]	0.80	[0.72,0.86]	0.98	[0.97,1.01]	0.98	[0.97,1.00]	226,792
Smelting and pressing of ferrous metals	0.10	[0.07,0.14]	0.03	[0.03,0.04]	0.85	[0.80,0.90]	0.98	[0.97,0.99]	0.98	[0.97,0.99]	29,102
Smelting and pressing of nonferrous metals	0.12	[0.08,0.16]	0.03	[0.03,0.04]	0.84	[0.79,0.88]	0.99	[0.99,1.00]	0.99	[0.99,1.00]	20,671
Metal products	0.17	[0.13,0.23]	0.09	[0.08,0.11]	0.71	[0.66,0.76]	0.97	[0.96,1.00]	0.97	[0.95,0.99]	117,081
Ordinary machinery	0.20	[0.16,0.26]	0.08	[0.06,0.09]	0.80	[0.73,0.85]	1.07	[1.06,1.09]	1.07	[1.06,1.08]	148,586
Special purpose equipment	0.24	[0.22,0.28]	0.08	[0.06,0.10]	0.79	[0.73,0.85]	1.13	[1.09,1.16]	1.13	[1.10,1.16]	77,157
Transport equipment	0.16	[0.11,0.22]	0.07	[0.06,0.09]	0.76	[0.69,0.82]	0.99	[0.99,1.00]	0.99	[0.98,1.00]	75,943
Electric equipment and machinery	0.15	[0.11,0.21]	0.06	[0.05,0.07]	0.79	[0.73,0.84]	1.00	[0.99,1.01]	1.00	[0.99,1.01]	63,631
Electronic and telecommunications equipment	0.23	[0.17,0.30]	0.10	[0.09,0.11]	0.73	[0.65,0.80]	1.06	[1.05,1.08]	1.06	[1.05,1.08]	48,716
Instruments, meters, cultural and office equipment	0.20	[0.13,0.29]	0.09	[0.07,0.10]	0.72	[0.63,0.79]	1.00	[0.97,1.04]	1.00	[0.96,1.03]	25,494
Other manufacturing	0.21	[0.14,0.29]	0.06	[0.04,0.07]	0.78	[0.70,0.84]	1.02	[1.00,1.06]	1.02	[1.00,1.05]	39,978

Notes: IQR means inter-quartile range. In Panel B, we calculate the r in $k^r Y = F(kK, kL, kM)$, where Y,K,L,M are output, capital, labor, and material, respectively. The calculation is local to the data values and our estimate. The columns under "double" and "triple" are the results when k is chosen to be 2 and 3, respectively.

the trade shares $\{\pi_{ni}\}$ and $\{x_n\}$ in our two-country framework that is consistent with our data.[48]

Note that the measure of goods is normalized to one in BEJK. For comparison, we also assume that there is a fixed measure of good γ in the BEJK model and use the moment " relative number of firms" to estimate this parameter. In the BEJK model, the counterpart to (16) is

$$\frac{N_1}{\overline{N}} = \frac{\gamma}{\overline{N}} \times \Pr\left[\frac{1}{\varphi^*_{1\omega}} < \frac{w\tau}{\varphi^*_{2\omega}}\right] = \frac{\gamma}{\overline{N}} \times \pi_{11}.$$

As mentioned in footnote 23, we scale both the numerator and denominator of the right-hand side of (16) by 1/10, and so $\overline{N}$ is set to 200, 000. Again, we have $\gamma \equiv$ *total_goods_baseline*$^*\tilde{\gamma}$, where *total_goods_baseline* is set to be 250,000. As $N_1/\overline{N} = 0.596$ and $\pi_{11} = 0.778$, $\tilde{\gamma} = 0.613$. As the "relative number of firms" is already used to estimate γ, it does not enter the SMM procedure to estimate the remaining parameters (θ, σ) in the BEJK model.

The estimation results are reported in Table A2. Following Simonovska and Waugh (2014), we report the p-value of the *J* statistics, which is used to test the null hypothesis that the model is correctly specified.[49] Unlike the situation where an optimal weighting matrix is used in the SMM procedure, the analytical form of the sampling distribution of the *J* statistic under the equal-weight weighting matrix is generally unknown. To overcome this difficulty, we bootstrap the *J* statistic for each of the four estimations presented in Table A2. Specifically, we randomly draw 200 samples from the dataset and calculate the set of moments for each sample. For each such set of moments, each model is re-estimated and the *J* statistic computed. This procedure yields a sampling distribution of the *J* statistic. The smaller the p-value, the more unlikely the model is correctly specified. Here, we see that the null hypothesis that the model is well-specified is not rejected for any of the models that we consider. Moreover, our model generates higher p-values than the BEJK counterparts, implying that our model overall fits the data moments better than the BEJK model.

Several more observations are in order. First, the most striking pattern of the estimates is that the σ estimates in both models increase significantly from 1.45 and 1.24 to around 3.6 when the BEJK moments are included. This is consistent with the BEJK estimate of σ at 3.79 and mainly because the BEJK moments emphasize sales. The BEJK estimate of θ is 3.6, whereas the estimates here are only slightly higher. The reason for the low σ in the first set is to allow a larger range whereby markups can vary so as to fit the markup distribution better. When σ becomes larger in the second set, the fit on the moments of markups becomes much worse. This tradeoff is a common feature of the two models.

