7-2013

Allocative Efficiency, Mark-ups, and the Welfare Gains from Trade

Thomas J. HOLMES
University of Minnesota

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

Sanghoon LEE
University of British Columbia

NBER WORKING PAPER SERIES

ALLOCATIVE EFFICIENCY, MARK-UPS, AND THE WELFARE GAINS FROM TRADE

Thomas J. Holmes
Wen-Tai Hsu
Sanghoon Lee

Working Paper 19273
http://www.nber.org/papers/w19273

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2013

This paper grew out of work initially circulated under the title, "Plants, Productivity, and Market Size, with Head-to-Head Competition." We are very grateful for discussion comments from Donald Davis and Marc Melitz on this earlier work. We have also benefited from discussions with various colleagues and in particular thank Sam Kortum and Jim Schmitz. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis, the Federal Reserve System, or the National Bureau of Economic Research.

Allocative Efficiency, Mark-ups, and the Welfare Gains from Trade
Thomas J. Holmes, Wen-Tai Hsu, and Sanghoon Lee
NBER Working Paper No. 19273
July 2013
JEL No. D61,F10,L13

ABSTRACT

This paper develops an index of allocative efficiency that depends upon the distribution of mark-ups across goods. It determines how changes in trade frictions affect allocative efficiency in an oligopoly model of international trade, decomposing the effect into the cost-change channel and the price-change channel. Formulas are derived shedding light on the signs and magnitudes of the two channels. In symmetric country models, trade tends to increase allocative efficiency through the cost-change channel, yielding a welfare benefit beyond productive efficiency gains. In contrast, the price-change channel has ambiguous effects on allocative efficiency.

Thomas J. Holmes
Department of Economics
University of Minnesota
4-101 Hanson Hall
1925 Fourth Street South
Minneapolis, MN 55455
and The Federal Reserve Bank of Minneapolis
and also NBER
holmes@umn.edu

Wen-Tai Hsu
School of Economics
Singapore Management University
90 Stamford Road
Singapore 178903
wentaihsu@smu.edu.sg

Sanghoon Lee
Sauder School of Business
University of British Columbia
2053 Main Mall
Vancouver, BC V6T1Z2
Canada
Sanghoon.Lee@sauder.ubc.ca

1 Introduction

When mark-ups are the same across all goods, first-best allocative efficiency is attained. The condition that the price ratio equals the marginal cost ratio, for any pair of goods, holds because the constant mark-ups in prices cancel out. In this paper, we develop an index W^A of allocative efficiency that can be calculated when mark-ups *differ* across goods, and the first-best *is not* attained. We focus on how international trade influences W^A. In particular, we distinguish effects on allocative efficiency from standard Ricardian gains from trade, which we account for through how trade affects an index of productive efficiency W^{Prod}. Our key result is a decomposition of the effect on allocative efficiency W^A into what we define as the *cost-change channel* and the *price-change channel*. The decomposition is useful because each channel has an intuitive formula that makes it possible to discuss conditions determining sign and magnitude. In important limiting cases, both terms are zero, and effects of international trade on W^A can be safely ignored. In general, however, the two terms are not zero, and the effect on allocative efficiency can be a significant component of the overall welfare analysis of trade.

The analysis is conducted in an oligopoly model of international trade where firms compete "head-to-head" in a Bertrand fashion. This approach follows Bernard, Eaton, Jensen, and Kortum (2003) (hereafter BEJK), and Atkeson and Burstein (2008).[1] Consider how W^A changes when a friction τ impeding trade between countries is *reduced*. To determine the cost-change channel, we evaluate the effect of lower τ on mark-ups, when we take into account effects on costs, but leave prices fixed. Lower τ affects only the costs of imported goods. Thus, holding prices fixed, lower τ raises mark-ups on imported goods. To determine how this change affects allocative efficiency W^A, the formula for the cost-change channel compares mark-ups for imported goods, with the average mark-up (foreign and domestic goods combined). The formula is intuitive, and is straightforward to calculate in empirical applications, if micro data on product-level mark-ups are available. Suppose, for example, that mark-ups on foreign goods are initially less than average. If τ is then lowered, it will enable foreign firms to raise mark-ups closer to the average, attenuating the initial distortion.

In the limiting case where the friction is small and countries are symmetric, the cost-change channel for the effect on W^A goes to zero, because mark-ups on imports and domestic goods are identical in the limit. The cost-change channel is also approximately zero in another limiting case where competing firms draw their productivities from a Pareto distri-

[1] Atkeson and Burstein (2008) focus on the Cournot version of their model, but also consider a Bertrand variant. See also Devereux and Lee (2001) and Neary (2003) for related Cournot versions.

bution, as is commonly assumed in the trade literature.[2] Holding fixed productivity draws, foreign firms incur trade costs that domestic firms do not, and everything else the same, this tends to lower mark-ups for foreign firms compared to domestic. However, foreign firms face a tougher selection process over productivity (since foreign firms must surmount the trade cost barrier), and everything else the same, higher productivity for the market leader raises mark-ups. Under the Pareto, these two offsetting forces cancel out, and mark-ups for imports are the same as overall, zeroing out the cost-change channel. The outcome is a consequence of the "fat-tailed" nature of the Pareto, which gives the selection effect great force. If instead we use a distribution with less of a fat tail, like the log-normal, the selection effect no longer "keeps up" as an offsetting force. Everything else the same (e.g. a symmetric setup where foreign firms draw from the same productivity distribution as domestic firms), foreign firms tend to have lower mark-ups than domestic because of the friction, and the cost-change channel for the effect of a reduction in τ on W^A is strictly positive.

To understand the price-change channel, consider the effect of lower τ on allocative efficiency W^A, when we take into account how prices change, but hold costs fixed. It turns out that in the two limiting cases just mentioned, the case of symmetry and negligible frictions, and the case where productivity draws are Pareto, the price-change channel is zero like the cost-change channel, and the overall effect of lower τ on W^A is zero. More generally the price-change channel is non-zero, and its sign and magnitude depend upon how mark-ups on goods whose prices decrease, when τ goes down, compare with mark-ups on goods whose prices remain the same. In symmetric cases when demand tends to be inelastic, the price-change channel tends to be positive, and thereby reinforces the positive effect of lower τ that comes through the cost-change channel. In contrast, when demand is elastic, the price-change channel tends to be negative, and in some cases can more than offset the positive effect from the cost-change channel. In these cases, allocative efficiency actually falls as trade frictions decline, as firms are less able to harmonize their mark-ups around the simple monopoly mark-up.

Our paper builds on long understood ideas about allocative efficiency. In particular, Robinson (Ch 27, 1934) showed that if there is a constant mark-up across all goods, first-best efficiency is achieved. The literature on the theory of the second best (e.g. Lipsey and Lancaster (1956-1957)) made the point that making one sector more competitive potentially reduces welfare if there already exists monopoly distortions elsewhere in the economy. Based on the insights of this old literature, it is clear that while increased trade might have "pro-competitive" effects in reducing mark-ups, the effect of trade on allocative efficiency will not

[2] A similar limiting case is when firms draw productivities from a fat-tailed distribution and the number of firms goes to infinity.

necessarily be positive.

The related modern literature is extensive and we start with BEJK. Our model is the same as BEJK, with BEJK making a particular functional form assumption for the productivity distribution. BEJK show in their setup that the mark-up distribution is the same for imports as it is for domestic goods, and that changes in trade frictions do not affect the distribution of mark-ups. The BEJK productivity distribution has a fat-tailed shape, and the cost-change and price-change channels are both zero, for how τ affects W^A, following our discussion above. Atkeson and Burstein (2008) and de Blas and Russ (2012) start with the BEJK model and show that with alternative assumptions on the distribution of productivity, foreign goods can have different mark-ups than domestic goods, and changes in trade frictions can affect the overall distribution of mark-ups. Our work is different from these papers, in our focus on allocative efficiency.

Recently, Edmond, Midrigan, and Xu (2012) consider a similar model and examine gains from trade achieved through the "pro-competitive" effect of how trade changes the distribution of mark-ups. The paper provides a quantitative analysis, with model parameters pinned down with Taiwanese manufacturing data. The key differences in our paper include (i) our approach in developing a formal measure of allocative efficiency, (ii) how we decompose the effects of trade on allocative efficiency into the cost-change and price-change components, and (iii) how we use the decomposition to shed light on the potential signs and magnitudes of the pro-competitive effect.

In a recent paper, Arkolakis, Costinot, and Rodriguez-Clare (2012), hereafter ACR, derive a condition summarizing the welfare gains from trade that is applicable in a variety of models, including BEJK. The condition depends upon the volume of observed trade. For example, in the ACR framework, a necessary condition for trade to have welfare effects is that there be positive trade flows. By focusing on the fat-tailed productivity draws included in the BEJK setup, for which the mark-up distribution is invariant to trade, the ACR approach shuts down any possibility of welfare effects through allocative efficiency. All welfare effects go through a *productive efficiency index* W^{Prod}. If instead we consider productivity distributions without a fat tail, then in general trade will affect both productive efficiency W^{Prod} and allocative efficiency W^A. In the end, if observed trade volume is zero, the ACR formula will determine that trade leads to no gains in W^{Prod}. However, even if there are no observed trade flows, the possibility of trade can affect the mark-up distribution, and hence overall welfare through W^A.[3]

There are now several models of monopolistic competition where trade affects mark-

[3] There is empirical evidence that the threat of competition from imports can influence domestic outcomes, even if in the end, imports don't come in. See Salvo (2010) and Schmitz (2005).

ups, including Melitz and Ottaviano (2008) and Behrens and Murata (2012).[4] Arkolakis, Costinot, Donaldson, and Rodriguez-Clare (2012) is particularly relevant as it generalizes the monopolistic competition version of the ACR framework to capture what they refer to as "the elusive pro-competitive effects of trade." The economics of the pro-competitive effect is very different in a monopolistic competition model than it is in the oligopoly model we consider. In monopolistic competition, a change in the trade friction only affects a domestic firm through general equilibrium effects that might shift or rotate the firm's demand curve. Depending on assumptions about the shape of the utility function, monopoly demand can become more or less elastic, and domestic mark-ups can go down or up. In contrast, in a Bertrand environment, the pro-competitive force of trade operates at the level of the particular good, not through general equilibrium. If trade frictions are lowered, a domestic firm limit pricing on a foreign rival will directly have to lower price (and mark-up) to meet competition.

