8-2017

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

Wen-Tai HSU
Singapore Management University, WENTAIHSU@smu.edu.sg

Yi LU

Guiying Laura WU

SMU ECONOMICS & STATISTICS

Competition, Markups, and Gains from Trade: A Quantitative Analysis of China Between 1995 and 2004

Wen-Tai Hsu, Yi Lu, Guiying Laura Wu

August 2017

Paper No. 12-2017

Competition, Markups, and Gains from Trade: A Quantitative Analysis of China Between 1995 and 2004

Wen-Tai Hsu Yi Lu Guiying Laura Wu*†

August 2, 2017

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 9.4%. The gains are decomposed into a Ricardian component and two pro-competitive ones. The pro-competitive effects account for 25.4% of the total gains. Moreover, the total gains from trade are $17-27\%$ larger than what would result from the formula provided by ACR (Arkolakis, Costinot, and Rodriguez-Clare 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. One methodological advantage of this paper's quantitative framework is that its application is not constrained by industrial or product classifications; thus it can be applied to countries of any size.

*Hsu: School of Economics, Singapore Management University. 90 Stamford Road, Singapore 178903. Email: wentaihsu@smu.edu.sg; Lu: School of Economics and Management, Tsinghua University, Beijing 100084, China. Email: luyi@sem.tsinghua.edu.cn; Wu: Division of Economics, School of Humanities and Social Sciences, Nanyang Technological University. 14 Nanyang Drive, Singapore 637332. Email: guiying.wu@ntu.edu.sg.

†For their helpful comments, we thank Costas Arkolakis, Issac Baley, Pao-Li Chang, Davin Chor, Jonathan Eaton, Tom Holmes, Chang-Tai Hsieh, Nicolas Jacquet, Sam Kortum, James Markusen, Sanghoon Lee, Hong Ma, Andres Rodriguez-Clare, Tom Sargent, Michael Zheng Song, Bob Staiger, Ping Wang, Yong Wang, Michael Waugh, and Daniel Yi Xu. We also thank Lianming Zhu, Yunong Li, and Xuan Luo for their excellent research assistance. Hsu gratefully acknowledge the financial support provided by the Sing Lun Fellowship of Singapore Management University.

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 a 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, we conduct 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 transport infrastructure was rapidly expanded, but joining World Trade Organization (WTO) in 2001 also drastically reduced trade barriers.[2] Recently, Brandt, Van Biesebroeck, Wang and Zhang (2012) 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.

To appreciate what we do, it is important to understand an ongoing debate regarding pro-competitive effects. It starts with Arkolakis, Costinot, and Rodriguez-Clare (2012; henceforth ACR), who show that for a class of influential trade models, welfare gains from trade (W'/W) can be simply calculated by $(v'/v)^{1/\epsilon}$, where v is domestic expenditure share, and ϵ is the trade elasticity. As both v and ϵ depend on trade flows, trade flows provide sufficient information regarding gains from trade. However, this class of mod-

[1]The idea of allocative efficiency dates back to Robinson (1934, Ch. 27) and Lipsey et al. (1956-57).

[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%.

els features no pro-competitive effects. To investigate pro-competitive effects, Edmond, Midgrigan, and Xu (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. On the other hand, Arkolakis, Costinot, Donaldson, and Rodriguez-Clare (2016) investigate the same issue in a monopolistic competitive model with a general preference that allows variable markup, and they find that pro-competitive effects are "elusive". What causes the difference? It seems market structure could play an important role.

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 (2016). As 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, Eaton, Jensen and Kortum (2003; henceforth BEJK). To help understand, we note three features of BEJK. First, the productivity of firms is heterogeneous and follows a Frechét 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 the trade cost. Later, Holmes, Hsu and Lee (2014) find that this invariance is due to the assumption that the productivity distribution is fat-tailed (Frechét). 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 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

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

such changes.[4]

We thus adopt Holmes et al. (2014) with the productivity draws from log-normal distributions. The log-normal distribution has been widely used in empirical applications; in particular, Head, Mayer, and Thoenig (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 the 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 can gauge the effect of "factual improvement in openness" by examining the effect of changing trade cost 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.

In Section 4, we gauge the gains from trade via various angles. First, we conduct a counter-factual analysis based on 2004 estimates with the trade cost reverted back to the level estimated using 1995 data to gauge the gains from the improved openness in this period. The gain is 9.4% of real income, and the contribution of the pro-competitive effects is 25.4%. The improvement of allocative efficiency accounts for the bulk of pro-competitive effects at 22.3%, whereas the relative markup effect accounts for the remaining 3.1%. The overall gains at 9.4% seems a relatively large number compared with those found in the literature, but this is partly due to the large reduction in trade cost during this period (from an iceberg cost of 2.31 to 1.66). Also, as shown by ACR, a smaller trade elasticity implies larger gains from trade. Simonovska and Waugh (2014b) 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 trade elasticity and hence larger gains. By accounting for markup dispersion in the data, our quantification also entails smaller trade elasticities, which also contributes to the larger gains.

The more intriguing finding is that even given trade elasticity local to the estimated models, the gains from trade are larger than those calculated using the ACR formula by 24.3% in 1995 and by 17.1% in 2004. For large change in trade cost, we compare with the ACR formula by integrating the local formula because trade elasticity is a variable in our model. In this case, the total gains from trade are 27.0% larger than the ACR formula. We investigate the reasoning behind this, and prove that pro-competitive effects are precisely

[4]Similarly, Feenstra (2014) 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.

the extra gains 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. (2017), Feenstra et al. (2016) and many others)[5] 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 various 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 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 may be hundreds of firms so that firms' market shares are typically much smaller compared with a similar data set for a small country. The problem here is that when firms' market shares are "diluted" by country size for a

[5]There is an extensive literature exploring properties of markups under monopolistic competition; see, for example, Dixit and Stiglitz (1977), Krugman (1979), Ottaviano, Tabuchi and Thisse (2002), Melitz and Ottaviano (2008), Behrens and Murata (2012), Zhelobodko, Kokovin, Parenti, and Thisse (2012), Feenstra (2014), Weinberg (2015), Feenstra and Weinstein (2016), and Dhingra and Morrow (2016).

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.

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 (2014b) 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 trade flows do not necessarily provide sufficient information for welfare. Melitz and Redding (2015) also show that the trade elasticity becomes a variable and trade flows do not provide sufficient information for welfare when the distribution of productivity deviate from untruncated Pareto in Melitz (2003). Obviously, their mechanism is different from ours.[6]

Our work is also related to de Blas and Russ (2012) and Goldberg, De Loecker, Khandelwal and Pavcnik (2015), who provide analyses of how trade affects the distribution of markup. But these papers do not address welfare gains from trade. By looking at alloca-

[6]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, Levchenko, and Zhang (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.

tive 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 $i = 1, 2$.[7] 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

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

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

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

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

[7]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 can not use a multiple-country framework because we do not have access to firm-level data in multiple countries.

country j are given by

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

and $\phi_{j\omega} \equiv \left(\frac{p_{j\omega}}{P_j}\right)^{1-\sigma}$ is country j's spending share on the good ω.

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

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

where $\varphi_{\omega,ik}$ is the Hicks-neutral productivity of firm $k \in \{1, 2, ..., n_{\omega,i}\}$, $n_{\omega,i}$ is the number of entrants in country i for good ω, and $\ell_{\omega,ik}$ 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.[8] 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

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

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

[8] Eaton, Kortum and Sotelo (2013) also model finite number of firms as a Poisson random variable, but for a very different purpose.

entry explicitly. By this probabilistic formulation, we let λ_i summarize the entry effort in each country. From (1), we see that the larger the mean numbers of firms λ_i, the larger the $\bar{\omega}$.

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.[9]

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

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

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 $\bar{p}_{j\omega} = \frac{\sigma}{\sigma-1} a^*_{j\omega}$. 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.

$$p_{j\omega} = \min\left(\bar{p}_{j\omega}, a^{**}_{j\omega}\right) = \min\left\{ \frac{\sigma}{\sigma-1} a^*_{j\omega}, a^{**}_{j\omega} \right\}. \tag{2}$$

The markup of good ω at j is therefore

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

Note that firms' markups may differ from the markups for consumers. A non-exporter's markup is the same as the one 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 con-

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

stant returns to scale,

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

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. Then, we have

$$\begin{aligned} M_i^{sell} &= \frac{R_i}{w_i L_i}=\frac{\int_{\left\{\omega:\,\chi_1^*(\omega)=i\right\}}\phi_{1\omega}R_1 d\omega+\int_{\left\{\omega:\,\chi_2^*(\omega)=i\right\}}\phi_{2\omega}R_2 d\omega}{\int_{\left\{\omega:\,\chi_1^*(\omega)=i\right\}}m_{1\omega}^{-1}\phi_{1\omega}R_1 d\omega+\int_{\left\{\omega:\,\chi_2^*(\omega)=i\right\}}m_{2\omega}^{-1}\phi_{2\omega}R_2 d\omega} \\ &= \left(\int_{\left\{\omega:\,\chi_1^*(\omega)=i\right\}}m_{1\omega}^{-1}\frac{\phi_{1\omega}R_1}{R_i}d\omega+\int_{\left\{\omega:\,\chi_2^*(\omega)=i\right\}}m_{2\omega}^{-1}\frac{\phi_{2\omega}R_2}{R_i}d\omega\right)^{-1}, \end{aligned} \tag{3}$$

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:

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

Let the inverses of markups be called cost shares, as they are the shares of costs in revenues. A harmonic mean of markups is the inverse of the weighted arithmetic mean of cost shares. Harmonic means naturally appear here precisely because the weights are revenue. 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 for our empirical work.

2.4 Wages and General Equilibrium

Labor demand in country i from a non-exporter that produces input ω is

$$\ell_{\omega,i}=\frac{q_{i\omega}}{\varphi_{\omega,i}^*}=\frac{1}{\varphi_{\omega,i}^*}\frac{R_i}{P_i}\left(\frac{p_{i\omega}}{P_i}\right)^{-\sigma}.$$

For an exporter at i, its labor demand is

$$\begin{aligned}
\ell_{\omega,1} &= \frac{q_{1\omega}+\tau q_{2\omega}}{\varphi^*_{\omega,1}} = \frac{1}{\varphi^*_{\omega,1}}\left[\frac{R_1}{P_1}\left(\frac{p_{1\omega}}{P_1}\right)^{-\sigma} + \frac{\tau R_2}{P_2}\left(\frac{p_{2\omega}}{P_2}\right)^{-\sigma}\right] \\
\ell_{\omega,2} &= \frac{\tau q_{1\omega}+q_{2\omega}}{\varphi^*_{\omega,2}} = \frac{1}{\varphi^*_{\omega,2}}\left[\frac{\tau R_1}{P_1}\left(\frac{p_{1\omega}}{P_1}\right)^{-\sigma} + \frac{R_2}{P_2}\left(\frac{p_{2\omega}}{P_2}\right)^{-\sigma}\right].
\end{aligned}$$

Labor market clearing in country i is

$$\int_{\omega\in\chi_i} \ell_{\omega,i}\, d\omega = L_i, \tag{4}$$

where χ_i is the set of ω produced at i.

