

NBER WORKING PAPER SERIES

UNDERSTANDING MARKUPS IN THE OPEN ECONOMY UNDER BERTRAND
COMPETITION

Beatriz de Blas
Katheryn Russ

Working Paper 16587
<http://www.nber.org/papers/w16587>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2010

The authors thank seminar and conference participants at the 2010 American Economics Association winter meetings; the Banco do Portugal; the 2011 European Economics Association/Econometrics Society Meetings; the 2010 European Trade Study Group; DEGIT XV at Goethe University of Frankfurt; the Federal Reserve Banks of Dallas, Richmond, and St. Louis; the 2011 NBER International Trade and Investment Group spring meeting; the 2012 International Finance and Macroeconomics Summer Meeting; Princeton University; Stanford University; the 2012 West Coast Trade Workshop; and the 2012 Western Economics Association International meetings. They are indebted to Andrew Bernard, Ariel Burstein, Jeffrey Campbell, Robert Feenstra, Stefania Garetto, Yingyao Hu, Oleg Itskhoki, Konstantin Kucheryavyy, Karen Lewis, Anthony Landry, Marc Melitz, Virgiliu Midrigan, Kanda Naknoi, Maurice Obstfeld, David Richardson, Barbara Spencer, Jonathan Vogel, and Nick Zolas for detailed comments in early stages of the project. Beatriz de Blas acknowledges financial support from the ECO2008-04073 project of the Spanish MEC, and the CCG08-UAM/HUM-4438. This paper circulated in an earlier form under the title "Teams of Rivals: Endogenous Markups in a Ricardian World." The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2010 by Beatriz de Blas and Katheryn Russ. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Understanding Markups in the Open Economy under Bertrand Competition
Beatriz de Blas and Katheryn Russ
NBER Working Paper No. 16587
December 2010
JEL No. F0,F1,F4

ABSTRACT

The purpose of this paper is to understand the effects of endogenous markups and trade costs on the pricing behavior of exporters when firms are heterogeneous in productivity. Using new analytical distributions for markups under Bertrand competition, we uncover Ricardian patterns of export pricing that generate higher markups and export price volatility when industrialized countries sell to developing countries. These Ricardian patterns dissipate when developing countries move from bilateral to multilateral trade liberalization. The results arise from a form of price rigidity for exports that arises endogenously due to cut-throat competition, even though prices are otherwise perfectly flexible.

Beatriz de Blas
Dpto. A.E.: Teoría e Historia Económica
Universidad Autónoma de Madrid
C/ Francisco Tomás y Valiente, 5
28049 Madrid, Spain
beatriz.deblas@uam.es

Katheryn Russ
Department of Economics
University of California, Davis
One Shields Avenue
Davis, CA 95616
and NBER
knruss@ucdavis.edu

1 Introduction

The purpose of this paper is to understand the effects of endogenous markups and trade costs on the pricing behavior of exporters when firms are heterogeneous in productivity. Using new analytical distributions for markups under Bertrand competition, we uncover Ricardian patterns of export pricing that generate higher markups and export price volatility when industrialized countries sell to developing countries. These Ricardian patterns dissipate when developing countries move from bilateral to multilateral trade liberalization. The results arise from a form of price rigidity for exports that arises endogenously due to cut-throat competition, even though prices are otherwise perfectly flexible.

Diminishing pass-through and volatility in traded goods prices over the past several decades, as documented by (Bergin and Feenstra (2008), Corsetti, Dedola, and Leduc (2008), and Gust, Leduc, and Vigfusson (2010)), present a puzzle for models of heterogeneous firms with endogenous markups. Atkeson and Burstein (2007 and 2008) demonstrate for the first time that trade costs reduce the market power of heterogeneous exporters, limiting the degree to which they can pass through shocks to marginal costs to consumers and thereby dampening export price volatility relative to other aggregates. Yet trade costs have been falling at the same time as pass-through and volatility in traded goods prices. This presents a conceptual challenge for current models of trade with heterogeneous firms. Both Edmond, Midrigan, and Xu (2012) and Arkolakis, Costinot, Donaldson, and Rodriguez-Clare (2012a) show for a wide class of models with imperfect competition and endogenous markups that trade liberalization can actually increase the market power of exporters as they internalize the drop in trade costs by charging higher markups, given that the liberalization takes place between symmetric countries or between only two countries with independent productivity distributions. Under these circumstances, theory would predict increased pass-through and volatility in an age of falling trade costs while the data suggests the opposite.

In this paper, we demonstrate with concise functional forms exactly when

trade produces pro-competitive effect under Bertrand competition with heterogeneous firms that would result in less ability for exporters to pass through shocks to marginal cost into the prices of their goods. We argue that developing countries must charge lower markups when exporting to technologically more advanced countries, which reduces their ability to pass through shocks to marginal costs into prices. Further, accession into multilateral trade agreements can reduce pass-through and volatility in the prices of goods exported to developing countries from industrialized countries. Bilateral trade liberalization allows exporters from a favored industrialized country to compete principally with domestic producers in the destination market, where the average foreign exporter would have a technological edge. However, *multilateral* trade liberalization in this case forces exporters from different industrialized countries to compete more directly with each other, so they have less ability to garner market power in the destination market and thus less ability to pass on cost shocks into prices.

We characterize the bilateral result, where exporters respond to trade liberalization by charging higher markups, resulting in increased volatility, as an anti-competitive effect. The decrease in markups charged by exporters in the favored nation when the destination market expands to multilateral liberalization is thus a pro-competitive effect. Pro-competitive effects also emerge among domestic producers when their native market opens to trade under any circumstances, as new competition from foreign rivals forces them to reduce markups.¹ The mechanisms behind pro-competitive effects have heretofore been a black box in international trade and macroeconomics— are they driven by technological differences, by domestic contestability, by geographic frictions? We show that they are a product of all three and show their relative contributions.

Understanding this balance is crucial in open economy models, as it gener-

¹There is ample empirical evidence from the trade literature that pro-competitive effects from trade exist and can be large. Among these are Levinsohn (1993), Harrison (1994), Roberts and Supina (1996), Bottasso and Sembenelli (2001), Novy (2010), and Feenstra and Weinstein (2010), Edmond, Midrigan, and Xu (2012), and de Loecker, Goldberg, Khandelwal, and Pavcnik (2012).

ates a type of price inertia distinct from the conventional forms of price rigidity assumed in standard macroeconomic models. Atkeson and Burstein (2007 and 2008) and Alessandria (2009) show that the special behavior of the endogenous markups combines with costly trade to generate a lower volatility of the terms of trade relative to the real exchange rate.² This rigidity can have other important macroeconomic manifestations, as in Obstfeld’s (2009) argument that incomplete pass-through in export prices following a nominal exchange rate appreciation broke the Harrod-Balassa-Samuelson effect during the appreciation of the Japanese yen from 1985-95. Recent evidence also suggests a Ricardian influence over the degree of pass-through: Bergin and Fenstra (2008) and Gust, Leduc, and Vigfusson (2010) show that pass-through is particularly limited for exporters from developing countries to the U.S. market, while Frankel, Parsley, and Wei (2012) document that pass-through is greater for exports to developing countries than to industrialized countries. We believe that the two sets of findings— lower markups under trade and relatively rigid export prices, with the degree of rigidity varying in a Ricardian pattern— are related through trade costs and Bertrand competition by heterogeneous firms.

Our model involves a finite number of firms competing within each industry. The most efficient firm in the industry ultimately becomes the sole supplier of that particular good, but only because it beats back its competitors by underselling them: it cannot charge a price higher than the marginal cost of its next best rival. We argue that trade costs make firms’ prices more likely to be bound by their next best rival when selling overseas compared to their home market. We show analytically that this cutthroat competition can generate reduced markups under trade, pricing to market, and imperfect pass-through by causing price rigidity even when prices are not set in advance due to menu costs or other constraints.

Bernard, Eaton, Jensen, and Kortum (2003, hereafter BEJK) pioneered the use of this form of price competition with heterogeneous firms and trade

²In particular, papers such as Baxter and Landry (2010), Berman, Martin, and Mayer (2012), Fitzgerald and Haller (2010), Schoenle (2010), Gopinath, Itshoki, and Rigobon (2010), and Gopinath and Itshoki (2010) show that firms often price to market and do not fully pass on changes in marginal costs and exchange rates to foreign buyers.

almost a decade ago. It is not clear whether they assumed the number of rivals to be infinitely large or Poisson distributed.³ In contrast, we generalize their approach by endogenizing and explicitly focusing on a finite number of entrants that compete in each industry so that markups are sensitive to market size and structure, in line with findings in the closed-economy literature such as those by Campbell and Hopenhayn (2005). When the number of competitors is infinitely large, our distribution of markups converges to that of BEJK. In particular, in this limiting case, there are no anti- or pro-competitive effects from trade, allowing Bertrand competition to fit within a gravity framework.

Because we explicitly include a finite number of rivals, we see that the distribution of markups is directly affected by the number of firms competing to be the low-cost supplier in the closed economy or in the case when both of the best potential suppliers to a market are in the same source country.⁴ One can conceptualize the number of rivals as an exogenous policy parameter, as in the numerical analyses by Atkeson and Burstein (2007 and 2008) and de Blas and Russ (2011), or endogenize it using a free entry condition as we do below, and as Holmes, Hsu, and Li (2011) have recently done, as well. The key is that unlike models using Chamberlinian monopolistic competition or limit pricing, such as Melitz (2003) and Melitz and Ottaviano (2008), the degree of entry embodied in the number of rivals changes the shape of the entire distribution of markups, costs, and firm size. Although the number of rivals appears similar to a scale parameter representing technological advancement in our model, it has a more complex effect on the distribution of markups through the joint distribution of the first and second order statistic for unit cost. The number of rivals itself is not critical, unless it becomes so large as to obscure the

³Their results nest within ours and within those of Holmes, Hsu, and Li (2011) if the number of rivals is infinite. However, in the statistics literature, it is known that a Poisson-distributed number of draws from the Weibull can be integrated out to focus on other parameters (Raftery 1987). Finally, the number of rivals also drops out of the distribution of markups in two special cases discussed below if the underlying distribution of efficiency levels is Pareto, rather than Fréchet.

⁴Claessens and Laeven (2004) and de Blas and Russ (2011) refer to this as “contestability,” a phenomenon with roots in the industrial organization literature. See in particular Baumol, Panzar and Willig (1982), related to the Panzar-Rossi measure of contestability used by Claessens and Laeven (2004).

effects of trade costs or technological differences on market structure. Most importantly, having a finite number of rivals allows other elements of market structure— including trade costs and differences in technology and wages across countries— to influence the distribution of markups. We use these new, closed-form distributions to draw out Ricardian effects on markups, as well as the impact of multi- versus bilateral trade liberalization on price volatility that have been observed empirically but not integrated into the theoretical literature on pass-through.

The paper is organized as follows. In Section 2, we discuss the relationship of our model with previous empirical and theoretical studies. Section 3 presents a simple closed economy model with analytical solutions for the distribution of markups and prices which include the number of rivals. We show the relationship between entry and the aggregate domestic price level. Section 4 considers the implications of trade in goods for these distributions given asymmetric trading partners. In Section 5, we show that our distribution of markups implies a higher degree of price adjustment (higher price volatility) in response to these shocks for domestic versus export sales. Section 6 concludes.

2 Related literature

The basic insight motivating our construction of the model stems from the numerical simulations of Atkeson and Burstein (2007 and 2008), Garetto (2012), and de Blas and Russ (2011). All of these numerical studies start with a Fréchet or lognormal distribution of firm efficiency levels, then build on BEJK by computing markups under Bertrand competition (also Cournot in the case of Atkeson and Burstein (2008)). Collectively, they note that the size of the markup shrinks under trade and that trade costs make firms less able to pass on shocks to marginal costs by raising export prices.⁵ They also note another

⁵A number of recent papers make advances using nominal rigidities to achieve the stickier export prices behind the low relative volatility of the terms of trade. These include Corsetti, Dedola, and Leduc (2008), Gopinath and Itshoki (2010), and Schoenle (2010). Alessandria (2009) and Drodz and Nosal (2012) use consumer-oriented frictions in distribution and marketing. In Section 5, we discuss nominal rigidities as complementary to the real rigidities

key feature— that the number of competitors within each industry, either the number of domestic competitors or foreign trading partners, affects both the size of the average markup and the degree or frequency of pass-through.

We introduce a finite, cutthroat group of rivals for each industry within the BEJK framework that generates the markup and pricing behavior observed both empirically and numerically in previous studies. We assume an independent Fréchet distribution of efficiency draws for each country, following Eaton and Kortum (2002). Our model achieves tractable analytical solutions for the distribution of markups under autarky and trade with a finite number of firms competing to supply the market. In short, we build a standard Ricardian model of trade with a fully specified distribution of markups that is consistent with observed domestic and export pricing behavior. The assumption in our model that rivals to the best firm in each industry are latent is not necessary to achieve these distributions (see de Blas and Russ 2012 for a search framework with similar numerical results). We maintain this assumption to nest within existing literature and keep the model as simple as possible while we illustrate their implications for firm-level and aggregate price adjustment.

Entry in this Ricardian model does not affect the number of goods produced, but rather the number of firms competing to be the low-cost supplier of a particular good. “Competing” in this sense means drawing an efficiency parameter from an identical distribution and being ready to jump into production if a chance arises to undersell an active firm. The most efficient firm will have the lowest cost— the first order statistic for costs in the industry⁶— and become the only active supplier. An increase in the number of firms that compete to be the low-cost supplier of a good changes the shape of the entire distribution of marginal costs and markups. Since the circulation of this paper, Holmes, Hsu, and Li (2011) have used a similar approach to analyze agglomeration and allocative efficiency, as well as to draw important new insights regarding the gains from trade. Zolas (2011) has shown that the number

arising from market structure.

⁶The first order statistic is the first (lowest) cost in a random sample arranged in ascending order of magnitude (see David and Nagaraja (2003)).

of potential competitors influences a firm's choice of where to file patents. Finally Peters (2011) has developed a related framework with dynamic entry to analyze growth in total factor productivity within a closed economy.

Edmond, Midrigan, and Xu (2012, hereafter EMX) in an important contribution extend our understanding of and rigorously quantify the degree of gains from trade due to pro-competitive effects based on the model of Cournot (quantity-based) competition in Atkeson and Burstein (2008). They use Taiwanese data to calibrate the model and estimate the gains, and push the frontier of Cournot modeling with heterogeneous firms forward by linking the pro-competitive effects of trade to changes in the dispersion of markups, as well as the average markup, within a two-country framework, with the second country representing the rest of the world. They run several experiments, but the one closest to our treatment of the Bertrand setting of price competition is their experiment where the distribution of productivity draws across countries is independent. In this case, they find that pro-competitive effects of trade are actually negative in their simulations. The logic behind their result is similar to the insight drawn by Arkolakis, Costinot, Donaldson, and Rodríguez-Clare (2012a) regarding other models of trade with endogenous markups: highly efficient exporters exploit reductions in trade costs by absorbing them as higher markups. Our multi-country setting provides a distinct and nontrivial result. In short, this anti-competitive effect is only likely if trade costs are not lowered for competing trade partners—thus, in unilateral, versus multilateral trade liberalization. Second, EMX attribute the result largely to the fatness of the tails in the distribution of efficiency levels, but it is difficult to see the exact relationship between trade barriers and the parameters of the distribution in a numerical setting. We show the precise likelihood that lower trade costs will be absorbed into the markup as a function of this dispersion by deriving an explicit distribution of markups. In addition, our model suggests that we can verify empirically whether anti- or pro-competitive effects arise after liberalization even when it is not possible to measure markups by measuring whether the degree of pass-through and import price volatility increases (anti) or decreases (pro).

