Competition, Markups, and the Gains from International Trade*

Chris Edmond† Virgiliu Midrigan‡ Daniel Yi Xu§

First draft: July 2011. This draft: June 2014

Abstract

We study the pro-competitive gains from international trade in a quantitative model with endogenously variable markups. We find that trade can significantly reduce markup distortions if two conditions are satisfied: (i) there is extensive misallocation and (ii) opening to trade exposes hitherto dominant producers to greater competitive pressure. We measure the extent to which these two conditions are satisfied in Taiwanese producer-level data. Versions of our model consistent with the Taiwanese data predict that opening up to trade strongly increases competition and reduces markup distortions by up to one-third, thus significantly reducing productivity losses due to misallocation.

Keywords: misallocation, markup dispersion, head-to-head competition.

JEL classifications: F1, O4.

*We thank our editor Penny Goldberg and four anonymous referees for valuable comments and suggestions. We have also benefited from discussions with Fernando Alvarez, Costas Arkolakis, Andrew Atkeson, Ariel Burstein, Vasco Carvalho, Andrew Cassey, Arnaud Costinot, Jan De Loecker, Dave Donaldson, Ana Cecilia Fieler, Oleg Itskhoki, Phil McCalman, Markus Poschke, Andrés Rodríguez-Clare, Barbara Spencer, Iván Werning, and from numerous conference and seminar participants. We also thank Andres Blanco, Jiwoon Kim and Fernando Leibovici for their excellent research assistance. We gratefully acknowledge support from the National Science Foundation, Grant SES-1156168.

†University of Melbourne, cedmond@unimelb.edu.au.

‡New York University and NBER, virgiliu.midrigan@nyu.edu.

§Duke University and NBER, daniel.xu@duke.edu.

1 Introduction

Can international trade significantly reduce product market distortions? We study this question in a quantitative trade model with endogenously variable markups. In such a model, markup dispersion implies that resources are *misallocated* and that aggregate productivity is low. By exposing producers to greater competition, international trade may reduce markup dispersion thereby reducing misallocation and increasing aggregate productivity. Our goal is to use producer-level data to quantify these *pro-competitive effects* of trade on misallocation and aggregate productivity.

We study these pro-competitive effects in the model of Atkeson and Burstein (2008). In this model, any given sector has a small number of producers who engage in *oligopolistic competition*. The demand elasticity for any given producer is decreasing in its market share and hence its markup is increasing in its market share. By reducing the market shares of dominant producers, international trade can reduce markups and markup dispersion. The Atkeson and Burstein (2008) model is particularly useful for us because it implies a *linear* relationship between (inverse) producer-level markups and market shares, which in turn makes the model straightforward to parameterize.

We find that trade can significantly reduce markup distortions if two conditions are satisfied: (i) there is extensive misallocation, and (ii) international trade does in fact put producers under greater competitive pressure. The first condition is obvious — if there is no misallocation, there is no misallocation to reduce. The second condition is more subtle. Trade has to increase the degree of effective competition amongst producers prevailing within the market. If both domestic and foreign producers have similar productivities within a given sector, then opening to trade exposes them to genuine *head-to-head competition* that reduces market power thereby reducing markups and markup dispersion. By contrast, if there are large cross-country differences in productivity within a given sector, then opening to trade may allow producers from one country to substantially increase their market share in the other country, thereby increasing markups and markup dispersion so that the pro-competitive 'gains' from trade are negative.

We quantify the model using 7-digit Taiwanese manufacturing data. We use this data to discipline two key determinants of the extent of misallocation: (i) the elasticity of substitution across sectors, and (ii) the equilibrium distribution of producer market shares. The elasticity of substitution across sectors plays a key role because it determines the extent to which producers that face little competition in their own sector can raise markups. We pin down this elasticity by requiring that our model fits the cross-sectional relationship between measures of markups and market shares that we observe in the Taiwanese data. We pin down the parameters of the producer-level productivity distribution and fixed costs of operating and exporting by requiring that the model reproduces key moments of the distribution of market shares within and across sectors in the Taiwanese data.

The Taiwanese data feature a large amount of dispersion and concentration in producer-level market shares, as well as a strong relationship between market shares and markups. Interpreted through the lens of the model, this implies a significant amount of misallocation and hence the possibility of significant productivity gains from reduced markup distortions.

Given this misallocation, the model predicts large pro-competitive gains if, within a given sector, domestic producers and foreign producers have relatively similar levels of productivity so that increased trade in fact increases the degree of effective competition amongst the producers prevailing within the market. This feature of the model is largely determined by the *cross-country correlation* in sectoral productivity draws. We choose the amount of correlation in sectoral draws so that the model reproduces standard estimates of the elasticity of trade flows with respect to changes in variable trade costs. As the amount of correlation increases, there is less cross-country variation in the productivity with which producers within a given sector operate. Consequently, small changes in trade costs have relatively larger effects on trade flows — in short, the *trade elasticity* is increasing in the amount of cross-country correlation. To match standard estimates of the trade elasticity, the benchmark model requires a relatively high 0.93 cross-country correlation in sectoral draws. This high correlation also allows the model to reproduce the strong positive relationship between a sector's share of domestic sales and its share of imports that we observe in the data — i.e., reproduces the fact that sectors with relatively large, productive firms are also sectors with relatively large import shares.

Given this high degree of correlation, opening to trade indeed reduces markup dispersion and increases aggregate productivity. For the benchmark model, calibrated to Taiwan's import share, opening to trade reduces markup distortions by about one-fifth and increases aggregate productivity by 12% relative to autarky. In short we find that, yes, opening to trade can lead to a quantitatively significant reduction in misallocation. We also find that these pro-competitive effects are strongest near autarky — the pro-competitive effects are more important for an economy opening from autarky to a 10% import share than for an economy increasing its openness from a 10% to 20% import share.

In the model, a given producer's productivity has both a sector-specific component and an idiosyncratic component, both drawn from Pareto distributions. In our benchmark model, the sectoral draws are correlated across countries but the idiosyncratic draws are uncorrelated. We also consider an extension of the model in which the idiosyncratic draws are also correlated across producers in a given sector in different countries. This extension is motivated by the observation that sectors with high concentration amongst domestic producers are also sectors with high import penetration. While our benchmark model cannot reproduce this feature of the data, our extension with correlated idiosyncratic draws can. This extension predicts an even larger role for trade in reducing markup distortions because countries import more of exactly those goods for which the domestic market is more distorted. In this version of the model, trade eliminates about one-third of the productivity losses from misallocation.

We consider a large number of robustness checks on our benchmark model — including allowing for heterogeneity in sector-level tariffs, introducing labor market distortions, and changing the mode of competition from Cournot to Bertrand, amongst others. Our main findings are robust to these alternative specifications. We also study an extension of the model in which we introduce capital and elastic labor supply and show that the pro-competitive gains from trade are even larger. Finally, we study a version of the model with free-entry and show that versions of the free-entry model that reproduce the salient features of the Taiwanese data continue to predict significant pro-competitive gains from trade.

Markups, misallocation, and trade. Recent papers by Restuccia and Rogerson (2008), Hsieh and Klenow (2009) and others show that misallocation of factors of production can substantially reduce aggregate productivity. We focus on the role of markup variation as a source of misallocation.[1] We find that, by reducing markup dispersion, trade can play a powerful role in reducing misallocation and can thereby increase aggregate productivity.

The possibility that opening an economy to trade may lead to welfare gains from increased competition is, of course, one of the oldest ideas in economics. But standard quantitative trade models, such as the perfect competition model of Eaton and Kortum (2002) or the monopolistic competition models with constant markups of Melitz (2003) and Chaney (2008), cannot capture this pro-competitive intuition.

Perhaps more surprisingly, existing trade models that *do* feature variable markups also do not generally predict pro-competitive gains from trade. For example, the Bernard, Eaton, Jensen and Kortum (2003, hereafter BEJK) model of Bertrand competition results in an endogenous distribution of markups, that, due to specific functional form assumptions, is exactly *invariant* to changes in trade costs and has exactly *zero* pro-competitive gains from trade.[2] Similarly, in the monopolistic competition models with non-CES demand[3] studied by Arkolakis, Costinot, Donaldson and Rodríguez-Clare (2012b, hereafter ACDR), the markup

[1]Two closely related papers are Peters (2013), who considers endogenous markups, as we do, in a closed economy quality-ladder model of endogenous growth and Epifani and Gancia (2011) who consider an open economy model but with exogenous markup dispersion.

[2]An important contribution by De Blas and Russ (2010) extends BEJK to allow for a finite number of producers in a given sector so that, as in our model, the distribution of markups varies in response to changes in trade costs. Holmes, Hsu and Lee (2011) study the impact of trade on productivity and misallocation in this setting. Relative to these theoretical papers, as well as to Devereux and Lee (2001) and Melitz and Ottaviano (2008), our main contribution is to *quantify* the pro-competitive gains from trade using micro data.

[3]Special cases of which include the non-CES demand systems used by Krugman (1979), Feenstra (2003), Melitz and Ottaviano (2008), and Zhelobodko, Kokovin, Parenti and Thisse (2012).

distribution is likewise invariant to changes in trade costs and there are in fact *negative* pro-competitive 'gains' from trade.

The reason models with variable markups yield conflicting predictions regarding the pro-competitive gains from trade is that, as emphasized by ACDR, what really matters for these effects is the *joint distribution* of markups and employment. The response of this joint distribution to a reduction in trade costs depends on details of the parameterization of the model, and in particular the amount of cross-country correlation in productivity draws. We show that versions of our model with low correlation do indeed predict negative pro-competitive gains. But such parameterizations also imply both (i) low aggregate trade elasticities, and (ii) a weak or negative relationship between a sector's share of domestic sales and its share of imports — and thus are inconsistent with empirical evidence.

Empirical literature on markups and trade. There is a large empirical literature on producer markups and trade, important early examples include Levinsohn (1993), Harrison (1994), and Krishna and Mitra (1998). Tybout (2003) reviews this literature and concludes that "in *every* country studied, relatively high sector-wide exposure to foreign competition is associated with lower price-cost margins, and the effect is concentrated in *larger* plants." More recently, Feenstra and Weinstein (2010) infer large markup reductions from observed changes in US market shares from 1992–2005. De Loecker, Goldberg, Khandelwal and Pavcnik (2012) study the effects of India's tariff reductions on both final goods and inputs and find that the *net* effect was in fact to *increase* markups — because input tariffs fell, so did the costs of final goods producers. When they condition on the effects of trade liberalization through inputs, however, De Loecker et al. find that the markups of final goods producers fall. In this sense, their results are consistent with our benchmark model.

There are important conceptual differences between the effects of trade in this literature and pro-competitive gains that operate through *reduced misallocation*. Documenting changes in the domestic markup distribution following a trade liberalization does not tell us whether misallocation has gone down or not. Again, what matters for misallocation is the response of the joint distribution of employment and markups of all producers, including exporters.

Trade flows and the gains from trade. Our focus on the gains from trade is related to the work of Arkolakis, Costinot and Rodríguez-Clare (2012a, hereafter ACR), who show that the total gains from trade are identical in a large class of models and are summarized by the aggregate trade elasticity. Interestingly, we find that the ACR formula provides an accurate approximation in our setup with variable markups. This is only the case, however, if we compute the trade elasticity as ACR do, namely as the responsiveness of trade flows to changes in *trade costs*, and not as the responsiveness of trade flows to changes in *relative* *prices* as is standard in the international macro literature. In our model, in contrast to standard trade models, the trade cost elasticity is generally lower than the relative price elasticity because variable markups imply incomplete pass-through from changes in trade costs to changes in prices.

That said, while the total gains from trade in our model are well approximated by the ACR formula, the *decomposition* of those gains into pro-competitive and other channels depends quite sensitively on the micro details of producer-level productivity and competition. Our model predicts, for example, that following a trade liberalization, an economy with very mild markup distortions will receive gains primarily through standard trade channels whereas an economy with extensive markup distortions may receive gains both through the pro-competitive channel *and* through standard trade channels.

The remainder of the paper proceeds as follows. Section 2 presents the model. Section 3 gives an overview of the data and Section 4 explains how we use that data to quantify the model. Section 5 presents our benchmark results on the gains from trade. Section 6 conducts a number of robustness checks. Section 7 presents results for two more significant extensions of our benchmark model, (i) trade between asymmetric countries, and (ii) free entry and an endogenous number of competitors per sector. Section 8 concludes.

2 Model

Our world consists of two symmetric countries, Home and Foreign. In keeping with standard assumptions in the trade literature, we assume a static environment with a single factor of production, labor, that is in inelastic supply and immobile between countries. We focus on describing the Home economy in detail. We indicate Foreign variables with an asterisk.

2.1 Final good producers

Perfectly competitive firms in each country produce a homogeneous final good for consumption. These final good firms produce using inputs from a continuum of *sectors*

$$Y = \left(\int_0^1 y(s)^{\frac{\theta-1}{\theta}} \, ds \right)^{\frac{\theta}{\theta-1}} , \tag{1}$$

where $\theta > 1$ is the elasticity of substitution *across* sectors $s \in [0,1]$. Importantly, each sector consists of a *finite* number of domestic and foreign intermediate producers. In sector s, output is produced using $n(s) \in \mathbb{N}$ domestic and $n(s)$ imported intermediate inputs

$$y(s) = \left(\sum_{i=1}^{n(s)} y_i^{\mathrm{H}}(s)^{\frac{\gamma-1}{\gamma}} + \sum_{i=1}^{n(s)} y_i^{\mathrm{F}}(s)^{\frac{\gamma-1}{\gamma}} \right)^{\frac{\gamma}{\gamma-1}} , \tag{2}$$

where $\gamma > \theta$ is the elasticity of substitution across goods i *within* a particular sector $s \in [0,1]$.

In our benchmark model, the number of potential producers $n(s)$ in sector s is exogenous and the same in both countries. In Section 7 below we consider an extension of the benchmark model with *free entry* that makes $n(s)$ endogenous.[4]

2.2 Intermediate goods producers

Intermediate producer i in sector s produces output with labor

$$y_i(s) = a_i(s) l_i(s)\,, \tag{3}$$

where producer-level productivity $a_i(s)$ is drawn from a distribution that we discuss in detail in Section 4 below.

Trade costs. An intermediate producer sells output to final goods producers located in both countries. Let $y_i^{\mathrm{H}}(s)$ denote the amount sold by a Home intermediate producer to Home final good producers and similarly let $y_i^{*\mathrm{H}}(s)$ denote the amount sold by a Home intermediate producer to Foreign final good producers. The resource constraint for Home intermediate producers is

$$y_i(s) = y_i^{\mathrm{H}}(s) + \tau\, y_i^{*\mathrm{H}}(s)\,, \tag{4}$$

where $\tau \geq 1$ is an iceberg trade cost, i.e., $\tau\, y_i^{*\mathrm{H}}(s)$ must be shipped for $y_i^{*\mathrm{H}}(s)$ to arrive abroad. Foreign intermediate producers face symmetric trade costs. We let $y_i^*(s)$ denote their output and note that the resource constraint facing Foreign intermediate producers is

$$y_i^*(s) = \tau\, y_i^{\mathrm{F}}(s) + y_i^{*\mathrm{F}}(s)\,, \tag{5}$$

where $y_i^{*\mathrm{F}}(s)$ denotes the amount sold by a Foreign intermediate producer to Foreign final good producers and $y_i^{\mathrm{F}}(s)$ denotes the amount sold by a Foreign intermediate producer to Home final good producers.

