Competition, Markups, and the Gains from International Trade: *Appendix*

Chris Edmond[*] Virgiliu Midrigan[†] Daniel Yi Xu[‡]

April 2012

Contents

[*]University of Melbourne, cedmond@unimelb.edu.au.

[†]Federal Reserve Bank of Minneapolis, New York University, and NBER, virgiliu.midrigan@nyu.edu.

[‡]Duke University and NBER, daniel.xu@duke.edu.

This appendix is organized as follows. In Table A1 we provide an example of the product classification in our 7-digit Taiwanese manufacturing data. In Appendix A we discuss how we use the Taiwanese data to identify the Pareto shape parameter μ_H governing the amount of concentration amongst top producers. In Appendix B we provide further details on our method for estimating the key cross-sector elasticity θ. In Appendix C we show that our markup estimates are similar to those obtained from modern IO techniques. In Appendix D we report two additional sets of results on the gains from trade: (i) we show for our benchmark model that the gains from tariff reductions are similar to the gains from trade cost reductions, and (ii) we show how the gains from trade depend on the correlation parameter $\tau(\rho)$. Finally, in Appendix E we provide further details on the robustness experiments mentioned in the main text.

A Identifying μ_H, the second Pareto tail

Double Pareto. In our model, the degree of market concentration plays a crucial role in pinning down the degree of markup dispersion. As discussed in the main text, we find that for the model to match the degree of concentration in the Taiwanese manufacturing data, we need to assume that productivities a are drawn from a very fat-tailed distribution. In particular, we assume that the (inverse) cumulative distribution $a = F^{-1}(u)$ is given by:

$$F^{-1}(u) = \begin{cases} (1-u)^{-\frac{1}{\mu_L}} & \text{if} \quad u < 1 - p_H \\ (1-u)^{-\frac{1}{\mu_H}} & \text{if} \quad u \geq 1 - p_H \end{cases} . \tag{1}$$

where, without loss of generality, $\mu_H \leq \mu_L$. We refer to this as the *double Pareto*. A fraction $1 - p_H$ of producers draw from a Pareto with relatively thin tails (and a low mean) governed by μ_L, while the remaining fraction p_H draw from a Pareto with fatter tails (and a high mean) governed by μ_H.[1]

Concentration statistics. Table A2 reports the key concentration statistics both in our Taiwanese data and for several alternate settings of the shape parameter μ_H for the second Pareto tail.[2] The last two columns on the right also report results for the case where $\mu_H = \mu_L = \mu$, i.e., where there is a single Pareto distribution. As discussed in the main text, the single Pareto cannot simultaneously match both properties of the conditional within-sector concentration statistics and the unconditional size distribution of producers. The single Pareto with $\mu = 3.4$ targets all of our statistics *except* the size distribution, while the alternative single Pareto with $\mu = 1.95$ targets *only* the size distribution. The former does very well at matching the within-sector concentration statistics, except at the very top.

[1] As noted in the main text, we have also redone our analysis with a distribution $F(a)$ that is a binomial mixture of Paretos and obtained essentially identical results. We prefer the specification above since it can match the data with one fewer parameter than the binomial mixture and also because an inverse is available in closed form, which is convenient when working with the copula specification that we use to control the cross-country productivity correlation.

[2] We have re-calibrated all other parameters when computing these statistics.

The remaining columns show double Pareto specifications with $\mu_H < \mu_L$. These specifications allow us to reduce μ_H so that top producers draw from a fatter tailed distribution. Consider the case where $\mu_H = 2.5$ and suppose $\mu_L = 3.83$, as in our benchmark. Then the average productivity draw from the high Pareto is $\mu_H/(\mu_H - 1) = 1.67$, some 24% higher than the average productivity drawn from the low Pareto, $\mu_L/(\mu_L - 1) = 1.35$. To match the concentration data we need the high Pareto to generate *much* larger productivity draws. If we reduce the shape parameter further, to $\mu_H = 1.5$, the average draw from the high Pareto increases substantially, to $\mu_H/(\mu_H - 1) = 3$, well more than twice the average draw from the low Pareto. As shown in Table A2, with $\mu_H = 1.5$ the model does a good job of matching the top concentration statistics as well as the unconditional size distribution. In particular, the specification with $\mu_H = 1.5$ has the lowest root mean squared (log) deviation from the data.

While the data very clearly require us to choose a μ_H considerably below μ_L, they are, however, somewhat less clear as to *how low* we should choose μ_H. Depending on which statistics in Table A2 are given most weight, a case could be made that we should a value for μ_H below 1.5. Lower values of μ_H would imply even more concentration and would lead the model to predict even larger gains from trade. To be conservative in our parameterization, we choose $\mu_H = 1.5$.

B Estimating θ, the across-sector elasticity of substitution

As discussed in the main text, with Cournot competition the model implies a linear relationship between sectoral shares and labor (and labor + capital) shares. We use this cross-sectional relationship to estimate the across-sector elasticity of substitution θ.

B.1 Single-product firms

Specifically, for single-product firms the relationship between sectoral shares and labor shares is given by

$$\frac{wl_i}{p_i y_i} = (1-\alpha)\left(1 - \frac{1}{\gamma}\right) - (1-\alpha)\left(\frac{1}{\theta} - \frac{1}{\gamma}\right)\omega_i + \frac{wF_d}{p_i y_i}, \tag{2}$$

where wl_i is the wage bill for producer i, $p_i y_i$ its value added, ω_i its sectoral share, and wF_d is the fixed cost, assumed common across all producers.