3 Autarky

The heart of the model lies in the production of intermediate goods by heterogeneous firms. For simplicity, we assume that producers of the final good are perfectly competitive and assemble the intermediate goods, with no additional capital or labor necessary. The continuum of intermediate goods j spans the fixed interval $[0,1]$. The assembly process uses a technology involving a constant elasticity of substitution across inputs, with aggregate output given by

$$Y = \left[\int_0^1 Y(j)^{\frac{\sigma-1}{\sigma}} dj \right]^{\frac{\sigma}{\sigma-1}} .$$

Output of the final good is purchased for immediate use by consumers or as an input into the production of intermediate goods. When used as a production input, it is fully expended—no inventories are carried over into future periods. We consider each intermediate input j as representing a different industry and assume that the price elasticity of substitution between output from different industries σ is greater than one. The demand for an individual input is downward sloping in its price, $Y(j) = \left(\frac{P(j)}{P}\right)^{-\sigma} Y$, and the aggregate price level P is given by

$$P = \left[\int_0^1 P(j)^{1-\sigma} dj \right]^{\frac{1}{1-\sigma}} . \tag{1}$$

Each producer of an intermediate good draws an efficiency parameter z from a cumulative distribution $F(z)$ with positive support over the interval $(0,\infty]$. Eaton and Kortum (2009, Chapter 4) describe a process whereby over time, $F(z)$ can emerge as a frontier distribution representing the efficiency levels associated with the best surviving ideas available to produce a particular good j . Being the distribution of the best surviving ideas, $F(z)$ naturally takes on an extreme value form and under mild assumptions, it can be characterized by a Fréchet distribution.⁷ Thus, we assume that an endogenous number of

⁷In particular, EK suppose that each period a group of new ideas emerges with the quality of these ideas distributed as Pareto. Over time, the distribution of the best (most

firms r each draw an efficiency parameter from a distribution given by

$$F(z) = e^{-Tz^{-\theta}}.$$

We assume that $T > 0$ and also that the shape parameter, θ , is positive. Only the most efficient firm with efficiency level $Z_1(j)$ in any industry supplies the market. This efficiency parameter increases the level of output a firm produces from one unit of a composite input Q :

$$Y(j) = Z_1(j)Q(j).$$

Marginal cost for this most efficient firm, $C_1(j)$, is inversely related to the efficiency parameter,

$$C_1(j) = \frac{wd}{Z_1(j)},$$

which accounts for both the cost of the composite input, w , and any frictions involved in sending intermediate goods to the assemblers of the final good, $d \geq 1$. We assume that both labor and the final good are used in the production of intermediate goods with constant cost shares: $w = \omega^\beta P^{1-\beta}$, ω being the labor wage rate and p the cost of a bundle of intermediate goods. The cost parameter drawn by any firm hoping to produce good j is distributed

$$G(c) = 1 - e^{-T(wd)^{-\theta} c^\theta}.$$

Given that some number of rivals r draw an efficiency parameter hoping to be

efficient) idea surviving from each period then becomes Fréchet, also known as an inverse Weibull (Pawlas and Szynal, 2000). Costs are inversely related to efficiency levels, so costs in this case are Weibull distributed, as in Eaton and Kortum (2002).

the low-cost supplier of industry j , the distribution of the lowest cost $C_1(j)$ is⁸

$$G_1(c_1) = 1 - e^{-rT(wd)^{-\theta}c_1^\theta}. \quad (2)$$

We assume that $d = 1$ under autarky in this section and for domestic sales in the open economy in Section 3. The limiting distribution— $G_1(c_1)$ given an infinitely large sample r — is well defined for any positive, finite T . (See Castillo (1988, p.116) and Castillo, Hadi, Balakrishnan, and Sarabia (2005, p.207) for the derivation of the limiting distribution of maxima drawn from a Fréchet distribution.)

3.1 The distribution of markups

Let $C_2(j)$ represent the unit cost of the second-best competitor in industry j , who sits inactive but ready to begin production instantly should the opportunity arise. Given the CES assembly technology for the final good, the lowest-cost firm producing good j would like to set a price using what we call the “unconstrained” markup, where marginal cost equals marginal revenue—the CES markup $\bar{m} \equiv \frac{\sigma}{\sigma-1} > 1$. However, if charging the CES markup results in a price that exceeds the marginal cost of the second-best competitor waiting in the wings, the lowest-cost supplier may find itself undersold. In short, no firm can charge a price that exceeds the unit cost of its next best rival. The low-cost supplier in each industry j takes the prices of the low-cost supplier in every other industry as given. The markup for industry j is then

$$M(j) = \min \left\{ \frac{C_2(j)}{C_1(j)}, \bar{m} \right\}.$$

⁸See Rinne (2009), p.237 for derivation. The assumption that BEJK use regarding the number of entrants to simplify their framework is not stated in the published or working-paper version of the text. Their results can be replicated by assuming that it is infinitely large. (See Holmes, Hsu, and Li (2011) for details regarding the derivation of the composite distribution of efficiency draws, or our discussion below with regard to markups.) Alternatively, the number would drop out of the analysis if one assumed that the number of firms competing in any industry is a random variable with a Poisson distribution and integrating over its domain. In contrast, we preserve the number of rivals in the following analysis.

With this formula for the markup, we compute the expected output-weighted price for any good j in several steps. First, note that the price for good j , $P(j)$, is given by

$$P(j) = \begin{cases} C_2(j) & \text{for } \frac{C_2(j)}{C_1(j)} \leq \bar{m} \\ \bar{m}C_1(j) & \text{for } \frac{C_2(j)}{C_1(j)} \geq \bar{m} \end{cases}$$

Thus, the pricing rule depends not only upon the distribution of the first and second order statistic of the marginal costs, but also upon the distribution of the ratio of the two order statistics. In Appendix A we use a straightforward Jacobian transformation on a result from Malik and Trudel (1982) to obtain the distribution of $\frac{C_2(j)}{C_1(j)}$, which is the distribution of the markup before imposing the unconstrained markup from the CES bundling of intermediate goods. Assuming that the frontier distribution of efficiency parameters is identical for every industry j , the probability density of the ratio $\frac{C_2(j)}{C_1(j)}$ is given by

$$h(m) = \begin{cases} \frac{r(r-1)\theta m^{-(\theta+1)}}{[(r-1)+m^{-\theta}]^2} & \text{for } 1 \leq m < \bar{m} \\ \int_{\bar{m}}^{\infty} \frac{r(r-1)\theta m^{-(\theta+1)}}{[(r-1)+m^{-\theta}]^2} dm & \text{for } m = \bar{m} \\ 0 & \text{for } m > \bar{m} \end{cases}, \quad (3)$$

At the unconstrained markup, there is a mass point, which we show in Figure 1.

Like the distribution of markups given in BEJK, this distribution is statistically independent of $C_1(j)$ and $C_2(j)$. In fact, for very large r , we have $\lim_{r \rightarrow \infty} h(m) = \theta m^{-\theta-1}$ for $1 \leq m \leq \bar{m}$, which is a Pareto density for markups identical to the one in BEJK. With the CES bundling technology, firms will never set a markup greater than \bar{m} , creating a mass point in the density at \bar{m} , since all cases where $\frac{C_2(j)}{C_1(j)}$ exceeds \bar{m} are assigned a value of \bar{m} . The probability of charging the unconstrained markup is simply

$$\Pr \left[\frac{C_2(j)}{C_1(j)} \geq \bar{m} \right] = \int_{\bar{m}}^{\infty} h(m) dm = \frac{r}{1 + (r-1)\bar{m}^{\theta}}. \quad (4)$$

Note that as \bar{m} goes from its own upperbound of ∞ (for $\sigma = 1$) to its lowerbound of 1 (for $\sigma \rightarrow \infty$), this probability moves monotonically from 0 to 1, so it is a well behaved cumulative distribution function over the range of possible markups.

Proposition 1 *The average markup is decreasing in the number of rivals r under autarky.*

Proof. For any given value $1 \leq m' \leq \bar{m}$, the probability that $M(j) \geq \frac{C_2(j)}{C_1(j)}$ is greater than or equal to m' is decreasing in r :

$$\frac{\partial \Pr [M(j) \geq m']}{\partial r} = \frac{\partial \left(\int_{m'}^{\infty} \frac{r(r-1)\theta m^{-(\theta+1)}}{[(r-1)+m^{-\theta}]^2} dm \right)}{\partial r} = \frac{-[(m')^\theta - 1]}{[1 + (r-1)(m')^\theta]^2} < 0.$$

Equivalently, we can say that the distribution of markups when r is low first-order stochastically dominates the distribution of markups with a higher r . First-order stochastic dominance implies a higher expected value; therefore $E[M(j)]$ must be decreasing in r . ■

It follows from Proposition 1 that the number of rivals affects the size of the mass of firms charging the unconstrained markups, as stated in the following corollary.

Corollary 1 *In expectation, the fraction of firms charging the unconstrained markup is decreasing in the number of rivals r under autarky.*

Proof. The proof of Proposition 1, combined with equation (4) shows that the probability of $\frac{C_2(j)}{C_1(j)}$ being at least as large as \bar{m} is decreasing in the number of rivals. Markups are set equal to \bar{m} whenever $M(j)$ would be greater than \bar{m} without the restriction of the CES upperbound. Thus, in expectation, the fraction of firms charging the unconstrained markup is decreasing in the number of rivals. ■

As the number of rivals in an industry j increases, both the average markup and the probability that any firm charges the unconstrained markup falls—increased rivalry squeezes markups. Intuitively, the result emerges because,

on average, increasing the number of rivals in our order-statistic framework diminishes the difference between the costs of the two best potential suppliers. This is not the case for a Pareto distribution of firm efficiency levels, as shown in Appendix A. When firms draw from a Pareto distribution of efficiency levels, markups are again Pareto distributed as in BEJK (and in our special case above with many competitors), with no impact from the number of rivals. To reinterpret BEJK’s sports analogy in our setup: with the distribution of costs in equation (2), a competitor running second in a race will run even faster relative to the winner when there are more competitors behind him. However, with Pareto efficiency draws, no matter how many additional competitors trail behind in the race, each runner maintains both his speed and spacing *relative* to the person in front of him.⁹ In economic terms, the Fréchet distribution implies diminishing returns to technological growth through entry: as the number of rivals increases, there is a greater chance that additional rivals’ efficiency draws will fall within the existing production possibilities frontier than that they will expand it outward.

To illustrate our new distribution of markups, Figure 1 shows the restricted distribution of markups when r equals its minimum value of 2, versus 20, the number of rivals chosen by Atkeson and Burstein (2007) calibrated to match U.S. industry concentration. We use $\theta=3.6$ and $\sigma=3.79$, as estimated by BEJK. The fraction of firms charging the unconstrained markup falls drastically, from one-half to just over one-third. We will discuss the implications of this statistic for price rigidity in Section 5. For now, we use the distribution of markups to compute the aggregate price level.

⁹We believe the key difference is that the value of any outcome z enters the inverse of the hazard function linearly, which is not the case for the Fréchet used here or the lognormal used in Atkeson and Burstein (2008). Or more simply, the Pareto mean is linear in its minimum, which is also the case for the uniform distribution. Their density functions are flat or convex, rather than being strictly concave around the mode as in our model.

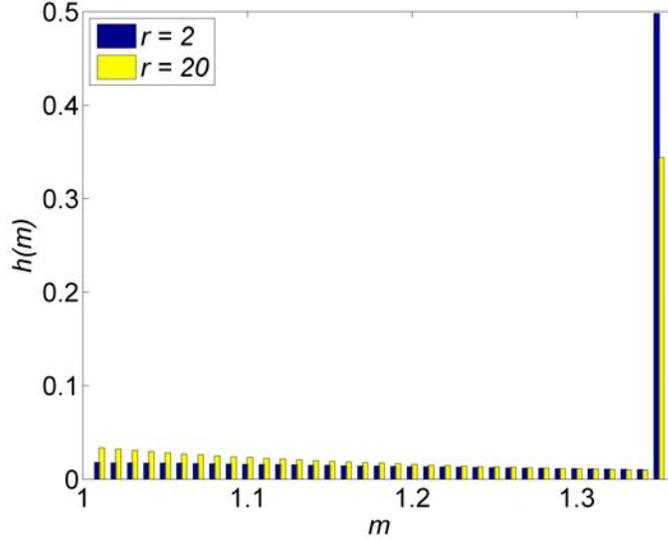


Figure 1: Increasing the number of rivals reduces markups

3.2 The distribution of prices

As shown in de Blas and Russ (2011), the joint distribution for the first and second order statistic also contains the number of rivals r :

$$g_{1,2}(c_1, c_2) = r(r-1) [\theta T w^{-\theta}]^2 c_1^{\theta-1} c_2^{\theta-1} e^{-T w^{-\theta} c_1^\theta} e^{-(r-1) T w^{-\theta} c_2^\theta}. \quad (5)$$

To find the marginal distribution for $C_1(j)$ ($C_2(j)$), one can integrate the joint distribution over values of c_2 (c_1).¹⁰ We find that increasing the number of rivals leads, on average, to lower costs in the industry. We compute the moment $1 - \sigma$, which appears in the formula for the aggregate price level Equation (1), for the first and second order statistics of marginal costs, so

¹⁰Integrating the joint distribution over c_2 from c_1 to ∞ , for instance, one obtains the marginal distribution $g_1(c_1)$ and sees immediately that it is equal to the first derivative of $G_1(c_1)$. To obtain the marginal for $C_2(j)$, one instead integrates over c_1 from zero to c_2 , as we do later for the open economy in Appendix E.

that we can use them below to construct the aggregate price level:

$$\begin{aligned} E[C_1(j)^{1-\sigma}] &= (rTw^{-\theta})^{\frac{\sigma-1}{\theta}} \Gamma\left(\frac{1-\sigma}{\theta} + 1\right), \\ E[C_2(j)^{1-\sigma}] &= (Tw^{-\theta})^{\frac{\sigma-1}{\theta}} \Gamma\left(\frac{1-\sigma}{\theta} + 1\right) \left[r(r-1)^{\frac{\sigma-1}{\theta}} - (r-1)r^{\frac{\sigma-1}{\theta}} \right]. \end{aligned}$$

Taking the derivative with respect to r , we see that these $1 - \sigma^{th}$ moments are increasing in r as long as $\theta \geq \sigma - 1$. Since we assume that σ is greater than 1, the first moments, $E[C_1(j)]$ and $E[C_2(j)]$, by implication are falling in r .¹¹

Proposition 2 *The aggregate price level P is decreasing in the number of rivals r under autarky as long as the dispersion in firm efficiency levels is large enough to balance the consumer's love of variety, $\theta \geq \sigma - 1$.*

Proof. Intuitively, Proposition 2 is true because an increase in r shifts the distribution of markups to the left at the same time it reduces the expected marginal cost of the best supplier. More rigorously, since firms in all industries draw from the same underlying distribution, using the law of large numbers one can calculate the aggregate price level,

$$P^{1-\sigma} = E\left[\int_0^1 P(j)^{1-\sigma} dj\right] = \int_0^1 E[P(j)^{1-\sigma}] dj = E[P(j)^{1-\sigma}].$$

Recall that $P(j) = M(j)C_1(j)$. Using this pricing rule and noting that the distribution of the markup is independent of outcomes for the individual order statistics $C_1(j)$ and $C_2(j)$, we have

$$P^{1-\sigma} = E[M(j)^{1-\sigma} C_1(j)^{1-\sigma}] = E[M(j)^{1-\sigma} C_1(j)^{1-\sigma}]$$

We show above that both $E[M(j)]$ and $E[C_1(j)]$ are decreasing in r as long as $\theta \geq \sigma - 1$, thus P is also falling in r . ■

¹¹We also know from Proposition 1 that $E[C_2(j)]$ falls faster in r than $E[C_1(j)]$, since the expected ratio, $E\left[\frac{C_2(j)}{C_1(j)}\right]$ is falling in r .