The literature discussion above focuses specifically on trade. We note our paper is also part of a broader literature on how allocative efficiency affects aggregate productivity, including Restuccia and Rogerson (2008), Hsieh and Klenow (2009), and Peters (2012).[5]

2 Model

There are a continuum of goods on the unit interval, each good indexed by ω. Let q_ω indicate a quantity of consumption good ω and let $\mathbf{q} = \{q_\omega : \omega \in [0,1]\}$ denote a particular consumption bundle. Assume preferences are represented by a homothetic utility function $\mathbf{U}(\mathbf{q})$ over a bundle $\mathbf{q}$. For most results, we impose the CES form,

$$\begin{aligned} \mathbf{U}(\mathbf{q}) &= \left(\int_0^1 q_\omega^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}}, \text{ for } \sigma > 0, \ \ \sigma \neq 1 \\ &= \exp\left(\int_0^1 \ln q_\omega d\omega \right), \text{ for } \sigma = 1 \text{ (Cobb-Douglas).} \end{aligned}$$

For simplicity of exposition, assume there are two countries, $i = 1, 2$. Each country has a measure L_i workers. Labor is the only factor of production.

The different goods potentially vary in the number of firms capable of producing the good and the productivity of the various firms. In particular, there are $n_{\omega,i}$ different firms capable of producing good ω at i. The total number of firms for good ω across the two

[4] See also Ottaviano, Tabuchi, and Thisse (2002) for a treatment in a regional context.

[5] Peters (2012) in particular uses an index of allocative efficiency to examine growth that coincides with our measure for the case of Cobb-Douglas.

countries is $n_\omega = n_{\omega,1} + n_{\omega,2}$. Assume that $n_\omega \geq 2$. Let $k \in \{1, 2, ..n_{\omega,i}\}$ index a particular firm located at i capable of producing good ω at i.

Define $b_{k,\omega,i}$ as labor requirement per unit produced of firm k located at i for good ω. The inverse of $b_{k,\omega,i}$ is the firm's productivity $x_{k,\omega,i}$. Suppose we rank all $n_{\omega,i}$ firms at i for good ω in terms of labor requirement, from lowest to highest. Let $b^*_{\omega,i}$ and $b^{**}_{\omega,i}$ be labor requirements of the firms with the first and second lowest values. If there is only one firm for good ω at i, then set $b^{**}_{\omega,i} = \infty$. If there are no firms for good ω at i (implying there are two or more in the other country) then set $b^*_{\omega,i} = \infty$ and $b^{**}_{\omega,i} = \infty$. The third lowest and beyond will not be relevant for pricing or production, following the standard logic of Bertrand competition (Grossman and Helpman (1991)), which will apply here.

We treat the labor requirements for a given good ω as random draws from a joint distribution $G_\omega(b^*_1, b^{**}_1, b^*_2, b^{**}_2)$ of the first and second lowest labor requirement at 1, and the first and second lowest at 2. Our main results do not impose the restriction that labor requirements be drawn independently across firms.

In addition to the labor requirement, the cost of a firm at i to deliver to j will depend upon the wage at i and the trade friction to deliver from i to j. Let w_i denote the wage at i. Assume an iceberg trade friction τ to ship from one location to the other, i.e., to deliver one unit, $\tau \geq 1$ units must be shipped. There is no shipping cost to deliver goods domestically. The total costs to deliver to location 1, for the two lowest cost producers at 1, and the two lowest cost producers at 2, are then

$$\left\{ w_1 b^*_{\omega,1}, w_1 b^{**}_{\omega,1}, \tau w_2 b^*_{\omega,2}, \tau w_2 b^{**}_{\omega,2} \right\}.$$

Let $a^*_{\omega,1}$ and $a^{**}_{\omega,1}$ be the lowest and second lowest elements of this set. Analogously, we can define the lowest and second lowest costs $a^*_{\omega,2}$ and $a^{**}_{\omega,2}$ to deliver to country 2.

Firms compete in price in a Bertrand fashion, market by market. For a particular good ω at j, we can derive a demand curve $D_{\omega,j}(p)$ for ω at j as a function of price p, holding other prices and income as fixed. Assume the underlying utility function is such that the demand curve is continuously differentiable and strictly downward sloping. Suppose hypothetically that the most efficient producer of ω at j were a monopolist. The monopoly price $\bar{p}_{\omega,j}$ would solve

$$\begin{aligned} \bar{p}_{\omega,j} &= \arg\max_p \left(p - a^*_{\omega,j}\right) D_{\omega,j}(p) \\ &= \frac{\varepsilon_{\omega,j}}{\varepsilon_{\omega,j} - 1} a^*_{\omega,j}, \end{aligned}$$

i.e., the standard mark-up rule applies, where $\varepsilon_{\omega,j}$ is the elasticity of demand for ω evaluated at the monopoly price. (In the CES case, which is our main focus, $\varepsilon_{\omega,j} = \sigma$.)

In the equilibrium outcome of Bertrand competition, price will equal the minimum of the monopoly price and the marginal cost $a^{**}_{\omega,j}$ of the second lowest cost firm to deliver to j, i.e.,

$$p_{\omega,j} = \min\left(\bar{p}_{\omega,j}, a^{**}_{\omega,j}\right).$$

Define $c_{\omega,j}$ to be the variable cost share of price for good ω at j,

$$c_{\omega,j} = \frac{a^{*}_{\omega,j}}{p_{\omega,j}}$$

The inverse of $c_{\omega,j}$ is the mark-up. It is notationally convenient here to focus on $c_{\omega,j}$ rather than its inverse, the mark-up, which is the typical focus in the literature.

3 Welfare

To define our welfare decomposition, we first introduce additional notation. Let $\mathbf{p}_j = \{p_{\omega,j} : \omega \in [0,1]\}$ denote a particular set of prices at j. Having already assumed a homothetic utility function, without loss of generality we can further assume utility is homogeneous of degree one. Define the price index P_j to be the minimum cost at j to construct a consumption bundle delivering a unit level of utility, i.e.,

$$P_j \equiv \int_0^1 p_{\omega,j} \tilde{q}_{\omega,j} d\omega,$$

where $\tilde{\mathbf{q}}_j = \{\tilde{q}_{\omega,j} : \omega \in [0,1]\}$ is the expenditure-minimizing consumption bundle that solves

$$\tilde{\mathbf{q}}_j \equiv \arg\min_{\mathbf{q}} \int_0^1 p_{\omega,j} q_\omega d\omega,$$

subject to

$$\mathbf{U}(\mathbf{q}) = \mathbf{1}.$$

Let $s_{\omega,j}$ be the share of spending at j on good ω,

$$s_{\omega,j} = \frac{p_{\omega,j}\tilde{q}_{\omega,j}}{P_j}.$$

Next, let R_i be total revenue of firms located at i across all goods, including both domestic sales and exports. This will equal the total income at i, which is divided between labor and profits,

$$R_i = w_i L_i + \Pi_i.$$

Define Ec_i^{sell} to be the revenue-weighted mean share of variable cost in revenue across goods with *source* at location i. This equals

$$Ec_i^{sell} = \frac{w_i L_i}{R_i} = \frac{\int_{\{\omega:\ \chi_1^*(\omega)=i\}} c_{\omega,1} s_{\omega,1} R_1 d\omega + \int_{\{\omega:\ \chi_2^*(\omega)=i\}} c_{\omega,2} s_{\omega,2} R_2 d\omega}{\int_{\{\omega:\ \chi_1^*(\omega)=i\}} s_{\omega,1} R_1 d\omega + \int_{\{\omega:\ \chi_2^*(\omega)=i\}} s_{\omega,2} R_2 d\omega}, \tag{1}$$

where $\chi_j^*(\omega) \in \{1, 2\}$ denotes the source country for any particular good ω at destination j. To understand (1), observe that for a good ω sold at 1 with source at 1, $s_{\omega,1} R_1$ is the spending at 1, by definition of $s_{\omega,1}$. Then $c_{\omega,1} s_{\omega,1} R_1$ is the total variable cost for such goods, by definition of $c_{\omega,1}$ as the variable cost share. This accounts for the numerator of (1). The denominator sums sales to 1 and 2, from source i.

Define Ec_i^{buy} to be the revenue-weighted mean share of variable cost in revenue across goods with *destination* at i.

$$Ec_i^{buy} = \int_0^1 c_{\omega,i} s_{\omega,i} d\omega.$$

Finally, we introduce additional notation for what prices would be under marginal cost pricing. Let the price of an individual good at marginal cost be denoted

$$m_{\omega,i} = a^*_{\omega,i},$$

and let M_i be the price index at i when all goods are priced at marginal cost. Assume that productivity distributions are not too fat-tailed so that both price indices P and M are finite.[6]

We now present our welfare decomposition. We can write welfare at location i as

$$\begin{aligned} W_i^{Total} &= \frac{R_i}{P_i} = \frac{w_i L_i}{Ec_i^{sell}} \frac{1}{P_i} \\ &= w_i L_i \times \frac{1}{M_i} \times \frac{Ec_i^{buy}}{Ec_i^{sell}} \times \frac{M_i}{P_i \times Ec_i^{buy}} \end{aligned} \tag{2}$$

To get the first line, note that because utility is homogeneous of degree one, we can write utility at i as income divided by the price index. To obtain the rest of the first line, we

[6] For example, if we assume Fréchet structure of productivity draws as in BEJK, then, as in any Eaton-Kortum models, $\theta > \sigma - 1$ is required to guarantee finite price index (and hence equilibrium existence), where θ is the shape parameter of Fréchet. The larger the θ, the thinner the tail.

substitute in (1). The second line is a straightforward manipulation. We can write (2) as

$$W_i^{Total} = w_i L_i \times W_i^{Prod} \times \frac{Ec_i^{buy}}{Ec_i^{sell}} \times W_i^A \tag{3}$$

for the *productive efficiency* index W_i^{Prod},

$$W_i^{Prod} \equiv \frac{1}{M_i},$$

and the *allocative efficiency* index W_i^A,

$$W_i^A \equiv \frac{M_i}{P_i \times Ec_i^{buy}} = \frac{\int_0^1 m_{\omega,i} \tilde{q}_{\omega,i}^m d\omega}{\int_0^1 m_{\omega,i} \tilde{q}_{\omega,i} d\omega}. \tag{4}$$

The manipulation results in a decomposition (3) of welfare into four terms. Without loss of generality we will focus on the welfare of country 1, and we will set the wage at 1 to be the numeraire, $w_1 = 1$. To save notation, we will then leave country subscript i implicit (as $i = 1$ always). As the labor supply L_i will be fixed in the analysis, the first term in the welfare decomposition is a constant that we will henceforth ignore.