Let b_0 and b_1 denote, respectively, the intercept and coefficient on sectoral share in this regression

$$b_0 = (1-\alpha)\left(1 - \frac{1}{\gamma}\right),$$

and

$$b_1 = -(1-\alpha)\left(\frac{1}{\theta} - \frac{1}{\gamma}\right).$$

Thus the *ratio* of the slope to intercept is *independent* of α and is given by:

$$\frac{b_1}{b_0} = -\frac{\frac{1}{\theta} - \frac{1}{\gamma}}{1 - \frac{1}{\gamma}}.$$

Hence for a given value of γ, we have that:

$$\theta = \left(\frac{1}{\gamma} - \frac{b_1}{b_0}\left(1 - \frac{1}{\gamma}\right)\right)^{-1}.$$

B.2 Multi-product firms

Many of our product-level observations are for firms that produce multiple products. For these firms (2) holds for each of $k = 1, ..., K_i$ products, namely:

$$\frac{wl_{ik}}{p_{ik}y_{ik}} = (1-\alpha)\left(1 - \frac{1}{\gamma}\right) - (1-\alpha)\left(\frac{1}{\theta} - \frac{1}{\gamma}\right)\omega_{ik} + \frac{wF_d}{p_{ik}y_{ik}}. \tag{3}$$

Multiplying both sides by each product's value added, $p_{ik}y_{ik}$, gives

$$wl_{ik} = \left((1-\alpha)\left(1 - \frac{1}{\gamma}\right) - (1-\alpha)\left(\frac{1}{\theta} - \frac{1}{\gamma}\right)\omega_{ik}\right)p_{ik}y_{ik} + wF_d,$$

and then summing over the K_i products and dividing by the producer's total value added ($p_iy_i = \sum_{k=1}^{K_i} p_{ik}y_{ik}$) gives

$$\frac{wl_i}{p_iy_i} = \sum_{k=1}^{K_i}\left((1-\alpha)\left(1 - \frac{1}{\gamma}\right) - (1-\alpha)\left(\frac{1}{\theta} - \frac{1}{\gamma}\right)\omega_{ik}\right)\frac{p_{ik}y_{ik}}{p_iy_i} + K_i\frac{wF_d}{p_iy_i},$$

or simply

$$\frac{wl_i}{p_iy_i} = (1-\alpha)\left(1 - \frac{1}{\gamma}\right) - (1-\alpha)\left(\frac{1}{\theta} - \frac{1}{\gamma}\right)\omega_i + K_i\frac{wF_d}{p_iy_i}, \tag{4}$$

where we now write

$$\omega_i = \sum_{k=1}^{K_i}\omega_{ik}\frac{p_{ik}y_{ik}}{p_iy_i}, \tag{5}$$

for the weighted average of each product's sectoral share. Observe that single-product firms are a special case of equation (4) with $K_i = 1$.

B.3 Exporters

Including observations from exporting firms in our sample is difficult because we do not observe their sectoral shares in the foreign markets. However, we do know that a version of (2) holds in both their *domestic* (d) market and their *foreign* (f) market. That is

$$\frac{wl_{ik}^d}{p_{ik}^d y_{ik}^d} = (1-\alpha)\left(1 - \frac{1}{\gamma}\right) - (1-\alpha)\left(\frac{1}{\theta} - \frac{1}{\gamma}\right)\omega_{ik}^d + \frac{wF_d}{p_{ik}y_{ik}}, \tag{6}$$

and

$$\frac{wl_{ik}^f}{p_{ik}^f y_{ik}^f} = (1-\alpha)\left(1 - \frac{1}{\gamma}\right) - (1-\alpha)\left(\frac{1}{\theta} - \frac{1}{\gamma}\right)\omega_{ik}^f + \frac{wF_f}{p_{ik}^f y_{ik}^f}. \tag{7}$$

Summing across markets and products

$$\frac{wl_i}{p_iy_i} = (1-\alpha)\left(1 - \frac{1}{\gamma}\right) - (1-\alpha)\left(\frac{1}{\theta} - \frac{1}{\gamma}\right)\left(\sum_{k=1}^{K_i}\omega_{ik}^d\frac{p_{ik}^d y_{ik}^d}{p_iy_i} + \omega_{ik}^f\frac{p_{ik}^f y_{ik}^f}{p_iy_i}\right) + K_i\frac{w(F_d + F_f)}{p_iy_i},$$

where p_iy_i again denotes total value added.

Correlated sales. We now assume that domestic and foreign sectoral shares are *correlated*

$$\omega_{ik}^f = \phi\omega_{ik}^d + \varepsilon_{ik},$$

where ϕ is a parameter and ε_{ik} is i.i.d. noise. Then the regression becomes:

$$\begin{aligned}\frac{wl_i}{p_i y_i} =&(1-\alpha)\left(1-\frac{1}{\gamma}\right) - (1-\alpha)\left(\frac{1}{\theta}-\frac{1}{\gamma}\right)\left(\sum_{k=1}^{K_i}\omega_{ik}^d \frac{p_{ik}^d y_{ik}^d}{p_i y_i} + \phi\omega_{ik}^d \frac{p_{ik}^f y_{ik}^f}{p_i y_i}\right) \\ &-(1-\alpha)\left(\frac{1}{\theta}-\frac{1}{\gamma}\right)\left(\sum_{k=1}^{K_i}\varepsilon_{ik}\frac{p_{ik}^f y_{ik}^f}{p_i y_i}\right) + K_i \frac{w(F_d+F_f)}{p_i y_i}.\end{aligned}$$

Now, analogous to (5), let ω_i^d and ω_i^f denote a weighted-average of each good's *domestic* sectoral share, ω_{ik}^d, with weights given, respectively, by that good's domestic or foreign value added:

$$\omega_i^d = \sum_{k=1}^{K_i}\omega_{ik}^d \frac{p_{ik}^d y_{ik}^d}{p_i^d y_i^d}, \tag{8}$$

and

$$\omega_i^f = \sum_{k=1}^{K_i}\omega_{ik}^d \frac{p_{ik}^f y_{ik}^f}{p_i^f y_i^f}. \tag{9}$$

We can therefore reduce the regression to:

$$\frac{wl_i}{p_i y_i} = (1-\alpha)\left(1-\frac{1}{\gamma}\right) - (1-\alpha)\left(\frac{1}{\theta}-\frac{1}{\gamma}\right)\left(\omega_i^d \frac{p_i^d y_i^d}{p_i y_i} + \phi\omega_i^f \frac{p_i^f y_i^f}{p_i y_i}\right) + K_i\frac{w(F_d+F_f)}{p_i y_i} + \hat{\varepsilon}_i. \tag{10}$$

The composite error term

$$\hat{\varepsilon}_i = (1-\alpha)\left(\frac{1}{\theta}-\frac{1}{\gamma}\right)\left(\sum_{k=1}^{K_i}\varepsilon_{ik}\frac{p_{ik}^f y_{ik}^f}{p_i y_i}\right),$$

is uncorrelated with the other variables, by assumption, but it is not homoscedastic; the variance increases with the export market share. From an OLS regression on (10) we can again recover θ from the coefficient on domestic value added. Again, observe that this nests the non-exporting firms as a special case with $p_i^f y_i^f = 0$.