3.3 The number of rivals

The variable in our model which BEJK normalize to suppress the effects of all other aspects of market structure on markups is the number of rivals r . We use a free entry condition to motivate the treatment of r as finite.¹² Following Melitz (2003), we assume that there is a uniform probability of death, $0 < \delta < 1$, in every period. Entrepreneurs must pay a fixed cost f in order to draw an efficiency parameter. This fixed cost is denominated in units of revenue here, but we can also specify f in units of labor as in Melitz (2003) without affecting our qualitative results below at all. When drawing, they can only see how many rivals there are in the industry, not the marginal cost of any rival. In equilibrium, the finite number of rivals, an integer, must be such that the expected present discounted value of output for an active producer is no smaller than the sunk cost of entry,

$$E_t \left[\sum_{s=0}^{\infty} (1 - \delta)^{t+s} (P_{t+s}(j)Y_{t+s}(j) - C_1(j)Y_{t+s}(j)) | r \right] \geq f \quad (6)$$

$$E_t \left[\sum_{s=0}^{\infty} (1 - \delta)^{t+s} (P_{t+s}(j)Y_{t+s}(j) - C_1(j)Y_{t+s}(j)) | r + 1 \right] < f.$$

We also use the labor market clearing condition to define market size Y . In steady state, it is

$$\omega L = \beta \lambda P Y, \quad (7)$$

where L is the number of workers, β is labor's cost share in the input bundle used to produce intermediate goods and λ is the share of variable costs as a fraction of total expenditures,

$$\lambda = \frac{E[C_1(j)Y(j)]}{E[P(j)Y(j)]} = \frac{E[M^{-\sigma}(j)]}{E[M^{1-\sigma}(j)]}. \quad (8)$$

Isolating Y in equation (7), normalizing the wage $w \equiv 1$, and then sub-

¹²Since the circulation of this paper, new working papers by Holmes, Hsu, and Li (2011) and Zolas (2011) have also begun to do so, with different applications relating to agglomeration and patenting.

stituting for Y and λ in the free entry condition, equation (6), yields the steady-state expressions

$$\begin{aligned} \frac{E[M^{1-\sigma}(j)|r]}{E[M^{-\sigma}(j)|r]} &\geq 1 + \frac{\beta\delta f}{L}, \\ \frac{E[M^{1-\sigma}(j)|r+1]}{E[M^{-\sigma}(j)|r+1]} &< 1 + \frac{\beta\delta f}{L}. \end{aligned} \quad (9)$$

We present this condition graphically in Figure 2, which is calibrated as in Figure 1, with the quantity $1 + \frac{\beta\delta f}{L}$ calibrated to show entry by 20 rivals in each industry, the level used in Atkeson and Burstein (2008). Recall that the probability of forced exit, δ , is independent of firm efficiency, and that the distribution of the markup is independent of the distribution of costs,¹³ so in Appendix B using Jensen’s inequality, we show that the free entry condition reduces to

$$E[\ln M(j)] \geq \ln \left(1 + \frac{\beta\delta f}{L} \right). \quad (10)$$

In Appendix B.2, we prove that the left hand side is decreasing in r , as shown in Figure 2, resulting in a unique equilibrium solution for r . This means that the number of rivals in each industry grows as the fixed cost f , and the exit rate δ fall, as well as when market size L is bigger.

The distribution of the markup derived above does not yield a closed-form solution for the expected markup $E[M(j)]$ or for the expected log markup, $E[\ln M(j)]$. However, we can determine an upper- and lowerbound for r . Noting from Jensen’s inequality that $E[M(j)] \geq \ln E[M(j)]$ and that $\ln E[M(j)] \geq E[\ln M(j)]$, the form of the free entry condition in equation (10) implies

$$E[M(j)] \geq \ln E[M(j)] \geq E[\ln M(j)] \geq \ln \left(1 + \frac{\beta\delta f}{L} \right). \quad (11)$$

Proposition 1 states that the mean markup, $E[M(j)]$, is decreasing in the number of rivals, r . In combination with this insight from Proposition 1, $E[M(j)] \geq \ln(1 + \frac{\beta\delta f}{L})$ implies that enough rivals simultaneously “enter” the

¹³To see this, recall that the cost parameters C_k do not enter into the expression for $h(m)$ for $k \in N$.

industry (i.e., draw a productivity parameter) such that expected profits are at least as large as the amortized fixed cost of entry.

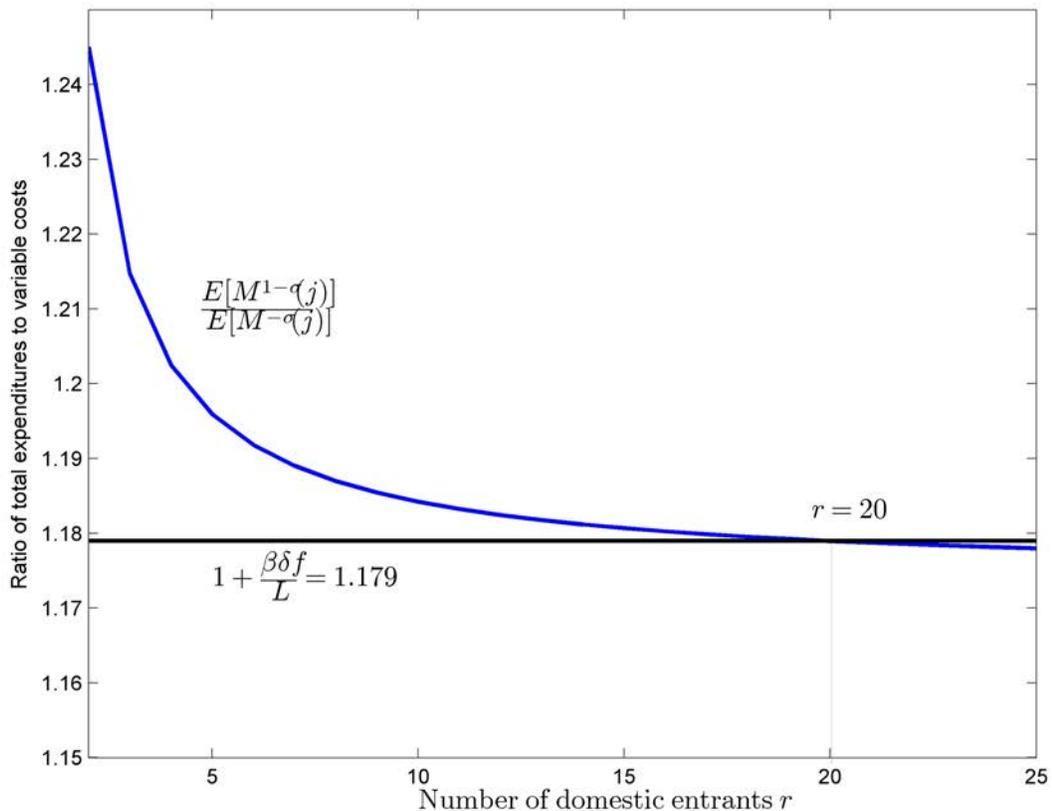


Figure 2: Free entry and the number of rivals under autarky

Appendix B.2 uses expression (9) to derive the upper- and lower- bounds for r given by

$$\frac{\ln\left(1 + \frac{\beta\delta f}{L}\right) (e^{\theta\bar{m}} - 1)}{\ln\left(1 + \frac{\beta\delta f}{L}\right) e^{\theta\bar{m}} - \bar{m}} \geq r \geq \frac{E[\ln M(j)](e^{\theta\bar{m}} - 1)}{E[\ln M(j)]e^{\theta\bar{m}} - \bar{m}}.$$

Notice that the number of rivals in each industry grows as the fixed cost f , the share of labor in the input bundle β , and the exit rate δ fall, as well as when

market size L is bigger. Campbell and Hopenhayn (2005) show a negative relationship between markups and market size in U.S. cities consistent with the scale effect implied by our free entry condition. Note also that entry is not proportional to changes in market size L , but can grow much faster than L , a departure from the Melitz model that will impact gains from trade below. This yields many interesting implications, for example, from equation (2) it is clear that increasing the number of rivals influences the distribution of costs exactly like an increase in the technology parameter T . Thus, reducing barriers to entry pushes out the technological frontier, in addition to lowering the average markup. Since λ and P are both falling in r , we can see from equation (7) that reducing either the fixed cost or the exit rate increases aggregate output Y by boosting the number of rivals.

4 Trade in goods

Trade in our model not only shifts production toward lower-cost producers in the classic Ricardian sense, but also reduces markups in countries with low contestability, lowering the aggregate price level. The reason is simple: all else equal, openness increases the number of firms competing to serve the domestic market. In addition, trade costs increase the marginal cost for exporters situated far from a destination market relative to their rivals, making it more likely that their price will be bounded by a geographically closer rival. Furthermore, openness has a second effect on the sensitivity of markups to changes in economic conditions. As trade costs eat into markups due to the direct competition of firms from closer locations, this relative gravity effect prohibits a larger fraction of exporters from being able to adjust prices in response to idiosyncratic or country-specific shocks and limits the degree to which those that can actually *do* adjust them.

Below, we focus principally on how trade changes the distribution of markups. It also has an impact on entry. We do not consider entry in a dynamic setting, as does Peters (2011), rather, we compare two worlds— one with trade and one without— and characterize the number of firms *ex ante* that would want to pay

a fixed cost to draw an efficiency parameter having been born into one versus the other.

4.1 The distribution of costs in the open economy

We follow BEJK's notation, adding the subscript n to the terms $C_k(j)$, $g_k(c_k)$, and $G_k(c_k)$ from the autarkic case to refer to the costs and distribution of costs for goods supplied to country n in the open economy. When the potential supplier is from country i we add the subscript i , so that the unit cost of the k^{th} most efficient firm from country i when supplying any good (j) to country n becomes $C_{kni}(j)$, drawn from the underlying cumulative distribution function $G_{kni}(c_k)$, with the corresponding probability density $g_{kni}(c_k)$. We assume that Eaton and Kortum's (2002) no arbitrage condition for trade costs holds: $d_{ni} < d_{ui}d_{nu}$, so that it is always cheaper to send a good directly to its destination market, never to re-route it through a third country.

Let $G_{1n}(c_1)$ be the probability that the low-cost supplier of a good j to the home country n has a marginal cost less than or equal to some level c_1 under trade. The probability is equal to one minus the probability that any other potential supplier—domestic or foreign—has a marginal cost greater than c_1 . The cumulative distribution for low-cost suppliers under trade is thus

$$\begin{aligned} G_{1n}(c_1) &= \Pr[C_{1n}(j) \leq c_1] = 1 - \prod_{i=1}^N [1 - G_{1ni}(c_1)] \\ &= 1 - e^{-\Phi_n c_1^\theta}, \end{aligned} \tag{12}$$

where $G_{1ni}(c_1)$ is the distribution of low-cost suppliers to n from country i , $\Phi_n = \sum_{i=1}^N T_i (w_i d_{ni})^{-\theta} r_i$, and $d_{ni} \geq 1$ is an iceberg trade cost involved in shipping goods from country i to country n for $i \neq n$. It is straightforward to show that the probability that a country exports to n is the same as in Eaton and Kortum (2002) and BEJK, but allowing for the number of rivals:

$$\pi_{ni} = \Pr[EXPORT_{ni}] = \frac{r_i T_i (w_i d_{ni})^{-\theta}}{\Phi_n}. \tag{13}$$

4.2 Geography and markups

In three steps, we can compute the full distribution of markups under costly trade with asymmetric countries. First, we consider the case that the best two rivals for a destination market originate in the same country. Let ψ_{ni} be the probability that the two best rivals to supply country n both originate in country i . Then, it must be that the two best rivals in a particular industry in country i are more efficient (have lower marginal costs) than any other potential suppliers of the good to country n . Let c_{2i} be the second-best cost draw for an industry in country i . Then the probability that it is lower than the best draw for the same industry in any country $u \neq i$ is

$$\begin{aligned}\psi_{ni} &= \int_0^\infty \int_{c_{1i}}^\infty g_{2ni}(c_{2i}) \prod_{u \neq i}^N [1 - G_{1nu}(c_{2i})] dc_{2i} \\ &= \pi_{ni} \psi'_{ni},\end{aligned}\tag{14}$$

where we define $\psi'_{ni} = \frac{(r_i-1)T_i(w_i d_{ni})^{-\theta}}{\Phi_n - T_i(w_i d_{ni})^{-\theta}}$, the probability that the second best producer in i will be the second best supplier to country n in the world market as a whole, given that the best producer of a good in i is also the best supplier to n worldwide.¹⁴ The distribution of markups in this case is a simple application of our autarkic distribution, renaming r in equation (3) as r_i (see Appendix C for further detail).