The productive efficiency index W^{Prod}, the second term of (3), is what the welfare index would be with no mark-up. It equals the inverse of M_i, which is the price index at marginal cost prices. The index varies when there is technical change determining the underlying levels of productivity. It also varies when trade cost declines, decreasing the cost of foreign firms to deliver goods locally. Terms of trade effects also show up in W^{Prod}, because a lower wage from a source country will raise the index.

The third and fourth terms depend upon mark-ups, or equivalently, the inverse of the mark-ups which are the cost shares. In ACR, and in the broader literature that it encompasses, trade has no effect on the distribution of mark-ups, and so it has no effect on the third and fourth terms. Thus in ACR, the welfare effects of trade operate entirely through the effects on the productive efficiency index W^{Prod}.

The third term is a "terms of trade" effect on mark-ups. Holding fixed the other components of welfare, total welfare is higher in country 1, if the cost share of price in the goods that it purchases tends to be high relative to the cost share of price in the goods that it sells. In a symmetric version of the model where the two countries are mirror images of each other, the "buy" cost share will equal to the "sell" cost share, and the third term will drop out.

The fourth term is the allocative efficiency index W^A, the main focus of this paper. The expression (4) presents two alternative ways to write W^A. The first way writes it as the ratio of the price indexes under marginal cost and actual pricing, divided by the mean variable cost share of price.

For the second way, consider two bundles, $\tilde{\mathbf{q}}^m$ and $\tilde{\mathbf{q}}$, that deliver one unit of utility. Bundle $\tilde{\mathbf{q}}^m$ minimizes expenditure at marginal cost pricing, while bundle $\tilde{\mathbf{q}}$ minimizes expenditure at actual prices. The second way of writing the allocative efficiency index is as the ratio of expenditure on these two bundles, evaluated at marginal cost pricing. It is immediate that $W^A \leq 1$, since bundle $\tilde{\mathbf{q}}^m$ minimizes expenditure at marginal cost pricing by definition. Suppose in the actual price vector, the mark-up is constant across all goods ω, i.e., that the ratio between any two actual prices is the same as what the ratio would be with marginal cost pricing. Then $\tilde{\mathbf{q}}^m = \tilde{\mathbf{q}}$ and $W^A = 1$. Otherwise $W^A < 1$.

4 Allocative Efficiency and Trade

When evaluating the effect of a change in the trade friction, it is convenient to take logs and conduct the analysis in elasticity terms,

$$\eta^{Total} \equiv \frac{d\ln W^{Total}}{d\ln \tau} = \eta^{Prod} + \eta^{c\_term\_trade} + \eta^A,$$

where η^k is the elasticity for component k. (The first term $w_1 L_1$ of (3) is a constant given the normalization $w_1 = 1$, so we ignore this term.) In particular, we define the allocative efficiency elasticity η^A to be

$$\eta^A \equiv \frac{d\ln W^A}{d\ln \tau}.$$

This section derives a formula for η^A that decomposes it into two components: the cost-change channel and the price-change channel. Note in the introduction, we referred to a *decrease* in τ, in order to discuss gains from trade. However, for expositional simplicity going forward, we refer to an increase in τ, because if we did otherwise, we would have to carry around an extra minus sign.

4.1 Decomposing η^A into the Cost-Change and Price-Change Channels

We need to introduce additional notation to allow us to distinguish various cases. When a good is being purchased from a domestic firm, there are three possibilities: the domestic firm may be limit pricing on another domestic firm, limit pricing on a foreign firm, or setting the unconstrained monopoly price. Analogously, there are three possibilities when a good is purchased from a foreign firm. We index these six different possibilities by the elements of $\{11, 12, \bar{1}, 21, 22, \bar{2}\}$, where "11" means a firm at 1 is limit pricing on a firm at 1 (i.e. domestic on domestic), "12" means a firm at 1 is limit pricing on a firm at 2, and $\bar{1}$ means a firm at 1 is setting the monopoly price. Analogously, "21," means a firm at 2 is limit pricing on a firm at 1, and so on. Let Ω_{11} be the set of products with event "11", Ω_{12} be the products with event "12," and so on.

For tractability, we assume CES utility in this and the next sections. For allocative efficiency W^A, only Ec^{buy} is involved, and hence we suppress superscript "buy" to save notation. Let Ec_Ω be the expected (or mean) value of the sales-weighted cost share c_ω, conditional on $\omega \in \Omega$, for some subset of goods Ω. Analogously, let $Ec_\Omega^{1-\sigma}$ be the sales-weighted expected value of $c^{1-\sigma}$ conditional on Ω.

Let $\mathbf{p}$ and $\mathbf{m}$ denote the vector of prices and marginal costs of the lowest cost firms supplying at 1, i.e., the element p_ω is the price of the lowest cost firm and m_ω is the marginal cost, including any trade friction incurred (and again $c_\omega = m_\omega/p_\omega$). Given $\mathbf{p}$ and $\mathbf{m}$, spending shares are determined by consumer choice, and we can calculate allocative efficiency W^A through equation (4). To understand how a change in the friction τ affects W^A, we need to first determine how a change in τ affects $\mathbf{p}$ and $\mathbf{m}$, and then how W^A changes with new values of $\mathbf{p}$ and $\mathbf{m}$.

In our decomposition, we separate out how a change in τ affects costs $\mathbf{m}$ (the *cost channel*) and how it affects prices $\mathbf{p}$ (the *price channel*). What an increase in τ does to $\mathbf{m}$ is straightforward: all foreign firms that are exporting to country 1 experience a proportionate increase in costs. Next consider what an increase in τ does to $\mathbf{p}$. A domestic firm that is limit pricing on a foreign firm's cost (event "12") will be able to increase price proportionately as τ increases. Analogously, a foreign firm limit pricing on another foreign firm (event "22") proportionately increases price. Finally, a foreign firm that is setting the monopoly price (event "$\bar{2}$") proportionately increase price (as price is a constant mark-up $\sigma/(\sigma-1)$ over cost). Denote the three events where price rises proportionately with τ as

$$\Omega^{p\uparrow} \equiv \Omega_{12} \cup \Omega_{22} \cup \Omega_{\bar{2}}. \tag{5}$$

Denote the relative wage of country 2 to country 1 as w. Let γ denote the proportion of increase of τw, i.e., let $\tau w = \gamma \tau^\circ w^\circ$, where τ° and w° are the initial trade friction and relative wage. A change in γ affects costs $\mathbf{m}$ such that

$$\begin{aligned} m_\omega(\gamma) &= \gamma m_\omega^\circ, \, \omega \in \Omega_2, \\ m_\omega(\gamma) &= m_\omega^\circ, \quad \omega \notin \Omega_2, \end{aligned} \tag{6}$$

where Ω_2 is the set of goods at 1 that are imported from country 2. Similarly, a change in γ affects prices $\mathbf{p}$ for the set $\Omega^{p\uparrow}$ such that

$$\begin{aligned} p_\omega(\gamma) &= \gamma p_\omega^\circ,\ \omega \in \Omega^{p\uparrow}, \\ p_\omega(\gamma) &= p_\omega^\circ, \quad \omega \notin \Omega^{p\uparrow}. \end{aligned} \tag{7}$$

We now present our main result that decomposes the overall effect of an increase in τ on allocative efficiency into two components, a cost channel and a price channel. Let $\eta^A_{cost} \equiv \left.\frac{d\ln W^A(1)}{d\gamma}\right|_{\mathbf{p}\text{ fixed}}$ denote the derivative of $\ln W^A(\gamma)$ evaluated at $\gamma = 1$ while keeping prices $\mathbf{p}$ fixed. Similarly, $\eta^A_{price} \equiv \left.\frac{d\ln W^A(1)}{d\gamma}\right|_{\mathbf{m}\text{ fixed}}$ is the derivative of $\ln W^A$ evaluated at $\gamma = 1$ while keeping costs $\mathbf{m}$ fixed. We have

Proposition 1

$$\begin{aligned} \eta^A &\equiv \frac{d\ln W^A}{d\ln\tau} \\ &= \left(\eta^A_{\text{cost}} + \eta^A_{\text{price}}\right)\frac{1}{w}\frac{d(\tau w)}{d\tau}, \end{aligned}$$

where

$$\eta^A_{\text{cost}} = s_2\left(\frac{Ec_2^{1-\sigma}}{Ec^{1-\sigma}} - \frac{Ec_2}{Ec}\right), \tag{8}$$

$$\eta^A_{\text{price}} = -\sigma s_{\Omega^{p\uparrow}}\left(1 - \frac{Ec_{\Omega^{p\uparrow}}}{Ec}\right), \tag{9}$$

and where the subset $\Omega^{p\uparrow}$ is defined by equation (5).

Proof. We sketch the proof for Cobb-Douglas ($\sigma = 1$) and symmetric countries. The appendix provides a formal proof for general σ and general asymmetry. Under Cobb-Douglas, we can write the log of allocative efficiency as

$$\ln W^A = \ln\left(\frac{M}{P \times Ec}\right) = [E\ln m - E\ln p] - \ln(Ec) = E\ln c - \ln(Ec), \tag{10}$$

where we use the price index formula $P = \exp(E\ln p)$ that applies for the Cobb-Douglas case. With symmetric countries, we can assume $w_1 = w_2 = 1$. We separate out the effects of changes in τ on costs, from effects through changes in prices. Following the notation of

(6) and keeping prices $\mathbf{p}$ fixed, we write

$$c_\omega(\gamma) = \begin{cases} \frac{\gamma m_\omega}{p_\omega} & \omega \in \Omega_2, \\ \frac{m_\omega}{p_\omega} & \omega \notin \Omega_2. \end{cases} \tag{11}$$

Note that $d\ln\tau = d\ln\gamma$. On account of the proportionate cost change for imports, we have

$$\begin{aligned} \frac{d\ln c_\omega}{d\ln\tau} &= \frac{d\ln c_\omega(1)}{d\gamma} = 1, \quad \frac{dc_\omega}{d\ln\tau} = \frac{dc_\omega(1)}{d\gamma} = c_\omega, \quad \omega \in \Omega_2, \\ \frac{d\ln c_\omega}{d\ln\tau} &= \frac{d\ln c_\omega(1)}{d\gamma} = 0, \qquad\qquad\qquad\qquad\qquad \omega \notin \Omega_2. \end{aligned}$$

