

**Teams of Rivals:
Endogenous Markups in a Ricardian World**

Beatriz de Blas, Universidad Autónoma de Madrid

Katheryn Niles Russ, University of California, Davis and NBER

First draft: 31 December 2009

This draft: 18 November 2010

Abstract

We show that an ostensibly disparate set of stylized facts regarding firm pricing behavior can arise in a Ricardian model with Bertrand competition. Generalizing the Bernard, Eaton, Jensen, and Kortum (2003) model allows firms' markups over marginal cost to fall under trade liberalization, but increase with FDI, matching empirical studies in international trade. We are able to mesh this dichotomy with the existence of pricing-to-market and imperfect pass-through, as well as to capture stylized facts regarding the frequency and synchronization of price adjustment across markets. The result is a well specified distribution for markups that previously could only be seen numerically and a way to quantify endogenous pricing rigidities emerging from a market structure governed by fierce competition among rivals.

The authors thank Paul Bergin, Andrew Bernard, Swati Dhingra, Thibault Fally, Cecilia Fieler, Ann Harrison, Samuel Kortum, Kala Krishna, James Markusen, Virgiliu Midrigan, Martine Quinzii, Adam Russ, Burkhard Schipper, Monika Schnitzer, Raphael Schoenle, Joaquim Silvestre, Ina Simonovska, Barbara Spencer, Deborah Swenson, Andreas Waldkirch, and especially Stefania Garetto, Robert Feenstra, and Peter Neary for helpful suggestions, as well as participants at the 2010 American Economics Association winter meetings, the 2010 European Trade Study Group, DEGIT XV at Goethe University of Frankfurt, and the 2010 Western Economics Association International meetings. Beatriz de Blas acknowledges financial support from ECO2008-04073 project of the Spanish MEC, and CCG08-UAM/HUM-4438.

Corresponding Email: knruss@ucdavis.edu

1 Introduction

Empirical studies in international macroeconomics and trade present a rich and seemingly disparate array of stylized facts regarding firm pricing behavior when selling for domestic and export sale. In particular, firms often price to market and do not fully pass changes in marginal costs and exchange rates through to foreign buyers. An impressive list of empirical studies in international trade¹ also demonstrates that trade liberalization is associated with firms charging lower markups over marginal costs when setting prices. Among other contributions, Atkeson and Burstein (2007) show that numerical simulation of a model based on price competition developed by Bernard, Eaton, Jensen, and Kortum (2003) (BEJK) can capture these three stylized facts. We derive explicit formulas for the distribution of markups to show why and when this is the case.

In addition, recent macroeconomic studies report longer lags in adjusting prices for export markets than when selling to domestic buyers, price changes that are more highly synchronized for domestic sales than for export sales, and higher rates of passthrough among firms that change their prices more frequently.² A small but burgeoning set of trade studies demonstrates that in contrast to trade in goods, foreign direct investment (FDI) generates higher markups or profit margins in target firms following foreign takeovers and greenfield FDI.

We build on insights from the numerical simulations of Atkeson and Burstein (2007 and 2008), Garetto (2009), and de Blas and Russ (2010) to present a generalized version of Bernard, Eaton, Jensen, and Kortum (2003) with free entry that can capture all of these stylized facts. With free entry, we explicitly introduce “teams of rivals” within the BEJK framework that generate the observed markup and pricing behavior.³ We are able to present tractable analytical solutions for the distribution of markups under autarky, trade, and FDI when the number of firms competing to supply the market is not fixed or filtered out. In short, we use free entry to build a standard Ricardian model of trade with a well specified distribution of markups that is consistent with observed domestic, export, and multinational pricing behavior.

¹Among these are Levinsohn (1993), Harrison (1994), Roberts and Supina (1996), Bottasso and Sembenelli (2001), Novy (2010), and Feenstra and Weinstein (2010).

²For instance, see Schoenle (2010), Haller and Fitzgerald (2010), and Gopinath and Itshokhi (2010).

³We take the term “team of rivals” to indicate a decision process governed by intense competition within a group from the bestseller *Team of Rivals* by Doris Kearns Goodwin (2006). Goodwin tells how in 1861, the newly elected President Abraham Lincoln installed all of his rivals for the Republican presidential nomination (as well as two Democrats) in his cabinet, producing policy amalgamated within a crucible of dissent rather than through harmonious coordination between like-minded individuals.

The intuition behind the results rests on a form of duopolistic competition modeled in the context of trade by BEJK. They use the Dixit and Stiglitz (1977) love of variety to limit the market shares of heterogeneous firms. Only one firm ends up supplying each good, similar to the Dixit-Stiglitz model of monopolistic competition. However, the supply side of the market for each good in the continuum is characterized by a fierce competition among a group of rival firms competing to be the sole producer. The most efficient firm in this group ultimately becomes the sole supplier of that particular good, but only because it beats back its competitors by underselling them: it can not charge a price higher than the marginal cost of its next best rival. The low-cost supplier can not automatically charge the Dixit-Stiglitz markup despite the constant elasticity of substitution (CES) overlaying the demand for their goods. Rather, if the competition is sufficiently strong, the best firm must charge a price equal to the marginal cost of its next best rival. The CES markup becomes the *maximum* markup that it might charge over its own marginal cost without jeopardizing profits; it is no longer the default markup.

The challenge is to fit this competitive effect into a tractable framework which also captures observed reductions in markups under trade, as the distribution of markups does not change with trade liberalization in the BEJK model. To do this, we generalize the BEJK setting to allow for an endogenous number of rival competitors in each industry. This entry 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 industry—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 lowers the expected value of the first order statistic for marginal costs. It also shifts the mass in the distribution of markups toward the lower end. Therefore, increased entry reduces the aggregate price level under autarky.

Openness to trade has a first-order effect that is similar to increasing domestic entry under autarky. Higher geographic frictions impede trade as in BEJK, but also increase suppliers’ market power, allowing them to charge higher markups conditional on the trade cost. Trade can also increase efficiency directly by encouraging more entry. In our order-statistic framework, increasing entry changes the entire shape of the distribution of efficiency levels among active firms, moving the mass toward the upper end of the distribution and depressing markups even further. Both because the distribution of markups is itself a function of trade costs and because the number of rivals alters the entire distribution of efficiency levels

and markups, the model includes gains from trade not subject to the critique in Arkolakis, Costinot, and Rodriguez-Clare (2010).

By introducing different shocks into the model, we are also able to explain stylized facts with regard to the frequency of price changes in domestic versus export sales. We can quantify the degree of price rigidity conditional on the intensity of competition. Firms with greater market power at the point of sale are able to adjust their prices more frequently in response to idiosyncratic cost shocks because their price is not bounded as strongly by their next-best rival. Industries where sectoral shocks are more prevalent than idiosyncratic shocks also will experience more frequent price changes, as shocks that affect the efficiency of all rivals do not affect markups. In the empirical literature, the connection between the frequency of price changes and market concentration is generally positive not robust. In our model, this lack of robustness would arise if sectoral shocks were less prevalent than idiosyncratic shocks in industries with greater market concentration (a lower number of rivals to produce each good).⁴

Cross-border takeovers increase markups and reduce prices in the host country, but foreign-owned firms charge higher markups than domestic firms. When a cross-border takeover transfers superior foreign technology to a local target firm, the target becomes even more efficient than its next best local competitor, increasing the markup. At the same time, the marginal cost of the next best local rival has not changed, so the acquired firm can not increase its price and may even end up cutting it, passing on some of its technological efficiency gains to consumers. A similar but slightly more subtle intuition applies in the case of greenfield FDI.

We also show the importance of the pre-existing level of domestic competition when evaluating the impact of FDI and trade on markups and prices. Higher domestic entry results in fewer firms charging the maximum markup, leaving less room for foreign competitors to challenge high profit margins in the domestic market. Thus, trade liberalization and FDI have a bigger effect on markups and prices in countries or industries with small pools of rivals, a situation that we call low contestability. The number of rivals in the market at the point of sale is critical in determining the degree of passthrough, as Atkeson and Burstein (2007 and 2008) and Garetto (2009) find in numerical simulations within a 2-country model. Atkeson and Burstein (2007) note that to simulate the model one must choose the number of rivals in each industry, an issue not touched upon in BEJK. By generalizing BEJK to include the number of rivals, we provide formulas to characterize the degree of passthrough based on the

⁴We are grateful to Claudia Buch for pointing out that although we can not measure the degree of entry, since it involves one active firm and an unobserved number of latent rivals, the level of market concentration is a sufficient statistic to capture it. Thus, previous empirical analyses of pricing behavior and market concentration are quite relevant to our model.

degree of competition. Using these formulas, we are able to show that imperfect competition also may explain differences in the relative frequency of price adjustments across markets.

In summary, we build a model that reconciles stylized facts on markups and prices from a diverse set of studies. The goal is to find an algebraic representation for the distribution of markups that permits us to produce markups that fall under trade liberalization, rise with foreign ownership, and are positively correlated with market concentration. We want markup behavior to result in prices that are set differently across countries, yield sluggish (less frequent) export price changes and incomplete pass-through for exports when marginal costs change— with greater pass-through for the most frequent price adjusters and which in some cases generates more frequent price changes for firms in markets with higher concentration. We want to accomplish this within a tractable framework allowing us to quantify algebraically and endogenously the proportion of firms likely to change their prices at any given time.

Section 2 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 difference between the effect of sectoral versus idiosyncratic shocks on the size and frequency of price changes. Section 3 considers the implications of trade in goods without FDI for these distributions. In Section 4, we contrast this case of trade in goods only with that of FDI when no goods are traded, illustrating the boost that FDI gives to markups in the host country. We briefly discuss why intra-firm trade might motivate strategic takeovers that would increase markups in the case of export-platform FDI. We save discussion of empirical studies of markups and pricing until Sections 2, 3, and 4 to show the relevance of our theoretical results. Section 5 concludes.

2 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 simply 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,

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

with elasticity of substitution σ 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

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

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.⁵ Thus, we assume that an endogenous number of firms r each draw an efficiency parameter from a distribution given by

$$F(z) = 1 - e^{-Tz^{-\theta}}.$$

We assume that $T > 0$ and also that the shape parameter, θ , is no less than $\sigma - 1$ to ensure the existence of certain moments of interest below. 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

⁵In 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 (lowest cost) idea from each period then becomes Weibull. More generally, BEJK (2003) state that if firms draw from this frontier distribution, the lowest cost (the first order statistic) takes on a Weibull distribution. We note that the first order statistic of a Weibull distribution is also Weibull, so the underlying distribution from which firms are drawing their cost parameters can be reasonably modeled as Weibull, as we do here. Costs and efficiency levels are simply the inverse of one another, so that assumption implies that efficiency levels take on an inverse Weibull distribution. The Fréchet distribution is isomorphic to an inverse Weibull distribution and we can equivalently describe the distribution from which firms draw their efficiency levels as Fréchet. We do only this to match the model with the EK and BEJK terminology— to show that our model is a generalization of their framework.

parameter,

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

which accounts for both the cost of the composite input $w = \omega^\alpha p^{1-\alpha}$, with ω being the labor wage rate and p the cost of a bundle of intermediate goods, and any frictions involved in sending intermediate goods to the assemblers of the final good, $d \geq 1$. 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 r rivals draw an efficiency parameter hoping to be the low-cost supplier of industry j , the distribution of the lowest cost $C_1(j)$ is⁶

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

We assume that d equals one under autarky in this section and for domestic sales in the open economy in Section 3.