To calculate the trade flows, observe that the total exports from country i to country j is

$$R_{j,i} = \int_{\left\{\omega:\, \chi^*_j(\omega)=i\right\}} E_{j\omega}\, d\omega = R_j \int_{\left\{\omega:\, \chi^*_j(\omega)=i\right\}} \left(\frac{p_{j\omega}}{P_j}\right)^{1-\sigma} d\omega. \tag{5}$$

where $\chi^*_j(\omega) \in \{1,2\}$ denotes the source country for any particular good ω at destination j. The balanced trade condition is therefore

$$R_{2,1} = R_{1,2}. \tag{6}$$

We choose country 1's labor as numeraire, and hence $w_1 = 1$, and $w \equiv w_2$ is also the wage ratio. Given $\{w, R_1, R_2\}$, the realization of $n_{i,\omega}$ for each i and ω, and the realization of $\{\varphi_{\omega,ik}\}$ for each firm $k \in \{1,2,...,n_{i,\omega}\}$, pricing, markups, consumption decisions, labor demand, and trade flows are all determined as described above. The two labor market clearing conditions in (4) and the balanced trade condition (6) thus determine $\{w, R_1, R_2\}$. For easier computation for our quantitative work, we use an algorithm of equilibrium computation that reduces the above-mentioned system of equations to one equation in one unknown. We describe such an algorithm in Appendix A1.

Similar to the literature, our benchmark model and estimation are based on the assumption of balanced trade. Nevertheless, we also gauge the robustness of our results by investigating the case where trade imbalance is allowed. See Section 4.5 for details.

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:

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

where $\tilde{\mathbf{q}}^a_j = \left\{ \tilde{q}^a_{j\omega} : \omega \in [0, \bar{\omega}] \right\}$ is the expenditure-minimizing consumption bundle that delivers one unit of utility. Total welfare is defined as real income R_j/P_j. As the product of producers' aggregate markup and labor income entails total revenue (3), 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

$$W_j^A \equiv \frac{A_j \times M_j^{buy}}{P_j} = \frac{\int_0^{\bar{\omega}} a^*_{j\omega} \tilde{q}^a_{j\omega} d\omega}{\int_0^{\bar{\omega}} a^*_{j\omega} \tilde{q}_{j\omega} d\omega} \leq 1. \tag{7}$$

The inequality follows from the fact that under marginal cost pricing, $\tilde{q}^a_{j\omega}$ 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 only requires homothetic preference and is thus applicable to all market structures.[10]

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 pointed out 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).[11] 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 $\epsilon_j^{ii'} \equiv$

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

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

$\partial \ln (X_{ij}/X_{jj}) / \partial \ln \tau_{i'j}$ 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 $\epsilon_j^{ii'}$ 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 $i \neq j$,

$$\epsilon_j^i = \frac{\partial \ln \left(\frac{X_{ij}}{X_{jj}}\right)}{\partial \ln \tau_{ij}} = \frac{\partial \ln \left(\frac{1-v_{jj}}{v_{jj}}\right)}{\partial \ln \tau_{ij}}. \tag{8}$$

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,

$$\begin{aligned} d \ln W_j &= -\left(v_{ij} \frac{d \ln v_{ij} - d \ln v_{jj}}{\epsilon_j^i} + v_{jj} \frac{d \ln v_{jj} - d \ln v_{jj}}{\epsilon_j^i}\right) \\ &= \frac{1}{\epsilon_j^i} d \ln v_{jj}. \end{aligned} \tag{9}$$

where the last line uses $v_{ij} + v_{jj} = 1$, which implies that $v_{ij} d \ln v_{ij} + v_{jj} d \ln v_{jj} = 0$.[12] 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

$$d \ln W_j^{ACR} = \frac{1}{\epsilon} d \ln v_{jj}. \tag{10}$$

Moreover, the global formula $W_j'/W_j = \left(v_{jj}'/v_{jj}\right)^{\frac{1}{\epsilon}}$ holds when R3 holds. We repeat the derivation in ACR in (9) 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 $\tau_{21} = \tau_{12} = \tau$.[13]

Now, back to our model, and we examine how productive efficiency in our model is related to the ACR formula. As $W_j^{Prod} = 1/A_j$, the price index under marginal cost pricing, ACR's proof of Proposition 1 for the perfect competition case actually applies up to Step

[12]The expression in (9) 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 $d \ln N_j = 0$.

[13]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.

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 (9), for any $i \neq j$, the above implies

$$d\ln W_j^{Prod} = -d\ln A_j = \frac{1}{\tilde{\epsilon}_j^i} d\ln \tilde{v}_{jj}. \tag{11}$$

Note that the ACR formula (10) should be applied using actual trade flow to calculate trade elasticity and domestic expenditure share (that is, actual pricing (2) should be used), whereas (11) 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:

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

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, $\tilde{v}_{jj} = v_{jj}$. By (8), $\tilde{\epsilon}_j^i = \epsilon_j^i$. In this case, $d\ln W_j^{ACR} = d\ln W_j^{Prod}$ with the trade elasticity being ϵ_j^i. But, as ϵ_j^i varies with trade shock $d\boldsymbol{\tau}$, where $\boldsymbol{\tau} = \{\tau_{ij}\}$, the global ACR formula does not apply.

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 $\boldsymbol{\tau}$, the change in the productive efficiency W_j^{Prod} can be expressed as*

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

where $\tilde{\epsilon}_j^i$ and $\tilde{v}_{jj}$ are 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 $d\ln W_j^{\text{ACR}} = d\ln W_j^{\text{Prod}}$.

In the case of $\sigma = 1$, this proposition says that for infinitesimal changes in $\boldsymbol{\tau}$, 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. Our quantitative analysis in Section 4.2 reveals that in the general case of $\sigma > 1$, $d\ln W_j^{ACR}$ is still relatively close to $d\ln W_j^{Prod}$; therefore the total gains $d\ln W_j^{Total}$ 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 $(\bar{\imath})$ 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 export because import is what is needed for the ACR formula. To look at pro-competitive effects, we look at only 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-deterence example mentioned in the introduction. In Case $(2,1)$, a lower τ increases the markup but does not affect the price and import because the foreign supplier is only constrained 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.

The above table shows how head-to-head competition separates markups and import. 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 of 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. (2017) 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

We use the following two steps to quantify the model. First, we estimate the markup distribution and infer the elasticity of substitution from such distribution. Then, given σ, measures of $\{w, R_1, R_2\}$, we use the moments of markups, trade flows, number of firms and fraction of exporters to estimate the remaining parameters by SMM. 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 non-exporters can help identify the different parameters of the two countries.

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.[14] 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

[14]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.

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 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,[15] 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.[16] The estimation of firm-level markup hinges on choosing an input X that is free of any adjustment costs, and the estimation of its output elasticity 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 output 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 A2.

Table 1 gives summary statistics of the markup distribution,[17] with breakdowns in each year and between exporters and non-exporters. Observe that the (unweighted)

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

[16] 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 Table A1 in the appendix. Interestingly, EMX also found similar results using Taiwanese firm-level data.

[17] Following the literature, e.g., Goldberg, De Loecker, Khandelwal and Pavcnik (2015) 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.

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 non-exporters' standard deviation is larger than the increase in exporter's standard deviation, 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 because allocative efficiency depends on consumers' markups rather than firms' 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 Elasticity of Substitution

As a preference parameter, we infer a common elasticity of substitution σ for both years. Note that the model implies that $m \in \left[1, \frac{\sigma}{\sigma-1}\right]$, 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.

As we examine the effects of markups, we infer σ using the upper bound of the markup distribution. Considering the possibility of measurement errors and outliers, we equate $\sigma/(\sigma-1)$ to the 99th percentile of the estimated markup distribution (using the pooled sample from 1995–2004). We obtain $\sigma = 1.40$, which reflects that the 99th percentile is around 3.5.[18] This calibrated $\sigma = 1.40$ is strikingly similar to the estimate of the same parameter (1.37) in Simonovska and Waugh (2014b) with the optimal weighting matrix in their method of moments procedure.

The inferred σ here is quite different from those estimates in models that feature constant markups (often a CES preference coupled with either monopolistic competition or

[18]Note that this estimate of σ is not sensitive to sample size. In our multi-sector exercise, σ_s is separately inferred for each sector s using the markup distribution of that sector. The unweighted mean of σ_s is 1.44, and 23 out of 29 σ_s are within one standard deviation from the mean, $(1.27, 1.61)$. See Section 5.1.

perfect competition). This is essentially because $\sigma/(\sigma-1)$ 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.

Note that in BEJK, the trade elasticity is given the tail index of the Frechét distribution, and is independent of the elasticity of substitution σ. In our model where the productivity draws deviate from Frechét, 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 (2014b) under the BEJK model.

3.4 Simulated Method of Moments

3.4.1 Method

We estimate the remaining parameters using SMM for 1995 and 2004 separately. It is important to allow all parameters to vary between the two years so that the changes in the environment of the Chinese economy can be reflected. 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 remaining parameters 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

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. In order to use SMM to estimate these seven parameters, we need at least seven moments. We use the following 12 moments: the import and export shares; relative number of firms; fraction of exporters; weighted mean and standard deviation of cost shares for both exporters and non-exporters; and the median and 95th percentile of cost shares for exporters and non-exporters.[19]

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

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

Divide both sides by $\bar{N}$, a large number that is chosen for normalization. The moment we use is the relative number of Chinese firms:

$$\frac{N_1}{\bar{N}} = \frac{\gamma\left(1 - e^{-\lambda_1}\right)}{\bar{N}} \times \Pr\left[\frac{1}{\varphi^*_{1\omega}} < \frac{w\tau}{\varphi^*_{2\omega}}\right], \tag{12}$$

The choice of $\bar{N}$ does not affect the estimates, but we must choose the same $\bar{N}$ for both 1995 and 2004 in order to gauge the increase in γ. For this purpose, we choose $\bar{N}$ to be 2 million.[20]

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.

Note that our model structure implies that all of the above-mentioned moments can be simulated for a given set of parameters and observed macro variables $\{w, R_1, R_2\}$. Thus, our SMM procedure essentially searches for the set of parameters that are most consistent with the observed micro moments given observed macro variables $\{w, R_1, R_2\}$.

[19] 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.

[20] Recall that we have $1,324,752$ firms in our 2004 data. Also, larger $\bar{N}$ generates a larger number of goods (and hence draws from both distributions in the model) in simulation, and the law of large numbers helps to generate more precise moments.

[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 a n-country model with $n \geq 2$ unless one can gather firm-level data for all of these countries, which is a daunting task.