C Markups estimated with IO methods

Overview. We now explore alternative approaches to estimating markups that are commonly used in the industrial organization literature. Our purpose is to establish to what extent these alternative approaches lead to estimates of markups that are substantially different from our benchmark approach (and hence may substantially affect our estimates of the key across-sector elasticity θ). We emphasize that these IO methods make auxiliary assumptions that are not fully consistent with our model; because of this, these alternative markup estimates should be viewed primarily as a robustness check on our results.

Cobb-Douglas benchmark. In our benchmark analysis we assume a Cobb-Douglas production function $y = ak^{\alpha}l^{1-\alpha}$ with a constant elasticity of output with respect to labor input of $1-\alpha$. Cost minimization with respect to the choice of labor input implies

$$W_t = v_{it}\frac{\partial y_{it}}{\partial l_{it}},$$

or

$$W_t\frac{l_{it}}{y_{it}} = v_{it}\frac{\partial \ln y_{it}}{\partial \ln l_{it}} = v_{it}(1-\alpha), \tag{11}$$

where v_{it} is marginal cost. Defining the markup $m_{it} \equiv p_{it}/v_{it}$, we then have:

$$\frac{W_t l_{it}}{p_{it}y_{it}} = \frac{1-\alpha}{m_{it}}.$$

Since the elasticity $1-\alpha$ is constant, the labor share distribution is directly translated into the (inverse) markup distribution. With this in hand, we can then estimate the key across-sector elasticity of substitution θ by relating markups (labor shares) to market shares, as in Appendix B above. If, however, the elasticity $1-\alpha$ is not constant, then the labor share distribution cannot be translated into the markup distribution in this direct way.

Translog production function. We now follow in the spirit of De Loecker and Warzynski (2012) and assume a more general *translog* production function:

$$\ln y_{it} = \alpha_l \ln l_{it} + \alpha_k \ln k_{it} + \alpha_{ll}\ln l_{it}^2 + \alpha_{kk}\ln k_{it}^2 + \alpha_{kl}\ln l_{it}\ln k_{it} + \tilde{a}_{it}.$$

This specification can also be viewed as an approximation of some arbitrary twice continuously differentiable production function. The Cobb-Douglas benchmark is the special case with $\alpha_{ll} = \alpha_{kl} = \alpha_{kk} = 0$.

With this translog specification, the elasticity of output with respect to labor input is

$$\frac{\partial \ln y_{it}}{\partial \ln l_{it}} = \alpha_l + 2\alpha_{ll}\ln l_{it} + \alpha_{kl}\ln k_{it} \equiv e_{it}^l,$$

hence the output elasticity e_{it}^l varies across firms. Plugging this elasticity into the cost-minimization condition (11) then gives:

$$\frac{W_t l_{it}}{p_{it}y_{it}} = \frac{\alpha_l + 2\alpha_{ll}\ln l_{it} + \alpha_{kl}\ln k_{it}}{m_{it}} = \frac{e_{it}^l}{m_{it}}. \tag{12}$$

We now use different estimation methods from the IO literature to estimate the parameters $\alpha_l, \alpha_{ll}, \alpha_{kl}$ and hence the elasticity e_{it}^l and markups m_{it} from which we can then recover θ from the relationship between markups m_{it} and market shares ω_{it} outlined above. These methods are primarily concerned with addressing various *endogeneity* problems; the most basic of which is that firms' choices of inputs will typically be correlated with the disturbance $\tilde{a}_{it}$, which will bias coefficient estimates. Ackerberg, Benkard, Berry and Pakes (2007) provide an excellent overview of these issues.

Fixed effects regression. We begin by assuming that productivity $\tilde{a}_{it} = a_i + \varepsilon_{it}$ is the sum of a permanent difference across firms a_i and i.i.d. noise ε_{it}. We then estimate the key elasticity e^l_{it} using a simple fixed-effects regression. This gives the following estimated distribution[3] of elasticities and markups:

percentile	5	25	50	75	95	dispersion
elasticity e^l	0.66	0.69	0.72	0.75	0.79	0.18
markups m	1.02	1.13	1.33	1.82	5.15	1.62

Translog fixed-effects markups

The median elasticity is quite close to our benchmark value 0.67 and the distribution of e^l_{it} is not very dispersed (we measure *dispersion* by the log of the ratio of 95th to 5th percentiles). If we impose a Cobb-Douglas specification with constant elasticity $\alpha_l = 0.67$ (and $\alpha_{ll} = \alpha_{kl} = \alpha_{kk} = 0$) we obtain:

percentile	5	25	50	75	95	dispersion
elasticity e^l			0.67			0
markups m	1.02	1.12	1.32	1.82	5.49	1.68

Cobb-Douglas markups

Observe that despite the lack of variation in the output elasticity, clearly the markup distribution implied by this Cobb-Douglas specification is very similar to that implied by the general translog specification estimated by fixed effects regression above.

We report estimates of the key across-sector elasticity of substitution implied by these fixed-effects translog markups in Table 2 in the main text. We find estimates of θ of approximately 1.2 or perhaps even lower, depending on how outliers are treated. Estimates of θ that are lower than our benchmark $\theta = 1.25$ imply larger gains from trade.