Differentiating (10) yields

$$\eta^A_{cost} = \left.\frac{d\ln W^A(1)}{d\gamma}\right|_{\mathbf{p}\text{ fixed}} = \int_{\Omega_2} 1 \cdot d\omega - \frac{\int_{\Omega_2} c_\omega d\omega}{Ec} = s_2\left(1 - \frac{Ec_2}{Ec}\right). \tag{12}$$

Calculating the price channel η^A_{price} is analogous, but opposite in sign, since price is in the denominator of the cost share of price. In the proof above, we have switched the order of differentiation and integration. In the appendix, we show why this can be done. Note we haven't mentioned how a change in τ affects the sets Ω_2 and $\Omega^{p\uparrow}$, because the changes in these sets have no first-order effects. ■

For Cobb-Douglas, the cost-change channel (8) boils down to the particularly intuitive expression (12). An increase in τ proportionately increases costs for all imported goods. To understand the effects on allocative efficiency, holding prices fixed, we only need to compare the mean cost share Ec_2 on imported goods, which all get a cost increase, with the overall cost share Ec. In particular, if Ec_2 is greater than Ec, the effect of higher τ on allocative efficiency through this channel is strictly negative. In this case, higher τ increases the mark-up discrepancy between the imported and domestic goods, exacerbating the distortion. Formula (12) also depends upon the spending share s_2 on imports, and it is intuitive that this should matter. Next note that for the general σ case, formula (8) for the cost-change channel is similar to what it is in the Cobb-Douglas case, with an additional term that takes into account that spending shares across goods are not constant, outside of the Cobb-Douglas case. Finally, the intuition for the price-change channel is similar to the intuition for the cost-change channel.

We make two comments about Proposition 1. The first is about the additional term $\frac{1}{w}d(\tau w)/d\tau$ in the decomposition in Proposition 1 taking into account that τ can affect the relative wage between countries. With symmetric countries, the relative wage $w = 1$, and this adjustment drops out (i.e., it equals one as a multiplicative term). More generally, we expect the sign of $d(\tau w)/d\tau$ to be strictly positive, i.e. an increase in the friction increases relative cost in the foreign country, even taking into account any potentially offsetting effect on the relative wage. Hence, the sign of η^A is the same as $\eta^A_{cost} + \eta^A_{price}$. In what follows we fill focus on the sign of $\eta^A_{cost} + \eta^A_{price}$, and its decomposition into the cost-change channel η^A_{cost} and price-change channel η^A_{price}.

The second comment concerns how it might be possible to estimate the cost and price channels in empirical applications. Suppose we have product-level data on mark-ups and sales volumes, where imported and domestic products are separately classified. This is a stringent data requirement, but one that is potentially attainable, as access to micro data sets has expanded, and data sets from different countries are combined. If we also have knowledge of the σ parameter (a common parameter needed in trade analyses), the cost-change channel given by equation (8) can be directly calculated. (It uses two moments of the data.) Calculating the price-change channel (9) is more daunting, because we need to know which firms would raise price if the friction increases. Perhaps data on how prices change after a tariff change would be of help here. In lieu of obtaining a point estimate of the price-change channel, in applications it may be possible to use the formula to bound the price-chance channel. Combining this bound with a point estimate on the cost-change channel will produce a bound on the overall effect on allocative efficiency.

4.2 Signing the Cost-Change and Price-Change Channels

In this subsection we discuss the signs of the two components. We begin by examining two limiting cases where we find that both the cost-change and price-change channels are zero. We then discuss why in general the channels are nonzero away from the limiting cases.

The first limiting case is the BEJK model which imposes a particular Fréchet structure on the distribution of productivities across firms. In particular, BEJK assume that the distribution of two top productivities each country i follows

$$\Pr\left[X_{1i} \leq x_1, X_{2i} \leq x_2\right] = \left[1 + T_i\left(x_2^{-\theta} - x_1^{-\theta}\right)\right] e^{-T_i x_2^{-\theta}} \qquad \text{for } 0 \leq x_2 \leq x_1. \tag{13}$$

As de Blas and Russ (2012) show, in a Bertrand oligopoly model like the one here, the distribution above is in fact the limiting distribution, when the number of firms goes to infinity, and firms draw their productivities from a fat-tailed distribution, such as the Pareto and Fréchet itself.[7] We obtain the following result about the allocative efficiency elasticity η^A and its components η^A_{cost} and η^A_{price} for this case:

[7] That is, a distribution that falls in the domain of attraction for Fréchet.

Proposition 2 *If firms draw their productivities from a fat-tailed distribution, in the limit where the number of firms goes to infinity, the overall* $\eta^A = 0$*, as well as the components,* $\eta^A_{\text{cost}} = 0$ *and* $\eta^A_{\text{price}} = 0$*.*

Proof. Under said conditions, (13) holds, and the limit model is BEJK. In BEJK, the equilibrium share of variable cost in revenue is invariant to source country, and in particular, is the same for imports as it is overall, i.e., $Ec_2 = Ec$. It is easily verified that $Ec^{1-\sigma} = Ec_2^{1-\sigma}$, using similar proof to that in BEJK. By (8), $\eta^A_{cost} = 0$. Furthermore, in BEJK, the overall distribution of mark-ups is invariant to τ, implying that $\eta^A = 0$. The decomposition in Proposition 1, and $\eta^A = \eta^A_{cost} = 0$, immediately imply that $\eta^A_{price} = 0$. ■

A key result of BEJK is that the mark-up for foreign firms are the same as for domestic firms. Foreign firms have to pay a trade friction that domestic firms avoid. Everything else the same, this would drive up a foreign firm's cost and lower the mark-up relative to domestic firms. However, a second consideration is a selection process over productivity. On average, foreign firms need to be more productive to overcome the trade friction disadvantage. In the BEJK Fréchet structure, the two forces exactly counterbalance, implying that cost share for imports Ec_2 is identical to the overall share Ec, which means $\eta^A_{\text{cost}} = 0$. It turns out that mark-ups on the set of goods $\Omega^{p\uparrow}$ where price increases in τ are also the same as the overall average, and so $\eta^A_{\text{price}} = 0$, as well.

We next consider a second limiting case where countries are symmetric and the friction is small.

Proposition 3 *If countries are symmetric then at* $\tau = 1$*,*

$$\eta^A = \eta^A_{cost} = \eta^A_{price} = 0,$$

while

$$\eta^{Prod} < 0.$$

Proof. With symmetry and $\tau = 1$, it is immediate that the distribution of the cost share of price is the same regardless of whether a good originates at country 1 or country 2. This implies $Ec_2 = Ec$, and $Ec_2^{1-\sigma} = Ec^{1-\sigma}$, and thus $\eta^A_{\text{cost}} = 0$. Next observe that from symmetry and $\tau = 1$, that $Ec_{12} = Ec_{21}$, and $s_{12} = s_{21}$. This implies $s_{\Omega^{p\uparrow}} = s_2$ and $Ec_{\Omega^{p\uparrow}} = Ec_2$, and thus $\eta^A_{\text{price}} = 0$. Next,

$$\begin{aligned}\eta^{\text{Prod}} &\equiv \frac{d\ln W^{\text{Prod}}}{d\ln \tau} = -\frac{d\ln M}{d\ln \tau} \\ &= -s_2^m,\end{aligned}$$

where s_2^m is the import share of spending at marginal cost prices. The first line above follows from the definition of W^{Prod}, and the second line is a straightforward application of the envelope theorem. Note we use the fact that on account of symmetry, $w = 1$ and hence $d(\tau w)/d\tau = 1$. The result that $\eta^{\text{Prod}} < 0$ then follows because $s_2^m = \frac{1}{2}$ at the limit of $\tau = 1$, on account of symmetry. ■

Proposition 3 says that adding a small friction to trade between symmetric countries has no first-order effect on allocative efficiency. To see the intuition, observe that at the limit with no trade frictions, foreign firms are just like domestic firms, and in particular, have the same distribution of cost shares, implying that the cost-change term is zero. Analogously, the price-change term is zero. While there is no first-order effect on allocative efficiency, there is a first-order effect on productive efficiency at the limit. Thus, in this part of the parameter space, the issue of allocative efficiency is negligible relative to the productivity effects of τ.[8] In other words, with symmetry and small frictions, to a first approximation is it safe to abstract from the issue of how trade frictions affect allocative efficiency.

We turn now to more general environments away from these limiting cases. We begin with a discussion of the price-change term. To focus the discussion, we consider an environment in which there are no imports in equilibrium, in which case the cost-change term (8) is zero. (This is because $s_2 = 0$; we also note that $\eta^{\text{Prod}} = 0$ for this case.) Specifically, assume there are two sectors, A and B. In sector A there is one firm for each good in each country, and assume that productivity is identical for the two firms. The domestic firm will limit price on the foreign firm, and the cost share in sector A equals $c_A = \frac{1}{\tau}$.[9] For sector B, we consider two possibilities. In the first case there are two firms in each country with equal productivity. With Bertrand competition, price equals marginal cost, and $c_B = 1$. In the second case there is a single firm in each country, and we assume the sector B good is nontradable, so the single firm is a monopolist. Assuming elastic demand for this case, $\sigma > 1$, the cost share in sector B is $c_B = (\sigma - 1)/\sigma$. In either case, price increases with τ for the sector A goods, and doesn't change for the sector B goods. It is convenient to rewrite the price-change channel in equation (9) as

$$\eta^A_{price} = -\sigma s_A s_B \left(\frac{Ec_B - Ec_A}{Ec} \right).$$

In the first case where sector B is competitive and $Ec_B = 1$, it follows that $\eta^A_{price} < 0$.

[8] This point is contingent on τ being a true resource friction as opposed to a tariff for which welfare weight is put on tariff revenue collections. In this case, following the usual logic, the first-order effect on W^{Prod} will be offset by a change in tariff revenue.

[9] Assume $\sigma < \tau/(\tau - 1)$. Otherwise domestic firms set the simple monopoly price and increases in τ are irrelevant.

Raising τ allows the sector A firms to lower the cost share of price even further below the competitive case (where it equals one), increasing distortions. In the second case where sector B is monopoly, and $Ec_B < Ec_A$, it follows that $\eta^A_{price} > 0$. Raising τ makes the sector A firms more similar to the monopoly firms in sector B. As the mark-ups in the two sectors get closer together, allocative efficiency increases. This discussion makes clear that while the price-change component is zero in the limiting cases considered in Propositions 2 and 3, in general the effect can go either way, depending on how mark-ups for the goods getting price increases compare with mark-ups for goods with unchanged prices.