Demand for intermediate inputs. Final good producers buy intermediate goods from Home producers at prices $p_i^{\mathrm{H}}(s)$ and from Foreign producers at prices $p_i^{\mathrm{F}}(s)$. Consumers buy the final good at price P. The problem of a final good producer is to choose intermediate inputs $y_i^{\mathrm{H}}(s)$ and $y_i^{\mathrm{F}}(s)$ to maximize profits:

$$PY - \int_0^1 \Big(\sum_{i=1}^{n(s)} p_i^{\mathrm{H}}(s) y_i^{\mathrm{H}}(s) + \tau \sum_{i=1}^{n(s)} p_i^{\mathrm{F}}(s) y_i^{\mathrm{F}}(s) \Big)\, ds\,, \tag{6}$$

[4] In the Appendix we also report results for a version of our model where the numbers of Home and Foreign producers per sector remain exogenous but are uncorrelated across countries.

subject to (1) and (2). The solution to this problem gives the demand functions:

$$y_i^{\mathrm{H}}(s) = \left(\frac{p_i^{\mathrm{H}}(s)}{p(s)}\right)^{-\gamma} \left(\frac{p(s)}{P}\right)^{-\theta} Y\,, \tag{7}$$

and

$$y_i^{\mathrm{F}}(s) = \left(\frac{\tau p_i^{\mathrm{F}}(s)}{p(s)}\right)^{-\gamma} \left(\frac{p(s)}{P}\right)^{-\theta} Y\,, \tag{8}$$

where the aggregate and sectoral price indexes are

$$P = \left(\int_0^1 p(s)^{1-\theta}\, ds\right)^{\frac{1}{1-\theta}}\,, \tag{9}$$

and

$$p(s) = \left(\sum_{i=1}^{n(s)} \phi_i^{\mathrm{H}}(s) p_i^{\mathrm{H}}(s)^{1-\gamma} + \tau^{1-\gamma} \sum_{i=1}^{n(s)} \phi_i^{\mathrm{F}}(s) p_i^{\mathrm{F}}(s)^{1-\gamma}\right)^{\frac{1}{1-\gamma}}\,, \tag{10}$$

and where $\phi_i^{\mathrm{H}}(s) \in \{0,1\}$ is an indicator function that equals one if a producer operates in the Home market (its domestic market) and likewise $\phi_i^{\mathrm{F}}(s) \in \{0,1\}$ is an indicator function that equals one if a Foreign producer operates in the Home market (its export market).

Market structure. An intermediate good producer faces the demand system given by equations (7)-(10) and engages in *Cournot competition* within its sector.[5] That is, each individual firm chooses a given quantity $y_i^{\mathrm{H}}(s)$ or $y_i^{*\mathrm{H}}(s)$ taking as given the quantity decisions of its competitors in sector s. Due to constant returns, the problem of a firm in its domestic market and its export market can be considered separately.

Fixed costs. A fixed cost f_d must be paid in order to operate in the domestic market and a fixed cost f_x must be paid in order to export. Both of these are denominated in units of domestic labor. A firm can choose to produce zero units of output for the domestic market to avoid paying the fixed cost f_d. Similarly, a firm can choose to produce zero units of output for the export market to avoid paying the fixed cost f_x.

Domestic market. Taking the wage W as given, the problem of a Home firm in its domestic market can be written

$$\pi_i^{\mathrm{H}}(s) := \max_{y_i^{\mathrm{H}}(s)\,,\,\phi_i^{\mathrm{H}}(s)} \left[\left(p_i^{\mathrm{H}}(s) - \frac{W}{a_i(s)}\right) y_i^{\mathrm{H}}(s) - W f_d\right] \phi_i^{\mathrm{H}}(s)\,, \tag{11}$$

[5] In Section 6 below we solve our model under the alternative assumption of *Bertrand competition* and find similar results.

subject to the demand system above. The solution to this problem is characterized by a price that is a markup over marginal cost

$$p_i^{\mathrm{H}}(s) = \frac{\varepsilon_i^{\mathrm{H}}(s)}{\varepsilon_i^{\mathrm{H}}(s) - 1} \frac{W}{a_i(s)}, \tag{12}$$

where $\varepsilon_i^{\mathrm{H}}(s) > 1$ is the demand elasticity facing the firm in its domestic market. With the nested CES demand system above and Cournot competition, it can be shown that this demand elasticity is a weighted *harmonic* average of the underlying elasticities of substitution θ and γ, specifically

$$\varepsilon_i^{\mathrm{H}}(s) = \left(\omega_i^{\mathrm{H}}(s)\frac{1}{\theta} + (1 - \omega_i^{\mathrm{H}}(s))\frac{1}{\gamma}\right)^{-1}, \tag{13}$$

where $\omega_i^{\mathrm{H}}(s) \in [0, 1]$ is the firm's share of sectoral revenue in its domestic market

$$\omega_i^{\mathrm{H}}(s) := \frac{p_i^{\mathrm{H}}(s) y_i^{\mathrm{H}}(s)}{\sum_{i=1}^{n(s)} p_i^{\mathrm{H}}(s) y_i^{\mathrm{H}}(s) + \tau \sum_{i=1}^{n(s)} p_i^{\mathrm{F}}(s) y_i^{\mathrm{F}}(s)} = \left(\frac{p_i^{\mathrm{H}}(s)}{p(s)}\right)^{1-\gamma}. \tag{14}$$

For short, we refer to $\omega_i^{\mathrm{H}}(s)$ as a Home firm's domestic *market share*.

Export market. The problem of a Home firm in its export market is essentially identical except that to export (operate abroad) it pays a fixed cost f_x rather than f_d so that its problem is

$$\pi_i^{*\mathrm{H}}(s) := \max_{y_i^{*\mathrm{H}}(s),\, \phi_i^{*\mathrm{H}}(s)} \left[\left(p_i^{*\mathrm{H}}(s) - \frac{W}{a_i(s)}\right) y_i^{*\mathrm{H}}(s) - W f_x\right] \phi_i^{*\mathrm{H}}(s), \tag{15}$$

subject to the demand system abroad. Prices are again a markup over marginal cost

$$p_i^{*\mathrm{H}}(s) = \frac{\varepsilon_i^{*\mathrm{H}}(s)}{\varepsilon_i^{*\mathrm{H}}(s) - 1} \frac{W}{a_i(s)}, \tag{16}$$

where $\varepsilon_i^{*\mathrm{H}}(s) > 1$ is the demand elasticity facing the firm in its export market

$$\varepsilon_i^{*\mathrm{H}}(s) = \left(\omega_i^{*\mathrm{H}}(s)\frac{1}{\theta} + (1 - \omega_i^{*\mathrm{H}}(s))\frac{1}{\gamma}\right)^{-1}, \tag{17}$$

and where $\omega_i^{*\mathrm{H}}(s) \in [0, 1]$ is the firm's share of sectoral revenue in its export market

$$\omega_i^{*\mathrm{H}}(s) := \frac{\tau p_i^{*\mathrm{H}}(s) y_i^{*\mathrm{H}}(s)}{\tau \sum_{i=1}^{n(s)} p_i^{*\mathrm{H}}(s) y_i^{*\mathrm{H}}(s) + \sum_{i=1}^{n(s)} p_i^{*\mathrm{F}}(s) y_i^{*\mathrm{F}}(s)}. \tag{18}$$

For short, we refer to $\omega_i^{*\mathrm{H}}(s)$ as a Home firm's export market share.

Market shares and demand elasticity. In general, each firm faces a different, endogenously determined, demand elasticity. The demand elasticity is given by a weighted average of the within-sector elasticity γ and the across-sector elasticity $\theta < \gamma$. Firms with a small market share within a sector (within a given country) compete mostly with other firms in their own sector and so face a relatively high demand elasticity, closer to the within-sector γ. Firms with a large market share face relatively more competition from firms in other sectors than they do from firms in their own sector and so face a relatively low demand elasticity, closer to the across-sector θ. The markup a firm charges is an increasing *convex* function of its market share. An infinitesimal firm charges a markup of $\gamma/(\gamma - 1)$, the smallest possible in this model. At the other extreme, a pure monopolist charges a markup of $\theta/(\theta - 1)$, the largest possible in this model. Because of the convexity, a mean-preserving spread in market shares will increase the average markup.

The extent of *markup dispersion* across firms depends both on the *gap* between θ and γ and on the extent of dispersion in market shares. In the special case where $\theta = \gamma$, the demand elasticity is constant and independent of the dispersion in market shares and the model collapses to a standard trade model with constant markups. But if θ is substantially smaller than γ, then even a modest change in market share dispersion can have a large effect on markup dispersion and hence a large effect on aggregate productivity.

Notice also that a firm operating in both countries will generally have different market shares in each country and consequently face different demand elasticities and charge different markups in each country.

Market shares and markups. The formula (13) for a firm's demand elasticity implies a linear relationship between a firm's *inverse markup* and its market share

$$\frac{1}{\mu_i^{\mathrm{H}}(s)} = \frac{\gamma - 1}{\gamma} - \left(\frac{1}{\theta} - \frac{1}{\gamma}\right) \omega_i^{\mathrm{H}}(s)\,. \tag{19}$$

where $\mu_i^{\mathrm{H}}(s) := \varepsilon_i^{\mathrm{H}}(s)/(\varepsilon_i^{\mathrm{H}}(s) - 1)$ denotes the firm's gross markup from (12). Since $\theta < \gamma$, the coefficient on the market share $\omega_i^{\mathrm{H}}(s)$ is negative. Within a sector s, firms with relatively high market shares have low demand elasticity and high markups. As discussed in Section 4 below, the strength of this relationship plays a key role in identifying plausible magnitudes for the gap between the elasticity parameters θ and γ.

Operating decisions. Each firm must pay a fixed cost f_d to operate in its domestic market and a fixed cost f_x to operate in its export market. A Home firm operates in its domestic market so long as

$$\left(p_i^{\mathrm{H}}(s) - \frac{W}{a_i(s)}\right) y_i^{\mathrm{H}}(s) \geq W f_d \tag{20}$$

Similarly, a Home firm operates in its export market so long as

$$\Big(p_i^{*\mathrm{H}}(s) - \frac{W}{a_i(s)}\Big) y_i^{*\mathrm{H}}(s) \geq W f_x \tag{21}$$

There are multiple equilibria in any given sector. Different combinations of firms may choose to operate, given that the others do not. As in Atkeson and Burstein (2008), within each sector s we place firms in the order of their physical productivity $a_i(s)$ and focus on equilibria in which firms sequentially decide on whether to operate or not: the most productive decides first (given that no other firm operates), the second most productive decides second (given that no other less productive firm operates), and so on.[6]

2.3 Market clearing

In each country there is a representative consumer that inelastically supplies one unit of labor and that consumes the final good. The labor market clearing condition is

$$\int_0^1 \Big(\sum_{i=1}^{n(s)} (l_i^{\mathrm{H}}(s) + f_d)\phi_i^{\mathrm{H}}(s) + \sum_{i=1}^{n(s)} (l_i^{*\mathrm{H}}(s) + f_x)\phi_i^{*\mathrm{H}}(s) \Big)\, ds = 1\,, \tag{22}$$

and the market clearing condition for the final good is simply $C = Y$.

2.4 Aggregate productivity and markups

Aggregation. The quantity of final output in each country can be written

$$Y = A\tilde{L}, \tag{23}$$

where A is the endogenous level of aggregate productivity and $\tilde{L}$ is the aggregate amount of labor employed *net of fixed costs*. Using the firms' optimality conditions and the market clearing condition for labor, it is straightforward to show that aggregate productivity is a *quantity-weighted* harmonic mean of firm productivities

$$A = \left(\int_0^1 \Big(\sum_{i=1}^{n(s)} \frac{1}{a_i(s)} \frac{y_i^{\mathrm{H}}(s)}{Y} + \tau \sum_{i=1}^{n(s)} \frac{1}{a_i(s)} \frac{y_i^{*\mathrm{H}}(s)}{Y} \Big)\, ds \right)^{-1}. \tag{24}$$

Now denote the aggregate (economy-wide) markup by

$$\mu := \frac{P}{W/A}\,, \tag{25}$$

[6]The exact ordering makes little difference quantitatively when we calibrate the model to match the strong concentration in the data. Productive firms always operate and unproductive ones never do, so the equilibrium multiplicity only affects the operating decisions of marginal firms that have a negligible effect on aggregates. Moreover, as we show in Section 6 below, our model's implications for markup dispersion are essentially unchanged when we set $f_d = f_x = 0$ so that all firms operate and the equilibrium is unique.

that is, aggregate price divided by aggregate marginal cost. It is straightforward to show that the aggregate markup is a *revenue-weighted* harmonic mean of firm markups

$$\mu = \left(\int_0^1 \Big(\sum_{i=1}^{n(s)} \frac{1}{\mu_i^{\mathrm{H}}(s)} \frac{p_i^{\mathrm{H}}(s) y_i^{\mathrm{H}}(s)}{PY} + \tau \sum_{i=1}^{n(s)} \frac{1}{\mu_i^{*\mathrm{H}}(s)} \frac{p_i^{*\mathrm{H}}(s) y_i^{*\mathrm{H}}(s)}{PY} \Big) \, ds \right)^{-1}, \tag{26}$$

where $\mu_i^{\mathrm{H}}(s)$ denotes a Home firm's markup in its domestic market and $\mu_i^{*\mathrm{H}}(s)$ denotes its markup in its export market (implied by equations (12) and (16), respectively).