2.1 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 that provides the maximum markup possible subject to demand—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

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

With this formula for the markup, we compute the expected output-weighted price for

⁶See Rinne (2009), p.237 for derivation. EK and BEJK simplify their frameworks by using the underlying assumption that the number of firms competing to be the low-cost supplier in any industry is a random variable with a Poisson distribution. It elegantly drops from the analysis, though one could possibly interpret an increase in the technology parameter T in their model as an increase in the mean number of competitors because T enters their Fréchet distribution of surviving ideas through the Poisson exponent.

any good j in several steps. First, note that the price for good j , $P(j)$, is given either by

$$P(j) = \frac{C_2(j)}{C_1(j)}C_1(j) = C_2(j) \quad \text{for} \quad \frac{C_2(j)}{C_1(j)} \leq \bar{m},$$

or by

$$P(j) = \bar{m}C_1(j) \quad \text{for} \quad \frac{C_2(j)}{C_1(j)} \geq \bar{m}.$$

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. Rinne (2009, p.243) provides a formula for the distribution of $\frac{C_1(j)}{C_2(j)}$. We apply a Jacobian transformation to obtain the distribution of $\frac{C_2(j)}{C_1(j)}$. Assuming that the frontier distribution of efficiency parameters is identical for every industry j , for values of the markup less than \bar{m} the probability density of the markup is given by

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

with a mass point at \bar{m} , since all cases where $\frac{C_2(j)}{C_1(j)}$ exceeds \bar{m} are assigned a value of \bar{m} .

Like the distribution of markups given in BEJK, this distribution is entirely independent of $C_1(j)$ and $C_2(j)$. In fact, for very large r , we have

$$\lim_{r \rightarrow \infty} h(m) = \theta m^{-\theta-1}, \quad (3)$$

a Pareto density for markups identical to the one in BEJK. However, because we explicitly include the number of rivals r — rather than elegantly integrating it out to focus on the role of gravity in a Ricardian setting as they do— we see that the distribution of markups is directly affected by the number of firms competing to be the low-cost supplier, a measure which we call contestability, drawing on work by Claessens and Laeven (2004) and de Blas and Russ (2010). One can conceptualize r as an exogenous policy parameter, as in the numerical analysis by Atkeson and Burstein (2007 and 2008) and de Blas and Russ (2010), or endogenize it using a free entry condition as in Melitz (2003). The key is that unlike models using a Pareto distribution of firm efficiency parameters, the degree of entry embodied in r changes the shape of the entire distribution of markups, costs, and firm size. Because \bar{m} is the upperbound imposed by the nested CES structure, integrating $h(m)$ over values from

\bar{m} to ∞ gives the probability that a firm will charge the maximum CES markup,

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

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 contestability r under autarky.

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

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

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, $E[M(j)]$ is decreasing in r . ■

Corollary 1: The fraction of firms charging the maximum markup is decreasing in contestability r under autarky.

Proof: The proof of Proposition 1 for any $1 \leq m' \leq \bar{m}$ also applies for \bar{m} and markups are set equal to \bar{m} whenever $M(j)$ would be greater than \bar{m} without the restriction of the CES upperbound. ■

As the number of rivals in an industry j increases, both the average markup and the probability that firms will be able to charge the maximum markup falls— increased rivalry squeezes markups. Intuitively, as we show below, the result emerges because on average, increasing the number of rivals diminishes the difference between the costs of the two best potential suppliers. Further, equation (2) reveals that our team of rivals is crucial in determining the shape of the distribution as long as the number of rivals is finite. To illustrate, Figure 1 shows the distribution of markups if r equals its minimum value of 2, versus 20, the number of rivals chosen by Atkeson and Burstein (2007) by calibration to match U.S. industry concentration. We use $\theta=3.6$ and $\sigma=3.79$, as estimated by BEJK. The number of firms charging the maximum markup falls drastically, from one-half to just over one-third. We will discuss the implications of this statistic for price rigidity but first, we use the distribution of

markups to compute the aggregate price level.

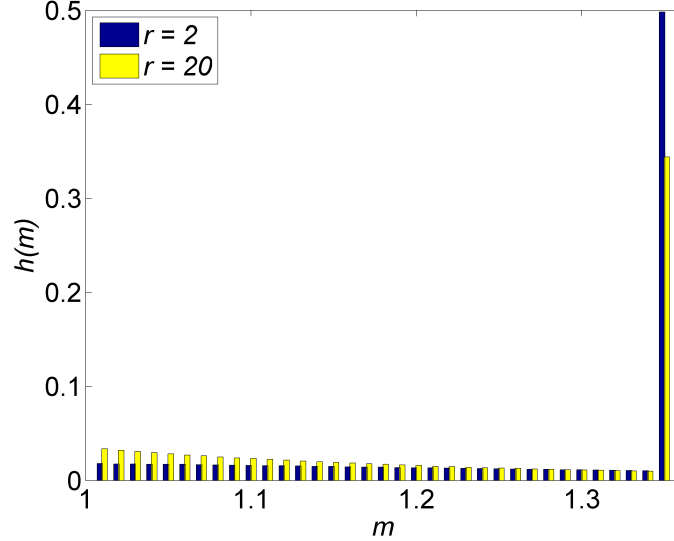


Figure 1: Increasing the number of rivals reduces markups

2.2 The distribution of prices

As shown in de Blas and Russ (2010), the joint distribution for the first and second order statistic also contains the contestability measure r :

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

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) from 0 to ∞ . We find that increasing the number of rivals leads, on average, to lower costs in the industry. We compute a particular moment of interest, $1 - \sigma$, for the first and second order statistics that will be used below to construct the aggregate price level

$$E[C_1(j)^{1-\sigma}] = \int_0^\infty c_1^{1-\sigma} g_1(c_1) dc_1 = (r T w^{-\theta})^{\frac{\sigma-1}{\theta}} \Gamma\left(\frac{1-\sigma+2\theta}{\theta}\right)$$

$$E[C_2(j)^{1-\sigma}] = \int_0^\infty c_2^{1-\sigma} g_2(c_2) dc_2 = r [T(r-1) w^{-\theta}]^{\frac{\sigma-1}{\theta}} \Gamma\left(\frac{1-\sigma+2\theta}{\theta}\right).$$

Taking the derivative of $E[C_1(j)^{1-\sigma}]$ and $E[C_2(j)^{1-\sigma}]$ with respect to r , we find that the

$(1 - \sigma)$ th moment of the second-lowest cost increases in r faster than the same moment for the lowest cost.

$$\frac{\frac{\partial E[C_2(j)^{1-\sigma}]}{\partial r}}{\frac{\partial E[C_1(j)^{1-\sigma}]}{\partial r}} = \left(\frac{r}{r-1}\right)^{\frac{\theta-(\sigma-1)}{\theta}} \left[r \left(1 + \frac{\theta}{\sigma-1}\right) - \frac{\theta}{\sigma-1} \right] > 1.$$

In other words, the second-lowest cost is falling in r faster than the lowest cost, demonstrating how increases in contestability can reduce markups. Because the distribution of the markup is independent of outcomes for the individual order statistics $C_1(j)$ and $C_2(j)$, we can compute the expected price $P(j)^{1-\sigma}$ as

$$E[P(j)^{1-\sigma}] = \Pr[M(j) > \bar{m}] \bar{m}^{1-\sigma} E[C_1(j)^{1-\sigma}] + \Pr[M(j) \leq \bar{m}] E[C_2(j)^{1-\sigma}],$$

which is also increasing in r . Since firms in all industries draw from the same underlying distribution, using the law of large numbers one can calculate the aggregate price level,

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

Proposition 2: The aggregate price level is decreasing in the level of contestability r under autarky for $\theta \geq 1$ and $\theta \geq \sigma - 1$.

Proof: Intuitively, this is true because an increase in r shifts the distribution of markups to the left at the same time it reduces the first- and second-lowest unit costs on average. More rigorously, taking the derivative of $p^{1-\sigma} = E[P(j)^{1-\sigma}]$ yields

$$\begin{aligned} \frac{\partial [p^{1-\sigma}]}{\partial r} &= \Pr[M(j) \geq \bar{m}] \frac{\partial E[(\bar{m}C_1(j))^{1-\sigma}]}{\partial r} + (1 - \Pr[M(j) \geq \bar{m}]) \frac{\partial [C_2(j)^{1-\sigma}]}{\partial r} \\ &\quad - \frac{\partial \Pr[M(j) \geq \bar{m}]}{\partial r} (E[C_2(j)^{1-\sigma}] - E[(\bar{m}C_1(j))^{1-\sigma}]). \end{aligned}$$

The first two terms on the right-hand side are positive, while it has been shown above that the probability of charging the maximum markup is falling in r , making its partial derivative negative. The derivative $\frac{\partial [p^{1-\sigma}]}{\partial r}$ is positive and $p^{1-\sigma}$ is increasing in r as long as $E[(\bar{m}C_1)^{1-\sigma}] \leq E[C_2^{1-\sigma}]$. Using the expressions for $E[C_2(j)^{1-\sigma}]$ and $E[C_1(j)^{1-\sigma}]$ derived above, we see this is possible whenever $r^{-\frac{\theta-(\sigma-1)}{\theta}} \leq 1$. Assuming that there are at least two competitors in each industry, this sufficient condition is satisfied by our assumption that $\theta \geq \sigma - 1$. ■

The assumption that $\theta \geq \sigma - 1$ is akin to the assumption in Ghironi and Melitz (2005)

that the Pareto shape parameter be at least as large as $\sigma - 1$ to ensure the existence of the moments used to compute the aggregate price level and plays the exact same role here. In economic terms, the assumption means that the dispersion in firm cost parameters can not be too great relative to the need for a variety of inputs into final goods production. (See Anderson, de Palma, and Thisse (1987) for a discussion in the context of consumption.) Otherwise, final goods producers would be too sensitive to variations in price across intermediate goods and one very efficient intermediate good manufacturer would absorb the entire market share, leading to a degenerate distribution of firm efficiency levels. Thus, under autarky, the aggregate price level p is decreasing in the number of rivals r under standard constraints on the parameters θ and σ .

2.3 The number of rivals

In this section, we present the free entry condition that pins down the solution for r . The distribution of costs is the same for firms competing in any market niche j within a particular country, so from this point we drop the index j . We follow the existing literature by using P and Y to refer to the price of an individual good and the output of an individual firm. The lowercase letters p and y refer to the aggregate price level and aggregate output.

A free entry condition limits the number of rivals r competing to be the low-cost producer of any good j . Following Melitz (2003), we assume that there is a uniform probability of death, $0 < \delta < 1$, in every period and that a startup cost must be paid by the active supplier of each good in the first period that the firm begins supplying the market. This startup cost is equal to a fraction of output, $0 < \kappa < 1$, in the first period of active production.⁷ In equilibrium, the number of rivals must be such that the expected present discounted value of output for an active producer equals the startup cost,

$$E_t \left[\sum_{s=0}^{\infty} \delta^{t+s} (P_{t+s} Y_{t+s} - C_1 Y_{t+s}) \right] \equiv \kappa C_1 Y_t,$$

yielding the steady-state expression

$$\frac{E[M^{1-\sigma}]}{E[M^{-\sigma}]} = (1 + \delta\kappa).$$

⁷Demidova, Kee, and Krishna (2006) report the average revenues of Bangladeshi textile manufacturers as just over \$4million, while Cherkashin, Demidova, Kee, and Krishna (2009) estimate the size of their upfront entry cost as \$408,500, suggesting a κ of 10%. It is likely to be much higher for capital-intensive industries. Dunne, Roberts, and Samuelson (1988) report average exit rates (δ) among US firms between 42% and 50% in a sample from 1963 and 1982. Thus, we would expect $1 + \delta\kappa$ to be no less than 1.05 and much higher for capital-intensive industries.

In Appendix A.1, we show that the left hand side is decreasing in r , resulting in a unique equilibrium solution. More intuitively, if an infinite number of entrepreneurs were to compete in the industry, the markup would fall to 1, with zero marginal profit and negative expected net profit because the most efficient producer would still have to pay the startup cost. Thus, in equilibrium, there must be some finite number of rivals such that expected net profit is zero.