In terms of the previous literature, we can interpret the price-change channel as capturing the "pro-competitive" effect of trade, as changes in τ yield changes in price, through Bertrand competition. Our discussion then shows how the pro-competitive effect of trade has an ambiguous effect on allocative efficiency.

We next turn to the cost-change component. We focus on the symmetric country case. In the literature it is standard to consider models where firms take *i.i.d.* productivity draws from either a log-normal or Pareto distribution. We follow that approach here, and consider both distributions. For the Pareto, the functional form is

$$F(x) = 1 - x^{-\theta},\ x \geq 1,$$

for productivity $x = \frac{1}{b}$ equal to the inverse of the labor requirement b. The parameter θ is the shape parameter, where the larger is θ, the more similar the draws. (There is also a scaling parameter that we normalize to one.) For the log-normal, we assume that $\log(x)$ (for $x > 0$) is distributed normal with unit mean and standard deviation $\frac{1}{\theta}$. For the numerical example we set $\theta = 5$ for the Pareto and $\theta = 3.87$ for the log-normal, which holds the coefficient of variation constant across the two cases. (Our qualitative results are similar for alternative values of the coefficient of variation.)

Table 1 presents numerical results for a variety of cases. In each example, the number of firms per good is the same across all goods, and we consider the case of 2 firms per good, 4 firms per good, and 6 firms per good, with the firms equally divided across the two countries. We consider various levels of σ. In the table, we have fixed $\tau = 1.5$ throughout. (We obtain qualitatively similar results for other values of τ, away from the limit of $\tau > 1$.) For each parameter set, the table reports in the last three columns the allocative efficiency elasticity η^A, and its decomposition into η^A_{cost} and η^A_{price}. For reference, the table also reports the import share s_2 and the productive efficiency elasticity η^{Prod}.

Table 1 provides another illustration of the point that η^A_{price} can be negative or positive. At high values of σ, η^A_{price} tends to be positive. For intuition, consider that when σ is high, a relatively large fraction of firms set the interior monopoly price. In such cases, when τ

increases, those domestic firms limit pricing on foreign firms are able to raise mark-ups closer to the monopoly mark-ups most other firms are setting, raising allocative efficiency. In the opposite case when σ is low, in particular $\sigma \leq 1$, all firms limit price, and η^A_{price} is strictly negative.

Next note that $\eta^A_{\text{cost}} \leq 0$ in each of the cases illustrated in Table 1. For log-normal, η^A_{cost} is negative throughout all range of parameters. For Pareto, η^A_{cost} is negative when there are two firms for each good (thus one in each country) and 0 when there are more than two firms. We have used numerical analysis to verify this pattern holds more broadly in the symmetric country model with each firm drawing from the Pareto or log-normal, and each good having the same number of firms.[10] Our numerical analysis has also looked at the uniform distribution and found that it behaves similarly to log-normal in terms of η^A_{cost}. The key point is that in this class of models, the cost share of price tends to be higher for imported goods. The direct effect of the friction on a higher cost share for imports is not completely offset by the selection effect of higher productivity, in contrast to BEJK, where these forces exactly counterbalance. Thus $Ec_2 > Ec$. Also, if $\sigma > 1$, we can expect $Ec_2^{1-\sigma} \leq Ec^{1-\sigma}$, which together with formula (8) imply that $\eta^A_{\text{cost}} < 0$. If $\sigma < 1$, the $Ec_2^{1-\sigma}/Ec^{1-\sigma}$ term goes the other way. However, in the numerical analysis, we find the net effect remains negative, $\eta^A_{\text{cost}} \leq 0$.

When η^A_{price} is also negative, it reinforces the negative effect of η^A_{cost}, and the combined effect $\eta^A = \eta^A_{\text{price}} + \eta^A_{\text{cost}}$ can be large in absolute value, and similar in magnitude to the productivity effect η^{Prod}. When η^A_{price} is positive, there are cases where it more than offsets the negative effect of η^A_{cost}, resulting in combined effect η^A that is positive. For example, this is true with 2 firms, log-normal draws, and $\sigma = 4$, where the overall allocative efficiency elasticity is .035. In this case, a higher trade friction *raises* allocative efficiency.

Our last point about Table 1 concerns the comparison between the log-normal and Pareto cases. When there are two firms for each good (again one in each country), the log-normal and Pareto cases are qualitatively the same. However, under the Pareto, when there two or more firms per good in each country, η^A_{cost} and η^A_{price} are both virtually zero. The log-normal is very different; while η^A_{cost} and η^A_{price} do shrink in magnitude as we add more firms, they do not collapse to zero. Recall from the above discussion that if we take the Pareto model with n draws, and make n large, the model goes to the Fréchet structure of BEJK. Therefore, with the Pareto in the limiting case of large n, the effects are zero (Proposition 3). It is interesting to see that as far as allocative efficiency is concerned (and specifically how τ relates to it), the Pareto model begins to approximate the limiting case with as little as two draws in each country.

[10] The appendix provides details about our numerical analysis.

The sharp differences in implications between the Pareto and log-normal, and the frequent use of both distributions in the literature, motivate the next section where we explore this issue further.

5 Allocative Efficiency with the Pareto and Log-normal

In this section we explore how allocative efficiency changes as the market becomes large and previously separated economies become integrated. In particular, we consider changes from $\tau = \infty$ where countries are in autarky, to $\tau = 1$, where different countries are completely integrated into one economy. Let there be n_1 firms for each good in country 1, and n_2 firms for each good in country 2. In autarky, only n_1 firms compete to sell in country 1. After integration, $n = n_1 + n_2$ compete to sell in country 1. Hence, integration is equivalent to an increase in the number of firms. In this section we ask: How does allocative efficiency W^A vary as we increase the number of firms and integrate economies? And how does the answer depend upon whether firms draw from Pareto or log-normal?

For a given product ω, let x^*_ω and x^{**}_ω be the first and second highest of n independent productivity draws. Let $r_\omega \equiv x^{**}_\omega / x^*_\omega$ be the ratio of the second highest to highest. It is well known that under the Pareto, the distribution of the ratio r_ω is invariant to the number of draws.[11] The Pareto maintains its relative shape, as we push out into the tail. For our next result, we need a slightly different statement of this property, where we condition on the second highest level of productivity x^{**}, as well as the number of draws n. Formally, let $R(r|n, x^{**})$ be the cumulative probability of this ratio conditional on n and x^{**}. The Pareto has the following property.

Lemma 1 *Under the Pareto, $R(r|n, x^{**}) = r^\theta$, $r \in [0, 1]$, that is, the distribution of the ratio of second to first best does not depend upon the number of draws or what the second best is.*

Proof. Observe that

$$R\left(r|x^{**}, n\right) = \frac{\int_{\frac{x^{**}}{r}}^{\infty} g\left(x^*, x^{**}\right) dx^*}{\int_{x^{**}}^{\infty} g\left(x^*, x^{**}\right) dx},$$

where g denotes the joint density of top two order statistics from a Pareto distribution, and the result is immediate by plugging in this density function. ∎

The cost share satisfies $c_\omega = r_\omega$, if $\sigma \leq 1$ and $c_\omega = \max\left\{r_\omega, \frac{\sigma-1}{\sigma}\right\}$ for $\sigma > 1$. Given Lemma 1, the distribution of mark-ups is independent of n and x^{**} and is a truncated Pareto, analogous to BEJK. Using this fact we can show

[11] Malika and Trudela (1982), for example, characterizes the distribution of the ratio of all order-statistics of the Pareto. See also the discussion in de Blas and Russ (2012) .

Proposition 4 *Under the Pareto, allocative efficiency is a constant* $W^A < 1$ *that does not vary with the number of firms* n. *Therefore, a trade opening that is equivalent to an increase in the number of firms has no effect on allocative efficiency.*

Proof. See the appendix. ■

The key step of the proof is that since the distribution of cost shares is invariant to n, the actual price index P and the price index M under marginal cost pricing decrease by the same proportion, as n increases. >From formula (4) for W^A, it is immediate that W^A remains constant. We note there is allocative inefficiency for any given value of n, i.e. $W^A < 1$. The point is that its level does not change as economies integrate and n increases.

The outcome is quite different for the log-normal.

Proposition 5 *If the distribution of the ratio* $r = x^{**}/x^{*}$, *given second best* x^{**} *and the number of firms* n, *is degenerate at* $r = 1$ *in the limit as* $n \to \infty$, *then*

$$\lim_{n \to \infty} W^A = 1.$$

Under the log-normal, the above condition is satisfied. Thus, under the log-normal, while $W^A < 1$ *for any finite* n, *allocative inefficiency is eliminated in the limit as the market becomes large.*

Proof. See the appendix. ■

For intuition, consider a case where firms draw from a bounded productivity distribution. When there are many draws, the highest and second highest will both tend to be close to the upper bound, and the ratio will be close to one. Limit prices will be close to the competitive level, with the cost shares of price close to one, where first-best allocative efficiency is achieved. This notion of compression at the top, when markets are large and there are many firms, is consistent with the discussion in Syverson (2004) and Combes et al. (2012). Proposition 5 shows that even though the log-normal is unbounded, the right tail is sufficiently thin enough that the implication for allocative efficiency in the limit is the same as for the bounded case. In contrast, the Pareto is very different. The tail is sufficiently fat that even with many draws, there is no sense that the first and second highest get compressed together. In fact, the relationship doesn't change at all.

6 Conclusion

We examine how trade affects a measure of allocative efficiency, deriving formulas for two components of welfare change, the *cost-change channel* and the *price-change channel*. Both channels are negligible if (i) frictions are negligible and countries are symmetric or (ii) firms draw productivity from Pareto-like distributions. More generally, trade affects allocative efficiency. In symmetric models, the cost-change channel leads to gains in allocative efficiency, that reinforce the standard production efficiency gains from trade. When demand tends to be inelastic, the price-change channel leads to additional gains in allocative efficiency. However, if demand is elastic, the price-change channel goes the other way.

We also consider a limit economy where we increase the number of competitors for each good, and find a sharp difference between what happens when firms draw from a Pareto-like distribution compared to when the distribution has less of a fat tail, like the log-normal. In the former case, distortions in allocative efficiency never go away, as the market gets big. In the later, the economy converges to first-best allocative efficiency. The result highlights the important role functional form assumptions can play, in quantitative models of the gains from trade.