Misallocation and markup dispersion. In this model, dispersion in markups reduces aggregate productivity, as in the work of Restuccia and Rogerson (2008) and Hsieh and Klenow (2009). To understand this effect, first notice that the expression (24) for aggregate productivity can be written

$$A = \left(\int_0^1 \Big(\frac{\mu(s)}{\mu} \Big)^{-\theta} a(s)^{\theta-1} \, ds \right)^{\frac{1}{\theta-1}}, \tag{27}$$

where $\mu(s) := p(s)/(W/a(s))$ denotes the sector-level markup and where sector-level productivity is given by

$$a(s) = \left(\sum_{i=1}^{n(s)} \Big(\frac{\mu_i^{\mathrm{H}}(s)}{\mu(s)} \Big)^{-\gamma} a_i(s)^{\gamma-1} \phi_i^{\mathrm{H}}(s) + \tau^{1-\gamma} \sum_{i=1}^{n(s)} \Big(\frac{\mu_i^{\mathrm{F}}(s)}{\mu(s)} \Big)^{-\gamma} a_i^*(s)^{\gamma-1} \phi_i^{\mathrm{F}}(s) \right)^{\frac{1}{\gamma-1}}. \tag{28}$$

First-best aggregate productivity. By contrast, the first-best level of aggregate productivity (the best attainable by a planner, subject to the trade cost τ) associated with an efficient allocation of resources is

$$A_{\text{efficient}} = \left(\int_0^1 a(s)^{\theta-1} \, ds \right)^{\frac{1}{\theta-1}}, \tag{29}$$

where sector-level productivity is

$$a(s) = \left(\sum_{i=1}^{n(s)} a_i(s)^{\gamma-1} \phi_i^{\mathrm{H}}(s) + \tau^{1-\gamma} \sum_{i=1}^{n(s)} a_i^*(s)^{\gamma-1} \phi_i^{\mathrm{F}}(s) \right)^{\frac{1}{\gamma-1}}, \tag{30}$$

with operating decisions $\phi_i^{\mathrm{H}}(s), \phi_i^{\mathrm{F}}(s) \in \{0, 1\}$ as dictated by the solution to the planning problem. If there is no markup dispersion (as occurs, for example, if $\theta = \gamma$), then aggregate productivity from (27)-(28) is at its first-best level. But with markup dispersion, the most productive producers employ a smaller share of the economy's labor than efficiency dictates, since markups and productivity are positively correlated. Markup dispersion lowers aggregate productivity relative to the first-best because it induces an inefficient allocation of resources across producers. If opening to trade reduces markup dispersion, then the losses due to misallocation will be smaller and there will be pro-competitive gains from trade. If opening to trade increases markup dispersion, then the losses due to misallocation will be larger and the pro-competitive 'gains' will be negative, as they are in Arkolakis, Costinot, Donaldson and Rodríguez-Clare (2012b).

2.5 Trade elasticity

In standard trade models, and as emphasized by Arkolakis, Costinot and Rodríguez-Clare (2012a), the gains from trade are largely determined by the elasticity of trade flows with respect to changes in *trade costs*. With constant markups, this elasticity with respect to trade costs is the same as the elasticity with respect to changes in international *relative prices*. But with variable markups, as in our model, these two concepts are not generally the same.

Trade elasticity with respect to international relative prices. Suppose all foreign prices uniformly change by a factor q (this may be because of changes in trade costs, or productivity, or labor supply etc). We define the trade elasticity with respect to international relative prices as

$$\sigma_{\text{relative prices}} := \frac{d\log\frac{1-\lambda}{\lambda}}{d\log q}\,, \tag{31}$$

where λ denotes the aggregate share of spending on domestic goods,

$$\lambda := \frac{\int_0^1 \sum_{i=1}^{n(s)} p_i^{\mathrm{H}}(s)y_i^{\mathrm{H}}(s)\,ds}{\int_0^1 \Big(\sum_{i=1}^{n(s)} p_i^{\mathrm{H}}(s)y_i^{\mathrm{H}}(s) + \tau\sum_{i=1}^{n(s)} p_i^{\mathrm{F}}(s)y_i^{\mathrm{F}}(s)\Big)\,ds} = \int_0^1 \lambda(s)\omega(s)\,ds\,, \tag{32}$$

and where $\lambda(s)$ denotes the sector-level share of spending on domestically produced goods and $\omega(s) := (p(s)/P)^{1-\theta}$ is that sector's share of aggregate spending. Some algebra shows that, in our model, the trade elasticity with respect to international relative prices is given by a weighted average of the underlying elasticities of substitution γ and θ, specifically[7]

$$\sigma_{\text{relative prices}} = \gamma\left(\int_0^1 \frac{\lambda(s)}{\lambda}\Big(\frac{1-\lambda(s)}{1-\lambda}\Big)\omega(s)\,ds\right) + \theta\left(1-\int_0^1 \frac{\lambda(s)}{\lambda}\Big(\frac{1-\lambda(s)}{1-\lambda}\Big)\omega(s)\,ds\right) - 1$$

[7] Our goal here is to obtain analytic results that aid in building intuition. To that end, in deriving (33) we abstract from the extensive margin and hold the set of producers in each country fixed. We relax this assumption and determine the set of operating firms endogenously when we compute the trade elasticity in our model. It turns out that treating the set of producers as fixed is, quantitatively, a good approximation in our model. In particular, as we show in Section 6 below, the quantitative implications of our model are almost identical when there are no fixed costs and all producers operate in both countries.

so that on collecting terms

$$\sigma_{\text{relative prices}} = (\gamma - 1) - (\gamma - \theta)\frac{\text{Var}[\lambda(s)]}{\lambda(1-\lambda)}\,, \tag{33}$$

where $\text{Var}[\lambda(s)]$ is the variance across sectors of the share of spending on domestic goods and λ is the aggregate share, as defined in (32). For short, we refer to the term $\text{Var}[\lambda(s)]/\lambda(1-\lambda)$ as our *index of import share dispersion.* Notice that this elasticity is generally less than $\gamma - 1$ and is decreasing in the index of import share dispersion. If there is no import share dispersion, $\lambda(s) = \lambda$ for all s, then $\text{Var}[\lambda(s)] = 0$ and the elasticity is relatively high, equal to $\gamma - 1$. Intuitively, if all sectors have identical import shares then there is no across-sector reallocation of expenditure and a uniform reduction in the relative price of foreign goods symmetrically increases import shares within each sector, an effect governed by γ. At the other extreme, if import shares are binary, $\lambda(s) \in \{0, 1\}$, then $\text{Var}[\lambda(s)] = \lambda(1-\lambda)$ and the elasticity is relatively low, equal to $\theta - 1$. Here there is *only* across-sector reallocation of expenditure and a uniform reduction in the relative price of foreign goods induces reallocation towards sectors with high import shares, an effect governed by θ.

The elasticity $\sigma_{\text{relative prices}}$ is the trade elasticity as typically defined in the international macro literature. We now contrast this with the trade elasticity with respect to trade costs.

Trade elasticity with respect to trade costs. We follow standard practice in the trade literature and define the trade elasticity with respect to trade costs as

$$\sigma_{\text{trade costs}} := \frac{d\log\frac{1-\lambda}{\lambda}}{d\log\tau}\,, \tag{34}$$

In a standard model, with constant markups, $d\log q = d\log\tau$ so that $\sigma_{\text{trade costs}} = \sigma_{\text{relative prices}}$. But in our model, with variable markups, there is *incomplete pass-through*: a 1% fall in trade costs reduces the relative price of foreign goods by less than 1%.

To derive the trade elasticity with respect to trade costs in our model, begin by noting that at the sector level the responsiveness of trade flows is given by

$$\frac{d\log\frac{1-\lambda(s)}{\lambda(s)}}{d\log\tau} = (\gamma-1)(1+\epsilon(s))\,,$$

where

$$\epsilon(s) := \sum_{i=1}^{n(s)}\frac{p_i^{\text{F}}(s)y_i^{\text{F}}(s)}{p^{\text{F}}(s)y^{\text{F}}(s)}\Big(\frac{d\log\mu_i^{\text{F}}(s)}{d\log\tau}\Big) - \sum_{i=1}^{n(s)}\frac{p_i^{\text{H}}(s)y_i^{\text{H}}(s)}{p^{\text{H}}(s)y^{\text{H}}(s)}\Big(\frac{d\log\mu_i^{\text{H}}(s)}{d\log\tau}\Big)\,,$$

denotes the elasticity with respect to trade costs of Foreign markups relative to Home markups and where $p^{\text{F}}(s)y^{\text{F}}(s)$ and $p^{\text{H}}(s)y^{\text{H}}(s)$ denote spending on Foreign goods and spending on Home goods in sector s. In general, the relative markup elasticity $\epsilon(s)$ is *negative*

— i.e., a reduction in trade costs tends to increase Foreign markups as their producers gain market share and to decrease Home markups as their producers lose market share.

The aggregate trade elasticity with respect to trade costs can then be written

$$\sigma_{\text{trade costs}} = (\gamma-\theta)\left(\int_0^1 \frac{\lambda(s)}{\lambda}\Big(\frac{1-\lambda(s)}{1-\lambda}\Big)(1+\epsilon(s))\omega(s)\,ds\right) \\ +(\theta-1)\left(\int_0^1 \Big(\frac{1-\lambda(s)}{1-\lambda}\Big)(1+\epsilon(s))\omega(s)\,ds\right). \tag{35}$$

Notice that in the special case where the relative markup elasticity is the same in each sector, $\epsilon(s)=\epsilon$ for all s, equation (35) reduces to

$$\sigma_{\text{trade costs}} = \left((\gamma-1)-(\gamma-\theta)\frac{\text{Var}[\lambda(s)]}{\lambda(1-\lambda)}\right)(1+\epsilon)$$

and comparing this with (33) we see that, for this special case, $\sigma_{\text{trade costs}} = \sigma_{\text{relative prices}}(1+\epsilon)$. In the further special case of $\gamma=\theta$, so that markups are constant, then $\epsilon=0$ (there is complete pass-through) so that the trade elasticity with respect to trade costs is the same as with respect to relative prices and both trade elasticities equal $\gamma-1$. With variable markups, the trade elasticity is generally less than $\gamma-1$, both because the elasticity with respect to relative prices is less than $\gamma-1$ and because the elasticity with respect to trade costs is less than that with respect to relative prices.

3 Data

We now describe the data we use. First we give a brief description of the Taiwanese dataset. We then highlight facts about producer concentration in this data that are crucial for our model's quantitative implications. Finally, we outline how we infer markups from this data.

3.1 Dataset

We use the Taiwan Annual Manufacturing Survey. This survey reports data for the universe of establishments[8] engaged in production activities. Our sample covers the years 2000 and 2002–2004. The year 2001 is missing because in that year a separate census was conducted.

Product classification. The dataset we use has two components. First, an *establishment*-level component collects detailed information on operations, such as employment, expenditure on labor, materials and energy, and total revenue. Second, a *product*-level component

[8]In the Taiwanese data, almost all firms are single-establishment. In our Appendix we show that using firm-level data rather than establishment-level data makes almost no difference to our results. If anything, using establishments rather than firms *understates* the extent of concentration among producers, a key feature that determines the gains from trade in our model.

reports information on revenues for each of the products produced at a given establishment. Each product is categorized into a 7-digit Standard Industrial Classification created by the Taiwanese Statistical Bureau. This classification at 7 digits is comparable to the detailed 5-digit SIC product definition collected for US manufacturing establishments as described by Bernard, Redding and Schott (2010). Panel A of Table A1 in the Appendix gives an example of this classification, while Panel B reports the distribution of 7-digit sectors within 4- and 2-digit industries. Most of the products are concentrated in the Chemical Materials, Industrial Machinery, Computer/Electronics and Electrical Machinery industries.

Import shares. We supplement the survey with detailed import data at the harmonized HS-6 product level. We obtain the import data from the WTO and then match HS-6 codes with the 7-digit product codes used in the Annual Manufacturing Survey. This match gives us disaggregated import penetration ratios for each product category.

3.2 Concentration facts

The amount of producer concentration in the Taiwanese manufacturing data is crucial for our model's quantitative implications.

Strong concentration within sectors. We measure a producer's market share by their share of domestic sales revenue within a given 7-digit sector. Panel A of Table 1 shows that producers within a sector are highly concentrated. The top producer has a market share of around 40 to 45%.[9] The median inverse Herfhindhal (HH) measure of concentration is about 3.9, much lower than 10 or so producers that operate in a typical sector. The distribution of market shares is skewed to the right and extremely *fat-tailed*. The median market share of a producer is just 0.5% while the average market share is 4%. The 95th percentile accounts for only 19% of sales while the 99th percentile accounts for 59% of sales. The overall pattern that emerges is consistently one of very strong concentration. Although quite a few producers operate in any given sector, most of these producers are small and a few large producers account for the bulk of the sector's domestic sales.

Strong unconditional concentration. Panel A of Table 1 also reports statistics on the distribution of sales revenue and the wage bill across sectors and across all producers. The top 1% of sectors alone accounts for 26% of aggregate sales and 11% of the aggregate wage bill. The top 5% of sectors accounts for fully half of all sales and a third of the wage bill. This pattern is reproduced at the producer level. The top 1% of producers accounts for 41% of sales and 24% of the wage bill, the top 5% of producers accounts for nearly two-thirds all sales

[9] We weight each sector by the sector's share of aggregate sales.

and nearly a half of the wage bill. Again, the overall pattern is thus of strong concentration both within and across sectors.

3.3 Inferring markups

In our model, as is standard in the trade literature, labor is the only factor of production and a producer's revenue productivity (which is observable) *is* its markup. But in comparing our model's implications for markups to the data, it is important to recognize that, in general, revenue productivity differs across producers not only because of 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 that surely exist across Taiwanese manufacturing industries.[10]

Controlling for heterogeneity in producer technology. To map our model into micro-level production data, we relax the assumptions of a single factor of production and constant returns to scale. In particular, we follow De Loecker and Warzynski (2012) and assume a *translog* gross production function

$$\log y_i = \alpha_l \log l_i + \alpha_k \log k_i + \alpha_m \log m_i + \alpha_{ll}(\log l_i)^2 + \alpha_{kk}(\log k_i)^2 + \alpha_{mm}(\log m_i)^2 \\ + \alpha_{lk}(\log l_i \log k_i) + \alpha_{lm}(\log l_i \log m_i) + \alpha_{km}(\log k_i \log m_i) + \log a_i$$

where l_i denotes labor, k_i denotes physical capital, m_i denotes material inputs and a_i is physical productivity. We estimate this translog specification for each 2-digit Taiwanese industry, giving us industry-specific coefficient estimates. Let $e_{l,i}$ denote the elasticity of output with respect to labor, that is

$$e_{l,i} := \frac{\partial \log y_i}{\partial \log l_i} = \alpha_l + 2\alpha_{ll} \log l_i + \alpha_{lk} \log k_i + \alpha_{lm} \log m_i \tag{36}$$

In a standard Cobb-Douglas specification, this elasticity is the constant α_l, but here it varies across firms depending on their input use. Cost minimization then implies that producer i sets

$$\frac{W l_i}{p_i y_i} = \frac{e_{l,i}}{\mu_i} \tag{37}$$

Thus variation in labor input cost shares across producers may be due to either variation in markups μ_i or to variation in output elasticities $e_{l,i}$. Moreover, output elasticities may themselves vary both because of different levels of input use and because of different coefficients

[10] More precisely, under the maintained assumptions of Hicks-neutral technology and constants returns to scale, our model's implications for aggregate productivity in (27)-(28) depend only on the joint distribution of physical productivity a_i and markups μ_i and *do not* depend on the precise details of the producer-level production technologies. But for this argument to hold, we must in fact be credibly measuring the producer level productivity and markups and to do that we *do* need to control for heterogeneity in technology. As it turns out, our estimated production functions are very close to constant returns.

(i.e., because producers are in different 2-digit industries). We now use data on labor input cost shares and production function estimates of $e_{l,i}$ to back out markups μ_i from (37).