Recall that the probability of forced exit is independent of firm efficiency and that the distribution of the markup is independent of the distribution of costs,⁸ so the free entry condition reduces to

$$E[\ln M] \geq \ln(1 + \delta\kappa).$$

Noting that $E[M] \geq \ln E[M]$ and using Jensen's inequality, we have

$$E[M] \geq \ln E[M] \geq E[\ln M] \geq \ln(1 + \delta\kappa) \tag{5}$$

Proposition 1 tells us that the mean markup, $E[M]$, is decreasing in the number of rivals. This means that rivals will keep “entering” the industry (i.e., draw a productivity parameter) as long as the markup a rival expects to charge, if it is the low-cost supplier, generates expected profits equal to the discounted startup cost. Since the mean markup is decreasing in the number of rivals r , it is clear that the maximum number of rivals is decreasing in the fixed cost parameter κ and the exit rate δ .

The distribution of the markup derived above does not yield a closed-form solution for the expected markup $E[M]$ or for the expected log markup, $E[\ln M]$. However, we can determine an upper- and lowerbound 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

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

⁸To see this, note that the cost parameters C_k do not enter into the expression for $h(m)$ for $k \in \mathbb{N}$.

Using a generalized version of Chebyshev's inequality⁹, we can characterize a lowerbound for the number of rivals:

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

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 (5) that the expected log gross markup $E[\ln M]$ must be at least as large as the gross log per-period cost of production, $\ln(1 + \delta\kappa)$, producing an upperbound for r . Thus, we know that r lies within these bounds:

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

When either the fixed cost parameter κ or the probability of default δ increase, the maximum number of rivals falls. Note also from equation (1) that increasing the number of rivals acts as a positive technology shock, just as increasing T would.¹⁰

2.4 Market concentration and pricing behavior

In Appendix B, we show that having a lower number of rivals not surprisingly results in higher market concentration. What is less obvious is that a lower number of rivals leads to more frequent price changes. The reason is clear from Figure 1. When r is low, more firms charge the maximum CES markup— their prices are not tightly bounded by the 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. Thus, we would expect a positive relationship between market concentration and frequency of price changes, as shown by Schoenle (2010, Tables 19, 20, and 21).¹¹ In addition, since firms will not change prices in response to an idiosyncratic

⁹See 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 $kP[g(X) \geq k] \leq E[g(X)]$.

¹⁰Normalizing the wage ω to equal 1, the model is easily closed by imposing a labor market clearing condition. We do this in Appendix C.1 to compare output levels under autarky versus free trade.

¹¹At the end of this section, we explain why the relationship between concentration and frequency may not always be robust.

shock unless they charge the maximum markup, Figure 1 suggests that at least half of firms will never be able to change their prices *ever*, unless they experience a shock common to all rivals and which affects all rivals at exactly the same time. 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.

Corollary 2: Increased contestability reduces firms' ability to adjust prices in response to idiosyncratic shocks.

Proof: For some random i.i.d. shock ε to firm-specific marginal cost with probability density $v(\varepsilon)$, we can compute the fraction of firms that will raise prices in response to an idiosyncratic increase in marginal costs. Suppose that we assume an increase in ε increases marginal cost for an active firm, but not its rivals. First, we note that only firms charging the maximum 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,

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

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

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

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.¹² 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.¹³ The intuition also applies for a downward cost shock, which is omitted

¹²That is, given the calculus used to prove Proposition 1, equation (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 .

¹³Our assumption that firms pay a fixed portion of first-period output when they become active prevents

here for the sake of brevity. In this case, all firms charging the maximum 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 would also lower prices, those for whom leaving the price at C_2 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. ■

2.5 The type of shocks

Taken together, Bils and Klenow (2004), Gopinath and Itskhoki (2010), and Hellerstein and Goldberg (2010) suggest that while large firms change prices more frequently, the link between frequency and market concentration is not robust. In our model, changes in marginal costs that affect an entire sector j — for instance, due to fluctuations in the labor wage or some sectoral efficiency parameter— in principle could apply equally and at the exact same time to all rivals in sector j , active and latent, and have no impact on markups at all. Sectoral shocks would be passed fully to the customer, resulting in more frequent price fluctuations than idiosyncratic shocks, where a fraction of firms is unable to raise prices if costs increase. The weak link between concentration and frequency of price changes might be explained if firms in highly concentrated industries are more subject to idiosyncratic rather than sectoral shocks. Empirical studies indicate that idiosyncratic shocks are likely to be prevalent and economically important: Gabaix (2010) 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 hitting plants have a standard deviation 5 times as large as that of industry-level productivity shocks. So idiosyncratic shocks are likely to be important from a macroeconomic perspective.

The discussion of pricing rigidities in the face of idiosyncratic shocks also pertains to shocks affecting the wider macroeconomy. A macroeconomic productivity shock may take time to filter through all sectors and to each firm in every sector, possibly never reaching some firms at all. Even a widespread technological advance could generate changes in relative costs consistent with a degree of price rigidity in some sectors, particularly if there are fixed costs involved in upgrading production techniques to take advantage of a scientific breakthrough

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.

or a sudden increase in labor efficiency. The same subtleties apply to monetary policy. A change in the Fed Funds rate may affect sectors and individual firms within sectors differently or with varying lags according to their level of capital intensiveness, the transaction costs involved in their mode of access to external credit, or the monopoly power involved in their individual relationships with banks. It is beyond the scope of this paper to model differential effects of macroeconomic shocks across sectors and firms. Our point is merely that to the degree macroeconomic shocks generate some idiosyncratic changes in relative costs between a first- and second- best rival, our formula for the distribution of markups suggests that we are likely to observe sluggish or severely limited price adjustment by a nontrivial fraction of firms under Bertrand competition regardless of the source of shocks. Below, we show analytically that exporters on average will have less ability to adjust their prices than domestic producers.

3 Trade in goods

Here we demonstrate that trade 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 for all trading partners. This squeeze on markups generates a gain from trade that is new to the BEJK framework. Trade also invites increased domestic entry (a higher r), which reduces markups, generating a second gain from trade that is new to the BEJK framework, though not to Bergin and Feenstra (2008), Melitz and Ottaviano (2005), or Rodriguez (2010). Furthermore, an increase in entry by itself can shift the distribution of efficiency levels among active firms to the right, an effect not captured by either BEJK, Bergin and Feenstra (2008), Melitz and Ottaviano (2005), or Rodriguez (2010). The increase in entry acts both as a technological advance and an increase in contestability. Thus when there are no multinationals present, trade always reduces the prices of imported goods relative to autarky *and* reduces the prices of domestically produced goods as increasing the number of domestic rivals increases average efficiency. We will also show how geography, in the form of trade frictions, interferes with all three of these sources of welfare gains.

Again, we drop the industry index j to simplify notation. This is possible because in both autarky and the open economy, the same distributions apply to each industry within a particular country. We add the subscript n to the notation $C_k(j)$, $g_k(c)$, etc., from the autarkic case to denote the supplier of a good j to country n from country i with the k^{th} lowest cost, so that the corresponding notation becomes C_{kni} , $g_{kni}(c)$, etc., where k is a positive integer. The variable C_{kn}^{TR} denotes the supplier to country n with the k^{th} lowest cost under trade, with corresponding cumulative distribution and probability density $G_{kn}^{TR}(c_k)$ and $g_{kn}^{TR}(c_k)$. We assume that Eaton and Kortum's no arbitrage condition for trade costs holds:

$d_{ni} < d_{ui}d_{nu}$.

$G_{1n}^{TR}(c_1)$ is 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 potential supplier (including the one in the home country) has a marginal cost greater than c_1 . The cumulative distribution for low-cost suppliers under trade is thus

$$\begin{aligned} G_{1n}^{TR}(c_1) &= \Pr[C_{1n}^{TR} \leq c_1] = 1 - \prod_{i=1}^N [1 - G_{1ni}(c_1)] \\ &= 1 - e^{-\Phi_n^{TR} c_1^\theta}, \end{aligned} \quad (7)$$

where $G_{1ni}(c_1)$ is the distribution of low-cost suppliers to n from country i , $\Phi_n^{TR} = \sum_{i=1}^N T_i (w_i d_{ni})^{-\theta} r_i$, 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:

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

3.1 Geography and markups

We can compute the full distribution of markups under costly trade with asymmetric countries. 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 . The probability that this occurs is

$$\begin{aligned} \psi_{ni} &= \int_0^\infty g_{2ni}(c_2) \prod_{u \neq i}^{N-1} [1 - G_{1nu}(c_2)] dc_2 \\ &= \frac{r_i (r_i - 1) T_i (w_i d_{ni})^{-\theta}}{\Phi_n^{TR}}. \end{aligned} \quad (8)$$

The full distribution of markups in country n under trade, $h_n^{TR}(m)$, is given by

$$h_n^{TR}(m) = \sum_{i=1}^N \psi_{ni} h_i(m) + \sum_{i=1}^N (1 - \psi_{ni}) h_{ni}^{TR,D}(m),$$

where $h_i(m)$ is simply the distribution of markups in country i as given under autarky and $h_{ni}^{TR,D}(m)$, derived in Appendix C, is the distribution of markups when the first- and second-best suppliers are from two different countries, i and $u \neq i$, respectively.

The distribution $h_{ni}^{TR,D}(m)$ itself is a compound distribution,

$$h_{ni}^{TR,D}(m) = \sum_{i \neq u}^{N-1} \psi_{niu} h_{niu}(m),$$

where $\psi_{niu} = \frac{r_u T_u (w_u d_{nu})^{-\theta}}{\Phi_n^{TR} - r_i T_i (w_i d_{ni})^{-\theta}}$ is the probability that the second-best rival to country i when supplying a particular good to country n is from country $u \neq i$, and $h_{niu}(m)$ is the distribution of markups given that the first- and second-best suppliers of a good to country n come from these two countries,¹⁴

$$h_{niu}(m) = \int_0^{\infty} c_1 g_{1ni}(c_1) g_{1nu}(m c_1) d c_1.$$

We derive this distribution in Appendix C, but here focus on the resulting probability that the supplier charges the maximum markup when its next-best rival is an exporter in a different country,

$$\Pr[M_{niu} \geq \bar{m}] = \frac{r_i T_i (w_i d_{ni})^{-\theta}}{r_i T_i (w_i d_{ni})^{-\theta} + r_u T_u (w_u d_{nu})^{-\theta} \bar{m}^{\theta}}. \quad (9)$$

One can see immediately that the supplier to country n exporting from country i will be more likely to charge the maximum markup when its next-best rival (1) resides in a country far from the destination country n (high d_{nu}), or (2) 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 maximum markup. If all countries are identical, this probability that a firm in i supplying country n charges the maximum markup when its next best rival is in a third country $u \neq i$ reduces to $\frac{1}{1+\bar{m}^{\theta}}$, which is easily shown to be lower than the probability under autarky in equation (4). Thus, the only way that markups would increase under trade is if the home country n opened its borders to trade with a world dominated by one country that was much closer than other trading partners (low d_{ni}) and were 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 (8) implies that reducing the trade cost d_{ni} for one particular country i increases the probability that a foreign suppliers from i will be able to charge their full autarkic markup when selling to country n , yielding an important argument for multilateral trade negotiations.

¹⁴See Mood, Graybill and Boes (1974, pp.187-88) for a description of the transformation method used to find the distribution of the quotient of two random variables.