Second, comparing the estimations of our model in the two sets, the parameter estimates generally do not vary much: the main differences are the increase in σ and the decreases in λ_1 and λ_2. Whereas σ increases to fit the sales moments, causing a narrow range of markups and the model implied markups to be generally much lower than the data counterparts, both λ_1 and λ_2 decrease so that markups can increase to improve the fit on this margin. This mechanism is lacking in the BEJK model, and this explains partially why our model fits the markup moments better than BEJK in the second set.

Third, the fact that the BEJK model does not fit these (mostly micro) moments as well is not totally surprising. The BEJK model is highly

[48] To more accurately compute total expenditure and trade shares, we recognize the trade surplus that China enjoys. Note that the import share entails $\phi_{1,2} = 0.222$. Then, $\phi_{1,1} = 1 - \phi_{1,2} = 0.778$. Recognizing China's trade surplus in 2004, $D_{2004} = 0.0263 \times R_1$, we compute total expenditure $Y_1 = R_1 - D$ and $Y_2 = R_1 + D$. We also compute $\phi_{2,1}$ from the following equation $(R_2 + D)\phi_{2,1} = (R_1 - D)\phi_{1,2} + D$.

[49] More specifically, this is calculated as the SMM objective function evaluated at the parameter estimates times $\frac{NS}{1+S}$, where *N* is the number of observations and *S* the number of simulation paths. In our implemention, *S* is set to 1.

Table A2
Comparison with the BEJK model.

	First comparison (14 moments)				Second comparison (20 moments)			
	Our model		BEJK model		Our model		BEJK model	
Moments for SMM	Data	Model	Data	Model	Data	Model	Data	Model
Import share	0.222	0.252	0.222	0.229	0.222	0.187	0.222	0.230
Export share	0.249	0.273	0.249	0.257	0.249	0.254	0.249	0.241
Relative number of firms	0.596	0.605	0.596	0.596	0.596	0.517	0.596	0.596
Fraction of exporters	0.105	0.064	0.105	0.049	0.105	0.107	0.105	0.049
Mean cost share for exporters	0.801	0.789	0.801	0.744	0.801	0.793	0.801	0.781
Std of cost share for exporters	0.142	0.139	0.142	0.161	0.142	0.073	0.142	0.070
p50 markup for exporters	1.168	1.224	1.168	1.277	1.168	1.288	1.168	1.238
p95 markup for exporters	2.183	2.207	2.183	2.352	2.183	1.389	2.183	1.383
p99 markup for exporters	3.364	3.225	3.364	3.511	3.364	1.389	3.364	1.383
Mean cost share for non-exporters	0.829	0.763	0.829	0.808	0.829	0.793	0.829	0.840
Std of cost share for non-exporters	0.139	0.161	0.139	0.148	0.139	0.093	0.139	0.099
p50 markup for non-exporters	1.213	1.285	1.213	1.188	1.213	1.342	1.213	1.160
p95 markup for non-exporters	2.400	2.193	2.400	1.943	2.400	1.389	2.400	1.383
p99 markup for non-exporters	3.523	2.735	3.523	2.488	3.523	1.389	3.523	1.383
Mean of export intensity					0.408	0.538	0.408	0.572
Std of export intensity					0.482	0.072	0.482	0.070
Size advantage of exporters					5.682	2.747	5.682	5.777
Productivity advantage of exporters					1.275	1.507	1.275	2.098
Std of log productivity					0.696	0.356	0.696	0.294
Std of log of domestic sales					1.444	0.844	1.444	0.620
Parameter values	Estimates	s.e.	Estimates	s.e.	Estimates	s.e.	Estimates	s.e.
τ, trade cost	1.782	0.007			1.809	0.004		
γ, relative measure of goods	0.659	0.003			0.634	0.005		
λ_1, Poisson parameter, China	2.618	0.017			2.000	0.052		
λ_2, Poisson parameter, ROW of log	5.024	0.048			3.940	0.032		
	−1.756	0.012			−1.763	0.009		
η_1, std of log productivity, China	0.425	0.002			0.471	0.002		
η_2, std of log productivity, ROW	0.357	0.011			0.398	0.007		
σ, elasticity of substitution	1.449	0.003	1.239	0.094	3.572	0.007	3.611	0.010
θ, Frechet shape parameter			3.754	0.012			4.361	0.011
P-value of the J test	0.92		0.78		0.61		0.24	

Notes: All the units, if any, are in billions USD, current price. The estimation is based on the 2004 data. The m14 columns are taken from Table 2.

tractable, much easier to quantify and apply to a multi-country setting, and explains the gravity equation. Much of its tractability comes from the fact that various country-specific parameters are absorbed into trade shares and total expenditures, which are treated as exogenous and can be taken from data. In short, there are four pre-determined parameters $\{x_1,x_2\}$ and $\{\pi_{11},\pi_{12},\pi_{21},\pi_{22}\}$, which are actually two parameters because $\sum_i \pi_{ni} = 1$, and two parameters to be estimated. In this sense, some data features are built into the BEJK model from the outset, and hence it is natural that the BEJK model fits the trade shares almost perfectly. Nevertheless, if there is any data pattern not accounted for by these pre-determined parameters, the BEJK model then relies on the adjustments in θ and σ to carry the load. In contrast, our model is more difficult to compute and quantify, but it speaks to micro moments with more flexibility with a richer micro structure. Interestingly, note that whereas our estimation does not input trade shares directly, our model fits the trade moments reasonably well; in particular, our model fits the fraction of exporters better than the BEJK model.