Estimating the translog production function. As is well-known, a key difficulty in estimating production functions is that input choices l_i, k_i, m_i will generally be correlated with true productivity a_i. We follow De Loecker and Warzynski (2012) and apply 'control' or 'proxy function' methods inspired by Olley and Pakes (1996), Levinsohn and Petrin (2003) and Ackerberg, Caves and Frazer (2006) to deal with this simultaneity. We give the full details of our implementation of this approach in the Appendix.

Estimation results. In Table 2, we report the median output elasticities and returns to scale for each of 21 Taiwanese manufacturing industries along with the inter-quartile range of output elasticities across producers within the same industry. Several points are worth noting: First, there is modest variation in output elasticities either within or across industries. For example, the 25th percentile of $e_{l,i}$ within industries is typically around 0.15 while the 75th percentile is typically around 0.4 with the standard deviation of median $e_{l,i}$ across industries being 0.04. Second, the median returns to scale within each industry is very close to 1 for almost all industries. In addition, the variation in returns to scale across producers within an industry is small, with the 25th percentile around 0.98 and the 75th percentile around 1.04. Third, the ranking of capital intensity across industries is intuitive, with Petroleum, Chemical Material, Computer, Machinery Equipment the most capital intensive, and Wood, Leather, Motor Vehicle Parts, Apparel the least.

Markup estimates. Given these estimates of $\widehat{e_{l,i}}$ for each producer for each industry, we recover 'measured inverse markups' $\widehat{1/\mu_i}$ from (37) as in De Loecker and Warzynski (2012). Panel A of Table 3 reports summary statistics of the distribution of markups obtained in this way. The estimated markups are highly dispersed, the 95th percentile markup is nearly 2.5 times the median markup and the 99th percentile markup is nearly 5 times the median. We also report the sector-level counterparts of these markup statistics; in accordance with the model, we measure sector-level markups as the revenue-weighted harmonic average of producer markups within a given sector. The sector-level markups are similarly dispersed.

Our theoretical model motivates a simple linear relationship between inverse markups and observed market shares, ω_i, namely

$$\widehat{1/\mu_i} = \beta_\mu + \beta_\omega \omega_i + \xi_{\mu,i} \tag{38}$$

One of the moments we will match in our model parameterization is the regression coefficient β_ω. In keeping with our theoretical model, we assume that the measured inverse markups are only systematically effected by producer market shares such that any residual markup variation, $\xi_{\mu,i}$, is orthogonal.[11] Under this assumption, the regression coefficient on market shares is simply $\beta_\omega = -(\frac{1}{\theta} - \frac{1}{\gamma})$. Given an estimate $\widehat{\beta_\omega}$ and a value for the within-sector elasticity γ, we can then calculate our estimate of the across-sector elasticity θ. Panel B of Table 3 reports the coefficient $\widehat{\beta_\omega}$ we obtain from regressing the De Loecker and Warzynski (2012) measured inverse markups $\widehat{1/\mu_i}$ on observed market shares ω_i using samples of single-product and multi-product producers. The market share coefficient is in a tight range around -0.66 to -0.69 across these regressions.

We also report moments for *projected markups*. These are moments of the inverse of the fitted values from (38), i.e., moments of $1/(\widehat{\beta_\mu} + \widehat{\beta_\omega}\omega_i)$, which we normalize by setting the intercept equal to its theoretical value, $\widehat{\beta_\mu} = \frac{\gamma-1}{\gamma}$ (the markup level does not affect allocations in our benchmark model). These projected markups are less dispersed, the 95th percentile sectoral markup is about 1.5 times the median and the 99th percentile is about 2.5 times the median. Since our model abstracts from any source of markup variation other than market share variation, we view these projected markups as being the natural empirical counterpart to the markups implied by our model. Moreover, since these projected markups are less dispersed than the measured markups, this choice means that, if anything, we will understate the amount of misallocation.

4 Quantifying the model

We now explain how we use the Taiwanese data to pin down the key parameters of our model.

4.1 Overview

In the model, the size of the gains from trade largely depends on two factors: (i) the extent of misallocation, and (ii) the responsiveness of that misallocation to changes in trade costs. In turn, these factors are largely determined by the joint distribution of productivity, both within and across countries, and on the elasticity of substitution parameters θ and γ. We discipline our model along these dimensions as follows.

We choose a *within-country* distribution of productivities so that our model reproduces the amount of concentration within and across sectors documented in the Taiwanese data. We choose the *gap* between the elasticities θ and γ so that our model reproduces the negative correlation between inverse markups and market shares. Together these determine the extent of misallocation in our benchmark economy. Given our within-country distribution of productivities, the cross-country joint distribution of productivities in our model is pinned down

[11] We consider an example with non-orthogonal residuals in Section 6 below where we allow for labor market distortions that are correlated with producer productivity and hence with producer market shares.

by one remaining parameter, the cross-country *correlation* in productivities at the producer level. We choose this correlation so that our model reproduces standard estimates of the trade elasticity.

4.2 Productivity distribution

The distribution of producer-level productivities $a_i(s)$ and $a_i^*(s)$ within sectors, across sectors, and across countries plays a key role in our analysis. Within a given country, the distribution of $a_i(s)$ determines the pattern of concentration within and across sectors and thus crucially shapes the extent of misallocation in the economy. Across countries, the correlation between $a_i(s)$ and $a_i^*(s)$ within a given sector determines the extent to which opening up to trade exposes highly productive domestic firms to competition from similarly productive foreign firms. If Home and Foreign productivities are strongly correlated within a sector, then opening up to trade implies that highly productive firms face strong competition that reduces their market share and hence reduces their markups. By contrast, if Home and Foreign productivities are weakly correlated then trade does not much affect the amount of competition and so has little effect on markups.

Within-country productivity distribution. We assume that across sectors the number of producers $n(s) \in \mathbb{N}$ is drawn IID Geometric with parameter $\zeta \in (0,1)$ so that $\text{Prob}[n] = (1-\zeta)^{n-1}\zeta$ and the average number of producers per sector is $1/\zeta$. We assume that an individual producer's productivity $a_i(s)$ is the product of a sector-specific component and an idiosyncratic component

$$a_i(s) = z(s)x_i(s)\,. \tag{39}$$

We assume $z(s) \geq 1$ is independent of $n(s)$ and across sectors is drawn IID Pareto with shape parameter $\xi_z > 0$. Within sector s, the $n(s)$ draws of the idiosyncratic component $x_i(s) \geq 1$ are IID Pareto across producers with shape parameter $\xi_x > 0$.

Cross-country productivity distribution. We assume that cross-country correlation in productivity arises through correlation in sectoral productivities. In particular, let $F_Z(z)$ denote the Pareto distribution of sector-specific productivities within each country and let $H_Z(z, z^*)$ denote the cross-country joint distribution of these sector-specific productivities. We write this cross-country joint distribution as

$$H_Z(z, z^*) = \mathcal{C}(F_Z(z), F_Z(z^*))\,, \tag{40}$$

where the *copula* $\mathcal{C}$ is the joint distribution of a pair of uniform random variables u, u^* on $[0,1]$. This formulation allows us to first specify the marginal distribution $F_Z(z)$ so as to match within-country productivity statistics and to then use the copula function to control the pattern of dependence between z and z^*.

Specifically, we assume that the marginal distributions are linked by a *Gumbel copula*, a widely used functional form that allows for dependence even in the right tails of the distribution

$$\mathcal{C}(u, u^*) = \exp\Big(- [(-\log u)^\rho + (-\log u^*)^\rho]^{1/\rho}\Big), \qquad \rho \geq 1\,. \tag{41}$$

The parameter ρ controls the pattern of dependence with higher values of ρ giving more dependence. If $\rho = 1$, the copula reduces to $\mathcal{C}(u, u^*) = uu^*$ so that the draws are independent. If $\rho \to \infty$ then, as is familiar from CES functions, the Copula approaches $\mathcal{C}(u, u^*) = \min[u, u^*]$ so that the draws are perfectly dependent. When working with heavy-tailed distributions it is standard to summarize dependence using the robust correlation coefficient known as *Kendall's tau*,[12] which we denote by $\tau(\rho)$ to distinguish it from the trade cost. With the Gumbel copula, this evaluates to $\tau(\rho) = 1 - 1/\rho$. Notice that $\tau(1) = 0$ (independence) and $\tau(\infty) = 1$ (perfect dependence). Once the within-country distribution $F_Z(z)$ has been specified, the single parameter $\tau(\rho)$ pins down the joint distribution $H_Z(z, z^*)$.

Finally, let $F_X(x)$ denote the Pareto distribution of idiosyncratic productivities within each sector and let $H_X(x, x^*)$ denote the associated joint distribution. For our benchmark model we assume these are independent across countries so that $H_X(x, x^*) = F_X(x)F_X(x^*)$.

4.3 Calibration

Elasticities of substitution. Following Atkeson and Burstein (2008), we directly assign the value $\gamma = 10$ to the within-sector elasticity of substitution.[13] We choose the across-sector elasticity of substitution θ so that our model reproduces the correlation between inverse markups and market shares implied by the regression (19). In particular, we choose $\theta = 1.28$ so that a regression of inverse markups on market shares gives a slope coefficient of $-(1/\theta - 1/\gamma) = -0.68$, squarely in the range of such coefficients we recover from the De Loecker and Warzynski (2012) procedure outlined above.

Given these elasticities of substitution, we then simultaneously choose the remaining parameters so that our model reproduces key features of the Taiwanese manufacturing data. Panel A of Table 1 reports the moments we target and the counterparts for our benchmark model. Panel B reports the parameter values that achieve this fit. We now briefly summarize the key features of the data that pin down the various parameters.

[12] Defined by:

$$\tau(\rho) := 4 \int_0^1 \int_0^1 \mathcal{C}(u, u^*)\, d\mathcal{C}(u, u^*),$$

which for the Gumbel copula in (41) evaluates to $\tau(\rho) = 1 - 1/\rho$.

[13] We discuss the robustness of our results to alternative values for γ in Section 6 below.

Number of producers, productivity, and fixed cost of operating. We choose the parameters ζ, ξ_z, ξ_x governing the within-country productivity distribution and the fixed cost f_d of operating in the domestic market to match key concentration statistics in the Taiwanese manufacturing data. Our model successfully reproduces the amount of concentration in the data. Within a given sector, the largest producer accounts for an average 45% of that sector's domestic sales. The model also reproduces the heavy concentration in the tails of the distribution of market shares with the 99th percentile share being 59% in both model and data. Moreover, the model also produces a fat-tailed size distribution of sectors and a fat-tailed size distribution of producers. The 99th percentile of sectors accounts for 24% of domestic sales (26% in the data) while the 99th percentile of producers accounts for 37% of domestic sales (41% in the data). The median number of producers per sector is a little too high (16 in the model, 10 in the data) but the model reproduces well the dispersion in the number of producers per sector (the 10th percentile is 3 producers in the model and 2 in the data, the 90th percentile is 47 producers in the model and 52 in the data).

The within-country joint distribution of productivity $a_i(s) = z(s)x_i(s)$ that generates this concentration is likewise very fat-tailed. This mostly comes from the sectoral productivity effect, $z(s)$, which has Pareto shape parameter $\xi_z = 0.56$. By contrast, the idiosyncratic productivity effect, $x_i(s)$, has relatively thin tails with Pareto shape parameter $\xi_x = 4.53$. The fixed cost to operate domestically is quite small, $f_d = 0.004$. This is about 0.16% of the average domestic producer's profits and 0.05% of their wage bill.

Trade costs. We choose the proportional trade cost τ and the fixed cost of operating in the export market f_x so that the model reproduces Taiwan's aggregate import share of 0.38 and aggregate fraction of firms that export of 0.25. The model achieves this with a trade cost of $\tau = 1.128$ (i.e., 1.128 units a good must be shipped for 1 unit to arrive) and a quite large fixed cost of operating in the export market, $f_x = 0.211$. This is about 3.39% of the average exporter's profits and 1.02% of their wage bill.

Trade elasticity and import share dispersion statistics. Finally, we choose the copula parameter $\tau(\rho)$ governing the degree of cross-country correlation in sectoral productivity so that, jointly with all of our other parameters, our model produces realistic values for (i) the trade elasticity, as well as (ii) the cross-sectional relationship between sector import shares and sector domestic size,[14] (iii) the amount of import share dispersion, and (iv) the amount of intra-industry trade. We target a trade elasticity of 4, a fairly standard estimate from aggregative data on trade flows — especially when one considers a two-country setting

[14]Specifically, the slope coefficient in a regression of sector imports out of total imports on sector domestic sales out of total domestic sales.

like ours. For the other import share statistics we simply target their counterparts in the Taiwanese data. Because the gains from trade depend crucially on the trade elasticity we assign a much larger weight to this moment, ensuring we match it exactly, and thus we slightly miss on the other statistics.

In the model, the trade elasticity is increasing in $\tau(\rho)$. This is because as the amount of correlation increases, there is less cross-country variation in the productivity with which producers within a given sector operate so that small changes in trade costs then have relatively larger effects on trade flows. To match a trade elasticity of 4 our model requires $\tau(\rho) = 0.93$ so that there is a high degree of correlation in productivity draws across countries. We discuss the sensitivity of our results to this value for $\tau(\rho)$ at length below.

4.4 Markup distribution

Table 4 reports moments of the distribution of markups $\mu_i(s)$ in our benchmark model and their counterparts in the data (these are the *projected markups*, implied by the fitted values from (38), as discussed above). We compare these to an economy that is identical except that we shut down international trade.

As shown in Panel A of Table 4, the benchmark model implies an average markup of 1.14, a median markup of 1.12 (just above the minimum $\gamma/(\gamma-1) = 1.11$) and a standard deviation of log markups of 0.08. These are very close to their data counterparts. Moreover, as in the data larger producers have considerably higher markups. The 95th percentile markup is 1.31 (compared to 1.20 in the data) and the 99th percentile markup is 1.67 (compared to 1.48 in the data) — though note that these are still a long way short of the $\theta/(\theta - 1) = 4.57$ markup a pure monopolist would charge in our model. Because large producers charge higher markups, the aggregate markup, which is a revenue-weighted harmonic average of the individual markups, is 1.31 — much higher than the simple average.

Let $\mu(s) = p(s)/(W/a(s))$ denote the aggregate markup in sector s. This sector-level markup $\mu(s)$ is likewise a revenue-weighted harmonic average of the producer-level markups $\mu_i(s)$ within that sector. Both in the model and in the data, these sector-level markups $\mu(s)$ are larger and more dispersed than their producer-level counterparts $\mu_i(s)$. In the model, the median sectoral markup is 1.29 as opposed to 1.12 for producers while the 99th percentile sectoral markup is 2.13 as opposed to 1.67 for producers. Thus, there are potentially large gains from reduced dispersion in markups across sectors as well as from reduced markup dispersion within sectors. Note however that the model fails to replicate the full extent of the across-sector variation in markups, especially in the tails. The 99th percentile markup in the data is 3.12, as opposed to 2.13 in the model. Since the actual dispersion in markups across sectors is considerably larger than in the model, this suggests we will, if anything, *understate* the true gains from reduced markup dispersion.