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

$$(p_n^{TR})^{1-\sigma} = E[(P_n^{TR})^{1-\sigma}] = \Pr [M_n^{TR} > \bar{m}] \bar{m}^{1-\sigma} E[(C_{1n}^{TR})^{1-\sigma}] + \Pr [M_n^{TR} \leq \bar{m}] E[(C_{2n}^{TR})^{1-\sigma}]$$

and note that

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

which is strictly greater than its counterparts 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} :

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

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. ■

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, $h_n^{TR}(m)$, reveals that they 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 cost on to the foreign consumer. We demonstrated that the probability of charging the maximum 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 Rodriguez-Clare (2010) 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. 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.¹⁵ Put more simply, the gains from trade liberalization can not be inferred from the value of aggregate flows alone because liberalization reduces markups, distorting the relationship between the trade cost and observed expenditures.

As in the variable-markup frameworks of Bergin and Feenstra (2000), Melitz and Otaviano (2005) and Feenstra and Weinstein (2009), entry changes the effective elasticity of demand (the price-elasticity of marginal revenues), even though the Dixit-Stiglitz elasticity of demand governing the upperbound for the markup is a constant. 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 in Appendix C.1, but save detailed analysis of gains from 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.

To the degree that trade induces new entry (increased r), 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.

Proposition 3: In a world with symmetric countries, free trade (a) reduces the aggregate price level, (b) increases the number of domestic rivals in each country, and (c) reduces the probability that firms will charge the maximum markup.

Proof: To illustrate more intuitively how trade affects the full distribution of markups, it is useful to suppose for a moment that countries are identical and 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 (7) reduces to the Weibull distribution

$$G_{1n}^{TR}(c_1) = 1 - e^{-rNTc_1^\theta},$$

which is observationally equivalent to a world with $R = rN$ rivals who all draw from an

¹⁵The reason is clear when one writes the formula for the expenditure by country n on goods from country i relative to domestically produced goods. Explicitly, the ratio of country n expenditures on any good from country i relative to any domestically produced good is

$$\frac{P_{ni}Y_{ni}}{P_{nn}Y_{nn}} = \left(\frac{M_{ni}C_{1ni}}{M_{nn}C_{1nn}} \right)^{1-\sigma}$$

where M_{ni} is the markup charged by a supplier from country i when selling in country n . Taking the expectation of this expression, the ratio of the expected markup only cancels out in the case of costless trade and varies with the level of the trade cost, separately from the unit costs. The derivative of the ratio with respect to trade costs is not constant, as it enters the distribution of markups charged by home and foreign firms differently, nonlinearly, and nonseparably.

underlying distribution that takes the same form as the distribution of cost parameters for any individual country, $G^{TR}(c) = 1 - e^{-Tc^\theta}$.¹⁶ The distribution of markups in this special case takes the form

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

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 maximum markup and, all else equal, the aggregate price level, which takes the same form as under autarky, but increasing the number of rivals to $R > r$. From here, all three pieces of Proposition 3 are straightforward:

Part a) $E[(C_{2n}^{TR})^{1-\sigma}]$ takes the exact form of its counterpart under autarky, only substituting $R > r^{autarky}$ for the number of rivals, making $E[(C_{2n}^{TR})^{1-\sigma}]$ greater than its counterpart under autarky.¹⁷ From the discussion in Lemma 1, we know that $E[(C_{1n}^{TR})^{1-\sigma}]$ must also be greater than its counterpart under autarky. Therefore, $(p_n^{TR})^{1-\sigma}$ must be greater than its counterpart under autarky, revealing that the aggregate price level falls under trade: $p_n^{TR} < p^{autarky}$.

Part b) To show that the number of rivals under trade equals a number $R > Nr^{autarky}$, we use the free entry condition. The probability of exporting to a country n can be used to formulate the free entry condition under trade:

$$\frac{1}{\delta} E \left[\sum_{i=1}^N \pi_{in} P_i Y_i - C_{1i} Y_i \right] = \kappa E [\pi_{nn} C_{1n} Y_n].$$

If all countries are identical and that trade is costless, we have

$$\frac{E[M^{1-\sigma}]}{E[M^{-\sigma}]} = 1 + \frac{\delta\kappa}{N}.$$

Since the left-hand side is decreasing in r and $[1 + \frac{\delta\kappa}{N}] < 1 + \delta\kappa$, it is clear that the possibility of exporting strictly increases entry.¹⁸ Thus, leaping from autarky to free trade increases the

¹⁶The distribution of first order statistics for samples drawn from a Weibull distribution is also Weibull.

¹⁷That is, normalizing the nominal wage to equal 1 in this special case,

$$E[(C_{2n}^{TR})^{1-\sigma}] = R [T(R-1)]^{\frac{\sigma-1}{\theta}} \Gamma \left(\frac{1-\sigma+2\theta}{\theta} \right).$$

¹⁸Obviously, trade also increases the upperbound for the number of entrants:

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

number of domestic rivals, in addition to reducing prices by reallocating production to more efficient producers. This is the case as long as the fixed costs of entering foreign markets are less than the fixed cost to enter domestic production, κ .¹⁹ Estimation of the fixed industry entry cost and the foreign market access cost by Cherashin, Demidova, Kee, and Krishna (2009) suggest this condition does hold in reality, as the industry entry cost is between 3 and 50 times larger than the foreign market access cost for their sample of Bangladeshi textile producers.

Part c) It follows directly from the calculus in Proposition 1 and the fact that $R > r^{autarky}$ that the average markup falls under trade. Similarly, the likelihood of charging the maximum markup falls when opening to trade. ■

The results from Proposition 3 echo those of Bergin and Feenstra (2000) and Melitz and Ottaviano (2008), but now within the *homothetic* preference structure of Bernard, Eaton, Jensen, and Kortum (2002). Atkeson and Burstein (2007 and 2008) show the results in Parts (a) and (c) numerically, while de Blas and Russ (2010) demonstrate that increasing the number of rivals 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 solutions calculated by Garetto (2010). 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.

3.2 Trade and prices

The expressions for markup behavior in the previous section 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 passthrough, matching them with data on pricing behavior. Here, we demonstrate similar results algebraically. First, pricing-to-market is evident the formula for $h_n^{TR}(m)$ 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

¹⁹Here, we assume the fixed cost of exporting is zero for simplicity.

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 (1) they are already charging the maximum markup and (2) 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.

Suppose again that there is a shock to marginal cost ε such that a drop in ε reduces efficiency and increases marginal costs. The probability that pass-through occurs under trade is now

$$Pr \left[M_n^{TR} \geq \varepsilon \left(\frac{C_{2n}}{C_{1n}} \right) \right] = \int_{-\infty}^{\infty} \int_{\bar{m}}^{\infty} h_n^{TR}(\varepsilon m) v(\varepsilon) dm d\varepsilon. \quad (10)$$

Under symmetry, equation (10) becomes

$$Pr \left[M^T \geq \varepsilon \left(\frac{C_{2n}}{C_{1n}} \right) \right] = \psi \int_{-\infty}^{\infty} \frac{r}{1 + (r-1)(\varepsilon \bar{m})^\theta} d\varepsilon + (1 - \psi) \int_{-\infty}^{\infty} \frac{1}{1 + (\varepsilon \bar{m})^\theta} v(\varepsilon) d\varepsilon,$$

where $\psi = \frac{r-1}{N}$. Since $\frac{r}{1+(r-1)(\varepsilon \bar{m})^\theta} > \frac{1}{1+(\varepsilon \bar{m})^\theta}$ for any $r \geq 2$ and $m \geq 1$, and r is at least as large under trade as under autarky, $Pr \left[M_n^{TR} \geq \varepsilon \left(\frac{C_{2n}}{C_{1n}} \right) \right]$ is less than $Pr \left[M \geq \varepsilon \left(\frac{C_{2n}}{C_{1n}} \right) \right]$, its autarkic counterpart given by equation (6). Therefore, the probability of full passthrough of cost shocks under trade must be less than the probability of full passthrough under autarky. The same can be shown for any degree of partial passthrough, as well.²⁰

For the case where countries are not symmetric, our markup formulas demonstrate results described in the numerical simulations of Garetto (2009). As we noted above from (9), the probability that a firm charges the maximum 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 (2009) argues— “firms should do less pricing-to-market when exporting to relatively more productive (richer) countries.”

3.2.1 Frequency and synchronization of price changes

On a related point from a separate literature, we can see that the frequency of price changes will be smaller in export markets than in domestic markets. Unless an exporting country has

²⁰Although several studies have shown that passthrough depends on the choice of currency invoicing, Goldberg and Tille (2009) demonstrate that this currency invoicing choice also depends on the degree of competition, so we view our market structure approach as quite relevant.

a huge advantage in the form of high T , high r , or low labor costs, it is harder for firms to charge the maximum markup in an export market compared to their native market due to the trade cost, which effectively increases their marginal cost relative to domestic firms in the destination country. Since firms must be charging the maximum 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). More intuitively, trade costs eat away a portion of firms' markups, giving them less leeway to adjust prices in response to increases in marginal costs.

Finally, Schoenle (2010) reports that export price changes are less synchronized than domestic price changes. Trade costs can generate this effect in the same way they reduce the frequency of price changes. However, the dichotomy could also come about because when marginal costs are subject to country-specific shocks, all domestic rivals experience the same shock, making it more likely that a firm can adjust its domestic price even if it is not charging the maximum markup. In contrast, rivals in the export markets more likely would be subject to shocks specific to their own native countries, forcing exporters to absorb domestic cost increases in the form of reduced profits when selling abroad. Notice that exchange rate fluctuations apply more broadly to exporters from multiple countries, and therefore are more likely to be passed on to buyers than source-country-specific shocks to marginal costs.²¹ The issue relates to observations by Gopinath and Itskhoki (2010), who find that the frequency of price adjustment is positively correlated with the degree of exchange rate passthrough. This is true in our model, as the most frequent price adjusters among exporters are those who are able to charge the maximum markup in the foreign market, giving them the greatest degree of passthrough. As discussed above, firms charging the maximum markup are most able to pass through shocks to their relative marginal costs, including those arising from exchange rate movements.

3.2.2 A note on generality

Atkeson and Burstein (2008) point out that if products are less than perfect substitutes, the endogenous markup behavior of firms depicted here breaks down and markups become the CES constant once again, questioning the generality of the BEJK framework. This is a reasonable criticism. However, we argue that at some level of disaggregation, products will

²¹This argument applies only when trade costs are positive, as positive trade costs make it more likely that markups in the domestic market are bounded by a domestic rival, rather than a foreign one. In a symmetric world with costless trade, the second-best rival is equally likely to be domestic or foreign firm and *on average* price changes would appear equally synchronized across markets.

be extremely close substitutes. Even if all upper layers of the CES nesting involve finite elasticities of substitution, the impact of endogenous markups from an underlying level of perfect substitutes would have similar implications for the frequency of observed price changes and the behavior of the aggregate price level, as long as the goods with prices characterized by endogenous markups were assembled in the country where they are consumed as part of the final good.²² The location where intermediate goods are assembled is a critical factor in determining how the endogenous markup behavior affects relative prices and pass-through across countries. Our model assumes, like BEJK and Atkeson and Burstein (2008), that intermediate goods are assembled in the same country where the final good is consumed.

How disaggregate do product categories have to be for goods to be close substitutes? In a structural estimation of the Atkeson and Burstein (2008) model, which uses quantity-based competition but the same nested CES framework as in BEJK and Atkeson and Burstein (2007), Edmond, Midrigan, and Xu (2009) estimate that the elasticity of substitution for the intermediate goods is 8.7 for goods *within* the 4-digit SIC level (goods are identified at the 7-digit SIC level) among firms in the Taiwan Annual Manufacturing Survey.²³ While the number 8.7 is obviously much lower than infinity, in terms of the implied markup it is not so far away: the markup shoots toward infinity when the elasticity of substitution between the goods of rival firms approaches 1, but it tends toward zero as the elasticity approaches infinity and already falls to 15% when the elasticity is 8. Their estimate of the elasticity is just under 8 for simulations with price-based competition of the form in BEJK. So even at an intermediate level of disaggregation, goods are already very close substitutes when analyzed in the context of a Ricardian model with endogenous markups.²⁴

We focus on the Weibull distribution here due to its frequent use in international trade and finance, the microfoundations for the emergence of such a technology laid out in Eaton and Kortum (2009), and its tractability in deriving the distribution of markups. However,

²²The idea of a constant markup being imposed at an upper retail level in the destination country is consistent with findings by Berger, Faust, Rogers, and Steverson (2009), who find that the markup added onto imported goods by distributors *after* the goods arrive at the dock is stable over time.