The second step is to compute the probability that the best supplier to n is from country i and the best rival supplier to supply country n is in country $u \neq i$, denoted ψ_{niu} . The unconditional probability that this occurs is the probability that the best supplier native to country u has some marginal cost c_{1u} , which lies between the first- and second-best draws in country i , c_{1i} and $c_{2i} > c_{1i}$, while the best rivals from all third countries ($v \neq u, i$) have a marginal cost that is larger than c_{1u} . See Appendix C for the full derivation

¹⁴Under symmetry, the probability ψ_{ni} collapses to the very intuitive expression $\psi = \frac{1}{N} * \frac{r-1}{Nr-1}$.

of this probability, given by

$$\begin{aligned}\psi_{niu} &= \int_0^\infty \int_{c_{1i}}^\infty \left(\int_{c_{1i}}^{c_{2i}} g_{1nu}(c_{1u}) \prod_{v \neq i, u}^{N-1} [1 - G_{1nv}(c_{1u})] dc_{1u} \right) g_{1ni,2ni}(c_{1i}, c_{2i}) dc_{2i} dc_{1i} \\ &= \psi'_{niu} \pi_{ni} (1 - \psi'_{ni}),\end{aligned}\tag{15}$$

where $\psi'_{niu} = \frac{r_u T_u (w_u d_{nu})^{-\theta}}{\Phi_n - r_i T_i (w_i d_{ni})^{-\theta}}$, the probability that the second best supplier to n is in country u conditional on the best supplier being from i while the second best is not. Note that $\sum_{u \neq i} \psi'_{niu} = 1$.

Finally, we compute the distribution of markups charged in country n given that the best rival to supply a good is in i and the second-best is in country u , which we call $h_{niu}(m)$. We use the formula for the distribution of the ratio of two independent random variables, $C_{1ni}(j)$ and $C_{1nu}(j)$, described by Mood, Graybill and Boes (1974, pp.187-88), given that $C_{1nu}(j)$ is greater than $C_{1ni}(j)$ but less than $C_{2ni}(j)$ and the best supplier to n from any third country:¹⁵

$$\begin{aligned}h_{niu}(m) &= \int_0^\infty \int_{mc_{1i}}^\infty c_{1i} \frac{g_{1ni,2ni}(c_{1i}, c_{2i})}{1 - G_{2ni}(mc_{1i})} \frac{g_{1nu}(mc_{1i})}{1 - G_{1nu}(c_{1i})} dc_{2i} dc_{1i} \\ &= \frac{\theta T_i (w_i d_{ni})^{-\theta} r_u T_u (w_u d_{nu})^{-\theta} m^{\theta-1}}{[T_i (w_i d_{ni})^{-\theta} + r_u T_u (w_u d_{nu})^{-\theta} (m^\theta - 1)]^2}\end{aligned}\tag{16}$$

Then, the full distribution of markups in country n under trade, $\tilde{h}_n(m)$, is given by

$$\tilde{h}_n(m) = \sum_{i=1}^N \psi_{ni} h_i(m) + \sum_{i=1}^N \sum_{u \neq i}^N \psi_{niu} h_{niu}.$$

It is easy to verify that the relevant weights sum to one: $\sum_{i=1}^N \psi_{ni} + \sum_{i=1}^N \sum_{u \neq i}^N \psi_{niu} =$

1. The important result for our purposes is the probability that the supplier charges the unconstrained markup when its next-best rival is an exporter in a

¹⁵We show in Appendix C that because the distribution of efficiency levels in each country is independent, terms for third countries cancel out of this conditional distribution. We can integrate the density over the domain $[1, \bar{m}]$, noting the mass point at \bar{m} and see that it forms a well behaved cumulative distribution function that integrates to one.

different country,¹⁶

$$\Pr[M_{niu} \geq \bar{m}] = \frac{T_i(w_i d_{ni})^{-\theta}}{T_i(w_i d_{ni})^{-\theta} + r_u T_u(w_u d_{nu})^{-\theta} (\bar{m}^\theta - 1)}. \quad (17)$$

One can see immediately that the supplier to country n exporting from country i will be more likely to charge the unconstrained markup when its next-best rival (a) resides in a country far from the destination country n (high d_{nu}), or (b) resides in a country with low contestability, low technology, or a high wage relative to country i . The country- i supplier's own distance from the destination country lowers the probability that it can charge the unconstrained markup. If all countries are identical, the terms representing technology, wages, and trade costs cancel out and this expression is easily shown to be lower than the probability under autarky in equation (4) for finite $r \geq 2$. We formalize this result in Proposition 3.

Proposition 3 *If a producer from country i exports to country n with the next-best rival to supply the same good to country n from a different country $u \neq i$, then the exporter's market power on average is increasing in country i 's technology relative to country u , $\frac{T_i}{T_u}$, but decreasing in country i 's relative wage, $\frac{w_i}{w_u}$; relative distance to the destination market, $\frac{d_{ni}}{d_{nu}}$; and the level of contestability in the rival's home country, r_u .*

Proof. From equation (17), for any markup m' , $1 \leq m' \leq \bar{m}$, the probability that the exporter's markup is greater than or equal to m' , given that its rival resides in a different country, is

$$\Pr[M_{niu} \geq m'] = \frac{T_i(w_i d_{ni})^{-\theta}}{T_i(w_i d_{ni})^{-\theta} + r_u T_u(w_u d_{nu})^{-\theta} (m'^\theta - 1)}.$$

All else equal, the derivative of this probability is increasing in $\frac{T_i}{T_u}$. By the property of stochastic dominance, the average markup is also increasing in $\frac{T_i}{T_u}$.

¹⁶More generally, the cumulative probability $\Pr[M(j) \leq m'] = 1 - \Pr[M_{niu} \geq m']$ ranges from 0 to 1 as m' increases from 1 to ∞ , so it is a well behaved cumulative distribution function for markups.

Similarly, the probability is decreasing in $\frac{w_i}{w_u}$, $\frac{d_{ni}}{d_{nu}}$, and r_u . Thus, the average markup in this case is also decreasing in these three factors. ■

Considering the special case where the next-best rival resides in the domestic market, $u = n$, demonstrates that opening to trade from autarky (reducing d_{ni} from infinitely large to a much lower level, while d_{nn} remains fixed at 1), we see that trade openness, on average, reduces the market power of domestic producers, forcing them to charge lower markups on sales in their native market. The only way that the average markup can increase under trade is if the home country n opened its borders to trade with a world dominated by one country that was both much closer than other trading partners (low d_{ni}) and this close neighbor was far superior to all other countries by having much lower labor input costs (low ω_i), or very advanced technology (high T_i). What is more, equation (17) implies that reducing the trade cost d_{ni} for only one particular country i increases the probability that a foreign supplier from i will be able to charge their full autarkic markup when selling to country n . Thus, our model nests the anti-competitive mechanism underlying results in EMX and Arkolakis, Costinot, Donaldson, and Rodríguez-Clare (2012), but also demonstrates that it can be superseded by multilateral, as opposed to unilateral, trade liberalization.

Corollary 2 *If an exporter to country n is located in a country $i \neq n$ with its next best rival to supply n in a third country $u \neq i, n$, then changing from bilateral to multilateral trade liberalization reduces the average markup.*

Proof. Lowering trade costs only between i and n (d_{ni}) reduces i 's geographic friction with respect to n relative to its competitor in country u , $\frac{d_{ni}}{d_{nu}}$. It follows from Proposition 3 that exporters from i to n can proceed to charge a higher markup on average than before the bilateral liberalization given that the next best rival is in country u . If afterward, country n lowers trade costs with respect to u , changing from bilateral to multilateral liberalization, then $\frac{d_{ni}}{d_{nu}}$ increases and by the same Proposition 3, the probability $\Pr[M_{niu} \geq m']$ falls. ■

Thus, bilateral liberalization can create an anti-competitive effect, increasing market power for exporters on average, while switching from bilateral to multilateral agreements can generate a pro-competitive effect. As in any Ricardian model, regardless of whether it is bilateral or multilateral, trade openness reduces prices, as seen in the following Lemma.

Lemma 1 *Trade lowers the aggregate price level.*

Proof. A country will never import a good with a higher price than it pays under autarky and the second-best competitor will never be less efficient than the second-best competitor under autarky. To quantify the impact on the aggregate price level, we can compute

$$\begin{aligned} P_n^{1-\sigma} &= E[P_n(j)^{1-\sigma}] \\ &= \Pr [M_n(j) > \bar{m}] \bar{m}^{1-\sigma} E[C_{1n}(j)^{1-\sigma}] + \Pr [M_n(j) \leq \bar{m}] E[C_{2n}(j)^{1-\sigma}] \end{aligned}$$

and note that the relevant moment of the supplier's expected marginal cost is given by

$$E[C_{1n}(j)^{1-\sigma}] = \int_0^{\infty} c_1^{1-\sigma} g_{1n}(c_1) dc_1 = (\Phi_n)^{\frac{\sigma-1}{\theta}} \Gamma \left(\frac{\theta - (\sigma - 1)}{\theta} \right).$$

Since $\sigma > 1$, this moment is strictly greater than its counterpart under autarky. We also can compute the same $1 - \sigma^{th}$ moment for the marginal cost of the second-best rival by using the probability that it is in the same source country i as the actual supplier, ψ_{ni} :

$$E[(C_{2n})^{1-\sigma}] = \sum_{i=1}^N \psi_{ni} E[(C_{2ni})^{1-\sigma}] + \sum_{i=1}^N \sum_{u \neq i}^N \psi_{niu} E[(C_{1nu})^{1-\sigma}],$$

which we know is at least as great as its counterpart under autarky because the second-best rival producer of a good j in the entire world (including the home country) by definition could not have a marginal cost any higher than the second-best rival under autarky. \blacksquare

Under costly trade, the markups that firms charge are different when they sell domestically compared to when they export. The formula for the distribution of markups, $\tilde{h}_n(m)$, reveals that firms internalize a portion of the trade cost, unless they are so technologically superior or have such a huge unit input cost advantage that they can pass the entire trade cost on to the foreign consumer. We demonstrated that the probability of charging the unconstrained markup is lower when one's next best rival is from a different country. The effect of incremental reductions in the trade cost on the import penetration ratio is no longer a constant, which Arkolakis, Costinot, and Rodríguez-Clare (2012b) report is the case for the BEJK model without entry. More formally, profits are no longer a constant share of revenues, independent of the variable trade cost.¹⁷ Instead, the share of profits in total revenues varies with the variable trade cost d , shrinking as d falls and firms are forced to charge lower markups due to competition from new foreign and possibly new domestic entrants. This violates the gravity restriction satisfied by many trade models, even though the probability of exporting to any country n , π_{ni} , appears very similar to the export equations in Eaton and Kortum (2002) and BEJK.

Put more simply, the gains from trade liberalization cannot be inferred from the value of aggregate flows alone because liberalization reduces markups, distorting the relationship between the trade cost and observed expenditures. Thus, trade liberalization has the potential to create welfare gains not only through productivity-based comparative advantage, but also by reducing firms' market power. We close the model and show output growth under free trade versus autarky under symmetry and free trade in Appendix E, but save detailed analysis of gains from costly trade with variable markups in this generalized Ricardian setting for future research and in order to focus our analysis on entry, pricing behavior, and the aggregate price level.

One potential question is whether the fact that rivals within each industry are latent— they do not produce and thus can not be observed for empirical analysis— is important in interpreting our results. The answer is a resounding

¹⁷The constant share in bilateral profit is also an important intermediate result in Arkolakis, Costinot, Donaldson, and Rodríguez-Clare (2012a).

“No.” De Blas and Russ (2011) generalize the type of Bertrand competition employed here to use a search framework in the open economy, which allows for a continuum of firms to produce the same good as long as search costs restrict the number of prices that buyers check before making a purchase, in the spirit of Burdett and Judd (1983) and Alessandria (2009). They use numerics rather than the analytics here, but all of the same intuition is reflected in the numerical results, suggesting fertile ground for future research applying our analytical distributions.

5 Price adjustment, volatility, and pass-through

Empirical studies indicate that idiosyncratic shocks are likely to be prevalent and economically important: Gabaix (2011) finds that a substantial portion of observed aggregate fluctuations in U.S. output can be explained by idiosyncratic shocks falling across a distribution of heterogeneous firms, while Foster, Haltiwanger, and Syverson (2008) determine that idiosyncratic shocks affecting plant-level output have a standard deviation five times as large as that of industry-level productivity shocks. Thus, recent literature indicates that idiosyncratic shocks are likely to be important from a macroeconomic perspective. In this context, our setup can shed some light on the pricing behavior of individual firms subject to idiosyncratic shocks in domestic versus foreign markets. In Appendix F, we analyze price adjustment in the closed economy with idiosyncratic shocks to fix show that the degree of contestability can affect firms’ ability to pass on idiosyncratic (firm-specific) shocks into prices. We show, drawing from the insights of Kucheryavy (2012), that perfect substitutability is not necessary to achieve this type of price rigidity.¹⁸

¹⁸Kucheryavy (2012) proves that the market share of each rival to the lowest-cost producer in an industry continuously falls to zero as the elasticity approaches infinity. By implication, he contradicts the common misconception that the number of firms charging a markup constrained by competition with rivals jumps to zero when this elasticity is finite. In fact, it converges continuously to zero as the elasticity falls to 1. We build on his insights to show numerically a continuous drop in pass-through as the elasticity approaches 1 in Appendix F. Thus, we disagree with the suggestion in Atkeson and Burstein (2008, p.2013) that the discontinuity in the number of firms, which jumps from $r > 0$ to 1 when

Here, we focus on pass-through in the open economy by applying a country-specific shock.

5.1 Trade and prices

The expressions for markup behavior in Section 3.2 yield pricing-to-market, incomplete pass-through, and the closely related facts that firms change prices on exported goods less frequently and with less synchronization relative to prices in the domestic market. Atkeson and Burstein (2007 and 2008) describe in brilliant detail the manner in which numerical simulations of BEJK and an innovative new quantity-based competitive framework result in pricing-to-market and incomplete pass-through, matching them with data on pricing behavior. Here, we demonstrate similar results algebraically. First, pricing-to-market is evident in the formula for $\tilde{h}_n(m)$ and both of its components, $h_i(m)$ and $h_{niu}(m)$. Unless trade is costless, firms can charge higher markups in their home markets than abroad because trade costs increase their domestic market power, as discussed above. The formulas also depict how firms set markups depending on the proximity of other export competitors in a particular destination market, if their next best rival is another exporter. Second, as under autarky, firms will only fully pass an increase in marginal cost to buyers in export market n if (a) they are already charging the unconstrained markup, and (b) the price increase would not surpass the marginal cost of the next-best rival to supply country n . Although the logic is quite general, we can show this mathematically if we again invoke symmetry, this time with costly trade.

Suppose again that there is a shock to marginal cost ε such that a shock $\varepsilon > 1$ reduces efficiency and increases the marginal cost of an industry's low-cost supplier in country n . Using equation (17), the probability that pass-

the elasticity reaches infinity, causes a discontinuity in pricing behavior which would make Bertrand competition less useful than Cournot or that assuming perfect substitutability is not a useful approximation to examine limited price flexibility.

through occurs under trade is now

$$\Pr[M_{niu}(j) = \frac{C_{2n}(j)}{\varepsilon C_{1n}(j)} \geq \bar{m}] = \frac{T_i(w_i d_{ni})^{-\theta}}{T_i(w_i d_{ni})^{-\theta} + r_u T_u(w_u d_{nu})^{-\theta} [(\varepsilon \bar{m})^\theta - 1]}.$$
(18)

By Proposition 3, it follows that pass-through is increasing in the technological advantage of the exporting country, $\frac{T_i}{T_u}$, and decreasing in $\frac{d_{ni}}{d_{nu}}$, $\frac{w_i}{w_u}$, and r_u —the same Ricardian factors that govern the market power of exporters.¹⁹

Our markup formulas demonstrate results described in the numerical simulations of Garetto (2012). As we noted above from equation (17), the probability that a firm charges the unconstrained markup (and as a result, the degree of passthrough) in an export market is greater when the exporting country has a higher level of technology T or a lower wage ω than its competitor’s source country. Thus, we show the point Garetto (2012) argues— that firms exhibit less market power in setting markups “when exporting to relatively more productive (richer) countries.”