Now consider what happens when we shut down all international trade. The median markup does not change, nor does the 75th percentile markup. Rather markups in the far *tails* of the distribution rise: the 95th percentile markup increases from 1.31 to 1.35 and the 99th percentile markup increases from 1.67 to 1.76. Markup dispersion increases, with the standard deviation of log markups rising from 0.08 in the benchmark to 0.10 under autarky, with almost all of this increase in markup dispersion coming from a fanning out of the tails. Even more significantly, the distribution of sector-level markups experiences a considerable increase in dispersion, with the 95th percentile sectoral markup increasing from 1.80 to 2.25 and the 99th percentile markup increasing from 2.13 to 4.57 as some sectors become pure monopolies. This increase in markup dispersion suggests there will be more misallocation under autarky than in the benchmark economy.

Indeed, as shown in Panel B of Table 4, the benchmark economy implies aggregate productivity 6.7% below the first-best level of productivity associated with the planning allocation. Under autarky, the economy is 8.5% below the first-best. In this sense, the extent of misallocation is considerably worse under autarky.

5 Gains from trade

We now calculate the aggregate productivity gains from trade in our benchmark model. As in Arkolakis, Costinot and Rodríguez-Clare (2012a), we focus on the gains from trade due to a permanent reduction in trade costs τ.

Total gains from trade. We measure the gains from trade by the log percentage change in aggregate productivity from one equilibrium to another (the percentage change in aggregate consumption is very similar). As reported in Panel B of Table 4, for our benchmark economy the total gains from trade are a 12.0% increase in aggregate productivity relative to autarky. This is, of course, an extreme comparison. In Table 5 we report the gains from trade for intermediate degrees of openness. In particular, holding all other parameters fixed, we change the proportional trade cost τ so as to induce import shares of 0 (autarky), 10%, 20%, 30% and 38% (the Taiwan benchmark).

The model predicts a 3.1% increase in aggregate productivity moving from autarky to an import share of 10%. Moving further to an import share of 20% adds another 2.8% so that the cumulative gain moving from autarky to 20% is $3.1 + 2.8 = 5.9\%$. Continuing all the way to Taiwan's openness gives the 12.0% benchmark gains (relative to autarky) discussed above. Local to the Taiwan benchmark, a 1% change in openness is associated with an approximately 0.38% change in aggregate productivity. Put differently, an increase in trade costs resulting in a relatively modest 1% fall in the import share lowers Taiwanese aggregate productivity by 0.38% relative to the benchmark.

Arkolakis, Costinot and Rodríguez-Clare (2012a) show that, in a large class of models, the gains from trade are summarized by the formula $\frac{1}{\sigma}\log(\lambda/\lambda')$ where σ is the trade elasticity with respect to variable trade costs, as in (34) above, and where λ and λ' denote the aggregate share of spending on domestic goods before and after the change in trade costs. According to this formula, moving from autarky to an import share of 10% with a trade elasticity of 4.2 (which is what our model implies for that degree of openness) gives gains of $\frac{1}{4.2}\log(1/0.9) = 0.025$ or 2.5%. This is reasonably close to the 3.1% we find in our model. Similarly, according to this formula, moving from autarky to Taiwan's import share gives total gains of 11.8%, remarkably close to the 12.0% we find in our model. In short, even though our model with variable markups is not nested by the ACR setup, we find that their formula still provides a good approximation to the total gains from trade in our setting.

Pro-competitive gains from trade. We now isolate the gains from trade that are attributable to pro-competitive effects. In our benchmark model, all pro-competitive effects operate through changes in misallocation (i.e., changes in markup dispersion). Thus the most straightforward summary of the pro-competitive effects of trade is the implied change in misallocation. Under autarky, the economy is 8.5% below the first-best level of productivity. With a 10% import share, the economy is 7.3% below the first-best. So, as reported in Table 5, with an import share of 10% misallocation relative to autarky is $7.3/8.5 = 0.86$. Opening further to an import share of 20% gives misallocation relative to autarky of 0.83. Opening to Taiwan's import share gives misallocation relative to autarky of 0.79, a reduction in misallocation of some 21%. Notice that the extent of the reduction in misallocation, and hence the strength of the pro-competitive effects, is largest near autarky and then diminishes in relative importance as the economy experiences increasing degrees of openness.

In Table 5 we also measure the pro-competitive gains from trade as the total gains from trade *less* the log percentage change in first-best productivity. In a model with constant markups, aggregate productivity equals first-best productivity (the equilibrium allocation is efficient) and hence there are zero pro-competitive gains. The pro-competitive gains will be positive if increased trade reduces misallocation so that the increase in aggregate productivity is larger than the increase in first-best productivity. The pro-competitive 'gains' will be negative if increased trade increases misallocation. For our benchmark model, opening to trade reduces misallocation so we see here that indeed aggregate productivity increases by more than first-best productivity so that there are positive pro-competitive gains. Opening from autarky to Taiwan's import share gives pro-competitive gains of 1.8%.

Finally, while the trade elasticity changes with the degree of openness the changes are in fact relatively modest, varying from 4.2 at an import share of 10% to 4 at the benchmark.

Domestic vs. import markups. As emphasized by Arkolakis, Costinot, Donaldson and Rodríguez-Clare (2012b), the overall sign of the pro-competitive effect depends on markup responses of producers both in their domestic market and in their export market. It can be the case that a reduction in trade barriers leads to lower domestic markups (as Home producers lose market share) combined with higher markups on imported goods (as Foreign producers gain market share) such that overall markup dispersion increases and misallocation is worse — in which case the pro-competitive 'gains' from trade would be negative. In short, looking only at the markups of domestic producers may be misleading. As reported in Table 5, we indeed see that markups on imported goods do increase as the economy opens to trade, the revenue-weighted harmonic average of markups on imported goods increases by 15.8% as the economy opens from autarky (where Foreign producers have infinitesimal market share) to an import share of 10% while the corresponding average for domestic (Home) markups falls by 1.4%. The latter fall receives much more weight in the economy-wide aggregate markup so that overall the aggregate markup falls 1.6%. Notice that the fall in the aggregate markup is larger than the fall in domestic markups alone. This is due to a compositional effect. In particular, although markups on imported goods are rising while domestic markups are falling, the *level* of domestic markups is higher than the level of markups on imported goods. As the economy opens, the aggregate markup falls both because the high domestic markups of Home producers are falling and because a greater share of spending is on low-markup imports from Foreign producers.

Role of cross-country correlation in productivity. To match an aggregate trade elasticity of 4, our benchmark model requires a quite high degree of cross-country correlation in sectoral productivity draws, $\tau(\rho) = 0.93$. This degree of sectoral correlation implies, that, following a reduction in trade barriers, there is a correspondingly high degree of *head-to-head competition* between producers within any given sector. In Panel A of Table 6, we show the sensitivity of our results to the extent of correlation in sectoral productivity. For each level of $\tau(\rho)$ shown, we recalibrate our model to match our original targets except for the trade elasticity and related import share dispersion statistics. As we reduce $\tau(\rho)$, the model trade elasticity falls monotonically, reaching values of less than 1. Corresponding to these low trade elasticities are extremely high total gains from trade. Mechanically, the trade elasticity falls because the index of import share dispersion $\text{Var}[\lambda(s)]/\lambda(1-\lambda)$, i.e., the coefficient on θ in equation (33) above, *rises* as $\tau(\rho)$ falls. That is, an increasing proportion of sectors are either completely dominated by domestic producers (with import shares close to 0) or completely dominated by foreign producers (with import shares close to 1) so that the trade elasticity depends more on the across-sector θ and less on the within-sector elasticity γ.

When the correlation $\tau(\rho)$ is high, sectoral productivity draws are similar across countries so that most trade is *intra-industry*. In this case, a given change in trade costs gives rise to relatively large changes in trade flows. Panel A of Table 6 shows that the Grubel and Lloyd (1971) index of intra-industry trade is monotonically decreasing in $\tau(\rho)$, falling from 0.5 for our benchmark model (meaning, 50% of trade is intra-industry) to less than 0.1 for $\tau(\rho) < 0.5$. We also note that, in our benchmark model, there is a strong positive relationship between a sector's share of domestic sales and its share of imports. In particular, the slope coefficient in a regression of sector imports as a share of total imports on sector domestic sales as a share of total domestic sales is about 0.77 — i.e., sectors with relatively large, productive firms are also sectors with relatively large import shares, which is suggestive of firms in these sectors facing a great deal of head-to-head competition. When we reduce $\tau(\rho)$ we find this regression coefficient falls, eventually becoming slightly negative, so that large sectors no longer have large import shares, suggesting domestic producers no longer face as much competition when $\tau(\rho)$ is low.

Importantly, when the correlation $\tau(\rho)$ is sufficiently low a reduction in trade costs actually increases misallocation so that, as in Arkolakis, Costinot, Donaldson and Rodríguez-Clare (2012b), the pro-competitive 'gains' from trade are negative. To understand this, begin with an economy with high correlation, $\tau(\rho) = 0.9$ (similar to our benchmark). As shown in Panel B of Table 6, the 99th percentile of the *domestic* markup distribution falls from about 1.76 to 1.61, a fall of some 8.7%. Since markups near the median of the distribution change very little, this also represents a substantial fall in markup dispersion across domestic producers. Ultimately this fall in markups at the top of the distribution is a consequence of these domestic producers losing substantial market share to foreign competition. By contrast, with less correlation in draws, say $\tau(\rho) = 0.1$, opening from autarky to trade reduces the 99th percentile of domestic markups by only 2.3%. With less correlation, these dominant domestic producers lose less market share and hence their markups fall by less than with high correlation. Since markups near the median again change very little, this means there is a smaller fall in markup dispersion across domestic producers. Indeed, with $\tau(\rho) = 0.1$ the fall in domestic markup dispersion is sufficiently small that it is dominated by the rise in markup dispersion for imported goods so that, overall, misallocation is actually worse. In this case, the increased misallocation subtracts about 0.5% from the total gains from trade (which are nonetheless large here, because of the counterfactually low trade elasticity with $\tau(\rho) = 0.1$).

In Panel A of Table 6, we also report the data counterparts of the index of import share dispersion, the Grubel and Lloyd index, and the coefficient of size on import shares. To match these, our model requires $\tau(\rho)$ in the range 0.8 to 1.0 (depending on how much weight is given to each measure) with the aggregate trade elasticity then being in the range 2.9 to 4.4. In short, to match the facts on import share dispersion and intra-industry trade, the model requires a high degree of cross-country correlation in productivity draws.

Alternative model: cross-country correlation in idiosyncratic draws. As a final way to see the importance of head-to-head competition, we provide results for an alternative version of our model where there is correlation in *both* sectoral productivities $z(s)$ and in producer-specific idiosyncratic draws $x_i(s)$. Specifically we assume $H_Z(z, z^*) = \mathcal{C}_Z(F_Z(z), F_Z(z^*))$ and $H_X(x, x^*) = \mathcal{C}_X(F_X(x), F_X(x^*))$ both linked via a Gumbel copula as in (41) but with distinct correlation coefficients, $\tau_z(\rho)$ and $\tau_x(\rho)$. The benchmark model is then the special case $\tau_z(\rho) = 0.93$ and $\tau_x(\rho) = 0$. We recalibrate this model targeting the same moments as our benchmark model plus one new moment that helps identify $\tau_x(\rho)$. In particular, we choose $\tau_x(\rho)$ so that our model reproduces the cross-sectional relationship between sectoral *import penetration* and sectoral *concentration amongst domestic producers* that we observe in the Taiwanese data. In the data, the slope coefficient in a regression of sector import penetration on sector domestic HH indexes is 0.21 — i.e., sectors with high import penetration are also sectors with relatively high concentration amongst domestic producers.[15] To match this, we require a modest degree of cross-country correlation in idiosyncratic draws, $\tau_x(\rho) = 0.22$. The required cross-country correlation in sectoral productivity is correspondingly slightly lower, $\tau_z(\rho) = 0.90$, down from the benchmark 0.93.

As reported in Panel B of Table 5, this alternative model implies very similar total gains from trade, 11.7% versus the benchmark 12%, but because dominant producers face more head-to-head competition there are now larger pro-competitive effects. Opening from autarky to Taiwan's import share now reduces misallocation by almost one-third and the the pro-competitive gains are 2.6%, up from the benchmark 1.8%. Here, trade plays a larger role in reducing markup distortions because countries import more of exactly those goods for which the domestic market is in fact more distorted.

Capital accumulation and elastic labor supply. In the benchmark model the only gains are from changes in aggregate productivity and hence the only source of pro-competitive gains is changes in markup dispersion. The aggregate markup falls 2.8% between autarky and the Taiwan benchmark but this change in the aggregate markup has no welfare implications. But with capital accumulation and/or elastic labor supply, the aggregate markup acts like a distortionary wedge affecting investment and labor supply decisions, and, because of this, a reduction in the aggregate markup increases welfare beyond the increases associated with a reduction in markup dispersion. In particular, suppose the representative consumer has intertemporal preferences $\sum_{t=0}^{\infty} \beta^t U(C_t, L_t)$ over aggregate consumption C_t and labor L_t and that capital is accumulated according to $K_{t+1} = (1-\delta)K_t + I_t$. Suppose also that individual producers have production function $y = ak^{\alpha}l^{1-\alpha}$. We solve this version of the model assuming

[15] For our benchmark model, the slope coefficient of sector import penetration on sector domestic HH indexes is 0.08, low relative to the 0.21 in the data. See the Appendix for more details.

utility function $U(C, L) = \log C - L^{1+\eta}/1 + \eta$, discount factor $\beta = 0.96$, depreciation rate $\delta = 0.1$, output elasticity of capital $\alpha = 1/3$ and various elasticities of labor supply η. We start the economy in autarky and then compute the transition to a new steady-state corresponding to the Taiwan benchmark. We measure the welfare gains as the consumption compensating variation taking into account the dynamics of consumption and employment during the transition to the new steady-state. For our benchmark experiment, TFP increases by 12% of which 1.8% is due to pro-competitive effects. As reported in Table 7, with capital accumulation and a Frisch elasticity of 1, the welfare gains are 17.6% of which 3.3% is due to pro-competitive effects.

6 Robustness experiments

We now consider variations on our benchmark model, each designed to examine the sensitivity of our results to parameter choices or other assumptions. For each robustness experiment we recalibrate the trade cost τ, export fixed cost f_x, and correlation parameter $\tau(\rho)$ so that the Home country continues to have an aggregate import share of 0.38, fraction of exporters 0.25 and trade elasticity 4, as in our benchmark model. A summary of these robustness experiments is given in Table 8. Further details and a full set of results for these experiments are reported in the Appendix.