One advantage of this approach is that we do not have to measure or estimate the difficult object country sizes $\{L_1, L_2\}$ as this may well depend on human capital in addition to population sizes and other things.[22] After we obtain parameter estimates, we simulate the model given observed macro variables $\{w, R_1, R_2\}$ to impute $\{L_1, L_2\}$ by simulating the labor demand across all goods and using the labor-market clearing conditions. Then, our counter-factual is based on the parameter estimates and the imputed $\{L_1, L_2\}$, and of course the macro variables $\{w, R_1, R_2\}$ are endogenous in counter-factuals.

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.[23] We then calculate 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).[24] 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.

3.4.2 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. For each year (1995 or 2004), we impute $\{L_1, L_2\}$ given the parameter estimates and observed macro variables $\{w, R_1, R_2\}$. Then, under parameter estimates and the imputed $\{L_1, L_2\}$, we simulate $\{w, R_1, R_2\}$. Note that the simulated $\{w, R_1, R_2\}$ and the observed ones need not be the same. Nevertheless, the bottom three rows in Table 2 show that they turn out to be quite close,[25] 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 1.66. The measure of goods γ more than triples from 0.26 to 0.85. This basically re-

[22]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.

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

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

[25]Here, the largest discrepancy between data values and simulated value is the total revenue of the ROW in 1995, which is about 10.5%. For all the other numbers, the discrepancies are all less than 5.2%.

flects 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 2.44 to 2.61, whereas in the ROW it increased from 5.29 to 5.83. China's mean log productivity (μ_1) relative to the ROW increased from -2.40 to -1.79. These numbers are negative, meaning that China's productivity is lower than that of the ROW (μ_2 is normalized to 0). Also, we see a slight decrease in the dispersion parameter of the productivity distribution in both countries (η_1, η_2). Interestingly, the productivity dispersion is larger in China than in the ROW, which is consistent with the finding by Hsieh and Klenow (2009).[26]

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, and 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.

Trade cost τ affects almost all moments significantly, and 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 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 λ_1 and λ_2, the 95th percentiles of markups and the relative number of active firms are crucial in identifying these two parameters, with the trade moments playing some role as well. The intuition is as follows. Fixing other parameters, when λ_i increases, the number of entrants per product in country i increases. Due to the non-fat-tailed nature of the productivity distribution, the ratio between the top two draws is narrowed, but since this

[26]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 6.9%. This impressive growth rate is actually similar to the 7.96% estimated by Brandt, Van Biesebroeck and Zhang (2012). Note that the 6.9% 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, Van Biesebroeck, Wang, and Zhang (2012) 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.

ratio is indeed the markup and since this is particularly pronounced for the top markups, the 95th percentiles are particularly useful in identifying these two parameters. The fact that we observe increases in λ_i during this period may reflect that the 95th percentiles of markups decrease during this period. Intuitively, the relative number of (active) Chinese firms is also useful for identifying λ_1, as seen clearly in (12).[27]

For the measure of goods γ, it is obvious that 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 moments of markups, including both means and standard deviations of the cost shares, than μ_1. Moreover, 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 almost always affect 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, number of firms and fraction of exporters. It decreases Chinese non-exporters' median and 95th percentile markups, but increases those of Chinese 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

[27]Trade flows are also useful, as an increase in λ_1 raises active firms' productivities in China, increasing the export share and reducing the import share. On the other hand, an increase in λ_2 raises active firms' productivities in the ROW, increasing the import share and reducing the export share in China.

Figure 1) and the increasing number of manufacturing firms are also consistent with the above-mentioned reforms.

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 an inhibitive value so that the economy becomes autarky.

The results are shown in Table 4. The welfare gains of changing τ from 1995's level to 2004's level are 9.4%, in which allocative efficiency accounts for 22.3% $(2.1/9.4)$ and relative markup effect accounts for 3.1% $(0.3/9.4)$. Thus, these pro-competitive effects jointly account for 25.4% 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 larger. Overall, although the relative markup effect is positive, it is relatively small, whereas the combined effect can account for about a quarter of the total gains. The total gains from autarky to 2004's τ are, of course, much larger, at 33.4%, but the decomposition is similar to the first analysis.

For the intuition regarding the source of the gains due to allocative efficiency, recall the six cases of markups facing Chinese consumers distinguished in Section 2.6. A change in trade cost affects markups only in cases $(1, 2)$ and $(2, 1)$. Trade liberalization depresses markups in case $(1, 2)$, but increases markups in case $(2, 1)$. Because foreign firms typically face fiercer competition due to the trade barrier, case $(1, 2)$ has on average higher markups relative to case $(2, 1)$; thus trade liberalization reduces markup dispersion and enhances allocative efficiency.

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 to 10%	3.1	1.9	1.2	38.7%
10% to 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 only includes 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 5.5% to 1.5%. 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 24%, which is quite close to the results reported in Panel A.

Looking at both panels together, the contribution of pro-competitive effects range from 23.3% to 27.6%, and the contribution of allocative efficiency ranges from 22.3% to 24.6%. 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, we 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.[28]

The trade elasticities local to our estimated model in 1995 and 2004 are -2.48 and -3.23; these are surprisingly close to the estimates of trade elasticity under the BEJK model in Simonovska and Waugh (2014b), which range from -2.74 to -3.32.[29] Looking at 2004 first, the welfare elasticity of trade cost is 0.409, meaning a 1% reduction in trade cost τ induces a 0.409% increase in real income; 19.3% of this elasticity is from pro-competitive effects. The ACR formula entails a welfare elasticity of 0.349, which is larger than but still relatively close to the elasticity of productive efficiency 0.330, as compared to its distance to the total elasticity. As a result, the totals gains from trade are larger than the gains predicted by the ACR formula by 17.1%, and most of these extra gains are from pro-competitive effects. In the case of the 1995 model, the contribution of pro-competitive effects and the additional gains over the ACR formula are slightly larger at 22.0% and 24.3%, respectively. Note that the contributions of pro-competitive effects in Table 5 are still relatively similar in magnitude to those reported in Table 4.

The total gains from trade in the benchmark result in Table 4, 9.4%, might seem large compared with many previous studies, but this is partly due to the reduction in τ being large between the two data years. 2004's τ is 28% off 1995's τ, and thus the two-point welfare elasticity to trade cost is $9.4/28 = 0.336$, which falls between the two welfare elasticities (0.249 and 0.409) in Table 5.

Local to each model at 1995 and 2005, the two channels for larger gains from trade are as follows. First, similar to Simonovska and Waugh (2014b), the trade elasticities in our estimated models with micro-level data are lower than those with only macro data such as trade flows. According to ACR, lower trade elasticities also lead to larger gains. Second, Proposition 1 and Table 5 show that given the same trade elasticity, head-to-head competition leads additional gains over the ACR formula. If the trade elasticity were an often adopted one, say 4 in Simonovska and Waugh (2014a), then the ACR welfare elasticity would be 0.124 and 0.281 in 1995 and 2004, respectively. Thus for 1995, 0.125 ($0.249 - 0.124$) is the gain due to the first and second above-said channels combined. The relative weight of these first and second channels are 61% ($\frac{0.2-0.124}{0.249-0.124}$) and 39%, respectively; that is, the difference in trade elasticity (-4 vs -2.74) contributes 61% of the two channels combined, whereas pro-competitive effects account for the remaining 39%. For 2004, these relative weights are 53% and 47%, respectively. Thus, we conclude that pro-competitive effects do contribute significantly to the larger gains in this model.

[28]Here, we set we set $h = 0.001$, 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.

[29]See Table 4 and 7 in their paper.

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. We relegate the calculation details to Appendix A3. The result is that the gains from trade according to the ACR formula are 7.4%. As the total gains from trade 9.4% (Table 4), they are 27.0% higher than the ACR formula. If the trade elasticity were a constant at -4, then the ACR formula predicts a welfare gain of 5.3%.[30] Thus, the first channel (due to difference in trade elasticity) accounts for 51.2% ($\frac{7.4-5.3}{9.4-5.3}$) of the additional gains, whereas the second channel (due to head-to-head competition) accounts for the remaining 48.8%. These numbers are similar to those found in the first comparison.

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 less data requirement. Nevertheless, ignoring cross-country differences may miss 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 although the standard errors here are somewhat smaller than those in the benchmark estimation, 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 misses out the general equilibrium effect in the adjustment of relative wages, which change from 10.5 to 5.3 (See Table 2), meaning that Chinese wages relative to the ROW almost double in this decade.

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. 2.7% versus 9.4%).

[30] Under the 2004 parameters, domestic consumption share drops from 0.9065 to 0.7368 when trade cost τ falls from the 1995 level to the 2004 level.

Both components also become much smaller. However, the contributions of the pro-competitive effects are still somewhat close to those at the benchmark case, although the variation is somewhat larger. 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 the autarky, we see that the allocative efficiency is much larger in the symmetric-country case than in the benchmark case (0.941 versus 0.897). As the allocative efficiency is larger to start with, it is not surprising that the gains in allocative efficiency are smaller (0.6% versus 2.1% and 2.5% versus 7.5%). The same rationale explains why we see a pronounced diminishing-returns (dropping from 32.5% to 13.3%) pattern in Panel B that is absent in the asymmetric-country case.

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 Effects of Other Parameters

In the previous subsections, we focus on the counter-factual exercises of trade cost. In order to understand better the mechanism of the model and welfare results, here we study the comparative statics of a variety of parameters.

First, we consider a closed economy and ask 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 capital welfare is $W_j^{PC} \equiv W_j^{Total}/L_j$ and that there is no markup level 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.[31] The gains in allocative efficiency here are relatively modest compared with the gains due to enhanced productivity.

	L and λ doubles	η doubles
$d\ln W_1^{PC}$	0.41	0.32
$d\ln W_1^{Prod}$	0.35	0.35
$d\ln W_1^{A}$	0.06	-0.03

In the case where the standard deviation of log-productivity doubles, both per capita welfare and productivity 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 open economy, and consider symmetric countries for clarity. We have seen the effect of trade liberalization in symmetric-country case in the previous subsection. Here, we want 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 8), as well as under different levels of λ (Panel B of Table 8).

The middle columns of both panels are the same as that reported in the 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 60% at 0.5 times η_0, the standard deviation at 2004, to as small as 14% at 1.5 times η_0.

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

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.

4.5 Robustness

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. In our first robustness check, we 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. For our third check, recall that we used the 99th percentile of the markup distribution to infer σ, but now we also report results based on the 97.5th percentile. In our fourth check, we return to the benchmark but allow for trade imbalance in the model.

The results are reported in Table 9. 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 between 5.8% and 7.7%, and the contribution of pro-competitive effects ranges from 19.4% to 27.3%, and that of allocative efficiency ranges from 19.0% to 22.4%. These indicate that the importance of pro-competitive effects remains similar, and the allocative efficiency still accounts for the bulk of gains from trade except the case with trade imbalance.

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. Both the overall gains and the pro-competitive effects are smaller in the first robustness check than in the benchmark. As China had smaller productivity and smaller entry in 1995, this indicates a complementary effect between trade liberalization and other fundamentals in the sense that there are more gains from trade when productivity and entry are higher.

Next, note that the σ inferred from raw markups is about 1.67, 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 effects are slightly less important under raw markups. This also explains why using the 97.5th percentile of the markup distribution to infer σ also induces smaller pro-competitive effects.