5.2 Export price rigidity and macroeconomic volatility

The macro-level manifestation of restricted price adjustment in export markets is reduced volatility in the terms of trade relative to the real exchange rate for high levels of contestability (r) in the host market or in the presence of trade costs. To illustrate the relationship between the micro and macro effects of relative cost shocks across countries, such as a small movement in the nominal exchange rate (see Burstein, Eichenbaum, and Rebelo (2005) for a discussion contrasting the impact of large versus small shocks on the real exchange rate), Table 1 lists the volatility of the terms of trade relative to the real exchange rate in U.S. data alongside results from simulated data for small cost shocks

¹⁹Although several studies have shown that pass-through depends on the choice of currency invoicing, Goldberg and Tille (2009) demonstrate that this currency invoicing choice also depends on the degree of competition in the destination market, so we view our market structure approach as quite relevant.

in our model.²⁰ We see in the table that the model delivers a variance in the terms of trade that is approximately one-half the variance of the real exchange rate when $d = 1.5$, between our lower-bound for trade costs of 1.25 from the survey by Russ and Valderrama (2010) and the higher value for trade costs (1.74) estimated by Anderson and van Wincoop (2004). This ratio of volatilities corresponds with the figure reported for the U.S. in Corsetti, Dedola, and Leduc (2008). The variance of the terms of trade relative to the real exchange rate increases monotonically as trade costs fall in the symmetric, two-country case. This relative variance is higher when the level of domestic rivalry is higher. High trade costs suppress exporters' ability to adjust prices more strongly than the level of domestic entry, which is quite similar to the results from the variable elasticity of substitution framework developed by Gust, Leduc, and Vigfusson (2010).

Thus, we interpret our framework as a complement to but not a replacement for nominal price stickiness analyzed in existing studies.²¹ However, if we

²⁰That is, to focus on the main mechanism of the Bertrand pricing behavior, these are small departures from a symmetric steady state without second-order effects on wages or entry. We use the same calibration as in Figure 1, with lognormal shocks that enter like ε above, but applied to all firms within a country. The shock is lognormally distributed with log of these shock distributed as normal with mean zero and variance 0.015, so that the standard deviation is equal to the standard deviation of aggregate technology shocks in the U.S. estimated by Basu, Fernald, and Kimball (2006), applied to each of two symmetric countries. Again, these are small deviations from steady state to focus on the workings of the pricing mechanism, apart from general equilibrium effects. We hold r constant in each case.

²¹Qualitative results for pass-through and relative macroeconomic volatilities do not depend on perfect substitutability between goods within a market niche. However, with this special limiting case of perfect substitutability, distributions in Section 3 imply that the frequency of price changes will be smaller in export markets than in domestic markets. Unless an exporting country has a huge advantage in the form of high T , high r , or low labor costs, it is harder for firms to charge the unconstrained markup in an export market compared to their native market. This is due to the trade cost, which effectively increases exporters' marginal cost relative to domestic firms in the destination country. Since firms must be charging the unconstrained markup in order to pass on idiosyncratic or country-specific shocks in the form of higher export prices, fewer firms will change prices in export markets (as compared to their native market) when marginal costs increase. As a consequence, the median and average frequency of price changes must be lower for exports, as shown by Schoenle (2010) and Fitzgerald and Haller (2010). Interestingly, it also replicates the frequency of export price changes relative to the frequency of domestic price changes documented by Schoenle (2010), but only for very low trade costs.

interpret country-specific shocks as exchange rate shocks, then our results coincide with the departure from Harrod-Balassa-Samuelson effects observed in Japan following the Plaza Accord by Obstfeld (2009). He documents that real exchange rate appreciation coincided with a drop in measured productivity in the traded goods sector during this period. If the appreciation of a country’s nominal exchange rate results in its exporters charging lower markups, this can reduce measured TFP, either because reduced markups may cause exporters to operate under capacity, as Obstfeld (2009) suggests, or due to empirical methodologies that confound changes in TFP with changes in markups. Using a numerical framework very similar to the one we present here, Garetto (2012) similarly finds that an appreciation of an exporter’s home currency relative to that of the destination country squeezes the exporter’s markups. In line with the Ricardian effects in our markup distributions discussed above, she also finds empirical evidence that pass-through of these shocks is more limited for exporters selling in countries that are relatively more technologically advanced— exactly where we predict their markups would be lower and their likelihood of pass-through less. Feenstra and Weinstein’s (2010) finding that exporters in low-wage countries, which have lower labor productivity, charge lower markups also supports this prediction. This is where our approach has a nontrivial nuance that is hard to reproduce with an assumption of nominal stickiness, unless we assume that degree of nominal stickiness is increasing in the technological development of the destination market relative to the source country.

Table 2 demonstrates that the volatility of the terms of trade in a 2-country model is higher for any level of trade costs when the home country has a lower level of available technology, which we interpret as a developing economy. Similarly, the volatility of both import and export prices is higher in the developing-country case. We calibrate the exercise from parameter estimates by Eaton and Kortum (2002), with $T_n = 0.5 < T_i = 1$. A decrease in trade costs in the 2-country case dramatically increases the volatility of import prices and, in more competitive environments ($r = 20$), it increases the volatility of import prices faster than export prices. This is compatible with empirical evi-

dence by Frankel, Parsley, and Wei (2012) demonstrating the higher volatility of import prices in developing countries and is the flip side of an analysis by Bergin and Feenstra (2009) that reproduces a drop in the volatility of import prices for the U.S. when trade increases with developing countries.

Moving one step beyond this Ricardian effect, one can see that reductions in trade barriers that replace bilateral treaties with multilateral treaties can actually reduce volatility in traded goods prices for developing countries. Our multi-country distribution in equation (16) shows the importance of asymmetry in a multi-country world, demonstrated in the last four columns of Table 2. In the 3-country world simulated in this table, lowering trade costs across all trading partners *reduces* the relative volatility of the terms of trade, as well as the volatility of import and export prices. The crucial difference here is that a bilateral trade liberalization—maintaining a low level of trade barriers with only one advanced country—can result in higher price volatility for a country with inferior technologies, as high-tech exporters from the advanced trading partner with favored treatment can take advantage of the relatively protected environment and inefficient domestic competitors by charging higher markups. Thus, they can pass along more of their own production cost shocks. However, multilateral liberalization—lowering trade barriers toward two or more advanced countries—can inhibit this markup effect (reduce the market power of foreign exporters) by forcing competition within the developing country between high-tech foreign suppliers from different advanced countries. This reduction in volatility for developing countries has some support in new findings by Flach and Cao (2011) that entry into multilateral trade agreements like the GATT/WTO or FTAs reduces import and export price volatility in developing countries.

6 Conclusions

The purpose of this paper is to understand the fundamentals of pricing behavior for heterogeneous firms in the open economy under Bertrand competition. We provide new distributions of markups which are sensitive to market struc-

Table 1: Volatility and passthrough, two identical countries

	U.S. data*	$d = 1.75$		$d = 1.5$		$d = 1.25$	
		r=2	r=20	r=2	r=20	r=2	r=20
$\frac{\sigma_{TOT}}{\sigma_{RER}}$	0.56	0.36	0.38	0.56	0.59	1.16	1.21
Avg. pass-through, exports	0.25	0.06	0.06	0.06	0.06	0.06	0.07
$\frac{\sigma_{PEX}}{\sigma_{RER}}$		0.70	0.68	0.87	0.84	1.47	1.40
$\frac{\sigma_{PIM}}{\sigma_{RER}}$		0.71	0.68	0.90	0.85	1.55	1.39

*U.S. figures for $\frac{\sigma_{TOT}}{\sigma_{RER}}$ are from Corsetti, Dedola, and Leduc (2008), frequencies from Schoenle (2010), and pass-through from Gopinath, Itshoki, and Rogobon (2010).

Table 2: Volatility and passthrough, developing country

	2 countries, $T_n < T_i$				3 countries, $T_n < T_i = T_j$			
	$d = 1.75$		$d = 1.25$		$d_{ni} = 1.25 < d_{nj}$		$d_{ni} = d_{nj} = 1.25$	
	r=2	r=20	r=2	r=20	r=2	r=20	r=2	r=20
$\frac{\sigma_{TOT}}{\sigma_{RER}}$	1.04	1.03	3.68	4.27	2.92	3.15	2.56	2.70
Avg. pass-through, imports	0.03	0.05	0.06	0.08	0.17	0.22	0.14	0.12
$\frac{\sigma_{PEX}}{\sigma_{RER}}$	1.55	1.58	4.12	4.25	3.20	3.23	2.71	2.71
$\frac{\sigma_{PIM}}{\sigma_{RER}}$	1.58	1.55	4.92	4.25	3.50	3.23	2.88	2.71

* $T_n = 0.5$ and $T_i = 1$ in the 2-country case. $T_j = 1$ also in the 3-country case, with $d_{nj} = 1.75$ in the 3-country case with asymmetric trade costs.

ture, demonstrating how market structure has important implications for gains from trade and both the level and volatility of prices. The distributions allow us to characterize in an analytically clean way firm markup behavior under trade, as well as the percentage of firms who can change their price in response to an idiosyncratic shock in any market, or in response to a source-country-specific shock in an export market. As in previous numerical studies using similar frameworks, key results include imperfect pass-through, pricing-to-market, and a lower volatility of the terms of trade relative to the real exchange rate. Our breakthrough is that we explicitly characterize in tractable formulas an endogenous degree of export price rigidity that depends on a multi-country market structure and varies across destination markets due to the degree of domestic entry and the level of trade costs.

References

- [1] Alessandria, George. 2009. Consumer Search, Price Dispersion, and International Relative Price Fluctuations. *International Economic Review* 50(3): 803-29.
- [2] Anderson, James E. and Eric van Wincoop. 2004. Trade Costs. *Journal of Economic Literature* 42:692-751.
- [3] Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare. 2012a. The Elusive Pro-Competitive Effects of Trade. Manuscript, University of California, Berkeley.
- [4] Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare. 2012b. New Trade Models, Same Old Gains? *American Economic Review* 102(1): 94-130.
- [5] Atkeson, Andrew and Ariel Burstein. 2007. Pricing to Market in a Ricardian Model of International Trade. *American Economic Review Papers and Proceedings* (May) 97(2): 362-367.
- [6] Atkeson, Andrew and Ariel Burstein. 2008. Pricing to Market, Trade Costs, and International Relative Prices. *American Economic Review* (December) 98(5): 1998-2031.
- [7] Basu, Susanto, John G. Fernald, and Miles S. Kimball. 2006. Are Technology Improvements Contractionary? *American Economic Review* 96(5): 1418-1448.

- [8] Baumol, W.J., Panzar, J.C., and Willig, R.D. 1982. *Contestable Markets and the Theory of Industry Organization*. San Diego: Harcourt-Brace-Jovanovich.
- [9] Baxter, Marianne and Anthony Landry. 2010. *IKEA: Products and Passthrough*. Manuscript, Boston University.
- [10] Bergin, Paul C. and Robert C. Feenstra. 2009. Pass-through of Exchange Rates and Competition between Floaters and Fixers. *Journal of Money Credit and Banking* 41(1):35-70.
- [11] Berman, Nicolas, Philippe Martin, and Thierry Mayer. 2012. How Do Different Exporters React to Exchange Rate Changes? *Quarterly Journal of Economics* 102: 437-492.
- [12] Bernard, Andrew, Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum. 2003. "Plants and Productivity in International Trade," *American Economic Review* 93(4): 1268-1290.
- [13] Bottasso, Anna, and Alessandro Sembenelli. 2001. Market Power, Productivity, and the EU Single Market Program: Evidence from a Panel of Italian Firms. *European Economic Review* 45(1): 167-186.
- [14] Burdett, Kenneth and Kenneth L. Judd. 1983. Equilibrium Price Dispersion. *Econometrica* 51(4): 955-969.
- [15] Burstein, Ariel, Martin Eichenbaum and Sergio Rebelo. 2005. Large Devaluations and the Real Exchange Rate. *Journal of Political Economy* 113(4): 742-784.
- [16] Campbell, Jeffrey R. and Hugo A. Hopenhayn. 2005. Market Size Matters. *Journal of Industrial Economics* v53(1): 1-25.
- [17] Castillo, Enrique. 1988. *Extreme Value Theory in Engineering*. San Diego, CA: Academic Press.
- [18] Castillo, Enrique, Ali S. Hadi, N. Balakrishnan, and Jose M. Sarabia. 2005. *Extreme Value and Related Models with Applications in Engineering and Science*. Hoboken, NJ: John Wiley & Sons, Inc.
- [19] Claessens, Stijn and Luc Laeven. 2004. What drives bank competition? *Journal of Money, Credit, and Banking* 36(3): 563-583.
- [20] de Blas, Beatriz and Katheryn N. Russ. 2011. *All Banks Great, Small and Global: Loan Pricing and Foreign Competition*. Manuscript, University of California, Davis.
- [21] Corsetti, Giancarlo, Luca Dedola and Sylvain Leduc. 2008. High Exchange Rate Volatility and Low Pass-Through. *Journal of Monetary Economics* 55(6): 1113-1128.
- [22] David, Herbert A. and Haikady N. Nagaraja. 2003. *Order Statistics*. Hoboken, NJ: John Wiley & Sons, Inc.