De Loecker and Warzynski approach. We next follow De Loecker and Warzynski (2012, DLW herafter) in applying 'control' or 'proxy function' methods inspired by Olley and Pakes (1996), Levinsohn and Petrin (2003) and Ackerberg, Caves and Frazer (2006) to estimate the elasticity of output with respect to labor. Briefly, we begin with the translog production function

$$\ln y_{it} = \alpha_l \ln l_{it} + \alpha_k \ln k_{it} + \alpha_{ll} \ln l_{it}^2 + \alpha_{kk} \ln k_{it}^2 + \alpha_{kl} \ln l_{it} \ln k_{it} + a_{it} + \varepsilon_{it},$$

where a_{it} is the firm's true productivity (which will typically be correlated with its input choices) and ε_{it} is i.i.d. noise. We then suppose that the firm's demand for materials x_{it} can be expressed as a function of the firms capital k_{it} and productivity a_{it} and that this *control function* $x_{it} = f(k_{it}, a_{it})$

[3] We truncate the distribution to enforce markups $m \geq 1$.

can be inverted to uniquely determine the productivity level a_{it} that corresponds to a given k_{it}, x_{it} configuration. Write this as:

$$a_{it} = h(k_{it}, x_{it}).$$

We can now write the conditional mean of log output as

$$\phi(k_{it}, l_{it}, x_{it}) \equiv \alpha_l \ln l_{it} + \alpha_k \ln k_{it} + \alpha_{ll} \ln l_{it}^2 + \alpha_{kk} \ln k_{it}^2 + \alpha_{kl} \ln l_{it} \ln k_{it} + h(k_{it}, x_{it}),$$

so that $\log y_{it} = \phi(k_{it}, l_{it}, x_{it}) + \varepsilon_{it}$ and we can estimate the conditional mean using polynomials. This gives us fitted values $\hat{\phi}_{it}$ and residuals $\hat{\varepsilon}_{it}$ from which we form an estimate of the productivity component

$$\hat{a}_{it}(\boldsymbol{\alpha}) = \hat{\phi}_{it} - \alpha_l \ln l_{it} - \alpha_k \ln k_{it} - \alpha_{ll} \ln l_{it}^2 - \alpha_{kk} \ln k_{it}^2 - \alpha_{kl} \ln l_{it} \ln k_{it}, \tag{13}$$

where $\boldsymbol{\alpha} \equiv (\alpha_l, \alpha_k, \alpha_{ll}, \alpha_{kk}, \alpha_{kl})$. We depart slightly from DLW by assuming that productivity follows an AR(1) process,

$$a_{it} = \rho a_{it-1} + \zeta_{it},$$

rather than allowing a more general nonparametric relationship between a_{it} and its lag. We do this to reduce the dimension of the estimation problem.

We now use our estimate of the productivity process $\hat{a}_{it}(\boldsymbol{\alpha})$ from the first stage (13) to estimate the productivity innovations $\zeta_{it}(\boldsymbol{\alpha})$ from this AR(1). We form the moment conditions:

$$\mathbb{E}\left[\zeta_{it}(\boldsymbol{\alpha})\begin{pmatrix} l_{it-1} \\ k_{it} \\ k_{it-1} \\ l_{it-1}^2 \\ k_{it}^2 \\ k_{it-1}^2 \\ l_{it-1}k_{it-1} \end{pmatrix}\right] = 0.$$

Observe that these moment conditions embed the assumption, standard in this literature, that capital is chosen before a_{it} is realized. We then solve for the estimates of $\boldsymbol{\alpha} \equiv (\alpha_l, \alpha_k, \alpha_{ll}, \alpha_{kk}, \alpha_{kl})$ using GMM. With these estimates in hand, we form the elasticity of output with respect to labor e_{it}^l and calculate the markup using the elasticity and the labor share of revenue as previously.

This gives the following estimated distribution of elasticities and markups:

percentile	5	25	50	75	95	dispersion
elasticity e^l	0.67	0.76	0.81	0.87	0.96	0.36
markups m	1.03	1.15	1.35	1.76	4.17	1.43

Translog DLW markups

Relative to the results from the simple fixed effects translog case, the median markup is almost identical while the amount of markup dispersion is slightly less due to smaller markups in the far right tail. Again, we report estimates of the key across-sector elasticity of substitution implied by these De Loecker and Warzynski (2012)-style markups in Table 2 in the main text. We find estimates of θ of approximately 1.2 or 1.3. These are essentially unchanged from the estimates of θ implied by our benchmark markup estimates.

In short, we find that alternative IO approaches to estimating markups lead to essentially unchanged implications for the inferred distribution of markups and the key across-sector elasticity.

D Gains from trade, additional results

Here we report two additional sets of results on the gains from trade. First, we show for our benchmark model that the gains from tariff reductions are similar to the gains from trade cost reductions. Second, we show how the gains from trade depend on the correlation parameter $\tau(\rho)$.

D.1 Changes in tariffs

Table A3 reports the effect on welfare and productivity of symmetric tariff reductions that increase the import share in both countries from 0 to 10%, 10 to 20%, and 20 to 30%. Panel A shows results for the case of identical Home and Foreign productivity draws, $\tau(\rho) = 1$. Panel B shows results for the case of independent Home and Foreign productivity draws, $\tau(\rho) = 0$.

For example, as shown in Panel A, for $\tau(\rho) = 1$ the model predicts very large gains from moving from autarky to a 10% import share: the welfare gains are equivalent to a 31.6% permanent increase in consumption. These gains are driven by a large (9.8%) drop in markups, but mostly by the very high (18.8%) increase in aggregate TFP associated with the reduction in misallocation. These numbers are very similar to the gains from a trade cost reduction that moves the economy from autarky to a 10% import share (as reported in Tables 5 and 6 in the main text).