A.3. Welfare Gains by the ACR Formula with Large Change in Trade Cost

We first calculate

$$\ln\frac{W_j^{ACR,2004}}{W_j^{ACR,1995}} = \int_{\tau_{1995}}^{\tau_{2004}} d\ \ln W_j^{ACR}(\tau) = \int_{\tau_{1995}}^{\tau_{2004}} d\left(\frac{\ln \nu_{jj}(\tau)}{\varepsilon(\tau)}\right).$$

To numerically calculate the above, we discretize the interval of $[\tau_{2004},\tau_{1995}]$ via the use of an n-grid such that $\tau_0 = \tau_{1995}$, $\tau_1 = \tau_{1995} - \frac{\tau_{1995}-\tau_{2004}}{n}$,..., $\tau_i = \tau_{1995} - i \times \frac{\tau_{1995}-\tau_{2004}}{n}$,..., and $\tau_n = \tau_{2004}$. The ACR formula for this large change in trade cost is thus calculated by

$$\ln\frac{W_j^{ACR,2004}}{W_j^{ACR,1995}} \approx \sum_{i=1}^{n} \frac{1}{\epsilon_i}\left[\ \ln \nu_{jj}(\tau_i) - \ \ln \nu_{jj}(\tau_{i-1})\right].$$

We calculate $\nu_{jj}(\tau_i)$ precisely at τ_i, and we calculate the trade elasticity ϵ_i on each i-th grid using the two-point formula mentioned in footnote 28 at $\tau = \frac{\tau_{i-1}+\tau_i}{2}$. For our numerical calculation, we use $n = 50$ so that the grid size is $(2.311 - 1.782)/50 = 0.01058$. Once we obtain $\ln\left(\frac{W_j^{ACR,2004}}{W_j^{ACR,1995}}\right)$, we can then calculate the percentage increase in welfare $\left(\frac{W_j^{ACR,2004}}{W_j^{ACR,1995}} - 1\right) \times 100\%$.

A.4. Comparative Statics of Some Parameters

First, we take a closed economy and consider what happens if population L doubles. The scale effect here can be interpreted as going from autarky to full integration among the countries. One quick result is that if the entry parameter λ remains fixed, then there is no effect on per capita welfare; only the total welfare scales up proportionally with the population. However, it is reasonable to assume that λ also scales up with L; as the number of firms in a free-trade world is more than each autarkic economy. Based on 2004 parameters, the result is reported in the following table. We denote the change of welfare by $d \ln W = \ln W' - \ln W$. Note that per capita welfare is $W_j^{PC} \equiv W_j^{Total}/L_j$ and that there is no relative markup effect for this exercise.

Here, we see that both per capita welfare and its components increase. As λ increases, there are more draws from the productivity distribution. Hence, there are gains due to increased productivity because "the best" now becomes better. There are also gains in allocative efficiency because of the compression of the ratio between the top two productivities when there are more draws from a non-fat-tailed distribution.[50] The gains in allocative efficiency here are relatively modest compared with the gains due to enhanced productivity.

	L and λ doubles	η doubles
$d \ln W_1^{PC}$	0.384	0.356
$d \ln W_1^{Prod}$	0.330	0.382
$d \ln W_1^{A}$	0.053	−0.026

In the case where the standard deviation of log-productivity doubles, both per capita welfare and productive efficiency increase, but the allocative efficiency decreases. The increase in productive efficiency is readily comprehensible. As η increases, not only does it increase the mean, but the top productivity is increased even more as the dispersion at the right-tail increases. In contrast, the increase in the dispersion at the right-tail enlarges the ratio between the top two productivities, and thus increases markup dispersion and reduces allocative efficiency. However, the effect on productive efficiency dominates and thus per capita welfare still increases.

Next, we return to an open economy, and consider symmetric countries for clarity. In Section 4.3, we saw the effect of trade liberalization in the symmetric-country case. Here, we seek to investigate the role of productivity dispersion (η) and the mean number of draws (λ, which reflects market structure) on gains from trade. As such, we replicate the exercise of gains from trade between 1995 and 2004, but under different levels of η (Panel A of Table A3), as well as under different levels of λ (Panel B of Table A3).

The middle columns of both panels are the same as that reported in panel A of Table 7. Panel A shows that the larger the dispersion of the productivity distribution, the larger the gains from trade in total and in productive efficiency. When the productivity draws are more dispersed and hence more skewed to the right, the best productivity in each country is therefore higher, increasing the gains from trade via productive efficiency. There are always positive gains from trade via improved allocative efficiency, but the magnitude is relatively stable. Thus, the contribution of allocative efficiency diminishes from as large as 54.6% at 0.5 times η_0, the standard deviation at 2004, to as small as 10.9% at 1.5 times η_0.