Appendix

Proof of Proposition 1

Observe that

$$\begin{aligned}\eta^A &\equiv \left.\frac{d\ln W^A}{d\ln\tau}\right|_{\tau=\tau^\circ,w=w^\circ} = \left(\frac{d\ln W^A}{\frac{d(\gamma\tau^\circ w^\circ)}{\gamma\tau^\circ w^\circ}}\frac{1}{\gamma w^\circ}\frac{d(\tau w)}{d\tau}\right)_{\gamma=1}\\ &= \left.\frac{d\ln W^A(1)}{d\gamma}\frac{1}{w^\circ}\frac{d(\tau w)}{d\tau}\right|_{\tau=\tau^\circ,w=w^\circ}.\end{aligned}$$

Note that for CES utility,

$$s_\omega = \frac{p_\omega^{1-\sigma}}{P^{1-\sigma}}, \tag{14}$$

$$\tilde{q}_\omega = \frac{p_\omega^{-\sigma}}{P^{-\sigma}} = \frac{p_\omega^{-\sigma}}{\left(\int_0^1 p_\omega^{1-\sigma}d\omega\right)^{\frac{\sigma}{\sigma-1}}}. \tag{15}$$

Using (4) and (15), we can express $\ln W^A$ and $\frac{d\ln W^A(1)}{d\gamma}$ as

$$\begin{aligned}\ln W^A &= -\frac{1}{\sigma-1}\ln\int_0^1 m_\omega^{1-\sigma}d\omega - \ln\int_0^1 m_\omega p_\omega^{-\sigma}d\omega + \frac{\sigma}{\sigma-1}\ln\int_0^1 p_\omega^{1-\sigma}d\omega,\\ \frac{d\ln W^A(1)}{d\gamma} &= -\frac{1}{\sigma-1}\frac{\frac{d}{d\gamma}\int_0^1 m_\omega^{1-\sigma}d\omega}{\int_0^1 m_\omega^{1-\sigma}d\omega} - \frac{\frac{d}{d\gamma}\int_0^1 m_\omega p_\omega^{-\sigma}d\omega}{\int_0^1 m_\omega p_\omega^{-\sigma}d\omega} + \frac{\sigma}{\sigma-1}\frac{\frac{d}{d\gamma}\int_0^1 p_\omega^{1-\sigma}d\omega}{\int_0^1 p_\omega^{1-\sigma}d\omega}.\end{aligned}$$

We will later show that

$$\frac{d}{d\gamma}\int_0^1 m_\omega^{1-\sigma} d\omega = \int_0^1 \frac{d}{d\gamma} m_\omega^{1-\sigma} d\omega, \tag{16}$$

$$\frac{d}{d\gamma}\int_0^1 p_\omega^{1-\sigma} d\omega = \int_0^1 \frac{d}{d\gamma} p_\omega^{1-\sigma} d\omega, \tag{17}$$

$$\frac{d}{d\gamma}\int_0^1 m_\omega p_\omega^{-\sigma} d\omega = \int_0^1 \frac{d}{d\gamma} m_\omega p_\omega^{-\sigma} d\omega. \tag{18}$$

Recall that $\int_0^1 m_\omega^{1-\sigma} d\omega = M^{1-\sigma}$, and that $\int_0^1 p_\omega^{1-\sigma} d\omega = P^{1-\sigma}$. By the above, we obtain

$$\begin{aligned}
\frac{d\ln W^A(1)}{d\gamma} &= -\frac{1}{\sigma-1}\frac{\int_0^1 \frac{d}{d\gamma} m_\omega^{1-\sigma} d\omega}{\int_0^1 m_\omega^{1-\sigma} d\omega} - \frac{\int_0^1 \frac{d}{d\gamma} m_\omega p_\omega^{-\sigma} d\omega}{\int_0^1 m_\omega p_\omega^{-\sigma} d\omega} + \frac{\sigma}{\sigma-1}\frac{\int_0^1 \frac{d}{d\gamma} p_\omega^{1-\sigma} d\omega}{\int_0^1 p_\omega^{1-\sigma} d\omega} \\
&= \left\{\frac{\int_{\Omega_2} m_\omega^{1-\sigma} d\omega}{\int_0^1 m_\omega^{1-\sigma} d\omega} - \frac{\int_{\Omega_2} m_\omega p_\omega^{-\sigma} d\omega}{\int_0^1 m_\omega p_\omega^{-\sigma} d\omega}\right\} + \left\{\sigma\frac{\int_{\Omega^{p\uparrow}} m_\omega p_\omega^{-\sigma} d\omega}{\int_0^1 m_\omega p_\omega^{-\sigma} d\omega} - \sigma\frac{\int_{\Omega^{p\uparrow}} p_\omega^{1-\sigma} d\omega}{\int_0^1 p_\omega^{1-\sigma} d\omega}\right\} \\
&= \left\{\frac{\int_{\Omega_2} p_\omega^{1-\sigma} d\omega}{P^{1-\sigma}}\left(\frac{\int_{\Omega_2} c_\omega^{1-\sigma}\frac{p_\omega^{1-\sigma}}{\int_\Omega p_\omega^{1-\sigma} d\omega} d\omega}{\int_0^1 c_\omega^{1-\sigma} s_\omega d\omega} - \frac{\int_{\Omega_2} c_\omega \frac{p_\omega^{1-\sigma}}{\int_{\Omega_2} p_\omega^{1-\sigma} d\omega} d\omega}{\int_0^1 c_\omega s_\omega d\omega}\right)\right\} \\
&\quad + \left\{\sigma\frac{\int_{\Omega^{p\uparrow}} p_\omega^{1-\sigma} d\omega}{P^{1-\sigma}}\frac{\int_{\Omega^{p\uparrow}} c_\omega \frac{p_\omega^{1-\sigma}}{\int_{\Omega^{p\uparrow}} p_\omega^{1-\sigma} d\omega} d\omega}{\int_0^1 c_\omega \frac{p_\omega^{1-\sigma}}{P^{1-\sigma}} d\omega} - \sigma\frac{\int_{\Omega^{p\uparrow}} p_\omega^{1-\sigma} d\omega}{P^{1-\sigma}}\right\} \\
&= \left\{s_2\left(\frac{Ec_\Omega^{1-\sigma}}{Ec^{1-\sigma}} - \frac{Ec_\Omega}{Ec}\right)\right\} + \left\{-\sigma s_{\Omega^{p\uparrow}}\left(1 - \frac{Ec_{\Omega^{p\uparrow}}}{Ec}\right)\right\}.
\end{aligned}$$

In each line above, the first large bracket is the cost channel and the second the price channel.

We are left to show that (16) to (18) hold. For this purpose, we first recall the Lebesgue dominated convergence theorem, which states that if $\{f_n\}$ is a sequence of Lebesgue-integrable functions on an interval I which converges almost everywhere on I to a limit function f, then $\int_I f = \lim_{n\to\infty}\int f_n$, provided that there exists a nonnegative, Lebesgue-integrable function g such that, for all n, $|f_n(x)| \leq g(x)$ almost everywhere on I.[12]

We start with proving (16). Define

$$f(\omega) \equiv \frac{d}{d\gamma} m_\omega^{1-\sigma}.$$

Consider a sequence $\{\gamma_n\}$ such that $\lim_{n\to\infty}\gamma_n = 1$. Define a sequence of functions $\{f_n\}$

[12] Theorem 10.27, Apostol, T. M., *Mathematical Analysis* (second edition), page 270.

such that

$$f_n(\omega) = \frac{m_{\gamma_n,\omega}^{1-\sigma} - m_\omega^{1-\sigma}}{\gamma_n - 1},$$

where $m_{\gamma,\omega}$ is m_ω associated with γ. Note that $m_\omega \equiv m_{1,\omega}$.

Since

$$\frac{d}{d\gamma}\int_0^1 m_\omega^{1-\sigma} d\omega = \lim_{n\to\infty} \frac{\int_0^1 m_{\gamma_n,\omega}^{1-\sigma} d\omega - \int_0^1 m_\omega^{1-\sigma} d\omega}{\gamma_n - 1} = \lim_{n\to\infty} \int_0^1 f_n(\omega)\, d\omega$$

and

$$\int_0^1 \frac{d}{d\gamma} m_\omega^{1-\sigma} d\omega = \int_0^1 \lim_{n\to\infty} f_n(\omega)\, d\omega = \int_0^1 f(\omega)\, d\omega,$$

we complete the proof by showing that the Lebesgue dominated convergence theorem holds. Note that $f_n(\omega)$ is integrable by assumption that the price index under marginal cost pricing M is finite. Also note that $f_n(\omega) \to f(\omega)$ by construction. Hence, we need to find an nonnegative, integrable g such that $|f_n(x)| \leq g(x)$ for all n.

Since $m_\omega = \min\left\{b_{1,\omega}^*, \tau w b_{2,\omega}^*\right\}$, when γ changes from 1, the proportion of change in $m_{\gamma,\omega}$ is less than or equal to γ. Specifically, when $\gamma_n > 1$, $m_{\gamma_n,\omega} \leq \gamma_n m_\omega$, and when $\gamma_n < 1$, $m_{\gamma_n,\omega} \geq \gamma_n m_\omega$. Hence,

$$|f_n(\omega)| = \left|\frac{m_{\gamma_n,\omega}^{1-\sigma} - m_\omega^{1-\sigma}}{\gamma_n - 1}\right| \leq \left|\frac{\gamma_n^{1-\sigma} - 1}{\gamma_n - 1}\right| m_\omega^{1-\sigma}. \tag{19}$$

By the mean value theorem, for each n there exists $\hat{\gamma}_n$ between 1 and γ_n such that

$$\frac{(\gamma_n)^{1-\sigma} - 1}{\gamma_n - 1} = \left.\frac{d\gamma^{1-\sigma}}{d\gamma}\right|_{\gamma=\hat{\gamma}_n} = (1-\sigma)\,\hat{\gamma}_n^{-\sigma}.$$

Thus, we obtain

$$|f_n(\omega)| \leq |1-\sigma| \times \sup_n \hat{\gamma}_n^{-\sigma} \times m_\omega^{1-\sigma} \equiv g(\omega).$$

That $g(\omega)$ is integrable follows from the fact that M is finite.

For the proof of (17), simply observe that for when $\gamma_n > 1$, $p_{\gamma_n,\omega} \leq \gamma_n p_\omega$, and when $\gamma_n < 1$, $p_{\gamma_n,\omega} \geq \gamma_n p_\omega$, and thus the key inequality (19) holds with m replaced with p. The rest of the proof is the same with m replaced with p (and M replaced with P).