- [23] de Loecker, Jan, Pinelopi Goldberg, Amit Khandelwal, and Nina Pavcnik. 2012. Prices, Markups, and Trade Reform. Manuscript, Columbia University.
- [24] 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.
- [25] Drodz, Lucasz and Jaromir Nosal. 2012. Understanding International Prices: Customers as capital. *American Economic Review* 102(1): 364-95.
- [26] Eaton, Jonathan and Samuel Kortum. 2002. "Technology, Geography, and Trade." *Econometrica* 70(5): 1741-1779.
- [27] Eaton, Jonathan and Samuel Kortum. 2009. Technology in the Global Economy: A Framework for Quantitative Analysis. Manuscript, University of Minnesota.
- [28] Edmond, Christopher, Virgiliu Midrigan, and Daniel Xu. 2012. Competition, Markups, and the Gains from Trade. NBER Working Paper No. 18041. (First draft: 2006.)
- [29] Feenstra, Robert C. and David E. Weinstein. 2010. Globalization, Markups, and the U.S. Price Level. NBER Working Paper No. 15749.
- [30] Feenstra, Robert C., Maurice Obstfeld, and Katheryn Russ. 2011. In Search of the Armington Elasticity. Manuscript, University of California, Davis.
- [31] Fitzgerald, Doireanne and Stephanie Haller. 2010. Pricing-to-Market: Evidence from Plant-Level Prices. Manuscript, Stanford University.
- [32] Flach, Lisandra and Vinh T. H. Cao. 2011. The Effect of GATT/WTO on Export and Import Price Volatility. Manuscript, University of Mannheim.
- [33] Foster, Lucia, John Haltiwanger and Chad Syverson. 2008. Reallocation, Firm Turnover and Efficiency: Selection on Productivity or Profitability?" *American Economic Review* (March).
- [34] Frankel, Jeffrey, David Parsley, and Shang-Jin Wei, 2012. Slow Pass-through Around the World: A New Import for Developing Countries? *Open Economies Review*, 23(2): 213-251 (April).
- [35] Gabaix, Xavier. The Granular Origins of Aggregate Fluctuations. 2011. *Econometrica* 79(3): 733-772.
- [36] Garetto, Stefania. 2012. Firm's Heterogeneity and Incomplete Pass-Through. Manuscript, Boston University.
- [37] Goldberg, Linda and Cedric Tille. 2009. Micro, Macro, and Strategic Forces in International Trade Invoicing. CEPR Discussion Paper 7534.
- [38] Gopinath, Gita and Oleg Itshoki. 2010. Frequency of Price Adjustment and Pass-through. *Quarterly Journal of Economics* 125(2): 675-727.

- [39] Gopinath, Gita, Oleg Itshoki, and Roberto Rigobon. 2010. Currency Choice and Exchange-Rate Pass-through. *American Economic Review* 100(1): 306-336.
- [40] Gopinath, Gita and Roberto Rigobon. 2008. Sticky Borders. *The Quarterly Journal of Economics* 123(2):531-575.
- [41] Gust, Christopher, Sylvain Leduc, and Robert Vigfusson. 2010. Trade Integration, Competition, and the Decline in Exchange-Rate Pass-Through. *Journal of Monetary Economics* 57: 309-324.
- [42] Harrison, Ann E. 1994. Productivity, Imperfect Competition and Trade Reform. *Journal of International Economics* 36(1-2): 53-73 (February).
- [43] Holmes, Thomas J., Wen-Tai Hsu, and Sanghoon Lee. 2011. Plants, Productivity, and Market Size with Head-to-Head Competition. Manuscript University of Minnesota (23 June).
- [44] Kucheryavyy, Konstantin. 2012. Continuity of a Model with a Nested CES Utility Function and Bertrand Competition. Forthcoming in *Economics Letters*.
- [45] Levinsohn, J. 1993. Testing the imports-as-market-discipline hypothesis. *Journal of International Economics* 35, pp. 122.
- [46] Malik, Henrick J. and Roger Trudel. 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.
- [47] Melitz, Marc J. 2003. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica* 71: 1695-1725 (November).
- [48] Melitz, Marc J. and Giancarlo I. P. Ottaviano. 2008. Market Size, Trade, and Productivity. *Review of Economic Studies* 75(1): 295-316.
- [49] Mood, Alexander McFarlane, Franklin A. Graybill, and Duane C. Boes. 1974. *Introduction to the Theory of Statistics*. McGraw Hill Companies.
- [50] Nakamura, Emi and Jon Steinsson. 2010. Lost in Transit: Product Replacement Bias and Pricing to Market. Manuscript, Columbia University.
- [51] Novy, Dennis. 2010. International Trade and Monopolistic Competition without CES: Estimating Translog Gravity. Manuscript, University of Warwick.
- [52] Obstfeld, Maurice. 2009. Time of Troubles: The Yen and Japan's Economy, 1985-2008. NBER Working Paper No. 14816.
- [53] Pawlas, Piotr and Dominik Szyal. 2000. *Applicationes Mathematicae* 27(2): 197-202.

- [54] Peters, Michael. 2011. Heterogeneous Firms and Endogenous Misallocation. Manuscript, Massachusetts Institute of Technology.
- [55] Raftery, Adrian E. 1987. Inference and Prediction for a General Order Statistic Model with Unknown Population Size. *Journal of the American Statistical Association*. 82(400): 1163-1168.
- [56] Rinne, Horst. 2009. *The Weibull Distribution: a Handbook*. New York: CRC Press.
- [57] Roberts, Mark and Dylan Supina. 1996. Output Price, Markups, and Producer Size. *European Economic Review* 40(3-5): 909-21.
- [58] Rodriguez, Jose Antonio Lopez. 2011. Prices and Exchange Rates: A Theory of Disconnect. *Review of Economic Studies* 78(3): 1135-1177.
- [59] Russ, Katheryn N. and Diego Valderrama. 2010. Financial Choice in a Ricardian Model of Trade. NBER Working Paper No.15528.
- [60] Schoenle, Raphael. 2010. International Menu Costs and Price Dynamics. Manuscript, Brandeis University.
- [61] Zolas, Nicholas J. 2011. Firms' Locational Patenting Decisions. Manuscript, University of California, Davis.

A Deriving the distribution of markups

Malik and Trudel (1982, equation (5.17)) use Mellin transforms to derive the following distribution for the ratio $z = \frac{c_1}{c_2}$, the ratio of the first order statistic to the second, given that the sample is from the Weibull distribution $G(c)$ in the main text. This distribution is also reported in Rinne (2009, p.244, equation (5.42c)):

$$\hat{h}(z) = \frac{r(r-1)\theta z^{\theta-1}}{[(r-1) + z^\theta]^2},$$

Since we specify the (unrestricted) markup as $m = \frac{c_2}{c_1}$, we note that m is a function of z , $m = \frac{1}{z}$, and apply a straightforward Jacobian transformation:

$$\begin{aligned} h(m) &= \left(\frac{r(r-1)\theta \left(\frac{1}{m}\right)^{\theta-1}}{[(r-1) + \left(\frac{1}{m}\right)^\theta]^2} \right) \left(\frac{1}{m^2} \right) \\ &= \frac{r(r-1)\theta m^{-(\theta+1)}}{[(r-1) + m^{-\theta}]^2}. \end{aligned}$$

Similarly, Malik and Trudel (1982) use a Mellin transform to derive the following distribution for $z = \frac{c_1}{c_2}$ given a Pareto distribution of efficiency draws (a power law distribution of cost draws)

$$\hat{h}(z) = \theta z^{\theta-1}.$$

Again, we specify the (unrestricted) markup as $m = \frac{c_2}{c_1}$, implying that $m = \frac{1}{z}$, and apply the Jacobian transformation:

$$\begin{aligned} h(m) &= \frac{\theta m^{1-\theta}}{m^2} \\ &= \theta z^{-(\theta+1)} \end{aligned}$$

B Free entry under autarky

Since the distribution of markups is the same for all goods j , we drop the goods index below for simplicity. Taking (natural) logs, the expression decomposes

into

$$\ln \left(1 + \frac{\beta \delta f}{L} \right) + \ln E[M^{-\sigma}] = \ln E[M^{1-\sigma}]. \quad (\text{B.1})$$

Since the natural log is a concave function, Jensen's inequality implies $E[\ln M^{1-\sigma}] \leq \ln E[M^{1-\sigma}]$ and $E[\ln M^{-\sigma}] \leq \ln E[M^{-\sigma}]$. The function $M^{-\sigma}$ has a greater degree of convexity than $M^{1-\sigma}$, so $\ln E[M^{-\sigma}] - E[\ln M^{-\sigma}] \geq \ln E[M^{1-\sigma}] - E[\ln M^{1-\sigma}]$.²² This last inequality implies that

$$E[\ln M^{1-\sigma}] \geq \ln \left(1 + \frac{\beta \delta f}{L} \right) + E[\ln M^{-\sigma}],$$

as taking the log inside the expectation reduces the right-hand side more than the left-hand side. We note that for any constant k , $E[\ln M^k] = kE[\ln M]$, yielding

$$E[\ln M] \geq \ln \left(1 + \frac{\beta \delta f}{L} \right).$$

B.1 Uniqueness

Standard properties of expectations tell us that $E[M(j)^{1-\sigma}] > E[M(j)^{-\sigma}]$ for $\infty > \sigma > 1$ and $M(j) \geq 1$. In Proposition 1, we showed that $E[M(j)]$ is decreasing in the number of rivals. Thus, $E[M(j)^{1-\sigma}]$ is increasing in r and $E[M(j)^{-\sigma}]$ is increasing even faster. Thus, $E[M(j)^{1-\sigma}]/E[M(j)^{-\sigma}]$ is greater than 1 and decreasing in r toward 1, meaning that there can only be one r for which the ratio equals the constant $(1 + \frac{\beta \delta f}{L})$

B.2 Upper- and lower- bounds for the number of rivals.

The distribution of the markup does not yield a closed-form solution for the expected markup $E[M]$ or for the expected log markup, $E[\ln M]$. However, we know from Proposition 1 that the mean markup $E[M]$ is decreasing in r .

²²Another way to see this is to note that $E[\ln M^{1-\sigma}]$, $\ln E[M^{1-\sigma}]$, $E[\ln M^{-\sigma}]$, and $\ln E[M^{-\sigma}]$ are all negative numbers, with $|\ln E[M^{1-\sigma}]| < |\ln E[M^{-\sigma}]| < |E[\ln M^{-\sigma}]|$ and $|\ln E[M^{1-\sigma}]| < |E[\ln M^{1-\sigma}]| < |E[\ln M^{-\sigma}]|$. Thus, switching the logs from outside to inside the expectation in equation (B.1) reduces the left hand side more than the right hand side.

Therefore, we determine an upper- and lower- bound for r . Specifically, we can express the minimum number of rivals as a function of the expected log markup and derive a clean closed-form solution for the maximum number of rivals. Let $V = \ln M$. Then the probability density for V is a simple transformation of $h(m)$,

$$\begin{aligned} h_V(v) &= e^v h(e^v) I_{\mathbb{R}_+}(v) \\ &= e^v \frac{r(r-1)\theta (e^v)^{-(\theta+1)}}{\left[(r-1) + (e^v)^{-\theta}\right]^2}. \end{aligned}$$

The probability that $V \geq \bar{m}$ (or any other positive constant) is then

$$\int_{\ln(\bar{m})}^{\infty} e^v \frac{r(r-1)\theta (e^v)^{-(\theta+1)}}{\left[(r-1) + (e^v)^{-\theta}\right]^2} dv = \frac{r}{1 + (r-1)e^{\theta\bar{m}}}.$$

Using a generalized version of Chebyshev's inequality²³, we can characterize a lower-bound for the number of rivals:

$$\begin{aligned} \bar{m} \Pr[\ln M \geq \bar{m}] &\leq E[\ln M] \\ \frac{r\bar{m}}{1 + (r-1)e^{\theta\bar{m}}} &\leq E[\ln M] \\ r &\geq \frac{E[\ln M](e^{\theta\bar{m}} - 1)}{E[\ln M]e^{\theta\bar{m}} - \bar{m}}. \end{aligned}$$

As noted previously, the expected markup and the number of rivals is inversely related, a relationship seen here in the lowerbound for r . When $E[M]$ falls, the lowerbound increases, reflecting the fact that more rivals will enter when the expected markup is high (and vice versa). We know from equation (11) that the expected log gross markup $E[\ln M]$ must be at least as large as the gross log per-period cost of production, $\ln(1 + \frac{\beta\delta f}{L})$, producing an upperbound

²³See Theorem 5 in Mood, Graybill, and Boes (1974, p.71): For a random variable X , a nonnegative function $g(\cdot)$, and a scalar $k > 0$, then $kP[g(X) \geq k] \leq E[g(X)]$.

for r . Thus, we know that r lies within the following bounds:

$$\frac{\ln\left(1 + \frac{\beta\delta f}{L}\right) (e^{\theta\bar{m}} - 1)}{\ln\left(1 + \frac{\beta\delta f}{L}\right) e^{\theta\bar{m}} - \bar{m}} \geq r \geq \frac{E[\ln M](e^{\theta\bar{m}} - 1)}{E[\ln M]e^{\theta\bar{m}} - \bar{m}}.$$

C The distribution of markups under trade

We calculate the distribution of markups using conditional densities and conditional probabilities. That is, one can compute the probability of an outcome m under costly trade and asymmetry as

$$\tilde{h}(m) = Pr[m|A1] * Pr[A1] + Pr[m|A2] * Pr[A2] + \dots Pr[m|Ax] * Pr[Ax],$$

where $A1\dots Ax$ are all the possible geographic locations of the first- and second-best supplier (both in country 1, one in country one and one in country 2, etc.). The weights in our expression for the compound distribution in the main text are the probability of any particular geographic pairing ($Pr[A1], \dots, Pr[Ax]$). Below, we show how to calculate the weights, then how we calculate the conditional densities for m .

To calculate the unconditional probability that both the first and second best suppliers of a good to country n are from country i , we start from the main text:

$$\psi_{ni} = \int_0^\infty \int_{c_1}^\infty g_{2ni}(c_{2i}) \prod_{u \neq i}^N [1 - G_{1nu}(c_{2i})] dc_{2i}$$

The first step is to derive the marginal distribution $g_{2ni}(c_{2i})$ from the joint distribution, which is analogous to the joint distribution under autarky but

including trade costs. Integrating from the lower limit c_{1i} , we have

$$\begin{aligned}
g_{2ni}(c_{2i}) &= \int_{c_{1i}}^{\infty} g_{1ni,2ni}(c_{1i}, c_{2i}) dc_{1i} \\
&= \int_{c_{1i}}^{\infty} r_i(r_i - 1) [\theta T_i(w_i d_{ni})^{-\theta}]^2 c_{1i}^{\theta-1} c_{2i}^{\theta-1} e^{-T_i(w_i d_{ni})^{-\theta} c_{1i}^{\theta}} e^{-(r_i-1)T_i(w_i d_{ni})^{-\theta} c_{2i}^{\theta}} dc_{1i} \\
&= r_i(r_i - 1) \theta T_i(w_i d_{ni})^{-\theta} c_{2i}^{\theta-1} e^{-(r_i-1)T_i(w_i d_{ni})^{-\theta} c_{2i}^{\theta}} \left(1 - e^{-T_i(w_i d_{ni})^{-\theta} c_{2i}^{\theta}}\right)
\end{aligned}$$