In Panel B, gains from trade in total and the two components are all decreasing in the level of λ. With a given distribution of productivity draws, the more draws suggest that the top productivity and the ratio between the top two are both operating at a righter part of the tail. The fact that the log-normal distribution is not fat-tailed implies that trade liberalization induces smaller increases in the productivities of actual suppliers when λ is higher because they were already quite high before trade liberalization. Also, when λ is higher, the same non-fat-tailed nature implies that trade liberalization induces a smaller reduction in the ratio of the top two productivities because they were already small before trade liberalization. The diminishing speed of these two components in λ are roughly the same, resulting in a relatively stable contribution of pro-competitive effects across different λ's.

A.5. Robustness Checks

We conduct four robustness checks. Recall that in the benchmark case, the counter-factual analyses are based on 2004 estimates and change τ back to the 1995 level. The first robustness check is to conduct a counter-factual analysis based on 1995 estimates and change τ to the 2004 level. In our second check, we use an alternative measure of markups to estimate the model and run counter-factuals. That is, by invoking the constant-returns-to-scale assumption, we calculate *raw markups* by taking the ratio of revenue to total costs.

Table A3
Comparative statics of other parameters.

Panel A: Comparative statics of η on gains from trade					
		Gains from trade (in percentage)			
η	$0.5\times\eta_0$	$0.75\times\eta_0$	$\eta_0=0.407$	$1.25\times\eta_0$	$1.5\times\eta_0$
Totalwelfare	0.53%	1.25%	2.35%	3.51%	4.42%
W Prod	0.24%	0.87%	1.89%	2.94%	3.94%
W_A	0.29%	0.37%	0.47%	0.51%	0.48%
Contribution of W_A	54.6%	29.4%	19.8%	14.5%	10.9%
Panel B: Comparative statics of λ on gains from trade					
		Gains from Trade (in percentage)			
λ	$0.5\times\lambda_0$	$0.75\times\lambda_0$	$\lambda_0=4.219$	$1.25\times\lambda_0$	$1.5\times\lambda_0$
Total welfare	4.64%	3.20%	2.36%	1.87%	1.56%
W Prod	3.71%	2.52%	1.89%	1.52%	1.31%
W_A	0.88%	0.66%	0.47%	0.33%	0.26%
Contribution of W_A	19.0%	20.6%	19.7%	17.6%	16.8%

Notes: Under symmetric countries, W_R = 1. In both panels, the analyses are done under 2004 estimates, and only the trade cost (τ) is changed to the level at 1995. The reported percentage increases in welfare are under the change from 1995'τ to 2004's τ. The contribution of allocative efficiency is the ratio of the percentage increase in allocative efficiency to that of total welfare.

There were substantial trade surpluses in China in both 1995 and 2004. They account for 2.25% of China's manufacturing sales in 1995 and 2.63% in 2004. Our third check is to accommodate trade imbalance in the model. To do this, we follow the literature by allowing an exogenous trade deficit D_i for each country i with the requirement that $D_1+D_2=0$.[51] With trade deficits, the total income in country i is R_i+D_i. As China has a trade surplus in both years, we can set here $D_2=D>0$ and $D_1=-D$, where D is the size of surplus in China. The details about the equilibrium conditions, the algorithm, and the implementation of SMM of this modified model can be found in Appendix A6.

Another potential concern on our results is that a substantial fraction of the Chinese trade is intermediate goods and "processing." In the benchmark, processing trade is included in the total import and export when calculating the import and export shares. Our fourth check is based on the export and import figures that exclude "processing trade." We first find the custom data on aggregate processing trade in each year. The processing-trade export is higher than the import because the processing-trade imports are intermediate inputs, whereas exports add some value-added from domestic inputs (including labor). Hence, we purge the processing-trade import from both total export and total import and re-calculate import and export shares. Further details and the SMM results are given in Appendix A7.

The results are reported in Table A4. We omit the numbers of the level of total welfare and its components and simply report the corresponding percentage changes. The total gains from trade between 1995 and 2004 range from 5.0% to 9.2%, and the contribution of pro-competitive effects ranges from 13.1% to 32.1%, and that of allocative efficiency ranges from 15.6% to 30.7%. These indicate that the importance of pro-competitive effects remains similar, and the allocative efficiency still accounts for the bulk of gains from trade.

The only difference between the first robustness check and the benchmark is that all parameters besides τ are fixed at the 1995 levels instead of at the 2004 levels. The overall gains become larger, but the pro-competitive effects remain similar. Next, note that the σ estimate under raw markups is about 1.81 in both years, which implies a smaller upper bound of markups and hence a smaller markup dispersion than the benchmark case. So it is not surprising that the pro-competitive

[50] Holmes et al. (2014) highlight this result.

[51] For example, see Caliendo and Parro (2015).

Table A4
Robustness check of counter-factual analyses.