²³There is a rich literature estimating the elasticity of substitution in the Dixit-Stiglitz (1977) framework with a constant elasticity of substitution, but we omit discussion of it here, as we are focusing on estimates implied by the type of model in this paper, a Ricardian model with endogenous markups.

²⁴The probability density of markups resulting from the quantity-based competition in Atkeson and Burstein (2008) for their preferred calibration also looks strikingly Pareto-like in shape. Edmond, Midrigan, and Xu (2009) show that their (quantity-based) Cournot competition produces a greater dispersion in markups than the price-based Bertrand from BEJK, closer to the actual dispersion observed in their sample. However, it is not clear whether this would be the case once one takes into account the Taiwanese producers shipping to export markets, or the import competition within Taiwan's own domestic market, or calibrates the number of rivals to match observed industry concentration as in Atkeson and Burstein (2008) rather than by counting the number of suppliers of 7-digit micro-industries. We find the study illuminating as the first to apply both Cournot and Bertrand competition in the Ricardian setting, running a horserace to explain observed markups in firm-level data.

any distribution will produce similar qualitative results as long as the ratio of its first two order statistics is decreasing in the sample size and for which the $1-\sigma$ th moments of the first two order statistics exists.

4 Foreign Direct Investment

In this section, we show a mechanism through which foreign-owned firms will have higher markups on average, consistent with existing empirical studies. We focus here on a stylized case without trade for clarity. The intuition originates in the work of Hymer (1960, 1976), who states two principal factors that would compel a firm to control an enterprise in a foreign country. The first is to exploit firm-specific competencies or technological advantages. The second is that it might be possible to eliminate the competition between them, increasing the acquiror's market power. We mathematically demonstrate how the first factor by itself can generate higher markups among foreign-owned firms selling in the host country compared to domestic firms (or compared to autarky). Then we discuss how introducing trade can activate the second factor, further boosting markups for foreign-owned firms exporting to the parent country or to third countries.

Advances in modern trade theory fomented several groundbreaking models to examine the effects of technological advantages, including Helpman, Melitz, and Yeaple (2003) and Nocke and Yeaple (2005). These models rely on a constant elasticity of substitution across a continuum of goods to limit the market share of any individual firm, even if it is far more efficient than its average rival. The love of variety prevents any firm from absorbing the entire market share no matter how superior its technology or how low its price. These studies provide an important window on the interesting tradeoff between exporting and investing abroad for greenfield foreign direct investment (FDI) and cross-border mergers and acquisitions (M&A), showing that FDI allows the most technologically efficient firms to capitalize upon their superiority. Tariff-jumping gives multinationals an additional cost advantage over exporters, boosting their market share above what the technological edge by itself would imply. The Dixit-Stiglitz (1977) preference structure they use imposes a constant markup in price setting. We show that FDI, whether in the form of M&As or greenfield ventures, can increase markups in the host country through the vehicle of technological transfer.

4.1 FDI and markups

Suppose that a country or industry is open to cross-border takeovers but not to trade. A foreign firm can acquire a domestic one, replacing the domestic technology with its own. An asterisk denotes prices, costs, and markups charged by foreign-owned firms in the host country. To clarify the intuition behind the increase in markups that occurs as a result, we first suppose that no trade in goods occurs, forcing all production to be for local sale. For acquired firms, the markup becomes

$$M_n^* = \min \left\{ \frac{C_{2n}}{C_{1n}^*}, \bar{m} \right\},$$

where C_{1n}^* is the lowest-cost draw among all foreign firms (originating in any of $N-1$ countries outside of the home country n) for industry j . C_{1n}^* must be lower than C_{1n} for an acquisition to be profitable for the parent firm, so the markup charged by a foreign-owned firm in the home country will always be at least as large as the pre-takeover markup. The only case where the markup would not increase after a takeover is when the target was already charging the maximum markup \bar{m} . A takeover can be profitable for a parent firm even in this case because the parent applies its superior technology in the acquired plant, resulting in a lower price ($\bar{m}C_{1n}^* < \bar{m}C_{1n}$) and greater sales, which allow it to buy out the target firm at a price equal to the profits it would have earned had it not sold out, given the level of the aggregate price index P that would prevail if all possible takeovers had occurred.

To calculate the distribution of the lowest cost firm among all foreign firms, we note the probability that C_{1n}^* is less than or equal to some level $c_1^* > 0$ is one minus the probability that any foreign country i has a firm in industry j with unit cost less than or equal to c_1^* . Let $G_{1ni}^*(c_1^*)$ be the cumulative distribution for the firm with the lowest cost among all foreign firms that could invest in the home country. Then this cumulative distribution takes the form²⁵

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

where $\Phi_n^* = \sum_{i \neq n} T_i \gamma_{ni}^{-\theta} r_i$, γ_{ni} represents frictions that prevent the seamless transfer of technology from country i to the home country n after a cross-border takeover, and $G_{1ni}^*(c_1^*) = 1 - e^{-T_i \gamma_{ni}^{-\theta} r_i w_n^{-\theta} c_1^{*\theta}}$ is the distribution of the lowest-cost draws for firms from country i pro-

²⁵Here, we make use of the fact that the distribution of the lowest-cost draw among all foreign firms is independent of draws by domestic rivals in the home country.

ducing in country n after a cross-border takeover. We assume that Eaton and Kortum's no arbitrage condition for trade costs also holds for technological hangups: $\gamma_{ni} < \gamma_{ui}\gamma_{nu}$.

Using the joint distribution, we exploit the methodology used in the case of trade above by applying a simple transformation to find the distribution of the markup for merged firms in country n , $\frac{C_{2n}}{C_{1n}^*}$. The distribution of markups among merged firms is given by

$$h_n^*(m) = \int_0^\infty c_1^* g_{1n}^*(c_1^*) g_{2n}(mc_1^*) dc_1^* \quad (11)$$

$$= \int_0^\infty r_n(r_n - 1) T_n \Phi_n^* w_n^{-\theta} \theta^2 c_1^{*\theta} (mc_1^*)^{\theta-1} e^{-[\Phi_n^* + (r_n - 1)T_n] w_n^{-\theta} (mc_1^*)^\theta} dc_1^* \quad (12)$$

$$= \frac{r_n(r_n - 1)\theta T_n \Phi_n^* m^{\theta-1}}{[\Phi_n^* + (r_n - 1)T_n m^\theta]^2}, \quad (13)$$

where $g_{1n}^*(c_1^*)dc_1^* = dG_{1n}^*(c_1^*)$ and $g_{2n}(\cdot)$ is defined exactly as in autarky, since the second-best competitor is necessarily a domestic firm when there is no goods trade.²⁶

Proposition 4: Foreign-owned firms acquired through M&As are more likely to charge the maximum markup and charge higher markups on average than domestically owned firms.

Proof: We can integrate from \bar{m} to ∞ to find the probability that a foreign-owned firm in n will charge the maximum markup,

$$\Pr [M_n^* \geq \bar{m}] = \int_{\bar{m}}^\infty \frac{r_n(r_n - 1)\theta T_n \Phi_n^* m^{\theta-1}}{[\Phi_n^* + (r_n - 1)T_n m^\theta]^2} dm = \frac{r_n \Phi_n^*}{\Phi_n^* + (r_n - 1)T_n \bar{m}^\theta}. \quad (14)$$

Markups are increasing in the level of technology and contestability outside the home country, and decreasing in the degree of technological hangups involved in transferring technology to foreign affiliates. This probability is greater than the probability for domestically owned firms (which is the same as under autarky and given in equation (4)) whenever $\Phi_n^* > T_n$. This condition holds as long as technological hangups are not too great and there is not too much of a native technological advantage in the host country. It always holds under symmetry in a world with seamless technology transfer. This logic also holds for firms charging any markup between 1 and \bar{m} , so we can invoke stochastic dominance as in Proposition 1 to prove that firms acquired by foreign parents charge higher markups on average. ■

²⁶Because the distribution of the ratio $\frac{C_{2n}}{C_{1n}^*}$ is independent of C_{1n} , the distribution of the ratio is also independent of the probability that C_{1n} is greater than C_{1n}^* . This means that the marginal distribution of cost parameters for the second-best domestic firms, $g_{2n}(c_2)$, is independent of the distribution of C_{1n}^* , $g_{1n}^*(c_1^*)$.

The intuition behind the higher markups among merged firms is very simple: Suppose the second-best firm supplying pencils in the home market n has marginal cost equal to 2, while the first-best has unit cost 1.5, with $\sigma = 3$, yielding $\bar{m} = 1.5$. Then the markup under autarky is simply $2/1.5 = 1.\bar{3}$. If a parent firm with marginal cost, including hangups, buys out the active domestic firm and transfers a lower unit cost, say 1.2, then the markup of the merged firm will be $\min\{2/1.2, \bar{m}\} = \bar{m}$.

Corollary: Foreign-owned firms founded through greenfield FDI are less likely to charge the maximum markup than foreign-owned firms established through M&As in any country and in some cases more likely than domestically owned firms.

Proof: The effect on markups is similar to that occurring with mergers and acquisitions. However, in this case we integrate over $g_{1n}(\cdot)$ instead of $g_{2n}(\cdot)$ when computing $h_n^*(m)$.²⁷ The result is

$$Pr \left[M_n^* \geq \bar{m} \right] = \frac{\Phi_n^*}{\Phi_n^* + r_n T_n \bar{m}^\theta},$$

which is strictly lower than the value for cross-border M&As given by equation (14). This probability is only lower than the probability for domestic firms if $\frac{\Phi_n^*}{r_n T_n} > \frac{\bar{m}^\theta}{\bar{m}^\theta - 1}$. Foreign firms are more likely to charge the higher markup as long as the host country does not have intense domestic rivalry or a keen technological advantage relative to foreigners and frictions involved in technology transfer are not too large. The greenfield foreign entrant competes against the best domestic firm in the host country niche, rather than the second-best, so it faces stiffer domestic competition. However, it still comes from an intense competition within the entire global pool of foreign firms. It not only has to be more efficient than the domestic incumbent, it must be more efficient than any other foreign firm in its market niche in order to undertake greenfield FDI without being driven out of the host country immediately by the entry of a more efficient foreign rival. So it is still possible that the greenfield entrant will be sufficiently superior to charge a higher markup than the domestic incumbent it pushes out

²⁷Specifically, we have

$$\begin{aligned} h_n^*(m) &= \int_0^\infty c_1^* g_{1n}^*(c_1^*) g_{1n}(m c_1^*) d c_1^* \\ &= \int_0^\infty r_n T_n \Phi_n^* w_n^{-\theta} \theta^2 c_1^{*\theta} (m c_1^*)^{\theta-1} e^{-[\Phi_n^* + r_n T_n] w_n^{-\theta} (m c_1^*)^\theta} d c_1^* \\ &= \frac{r_n \theta T_n \Phi_n^* m^{\theta-1}}{[\Phi_n^* + r_n T_n m^\theta]^2}. \end{aligned}$$

of business, particularly in a competitively or technologically disadvantaged country. ■

Merged and greenfield firms have lower costs, since the parent must be more efficient than the target to afford the takeover. It is important to note that although the markup may increase after a takeover or greenfield entry, the price charged for the good will never exceed $\min\{C_{2n}, \bar{m}C_{1n}^*\}$. Since C_{2n} has not changed and $C_{1n}^* < C_{1n}$, the price charged for good j in the host country may fall and will never increase, even if the markup does. The influence of domestic entry on markups under FDI is slightly different in the case of M&As versus greenfield. An increase in domestic entry in the host country (r_n) increases the probability that a merged foreign-owned firm will charge the maximum markup, only because it increases the probability that the target firm already charged the maximum markup. In contrast, an increase in domestic entry reduces the probability that a greenfield foreign investor will charge the maximum markup, as greater domestic rivalry increases the expected efficiency of the best domestic firms.