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. To accommodate trade imbalance, we follow the literature by allowing an exogenous trade deficit D_i for each country i with the requirement that $D_1 + D_2 = 0$.[32] 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 A4. The contribution from pro-competitive effects remain similar to the benchmark results in Table 4. Even though the contribution from allocative efficiency remain sizable at 19%, the relative markup effect accounts for 8.3% of the total welfare gains, larger than that in all of the previous cases. The intuition is that *compared with the case with balanced trade*, trade surplus/deficit makes the wage ratio w higher, making the marginal cost of foreign firms more expensive and hence softening competition that domestic firms face. Thus, Chinese consumers face a larger aggregate markup, whereas Chinese firms' aggregate markup is lower because the increase in foreign income ($R_2 + D$) increases the weight on the markups earned in the foreign market, which are lower than domestic ones. This difference causes the W^R term to be significantly lower than 1, as compared with the case with balanced trade. With smaller trade cost, the gap between the two aggregate markups is smaller because w is closer to 1 and because the differences between exporters' domestic and foreign markups are smaller. Hence, trade liberalization brings a positive relative markup effect.[33]

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

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

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

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

Model Modification There are S sectors, which are indexed by $s = 1, 2, ..., S$. 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:

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

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

$$\begin{aligned} P_j &= \Pi_{s=1}^{S} \left(\frac{P_{js}}{\alpha_s} \right)^{\alpha_s} \\ P_{js} &\equiv \left(\int_0^{\bar{\omega}_s} p_{js\omega}^{1-\sigma_s} d\omega \right)^{\frac{1}{1-\sigma_s}}. \end{aligned}$$

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

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

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

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

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.

Wages and General Equilibrium The labor demand for a non-exporter at i that produces good $s\omega$ is

$$\ell_{s\omega,i} = \frac{q_{is\omega}}{\varphi^*_{s\omega,i}} = \frac{1}{\varphi^*_{s\omega,i}} \frac{\alpha_s R_i}{P_{is}} \left(\frac{p_{is\omega}}{P_{is}} \right)^{-\sigma_s}.$$

The labor demand for an exporter at $i = 1$ or 2 is

$$\begin{aligned}
\ell_{s\omega,1} &= \frac{q_{1s\omega} + \tau q_{2s\omega}}{\varphi^*_{s\omega,1}} = \frac{1}{\varphi^*_{s\omega,1}} \left[\frac{\alpha_s R_1}{P_{1s}} \left(\frac{p_{1s\omega}}{P_{1s}} \right)^{-\sigma_s} + \frac{\tau_s \alpha_s R_2}{P_{2s}} \left(\frac{p_{2s\omega}}{P_{2s}} \right)^{-\sigma_s} \right], \\
\ell_{s\omega,2} &= \frac{\tau_s q_{1s\omega} + q_{2s\omega}}{\varphi^*_{s\omega,2}} = \frac{1}{\varphi^*_{s\omega,2}} \left[\frac{\tau_s \alpha_s R_1}{P_{1s}} \left(\frac{p_{1s\omega}}{P_{1s}} \right)^{-\sigma_s} + \frac{\alpha_s R_2}{P_{2s}} \left(\frac{p_{2s\omega}}{P_{2s}} \right)^{-\sigma_s} \right].
\end{aligned}$$

Labor market clearing in country i is

$$\sum_{s=1}^{S} \int_{\omega \in \chi_{s,i}} \ell_{s\omega,i} d\omega = L_i,$$

where $\chi_{s,i}$ is the set of $s\omega$ produced at i.

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:

$$\phi_{j,i} = \sum_{s=1}^{S} \alpha_s \int_{\left\{ s\omega :\, \chi^*_j(s\omega) = i \right\}} \phi_{js\omega} d\omega. \tag{13}$$

The balanced trade condition $R_{2,1} = R_{1,2}$ holds in equilibrium.

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

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

$$M_i^{buy} = \left(\sum_{s=1}^S \alpha_s \left(M_{is}^{buy}\right)^{-1}\right)^{-1}, \qquad M_i^{sell} = \frac{R_i}{w_i L_i} = \left(\sum_{s=1}^S \frac{R_{s,i}}{R_i}\left(M_{is}^{sell}\right)^{-1}\right)^{-1}, \tag{14}$$

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

$$M_{is}^{sell} = \left(\int_{\{\omega:\, \chi_{s1}^*(\omega)=i\}} m_{1s\omega}^{-1} \frac{\alpha_s R_1 \phi_{1s\omega}}{R_{s,i}} d\omega + \int_{\{\omega:\, \chi_{s2}^*(\omega)=i\}} m_{2s\omega}^{-1} \frac{\alpha_s R_2 \phi_{2s\omega}}{R_{s,i}} d\omega\right)^{-1}. \tag{15}$$

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, $w_i L_i = \sum_s \frac{R_{is}}{M_{is}^{sell}}$.

Quantifying the Model To quantify the model, we focus on 29 2-digit manufacturing sectors in Chinese Industrial Classifications (CIC).[34] We first calibrate $\{\alpha_s\}_{s=1}^S$. Recall that $P_{1s}Q_{1s} = \alpha_s R_1$. We use information about expenditure share in China's 1997 and 2002 input-output table to calibrate α_{st}, where $t = 1997, 2002$. We then set α_s to be the average between two years.[35] We then follow the same procedure as in the one-sector economy case to infer the elasticity of substitution σ_s and estimate the remaining parameters by SMM using sectoral firm-level data. Note that one convenience in our framework is that to implement SMM, moments are generated given wages w and total revenue R_1 and R_2, and each sector is actually estimated separately, which largely simplifies the estimation and equilibrium computation for counter-factuals.

The parameter estimates are shown in Tables 10A and 10B. 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 all consistent with the pattern observed in the one-sector case, except for the parameter η_1 (see Table 2). In particular, all esti-

[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] Specifically, we first map the input-output code to 2-digit CIC sectors. Then, we calculate the expenditure share for each 2-digit CIC sector, where the expenditure is calculated by subtracting exports from total use, which already includes imports.

mated trade costs decrease except for Tobacco Processing.[36] Also observe that the mean σ_s is 1.44, which is quite close to our benchmark in the one-sector economy, and σ_s in most industries (23 out of 29) are within one standard deviation from the mean, $(1.27, 1.61)$.

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 11. 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 7.2% between 1995 and 2004 and 28.2% from autarky. The contribution of pro-competitive effects here is around 20%, which is slightly smaller than the numbers in Table 4. Similarly, allocative efficiency accounts for almost all of the pro-competitive effects.

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 (14) 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.

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 weighted average of the M_{1s}^{buy} are then 1.21 and 1.36, respectively. The corresponding weighted average of the changes in trade costs τ_s (i.e, $\Delta\tau_s = \tau_{s,2004} - \tau_{s,1995}$) are -0.446 and -0.856, respectively. An alternative measure of trade liberalization is the changes in sectoral import tariffs,[37] which directly relate to the WTO entry but do not account for other factors of trade liberalization. In this case, the corresponding changes are -0.162 and -0.215, respectively. These simple statistics show

[36] This is mainly because the import and export shares decrease from 0.021 and 0.052 to 0.010 and 0.016 in this sector.

[37] For details of how the sectoral import tariffs are calculated, see Appendix A2.

a tendency where the higher the initial level of sectoral markups, the larger the reduction in trade costs (or import tariffs).[38]

Columns 1 and 5 of Table 12 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 try 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 11) and the results in this subsection.

[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 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 4, which is marginally insignificant (with a p-value at 0.11). Nevertheless, the value of this coefficient is similar to those in columns 1-3. Also, as the sample size is small (29), one should use caution when interpreting the significance levels.

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 9.4%. The pro-competitive effects account for 25.4% of the total gains from trade from 1995 to 2004 and 23.3% from autarky to 2004. Allocative efficiency plays a much more important role than the relative markup effect.

Local to the estimated models in 1995 and 2004, the total gains from trade are larger than the gains predicted by the ACR formula by 17% and 24%, respectively. The total gain from the change in trade cost between 1995's and 2004's levels is 27% larger than the ACR formula. These additional gains are mostly from pro-competitive effects. This is a result that is absent in models when a firm monopolizes a variety, such as in Arkolakis et al. (2017), EMX, Feenstra and Weinstein (2016), and other monopolistic competitive models. Head-to-head competition is the main reason driving this difference, as detailed in Sections 2.6 and 4.2.

Overall speaking, total gain from trade is relatively large compared with other estimates in the literature. Besides 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 in a similar fashion to Simonovska and Waugh (2014b). We find that both channels are important, as pro-competitive effects account for 39% and 47% of the two channels combined at 1995 and 2004, respectively. Similarly, for the change of trade cost from 1995's to 2004's levels, pro-competitive effects account for 49% of the two channels combined.

When comparing with the symmetric-country case, we find that the gains from trade and its components are substantially smaller in the symmetric-country case, indicating the important role played by the differences in productivities and markups. 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 (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. In particular, from autarky to a fully integrated world, λ 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.

Appendix

A1. Algorithm of Computing Equilibrium

We describe a procedure that reduces three equilibrium conditions in three unknowns $\{w, R_1, R_2\}$ to one equation in one unknown w. This is useful for faster computation.

One-Sector Economy First, observe from the definition of the producers' aggregate markup for country 1:

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

in which the second line uses the balanced trade condition $\frac{R_2}{R_1} = \frac{\phi_{1,2}}{\phi_{2,1}}$, where $\phi_{j,i}$ denote the total spending share of country j's consumers on good from country i. Note that $\phi_{j,i} = \int_{\{\omega:\, \chi_j^*(\omega)=i\}} \phi_{j\omega} d\omega$ only depends on relative wage w, but not on R_1 and R_2. Hence, M_1^{sell} becomes a function of w only. For any given w, we can calculate $M_1^{sell}(w)$. Then, given $w_1 = 1$ and L_1, we get $R_1(w) = M_1^{sell}(w) L_1$. For R_2, we use the balanced trade condition again:

$$R_2(w) = \frac{\phi_{1,2}}{\phi_{2,1}}(w) \times R_1(w).$$

In fact, $M_i^{sell} = \frac{R_i}{w_i L_i}$ is equivalent to the labor market clearing condition of country i. Next, we calculate

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

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, we obtain R_1 and R_2 via the $R_i(w)$ formula above.