Substituting into the formula for ψ_{ni} yields

$$\begin{aligned}
\psi_{ni} &= \int_0^{\infty} r_i(r_i - 1) \theta T_i(w_i d_{ni})^{-\theta} c_{2i}^{\theta-1} e^{-(r_i-1)T_i(w_i d_{ni})^{-\theta} c_{2i}^{\theta}} \left(1 - e^{-T_i(w_i d_{ni})^{-\theta} c_{2i}^{\theta}}\right) \\
&\quad \times e^{-[\Phi_n - r_i T_i(w_i d_{ni})^{-\theta}] c_{2i}^{\theta}} dc_{2i} \\
&= \int_0^{\infty} r_i(r_i - 1) \theta T_i(w_i d_{ni})^{-\theta} c_{2i}^{\theta-1} e^{-[\Phi_n - r_i T_i(w_i d_{ni})^{-\theta} + (r_i-1)T_i(w_i d_{ni})^{-\theta}] c_{2i}^{\theta}} dc_{2i} \\
&\quad - \int_0^{\infty} r_i(r_i - 1) \theta T_i(w_i d_{ni})^{-\theta} c_{2i}^{\theta-1} e^{-[\Phi_n - r_i T_i(w_i d_{ni})^{-\theta} + (r_i-1)T_i(w_i d_{ni})^{-\theta} + T_i(w_i d_{ni})^{-\theta}] c_{2i}^{\theta}} dc_{2i} \\
&= \frac{r_i(r_i - 1) T_i(w_i d_{ni})^{-\theta}}{\Phi_n - T_i(w_i d_{ni})^{-\theta}} (-1) e^{-[\Phi_n - T_i(w_i d_{ni})^{-\theta}] c_{2i}^{\theta}} \Big|_0^{\infty} - \frac{r_i(r_i - 1) T_i(w_i d_{ni})^{-\theta}}{\Phi_n} (-1) e^{-\Phi_n c_{2i}^{\theta}} \Big|_0^{\infty} \\
&= r_i(r_i - 1) T_i(w_i d_{ni})^{-\theta} \left(\frac{1}{\Phi_n - T_i(w_i d_{ni})^{-\theta}} - \frac{1}{\Phi_n} \right) \\
&= \frac{r_i T_i(w_i d_{ni})^{-\theta} (r_i - 1) T_i(w_i d_{ni})^{-\theta}}{\Phi_n [\Phi_n - T_i(w_i d_{ni})^{-\theta}]}.
\end{aligned}$$

We can also derive the unconditional probability that the first and second best rivals to supply a good to country n are, respectively, from i and $u \neq i$. We start with the formula (equation (15)) in the main text,

$$\psi_{niu} = \int_0^{\infty} \int_{c_1}^{\infty} \left(\int_{c_{1i}}^{c_{2i}} g_{1nu}(c_{1u}) \prod_{v \neq i, u}^{N-1} [1 - G_{1nv}(c_{1u})] dc_{1u} \right) g_{1ni,2ni}(c_{1i}, c_{2i}) dc_{2i} dc_{1i}.$$

Define A as the inner integral,

$$\begin{aligned}
A &:= \int_{c_{1i}}^{c_{2i}} g_{1nu}(c_{1u}) \prod_{v \neq i, u}^N [1 - G_{1nv}(c_{1u})] dc_{1u} \\
&= \int_{c_{1i}}^{c_{2i}} \theta r_u T_u c_{1u}^{\theta-1} (w_u d_{nu})^{-\theta} e^{-r_u T_u (w_u d_{nu})^{-\theta}} c_{1u}^\theta e^{-[\Phi_n - r_i T_i (w_i d_{ni})^{-\theta} - r_u T_u (w_u d_{nu})^{-\theta}] c_{1u}^\theta} dc_{1u} \\
&= \frac{-r_u T_u (w_u d_{nu})^{-\theta}}{\Phi_n - r_i T_i (w_i d_{ni})^{-\theta}} e^{-[\Phi_n - r_i T_i (w_i d_{ni})^{-\theta}] c_{1u}^\theta} \Big|_{c_{1i}}^{c_{2i}} \\
&= \psi'_{niu} \left(e^{-[\Phi_n - r_i T_i (w_i d_{ni})^{-\theta}] c_{1i}^\theta} - e^{-[\Phi_n - r_i T_i (w_i d_{ni})^{-\theta}] c_{2i}^\theta} \right), \tag{D.1}
\end{aligned}$$

where we define $\psi'_{niu} = \frac{r_u T_u (w_u d_{nu})^{-\theta}}{\Phi_n - r_i T_i (w_i d_{ni})^{-\theta}}$ as in the main text.

We then define B as the integral with the first half of A ,

$$\begin{aligned}
B &:= \psi'_{niu} \int_0^\infty \int_{c_1}^\infty e^{-[\Phi_n - r_i T_i (w_i d_{ni})^{-\theta}] c_{1i}^\theta} g_{1ni, 2ni}(c_{1i}, c_{2i}) dc_{2i} dc_{1i} \\
&= \psi'_{niu} \left(\frac{(r_i - 1) T_i (w_i d_{ni})^{-\theta}}{\Phi_n - T_i (w_i d_{ni})^{-\theta}} \right) \left(\frac{r_i T_i (w_i d_{ni})^{-\theta}}{\Phi_n} \right) \\
&= \psi'_{niu} \psi'_{ni} \pi_{ni} \tag{D.2}
\end{aligned}$$

with $\psi'_{ni} = \frac{(r_i - 1) T_i (w_i d_{ni})^{-\theta}}{\Phi_n - T_i (w_i d_{ni})^{-\theta}}$, also as in the main text. Let C be the integral with the second half of A , given by

$$\begin{aligned}
C &:= \psi'_{niu} \int_0^\infty \int_{c_1}^\infty e^{-[\Phi_n - r_i T_i (w_i d_{ni})^{-\theta}] c_{2i}^\theta} g_{1ni, 2ni}(c_{1i}, c_{2i}) dc_{2i} dc_{1i} \\
&= \psi'_{niu} \left(\frac{r_i T_i (w_i d_{ni})^{-\theta}}{\Phi_n} \right) \\
&= \psi'_{niu} \pi_{ni} \tag{D.3}
\end{aligned}$$

Finally we combine the two components to compute ψ_{niu} ,

$$\psi_{niu} = C - B = \psi'_{niu} \pi_{ni} (1 - \psi'_{ni})$$

Since $\sum_{u \neq i}^N \psi'_{niu} = 1$, it is clear that

$$\begin{aligned}
\sum_{i=1}^N \psi_{ni} + \sum_{i=1}^N \sum_{u \neq i}^N \psi'_{niu} \pi_{ni} (1 - \psi'_{ni}) &= \sum_{i=1}^N \psi_{ni} + \sum_{i=1}^N \pi_{ni} (1 - \psi'_{ni}) \\
&= \sum_{i=1}^N \psi_{ni} + \sum_{i=1}^N \pi_{ni} - \sum_{i=1}^N \psi_{ni} \\
&= \sum_{i=1}^N \pi_{ni} \\
&= 1
\end{aligned}$$

To compute the distribution of the (restricted) markup conditional on the best and second-best supplier of a good to country n being from two different countries i and u , respectively, with some arbitrary number V of third countries we call v , $v \in 1, \dots, V$, we compute the distribution of the ratio $\frac{c_{1nu}}{c_{1ni}}$, given that $C_{1nu} < c_{2ni}$ and $C_{1nu} < c_{1nv}$ for all v . Note that all of the third-country terms cancel out, since each country's firms draw efficiency levels simultaneously from independent distributions. The conditional distribution for this ratio is

thus,

$$\begin{aligned}
h_{niu}(m) &= \int_0^\infty \int_{mc_{1i}}^\infty \int_{mc_{1i}}^\infty \dots \int_{mc_i}^\infty c_{1i} \frac{g_{1ni,2ni}(c_{1i}, c_{2i})}{1 - G_{2ni}(mc_{1i})} \frac{g_{1nu}(mc_{1i})}{1 - G_{1nu}(c_{1i})} \\
&\times \prod_{v \neq i, u} \frac{g_{1nv}(c_{1v})}{1 - G_{1nv}(mc_{1i})} dc_{1v} \dots dc_{1v} dc_{2i} dc_{1i} \\
&= \int_0^\infty \int_{mc_i}^\infty c_{1i} \frac{r_i(r_i - 1) [\theta T_i(w_i d_{ni})^{-\theta}]^2 c_{1i}^{\theta-1} c_{2i}^{\theta-1} e^{-T_i(w_i d_{ni})^{-\theta} c_{1i}^\theta} e^{-(r_i-1)T_i(w_i d_{ni})^{-\theta} c_{2i}^\theta}}{r_i e^{-(r_i-1)T_i(w_i d_{ni})^{-\theta} (mc_{1i})^\theta}} \\
&\times \frac{r_u T_u(w_u d_{nu})^{-\theta} \theta m^{\theta-1} c_{1i}^{\theta-1} e^{-r_u T_u(w_u d_{nu})^{-\theta} m^\theta c_{1i}^\theta}}{e^{-r_u T_u(w_u d_{nu})^{-\theta} c_{1i}^\theta}} dc_{2i} dc_{1i} \\
&= \int_0^\infty \int_{mc_i}^\infty (r_i - 1) [\theta T_i(w_i d_{ni})^{-\theta}]^2 c_{1i}^\theta c_{2i}^{\theta-1} r_u T_u(w_u d_{nu})^{-\theta} \theta m^{\theta-1} c_{1i}^{\theta-1} \\
&\times \exp^{-[T_i(w_i d_{ni})^{-\theta} + r_u T_u(w_u d_{nu})^{-\theta} (m^\theta - 1) - (r_i - 1)T_i(w_i d_{ni})^{-\theta} m^\theta] c_{1i}^\theta} e^{-(r_i-1)T_i(w_i d_{ni})^{-\theta} c_{2i}^\theta} dc_{2i} dc_{1i} \\
&= \int_0^\infty \theta T_i(w_i d_{ni})^{-\theta} c_{1i}^\theta r_u T_u(w_u d_{nu})^{-\theta} \theta m^{\theta-1} c_{1i}^{\theta-1} \\
&\times e^{-[T_i(w_i d_{ni})^{-\theta} + r_u T_u(w_u d_{nu})^{-\theta} (m^\theta - 1)] c_{1i}^\theta} \\
&= \frac{\theta T_i(w_i d_{ni})^{-\theta} r_u T_u(w_u d_{nu})^{-\theta} m^{\theta-1}}{[T_i(w_i d_{ni})^{-\theta} + r_u T_u(w_u d_{nu})^{-\theta} (m^\theta - 1)]^2}
\end{aligned}$$

The same process results in the distribution of the markups when both of the two best rival suppliers to n are from the same country i reveals that the distribution is the same as in autarky, with third-country terms cancelling out due to the independence of each country's distribution:

$$h_{ni}(m) = \frac{r_i(r_i - 1)\theta m^{-(\theta+1)}}{[(r_i - 1) + m^{-\theta}]}$$

The difference between $h_{niu}(m)$ and $h_{ni}(m)$ arises due to the fact that outcomes c_{1nu} and c_{1ni} come from the best of r_u draws from country u 's distribution, which is independent of the realizations of the r_i draws in country i from which the best firm in country i emerges. In contrast, the distribution of two ordered draws in country i is not independent and thus the difference between them can not be constructed from two independent distributions, rather from one joint distribution.

D Gains from trade

To close the model under autarky or trade, we use a market clearing condition. Let λ_D be the share of variable costs in expenditures for each country, given the vector of trade costs D that it faces when exporting. We can use the free entry condition to show that under autarky, λ_D equals $\frac{1}{1+\frac{\delta\beta f}{L}}$. Similarly, under free trade with symmetric countries, λ equals $\frac{1}{1+\frac{\delta\beta f}{L(1+\frac{N-1}{N})}}$. Given our unit cost specification, the share of labor in these variable costs is β . Then, the labor market clearing condition stipulates that payments to labor equal labor's share in production costs:

$$\omega_n L_n = \beta \lambda P_n Y_n.$$

We use the wage as our numeraire, $\omega \equiv 1$. Then, we can compare output under autarky with output under free trade in a world with N symmetric countries:

$$\frac{Y^t}{Y^a} = \left(\frac{1 + \frac{\delta\beta f}{L}}{1 + \frac{\delta\beta f}{L(1+\frac{N-1}{N})}} \right) \left(\frac{P^a}{P^t} \right).$$

The first term on the right-hand side is greater than one and reflects the fact that aggregate revenues and average firm profits fall under trade versus autarky because opening to foreign competition squeezes markups. We already know from Propositions 2 and 3a that the autarkic price level is greater than the price level under free trade. To find out how much greater, we must substitute in our formulas for the aggregate price level under autarky and free trade,²⁴

$$\frac{P^a}{P^t} = \left(\frac{(1 + (R - 1)\bar{m}^\theta) \left\{ \bar{m}^\theta r^{\frac{\sigma-1}{\theta}+1} + (r - 1)(\bar{m}^\theta - 1) \left[r(r - 1)^{\frac{\sigma-1}{\theta}} - (r - 1)r^{\frac{\sigma-1}{\theta}} \right] \right\}}{(1 + (r - 1)\bar{m}^\theta) \left\{ \bar{m}^\theta R^{\frac{\sigma-1}{\theta}+1} + (R - 1)(\bar{m}^\theta - 1) \left[R(R - 1)^{\frac{\sigma-1}{\theta}} - (R - 1)R^{\frac{\sigma-1}{\theta}} \right] \right\}} \right)^{\frac{1}{1-\sigma}}$$

Even under symmetry, the level of gains from trade clearly depends upon the number of domestic rivals before liberalization. In Figure 3, we show that they are lower for countries with a high level of contestability (r^a) to begin

²⁴If we do not normalize the wage ω to equal 1, this expression is the ratio of the real wage under trade, relative to the real wage under autarky.

with, as these countries already have lower average markups than their trading partners.

Trade increases contestability for any given market, which reduces markups, generating a gain from trade that is new to the BEJK framework, though not to models with alternative environments with imperfect competition, such as Bergin and Feenstra (2009), Devereux and Lee (2001), Melitz and Ottaviano (2008), and Rodriguez (2011). However, in our model, an increase in entry due to market scale effects can shift the distribution of efficiency levels among active firms to the right at the same time it changes the shape of the distribution of markups, a combination not captured by any of these papers. Thus, the increase in entry acts both as a technological advance and an increase in the intensity of competition. Geography, in the form of trade frictions, interferes with welfare gains from both Ricardian efficiency effects and contestability.

E Gains from free trade vs. autarky

To the degree that free trade results in an increased number of rivals for any particular market, it shifts the entire distribution of marginal costs to the left, similar to an innovation in available technology T . A particularly clean case occurs when countries are identical and that trade is costless, so that $T_i = T$, $\omega_i = \omega \equiv 1$, and $d_{ni} = 1$ for all i . Then we see that the distribution for the lowest unit cost among all potential suppliers to any country n in equation (12) reduces to the Weibull distribution

$$G_{1n}(c_1) = 1 - e^{-rN T c_1^\theta},$$

which is observationally equivalent to a world with $R = rN$ rivals who all draw from an underlying distribution that takes the same form as the distribution of cost parameters for any individual country, $G(c) = 1 - e^{-Tc^\theta}$.²⁵ The

²⁵The distribution of first order statistics for samples drawn from a Weibull distribution is also Weibull.

distribution of markups in this special case takes the form

$$\tilde{h}(m) = \frac{R(R-1)\theta m^{-(\theta+1)}}{[(R-1) + m^{-\theta}]^2}.$$

The implication is clear: trade has the same effect on the distribution of markups as increasing contestability and therefore reduces the number of firms charging the unconstrained markup and, all else equal, the aggregate price level, which takes the same form as under autarky, but with the total number of rivals for each market, R . Defining r^a as the number of rivals under autarky, we will show below that $R > r^a$. Proposition 3 follows directly from this increase in contestability under trade.