Heterogeneous labor market distortions. Our benchmark model focuses on the importance of product market distortions but ignores the role of labor market distortions. We now show that this is not essential for our main results. We assume that there is a distribution of producer-level labor market distortions that act like labor input taxes, putting a *wedge* between labor's marginal product and its factor cost. Specifically, a producer with productivity a also faces an input tax $t(a)$ on its wage bill so that it pays $(1+t(a))W$ for each unit of labor hired. We assume $t(a) = \frac{a\tau_l}{1+a\tau_l}$ and choose the parameter τ_l governing the sensitivity of the labor distortion to producer productivity so that our model matches the spread between the average producer labor share and the aggregate labor share that we observe in the data. In the data, the average producer labor share is 1.35 times the aggregate labor share. Since the latter is a weighted version of the former, this tells us that large producers tend to have low measured labor shares. To match this, our model requires $\tau_l = 0.003$, so indeed producers with high productivities are also producers with relatively high labor distortions.[16]

These labor market distortions significantly reduce aggregate productivity relative to the benchmark economy — the level of productivity turns out to be only about two-thirds that

[16]For our benchmark model the average labor share is also greater than the aggregate labor share, but the spread is 1.16, somewhat lower than the 1.35 in the data.

of the benchmark. In this sense, total misallocation is much larger in this economy. But this is because there are now *two* sources of misallocation — labor market distortions and markup distortions. The amount of misallocation due to markup distortions alone is roughly the same as in the benchmark economy. To see this, notice that the level of productivity associated with a planner who faces the same labor distortions but can otherwise reallocate across producers is 6.4% higher than the equilibrium level of productivity, very close to the corresponding 6.7% gap in the benchmark economy.

Given that there are similar amounts of misallocation due to markups, it is not then surprising that the gains from trade turn out to be similar as well. The aggregate gains from trade are 12.1% versus the benchmark 12% while the pro-competitive gains are about 1.8% in both cases. Importantly, we find that allowing for labor market distortions *does not* change the estimate of θ implied by equation (19) above. We continue to find that a regression of inverse markups on market shares gives a coefficient of about -0.68, which is consistent with $\theta = 1.28$, just as in the benchmark. Essentially, this is because with $\tau_l = 0.003$ there is in fact only a weakly positive relationship between a producer's productivity and their labor market distortion and as a result the relationship between inverse markups and market shares is mostly driven by the product market distortions, as in the benchmark model.

Heterogeneous tariffs. In our benchmark model, the only barriers to trade are the physical trade costs τ and f_x and these are the same for every producer in every sector. We consider a version of our model where in addition to these trade costs there is a *sector-specific* distortionary tariff that is levied on the value of imported goods. For simplicity we assume the tariff revenues are rebated lump-sum to the representative consumer. We assume the tariff rates are drawn from a Beta distribution on $[0, 1]$ with parameters estimated by maximum likelihood using the Taiwanese micro data. These estimates imply a mean tariff rate of 0.062 with cross-sectional standard deviation of 0.039. With a mean tariff of 0.062, the trade cost required to match the aggregate import share is correspondingly lower, 1.067 down from the benchmark 1.128.

Perhaps surprisingly, we find the total gains from trade are somewhat larger than in the benchmark, 14% as opposed to 12%, with the pro-competitive gains being similarly larger, 3.9% as opposed to 1.8% in the benchmark. One might expect that, for a given distribution of tariffs, a symmetric reduction in trade costs would make the cross-sectoral misallocation due to tariffs worse and thereby reduce the gains from trade (relative to an economy without tariffs). In this experiment, we find the opposite. This is due to a kind of 'second best' effect — i.e., in the presence of two distortions, increasing one distortion does not necessarily reduce welfare. In particular, the additional cross-sectoral misallocation due to tariffs is more than offset by strong reductions in within-sector market share dispersion.

Bertrand competition. In our benchmark model, firms engage in Cournot competition. If we assume instead that firms engage in Bertrand competition, then the model changes in only one respect. The demand elasticity facing producer i in sector s is no longer a harmonic weighted average of θ and γ, as in equation (13), but is now an arithmetic weighted average, $\varepsilon_i(s) = \omega_i(s)\theta + (1 - \omega_i(s))\,\gamma$. With this specification the results are similar to the benchmark. The total gains from trade are 13.1%, up slightly from the benchmark 12%, and the pro-competitive gains are 2%, likewise up slightly from the benchmark 1.8%. As shown in the last two columns of Table 3, the Bertrand model implies somewhat lower markup dispersion than the Cournot model. But it also implies a larger *change* in markup dispersion when opening to trade and hence a larger reduction in misallocation. Opening from autarky to Taiwan's import share implies misallocation falls by one-half, up from the benchmark one-fifth fall. Perhaps not surprisingly, the competitive pressure on dominant firms following a trade liberalization is greater with Bertrand competition than with Cournot. Consequently, the Bertrand model implies, if anything, larger pro-competitive effects than the benchmark.

Sensitivity to γ. In our benchmark model, we set the within-sector elasticity of substitution to $\gamma = 10$, following Atkeson and Burstein (2008). We assess the sensitivity of our results to higher or lower values of γ by recalibrating out model for $\gamma = 5$ and $\gamma = 20$. With $\gamma = 5$ we find that the model cannot produce a trade elasticity of 4, even setting $\tau(\rho) = 1$ (perfect correlation) gives a low trade elasticity of 2.38. With this low trade elasticity, the model implies much higher total gains from trade, 19% of aggregate productivity of which 2.3% are pro-competitive gains. With $\gamma = 20$ we can match a trade elasticity of 4 with less correlation, $\tau(\rho) = 0.85$ than our benchmark. With less correlation, the pro-competitive gains are smaller, about 0.8% as opposed to 1.8% in the benchmark, but still sizeable. As discussed in the Appendix, compared to the benchmark model the $\gamma = 20$ model also implies a weaker relationship between sector size and import shares than we see in the data.

No fixed costs. To assess the role of the fixed costs f_d and f_x we compute results for a version of our model with $f_d = f_x = 0$. In this specification, all firms operate in both their domestic and export markets. Hence the equilibrium number of producers in a sector is simply pinned down by the Geometric distribution for $n(s)$. This version of the model yields almost identical results to the benchmark. Shutting down these extensive margins makes little difference because the typical producer near the margin of operating or not is small and has negligible impact on the aggregate outcomes.

Gaussian copula. Our benchmark model uses the Gumbel copula (41) to model cross-country correlation in sectoral productivity draws. To examine the sensitivity of our results to this functional form, we resolve our model using a Gaussian copula, namely

$$\mathcal{C}(u,u^*) = \Phi_{2,\rho}(\Phi^{-1}(u), \Phi^{-1}(u^*)) \tag{42}$$

where $\Phi(x)$ denotes the CDF of the standard Normal distribution and $\Phi_{2,\rho}(x,x^*)$ denotes the standard bivariate Normal distribution with linear correlation coefficient $\rho \in (-1,1)$. To compare results to the Gumbel case, we map the linear correlation coefficient into our preferred Kendall correlation coefficient, which for the Gaussian copula is $\tau(\rho) = 2\arcsin(\rho)/\pi$. To match a trade elasticity of 4 requires $\tau(\rho) = 0.97$, up slightly from the benchmark 0.93 value. This version of the model also yields very similar results to the benchmark. Conditional on choosing the amount of correlation to match the trade elasticity, the total gains from trade are 11.5% with pro-competitive gains of 1.5%, both quite close to their benchmark values. In short, our results are not sensitive to the assumed functional form of the copula.

7 Extensions

7.1 Asymmetric countries

Our benchmark model makes the stark simplifying assumption of trade between two symmetric countries. We now relax this and consider the gains from trade between countries that differ in size and/or productivity. Specifically, we normalize the Home country labor force to $L = 1$ and vary the Foreign labor force L^*. Home producers continue to have production function $y_i(s) = a_i(s)l_i(s)$, as in (3) above, and Foreign producers now have the production function $y_i^*(s) = \bar{A}^* a_i^*(s) l_i^*(s)$ with productivity scale parameter $\bar{A}^*$. We again recalibrate key parameters of the model so that for the Home country we reproduce the degree of openness of the Taiwan benchmark — in particular, we choose the proportional trade cost τ, export fixed cost f_x, and correlation parameter $\tau(\rho)$ so that the Home country continues to have an aggregate import share of 0.38, fraction of exporters 0.25 and trade elasticity 4.

Larger trading partner. The top panel of Table 9 shows the gains from trade when the Foreign country has labor force $L^* = 2$ and $L^* = 10$ times as large as the Home country. For the Home country, the total gains from trade are slightly smaller than under symmetry. And when the Foreign country is larger, its total gains from trade are smaller than the Home country gains. For example, when the Foreign country is 10 times as large as the Home country, the Home gains are 11.1% (down from 12% in the symmetric benchmark) whereas the Foreign gains are down to 1.8%. The Home country has much more to gain from integration with a large trading partner than the Foreign country has to gain from integration with a small trading partner. The pro-competitive gains are also slightly lower for both countries. When $L^* = 10$, the Home pro-competitive gains are 1.2% (down from 1.8%

in the symmetric benchmark) whereas the Foreign pro-competitive gains are down to 1.1%. Interestingly, the pro-competitive gains account for a high share of the Foreign country's total gains, 1.1% out of 1.8%. In this calibration, the Foreign country is considerably less open than the Home country, with an aggregate import share of 0.05 (as opposed to 0.38) and a fraction of exporters of 0.06 (as opposed to 0.25). Despite the lower openness, we see that Foreign consumers still gain considerably from exposing their producers to greater competition (Home consumers gain even more), and that failing to account for pro-competitive effects can seriously understate the gains from integration, even for a large country.

More productive trading partner. The bottom panel of Table 9 shows the gains from trade when the Foreign country has productivity scale $\bar{A}^* = 2$ and $\bar{A}^* = 10$ times that of the Home country but has the same size, $L^* = 1$. Not surprisingly, for the Home country the total gains from trade are considerably larger than under symmetry. For example, when $\bar{A}^* = 10$, the Home gains are 30.1% (up from 12% in the symmetric benchmark). But these very large gains are almost entirely due to increases in the first-best level of productivity. The pro-competitive gains are 1.2%, and hence relative to the symmetric benchmark are both smaller in absolute terms and smaller as a share of the total gains. The more productive Foreign country has smaller total gains (and so benefits less from trade than the less productive Home country) and smaller pro-competitive gains.

The correlation in cross-country productivity required to reproduce a Home trade elasticity of 4 is $\tau(\rho) = 0.55$, considerably lower than the benchmark $\tau(\rho) = 0.93$. With large productivity differences between countries, import shares are more responsive to changes in trade costs than under symmetry. But because there is less correlation, there is also less head-to-head competition and because of this the pro-competitive gains are smaller.

7.2 Free entry

In our benchmark model there is an exogenous number of firms in each sector, a subset of which choose to pay the fixed cost f_d and operate. Some of the firms that do operate make substantial economic profits and thus there is an incentive for other firms to try to enter. We now relax the no-entry assumption and assume instead that there is *free entry* subject to a sunk cost. In equilibrium, the expected profits simply compensate for this initial sunk cost.

To keep the analysis tractable, we assume that entry is not directed at a particular sector. After paying its sunk cost, a firm learns the productivity with which it operates, as in Melitz (2003), as well as the sector to which it is randomly assigned.[17] We also assume that there are no fixed costs of operating or exporting in any given period. Instead, we assume that a

[17] An unappealing implication of allowing *directed entry* is that the resulting model would predict low dispersion in sectoral markups, in stark contrast to the very high dispersion in sectoral markups in the data.

firm's productivity is drawn from a discrete distribution which includes a mass point at zero, thus allowing the model to generate dispersion in the number of firms that operate.

Computational issues. Given the structure of our model, the *expected* profits of a potential entrant (which, due to free entry, equals the sunk cost) are not equal to the *average* profits across those firms that operate. One reason for this difference is that a potential entrant recognizes the effect its entry will have on its own profits and those of the incumbents. An additional reason is that the measure of producers of different productivities in a given sector is correlated with the profits a particular firm makes in that sector. Computing the expected profits of a potential entrant is thus a computationally challenging task: we need to integrate the distribution (across sectors) of the measures of firms (over their productivities) — a finite, but high-dimensional object. In addition, a potential entrant must re-solve for the distribution of markups that would arise if it enters. Given that the number of firms that enter each sector is small, the law of large number fails, and the algorithm to compute an equilibrium is involved. For this reason, we make a number of additional simplifying assumptions relative to our benchmark model without entry. In particular, we use a coarse productivity distribution and set the operating and exporting fixed costs to $f_d = f_x = 0$.

Setup. The productivity of a firm in sector $s \in [0, 1]$ is now given by a world component, common to both countries, $z(s)$, and a firm-specific component. In addition, we assume a *gap* $u(s)$ between the productivity with which firms produce for their domestic market and that with which they produce for their export market. Specifically, let $u(s)$ denote the productivity gap for Home producers in sector s and let $u^*(s)$ denote the productivity gap for Foreign producers in sector s. There is an unlimited number of potential entrants. To enter, a firm pays a sunk cost f_e that allows it to draw (i) a sector s in which to operate, and (ii) idiosyncratic productivity $x_i(s) \in \{0, 1, \bar{x}\}$. To summarize, a Home firm in sector s with idiosyncratic productivity $x_i(s)$ produces for its domestic market with overall productivity $a_i^{\mathrm{H}}(s) = z(s)u(s)x_i(s)$ and produces for its export market with overall productivity $a_i^{*\mathrm{H}}(s) = z(s)x_i(s)/\tau$ where τ is the gross trade cost. Similarly, a Foreign firm in sector s with idiosyncratic productivity $x_i^*(s)$ produces for its domestic market with overall labor productivity $a_i^{*\mathrm{F}}(s) = z(s)u^*(s)x_i^*(s)$ and produces for its export market with overall productivity $a_i^{\mathrm{F}}(s) = z(s)x_i^*(s)/\tau$.

Cross-country correlation and head-to-head competition. In this version of the model, the amount of head-to-head competition can now be varied flexibly by changing the amount of dispersion in $u(s)$ across sectors. Greater dispersion in $u(s)$ reduces the amount of head-to-head competition between Home and Foreign producers and thereby lowers the aggregate trade elasticity.

Parameterization. The Taiwanese data feature a high degree of across-sector dispersion in markups, in the number of producers, and in market concentration. We match this across-sector dispersion by assuming that the probability that a firm draws idiosyncratic productivity $x_i(s) \in \{0, 1, \bar{x}\}$ varies with s (but is the same across countries for a given sector). In particular, we assume a non-parametric distribution $\text{Prob}[x_i(s) \,|\, s]$ across sectors and calibrate this distribution to match the same set of moments we targeted for our benchmark model (we have found that allowing for 9 types of sectors produces a good fit; in our Appendix we also discuss results for a simpler model with a single sector type).

We assume that the gaps $u(s)$ are drawn from a lognormal distribution with variance σ_u^2 and that the worldwide sectoral productivities $z(s)$ are drawn from a Pareto distribution with shape parameter ξ_z.