Multiple-Sector Economy The algorithm for calculating an equilibrium in a multiple-sector economy is similar. From (14) and (15), we can derive the following formula for M_1^{sell} and M_2^{sell}:

$$\begin{aligned} M_1^{sell} &= \left[\sum_{s=1}^{S} \alpha_s \left(\int_{\{\omega:\, \chi_{s1}^*(\omega)=1\}} m_{1s\omega}^{-1} \phi_{1s\omega} d\omega + \int_{\{\omega:\, \chi_{s2}^*(\omega)=1\}} m_{2s\omega}^{-1} \phi_{2s\omega} \frac{\phi_{1,2}}{\phi_{2,1}} d\omega \right) \right]^{-1} \\ M_2^{sell}(w) &= \left[\sum_{s=1}^{S} \alpha_s \left(\int_{\{\omega:\, \chi_{s1}^*(\omega)=2\}} m_{1s\omega}^{-1} \phi_{1s\omega} \frac{\phi_{2,1}}{\phi_{1,2}}(w)\, d\omega + \int_{\{\omega:\, \chi_{s2}^*(\omega)=2\}} m_{2s\omega}^{-1} \phi_{2s\omega} d\omega \right) \right]^{-1}, \end{aligned}$$

in which $\phi_{j,i}$ is the total spending share of j on i's goods given in (13). Then, we still calculate $R_1(w) = M_1^{sell}(w) L_1$, $R_2(w) = \frac{\phi_{1,2}}{\phi_{2,1}}(w) \times R_1(w)$, and $M_2^{sell}(w) = \frac{R_2(w)}{wL_2}$; the last is used to pin down equilibrium wage ratio w.

A2. 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]

$$Q_{it} = F_{it}(L_{it}, K_{it}, M_{it}, \omega_{it}), \tag{16}$$

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.

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

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

$$\min_{\{L_{it},K_{it},M_{it}\}} w_{it}L_{it} + r_{it}K_{it} + p_{it}^{m}M_{it} \tag{17}$$

$$s.t.\ F_{it}\left(L_{it},K_{it},M_{it},\omega_{it}\right) \geq Q_{it},$$

where w_{it}, r_{it}, and p_{it}^{m} denote the wage rate, rental price of capital and the price of intermediate inputs, respectively; and 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 its output elasticity. As labor is largely not freely chosen in China (particularly state-owned enterprises) and capital is often considered a dynamic input (which makes its output 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 (17) can be written as

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

Hence, the first-order condition for intermediate materials is

$$\frac{\partial\mathcal{L}}{\partial M_{it}} = p_{it}^{m} - \lambda_{it}\frac{\partial F_{it}}{\partial M_{it}} = 0. \tag{18}$$

Rearranging equation (18) and multiplying both sides by $\frac{M_{it}}{Q_{it}}$ yield

$$\begin{aligned}\frac{\partial F_{it}}{\partial M_{it}}\frac{M_{it}}{Q_{it}} &= \frac{1}{\lambda_{it}}\frac{p_{it}^{m}M_{it}}{Q_{it}} \\ &= \frac{P_{it}}{\lambda_{it}}\frac{p_{it}^{m}M_{it}}{P_{it}Q_{it}},\end{aligned} \tag{19}$$

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

Note that $\lambda_{it} = \frac{\partial\mathcal{L}}{\partial Q_{it}} = mc_{it}$ 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 (19) leads to the following estimation expression of firm markup[43]

$$\mu_{it} = \theta_{it}^{m}\left(\alpha_{it}^{m}\right)^{-1}, \tag{20}$$

[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.

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 (16). 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, Benkard, Berry and Pakes 2007). The solutions range from the instrumental variable estimation to the GMM estimation, and to the control function approach proposed by Olley and Pakes (1996). We adopt the control function approach developed by Ackerberg, Caves and Frazier (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

$$\begin{aligned} q_{it} \;=\;& \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \beta_{ll} l_{it}^2 + \beta_{kk} k_{it}^2 + \beta_{mm} m_{it}^2 \\ & + \beta_{lk} l_{it} k_{it} + \beta_{km} k_{it} m_{it} + \beta_{lm} l_{it} m_{it} \\ & + \beta_{lkm} l_{it} k_{it} m_{it} + \omega_{it} + \varepsilon_{it}, \end{aligned} \tag{21}$$

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. $\boldsymbol{\beta} = (\beta_l, \beta_k, \beta_m, \beta_{ll}, \beta_{kk}, \beta_{mm}, \beta_{lk}, \beta_{km}, \beta_{lm}, \beta_{lkm})$ is the vector of production function coefficients.

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

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

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

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

In the first stage, we estimate the following equation

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

where

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

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 $\boldsymbol{\beta}$ in the second stage, we model firm productivity as following a first-order Markov movement, i.e.

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

where ξ_{it} is an idiosyncratic shock.

From the first stage, the productivity for any given value of $\boldsymbol{\beta}$ can be computed as

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

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

To identify the coefficients of the production function, Ackerberg, Caves and Frazier (2006) assume that capital is determined one period beforehand and hence is not correlated with $\xi_{it} \left(\boldsymbol{\beta} \right)$. 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

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

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 (21) 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}, the logarithm of the net value of fixed assets deflated by investment price indices to measure k_{it}, and the logarithm of intermediate materials[44] deflated by input price indices to measure m_{it}; both price indices are provided by Brandt, Van Biesebroeck and Zhang (2012).

[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).

Once $\hat{\boldsymbol{\beta}} = \left(\hat{\beta}_l, \hat{\beta}_k, \hat{\beta}_m, \hat{\beta}_{ll}, \hat{\beta}_{kk}, \hat{\beta}_{mm}, \hat{\beta}_{lk}, \hat{\beta}_{km}, \hat{\beta}_{lm}, \hat{\beta}_{lkm}\right)$ is obtained, we can readily calculate the firm markup using equation (20), i.e.

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

where $\hat{\theta}_{it}^{m} = \hat{\beta}_m + 2\hat{\beta}_{mm} m_{it} + \hat{\beta}_{lm} l_{it} + \hat{\beta}_{km} k_{it} + \hat{\beta}_{lmk} l_{it} k_{it}$. Production estimates are reported in Table A1.

A3. Welfare Gains by the ACR Formula with Large Change in Trade Cost

We first calculate

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

To numerically calculate the above, we discretize the interval of $[\tau_{2004}, \tau_{1995}]$ by having an n-grid so 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

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

We calculate $v_{jj}(\tau_i)$ 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 $(2.311 - 1.664)/50 = 0.01294$. Once we obtain $\ln\left(\frac{W_j^{ACR,2004}}{W_j^{ACR,1995}}\right)$, we can then calculate the percentage increase in welfare $\frac{W_j^{ACR,2004}}{W_j^{ACR,1995}} - 1$.

A4. 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.

Labor demand in country i from a non-exporter that produces input ω is

$$\ell_{\omega,i} = \frac{q_{i\omega}}{\varphi_{\omega,i}^{*}} = \frac{1}{\varphi_{\omega,i}^{*}} \frac{R_i + D_i}{P_i} \left(\frac{p_{i\omega}}{P_i}\right)^{-\sigma}.$$

For an exporter at i, its labor demand is

$$\begin{aligned}
\ell_{\omega,1} &= \frac{q_{1\omega}+\tau q_{2\omega}}{\varphi^*_{\omega,1}} = \frac{1}{\varphi^*_{\omega,1}}\left[\frac{R_1-D}{P_1}\left(\frac{p_{1\omega}}{P_1}\right)^{-\sigma}+\frac{\tau\left(R_2+D\right)}{P_2}\left(\frac{p_{2\omega}}{P_2}\right)^{-\sigma}\right]\\
\ell_{\omega,2} &= \frac{\tau q_{1\omega}+q_{2\omega}}{\varphi^*_{\omega,2}} = \frac{1}{\varphi^*_{\omega,2}}\left[\frac{\tau\left(R_1-D\right)}{P_1}\left(\frac{p_{1\omega}}{P_1}\right)^{-\sigma}+\frac{R_2+D}{P_2}\left(\frac{p_{2\omega}}{P_2}\right)^{-\sigma}\right].
\end{aligned}$$

Labor market clearing in country i is again $\int_{\omega\in\chi_i}\ell_{\omega,i}d\omega = L_i$. The previous balanced trade condition (6) is now modified as $\left(R_2+D\right)\phi_{2,1} = \left(R_1-D\right)\phi_{1,2}+D$, or equivalently,

$$R_2\phi_{2,1} = R_1\phi_{1,2} + D\left(\phi_{2,2}-\phi_{1,2}\right), \tag{22}$$

where $\phi_{j,i}$ is the total spending share of j on i's goods:

$$\phi_{j,i} = \int_{\left\{\omega:\,\chi^*_j(\omega)=i\right\}}\phi_{j\omega}d\omega.$$

Similar to Section 2.4, the equilibrium $\{w, R_1, R_2\}$ is determined by the two labor market clearing conditions and (22).

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:

$$\begin{aligned}
M_1^{sell} &= \frac{R_1}{w_1L_1} = \frac{\int_{\left\{\omega:\,\chi^*_1(\omega)=1\right\}}\phi_{1\omega}Y_1d\omega+\int_{\left\{\omega:\,\chi^*_2(\omega)=1\right\}}\phi_{2\omega}Y_2d\omega}{\int_{\left\{\omega:\,\chi^*_1(\omega)=1\right\}}m_{1\omega}^{-1}\phi_{1\omega}Y_1d\omega+\int_{\left\{\omega:\,\chi^*_2(\omega)=1\right\}}m_{2\omega}^{-1}\phi_{2\omega}Y_2d\omega}\\
&= \left(\int_{\left\{\omega:\,\chi^*_1(\omega)=1\right\}}m_{1\omega}^{-1}\frac{\phi_{1\omega}\left(R_1-D\right)}{R_1}d\omega+\int_{\left\{\omega:\,\chi^*_2(\omega)=1\right\}}m_{2\omega}^{-1}\frac{\phi_{2\omega}\left(R_2+D\right)}{R_1}d\omega\right)^{-1}\\
&= \left(\left(1-\frac{D}{R_1}\right)\int_{\left\{\omega:\,\chi^*_1(\omega)=1\right\}}m_{1\omega}^{-1}\phi_{1\omega}d\omega+\left(\frac{\phi_{1,2}}{\phi_{2,1}}+\frac{D}{R_1}\frac{1-\phi_{1,2}}{\phi_{2,1}}\right)\int_{\left\{\omega:\,\chi^*_2(\omega)=1\right\}}m_{2\omega}^{-1}\phi_{2\omega}d\omega\right)^{-1}
\end{aligned}$$

Recall that $\phi_{j,i} = \int_{\left\{\omega:\,\chi^*_j(\omega)=i\right\}}\phi_{j\omega}d\omega$ only depends 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 $M_1^{sell}\left(w,R_1\right)$. Then, given $w_1=1$ and L_1, we can solve for R_1 as a fixed point in $R_1 = M_1^{sell}\left(w,R_1\right)L_1$ and obtain $R_1\left(w\right)$. For R_2, we use (22) again:

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

In fact, $M_i^{sell} = \frac{R_i}{w_i L_i}$ is equivalent to the labor market clearing condition of country i. Next, we calculate

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

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 A2.

References

Ackerberg, D., C. L. Benkard, S. Berry, and A. Pakes (2007), “Econometric Tools for Analyzing Market Outcomes" in *Handbook of Econometrics*, Volume 6A, ed. James J. Heckman and Edward E. Leamer, 4171–276. Amsterdam and Boston: Elsevier, North-Holland.