Robustness check 1: Based on 1995 estimates						
	Welfare				Contribution to total welfare	
	Total welfare	W_Prod	W_A	W_R	W_A and W_R	W_A
% change from τ at 1995 to τ at 2004	9.2%	7.3%	1.7%	0.1%	19.2%	18.2%
% change from autarky to τ at 1995	23.8%	18.0%	5.0%	-0.1%	20.6%	20.9%
Robustness check 2: Under raw markups						
	Welfare				Contribution to total welfare	
	Total welfare	W_Prod	W_A	W_R	W_A and W_R	W_A
% change from τ at 1995 to τ at 2004	5.0%	4.3%	0.8%	-0.1%	13.1%	15.6%
% change from autarky to τ at 2004	14.1%	11.6%	2.5%	-0.2%	16.3%	17.8%
Robustness check 3: Model with trade imbalance						
	Welfare				Contribution to total welfare	
	Total welfare	W_Prod	W_A	W_R	W_A and W_R	W_A
% change from τ at 1995 to τ at 2004	7.2%	5.6%	1.6%	0.0%	21.1%	21.7%
Robustness check 4: Under alternative estimates when processing trade is considered						
	Welfare				Contribution to total welfare	
	Total welfare	W_Prod	W_A	W_R	W_A and W_R	W_A
% change from τ at 1995 to τ at 2004	5.1%	3.4%	1.6%	0.1%	32.1%	30.7%
% change from autarky to τ at 2004	21.1%	15.2%	5.0%	0.2%	24.4%	23.7%

Notes: In the first robustness check, the analysis is based on the 1995 estimate and we change τ to the 2004 level. In the next three robustness checks, analyses are done based on 2004 estimates, as in the benchmark case.

effects are slightly less important under raw markups. The total gains from trade and the contribution from pro-competitive effects in the third robustness check are quite similar to the benchmark results in Table 4.[52] In the fourth check, as the import and export shares decrease, the estimated trade costs increase from 2.311 to 2.674 in 1995 and from 1.782 to 2.036 in 2004. However, the percentage decrease in trade cost between the two years remains similar to the benchmark case. Compared with Table 4, the total welfare gains from trade are reduced, and this is likely due to the higher trade costs. However, the relative contributions of the pro-competitive effects increase significantly (32% and 24%).

A.6. Computation in the Model with Trade Imbalance

To model trade imbalance, we follow the literature by allowing an exogenous trade deficit D_i for each country i with the requirement that $D_1 + D_2 = 0$. The total income in country i is therefore $Y_i = R_i + D_i$. As China has a trade surplus in both years, we can set here $D_2 = D > 0$ and $D_1 = -D$, where D is the size of surplus in China.

Observe that the total imports of country j from country i is

$$R_{j,i} = \int_{\left\{\omega:\chi_j^*(\omega)=i\right\}} E_{j\omega} d\omega = Y_j \int_{\left\{\omega:\chi_j^*(\omega)=i\right\}} \left(\frac{p_{j\omega}}{P_j}\right)^{1-\sigma} d\omega \equiv Y_j \phi_{j,i}. \quad (31)$$

where $\chi_j^*(\omega) \in \{1,2\}$ denotes the source country for any particular good ω at destination j and $\phi_{j,i}$ denote the spending share of country j's consumers on the goods produced in i. The previous balanced trade condition (9) is now modified as $(R_2 + D)\phi_{2,1} = (R_1 - D)\phi_{1,2} + D$, or equivalently,

$$R_2\phi_{2,1} = R_1\phi_{1,2} + D(\phi_{2,2} - \phi_{1,2}). \quad (32)$$

The algorithm for computing equilibrium is more complicated than the benchmark model. First, observe from the definition of the producers' aggregate markup for country 1:

$$\begin{aligned} M_1^{sell} &= \frac{R_1}{w_1 L_1} = \frac{\int_{\{\omega:\chi_1^*(\omega)=1\}} \phi_{1\omega} Y_1 d\omega + \int_{\{\omega:\chi_2^*(\omega)=1\}} \phi_{2\omega} Y_2 d\omega}{\int_{\{\omega:\chi_1^*(\omega)=1\}} m_{1\omega}^{-1} \phi_{1\omega} Y_1 d\omega + \int_{\{\omega:\chi_2^*(\omega)=1\}} m_{2\omega}^{-1} \phi_{2\omega} Y_2 d\omega} \\ &= \left(\int_{\{\omega:\chi_1^*(\omega)=1\}} m_{1\omega}^{-1} \frac{\phi_{1\omega}(R_1 - D)}{R_1} d\omega + \int_{\{\omega:\chi_2^*(\omega)=1\}} m_{2\omega}^{-1} \frac{\phi_{2\omega}(R_2 + D)}{R_1} d\omega \right)^{-1} \\ &= \left(\frac{R_1 - D}{R_1} \int_{\{\omega:\chi_1^*(\omega)=1\}} m_{1\omega}^{-1} \phi_{1\omega} d\omega + \phi_{1,2} + \frac{\frac{D}{R_1}\phi_{1,1}}{\phi_{2,1}} \int_{\{\omega:\chi_2^*(\omega)=1\}} m_{2\omega}^{-1} \phi_{2\omega} d\omega \right)^{-1} \end{aligned} \quad (33)$$