4.2 FDI inflows

A prospective foreign entrant will prefer mergers to greenfield FDI whenever a merger will increase the markup. It can engage in greenfield FDI whenever it has superior technology. Therefore, the probability of foreign entry of some type is simply the probability that some foreign country has a low-cost supplier with unit cost c_1 (including the technological hangup γ_{ni} involved in mergers) given that the home country has a low-cost supplier of good j with unit cost greater than c_1 . The probability that a foreign firm sets up an affiliate in country n by some means is given by

$$\begin{aligned}\pi_n^* &= \int_0^\infty g_{1n}^*(c_1)[1 - G_{1n}(c_1)]dc_1 \\ &= \int_0^\infty \theta\Phi_n^*w_n^{-\theta}c_1^{\theta-1}e^{-(\Phi_n^*+T_n)w_n^{-\theta}c_1^\theta}dc_1 \\ &= \frac{\Phi_n^*}{r_nT_n + \Phi_n^*}.\end{aligned}$$

Similarly, the probability that a firm from country i sets up an affiliate in country n is

$$\pi_{ni}^* = \frac{r_iT_i\gamma_{ni}^{-\theta}}{r_nT_n + \Phi_n^*}$$

4.3 Markups with trade and FDI

The basic logic underlying Proposition 4— that foreign-owned firms will charge higher markups than other firms— still follows in a setting with trade in goods. There are several reasons for this. First, in the case of production for local sale, jumping the tariff barrier gives a foreign-owned firm a cost advantage over arms-length exporters, allowing it to charge a higher markup than if it served a destination through exporting instead of a local affiliate. In addition, the basic efficiency advantage a foreign-owned firm has over domestically owned firms in the host country shown in the case without trade still holds. Both factors lead to higher markups for “horizontal” direct investment.

The case of FDI for export is even more interesting. A firm may invest abroad lured by cost advantages in a country with low wages. In this case, it may begin serving its native markets or third-country markets with much lower costs, but not much lower prices. The degree to which the cost advantage is absorbed in the form of higher markups depends on whether the firm was already charging the maximum markup. In markets where it already was charging the maximum markup, it will have to pass on the entire cost advantage to consumers. In markets where its price is set as the marginal cost of its next-best rival, then it can absorb at least part of the new labor cost savings in the form of higher markups.

Finally, a firm may use *strategic* takeovers to eliminate its best foreign rivals. We define a strategic motive as arising from the opportunity to increase markups in any market outside the host country, including in the parent country. Neary (2008) and Pac (2010) provide substantial evidence that strategic motives are important to understand cross-border M&A activity and build innovative models of this behavior in the context of a representative firm, showing that it increases markups. The Neary (2008) model is really remarkable as the first mathematical framework we are aware of that predicts increases in markups from FDI, in the spirit of Hymer (1960) as we have done here.²⁸

Allowing for strategic takeovers is consistent with the results of our model— it need not result in all firms charging the maximum markup— for two reasons. First, takeovers are costly and involve technological hangups that would prevent firms from always being able to attain the maximum markup by taking over their best rivals, either in autarky or in the open economy. Second, the occurrence of strategic takeovers under autarky would be likely to generate short-term increases in expected profitability that would eventually encourage additional entry, restoring the original equilibrium distribution of markups. In either autarky or the open economy, the result is an equilibrium with dispersion in the distribution of

²⁸With Dixit-Stiglitz preferences, entry by multinationals has no effect on markups even in the host country and additional entry under oligopoly with identical firms would reduce markups.

markups that is governed by the number of rivals in each industry.²⁹

4.4 FDI, Markups and Competition: Stylized facts

Our analysis of foreign-owned firms' pricing behavior after mergers is consistent with the few existing studies of markups and cross-border mergers and acquisitions among manufacturing firms.³⁰ Sembenelli and Siotis (2008) show that an increase in markups is associated with technology transfer to the target after a cross-border merger in Spanish manufacturing industries. In the industries most intensive in research and development (R&D), "FDI has a positive long-run effect on the mark-ups of target firms (p.108)." They argue that the key role of R&D in predicting the behavior of pre- versus post-takeover markups implies a key role for technology transfer between parents and subsidiaries in augmenting market power. In these sectors, they interpret their findings as support for "the fact that MNCs possess firm-specific advantages that can be transferred" so that after a foreign takeover, targeted firms "enjoy greater levels of efficiency, and therefore mark-ups (p.115)." Chari, Ouimet, and Tesar (2010) also find that shareholders expect a bigger return when a firm in an industrialized country takes over a firm in a developing country than when it takes over a firm in its native market—an effect that they attribute to the transfer of managerial know-how.

The most extensive set of relevant studies analyzes foreign takeovers and markups in the banking sector. Many, including Barajas, Steiner and Salazar (1999), Claessens, Demirgüç-Kunt, and Huizinga (2001), Goldberg (2007), and Vera, Zambrano-Sequin, and Faust (2007), demonstrate that the net interest margin—which de Blas and Russ (2009) show is equivalent to the log of a markup in a BEJK framework—increases in targeted banks following foreign takeovers, while costs tend to fall.³¹ Thus, existing micro evidence from both financial and non-financial firms points to improvements in efficiency that lead to higher markups following foreign takeovers, the precise mechanism at work in our model.

²⁹We show in the Appendix D that the impact of FDI on entry is somewhat ambiguous, as the prospect of reaping a share of the profits from a foreign takeover entices entry, while the prospect of competing with foreign-owned firms without the protection of trade costs reduces expected profits and deters entry.

³⁰As authors since Caves (1974) have pointed out, it is difficult to disentangle the impacts of technology transfer from changes in market competitiveness when foreign-owned firms enter a market. Authors such as Chung (2001), Arnold and Javorcik (2005), Alfaro, Kalemli-Ozcan, and Sayek (2009) have documented the technological transfer and spillovers that accompany foreign takeovers or inflows of foreign direct investment.

³¹We are not aware of empirical work documenting the frequency of price adjustment by affiliates of multinational firms and thus omit further discussion of pricing. Our results would predict that in the absence of transfer pricing, multinational affiliates would be more frequent price adjusters because they are more likely to charge the maximum markup than firms under autarky or trade.

5 Conclusions

The principal contribution of this paper is to provide well specified formulas for the distribution of markups under autarky, trade, and FDI in a Ricardian setting. To accomplish this, we generalize the BEJK framework with endogenous markups and heterogeneous firms to allow a role for domestic entry and foreign direct investment. In doing so, we link the theoretical work of BEJK to numerical results in more recent papers and to a large body of empirical work on markups and pricing in the open economy. Our formulas for the distribution of markups demonstrate that previous numerical results are quite general and that the markup formulas in BEJK are a special case of our generalized framework.

The distributions allow us to characterize in an analytically clean way firm markup behavior under trade and FDI, 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 the numerical simulations of Atkeson and Burstein (2007 and 2008) and Garetto (2009), key results include imperfect passthrough and pricing-to-market. In addition, we can characterize an endogenous degree of pricing rigidity that depends on market structure and varies across destination markets due to the degree of domestic entry and the level of trade costs. We also show how the same mechanisms that generate pricing-to-market and imperfect passthrough allow foreign-owned firms to charge higher markups than domestic firms. Thus, our model captures an array of previously disparate stylized facts regarding markup behavior under trade and FDI within a modern Ricardian framework.

References

- [1] Alfaro, Laura, Sebnem Kalemli-Ozcan, and Selin Sayek. 2009. FDI, Productivity and Financial Development. *World Economy* 32(1): 111-135.
- [2] Anderson, Simon P., Andre de Palma and Jacques-Francois Thisse. 1987. The CES is a Discrete Choice Model? *Economics Letters* 24: 139-40.
- [3] Arkolakis, Costas, Arnaud Costinot, and Andres Rodriguez-Clare. 2010. New Models, Same Old Gains from Trade? Forthcoming, *American Economic Review*.
- [4] Arnold, Jens Matthias and Beata Smarzynska Javorcik. 2005. Gifted Kids or Pushy Parents? Foreign Acquisitions and Plant Performance in Indonesia. CEPR Discussion Paper No. 5065.

- [5] Atkeson, Andrew and Ariel Burstein. 2007. Pricing to Market in a Ricardian Model of Trade. *American Economic Review Papers and Proceedings* (May).
- [6] Atkeson, Andrew and Ariel Burstein. 2008. Pricing to Market, Trade Costs, and International Relative Prices. *American Economic Review* (December).
- [7] Barajas, Adolfo, Roberto Steiner, and Natalia Salazar. 1999. Interest Spreads in Banking in Colombia, 1974-1996. *IMF Staff Papers* 46(2): 196-224.
- [8] Berger, David, Jon Faust, John H. Rogers, and Kai Steverson. 2009. Border Prices and Retail Prices. Federal Reserve Board of Governors International Finance Discussion Paper No. 2009-972.
- [9] Bergin, Paul C. and Robert C. Feenstra. 2008. Pass-through of Exchange Rates and Competition between Floaters and Fixers. *Journal of Money Credit and Banking* 41(1):35-70.
- [10] Bernard, Andrew, Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum. 2003. "Plants and Productivity in International Trade," *American Economic Review* 93(4): 1268-1290.
- [11] Bills, Mark and Peter Klenow. 2004. Some Evidence on the Importance of Sticky Prices. *Journal of Political Economy* 112(5): 947-85.
- [12] 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.
- [13] Caves, Richard. 1974. Multinational firms, competition, and productivity in host-country markets. *Economica* 41(162): 176-193.
- [14] Chari, Anusha, Paige Ouimet, and Linda Tesar. 2010. The Value of Control in Emerging Markets. *Review of Financial Studies* 23(4): 1741-1770.
- [15] Cherashin, Ivan, Svetlana Demidova, Hiau Looi Kee, and Kala Krishna. 2009. Firm Heterogeneity and Trade: A New Estimation Strategy and Policy Experiments. Manuscript, Pennsylvania State University.
- [16] Chung, Wilbur. 2001. Identifying Technology Transfer in Foreign Direct Investment: Influence of industry conditions and investing firm motives. *Journal of International Business* 32(2): 211-229.