Ackerberg, D., K. Caves, and G. Frazer (2006), “Structural Identification of Production Functions", working paper, UCLA.

Alvarez, Fernando and Robert Jr. Lucas (2007), "General Equilibrium Analysis of the Eaton-Kortum model of International Trade," *Journal of Monetary Economics*, 54(6): 1726-1768.

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012), “New Trade Models, Same Old Gains?,” *American Economic Review*, 102(1): 94–130.

Arkolakis, C., A. Costinot, D. Donaldson, and A. Rodriguez-Clare (2017), “The Elusive Pro-Competitive Effects of Trade,” working paper, Yale University, February 2017.

Asturias, Jose, Manuel García-Santana, and Roberto Ramos (2017), “Competition and the Welfare Gains from Transportation Infrastructure: Evidence from the Golden Quadrilateral of India,” working paper, Georgetown University Qatar.

Atkeson, A., and A. Burstein (2008), “Pricing to Market, Trade Costs, and International Relative Prices,” *American Economic Review*, 98(5): 1998–2031.

Behrens, K., and Y. Murata, (2012) “Trade, Competition, and Efficiency,” *Journal of International Economics*, 87(1): 1-17.

Bernard, A. B., J. Eaton, J. B. Jensen, and S. Kortum (2003), “Plants and Productivity in International Trade,” *American Economic Review,* 93(4): 1268–1290.

Brandt, L., J. Van Biesebroeck, L. Wang and Y. Zhang (2012), “WTO Accession and Performance of Chinese Manufacturing Firms,” working paper, Department of Economics, University of Toronto.

Brandt, L., J. Van Biesebroeck, Y. Zhang (2012), "Creative Accounting or Creative Destruction: Firm Level Productivity Growth in Chinese Manufacturing", *Journal of Development Economics*, 97(2): 339–351.

Caliendo, L. and F. Parro (2015), “Estimates of the Trade and Welfare Effects of NAFTA,” *Review of Economic Studies*, 82(1): 1-44.

de Blas, B., and K. Russ (2015), “Understanding Markups in the Open Economy under Bertrand Competition,” *American Economic Journal: Macroeconomics*, 7(2): 157-80.

di Giovanni, J., A. A. Levchenko and J. Zhang (2014), “The Global Welfare Impact of China: Trade Integration and Technological Change,” *American Economic Journal: Macro-*

economics, 6(3): 153-183.

De Loecker, J. and F. Warzynski (2012), "Markups and Firm-level Export Status," *American Economic Review*, 102(6): 2437-2471.

Devereux, M. B and K. M. Lee (2001), "Dynamic Gains from International Trade with Imperfect Competition and Market Power," *Review of Development Economics*, 5(2): 239-255.

Dixit, A. K. and J. E. Stiglitz (1977), "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 67(3): 297-308.

Dhingra, S., and Morrow, J. (2016), "Monopolistic Competition and Optimum Product Diversity Under Firm Heterogeneity," working paper, London School of Economics.

Eaton, J. and S. Kortum (2002), "Technology, Geography, and Trade," *Econometrica*, 70(5): 1741–1779.

Eaton, J., S. Kortum, and S. Sotelo (2013), "International Trade: Linking Micro and Macro," in *Advances in Economics and Econometrics*: 1st ed. Vol 2. Cambridge: Cambridge University Press: 329–370.

Edmond, C., V. Midrigan, and D. Y. Xu (2015), "Competition, Markups, and the Gains from International Trade," *American Economic Review,* 105(10): 3183–3221.

Epifani, P., and G. Gancia (2011), "Trade, markup heterogeneity and misallocations," *Journal of International Economics*, 83(1): 1-13.

Fajgelbaum, P. D., and Khandelwal, A. K. (2016), "Measuring the unequal gains from trade," *Quarterly Journal of Economics*, 131(3): 1113-1180.

Feenstra, R. C. (2014), "Restoring the Product Variety and Pro-competitive Gains from Trade with Heterogeneous Firms and Bounded Productivity," NBER Working Paper No. 19833, January 2014.

Feenstra, R. C., and Weinstein, D. E. (2016). "Globalization, Markups and U.S. Welfare," *Journal of Political Economy*, forthcoming.

Goldberg, P. K., J. De Loecker, A. K. Khandelwal, and N. Pavcnik (2015), "Prices, Markups and Trade Reform," *Econometrica*, forthcoming.

Head, K., T. Mayer, and M. Thoenig (2014), "Welfare and Trade without Pareto," *American Economic Review (Papers and Proceedings)*, 104(5): 310–316.

Helpman, E., M. Melitz, and Y. Rubinstein (2008), "Estimating Trade Flows: Trading Partners and Trading Volumes." *Quarterly Journal of Economics* 123: 441–487.

Holmes, T. J., W.-T. Hsu, and S. Lee (2014), "Allocative Efficiency, Mark-ups, and the Welfare Gains from Trade," *Journal of International Economics*, 94(2): 195–206.

Hsieh, C.-T. and P. J. Klenow (2009), "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124(4): 1403–1448.

Hsieh, C. T., and Ossa, R. (2016), “A global view of productivity growth in China,” *Journal of International Economics*, 102: 209-224.

Karabarbounis, L., and B. Neiman (2014), “The Global Decline of the Labor Share,” *Quarterly Journal of Economics*, 129(1): 61–103.

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

Levinsohn, J. A., and A. K. Petrin (2003), “Estimating Production Functions Using Inputs to Control for Unobservables,” *Review of Economic Studies*, 70(2): 317–342.

Lipsey, R. G. and Kelvin Lancaster (1956-1957), “The General Theory of Second Best,” *Review of Economic Studies*, 24(1): 11-32.

Lu, Y. and L. Yu (2015), “Trade Liberalization and Markup Dispersion: Evidence from China’s WTO Accession,” *American Economic Journal: Applied Economics*, 7(4): 221–253.

Markusen, J. (1981), “Trade and the Gains from Trade with Imperfect Competition,” *Journal of International Economics*, 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., and G. Ottaviano (2008), “Market Size, Trade, and Productivity,” *Review of Economic Studies*, 75(1): 295–316.

Melitz, M. J., and S. J. Redding (2015), “New Trade Models, New Welfare Implications,” *American Economic Review*, 105(3): 1105–46.

Olley, G. S., and A. Pakes (1996), “The Dynamics of Productivity in the Telecommunications Equipment Industry,” *Econometrica* 64, 1263–1298

Ottaviano, G., T. Tabuchi, and J.-F. Thisse (2002), “Agglomeration and Trade Revisited," *International Economic Review*, 43(2): 409–435.

Restuccia, D., and R. Rogerson (2008), “Policy Distortions and Aggregate Productivity with Heterogeneous Plants,” *Review of Economic Dynamics*, 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 Journal of Economics*, 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,” *Journal of Political Economy*, 113(3): 582–625.

Simonovska, I., and Waugh, M. E. (2014a), “The elasticity of trade: Estimates and evidence,” *Journal of International Economics*, 92(1): 34-50.

Simonovska, I., and Waugh, M. E. (2014b), “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," *American Economic Review*, 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., and Thisse, J. F. (2012), "Monopolistic competition: Beyond the constant elasticity of substitution," *Econometrica*, 80(6): 2765-2784.

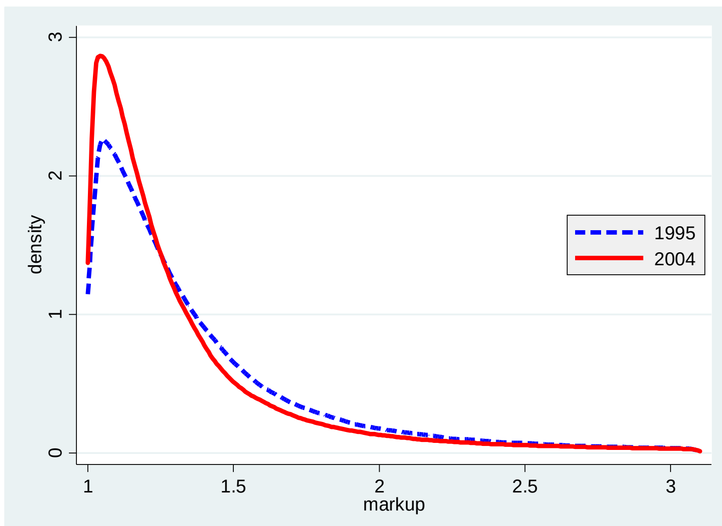

Figure 1: Markup Distributions (1995 versus 2004)

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

Table 2: SMM Results

		1995		2004	
Predetermined					
w	Relative wages (the ROW to China)	10.5		5.3	
R1	China's manufacturing sales ($b)	918,291		2,343,328	
R2	ROW's manufacturing sales ($b)	9,397,500		14,737,500	
σ	Inferred from p99 markup	1.40		1.40	
Moments for SMM		**Data**	**Model**	**Data**	**Model**
import share		0.130	0.144	0.222	0.262
export share		0.153	0.190	0.249	0.277
relative number of firms		0.210	0.219	0.596	0.616
fraction of exporters		0.044	0.024	0.105	0.062
mean cost share for exporters		0.845	0.798	0.801	0.804
std of cost share for exporters		0.135	0.142	0.142	0.124
p50 markup for exporters		1.120	1.212	1.168	1.203
p95 markup for exporters		2.199	2.457	2.183	1.839
mean cost share for non-exporters		0.789	0.712	0.829	0.775
std of cost share for non-exporters		0.147	0.187	0.139	0.152
p50 markup for non-exporters		1.266	1.391	1.213	1.264
p95 markup for non-exporters		2.475	2.775	2.400	2.056
Parameter values		**Estimates**	**s.e.**	**Estimates**	**s.e.**
τ, trade cost		2.311	0.020	1.664	0.005
$\gamma/\bar{N}$, measure of goods relative to $\bar{N}$		0.261	0.002	0.849	0.005
λ_1, Poisson parameter, China		2.442	0.062	2.607	0.040
λ_2, Poisson parameter, ROW		5.286	0.061	5.828	0.063
μ_1, mean of log productivity, China relative to ROW		-2.401	0.024	-1.785	0.009
η_1, std of log productivity, China		0.444	0.008	0.410	0.001
η_2, std of log productivity, ROW		0.349	0.016	0.293	0.011
Simulated macro variables under estimated parameters		**Data**	**Model**	**Data**	**Model**
w		10.5	10.3	5.3	5.3
R1		918,291	954,812	2,343,328	2,398,028
R2		9,397,500	8,410,637	14,737,500	13,974,893

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 3: Jacobian Matrix

moments	τ	γ	λ_1	λ_2	μ_1	η_1	η_2
import share	-0.514	0.005	-0.076	0.002	-0.246	-0.004	0.586
export share	-0.977	0.007	0.087	-0.070	1.169	0.726	-0.345
relative number of firms	0.316	0.775	0.110	-0.013	0.479	0.212	-0.922
fraction of exporters	-0.214	0.001	0.007	-0.012	0.154	0.190	-0.026
mean cost share for exporters	-0.024	0.001	0.013	0.006	0.026	-0.102	-0.062
std of cost share for exporters	0.009	-0.012	-0.009	-0.012	-0.022	-0.139	0.074
p50 markup for exporters	0.038	0.009	-0.020	-0.005	-0.029	0.168	0.034
p95 markup for exporters	0.369	-0.132	-0.185	-0.164	-0.594	-12.864	1.008
mean cost share for non-exporters	-0.109	-0.001	0.019	0.002	-0.066	-0.367	0.092
std of cost share for non-exporters	0.070	-0.001	-0.006	-0.001	0.031	0.186	-0.029
p50 markup for non-exporters	0.176	0.007	-0.037	-0.004	0.141	0.619	-0.189
p95 markup for non-exporters	1.019	-0.031	-0.111	-0.025	0.583	2.772	-0.522

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.