Recall that $\phi_{j,i} = \int_{\{\omega:\chi_j^*(\omega)=i\}} \phi_{j\omega} d\omega$ depends only on relative wage w, but not on R_1 and R_2. Hence, M_1^{sell} becomes a function of w and R_1 only. For any given R_1 and w, we can calculate $M_1^{sell}(w, R_1)$. Then, with $w_1 = 1$, we can plug $R_1 = M_1^{sell}(w, R_1) L_1$ into (33) and solve for $M_1^{sell}(w)$. We have

$$\begin{aligned} &M_1^{sell}(w) \\ &= L_1 - D \left[\frac{\frac{1-\phi_{1,2}}{\phi_{2,1}} \int_{\{\omega:\chi_2^*(\omega)=1\}} m_{2\omega}^{-1} \phi_{2\omega} d\omega - \int_{\{\omega:\chi_1^*(\omega)=1\}} m_{1\omega}^{-1} \phi_{1\omega} d\omega]}{L_1 \left[\int_{\{\omega:\chi_1^*(\omega)=1\}} m_{1\omega}^{-1} \phi_{1\omega} d\omega + \frac{\phi_{1,2}}{\phi_{2,1}} \int_{\{\omega:\chi_2^*(\omega)=1\}} m_{2\omega}^{-1} \phi_{2\omega} d\omega \right]} \right. . \end{aligned}$$

With $R_1 = M_1^{sell}(w) L_1$, we use (32) again to calculate $R_2(w)$:

$$R_2(w) = R_1(w) \frac{\phi_{1,2}}{\phi_{2,1}} + D\left(\frac{1-\phi_{1,2}}{\phi_{2,1}} - 1\right).$$

[52] Autarky is inconsistent with trade imbalance; hence in this case there is no result for the counter-factual based on autarky.

Table A5
SMM results in the model with trade imbalance.

		1995		2004	
Predetermined					
W	Relative wages (the ROW to China)	10.25		5.18	
R1	China's manufacturing sales ($b)	918,291		2,343,328	
R2	ROW's manufacturing sales ($b)	9,397,500		14,737,500	
Moments for SMM		Data	Model	Data	Model
Import share		0.130	0.137	0.222	0.235
Export share		0.153	0.181	0.249	0.293
Relative number of firms		0.210	0.197	0.596	0.601
Fraction of exporters		0.044	0.025	0.105	0.066
Mean cost share for exporters		0.845	0.799	0.801	0.790
Std of cost share for exporters		0.135	0.142	0.142	0.137
p50 markup for exporters		1.196	1.217	1.168	1.223
p95 markup for exporters		2.199	2.169	2.183	2.115
p99 markup for exporters		3.299	3.176	3.364	3.317
Mean cost share for non-exporters		0.789	0.737	0.829	0.763
Std of cost share for non-exporters		0.147	0.178	0.139	0.161
p50 markup for non-exporters		1.266	1.334	1.213	1.285
p95 markup for non-exporters		2.475	2.558	2.400	2.196
p99 markup for non-exporters		3.537	3.176	3.523	2.740
Parameter values		Estimates	s.e.	Estimates	s.e.
τ, trade cost		2.337	0.024	1.782	0.009
$\gamma/\overline{N}$, measure of goods relative to $\overline{N}$		0.187	0.001	0.643	0.003
λ_1, Poisson parameter, China		2.985	0.058	2.710	0.028
λ_2, Poisson parameter, ROW		5.508	0.148	5.024	0.043
μ_1, mean of log productivity, China relative to ROW		-2.391	0.037	-1.751	0.015
η_1, std of log productivity, China		0.450	0.008	0.425	0.003
η_2, std of log productivity, ROW		0.437	0.012	0.357	0.012
σ, elasticity of substitution		1.459	0.004	1.432	0.002
Simulated R2/R1		Data	Model	Data	Model
R2/R1		10.234	9.388	6.289	5.704

Notes: All units, if any, are in billions USD, current price. The import share is the import penetration ratio, i.e. IM/(R1-EX+IM), and the export share is the total export divided by the same denominator. All the cost share moments are weighted by firms' revenues. Recall that a firm's cost share is the inverse of its markup. p# denotes the #-th percentile.

Table A6
SMM results when processing trade is considered.

		1995		2004	
Predetermined					
w	Relative wages (the ROW to China)	10.25		5.18	
R1	China's manufacturing sales ($b)	918,291		2,343,328	
R2	ROW's manufacturing sales ($b)	9,397,500		14,737,500	
Moments for SMM		Data	Model	Data	Model
Import share		0.073	0.079	0.135	0.150
Export share		0.096	0.109	0.162	0.185
Relative number of firms		0.210	0.239	0.596	0.586
Fraction of exporters		0.044	0.019	0.105	0.039
Mean cost share for exporters		0.845	0.785	0.801	0.788
Std of cost share for exporters		0.135	0.146	0.142	0.136
p50 markup for exporters		1.196	1.256	1.168	1.226
p95 markup for exporters		2.199	2.372	2.183	2.286
p99 markup for exporters		3.299	2.372	3.364	3.218
Mean cost share for non-exporters		0.789	0.728	0.829	0.741
Std of cost share for non-exporters		0.147	0.176	0.139	0.173
p50 markup for non-exporters		1.266	1.379	1.213	1.327
p95 markup for non-exporters		2.475	2.372	2.400	2.440
p99 markup for non-exporters		3.537	2.372	3.523	3.218
Parameter values		Estimates	s.e.	Estimates	s.e.
τ, trade cost		2.674	0.052	2.036	0.014
$\gamma/\overline{N}$, measure of goods relative to $\overline{N}$		0.215	0.001	0.563	0.002
λ_1, Poisson parameter, China		2.784	0.114	2.667	0.038
λ_2, Poisson parameter, ROW		4.924	0.062	4.982	0.018
μ_1, mean of log productivity, China relative to ROW		-2.389	0.037	-1.735	0.010
η_1, std of log productivity, China		0.434	0.014	0.425	0.002
η_2, std of log productivity, ROW		0.399	0.017	0.321	0.015
σ, elasticity of substitution		1.729	0.005	1.451	0.003
Simulated R2/R1		Data	Model	Data	Model
R2/R1		10.234	7.874	6.289	5.091

Notes: All units, if any, are in billions USD, current price. The import share is the import penetration ratio, i.e. IM/(R1-EX+IM), and the export share is the total export divided by the same denominator. All the cost share moments are weighted by firms' revenues. Recall that a firm's cost share is the inverse of its markup. p# denotes the #-th percentile.