For the proof of (18), define $f(\omega) = \frac{d}{d\gamma}\left(m_\omega p_\omega^{-\sigma}\right)$ and $f_n(\omega) = \frac{m_{\gamma_n,\omega} p_{\gamma_n,\omega}^{-\sigma} - m_\omega p_\omega^{-\sigma}}{\gamma_n - 1}$. Because

$m_\omega \le p_\omega$,

$$\left|f_n(\omega)\right| = \left|\frac{m_{\gamma_n,\omega}\left(p_{\gamma_n,\omega}^{-\sigma} - p_\omega^{-\sigma}\right) + \left(m_{\gamma_n,\omega} - m_\omega\right)p_\omega^{-\sigma}}{\gamma_n - 1}\right| \le \left(\left|\frac{\gamma_n^{-\sigma}-1}{\gamma_n - 1}\right|\gamma_n + 1\right)p_\omega^{1-\sigma}.$$

The rest of the proof is similar.

Proof of Proposition 4

As the distribution of r is not degenerate, $W^A < 1$. We want to show that W^A is independent of the number of firms n. Given n, denote the joint density of the first and second highest productivities as $h_n(x^*, x^{**})$, and the marginal density of the first and second highest productivity as $h_{1n}(x^*)$ and $h_{2n}(x^{**})$, respectively. Let $v(r)$ be an arbitrary continuous function of $r = x^{**}/x^*$. Then,

$$\int_{x^*} v\left(\frac{x^{**}}{x^*}\right)\frac{h_n(x^*, x^{**})}{h_{2n}(x^{**})}dx^* = E_{x^*}\left(v\left(\frac{x^{**}}{x^*}\right)|x^{**}, n\right) = E_r\left(v(r)|x^{**}, n\right) = \int_r v(r)\, dR(r|n, x^{**}), \tag{20}$$

which is independent of x^{**} and n by Lemma 1.

For CES utility, $p(x^*, x^{**}) = \frac{1}{x^{**}}\min\left\{\frac{\sigma}{\sigma-1}\frac{x^{**}}{x^*}, 1\right\}$, and

$$P^{1-\sigma} = \int h_{2n}(x^{**})\left(\frac{1}{x^{**}}\right)^{1-\sigma}\left[\int\left(\min\left\{\frac{\sigma}{\sigma-1}\frac{x^{**}}{x^*}, 1\right\}\right)^{1-\sigma}\frac{h_n(x^*, x^{**})}{h_{2n}(x^{**})}dx^*\right]dx^{**} \tag{21}$$

$$= \left[\int\left(\min\left\{\frac{\sigma}{\sigma-1}\frac{x^{**}}{x^*}, 1\right\}\right)^{1-\sigma}\frac{h_n(x^*, x^{**})}{h_{2n}(x^{**})}dx^*\right]\int\left(\frac{1}{x^{**}}\right)^{1-\sigma}h_{2n}(x^{**})\,dx^{**}, \tag{22}$$

where the last equality follows from the fact that the bracket term is independent of x^{**}, as (20) explains. Note that $c = \max\left\{\frac{x^{**}}{x^*}, \frac{\sigma-1}{\sigma}\right\}$. Using the definition of Ec, (14), (22), and Lemma 1,

$$\begin{aligned} & Ec \\ = \ & \frac{\int\left(\frac{1}{x^{**}}\right)^{1-\sigma}h_{2n}(x^{**})\,dx^{**}}{P^{1-\sigma}}\int \max\left\{\frac{x^{**}}{x^*}, \frac{\sigma-1}{\sigma}\right\}\left(\min\left\{\frac{\sigma}{\sigma-1}\frac{x^{**}}{x^*}, 1\right\}\right)^{1-\sigma}\frac{h_n(x^*, x^{**})}{h_{2n}(x^{**})}dx^* \\ = \ & \frac{\int \max\left\{\frac{x^{**}}{x^*}, \frac{\sigma-1}{\sigma}\right\}\left(\min\left\{\frac{\sigma}{\sigma-1}\frac{x^{**}}{x^*}, 1\right\}\right)^{1-\sigma}\frac{h_n(x^*, x^{**})}{h_{2n}(x^{**})}dx^*}{\int\left(\min\left\{\frac{\sigma}{\sigma-1}\frac{x^{**}}{x^*}, 1\right\}\right)^{1-\sigma}\frac{h_n(x^*, x^{**})}{h_{2n}(x^{**})}dx^*}, \end{aligned}$$

which is independent of x^{**} and n. As $W^A \equiv \frac{M}{P\times Ec}$, we are done if M/P is also independent of n. From (22),

$$\left(\frac{M}{P}\right)^{1-\sigma}=\frac{\int m_{\omega}^{1-\sigma}d\omega}{\int p_{\omega}^{1-\sigma}d\omega}=\frac{\int\left(\frac{1}{x^{*}}\right)^{1-\sigma}h_{1n}\left(x^{*}\right)dx^{*}}{\left[\int\left(\min\left\{\frac{\sigma}{\sigma-1}\frac{x^{**}}{x^{*}},1\right\}\right)^{1-\sigma}\frac{h_{n}(x^{*},x^{**})}{h_{2n}(x^{**})}dx^{*}\right]\int\left(\frac{1}{x^{**}}\right)^{1-\sigma}h_{2n}\left(x^{**}\right)dx^{**}}.$$

As mentioned, the bracket term in the denominator is independent of x^{**} and n. By Lemma 1,

$$\begin{aligned}\frac{\int\left(\frac{1}{x^{*}}\right)^{1-\sigma}h_{1n}\left(x^{*}\right)dx^{*}}{\int\left(\frac{1}{x^{**}}\right)^{1-\sigma}h_{2n}\left(x^{**}\right)dx^{**}} &= \frac{\int\left[\int\left(\frac{x^{**}}{x^{*}}\right)^{1-\sigma}\frac{h_{n}(x^{*},x^{**})}{h_{2n}(x^{**})}dx^{*}\right]\left(\frac{1}{x^{**}}\right)^{1-\sigma}h_{2n}\left(x^{**}\right)dx^{**}}{\int\left(\frac{1}{x^{**}}\right)^{1-\sigma}h_{2n}\left(x^{**}\right)dx^{**}}\\ &= \int\left(\frac{x^{**}}{x^{*}}\right)^{1-\sigma}\frac{h_{n}\left(x^{*},x^{**}\right)}{h_{2n}\left(x^{**}\right)}dx^{*},\end{aligned}$$

which is independent of x^{**} and n. Thus, M/P is, indeed, independent of n.

Proof of Proposition 5

We first show that if $R\left(r|n,x^{**}\right)$ is degenerate at $r=1$ as $n\rightarrow\infty$, that is, if $R\left(r|n,x^{**}\right)\rightarrow 0$ for any $r<1$ as $n\rightarrow\infty$, then $W^{A}\rightarrow 1$. Under the said condition, $Ec=E\left(\max\left\{\frac{1}{r},\frac{\sigma-1}{\sigma}\right\}\right)\rightarrow 1$ as $n\rightarrow\infty$. As $W^{A}\equiv\frac{M}{P\times Ec}$, $W^{A}\rightarrow 1$ if $M/P\rightarrow 1$ when $n\rightarrow\infty$.

Use the same notation of h_n, h_{1n}, h_{2n}, and v as those in the proof of Proposition 4. Note that $\int_{x^{*}}v\left(\frac{x^{**}}{x^{*}}\right)\frac{h_{n}(x^{*},x^{**})}{h_{2n}(x^{**})}dx^{*}=\int_{r}v\left(r\right)R\left(r|n,x^{**}\right)dr$ approaches $v\left(1\right)$ when $n\rightarrow\infty$. Hence, using (21), we have

$$\begin{aligned}\left(\frac{M}{P}\right)^{1-\sigma} &= \frac{\int\left(\frac{1}{x^{*}}\right)^{1-\sigma}h_{1n}\left(x^{*}\right)dx^{*}}{\int\left(\frac{1}{x^{**}}\right)^{1-\sigma}h_{2n}\left(x^{**}\right)\left[\int\left(\min\left\{\frac{\sigma}{\sigma-1}\frac{x^{**}}{x^{*}},1\right\}\right)^{1-\sigma}\frac{h_{n}(x^{*},x^{**})}{h_{2n}(x^{**})}dx^{*}\right]dx^{**}} \\ &\rightarrow \frac{\int\left(\frac{1}{x^{*}}\right)^{1-\sigma}h_{1n}\left(x^{*}\right)dx^{*}}{\int\left(\frac{1}{x^{**}}\right)^{1-\sigma}h_{2n}\left(x^{**}\right)dx^{**}}\rightarrow 1,\end{aligned} \tag{23}$$

as $r\rightarrow 1$ in distribution implies that distributions of x^{*} and x^{**} are arbitrarily close to each other when n is arbitrarily large.[13]

[13] To see this, note that the difference between two distributions is

$$\Pr\left[X^{**}<x|n\right]-\Pr\left[X^{*}<x|n\right]=\Pr\left[rX^{*}<x|n\right]-\Pr\left[X^{*}<x|n\right]\rightarrow 0,$$

since r degenerates to 1 as $n\rightarrow\infty$.