Taiwan calibration revisited. We fix $\gamma = 10$ and $\theta = 1.28$, as in our benchmark model. We calibrate the new parameters $f_e, \bar{x}, \sigma_u, \xi_z$, the distribution $\text{Prob}[x_i(s) \,|\, s]$ across sectors, and the trade cost τ targeting the same moments as in our benchmark model. The full set of results for this calibration are reported in our Appendix.

Gains from trade with free entry. Panel A of Table 10 shows the gains from trade in this economy. With free entry, 168 firms pay the sunk cost and enter any individual sector. The economy is about 2% away from the first-best level of aggregate productivity. Thus, although we target the same concentration moments and have the same elasticities θ and γ as in the benchmark model, with free entry there is less misallocation.

Aggregate productivity is 7.2% above its autarky level and opening to trade reduces misallocation by just over one-third, from 3.2% to 2%. This reduction in misallocation implies pro-competitive gains of 1.2%, somewhat lower than the benchmark pro-competitive gains of 1.8%. Note that there are 187 firms attempting to enter under autarky, more than in the open economy. For a given number of firms, expected profits are higher under autarky and so more firms enter until the free-entry condition is satisfied. If we hold the number of firms fixed at the autarky level of 187 but otherwise open the economy to trade, aggregate productivity rises by 8.2%, larger than the 7.2% with free entry, and the pro-competitive gains are correspondingly larger at 1.4% as opposed to 1.2%.

To summarize, even with free-entry there is a quantitatively significant reduction in misallocation. Importantly, the somewhat weaker pro-competitive effects reflect the alternative calibration of the model which implies less initial misallocation, not the free entry itself. In particular, the model predicts much less dispersion in sectoral markups — e.g., the ratio of the 90th percentile to the median is 1.14 (compared to 1.31 in the data and 1.23 in the benchmark), and the ratio of the 95th percentile to the median is 1.16 (1.56 in the data and 1.40 in the benchmark). We address this discrepancy between the model and the data next.

Collusion. Given this failure to match the dispersion of sectoral markups in the data, we now consider a slight variation on the free-entry model designed to bridge the gap between the model and the data along this dimension. We suppose that with probability ψ all the high-productivity firms (those with $x_i(s) = \bar{x} > 1$) within a given sector are able to *collude*.[18] These colluding firms choose a single price to maximize their *group profits*. Since their collective market share is larger than their individual market shares, the price set by colluding firms is higher than the price they would charge in isolation and hence their collective markup is also correspondingly larger. Since this version of the model features more dispersion in markups, it also features more misallocation.

Panel B of Table 10 shows results for this model with $\psi = 0.25$. Even with free entry this version of the model features productivity losses of 4.3% relative to the first-best. The reason these productivity losses are greater is that now the dispersion in sectoral markups is greater. For example, the ratio of the 90th percentile to the median is 1.22 (compared to 1.14 absent collusion) and the ratio of the 95th percentile to the median is 1.30 (compared to 1.16 absent collusion). Thus this version of the model produces sectoral dispersion in markups more in line with our benchmark model and hence closer to the data.

Consequently, the model now predicts larger total gains from trade of 11.2%, of which 3.9% are pro-competitive gains — i.e., the model with free-entry and collusion implies larger pro-competitive gains than our benchmark model. With wide-spread collusion amongst domestic producers, opening to foreign competition provides an import source of market discipline. Notice also that the number of producers change very little (from 162 in autarky to 160 in the open economy) despite the reduction in firm markups (the aggregate markup falls from 1.34 to 1.27). The reason the number of firms does not change much is an externality akin to that in Blanchard and Kiyotaki (1987). Although an individual firm loses profits if its own markup falls, it benefits when other firms reduce markups due to the increase in aggregate output and the reduction in the aggregate price level. Overall, these two effects on expected profits roughly cancel each other out so that there is little effect on the gains from trade.

In short, with free entry and collusion the model implies strong pro-competitive effects. In our Appendix we report results for a wide range of collusion probabilities ψ and show that the same basic pattern holds. For example, if the collusion probability is $\psi = 0.15$ instead of $\psi = 0.25$ then the total gains from trade are 12.1% of which 3.8% are pro-competitive gains.

[18] Alternatively, this can be thought of as the result of mergers or acquisitions.

The results from the model with collusion reinforce our main message: the pro-competitive gains from trade are larger when product market distortions are large to begin with.

8 Conclusions

We study the pro-competitive gains from international trade in a quantitative model with endogenously variable markups. We find that trade can significantly reduce markup distortions if two conditions are satisfied: (i) there must be large inefficiencies associated with markups, i.e., extensive misallocation, and (ii) trade must in fact expose producers to greater competitive pressure. The second condition is satisfied if there is a high degree of cross-country correlation in the productivity with which producers within a given sector operate.

We calibrate our model using Taiwanese producer-level data and find that these two conditions are satisfied. The Taiwanese data is characterized by a large amount of dispersion and concentration in producer market shares and a strong cross-sectional relationship between producer market shares and markups, which implies extensive misallocation. Moreover to match standard estimates of the trade elasticity, and at the same time match key facts on import share dispersion, intra-industry trade and the cross-sectional relationship between import penetration and domestic concentration, the model requires a high degree of cross-country correlation in productivity. Consequently, the model implies that opening to trade does in fact expose producers to considerably greater competitive pressure.

We find that opening to trade reduces misallocation by about one-fifth in our benchmark model with Cournot competition, reduces misallocation by about one-third in our alternate model that matches the correlation between domestic concentration and import penetration, and reduces misallocation by about one-half in our model with Bertrand competition. In this sense, we find that, indeed, trade can significantly reduce product market distortions.

We conclude by noting that, from a policy viewpoint, our model suggests that obtaining large welfare gains from an improved allocation of resources may not require the detailed, perhaps impractical, scheme of producer-specific subsidies and taxes that reduce the distortions associated with variable markups. Instead, simply opening an economy to trade may provide an excellent practical alternative that substantially improves productivity and welfare. Conversely, our model also predicts that countries which open up to trade after having already implemented policies aimed at reducing markup distortions may benefit less from trade than countries with large product market distortions. The former countries would mostly receive the standard gains from trade, while the latter would also benefit from the reduction in markup distortions.

References

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer, "Structural Identification of Production Functions," 2006. UCLA working paper.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare, "New Trade Models, Same Old Gains?," *American Economic Review*, February 2012, *102* (1), 94–130.

_ , _ , Dave Donaldson, and Andrés Rodríguez-Clare, "The Elusive Pro-Competitive Effects of Trade," March 2012. MIT working paper.

Atkeson, Andrew and Ariel Burstein, "Pricing-to-Market, Trade Costs, and International Relative Prices," *American Economic Review*, 2008, *98* (5), 1998–2031.

Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum, "Plants and Productivity in International Trade," *American Economic Review*, September 2003, *93* (4), 1268–1290.

_ , Stephen J. Redding, and Peter Schott, "Multiple-Product Firms and Product Switching," *American Economic Review*, March 2010, *100* (1), 70–97.

Blanchard, Olivier Jean and Nobuhiro Kiyotaki, "Monopolistic Competition and the Effects of Aggregate Demand," *American Economic Review*, 1987, *77* (4), 647–666.

Chaney, Thomas, "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, September 2008, *98* (4), 1707–1721.

De Blas, Beatriz and Katheryn Russ, "Teams of Rivals: Endogenous Markups in a Ricardian World," 2010. NBER Working Paper 16587.

De Loecker, Jan and Frederic Warzynski, "Markups and Firm-Level Export Status," *American Economic Review*, October 2012, *102* (6), 2437–2471.

_ , Pinelopi K. Goldberg, Amit K. Khandelwal, and Nina Pavcnik, "Prices, Markups and Trade Reform," 2012. Princeton University working paper.

Devereux, Michael B. and Khang Min Lee, "Dynamic Gains from International Trade with Imperfect Competition and Market Power," *Review of Development Economics*, 2001, *5* (2), 239–255.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography, and Trade," *Econometrica*, September 2002, *70* (5), 1741–1779.

Epifani, Paolo and Gino Gancia, "Trade, Markup Heterogeneity, and Misallocations," *Journal of International Economics*, January 2011, *83* (1), 1–13.

Feenstra, Robert C., "A Homothetic Utility Function for Monopolistic Competition Models, Without Constant Price Elasticity," *Economics Letters*, January 2003, *78* (1), 79–86.

_ **and David E. Weinstein**, "Globalization, Markups, and the U.S. Price Level," 2010. NBER Working Paper 15749.

Grubel, Herbert G. and Peter J. Lloyd, "The Empirical Measurement of Intra-Industry Trade," *Economic Record*, 1971, *47* (4), 494–517.

Harrison, Ann E., "Productivity, Imperfect Competition and Trade Reform: Theory and Evidence," *Journal of International Economics*, February 1994, *36* (1), 53–73.

Holmes, Thomas J., Wen-Tai Hsu, and Sanghoon Lee, "Monopoly Distortions and the Welfare Gains from Trade," November 2011. FRB Minneapolis working paper.

Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, November 2009, *124* (4), 1403–1448.

Krishna, Pravin and Devashish Mitra, "Trade Liberalization, Market Discipline, and Productivity Growth: New Evidence from India," *Journal of Development Economics*, August 1998, *56* (2), 447–462.

Krugman, Paul R., "Increasing Returns, Monopolistic Competition, and International Trade," *Journal of International Economics*, November 1979, *9* (4), 469–479.

Levinsohn, James A., "Testing the Imports-as-Market-Discipline Hypothesis," *Journal of International Economics*, August 1993, *35* (1), 1–22.

Levinsohn, James and Amil Petrin, "Estimating Production Functions using Inputs to Control for Unobservables," *Review of Economic Studies*, April 2003, *70* (2), 317–342.

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

_ **and Giancarlo I.P. Ottaviano**, "Market Size, Trade, and Productivity," *Review of Economic Studies*, 2008, *75* (1), 295–316.

Olley, G. Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry.," *Econometrica*, November 1996, *64* (6), 1263–1297.

Peters, Michael, "Heterogeneous Mark-Ups, Growth and Endogenous Misallocation," September 2013. LSE working paper.

Restuccia, Diego and Richard Rogerson, "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments," *Review of Economic Dynamics*, October 2008, *11* (4), 707–720.

Tybout, James R., "Plant- and Firm-Level Evidence on "New" Trade Theories," in E. Kwan Choi and James Harrigan, eds., *Handbook of International Trade*, Basil Blackwell. Oxford. 2003.

Zhelobodko, Evgeny, Sergey Kokovin, Mathieu Parenti, and Jacques-François Thisse, "Monopolistic Competition: Beyond the Constant Elasticity of Substitution," *Econometrica*, November 2012, *80* (6), 2765–2784.

Table 1: Parameterization

Panel A: Moments

	Data	Model		Data	Model
Within-sector concentration, domestic sales			*Size distribution sectors, domestic sales*		
mean inverse HH	7.25	4.43	fraction sales by top 0.01 sectors	0.26	0.24
median inverse HH	3.92	3.82	fraction sales by top 0.05 sectors	0.52	0.36
mean top share	0.45	0.45	fraction wages (same) top 0.01 sectors	0.11	0.25
median top share	0.40	0.41	fraction wages (same) top 0.05 sectors	0.32	0.37
Distribution of sectoral shares, domestic sales			*Size distribution producers, domestic sales*		
mean share	0.04	0.05	fraction sales by top 0.01 producers	0.41	0.37
median share	0.005	0.006	fraction sales by top 0.05 producers	0.65	0.64
p75 share	0.02	0.03	fraction wages (same) top 0.01 producers	0.24	0.35
p95 share	0.19	0.27	fraction wages (same) top 0.05 producers	0.47	0.60
p99 share	0.59	0.59			
std dev share	0.11	0.11			
Across-sector concentration			*Strength of markup, market share relationship*		
p10 inverse HH	1.17	1.74	coefficient, inv. markup on market share	-0.66	-0.68
p50 inverse HH	3.73	3.82		to -0.69	
p90 inverse HH	13.82	7.94	*Import and export statistics*		
p10 top share	0.16	0.23	aggregate fraction exporters	0.25	0.25
p50 top share	0.41	0.41	aggregate import share	0.38	0.38
p90 top share	0.92	0.74	trade elasticity	4	4
p10 number producers	2	3	coefficient, share imports on share sales	0.81	0.77
p50 number producers	10	16	index import share dispersion	0.38	0.23
p90 number producers	52	47	index intraindustry trade	0.37	0.50

Panel B: Parameter Values

γ	10	within-sector elasticity of substitution
θ	1.28	across-sector elasticity of substitution
ξ_x	4.53	Pareto shape parameter, idiosyncratic productivity
ξ_z	0.56	Pareto shape parameter, sector productivity
ζ	0.043	Geometric parameter, number producers per sector
f_d	0.004	fixed cost of domestic operations
f_x	0.211	fixed cost of export operations
τ	1.128	gross trade cost
$\tau(\rho)$	0.93	Kendall correlation for Gumbel copula

Table 2: Production Function Estimates

		Panel A: Output Elasticity With Respect to ...						Panel B: Returns to Scale		
		Labor		Capital		Materials				
TW SIC 2	Sector	Median	IQR	Median	IQR	Median	IQR	Median	IQR	Obs.
11	Textile	0.27	[0.17,0.39]	0.04	[0.02,0.06]	0.69	[0.57,0.79]	1.00	[0.98,1.02]	5982
12	Apparel	0.26	[0.11,0.43]	0.03	[0.00,0.06]	0.68	[0.56,0.80]	0.97	[0.92,1.03]	3790
13	Leather	0.30	[0.21,0.39]	0.02	[0.01,0.03]	0.66	[0.59,0.74]	0.98	[0.96,1.01]	4585
14	Wood	0.30	[0.27,0.34]	0.01	[0.00,0.01]	0.69	[0.64,0.74]	1.00	[0.99,1.01]	4765
15	Paper	0.23	[0.13,0.34]	0.05	[0.03,0.07]	0.70	[0.61,0.78]	0.98	[0.96,1.00]	4919
16	Printing	0.35	[0.24,0.46]	0.05	[0.03,0.07]	0.62	[0.52,0.73]	1.03	[1.01,1.05]	7744
17	Petroleum	0.20	[0.07,0.37]	0.10	[0.03,0.17]	0.68	[0.53,0.83]	0.99	[0.95,1.04]	3337
18	Chemical Material	0.30	[0.21,0.41]	0.08	[0.04,0.13]	0.63	[0.55,0.70]	1.02	[0.98,1.06]	6860
19	Chemical Prod	−0.11	[−0.24,0.04]	0.18	[0.06,0.28]	0.86	[0.77,0.95]	0.94	[0.85,1.03]	706
20	Pharmaceutical	0.33	[0.22,0.43]	0.04	[0.01,0.06]	0.64	[0.53,0.75]	1.00	[0.98,1.03]	3424
21	Rubber	0.29	[0.20,0.39]	0.05	[0.03,0.07]	0.66	[0.58,0.74]	1.00	[0.97,1.03]	23813
22	Plastic	0.23	[0.09,0.37]	0.05	[0.04,0.07]	0.73	[0.59,0.87]	1.01	[0.98,1.04]	8041
23	Non-metallic Mineral	0.41	[0.30,0.53]	0.09	[0.04,0.14]	0.52	[0.43,0.61]	1.03	[0.97,1.08]	7693
24	Basic Metal	0.30	[0.20,0.41]	0.05	[0.03,0.06]	0.67	[0.57,0.77]	1.02	[0.99,1.04]	35622
25	Fabricated Metal	0.29	[0.19,0.41]	0.04	[0.03,0.05]	0.66	[0.56,0.76]	1.00	[0.98,1.02]	52159
26	Electronic Parts & Components	0.30	[0.17,0.42]	0.07	[0.03,0.11]	0.63	[0.53,0.73]	0.99	[0.96,1.03]	6772
27	Computer, Electronic, Optical	0.34	[0.23,0.44]	0.10	[0.06,0.13]	0.62	[0.51,0.72]	1.05	[1.02,1.07]	8723
28	Electrical Equipment	0.25	[0.10,0.42]	0.06	[0.03,0.09]	0.69	[0.55,0.83]	1.00	[0.97,1.04]	11316
29	Machinery and Equipment	0.28	[0.20,0.37]	0.08	[0.05,0.11]	0.67	[0.60,0.73]	1.03	[1.01,1.04]	12708
30	Motor Vehicle and Parts	0.38	[0.27,0.48]	0.03	[0.01,0.06]	0.57	[0.47,0.67]	0.97	[0.90,1.06]	3923
31	Trans. Equipment and Parts	0.31	[0.20,0.42]	0.04	[0.01,0.06]	0.65	[0.56,0.74]	1.00	[0.98,1.01]	10288