Next, we calculate

$$M_2^{sell}(w) = \left(\frac{R_1(w)-D}{R_2(w)}\int_{\{\omega:\chi_1^*(\omega)=2\}} m_{1\omega}^{-1}\phi_{1\omega}d\omega + \frac{R_2(w)+D}{R_2(w)}\int_{\{\omega:\chi_2^*(\omega)=2\}} m_{2\omega}^{-1}\phi_{2\omega}d\omega\right)^{-1},$$

Finally, given L_2, we can use the market clearing condition of country 2 to solve for w:

$$M_2^{sell}(w) = \frac{R_2(w)}{wL_2}.$$

Given the solution of w, equilibrium R_1 and R_2 can be obtained using the above procedure. The SMM result of the modified model is presented in Table A5.

A.7. SMM Results Considering Processing Trade

When considering processing trade, the only changes from the benchmark SMM procedure are the import and export shares. From the custom data, we first obtain the fraction of processing-trade import in total import. We then use the formula for import and export shares, as well as R_1 given in Table 2, to re-calculate export and import shares. The new numbers for 1995 are 0.073 and 0.096 for import and export shares, respectively. The corresponding numbers for 2004 are 0.135 and 0.162. The SMM results are given in Table A6.

References

Ackerberg, D., Caves, K., Frazer, G., 2006. Structural identification of production functions. Working Paper. UCLA.

Ackerberg, D., Benkard, C.L., Berry, S., Pakes, A., 2007. Econometric tools for analyzing market outcomes. In: Heckman, James J., Leamer, Edward E. (Eds.), Handbook of Econometrics. 6A. Elsevier, North-Holland, Amsterdam and Boston, pp. 4171–4276.

Adda, J., Cooper, R.W., 2003. Dynamic Economics: Quantitative Methods and Applications. MIT Press.

Alvarez, Fernando, Lucas, Robert Jr., 2007. General equilibrium analysis of the Eaton-Kortum model of international trade. J. Monetary Econ. 54 (6), 1726–1768.

Arkolakis, C., Costinot, A., Rodríguez-Clare, A., 2012. New trade models, same old gains? Am. Econ. Rev. 102 (1), 94–130.

Arkolakis, C., Costinot, A., Donaldson, D., Rodríguez-Clare, A., 2019. The elusive pro-competitive effects of trade. Rev. Econ. Stud. 86 (1), 46–80.

Asturias, Jose, García-Santana, Manuel, Ramos, Roberto, 2017. Competition and the welfare gains from transportation infrastructure: evidence from the golden quadrilateral of India. Working Paper. Georgetown University, Qatar.

Atkeson, A., Burstein, A., 2008. Pricing to market, trade costs, and international relative prices. Am. Econ. Rev. 98 (5), 1998–2031.

Behrens, K., Murata, Y., 2012. Trade, competition, and efficiency. J. Int. Econ. 87 (1), 1–17.

Bernard, A.B., Eaton, J., Jensen, J.B., Kortum, S., 2003. Plants and productivity in international trade. Am. Econ. Rev. 93 (4), 1268–1290.

Brandt, L., Van Biesebroeck, J., Zhang, Y., 2012. Creative accounting or creative destruction: firm level productivity growth in Chinese manufacturing. J. Dev. Econ. 97 (2), 339–351.

Brandt, L., Van Biesebroeck, J., Wang, L., Zhang, Y., 2017. WTO accession and performance of Chinese manufacturing firms. Am. Econ. Rev. 107 (9), 2784–2820.

Caliendo, L., Parro, F., 2015. Estimates of the trade and welfare effects of NAFTA. Rev. Econ. Stud. 82 (1), 1–44.

Costinot, A., Rodríguez-Clare, A., 2014. Trade theory with numbers: Quantifying the consequences of globalization. Handbook of International Economics. Vol. 4. Elsevier, pp. 197–261.

de Blas, B., Russ, K., 2015. Understanding markups in the open economy under Bertrand competition. Am. Econ. J. Macroecon. 7 (2), 157–180.

De Loecker, J., Warzynski, F., 2012. Markups and firm-level export status. Am. Econ. Rev. 102 (6), 2437–2471.

De Loecker, J., Goldberg, P.K., Khandelwal, A.K., Pavcnik, N., 2016. Prices, markups and trade reform. Econometrica 84 (2), 445–510.

Devereux, M.B., Lee, K.M., 2001. Dynamic gains from international trade with imperfect competition and market power. Rev. Dev. Econ. 5 (2), 239–255.