Proposition 4 *In a world with symmetric countries, free trade (a) increases the total number of rivals competing to supply a destination market, (b) reduces the aggregate price level, and (c) reduces the expected markup, as well as the probability that firms will charge the unconstrained markup.*

Proof.

Part a) To show that the total number of rivals under trade equals a number $R > r^a$, we use the open economy version of the free entry condition and a labor market clearing condition that takes the same form as the closed economy version in equation (7) for each country. If all countries are identical and trade is costless, we have

$$E_t \left[\sum_{s=0}^{\infty} (1-\delta)^{t+s} \left(1 + \frac{N-1}{N} \right) (P_{t+s}(j)Y_{t+s}(j) - C_1(j)Y_{t+s}(j)) | R \right] \quad (\mathcal{D}.4)$$

$$E_t \left[\sum_{s=0}^{\infty} (1-\delta)^{t+s} \left(1 + \frac{N-1}{N} \right) (P_{t+s}(j)Y_{t+s}(j) - C_1(j)Y_{t+s}(j)) | R + 1 \right] < f$$

The condition simplifies to

$$\begin{aligned} \frac{E[M^{1-\sigma}|r]}{E[M^{-\sigma}|r]} &\geq 1 + \frac{\beta\delta f}{(1 + \frac{(N-1)}{N})L} \\ \frac{E[M^{1-\sigma}|r]}{E[M^{-\sigma}||r+1]} &< 1 + \frac{\beta\delta f}{(1 + \frac{(N-1)}{N})L}. \end{aligned} \tag{D.5}$$

Since the left-hand side is decreasing in r and $\frac{\beta\delta f}{[1 + \frac{(N-1)}{N}]L} < 1 + \frac{\beta\delta f}{L}$, it is clear that the possibility of exporting strictly increases entry. Thus, leaping from autarky to free trade increases the number of rivals competing to produce any good ($R = Nr > r^a$), in addition to reducing prices by reallocating production to more efficient producers.²⁶

Part b) $E[C_2(j)^{1-\sigma}]$ under free trade and symmetry takes the exact form of its counterpart under autarky, only substituting $R > r^a$ for the number of rivals, making $E[C_2(j)^{1-\sigma}]$ greater than its counterpart under autarky. From the discussion in Lemma 1, we know that $E[C_{1n}(j)^{1-\sigma}]$ must also be greater than its counterpart under autarky. Therefore, $(P)^{1-\sigma}$ must be greater than its counterpart under autarky $(P^a)^{1-\sigma}$, revealing that the aggregate price level falls under trade: $P < P^a$.

Part c) It follows directly from the derivative in Proposition 1 and the fact that $R > r^a$ that the average markup falls under trade. Similarly, the likelihood of charging the unconstrained markup falls when opening to trade.

■

The results from Proposition 3 echo those of Bergin and Feenstra (2009) and Melitz and Ottaviano (2008), but now within the *homothetic* preference structure of Bernard, Eaton, Jensen, and Kortum (2003). Atkeson and Burstein (2007 and 2008) show the results in Parts (b) and (c) numerically, while de Blas and Russ (2011) demonstrate that having a large number of rivals under autarky reduces the impact of trade liberalization on markups. Note also that increasing the number of trading partners has a similar effect to increasing the number of rivals in any trading partner, seen in numerical so-

²⁶We assume the fixed cost of exporting is zero for simplicity, but one can also derive a reasonable restriction on the size of a fixed cost of exporting that preserves this result.

lutions calculated by Garetto (2012). Under costless trade, it does not matter how the rivals are distributed across countries. Markups respond as though all entrants worldwide compete on equal footing to be the low-cost supplier. As in classic studies of trade and endogenous market structure, geographic frictions here increase market power, dampening the effect of foreign industrial structure on domestic markups and prices.

In Figure 3, we show an estimate of the gains for a country opening to free trade with an arbitrary number of partners identical to itself. In this stylized case, the number of rivals competing to supply the domestic market under trade relative to autarky determines the increase in aggregate output. Each line in the graph represents an increase in the number of rival suppliers compared to autarky by 1, 5, or 10. The ratio $\frac{Y^t}{Y^a}$ on the vertical axis is computed as in Appendix E. The number of total competitors for a market under trade, R , lies on the horizontal axis. Where R equals 20 under trade, for instance, we see the gains from trade if there were 10, 15, or 19 domestic rivals under autarky. For example, a country which under free trade has 20 potential suppliers of a product experiences an increase in aggregate output of less than 3 percent if it already had 19 competitors under autarky, but greater than 40 percent, if it had only 10 competitors under autarky. Equation (9) implies that small countries, ones with smaller market size L , will have fewer rivals in each industry under autarky. Thus, gains from trade are greatest for small countries opening to trade, even if a pair of small countries establishes a free trade agreement, echoing a result in representative firm models such as Devereux and Lee (2001) under Cournot competition or Novy (2010) with translog preferences.

F Price adjustment in the closed economy

In the simple, closed-economy framework, a lower number of rivals leads to more frequent price changes in response to idiosyncratic shocks to marginal costs. The reason is clear from Figure 1. When r is low, more firms charge the unconstrained CES markup— their prices are not tightly bounded by the

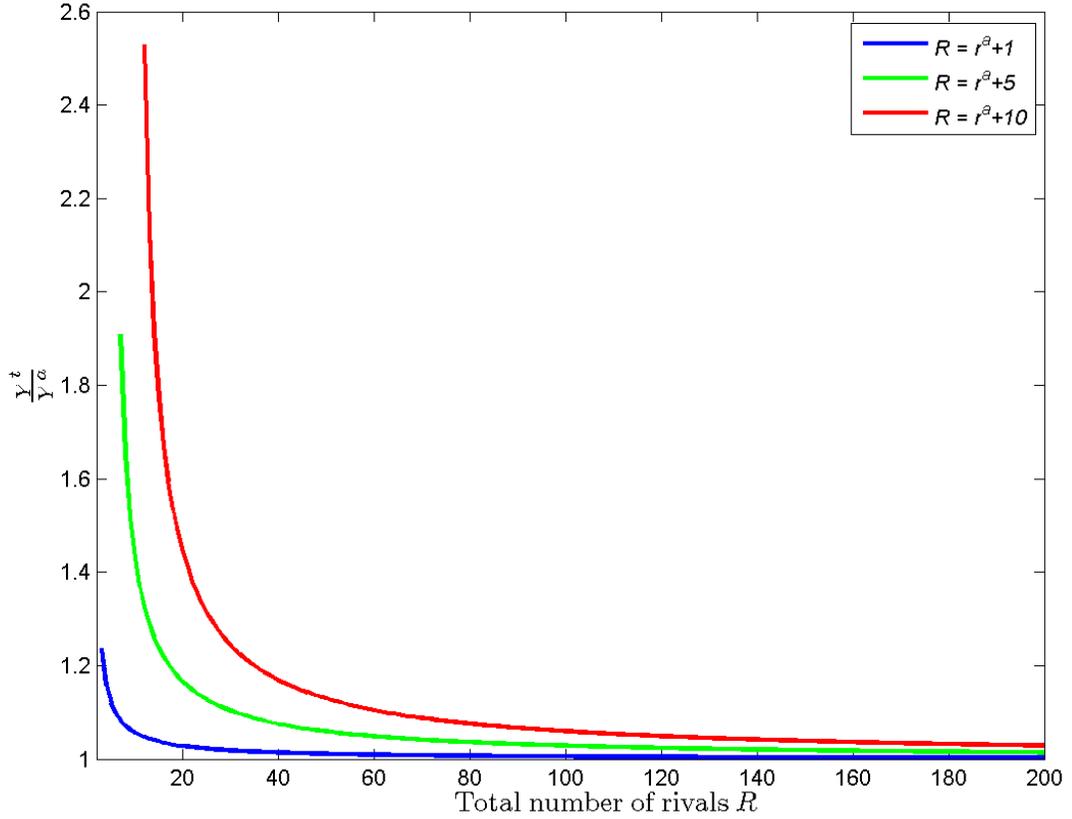


Figure 3: Gains from trade depend on the number of additional rivals

marginal costs of their next-best rival, so they are better able to pass on idiosyncratic increases in marginal cost to their customers. The fraction of firms that set their price equal to the marginal cost of the next-best rival are unable to do this. Since firms will not change prices in response to an idiosyncratic shock unless they charge the unconstrained markup, Figure 1 suggests that at least half of firms will never be able to adjust their prices upward *ever*, unless they experience a shock common to all rivals and which affects all rivals at exactly the same time. We apply a lognormally distributed idiosyncratic shock with the log of the shock being distributed $N(0,10)$, so that the standard

deviation of the shock is 10%.²⁷ After 1000 simulations, using the same parameters as in Figure 1, we compute that, all else equal, 73.3% of firms adjust their price in response to a shock when $r = 2$, while the figure falls to 64.8% when $r = 20$. This is consistent with results from Nakamura and Steinsson (2010), who find that no price changes are observed for 40% of products over the period 1982-2007, as well as Gopinath and Rigobon (2008) and Gopinath, Itskhoki, and Rigobon (2010), who report static prices for approximately 30% of their sample. The following corollary to Proposition 1 formalizes this result.

Corollary 3 *The probability that shocks to marginal cost are reflected in a firm's price is falling in the number of rivals r , as is price volatility.*

Proof. For some random i.i.d. shock ε to firm-specific marginal cost with probability density $\eta(\varepsilon)$ over the domain $(0, \bar{\varepsilon})$, we can compute the fraction of firms that will raise prices in response to an idiosyncratic increase in marginal costs. Suppose a shock occurs such that $\varepsilon > 1$, increasing the marginal cost for a particular active firm but not its rivals in the industry.

First, we note that only firms charging the unconstrained CES markup would be able to increase their prices, since firms setting prices bounded by the marginal cost of their next-best rival can not. Then, the probability that a firm will pass an idiosyncratic increase in marginal cost fully to buyers by raising its price is equal to the probability that the current price (\bar{m} times marginal cost) times the shock does not exceed the marginal cost of the next best rival,

$$\Pr [\bar{m}\varepsilon C_1(j) \leq C_2(j)] = \Pr \left[\frac{C_2(j)}{C_1(j)} \geq \bar{m}\varepsilon \right] = \Pr [M(j) \geq \bar{m}\varepsilon].$$

²⁷This is in line with calibration by Feenstra, Obstfeld, and Russ (2011) for micro-level shocks, drawing on empirical estimates by Basu, Fernald, and Kimball (2006) and Foster, Haltiwanger, and Syverson (2008).

Since the distribution of markups is independent of ε , we can compute this probability as

$$\begin{aligned} \Pr [M(j) \geq \bar{m}\varepsilon] &= \int_{-\infty}^{\infty} \int_{\bar{m}}^{\infty} h(\varepsilon m) \eta(\varepsilon) dm d\varepsilon \\ &= \int_{-\infty}^{\infty} \frac{r}{1 + (r-1)(\varepsilon \bar{m})^\theta} \eta(\varepsilon) d\varepsilon. \end{aligned} \quad (\text{D.6})$$

It follows from Corollary 1 that regardless of the probability distribution for ε , as long as the marginal cost shock is independent of the markup, the probability of full pass-through under autarky is decreasing in the number of rivals.²⁸ Multiplying ε above by some positive constant less than one, we see that the result is general to any degree of pass-through, not just full pass-through.²⁹ The intuition also applies for a downward cost shock, which is omitted here for the sake of brevity. In this case, all firms charging the unconstrained markup would have to lower their prices, otherwise their markup would rise above \bar{m} , implying marginal revenues less than marginal costs. Further, some portion of firms charging a price equal to $C_2(j)$ would also lower prices, namely those for whom leaving the price at $C_2(j)$ resulted in a markup greater than \bar{m} . Thus, downward adjustment is most likely when firms are more likely to have relatively inefficient rivals, which is the case when r is low. Note that having less complete and less frequent price adjustment in response to idiosyncratic shocks implies lower price volatility. ■

Limited pass-through of the marginal cost shock does not depend on the elasticity of substitution being infinite. In Figure F, we vary the elasticity of

²⁸That is, given the calculus used to prove Proposition 1, equation (D.6) implies that the probability of the markup being high enough to permit adjustment to positive price shocks is decreasing in the number of rivals r .

²⁹Our assumption that firms pay a fixed cost when they become active prevents the lowest-cost producer from having to adjust prices in response to temporary idiosyncratic shocks hitting its next-best rivals. The rivals will not find it profitable to try to undercut an existing producer unless they experience a transitory shock large enough to cover the entire fixed cost. We assume that the variance of costs is small enough that the likelihood of such a large shock is negligible.

substitution between varieties *within* an industry j from 4 (just greater than our calibrated value for σ) to 10,000 and show that pass-through by the low-cost producer is limited even at this low elasticity and decreases continuously as the elasticity of substitution increases, in response to a large shock ($\varepsilon = 1$) to marginal cost, given two rivals in each industry. The degree of pass-through varies the size of the shock and the number of rivals, but always displays this continuous decreasing behavior.³⁰

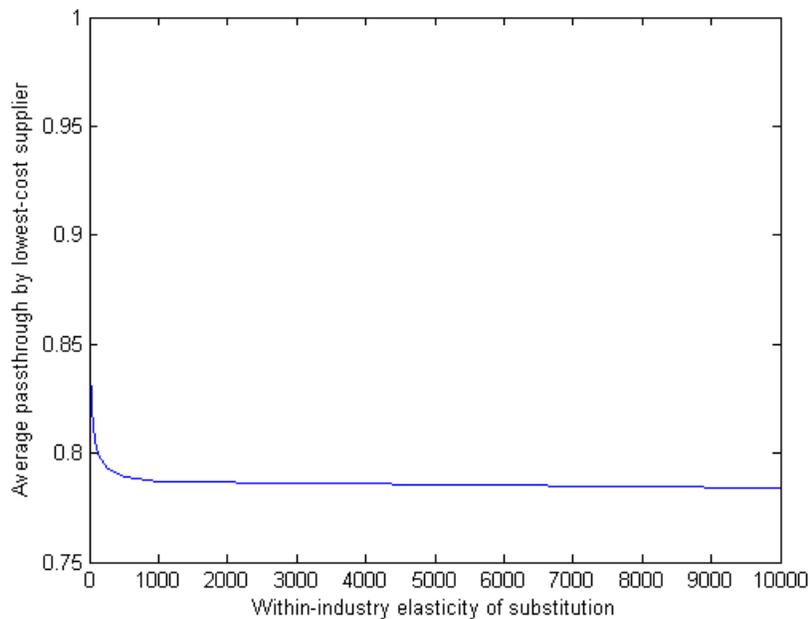


Figure 4: Passthrough declines continuously in the elasticity of substitution between rival's goods within an industry j

³⁰We are very grateful to Ariel Burstein for alerting us to the fact that this continuity can be shown numerically and Konstantin Kucheryavyy for sharing insights and code for a numerical illustration of markup behavior within his proof in Kucheryavyy (2012).