D.2 General amounts of dependence

In our general model, the joint distribution of productivities $H(a, a^*)$ is given by

$$H(a, a^*) = \mathcal{C}(F(a), F(a^*)), \tag{14}$$

where the (inverse) marginal distribution $a = F^{-1}(u)$ is given by (1) above and where $\mathcal{C}(u, u^*)$ is the *Gumbel* copula:

$$\mathcal{C}(u, u^*) = \exp\left(-\left[(-\log u)^\rho + (-\log u^*)^\rho\right]^{\frac{1}{\rho}}\right), \quad \rho \geq 1. \tag{15}$$

The parameter ρ controls the amount of dependence between u and u^*. In particular, for the Gumbel copula the robust measure of correlation known as *Kendall's* τ is given by:

$$\tau(\rho) = 1 - 1/\rho.$$

For $\tau(\rho) = 0$, $C(u, u^*) = uu^*$, i.e., the distributions are *independent*. For $\tau(\rho) = 1$, $C(u, u^*) = \min(u, u^*)$, i.e., the distributions are *identical* (perfectly dependent).

Table A4 shows the key results for our model for correlation coefficients $\tau(\rho)$ ranging from 0 to 1. We report the welfare gains, productivity gains and change in markups for changes in iceberg trade costs that induce import shares to move from zero to 10%, 10% to 20%, and 20% to 30%. We also report measures of intra-industry trade, namely the index of import share dispersion (weight on γ in the Armington elasticity) and the Grubel and Lloyd (1971) index. These intra-industry trade measures are calculated using the post-liberalization import shares.

We compute welfare as the equivalent percentage change in lifetime consumption including the transition path to the new steady state. So, for example, when we compute the gains from trade obtained by moving from an import share of 10% to an import share of 20%, we imagine a once-and-for-all reduction in tariffs such that the import share is 20% in the new steady state and we compute the equivalent percentage change in lifetime consumption inclusive of the transition path to this new steady state.

E Robustness experiments

Here we provide further details of our robustness experiments. For each experiment, we recalibrate parameters to match the target moments unless stated otherwise. The parameters used for each experiment are shown in Table A5. The target moments and the moments implied by each model are shown in Table A6. In these tables we also record the parameter values used in our calibrated standard trade model (with constant markups), as discussed in Section 7 in the main text. The welfare results themselves are given in Table A7. Finally, to establish the role of initial misallocation in generating welfare gains, Table A8 shows the distance between the Autarky and First-Best versions of each of these economies.

In this appendix we focus on the results for changes in the tariff rate for the case of identical productivity draws, $\tau(\rho) = 1$. We change the tariff ξ so that the model reproduces import shares of 0% (autarky), 10%, 20% and 30%.

E.1 Within-sector elasticity γ

For this exercise we reduce the within-sector elasticity γ from its benchmark value 8.5 to $\gamma = 5.2$ so that the model produces a lower Armington elasticity of 5. As shown in Table A7, the gains from trade are large and similar to the benchmark economy.

E.2 Across-sector elasticity θ

For these exercises we consider either a low cross-sector elasticity, $\theta = 1$ (i.e., constant sectoral expenditure shares), as in Atkeson and Burstein (2008), or a high across-sector elasticity, $\theta = 3$. As shown in Table A7, we find that the gains from trade are considerably larger than in our benchmark model when we use the Atkeson and Burstein value of $\theta = 1$. As shown in inTable A8 this largely because with $\theta = 1$ there is considerably more initial misallocation; in particular, markup dispersion is much larger than in the benchmark. Consistent with our results on the importance of initial misallocation, we also find smaller gains when we use $\theta = 3$ for which the initial amount of markup dispersion is much smaller.

Recall, however, that a value of $\theta = 3$ is more than twice as high as any value implied by the cross-sectional relationship between market shares and labor and capital shares in our data (see the main text for further discussion).

E.3 Bertrand competition

For this exercise we re-solve the model under the assumption that producers compete by simultaneously choosing prices (Bertrand) rather than simultaneously choosing quantities (Cournot). As discussed in the main text, this changes the model set-up in only one way. The demand elasticity facing a producer is now an arithmetic, rather than harmonic, weighted average of the substitution elasticities γ and θ. As shown in Table A7, the gains from trade with Bertrand competition are almost identical to the gains with Cournot competition. When opening an economy from autarky to an import share of 10%, the gains are 32.3% with Bertrand as opposed to 31.6% with Cournot. The marginal gains from increasing the import share from 10% to 20% or 20% to 30% are slightly lower under Bertrand competition, however.

E.4 5-digit sectors

For this exercise we re-calibrate the model to target statistics for *5-digit* Taiwanese manufacturing data. The target moments are shown in the second-last column of Table A6. The gains from trade at this higher level of aggregation are smaller than in our benchmark. When opening an economy from autarky to an import share of 10%, the gains are 14.4% in the 5-digit model as opposed to 31.6% in the 7-digit benchmark. Note however that in this exercise we keep the across-sector elasticity θ fixed at its benchmark value of 1.25. The gains from trade would be larger if we reduced θ to reflect less substitution between sectors at this higher level of aggregation.

E.5 No fixed costs

Next, we solve our model assuming that fixed costs are zero, $F_d = F_f = 0$. In this version of the model there is no entry or exit by domestic producers. As shown in Table A7, this version of the model produces almost identical results to our benchmark model. The presence of many tiny unproductive producers does not affect the market power of the largest, most productive producers. In this sense, the fixed costs have very little consequence for the model's aggregate implications.

E.6 Single Pareto

Finally, we solve our model when productivity draws are taken from a single Pareto distribution rather than the double Pareto we use for our other experiments. As shown in Table A6, the single Pareto cannot simultaneously account for both the lower and upper tails of the size distribution of producers. We consider two alternative calibration strategies. First, we study a single Pareto economy that targets all the moments *except* the size distribution of establishments. Second, we study a single Pareto economy that targets *only* the size distribution of establishments.

We find that the welfare gains from trade are smaller in these alternative economies than in our original setup. For example, the welfare gains from a reduction in tariffs that increases the import share from 0 to 10% leads to a 31.6% welfare gain in our benchmark model, and only a 4.1% and 8.2% gain, respectively, in the two single Pareto calibrations we consider. Likewise, the gains from an increase in the import share from 10 to 20% are 6.5% in our benchmark model, and only 3.6% and 5.8% in the single Pareto economies.