- [17] Claessens, Stijn, Asli Demirguc-Kunt, and Harry Huizinga. 2001. How Does Foreign Entry Affect Domestic Banking Markets? *Journal of Banking and Finance* 25: 891-911.
- [18] Demidova, Svetlana, Hiau Looi Kee, and Kala Krishna. 2006. Does Trade Policy Induce Sorting? Theory and Evidence from Bangladeshi Apparel Exporters. NBER Working Paper No.
- [19] Dixit, Avinash K. and Joseph E. Stiglitz. 1977. Monopolistic Competition and Optimum Product Diversity. *American Economic Review* 67: 297-308.
- [20] Dunne, Timothy, Mark J. Roberts and Larry Samuelson. 1988. Patterns of Firm Entry and Exit in U.S. Manufacturing Industries. *RAND Journal of Economics* 19(4): 495-515.
- [21] Eaton, Jonathan and Samuel Kortum. 2002. "Technology, Geography, and Trade." *Econometrica* 70(5): 1741-1779.
- [22] Eaton, Jonathan and Samuel Kortum. 2009. Technology in the Global Economy: A Framework for Quantitative Analysis. Manuscript, University of Minnesota.
- [23] Edmond, Christopher, Virgiliu Midrigan, and Daniel Xu. Competition and markups: Micro evidence and aggregate implications. Presentation for the July 2006 Meetings of the Society for Economic Dynamics, Istanbul.
- [24] Feenstra, Robert C. and David E. Weinstein. 2010. Globalization, Markups, and the U.S. Price Level. NBER Working Paper No. 15749.
- [25] Fitzgerald, Doireanne and Stephanie Haller. 2010. Pricing-to-Market: Evidence from Plant-Level Prices. Manuscript, Stanford University.
- [26] Foster, Lucia, John Haltiwanger and Chad Syverson. 2008. Reallocation, Firm Turnover and Efficiency: Selection on Productivity or Profitability?" *American Economic Review* (March).
- [27] Gabaix, Xavier. The Granular Origins of Aggregate Fluctuations. 2010. Forthcoming, *Econometrica*.
- [28] Garetto, Stefania. 2009. Firm Heterogeneity and Incomplete Pass-Through. Manuscript, Boston University.
- [29] Ghironi, Fabio and Marc J. Melitz. 2005. International Trade and Macroeconomic Dynamics with Heterogeneous Firms. *Quarterly Journal of Economics* 120: 865-915 (August).

- [30] Goldberg, Linda. 2007. Financial Sector FDI and Host Countries: New and Old Lessons. Federal Reserve Bank of New York Economic Policy Review. March: 1-17.
- [31] Goldberg, Linda and Cedric Tille. 2009. Micro, Macro, and Strategic Forces in International Trade Invoicing. CEPR Discussion Paper 7534.
- [32] Gopinath, Gita and Oleg Itshoki. 2010. Frequency of Price Adjustment and Pass-through. Quarterly Journal of Economics, forthcoming.
- [33] Gopinath, Gita, Oleg Itshoki, and Roberto Rigobon. 2010. Currency Choice and Exchange-Rate Pass-through. American Economic Review 100(1): 306-336.
- [34] Goodwin, Doris Kearns. 2006. Team of Rivals. New York: Simon and Schuster Paperbacks.
- [35] Harrison, Ann E. 1994. Productivity, Imperfect Competition and Trade Reform. Journal of International Economics 36(1-2): 53-73 (February).
- [36] Hymer, Stephen Herbert. 1976. The International Operations of National Firms: A study of foreign direct investment. The MIT Press: Cambridge, MA. (1960 doctoral thesis submitted posthumously for publication by Charles P. Kindleberger).
- [37] Levinsohn, J. 1993. Testing the imports-as-market-discipline hypothesis. Journal of International Economics 35, pp. 122.
- [38] Melitz, Marc J. 2003. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. Econometrica 71: 1695-1725 (November).
- [39] Melitz, Marc J. and Giancarlo I. P. Ottaviano. 2008. Market Size, Trade, and Productivity. Review of Economic Studies 75(1): 295-316.
- [40] Mood, Alexander McFarlane, Franklin A. Graybill, and Duane C. Boes. 1974. McGraw Hill Companies.
- [41] Nakamura, Emi and Jon Steinsson. 2010. Lost in Transit: Product Replacement Bias and Pricing to Market. Manuscript, Columbia University.
- [42] Neary, J. Peter. 2007. Cross-border mergers as instruments of comparative advantage. Review of Economic Studies 74(4): 1229-1257.
- [43] Nocke, Volker and Stephen R. Yeaple. 2007. "Cross-Border Mergers and Acquisitions versus Greenfield Foreign Direct Investment: The Role of Firm Heterogeneity," Journal of International Economics 72(2): 336-365.

- [44] Novy, Dennis. 2010. International Trade and Monopolistic Competition without CES: Estimating Translog Gravity. Manuscript, University of Warwick.
- [45] Pac, Grzegorz T. 2010. The Role of Multinational Enterprises in the Shutdown of Europe. FREIT Working Paper No. 234.
- [46] Rinne, Horst. 2009. The Weibull Distribution: a Handbook. New York: CRC Press.
- [47] Roberts, Mark and Dylan Supina. 1996. Output Price, Markups, and Producer Size. *European Economic Review* 40(3-5): 909-21.
- [48] Rodriguez, Jose Antonio Lopez. 2010. Prices and Exchange Rates: A Theory of Disconnect. Forthcoming, *Review of Economic Studies*.
- [49] Sembenelli, Alessandro and Georges Siotis. 2008. Foreign Direct Investment and Markup Dynamics: Evidence from Spanish firms. *Journal of International Economics* 76: 107-115.
- [50] Schoenle, Raphael. 2010. International Menu Costs and Price Dynamics. Manuscript, Brandeis University.
- [51] Simonovska, Ina and Michael E. Waugh. 2010. The Elasticity of Trade: Estimates and Evidence. Manuscript, University of California, Davis.
- [52] Vera, Leonardo, Luis Zambrano-Sequin, and Andreas Faust. 2007. The Efficiency-Stability Trade-Off: The case of high interest rate spreads in Venezuela. *The Developing Economies* 45(1): 1-26.

A Free entry under autarky

Substituting the pricing rule into the free entry condition, we have

$$\begin{aligned} \frac{1}{\delta} E [M(j)^{1-\sigma} C_1(j)^{1-\sigma} P^\sigma Y - M(j)^{-\sigma} C_1(j)^{1-\sigma} P^\sigma Y] &= \kappa E [M(j)^{-\sigma} C_1(j)^{-\sigma} P^\sigma Y] \\ \frac{1}{\delta} E [M(j)^{1-\sigma} - M(j)^{-\sigma}] &= \kappa E [M(j)^{-\sigma}] \\ E [M(j)^{1-\sigma}] &= (1 + \delta\kappa) E [M(j)^{-\sigma}], \end{aligned}$$

where the second equality comes from the fact that the distribution of markups is independent of any other variable, including unit cost, and the assumption that firms take the behavior of all firms outside their own niche j (and thus aggregate variables) as given.

A.1 A lowerbound for the log markup.

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

$$\ln(1 + \delta\kappa) + \ln E[M^{-\sigma}] = \ln E[M^{1-\sigma}]. \quad (15)$$

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 $\ln E[M^{-\sigma}] - E[\ln M^{-\sigma}] \geq \ln E[M^{1-\sigma}] - E[\ln M^{1-\sigma}]$.³² This last inequality implies that

$$E[\ln M^{1-\sigma}] \geq \ln(1 + \delta\kappa) + E[\ln M^{-\sigma}],$$

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

$$E[\ln M] \geq \ln(1 + \delta\kappa).$$

A.2 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 $(1 + \delta\kappa)$.

A.3 Deriving the lowerbound for r

Beginning from the generalization of the Chebyshev inequality in the text, we have

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

³²Another 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 (15) reduces the left hand side more than the right hand side.

B Proving that the herfindahl index is decreasing in the number of rivals

B.1 Preliminaries

$$1 - \Pr[M > \bar{m}] = \frac{(r-1)(\bar{m}^\theta - 1)}{1 + (r-1)\bar{m}^\theta}, \quad \frac{\partial \Pr[M > \bar{m}]}{\partial r} = - \left(\frac{\bar{m}^\theta - 1}{r[1 + (r-1)\bar{m}^\theta]} \right) \Pr[M > \bar{m}]$$

$$\begin{aligned} E[C_1^{-2\sigma}] &= (rTw)^{\frac{2\sigma}{\theta}} \Gamma\left(\frac{2\theta + 2\sigma}{\theta}\right), & \frac{\partial E[C_1^{-2\sigma}]}{\partial r} &= \left(\frac{2\sigma}{\theta r}\right) E[C_1^{-2\sigma}] \\ E[C_2^{-2\sigma}] &= r \left(\frac{r-1}{r}\right)^{\frac{2\sigma}{\theta}} E[C_1^{-2\sigma}], & \frac{\partial E[C_2^{-2\sigma}]}{\partial r} &= \left(\frac{2\sigma r + \theta(r-1)}{\theta r(r-1)}\right) E[C_2^{-2\sigma}] \end{aligned}$$

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

B.2 The Herfindahl index

Let H represent the Herfindahl index. Then,

$$\begin{aligned} H &= \int_0^1 \left(\frac{Y(j)}{y}\right)^2 dj = \int_0^1 \left(\frac{P(j)}{p}\right)^{-2\sigma} dj = p^{2\sigma} \int_0^1 P(j)^{-2\sigma} dj = p^{2\sigma} E[P(j)^{-2\sigma}] \\ \frac{\partial H}{\partial r} &= \frac{\partial [p^{2\sigma}]}{\partial r} E[P(j)^{-2\sigma}] + p^{2\sigma} \frac{\partial E[P(j)^{-2\sigma}]}{\partial r} \end{aligned}$$

Further, $\frac{\partial H}{\partial r} < 0$ whenever

$$-\frac{\partial [p^{2\sigma}]}{\partial r} / p^{2\sigma} > \left(\frac{\partial E[P(j)^{-2\sigma}]}{\partial r} \right) / E[P(j)^{-2\sigma}].$$

Substituting relevant expressions from above, we find that

$$\begin{aligned}
\frac{\partial [p^{2\sigma}]}{\partial r} &= - \left(\frac{2\sigma}{\sigma-1} \right) p^{3\sigma-1} \left(\frac{\partial \Pr[M > \bar{m}]}{\partial r} \bar{m}^{1-\sigma} E [C_1^{1-\sigma}] + \Pr[M > \bar{m}] \bar{m}^{1-\sigma} \frac{\partial E [C_1^{-2\sigma}]}{\partial r} - \right. \\
&\quad \left. \frac{\partial \Pr[M > \bar{m}]}{\partial r} E [C_2^{1-\sigma}] + (1 - \Pr[M > \bar{m}]) \frac{\partial E [C_2^{-2\sigma}]}{\partial r} \right) \\
&= - \left(\frac{2\sigma}{\sigma-1} \right) p^{3\sigma-1} \Pr[M > \bar{m}] E [C_1^{1-\sigma}] * \\
&\quad \left[\begin{aligned} &\bar{m}^{1-\sigma} \left(\frac{\sigma-1}{\theta r} - \frac{\bar{m}^\theta - 1}{r[1+(r-1)\bar{m}^\theta]} \right) + \\ &\left(\frac{\bar{m}^\theta - 1}{r} \right) \left(\frac{r-1}{r} \right)^{\frac{\sigma-1}{\theta}} \left(\frac{\theta(r-1)+(\sigma-1)r}{\theta} + \frac{r}{1+(r-1)\bar{m}^\theta} \right) \end{aligned} \right].
\end{aligned}$$

Dividing by $-p^{2\sigma}$, we have

$$\begin{aligned}
-\frac{\frac{\partial [p^{2\sigma}]}{\partial r}}{p^{2\sigma}} &= \left(\frac{2\sigma}{\sigma-1} \right) p^{\sigma-1} \Pr[M > \bar{m}] E [C_1^{1-\sigma}] * \\
&\quad \left[\begin{aligned} &\bar{m}^{1-\sigma} \left(\frac{\sigma-1}{\theta r} - \frac{\bar{m}^\theta - 1}{r[1+(r-1)\bar{m}^\theta]} \right) + \\ &\left(\frac{\bar{m}^\theta - 1}{r} \right) \left(\frac{r-1}{r} \right)^{\frac{\sigma-1}{\theta}} \left(\frac{\theta(r-1)+(\sigma-1)r}{\theta} + \frac{r}{1+(r-1)\bar{m}^\theta} \right) \end{aligned} \right].
\end{aligned}$$