Table 3: Markup Estimates

Panel A: Markup Distribution

	DLW	Projected
Unconditional markup distribution		
p75/p50	1.24	1.01
p90/p50	1.74	1.04
p95/p50	2.46	1.08
p99/p50	4.84	1.33
std dev log	0.38	0.06
Across-sector markup distribution		
p75/p50	1.30	1.10
p90/p50	1.99	1.31
p95/p50	2.81	1.56
p99/p50	4.56	2.58
std dev log	0.41	0.20

Panel B: Inverse Markup Regressions

Regression of DLW inverse markups on market shares	
multi-product	−0.69
	[0.01]
single-product	−0.66
	[0.02]

Notes: Panel A summarizes the distribution of markups and sectoral markups estimated from the Taiwanese data. DLW refers to markups obtained using the De Loecker and Warzynski (2012) procedure. Panel B reports the slope coefficient $\widehat{\beta_\omega}$ obtained from regressing the DLW inverse markups on observed market shares, with standard errors in brackets. The projected markups are the inverse of the fitted values from this regression, which we normalize by setting the intercept equal to its theoretical value $\frac{\gamma-1}{\gamma}$. See the text for further discussion.

Table 4: Markups in Data and Model

Panel A: Markup Moments

	Data	Benchmark		Bertrand	
		Taiwan	Autarky	Taiwan	Autarky
aggregate markup		1.31	1.35	1.21	1.23
Unconditional markup distribution					
mean	1.13	1.14	1.15	1.12	1.12
p50	1.11	1.12	1.12	1.11	1.11
p75	1.12	1.14	1.14	1.11	1.11
p90	1.15	1.21	1.23	1.13	1.13
p95	1.20	1.31	1.35	1.15	1.16
p99	1.48	1.67	1.76	1.33	1.37
std dev log	0.06	0.08	0.10	0.04	0.08
log p95/p50	0.08	0.16	0.19	0.04	0.04
Across-sector markup distribution					
mean	1.32	1.34	1.39	1.22	1.26
p50	1.21	1.29	1.32	1.18	1.18
p75	1.33	1.41	1.45	1.24	1.26
p90	1.59	1.59	1.79	1.39	1.48
p95	1.89	1.80	2.25	1.53	2.14
p99	3.12	2.13	4.57	2.07	4.57
std dev log	0.20	0.13	0.28	0.12	0.29
log p95/p50	0.45	0.33	0.54	0.26	0.60

Panel B: Aggregate Implications

import share	0.38	0.38	0	0.38	0.38
fraction exporters	0.25	0.25	0	0.25	0.25
TFP loss, %		6.7	8.5	1.9	3.9
gains from trade, %		12.0	–	13.1	–
pro-competitive gains, %		1.8	–	2.0	–

Notes: Data markup moments are for the projected markups reported in Table 3. The benchmark model features Cournot competition. TFP losses are the percentage gap between the level of aggregate productivity and the first-best level of aggregate productivity associated with the planning allocation (subject to the same trade costs). The gains from trade are the percentage change in aggregate productivity relative to autarky. The pro-competitive gains from trade are the percentage change in aggregate productivity less the percentage change in first-best productivity.

Table 5: Gains from Trade

Panel A: Benchmark Model

Change in import share	0 to 10%	10 to 20%	20 to 30%	30% to Taiwan	0 to Taiwan
change TFP, %	3.1	2.8	3.3	2.8	12.0
change first-best TFP, %	1.9	2.5	3.1	2.7	10.2
pro-competitive gains, %	1.2	0.3	0.2	0.1	1.8
misallocation relative to autarky	0.86	0.83	0.80	0.79	0.79
change aggregate markup, %	−1.6	−0.6	−0.4	−0.2	−2.8
domestic	−1.4	−0.5	−0.5	−0.3	−2.7
import	15.8	0.2	0.3	0.2	16.5
change p99/p50, %	−0.9	−1.1	−0.7	−0.1	−2.8
domestic	−1.6	−1.6	−1.4	−1.2	−5.8
import	23.4	0.3	−0.5	0.0	23.2
trade elasticity (ex post)	4.2	4.1	4.0	4.0	4.0
ACR gains, %	2.5	2.9	3.3	2.8	11.8

Panel B: Alternative Model with Correlated $x_i(s), x_i^*(s)$

Change in import share	0 to 10%	10 to 20%	20 to 30%	30% to Taiwan	0 to Taiwan
change TFP, %	3.0	2.6	3.2	2.8	11.7
change first-best TFP, %	1.6	2.1	2.8	2.6	9.0
pro-competitive gains, %	1.5	0.5	0.4	0.2	2.6
misallocation relative to autarky	0.82	0.76	0.71	0.69	0.69
change aggregate markup, %	−2.0	−0.8	−0.5	−0.3	−3.6
domestic	−1.6	−0.7	−0.6	−0.4	−3.3
import	14.5	0.0	0.4	0.3	15.2
change p95/p50, %	−0.8	−1.1	−0.8	−0.2	−2.9
domestic	−1.3	−1.5	−1.5	−1.0	−5.3
import	21.7	0.1	−0.1	0.2	21.9
trade elasticity (ex post)	4.2	4.1	4.1	4.0	4.0
ACR gains, %	2.5	2.8	3.3	2.8	11.8

Notes: Panel A shows the gains from trade for our benchmark model. Panel B shows the gains from trade for our alternative model with correlation in idiosyncratic draws $x_i(s), x_i^*(s)$ chosen to match the cross-sectional relationship between import penetration and domestic producer concentration, as discussed in the main text. For our benchmark model $x_i(s), x_i^*(s)$ are independent and there is cross-country correlation in productivity only through correlation in sectoral productivity $z(s), z^*(s)$.

Table 6: Importance of Head-to-Head Competition

Panel A: Sensitivity to Cross-Country Correlation, $\tau(\rho)$

	$\tau(\rho)$	Trade elasticity	Import share dispersion	Intraindustry trade	Share imports wrt share sales	Pro-competitive gains, %	Total gains, %
	1.00	4.41	0.15	0.64	0.93	2.0	11.1
	0.90	3.77	0.28	0.44	0.69	1.7	12.6
	0.80	2.91	0.46	0.28	0.40	1.4	16.5
	0.70	2.22	0.60	0.18	0.24	1.2	21.3
	0.60	1.77	0.69	0.12	0.15	0.9	26.7
	0.50	1.44	0.76	0.09	0.08	0.7	32.9
	0.40	1.19	0.81	0.07	0.03	0.5	39.8
	0.30	1.01	0.85	0.05	0.00	0.2	47.8
	0.20	0.86	0.88	0.04	−0.03	−0.1	57.4
	0.10	0.74	0.90	0.03	−0.05	−0.5	69.3
	0.00	0.66	0.92	0.02	−0.04	−0.9	86.1
benchmark	0.93	4.00	0.23	0.50	0.77	1.8	12.0
data		4.00	0.38	0.37	0.81		

Panel B: Markup Dispersion and Cross-Country Correlation

	$\tau(\rho) = 0.9$			$\tau(\rho) = 0.1$		
	Autarky	Taiwan	Change, %	Autarky	Taiwan	Change, %
All markups						
aggregate markup	1.35	1.31	−2.7	1.39	1.39	0.3
markup p50	1.12	1.12	0.0	1.12	1.12	−0.2
markup p95	1.35	1.31	−2.6	1.35	1.31	−2.6
markup p99	1.76	1.68	−4.7	1.76	1.69	−3.8
p95/p50	1.21	1.18	−2.6	1.21	1.18	−2.5
p99/p50	1.57	1.50	−4.7	1.57	1.51	−3.7
Domestic markups						
aggregate markup	1.35	1.32	−2.6	1.39	1.39	0.1
markup p50	1.12	1.11	−0.3	1.12	1.12	−0.1
markup p95	1.35	1.27	−5.8	1.35	1.32	−1.9
markup p99	1.76	1.61	−8.7	1.75	1.71	−2.3
p95/p50	1.21	1.14	−5.5	1.21	1.19	−1.8
p99/p50	1.57	1.45	−8.5	1.57	1.53	−2.2
Import markups						
aggregate markup	1.11	1.31	16.7	1.11	1.40	22.8
markup p50	1.11	1.14	2.1	1.11	1.11	0.3
markup p95	1.11	1.44	25.6	1.11	1.29	15.0
markup p99	1.11	1.81	49.0	1.11	1.65	39.5
p95/p50	1.00	1.26	23.5	1.00	1.16	14.7
p99/p50	1.00	1.60	46.8	1.00	1.48	39.2

Table 7: Gains from Trade with Elastic Factors

	Constant Markups	Variable Markups		
		Frisch elasticity of labor supply ($1/\eta$)		
		0	1	∞
change TFP, %	10.2	12.0	12.0	12.0
change markup, %	0	−2.8	−2.8	−2.8
change C, %	15.3	18.7	20.4	22.1
change K, %	15.3	22.1	23.8	25.5
change Y, %	15.3	19.3	21.0	22.7
change L, %	0	0	1.7	3.4
change welfare, % (including transition)	14.3	17.4	17.6	18.2
pro-competitive welfare gains, %	0	3.1	3.3	3.9

Notes: Representative consumer has preferences $\sum_{t=0}^{\infty} \beta^t U(C_t, L_t)$ over aggregate consumption C_t and labor L_t with $U(C, L) = \log C - L^{1+\eta}/1 + \eta$. Capital is accumulated according to $K_{t+1} = (1-\delta)K_t + I_t$. Individual producers have production function $y = ak^{\alpha}l^{1-\alpha}$. We set discount factor $\beta = 0.96$, depreciation rate $\delta = 0.1$, output elasticity of capital $\alpha = 1/3$ and elasticities of labor supply η as shown.

Table 8: Robustness Experiments

	Benchmark	Labor wedges	Tariffs	Bertrand	Low γ	High γ	No fix costs	Gauss. copula
trade elasticity	4.00	4.00	4.00	4.00	2.38	4.00	4.00	4.00
import share	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38
fraction exporters	0.25	0.25	0.25	0.25	0.25	0.25	1	0.25
TFP loss autarky, %	8.5	8.2	8.5	3.9	9.1	9.0	8.5	8.6
TFP loss Taiwan, %	6.7	6.4	6.7	1.9	6.8	8.2	6.7	7.1
gains from trade, %	12.0	12.1	14.0	13.1	19.0	11.1	11.5	11.5
pro-competitive gains, %	1.8	1.8	3.9	2.0	2.3	0.8	1.8	1.5
Key parameters								
$\tau(\rho)$	0.93	0.91	0.92	0.90	1.00	0.85	0.93	0.97
τ	1.128	1.128	1.067	1.132	1.214	1.137	1.137	1.129
f_x	0.211	0.243	0.199	0.109	0.710	0.018	0	0.195
Additional moments								
average/aggregate labor share	1.16	1.35	1.16	1.08	1.20	1.17	1.18	1.17
mean tariff			0.062					
std dev tariff			0.039					

Table 9: Gains from Trade with Asymmetric Countries

Panel A: Larger Trading Partner

	Benchmark		$L^* = 2L$		$L^* = 10L$	
	Home	Foreign	Home	Foreign	Home	Foreign
$\tau(\rho)$	0.93	0.93	0.93	0.93	0.94	0.94
trade elasticity	4.00	4.00	4.00	4.14	4.00	4.26
import share	0.38	0.38	0.38	0.21	0.38	0.05
fraction exporters	0.25	0.25	0.25	0.18	0.25	0.06
TFP loss autarky, %	8.5	8.5	8.5	8.5	8.5	8.5
TFP loss Taiwan, %	6.7	6.7	7.0	6.9	7.3	7.4
gains from trade, %	12.0	12.0	11.7	6.1	11.1	1.8
pro-competitive gains, %	1.8	1.8	1.5	1.7	1.2	1.1

Panel B: More Productive Trading Partner

	$\bar{A}^* = 2\bar{A}$		$\bar{A}^* = 10\bar{A}$	
	Home	Foreign	Home	Foreign
$\tau(\rho)$	0.83	0.83	0.55	0.55
trade elasticity	4.00	3.15	4.00	1.27
import share	0.38	0.21	0.38	0.07
fraction exporters	0.25	0.16	0.25	0.03
TFP loss autarky, %	8.5	8.5	8.7	8.7
TFP loss, %	7.3	7.0	7.5	8.4
gains from trade, %	15.0	7.9	30.1	6.2
pro-competitive gains, %	1.3	1.5	1.2	0.3

Table 10: Entry and Collusion

	Panel A: No Collusion			Panel B: 25% Collusion		
	No entry	Free entry	Autarky	No entry	Free entry	Autarky
number of firms trying to enter	187	168	187	162	160	162
TFP loss, %	1.8	2.0	3.2	4.2	4.3	8.1
total fixed costs	0.46	0.41	0.46	0.31	0.31	0.31
aggregate profits	0.46	0.47	0.51	0.38	0.38	0.42
aggregate markup	1.25	1.26	1.32	1.27	1.27	1.34
expected profits, entrants	0.22	0.25	0.24	0.19	0.19	0.19
gains from trade, %	8.2	7.2		10.9	11.2	
pro-competitive gains, %	1.4	1.2		3.8	3.9	
misallocation relative to autarky	0.56	0.63		0.52	0.53	
change markup, %	−5.5	−4.7		−6.1	−6.0	