The welfare gains in these economies are smaller precisely because of these models' inability to account for the pattern of concentration in the data. The failure to account for the pattern of concentration implies that the distortions arising due to markups are much smaller now and so there is less scope for trade to improve allocations. For example, as shown in Table A8, the level of TFP in the two single Pareto economies is 8.9% and 15.4% away from its first-best level, and 24.6% away from it first-best in our benchmark economy. These results thus reinforce our conclusion that the size of the pro-competitive gains from trade is larger, the larger is the amount of misallocation due markups in the economy.

References

Ackerberg, Daniel A., C. Lanier Benkard, Steven Berry, and Ariel Pakes, "Econometric Tools for Analyzing Market Outcomes," in James J. Heckman and Edward E. Leamer, eds., *Handbook of Econometrics*, Vol. 6A, Elsevier, 2007, chapter 63, pp. 4173–4276.

_ **, Kevin Caves, and Garth Frazer**, "Structural Identification of Production Functions," 2006. UCLA working paper.

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

De Loecker, Jan and Frederic Warzynski, "Markups and Firm-Level Export Status," *American Economic Review*, 2012, *forthcoming.*

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

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

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

Table A1: Data Description

Panel A: *An example of product classification*

3-digit	314 - computers and storage equipment
5-digit	31410 - computers
7-digit	3141000 - mini-computer
	3141010 - work-station
	3141021 - desktop computer
	3141022 - laptop computer
	3141023 - notebook computer
	3141024 - palmtop computer
	3141025 - pen-based computer
	3141026 - hand held computer
	3141027 - electronic dictionary

Panel B: *Distribution of products and industries*

2-digit	4-digit (industry)	7-digit (sector)
textile	16	76
apparel	10	39
leather	4	33
lumber	6	15
furniture	4	12
paper	6	23
printing	3	4
chemical materials	7	152
chemical products	9	83
petroleum	2	12
rubber	3	16
plastics	7	34
clay/glass/stone	18	47
primary metal	14	99
fabricated metal	14	65
industrial machinery	29	163
computer/electronics	11	136
electronic parts	6	72
electrical machinery	11	125
transportation	12	99
instruments	7	70

Table A2: Identifying μ_H

		Double Pareto			Single Pareto	
	Data	$\mu_H = 1.05$	$\mu_H = 1.5$	$\mu_H = 2.5$	$\mu = 3.4$	$\mu = 1.95$
Panel A: *Within-sector concentration, domestic producers*						
mean inv HH	7.25	5.28	5.50	5.13	5.07	2.75
median inv HH	3.95	4.86	5.08	4.71	4.31	2.35
mean highest share	0.45	0.50	0.49	0.50	0.43	0.60
median highest share	0.39	0.36	0.35	0.37	0.39	0.56
Panel B: *Distribution of sectoral shares, including importers*						
mean share	0.029	0.027	0.028	0.028	0.028	0.019
s.d. share	0.087	0.074	0.074	0.075	0.069	0.076
median share	0.0036	0.0040	0.0045	0.0039	0.0041	0.0000
75th p.c. share	0.016	0.015	0.016	0.015	0.016	0.001
95th p.c share	0.141	0.141	0.143	0.145	0.154	0.108
99th p.c. share	0.461	0.464	0.460	0.458	0.382	0.465
Panel C: *Size distribution of establishments*						
fraction value added by top 1%	0.48	0.53	0.40	0.30	0.21	0.38
fraction value added by top 5%	0.67	0.76	0.68	0.64	0.57	0.87
fraction wage bill by top 1%	0.32	0.44	0.32	0.23	0.17	0.32
fraction wage bill by top 5%	0.53	0.70	0.62	0.57	0.52	0.83
root mean squared log deviation		0.17	0.14	0.20	0.31	1.51

Target moments in the 7-digit Taiwanese manufacturing data and the implications of our model for various values of the second Pareto tail parameter μ_H. To compute *within-sector* concentration statistics, we first compute the statistic for each sector and then take averages over those sector-specific statistics. See the text for further details.

Table A3: Gains from Trade Due to Reductions in Tariffs

Δ import share	0% to 10%	10% to 20%	20% to 30%
Panel A: *Identical Home and Foreign productivity*			
welfare gains, %	31.6	6.5	3.2
markup change, %	−9.8	−4.3	−4.4
TFP gains, %	18.8	3.1	0.8
Armington Elast (post)	7.4	7.8	8.2
Panel B: *Independent Home and Foreign productivity*			
welfare gains, %	14.9	27.5	11.1
markup change, %	10.9	6.4	−6.2
TFP gains, %	15.5	22.9	5.1
Armington Elast (post)	4.1	2.4	3.8

Gains from trade due to changes in tariffs ξ that change the average import share. Welfare gains measured in consumption equivalents including transition. Panel A shows results for our benchmark model with identical Home and Foreign productivity draws, $\tau(\rho) = 1$. Panel B shows results for the same model but with independent Home and Foreign productivity draws, $\tau(\rho) = 0$.

Table A4: Gains from Trade due to Reductions in Trade Costs, General Case

Gumbel Copula with Kendall's Correlation $\tau(\rho)$

Panel A: 0 to 10% import share

	I: Welfare Gains			II: TFP Gains			III: Markup Change	IV: ACR Formulas		V: Intra-Industry Trade	
$\tau(\rho)$	var. markups	import shares	first-best	var. markups	import shares	first-best		single-sector	multi-sector	weight on γ	GL index
1	26.1	2.8	0.3	13.4	2.0	0.1	−14.5	1.4	2.0	0.85	1.00
0.9	25.6	3.2	4.1	13.9	2.4	2.9	−11.9	1.4	2.3	0.76	0.83
0.8	24.9	3.7	7.7	14.3	2.7	5.5	−9.4	1.4	2.7	0.67	0.64
0.7	23.8	4.4	12.0	14.9	3.2	8.7	−5.8	1.4	3.1	0.54	0.41
0.6	21.1	4.7	15.3	14.6	3.4	11.0	−1.8	1.4	3.4	0.43	0.30
0.5	17.9	4.6	16.0	13.5	3.4	11.6	0.9	1.4	3.3	0.35	0.20
0.4	13.2	4.0	15.0	11.1	2.9	10.8	3.4	1.4	2.9	0.31	0.08
0.3	11.0	3.6	13.9	9.7	2.7	10.0	3.9	1.4	2.6	0.34	0.04
0.2	9.7	3.5	12.8	8.8	2.6	9.2	3.9	1.4	2.5	0.36	0.00
0.1	9.3	3.5	12.3	8.4	2.6	8.9	3.8	1.4	2.5	0.38	0.00
0	8.9	3.7	12.2	8.2	2.7	8.8	3.9	1.4	2.6	0.39	0.00