Here, substituting relevant expressions from above for $\partial E [C_2^{-2\sigma}]$ and $\Pr[M > \bar{m}]$, we note that

$$p^{\sigma-1} = \left\{ \Pr[M > \bar{m}] E [C_1^{1-\sigma}] \left[\bar{m}^{1-\sigma} + (r-1) (\bar{m}^\theta - 1) \left(\frac{r-1}{r} \right)^{\frac{\sigma-1}{\theta}} \right] \right\}^{-1},$$

so that

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

Using a similar process, we find that the right-hand side can also be reduced to

$$\frac{\frac{\partial E [P(j)^{-2\sigma}]}{\partial r}}{E [P(j)^{-2\sigma}]} = \frac{\left[\begin{aligned} &\bar{m}^{-2\sigma} \left(\frac{2\sigma}{\theta r} - \frac{\bar{m}^\theta - 1}{r[1+(r-1)\bar{m}^\theta]} \right) + \\ &\left(\frac{\bar{m}^\theta - 1}{r} \right) \left(\frac{r-1}{r} \right)^{\frac{2\sigma}{\theta}} \left(\frac{\theta(r-1)+2\sigma r}{\theta} + \frac{r}{1+(r-1)\bar{m}^\theta} \right) \end{aligned} \right]}{\bar{m}^{-2\sigma} + (r-1) (\bar{m}^\theta - 1) \left(\frac{r-1}{r} \right)^{\frac{2\sigma}{\theta}}}.$$

Using these reduced expressions, it is straightforward to check numerically that $-\frac{\partial [p^{2\sigma}]}{\partial r} / p^{2\sigma} >$

$\left(\frac{\partial E[P(j)^{-2\sigma}]}{\partial r}\right) / E[P(j)^{-2\sigma}]$ whenever $\theta \geq \sigma - 0.5$, a slightly stronger assumption than used in Proposition 2, which was explained in detail in Section 2.3, but one that still fits BEJK's estimates of σ and θ .

C Markups under trade

We begin by calculating the probability that the first- and second-best rivals are from the same country, so that they come from a joint distribution. Let $G_{kni}(c_k)$ and $g_{kni}(c_k)$ be, respectively, the cumulative distribution and probability density of unit costs for the k^{th} -best firm from country i competing to supply country n .³³ Then we calculate the probability that the unit cost for the second-best rival to supply country n with some good j (we have omitted the j index) is less than the unit cost of the best firm that could supply j to country n from any country $u \neq i$ as

$$\begin{aligned}
\psi_{ni} &= \int_0^{\infty} g_{2ni}(c_2) \prod_{u \neq i}^{N-1} [1 - G_{1nu}(c_2)] dc_2 \\
&= \int_0^{\infty} \theta r_i (r_i - 1) T_i(w_i d_{ni})^{-\theta} c_2^{\theta-1} e^{-(r_i-1)T_i(w_i d_{ni})^{-\theta} c_2^{\theta}} e^{-\tilde{\Phi}_{nu}^T c_2^{\theta}} dc_2 \\
&= \frac{r_i(r_i - 1)T_i(w_i d_{ni})^{-\theta}}{\Phi_n^T} \int_0^{\infty} \theta (r_i - 1) T_i(w_i d_{ni})^{-\theta} c_2^{\theta-1} e^{-\Phi_n^T c_2^{\theta}} dc_2 \\
&= \frac{r_i(r_i - 1)T_i(w_i d_{ni})^{-\theta}}{\Phi_n^T} \int_0^{\infty} \Phi_n^T c_2^{\theta-1} e^{-\Phi_n^T c_2^{\theta}} dc_2 \\
&= \left(\frac{r_i(r_i - 1)T_i(w_i d_{ni})^{-\theta}}{\Phi_n^T} \right) e^{-\Phi_n^T c_2^{\theta}} \Big|_0^{\infty}
\end{aligned}$$

Then, we use the formula for the distribution of the ratio of two random variables, C_{1ni}

³³Explicitly,

$$g_{1ni} = \theta r_i T_i(w_i d_{ni})^{-\theta} c_1^{\theta-1} e^{-r_i T_i(w_i d_{ni})^{-\theta} c_1^{\theta}}$$

and

$$g_{2ni} = \theta r_i (r_i - 1) T_i(w_i d_{ni})^{-\theta} c_2^{\theta-1} e^{-(r_i-1)T_i(w_i d_{ni})^{-\theta} c_2^{\theta}}$$

and C_{1nu} ,

$$\begin{aligned}
h_{niu}^{TR,D}(m) &= \int_0^{\infty} \theta c_1 g_{1ni}(c_1) g_{1nu}(mc_1) dc_1 \\
&= \int_0^{\infty} \theta^2 r_i T_i(w_i d_{ni})^{-\theta} c_1^\theta r_u T_u(w_u d_{nu})^{-\theta} (mc_1)^{\theta-1} e^{-[r_i T_i(w_i d_{ni})^{-\theta} + r_u T_u(w_u d_{nu})^{-\theta} m^\theta]} c_1^\theta dc_1 \\
&= \frac{\theta r_i T_i(w_i d_{ni})^{-\theta} r_u T_u(w_u d_{nu})^{-\theta} (m)^{\theta-1}}{[r_i T_i(w_i d_{ni})^{-\theta} + r_u T_u(w_u d_{nu})^{-\theta} m^\theta]^2}
\end{aligned}$$

The probability that the second-best rival to supply a good to country n comes from country $u \neq i$ is equal to the probability that u has marginal cost c_1 given that the best rivals from all third countries ($v \neq u, i$) have a marginal cost of supplying the good to n that is larger than c_1 ,³⁴

$$\begin{aligned}
\psi_{niu}^T &= \int_0^{\infty} g_{1nu}(c_1) \Pi_{v \neq u, i} [1 - G_{1nv}(c_1)] dc_1 \\
&= \theta r_u T_u(w_u d_{nu})^{-\theta} \int_0^{\infty} c_1^{\theta-1} e^{-\tilde{\Phi}_{ni}^T c_1^\theta} dc_1 \\
&= \frac{\theta r_u T_u(w_u d_{nu})^{-\theta}}{\tilde{\Phi}_{ni}^T} e^{-\tilde{\Phi}_{ni}^T c_1^\theta} \Big|_0^\infty.
\end{aligned}$$

Note that $\sum_{u \neq i} \psi_{niu} = 1$.

C.1 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 profits 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 $\frac{1}{1+\delta\kappa}$. Similarly, under free trade with symmetric countries, λ_D equals $\frac{1}{1+(\delta\kappa/N)}$. 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

³⁴We can ignore the second-best producer from i because the probability that both the first and second-best producers are from the same country, psi_{ni} is independent of other countries' best draw (all countries are drawing from independent distributions). This means we are completely ignoring the possibility of multinational firms for the moment.

costs:

$$\omega_n L_n = \beta \lambda_D p_n y_n.$$

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:

$$\frac{y^{TR}}{y^{autarky}} = \left(\frac{1 + \delta\kappa}{1 + (\delta\kappa/N)} \right) \left(\frac{p^{autarky}}{p^{TR}} \right).$$

The first term on the right-hand side is less 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,

$$\frac{p^{autarky}}{p^{TR}} = \left(\frac{\frac{R}{1+(R-1)\bar{m}^\theta} \left(\bar{m}^\theta R^{\frac{\sigma-1}{\theta}} + (R-1)^{1+\frac{\sigma-1}{\theta}} (\bar{m}^\theta - 1) \right)}{\frac{r}{1+(r-1)\bar{m}^\theta} \left(\bar{m}^\theta r^{\frac{\sigma-1}{\theta}} + (r-1)^{1+\frac{\sigma-1}{\theta}} (\bar{m}^\theta - 1) \right)} \right)^{\frac{1}{\sigma-1}}$$

Even under symmetry, the level of gains from trade clearly depends upon the number of domestic rivals before liberalization. In particular, they are lower for countries with a high level of contestability ($r^{autarky}$) to begin with, as these countries already have lower average markups than their trading partners.

D FDI and entry

To examine the impact of FDI on entry, here we discuss the free entry condition in the world with FDI and no trade. We assume for simplicity that with a cross-border takeover, the parent firm pays a fixed fraction ε of its first period output to the target firm each period as a condition of the buyout.³⁵ Then, the expected value of profits earned domestically, combined with expected profits earned from overseas operations given that a firm becomes a multinational or the expected takeover fees given that a firm sells in a cross-border takeover,

³⁵de Blas and Russ (2010) fully endogenize the cost of the takeover—making it equal the option value for the target firm if it does not merge—with no substantive change in the behavior of parents or targets. The simpler assumption here allows greater tractability without loss of generality. The fraction ε must be large enough that the takeover price is at least as large as the startup cost.

must equal the startup cost:³⁶

$$\sum_{s=0}^{\infty} \delta^{t+s} \left[\begin{array}{l} (1 - \Pr[FDI_n]) (P_n Y_n - C_{1n} Y_n) + \\ (1 - \varepsilon) \sum_{\substack{i \neq n \\ i=1}}^{N-1} \pi_{in}^* (P_i Y_i - \gamma_{in}^{-\theta} C_{1n} Y_i) + \\ \varepsilon \sum_{\substack{i \neq n \\ i=1}}^{N-1} \pi_{ni}^* (P_n^* Y_n^* - \gamma_{ni}^{-\theta} C_{1i} Y_n^*) \end{array} \right] \equiv \kappa (1 - \Pr[FDI_n]) C_{1n} Y_n.$$

In the case of symmetry, the expression reduces to

$$(1 - \Pr[FDI]) \left(\begin{array}{c} E[M^{1-\sigma}] \\ -(1 + \delta\kappa)E[M^{-\sigma}] \end{array} \right) + \gamma^{1-\sigma} \Pr[FDI]^{\frac{1}{N-1}} (N-1) \left(\begin{array}{c} E[(M^*)^{1-\sigma}] \\ -(1 + \delta\kappa)E[(M^*)^{-\sigma}] \end{array} \right) \equiv 0$$

From the autarkic case, we know that $E[M^{1-\sigma}] - (1 + \delta\kappa)E[M^{-\sigma}]$ is monotonically decreasing in r . But the direction of $E[(M^*)^{1-\sigma}] - E[(M^*)^{-\sigma}]$ as r increases depends on several factors. Using an approach similar to that in Appendix B, we find that in the case of symmetry, where all countries are identical *ex ante*, this second value is decreasing in r only if

$$\left[\left(\frac{r-1}{2r-1} \right) (N-1) \right]^{\frac{1}{\theta}} < m, \text{ for } 1 \leq m \leq \bar{m}.$$

This means that the expected profit from overseas operations and buyout fees actually increases in r unless θ is sufficiently large relative to the size of the market and the number of entrants. A low shape parameter, θ , means more dispersion in firm productivity— thus, greater potential for cross-border mergers to occur. Increasing the number of draws augments this comparative advantage dimension of the BEJK Ricardian framework, increasing the possibility of a cross-border takeover and thus the expected gains from a takeover or buyout from the point of view of potential entrants deciding whether to draw a cost parameter. When there is little dispersion between draws to begin with (a high θ), having additional entrants (higher r) increases the probability of an active firm benefitting from a buyout or acquisition less than it reduces the expected markup through the competitive effects seen in the autarkic case.

The free entry condition has the same structure for the case of greenfield FDI, but in this case the parent firm need not divide its profits with the target, with the cost that its markup is still subject to competition from the best domestic rival in the host country, rather than the second-best domestic rival. In countries with poor technology or high costs for domestic

³⁶The firm does not pay the startup cost if it sells out to the foreign firm. The parent firm pays only a startup cost in its native country. Thus, there are firm-level economies of scale from multinational production.

entry (few domestic rivals), we would thus expect more greenfield FDI as opposed to cross-border takeovers. In such countries, it is less likely that the efficiency of the best domestic rival in the host country would be bounding the markup of a foreign-owned affiliate.