Table 4: Counter-factual Analysis

Panel A: Counter-factual from 2004 estimates

	Under 2004 estimates	τ at 1995	% change	Autarky	% change
τ, trade cost	1.664	2.311		1,000,000	
Welfare					
Total Welfare	1.90E+21	1.73E+21	9.4%	1.42E+21	33.4%
W_Prod	1.04E+15	9.78E+14	6.8%	8.44E+14	23.8%
W_A	0.965	0.945	2.1%	0.897	7.5%
W_R	1.003	1.000	0.3%	1.000	0.3%
Contribution to total welfare					
W_A and W_R			25.4%		23.3%
W_A			22.3%		22.4%

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	2.252		1.810	
Welfare					
Total Welfare	1.42E+21	1.74E+21	22.6%	1.85E+21	6.0%
W_Prod	8.44E+14	9.81E+14	16.2%	1.02E+15	4.3%
W_A	0.897	0.946	5.5%	0.960	1.5%
W_R	1.000	1.000	0.0%	1.002	0.2%
Contribution to total welfare					
W_A and W_R			24.2%		27.6%
W_A			24.2%		24.6%

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.

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	0.194	0.055	22.0%	-2.48	0.200	24.3%
2004	0.409	0.330	0.079	19.3%	-3.23	0.349	17.1%

Notes: All the welfare gains here are calculated in terms of welfare elasticity to trade cost, i.e., dln(W)/dln(tau), 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 = 0.001. 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(tau21) 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.

`

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	
σ	Inferred from p99 markup	1.40		1.40	
Moments		**Data**	**Model**	**Data**	**Model**
import share		0.130	0.053	0.222	0.117
export share		0.153	0.049	0.249	0.114
relative number of firms		0.210	0.213	0.596	0.611
fraction of exporters		0.044	0.064	0.105	0.140
mean cost share for exporters		0.845	0.731	0.801	0.747
std of cost share for exporters		0.135	0.158	0.142	0.142
p50 markup for exporters		1.120	1.370	1.168	1.334
p95 markup for exporters		2.199	2.564	2.183	2.052
mean cost share for non-exporters		0.789	0.759	0.829	0.793
std of cost share for non-exporters		0.147	0.170	0.139	0.148
p50 markup for non-exporters		1.266	1.289	1.213	1.230
p95 markup for non-exporters		2.475	2.399	2.400	1.995
Parameter values		**Estimate**	**s.e.**	**Estimate**	**s.e.**
τ, trade cost		2.329	0.008	1.738	0.003
$\gamma/\bar{N}$, measure of goods relative to $\bar{N}$		0.228	0.002	0.699	0.003
λ, Poisson parameter		3.635	0.010	4.219	0.080
η, std. of log productivity		0.399	0.003	0.407	0.005

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

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	1.738	2.329		1,000,000	
Welfare					
Total Welfare	2.30E+19	2.24E+19	2.7%	2.13E+19	8.1%
W_Prod	1.31E+13	1.28E+13	2.1%	1.24E+13	5.4%
W_A	0.964	0.958	0.6%	0.941	2.5%
Contribution to total welfare					
W_A			23.7%		31.4%

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	1.815		1.465	
Welfare					
Total Welfare	2.13E+19	2.29E+19	7.5%	2.37E+19	3.6%
W_Prod	1.24E+13	1.30E+13	5.0%	1.34E+13	3.1%
W_A	0.941	0.964	2.4%	0.9681	0.5%
Contribution to total welfare					
W_A			32.5%		13.3%

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.

Table 8: Comparative Statics of Other Parameters

Panel A: Comparative Statics of η on Gains from Trade

	Gains from Trade (in percentage)				
η	$0.5\times\eta_0$	$0.75\times\eta_0$	η_0=0.407	$1.25\times\eta_0$	$1.5\times\eta_0$
Total Welfare	1.0%	1.6%	2.7%	4.0%	5.1%
W_Prod	0.4%	1.0%	2.1%	3.3%	4.4%
W_A	0.6%	0.6%	0.6%	0.7%	0.7%
Contribution of W_A	60.0%	37.5%	22.2%	17.5%	13.7%

Panel B: Comparative Statics of λ on Gains from Trade

	Gains from Trade (in percentage)				
λ	$0.5\times\lambda_0$	$0.75\times\lambda_0$	λ_0=4.219	$1.25\times\lambda_0$	$1.5\times\lambda_0$
Total Welfare	5.8%	3.9%	2.7%	2.0%	1.6%
W_Prod	4.6%	2.9%	2.1%	1.6%	1.3%
W_A	1.2%	0.9%	0.6%	0.4%	0.3%
Contribution of W_A	20.7%	23.1%	22.2%	20.0%	18.8%

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.

Table 9: 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	7.7%	6.0%	1.5%	0.1%	20.8%	19.3%
% change from autarky to τ at 1995	28.7%	21.2%	6.1%	0.1%	21.6%	21.1%

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	5.8%	4.5%	1.2%	0.0%	21.1%	21.1%
% change from autarky to τ at 2004	19.2%	14.9%	4.0%	-0.3%	19.6%	21.1%

Robustness Check 3: Using the 97.5th percentile to Infer Sigma

	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	6.4%	5.1%	1.2%	0.0%	19.4%	19.4%
% change from autarky to τ at 2004	21.8%	15.7%	4.9%	0.4%	24.0%	22.4%

Robustness Check 4: 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	6.7%	4.8%	1.3%	0.6%	27.3%	19.0%

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.

Table 10A: Estimation Result in Multi-Sector Model (Part A)

		Predetermined		γ			τ			Tariff			Non-tariff τ		
cic2d	**Industry definition**	**σ**	**α**	**1995**	**2004**	**% change**	**1995**	**2004**	**% change**	**1995**	**2004**	**% change**	**1995**	**2004**	**% change**
13	Food processing	1.51	0.049	0.018	0.044	139.1	2.47	2.347	-4.9	25.6	16.6	-35.2	1.97	2.01	2.5
14	Food manufacturing	1.33	0.017	0.009	0.018	109.2	4.62	2.49	-46.1	17.5	9.9	-43.6	3.93	2.27	-42.3
15	Beverage manufacturing	1.21	0.014	0.009	0.015	72.4	4.80	3.33	-30.7	25.3	7.7	-69.8	3.83	3.09	-19.3
16	Tobacco processing	1.22	0.014	0.0003	0.0002	-29.8	4.49	4.81	6.9	37.9	9.8	-74.3	3.26	4.38	34.4
17	Textile industry	1.49	0.059	0.013	0.049	269.5	1.88	1.69	-10.4	19.7	7.6	-61.4	1.57	1.57	-0.3
18	Garments & other fiber products	1.37	0.023	0.009	0.028	204.7	3.52	2.91	-17.2	10.8	9.2	-15.0	3.18	2.67	-16.0
19	Leather, furs, down & related products	1.39	0.016	0.006	0.014	136.9	1.96	1.61	-18.0	9.9	5.5	-44.3	1.79	1.53	-14.6
20	Timber processing, bamboo, cane, palm fiber & straw products	1.40	0.011	0.009	0.026	198.8	2.10	1.64	-21.9	7.8	2.6	-67.1	1.95	1.60	-17.9
21	Furniture manufacturing	1.26	0.008	0.005	0.015	215.2	2.47	1.92	-22.2	8.3	1.0	-88.0	2.28	1.90	-16.6
22	Papermaking & paper products	1.48	0.020	0.008	0.025	200.5	2.59	2.16	-16.3	23.7	4.0	-83.0	2.09	2.08	-0.5
23	Printing industry	1.29	0.009	0.010	0.027	157.5	2.68	2.28	-15.1	5.3	0.9	-83.5	2.55	2.26	-11.4
24	Cultural, educational & sports goods	1.35	0.007	0.002	0.009	254.9	2.11	1.72	-18.7	4.1	1.5	-64.3	2.03	1.70	-16.6
25	Petroleum processing & coking	1.45	0.050	0.001	0.006	346.7	1.96	1.54	-21.6	8.6	5.0	-42.2	1.80	1.46	-18.9
26	Raw chemical materials & chemical products	1.50	0.072	0.015	0.073	382.9	2.51	1.74	-30.5	14.6	7.2	-51.0	2.19	1.62	-25.7
27	Medical & pharmaceutical products	1.33	0.017	0.004	0.007	71.9	4.43	2.76	-37.7	6.9	3.8	-44.9	4.15	2.66	-35.8
28	Chemical fiber	2.01	0.010	0.001	0.003	337.5	3.16	2.23	-29.4	22.0	4.9	-77.7	2.59	2.12	-17.9
29	Rubber products	1.50	0.010	0.003	0.009	237.4	2.09	1.84	-11.8	20.2	11.0	-45.6	1.74	1.66	-4.5
30	Plastic products	1.54	0.027	0.011	0.045	320.7	1.76	1.72	-2.4	13.9	5.4	-61.0	1.55	1.63	5.4
31	Nonmetal mineral products	1.35	0.050	0.035	0.094	172.8	4.62	2.39	-48.3	12.8	5.9	-54.0	4.10	2.26	-44.9
32	Smelting & pressing of ferrous metals	1.78	0.092	0.005	0.013	159.1	2.46	2.17	-11.6	10.9	4.9	-55.2	2.22	2.07	-6.6
33	Smelting & pressing of nonferrous metals	1.67	0.031	0.002	0.009	294.3	2.12	1.83	-13.9	7.7	3.9	-49.4	1.97	1.76	-10.7
34	Metal products	1.43	0.032	0.014	0.048	249.6	1.98	1.76	-10.8	13.2	4.0	-69.9	1.75	1.70	-2.8
35	Ordinary machinery	1.50	0.052	0.017	0.084	393.0	3.07	1.64	-46.6	17.5	5.1	-71.0	2.61	1.56	-40.3
36	Special purpose equipment	1.35	0.030	0.011	0.038	241.5	2.41	1.61	-33.2	16.6	5.3	-68.2	2.07	1.53	-26.0
37	Transport equipment	1.36	0.076	0.014	0.036	161.0	2.62	2.18	-16.8	43.5	12.7	-70.8	1.83	1.93	5.9
39	Electric equipment & machinery	1.51	0.061	0.012	0.037	197.9	1.71	1.53	-10.5	11.3	3.0	-73.0	1.54	1.48	-3.4
40	Electronic & telecommunications equipment	1.34	0.121	0.005	0.017	271.3	2.19	1.51	-31.1	13.5	1.3	-90.5	1.93	1.49	-22.8
41	Instruments, meters, cultural & office equipment	1.35	0.013	0.004	0.012	194.4	1.85	1.52	-18.2	15.7	4.3	-72.6	1.60	1.45	-9.2
42	Other manufacturing	1.32	0.009	0.006	0.016	173.0	2.30	1.68	-26.9	8.8	2.8	-67.8	2.12	1.64	-22.7
Mean		1.44	0.034	0.01	0.03	211.52	2.72	2.09	-21.23	15.64	5.74	-61.87	2.35	1.97	-13.77
Standard deviation		0.17	0.029	0.01	0.02	95.89	0.96	0.70	13.23	9.12	3.66	17.25	0.80	0.62	16.52
Max		2.01	0.121	0.03	0.09	392.98	4.80	4.81	6.95	43.46	16.57	-14.99	4.15	4.38	34.39
Min		1.21	0.007	0.00	0.00	-29.81	1.71	1.51	-48.28	4.08	0.87	-90.54	1.54	1.45	-44.90