Panel B: 10 to 20% import share

	I: Welfare Gains			II: TFP Gains			III: Markup Change	IV: ACR Formulas		V: Intra-Industry Trade	
$\tau(\rho)$	var. markups	import shares	first-best	var. markups	import shares	first-best		single-sector	multi-sector	weight on γ	GL index
1	7.6	2.2	1.2	4.7	1.6	0.8	−2.7	1.6	1.6	0.91	1.00
0.9	8.5	2.6	2.1	5.2	1.9	1.5	−2.9	1.6	1.9	0.85	0.89
0.8	10.3	3.2	3.7	6.3	2.4	2.6	−3.3	1.6	2.4	0.79	0.78
0.7	13.4	4.4	6.6	8.3	3.2	4.7	−3.7	1.6	3.3	0.69	0.67
0.6	18.0	6.3	11.0	11.5	4.6	7.9	−4.1	1.6	4.7	0.60	0.57
0.5	23.1	8.6	16.5	15.1	6.3	11.9	−3.8	1.6	6.3	0.52	0.48
0.4	30.5	12.6	26.8	21.5	9.2	19.4	−1.2	1.6	9.0	0.39	0.38
0.3	34.4	15.8	34.7	25.9	11.5	25.0	2.2	1.6	11.1	0.30	0.31
0.2	37.1	18.7	41.7	29.6	13.6	30.1	6.0	1.6	13.0	0.21	0.23
0.1	36.8	18.9	43.5	30.6	13.8	31.4	8.4	1.6	13.1	0.14	0.11
0	28.8	14.5	36.4	25.4	10.6	26.3	9.6	1.6	10.3	0.13	0.05

Panel C: 20 to 30% import share

	I: Welfare Gains			II: TFP Gains			III: Markup Change	IV: ACR Formulas		V: Intra-Industry Trade	
$\tau(\rho)$	var. markups	import shares	first-best	var. markups	import shares	first-best		single-sector	multi-sector	weight on γ	GL index
1	5.8	2.2	2.4	3.9	1.7	1.7	−1.2	1.8	1.7	0.97	1.00
0.9	5.9	2.4	2.9	4.0	1.8	2.1	−1.2	1.8	1.8	0.92	0.90
0.8	6.3	2.7	3.7	4.2	2.0	2.6	−1.2	1.8	2.0	0.86	0.81
0.7	6.9	3.1	4.7	4.6	2.3	3.4	−1.2	1.8	2.4	0.79	0.72
0.6	7.7	3.7	5.9	5.2	2.7	4.2	−1.2	1.8	2.8	0.72	0.63
0.5	8.4	4.2	6.9	5.7	3.1	5.0	−1.1	1.8	3.2	0.65	0.55
0.4	10.3	5.6	9.1	7.0	4.1	6.5	−1.0	1.8	4.2	0.55	0.47
0.3	12.2	7.0	11.4	8.4	5.1	8.2	−0.9	1.8	5.3	0.48	0.40
0.2	14.7	8.9	14.3	10.4	6.5	10.3	−0.6	1.8	6.6	0.41	0.34
0.1	20.7	12.9	20.9	15.0	9.3	15.1	−0.1	1.8	9.4	0.34	0.29
0	35.3	22.6	37.7	26.6	16.4	27.1	2.3	1.8	16.0	0.27	0.24

Table A5: Parameterization of Robustness Experiments

	Double Pareto								Single Pareto	
	Benchmark	Arm = 5	$\theta = 1$	$\theta = 3$	Bertrand	5-digit	No fixed	Standard	All but size	Only size
N, number competitors	95	95	95	60	90	95	95	65	90	40
μ_L, first Pareto tail	3.83	2.60	3.60	4.60	4.60	4.90	3.83	4.80	3.40	1.95
μ_H, second Pareto tail	1.50	1.35	1.00	2.10	2.10	2.80	1.50	2.00	–	–
p_H, mass second Pareto	0.0035	0.0035	0.0035	0.0035	0.0035	0.0035	0.0035	0.002	0	0
F_f, fixed cost export, % S.S. labor	4.5	12.0	5.0	0.2	6.0	1.8	0	17	5	0
F_d, fixed cost, % S.S. labor	2.2	6.0	2.5	0.1	2.5	0.9	0	6	2	0.0004
γ, within-sector elasticity	8.45	5.20	8.50	8.00	8.20	8.50	8.45	8.02	8.20	8.45
θ, across-sector elasticity	1.25	1.25	1	3	1.25	1.25	1.25	1.25	1.25	1.25
τ, iceberg trade cost	0.21	0.330	0.25	0.11	0.12	0.17	0.21	0.07	0.21	0.33
Common to all experiments										
β, discount factor	0.96									
δ, capital depreciation	0.10									
α, capital intensity	0.33									
ξ, baseline tariff rate	0.064									

The parameters used in all of the robustness experiments. The target moments in the data and the corresponding moments implied by these models are given in Table A6 below. For all of these models we assume identical Home and Foreign productivity draws, $\tau(\rho) = 1$. See the text for further discussion.