Dhingra, S., Morrow, J., 2016. Monopolistic competition and optimum product diversity under firm heterogeneity. Working Paper. London School of Economics.

di Giovanni, J., Levchenko, A.A., Zhang, J., 2014. The global welfare impact of China: Trade integration and technological change. Am. Econ. J. Macroecon. 6 (3), 153–183.

Dixit, A.K., Stiglitz, J.E., 1977. Monopolistic competition and optimum product diversity. Am. Econ. Rev. 67 (3), 297–308.

Eaton, J., Kortum, S., 2002. Technology, geography, and trade. Econometrica 70 (5), 1741–1779.

Eaton, J., Kortum, S., Sotelo, S., 2013. International trade: Linking micro and macro. Advances in Economics and Econometrics, 1st ed Vol 2. Cambridge University Press, pp. 329–370.

Edmond, C., Midrigan, V., Xu, D.Y., 2015. Competition, markups, and the gains from international trade. Am. Econ. Rev. 105 (10), 3183–3221.

Epifani, P., Gancia, G., 2011. Trade, markup heterogeneity and misallocations. J. Int. Econ. 83 (1), 1–13.

Fajgelbaum, P.D., Khandelwal, A.K., 2016. Measuring the unequal gains from trade. Quart. J. Econ. 131 (3), 1113–1180.

Feenstra, R.C., Weinstein, D.E., 2017. Globalization, markups, and US welfare. Journal of Political Economy 125 (4), 1040–1074.

Feenstra, R.C., 2018. "Restoring the Product Variety and Pro-competitive Gains from Trade with Heterogeneous Firms and Bounded Productivity". Journal of International Economics 110, 16–27.

Head, K., Mayer, T., Thoenig, M., 2014. Welfare and trade without pareto. Am. Econ. Rev. (Papers and Proc.) 104 (5), 310–316.

Helpman, E., Melitz, M., Rubinstein, Y., 2008. Estimating trade flows: Trading partners and trading volumes. Quart. J. Econ. 123, 441–487.

Holmes, T.J., Hsu, W.-T., Lee, S., 2014. Allocative efficiency, mark-ups, and the welfare gains from trade. J. Int. Econ. 94 (2), 195–206.

Hsieh, C.-T., Klenow, P.J., 2009. Misallocation and manufacturing TFP in China and India. Quart. J. Econ. 124 (4), 1403–1448.

Hsieh, C.T., Ossa, R., 2016. A global view of productivity growth in China. J. Int. Econ. 102, 209–224.

Karabarbounis, L., Neiman, B., 2014. The global decline of the labor share. Quart. J. Econ. 129 (1), 61–103.

Krugman, P.R., 1979. Increasing returns, monopolistic competition, and international trade. J. Int. Econ. 9 (4), 469–479.

Levinsohn, J.A., Petrin, A.K., 2003. Estimating production functions using inputs to control for unobservables. Rev. Econ. Stud. 70 (2), 317–342.

Lipsey, R.G., Lancaster, Kelvin, 1956-1957. The general theory of second best. Rev. Econ. Stud. 24 (1), 11–32.

Lu, Y., Yu, L., 2015. Trade liberalization and markup dispersion: Evidence from China's WTO accession. Am. Econ. J. Appl. Econ. 7 (4), 221–253.

Markusen, J., 1981. Trade and the gains from trade with imperfect competition. J. Int. Econ. 11, 531–551.

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

Melitz, M.J., Ottaviano, G., 2008. Market size, trade, and productivity. Rev. Econ. Stud. 75 (1), 295–316.

Melitz, M.J., Redding, S.J., 2015. New trade models, new welfare implications. Am. Econ. Rev. 105 (3), 1105–1146.

Olley, G.S., Pakes, A., 1996. The dynamics of productivity in the telecommunications equipment industry. Econometrica 64, 1263–1298.

Ottaviano, G., Tabuchi, T., Thisse, J.-F., 2002. Agglomeration and trade revisited. Int. Econ. Rev. 43 (2), 409–435.

Restuccia, D., Rogerson, R., 2008. Policy distortions and aggregate productivity with heterogeneous plants. Rev. Econ. Dyn. 11, 707–720.

Robinson, Joan, 1934. The economics of imperfect competition. Macmillan.

Salvo, A., 2010. Inferring market power under the threat of entry: The case of the Brazilian cement industry. RAND J. Econ. 41 (2), 326–350.

Schmitz, Jr J.A., 2005. What determines productivity? Lessons from the dramatic recovery of the U.S. and Canadian iron ore industries following their early 1980s crisis. J. Polit. Econ. 113 (3), 582–625.

Simonovska, I., Waugh, M.E., 2014. Trade models, trade elasticities, and the gains from trade. NBER Working Paper w20495.

Trefler, D., 2004. The long and short of the Canada-U.S. free trade agreement. Am. Econ. Rev. 94 (4), 870–895.

Weinberger, A., 2015. Markups and misallocation with trade and heterogeneous firms. Globalization and Monetary Policy Institute Working Paper No. 251. Federal Reserve Bank of Dallas.

Zhelobodko, E., Kokovin, S., Parenti, M., Thisse, J.F., 2012. Monopolistic competition: Beyond the constant elasticity of substitution. Econometrica 80 (6), 2765–2784.