Now, we turn to the log-normal case. With the log-normal density,

$$\begin{aligned} R\left(r|n,x^{**}\right) &= \frac{\Pr[X^{*}\geq \frac{x^{**}}{r}, X^{**}=x^{**}]}{\Pr[X^{**}=x^{**}]} = \frac{F^{n-2}\left(x^{**}\right)\left[1-F\left(\frac{x^{**}}{r}\right)\right]}{F^{n-2}\left(x^{**}\right)\left[1-F\left(x^{**}\right)\right]} \\ &= \frac{1-F\left(\frac{x^{**}}{r}\right)}{1-F\left(x^{**}\right)} = \frac{\int_{\frac{x^{**}}{r}}^{\infty}\frac{1}{x}e^{-\frac{[\ln(x)-\mu]^2}{2\sigma^2}}dx}{\int_{x^{**}}^{\infty}\frac{1}{x}e^{-\frac{[\ln(x)-\mu]^2}{2\sigma^2}}dx}. \end{aligned}$$

Using L'Hopital's rule and noting that $x^{**}\to\infty$ when $n\to\infty$, for any $r<1$, we have

$$\lim_{x^{**}\to\infty} R\left(r|n,x^{**}\right) = \lim_{x^{**}\to\infty}\exp\left(\frac{\left[2\ln\left(x^{**}\right)-2\mu-\ln\left(r\right)\right]\ln\left(r\right)}{2\sigma^2}\right) = 0.$$

Numerical Analysis of sign of $\eta^A_{\mathbf{cost}}$

We calculate η^A_{cost} numerically for the following parameter values. For the distribution of a productivity draw x, we use log-normal, uniform, and the Pareto distributions. A log-normal distribution is characterized by mean and standard deviation of $\ln x$. The mean does not affect η^A_{cost}, thus we normalize it to 1. For the standard deviation we use the following values: 0.25, 0.5, 0.75, 1, 1.5, 2, 3, 5, and 8. A uniform distribution is characterized by lower and upper bound for x. Rescaling x does not change η^A_{cost}, thus we set the upper bound to 1 and vary the lower bound using the following values: 0 to 0.9 with 0.1 increment. A Pareto distribution is characterized by scale parameter and shape parameter. The scale parameter does not affect η^A_{cost}, thus we normalize it to 1. For the shape parameter, we use 0.25, 0.5, 0.75, 1, 1.5, 2, 3, 5, and 8. Beside the productivity draw distribution, we need to specify the values for trade friction τ, preference parameter σ, and the total number of firms (or potential entrants) for each product. For τ, we use 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.2, 2.4, 2.6, 2.8, 3, 3.5, and 4. For σ, we use 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0, 2.3, 2.6, 3, 3.5, 4, 5, 7, and 10. For the total number of firms (at both locations combined), we use 2,4,6,8,10, 14, and 20. The parameter grid generates 19,278 combinations for log-normal and for Pareto distributions and 21,420 combinations for uniform distribution.

We use 6 million draws to approximate the continuum of products featured in the model. The more number of products we use, the more precise the numerical approximation becomes. In order to see the sign of numerical errors, we also report simulation outcomes with smaller numbers of products: 10,000 to 3 million. As the number of products increases, the simulated values converge toward true values.

For Pareto, we distinguish the case of one firm at each location (two firms altogether) from the case where there are two or more firms in one location. For the log-normal and uniform distribution, we consider the entire range of firm counts. The numerical analysis summarized in Table A1 shows that the basic pattern revealed in Table 1 of the paper for a particular selection of model parameters holds throughout the wide range of parameters considered in the analysis. In particular, for log-normal and uniform for general firms, and the Pareto with one firm at one location, we have $\eta^{A}_{\text{cost}} \leq 0$. Note that in the simulations with finite draws we do get realizations where $\eta^{A}_{\text{cost}} > 0$. However, as shown in the table, the maximum realized value gets very close to zero as the number of draws increases. Also when we fix a cutoff value equal to .00001, we see that the fraction of deviations where the realized value goes above this small cutoff goes to zero as the number of draws increases.

For Pareto with two or more firms at each location, the pattern is different in Table 1. Rather than take a negative value, we see that η^{A}_{cost} is close to zero. Table A2 shows this pattern holds throughout the parameter range considered in the numerical analysis.

References

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

Arkolakis, C., A. Costinot, Dave Donaldson, and A. Rodriguez-Clare (2012), "The Elusive Pro-Competitive Effects of Trade," working paper, June 2012.

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

Behrens, Kristian and Yasusada Murata, (2012) "Trade, Competition, and Efficiency," *Journal of International Economics*, 87(1): 1-17.

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

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sébastien Roux (2012), "The Productivity Advantages of Large Cities: Distinguishing Agglomeration from Firm Selection," *Econometrica*, 80(6): 2543-2594.

de Blas, Beatriz , and Katheryn Russ (2012), "Understanding Markups in the Open Economy under Bertrand Competition," working paper, July 2012.

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

Edmond, Chris , Virgiliu Midrigan and Daniel Yi Xu (2012), "Competition, Markups, and the Gains from International Trade," NBER working paper 18041.

Grossman, Gene M. and Helpman, Elhanan (1991), "Quality Ladders in the Theory of Growth." *Review of Economic Studies*, 58(1): 43-61.

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

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

Malika, Henrick J. and Roger Trudela (1982), “Probability Density Function of Quotient of Order Statistics from the Pareto, Power and Weibull Distributions,” *Communications in Statistics* - Theory and Methods, 11(7): 801-814.

Melitz, Marc J. and G. Ottaviano (2008), “Market Size, Trade, and Productivity,” *Review of Economic Studies,* 75(1): 295-316.

Neary, J. Peter (2003), “Globalization and Market Structure,” *Journal of the European Economic Association*, 1: 245-271.

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

Peters, Michael (2012) “Heterogenous Mark-Ups and Endogenous Misallocation,” Yale University manuscript.

Restuccia, D. and R. Rogerson (2008), “Policy Distortions and Aggregate Productivity with Heterogeneous Plants,” *Review of Economic Dynamics*, 11: 707–720.

Robinson, Joan (1934), *The Economics of Imperfect Competition,* Macmillan.

Salvo, Alberto (2010), “Inferring Market Power under the Threat of Entry: The Case of the Brazilian Cement Industry,” *RAND Journal of Economics.* 41(2): 326-350.

Schmitz, Jr. James A. (2005), “What Determines Productivity? Lessons from the Dramatic Recovery of the U.S. and Canadian Iron Ore Industries Following Their Early 1980s Crisis,” *Journal of Political Economy*, 113(3): 582-625.

Syverson, C. (2004), “Market Structure and Productivity: A Concrete Example,” *Journal of Political Economy,* 112(6): 1181-1222.

Table 1
The Decomposition of the Allocative Efficiency Elasticity for Various Alternative Parameters

Panel A: LogNormal Distribution

Number of Firms	σ (elas.)	s_2 (import share)	η^{Prod}	η^{A}	Decomposition η^{A}_{cost}	Decomposition η^{A}_{price}
2	0.50	0.12	-0.14	-0.036	-0.018	-0.018
2	1.00	0.13	-0.13	-0.072	-0.036	-0.036
2	1.50	0.14	-0.12	-0.088	-0.054	-0.034
2	2.00	0.15	-0.12	-0.015	-0.063	0.048
2	4.00	0.11	-0.08	0.035	-0.038	0.074
4	0.50	0.09	-0.09	-0.009	-0.005	-0.004
4	1.00	0.09	-0.09	-0.018	-0.010	-0.008
4	1.50	0.09	-0.08	-0.025	-0.014	-0.012
4	2.00	0.09	-0.08	-0.026	-0.018	-0.008
4	4.00	0.07	-0.06	0.005	-0.019	0.023
6	0.50	0.07	-0.07	-0.005	-0.003	-0.002
6	1.00	0.07	-0.07	-0.009	-0.005	-0.004
6	1.50	0.07	-0.07	-0.014	-0.008	-0.006
6	2.00	0.07	-0.06	-0.016	-0.010	-0.005
6	4.00	0.06	-0.05	0.000	-0.013	0.013

Panel B: Pareto

Number of Firms	σ (elas.)	s_2 (import share)	η^{Prod}	η^{A}	Decomposition η^{A}_{cost}	Decomposition η^{A}_{price}
2	0.50	0.06	-0.06	-0.015	-0.008	-0.007
2	1.00	0.07	-0.07	-0.033	-0.017	-0.017
2	1.50	0.08	-0.07	-0.041	-0.026	-0.015
2	2.00	0.09	-0.07	-0.003	-0.035	0.032
2	4.00	0.11	-0.09	0.033	-0.025	0.058
4	0.50	0.08	-0.08	0.000	0.000	0.000
4	1.00	0.09	-0.09	0.000	0.000	0.000
4	1.50	0.09	-0.09	0.000	0.000	0.000
4	2.00	0.09	-0.09	0.000	0.000	0.000
4	4.00	0.10	-0.10	0.000	0.000	0.000
6	0.50	0.09	-0.09	0.000	0.000	0.000
6	1.00	0.09	-0.09	0.000	0.000	0.000
6	1.50	0.10	-0.10	0.000	0.000	0.000
6	2.00	0.10	-0.10	0.000	0.000	0.000
6	4.00	0.11	-0.11	0.000	0.000	0.000

The calculations set $\tau = 1.5$ throughout and $\theta=5$ for Pareto and $\theta=3.87$ for the lognormal.

Table A1. Numerical Analysis of η^{A}_{cost}

Number of Products	Count of Parameter Combinations	Min	Mean	Max	Count of $\eta^{A}_{\text{cost}} \geq 0.00001$	Share of $\eta^{A}_{\text{cost}} \geq 0.00001$
Panel A: LogNormal Distribution						
6,000,000	19,278	-1.26E-01	-7.45E-03	8.30E-14	0	0.00%
3,000,000	19,278	-1.27E-01	-7.72E-03	1.52E-08	0	0.00%
1,000,000	19,278	-1.27E-01	-7.33E-03	1.13E-05	1	0.01%
100,000	19,278	-1.26E-01	-7.22E-03	1.28E-03	71	0.37%
10,000	19,278	-1.29E-01	-7.27E-03	1.37E-02	378	1.96%
Panel B: Uniform						
6,000,000	21,420	-2.63E-02	-1.53E-03	4.76E-11	0	0.00%
3,000,000	21,420	-2.63E-02	-1.53E-03	1.51E-11	0	0.00%
1,000,000	21,420	-2.62E-02	-1.52E-03	6.58E-12	0	0.00%
100,000	21,420	-2.58E-02	-1.52E-03	7.56E-07	0	0.00%
10,000	21,420	-2.61E-02	-1.52E-03	5.29E-06	0	0.00%
Panel C: Pareto, with one firm in each location						
6,000,000	2,754	-9.35E-02	-1.27E-02	1.64E-11	0	0.00%
3,000,000	2,754	-1.29E-01	-2.42E-02	1.08E-11	0	0.00%
1,000,000	2,754	-9.32E-02	-1.18E-02	4.19E-12	0	0.00%
100,000	2,754	-9.45E-02	-1.09E-02	1.45E-03	5	0.18%
10,000	2,754	-9.18E-02	-9.69E-03	1.08E-03	11	0.40%

Table A2. Numerical Analysis of η^{A}_{cost}
Pareto distribution, with the number of firms at each location ≥ 2

Number of Products	Count of Parameter Combinations	Min	Mean	Max
6,000,000	16,524	-6.93E-04	-9.41E-06	2.56E-04
3,000,000	16,524	-1.40E-03	1.67E-05	9.63E-04
1,000,000	16,524	-9.98E-04	6.36E-05	1.34E-03
100,000	16,524	-2.27E-03	2.01E-05	3.55E-03
10,000	16,524	-8.40E-03	-1.97E-05	1.48E-02