Table A6: Moments Implied by Robustness Experiments

		Double Pareto							Single Pareto		5-Digit	
	Data	Benchmark	Arm = 5	$\theta = 1$	$\theta = 3$	Bertrand	No fixed	Standard	All but size	Size only	Data	Model
Panel A: *Within-sector concentration statistics, domestic producers*												
mean inv HH	7.25	5.50	5.87	5.66	5.53	5.10	5.73	4.94	5.07	2.75	14.90	8.66
median inv HH	3.95	5.08	5.15	5.38	4.51	4.03	5.27	3.60	4.31	2.35	7.97	8.13
mean highest share	0.45	0.49	0.49	0.48	0.48	0.52	0.49	0.52	0.43	0.60	0.30	0.43
median highest share	0.39	0.35	0.35	0.33	0.39	0.43	0.34	0.46	0.39	0.56	0.25	0.26
Panel B: *Distribution of sectoral shares, including importers*												
mean share	0.029	0.028	0.028	0.028	0.029	0.027	0.008	0.026	0.028	0.019	0.011	0.011
s.d. share	0.087	0.074	0.075	0.072	0.078	0.079	0.041	0.081	0.069	0.076	0.040	0.044
median share	0.0036	0.0045	0.0055	0.0044	0.006	0.0044	0.0003	0.0045	0.0041	0.0000	0.0014	0.0018
mean 75th p.c. share	0.016	0.016	0.016	0.017	0.017	0.013	0.001	0.012	0.016	0.001	0.006	0.005
mean 95th p.c. share	0.141	0.143	0.132	0.143	0.135	0.127	0.027	0.118	0.154	0.108	0.044	0.045
mean 99th p.c. share	0.461	0.460	0.462	0.439	0.45	0.487	0.183	0.497	0.382	0.465	0.173	0.187
Panel C: *Size distribution of establishments*												
fraction v.a. top 1%	0.48	0.40	0.39	0.26	0.97	0.37	0.61	0.33	0.21	0.38		0.43
fraction v.a. top 5%	0.67	0.68	0.67	0.59	0.99	0.69	0.88	0.67	0.57	0.87		0.71
fraction wl top 1%	0.32	0.32	0.31	0.17	0.97	0.35	0.54	0.33	0.17	0.32		0.35
fraction wl top 5%	0.53	0.62	0.61	0.51	0.98	0.67	0.85	0.67	0.52	0.83		0.66
Panel D: *Markups and additional statistics*												
aggregate markup		1.48	1.60	1.48	1.34	1.23	1.48	1.16	1.36	1.53		1.37
mean markup		1.17	1.27	1.18	1.15	1.14	1.14	1.14	1.17	1.16		1.15
s.d. markup		0.11	0.11	0.16	0.02	0.02	0.06	0	0.09	0.11		0.06
median markup		1.14	1.24	1.14	1.14	1.14	1.14	1.14	1.14	1.13		1.14
75th p.c. markup		1.15	1.25	1.15	1.15	1.14	1.14	1.14	1.15	1.14		1.14
95th p.c. markup		1.28	1.38	1.32	1.18	1.16	1.16	1.14	1.29	1.24		1.17
99th p.c. markup		1.76	1.90	2.00	1.28	1.26	1.32	1.14	1.62	1.77		1.33
mean labor share	0.61	0.62	0.60	0.61	0.63	0.63	0.58	0.74	0.61	0.58		0.62
median labor share	0.65	0.62	0.60	0.61	0.62	0.63	0.58	0.72	0.61	0.59		0.62
aggregate labor share	0.43	0.45	0.42	0.45	0.50	0.54	0.45	0.60	0.49	0.44		0.49
s.d. labor product	0.68	0.10	0.13	0.11	0.07	0.06	0.03	0.11	0.09	0.11		0.07
mean import share	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.25	0.26	0.26	0.26
fraction export	0.25	0.25	0.25	0.25	0.25	0.25	1	0.25	0.26	0.25	0.25	0.25
Armington elasticity	8	8	5	8	8	8	8.1	8	8	8	8	8.1
mean number producers	26	26	27	27	26	28	95	28	27	40		65
median number producers	11	29	29	29	28	29	95	27	25	40		94

Table A7: Gains from Trade, Robustness Experiments

Welfare gains from trade due to reductions in tariffs, correlated productivity

Δ import share	0% to 10%	10% to 20%	20% to 30%
		Double Pareto	
Benchmark	31.6	6.5	2.2
Armington = 5	29.3	8.1	4.2
$\theta = 1$	53.1	7.4	3.1
$\theta = 3$	8.9	2.7	1.4
Bertrand	32.3	3.5	1.7
5-digit	14.4	5.6	2.7
No fixed costs	30.6	7.1	3.1
		Single Pareto	
All but size moments	4.1	3.6	5.1
Only size moments	8.2	5.8	6.4

Gains from trade due to changes in tariffs ξ that change the average import share. For each robustness exercise we recalibrate the model with parameters given in Table A5 and with target and model moments given in Table A6 above. Welfare gains measured as consumption-equivalent variations including the transition path. See the text for further discussion.

Table A8: Losses Due to Markups

	Autarky		Autarky to First-Best	
	aggregate markup	s.d. markup	TFP gains, %	welfare gains, %
		Double Pareto		
Benchmark	1.75	0.33	24.6	53.0
Armington = 5	1.87	0.31	21.0	52.0
$\theta = 1$	1.77	8.01	84.2	53.1
$\theta = 3$	1.48	0.03	0.4	10.7
Bertrand	1.47	0.25	17.6	24.1
5-digit	1.44	0.14	15.2	29.9
No fixed costs	1.74	0.20	24.8	53.0
		Single Pareto		
All but size moments	1.39	0.12	8.9	19.7
Only size moments	1.67	0.20	15.4	37.6

For each of the robustness exercises, we raise the tariff rate to $\xi = 1$ to shut down all trade. In the first two columns we report the aggregate markup and standard deviation of markups under Autarky. We then calculate the distance from the Autarky economy to the corresponding first-best economy. We implement the first-best by choosing producer-specific subsidies or taxes such that all producers set a markup of 1. The subsidies/taxes are financed with a lump-sum tax on the representative consumer. Welfare gains are measured as consumption-equivalent variations including the transition path. See the text for further discussion.