Table 10B: 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	2.78	3.05	9.9	5.40	6.9	26.8	-2.37	-1.67	29.8	0.43	0.46	8.2	0.38	0.37	-4.8
14	Food manufacturing	3.03	3.11	2.5	4.25	4.6	8.7	-1.51	-1.50	0.7	0.15	0.34	136.2	0.33	0.34	0.2
15	Beverage manufacturing	3.05	3.24	6.2	5.82	4.5	-23.1	-1.20	-1.23	-2.7	0.11	0.22	95.7	0.09	0.16	82.1
16	Tobacco processing	2.75	3.18	15.8	6.00	5.6	-6.3	-2.38	-1.33	44.0	0.34	0.44	28.6	0.27	0.20	-23.6
17	Textile industry	3.04	3.25	6.9	5.73	6.3	10.2	-2.36	-1.78	24.6	0.35	0.38	9.2	0.35	0.17	-50.8
18	Garments & other fiber products	3.03	3.22	6.2	5.27	6.6	25.5	-2.15	-0.93	56.7	0.55	0.42	-23.9	0.11	0.42	269.7
19	Leather, furs, down & related products	2.85	3.21	12.7	5.13	4.6	-10.4	-2.19	-1.75	20.1	0.33	0.37	11.1	0.44	0.26	-40.3
20	Timber processing, bamboo, cane, palm fiber & straw products	2.85	2.98	4.7	5.26	5.4	2.7	-2.40	-1.74	27.5	0.41	0.32	-20.0	0.40	0.25	-38.7
21	Furniture manufacturing	2.51	2.87	14.4	5.53	5.0	-9.6	-2.03	-1.66	18.1	0.24	0.37	56.4	0.26	0.08	-70.6
22	Papermaking & paper products	3.02	2.76	-8.4	5.59	6.1	9.1	-2.33	-1.79	23.3	0.29	0.41	43.3	0.52	0.44	-16.4
23	Printing industry	2.97	2.61	-12.2	6.41	5.6	-12.7	-2.37	-1.78	24.9	0.35	0.42	18.7	0.09	0.18	106.3
24	Cultural, educational & sports goods	2.44	3.02	23.9	5.06	4.6	-8.6	-2.14	-1.70	20.5	0.38	0.42	9.0	0.28	0.15	-46.4
25	Petroleum processing & coking	2.70	2.87	6.3	5.37	6.1	13.8	-2.25	-1.84	18.3	0.21	0.32	51.8	0.36	0.32	-12.3
26	Raw chemical materials & chemical products	2.63	1.97	-25.1	4.84	6.7	39.0	-2.52	-1.92	23.7	0.43	0.51	17.1	0.52	0.36	-31.1
27	Medical & pharmaceutical products	2.99	2.66	-11.1	6.23	5.0	-19.4	-2.09	-1.61	23.1	0.54	0.50	-8.3	0.17	0.39	130.7
28	Chemical fiber	2.44	2.89	18.2	5.47	4.5	-17.6	-2.59	-1.75	32.5	0.46	0.26	-43.3	0.51	0.44	-14.5
29	Rubber products	3.36	2.61	-22.3	5.57	4.6	-18.2	-2.27	-1.72	24.0	0.30	0.37	22.5	0.15	0.15	5.1
30	Plastic products	2.73	3.13	14.6	5.31	5.8	9.1	-2.39	-1.67	30.2	0.31	0.34	11.0	0.21	0.26	22.6
31	Nonmetal mineral products	3.03	3.03	0.1	4.95	5.6	12.5	-1.59	-1.39	12.7	0.25	0.28	11.0	0.20	0.34	65.3
32	Smelting & pressing of ferrous metals	2.87	3.16	10.1	6.50	4.4	-32.5	-2.40	-1.60	33.3	0.06	0.10	59.4	0.34	0.40	20.0
33	Smelting & pressing of nonferrous metals	2.29	2.54	11.2	5.83	5.8	-0.6	-2.46	-1.77	28.1	0.36	0.40	9.5	0.23	0.33	41.7
34	Metal products	3.03	2.83	-6.7	5.54	5.8	5.2	-2.42	-1.82	24.6	0.37	0.39	5.1	0.24	0.09	-60.4
35	Ordinary machinery	2.24	2.50	11.6	4.64	7.3	56.7	-2.44	-1.79	26.8	0.40	0.42	3.8	0.52	0.31	-39.9
36	Special purpose equipment	2.12	2.49	17.3	5.17	6.9	33.3	-2.63	-1.80	31.4	0.34	0.40	19.4	0.52	0.42	-19.5
37	Transport equipment	2.74	2.53	-7.6	6.12	5.9	-3.9	-2.38	-1.72	27.5	0.38	0.42	8.6	0.29	0.30	6.2
39	Electric equipment & machinery	2.44	2.95	20.9	5.93	5.9	0.0	-2.43	-1.77	27.0	0.31	0.40	28.5	0.16	0.34	114.4
40	Electronic & telecommunications equipment	2.41	2.47	2.5	5.81	5.9	1.8	-2.36	-1.77	25.1	0.51	0.53	3.5	0.52	0.44	-16.0
41	Instruments, meters, cultural & office equipment	2.38	2.18	-8.7	4.95	5.8	16.2	-2.41	-1.75	27.5	0.45	0.46	1.9	0.51	0.44	-13.4
42	Other manufacturing	2.73	3.20	17.4	5.36	5.8	8.7	-2.25	-1.74	22.7	0.52	0.43	-17.6	0.34	0.16	-52.4
Mean		2.74	2.84	4.52	5.48	5.64	4.02	-2.25	-1.67	25.03	0.35	0.38	19.19	0.32	0.29	10.80
Standard deviation		0.30	0.34	12.66	0.52	0.81	19.75	0.32	0.21	10.82	0.12	0.09	35.61	0.14	0.11	72.80
Max		3.36	3.25	23.89	6.50	7.27	56.75	-1.20	-0.93	56.74	0.55	0.53	136.22	0.52	0.44	269.74
Min		2.12	1.97	-25.11	4.25	4.39	-32.53	-2.63	-1.92	-2.72	0.06	0.10	-43.25	0.09	0.08	-70.59

Table 11: 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	7.2%	5.7%	1.4%	0.0%	20.0%	19.7%
% change from autarky to τ at 2004	28.2%	21.1%	5.9%	-0.1%	20.6%	20.9%

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 τ to 2004's.

Table 12: 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	-2.109**	-1.774**	-1.980*	-1.856	-0.343**	-0.378**	-0.581*	-0.596*
	(0.799)	(0.833)	(1.063)	(1.117)	(0.154)	(0.164)	(0.296)	(0.303)
SOE share		0.242		-0.967		-0.029		0.042
		(0.157)		(0.728)		(0.032)		(0.171)
Log wage at 1995			-0.072	-0.056			-0.065	-0.064
			(0.209)	(0.218)			(0.066)	(0.068)
Log employment at 1995			-0.153	-0.352			-0.021	-0.011
			(0.125)	(0.218)			(0.029)	(0.056)
Log export at 1995			0.166**	0.181**			-0.045***	-0.047**
			(0.077)	(0.087)			(0.016)	(0.017)
Log import at 1995			-0.035	0.060			0.047**	0.044*
			(0.063)	(0.116)			(0.020)	(0.024)
R^2	0.169	0.186	0.386	0.449	0.108	0.114	0.363	0.366

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.*

Table A1: Production Function Estimates

Industry	Panel A: Output Elasticity With Respect to … Labor Median	Labor IQR	Capital Median	Capital IQR	Materials Median	Materials IQR	Panel B: Returns to Scale Double Median	Double IQR	Triple Median	Triple 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^rY=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.

Table A2: SMM Results in the Model with Trade Imbalance

		1995		2004	
Predetermined					
w	Relative wages (the ROW to China)	10.53		5.29	
R1	China's manufacturing sales ($b)	918,291		2,343,328	
R2	ROW's manufacturing sales ($b)	9,397,500		14,737,500	
σ	Inferred from p99 markup	1.40		1.40	
Moments for SMM		**Data**	**Model**	**Data**	**Model**
import share		0.130	0.141	0.222	0.237
export share		0.153	0.184	0.249	0.290
relative number of firms		0.210	0.218	0.596	0.590
fraction of exporters		0.044	0.023	0.105	0.064
mean cost share for exporters		0.845	0.813	0.801	0.798
std of cost share for exporters		0.135	0.135	0.142	0.132
p50 markup for exporters		1.195	1.188	1.168	1.213
p95 markup for exporters		2.166	2.069	2.197	1.962
mean cost share for non-exporters		0.789	0.733	0.829	0.764
std of cost share for non-exporters		0.147	0.177	0.139	0.159
p50 markup for non-exporters		1.264	1.344	1.213	1.283
p95 markup for non-exporters		2.411	2.518	2.418	2.158
Parameter values		**Estimates**	**s.e.**	**Estimates**	**s.e.**
τ, trade cost		2.131	0.012	1.739	0.010
$\gamma/\bar{N}$, measure of goods relative to $\bar{N}$		0.260	0.001	0.790	0.003
λ_1, Poisson parameter, China		2.442	0.035	2.591	0.045
λ_2, Poisson parameter, ROW		5.779	0.025	5.337	0.048
μ_1, mean of log productivity, China relative to ROW		-2.363	0.005	-1.763	0.014
η_1, std of log productivity, China		0.400	0.001	0.417	0.002
η_2, std of log productivity, ROW		0.329	0.009	0.322	0.014
Simulated macro variables under estimated parameters		**Data**	**Model**	**Data**	**Model**
w		10.53	10.37	5.29	5.23
R1		918,291	900,350	2,343,328	2,319,500
R2		9,397,500	9,167,500	14,737,500	14,271,000

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.