

# The Elusive Pro-Competitive Effects of Trade

Costas Arkolakis  
Yale and NBER

Arnaud Costinot  
MIT and NBER

Dave Donaldson  
MIT and NBER

Andrés Rodríguez-Clare  
UC Berkeley, Penn State and NBER

June 18, 2012

## Abstract

We study the pro-competitive effects of international trade, or lack thereof, in models with monopolistic competition, firm-level heterogeneity, and variable markups. Under standard restrictions on consumers' demand and the distribution of firms' productivity, we show that gains from trade liberalization are weakly lower than those predicted by the models with constant markups considered in [Arkolakis, Costinot, and Rodríguez-Clare \(2012\)](#). Our preferred estimates suggest that the decrease in the gains from trade liberalization due to variable markups is between 6 and 11%.

# 1 Introduction

How large are the gains from trade liberalization? In earlier work, [Arkolakis, Costinot, and Rodríguez-Clare \(2012\)](#), ACR hereafter, have shown that in an important class of trade models, the answer is pinned down by two statistics: (i) the share of expenditure on domestic goods,  $\lambda$ ; and (ii) an elasticity of imports with respect to variable trade costs,  $\varepsilon$ , which we refer to as the trade elasticity. If a small change in trade costs raises trade openness in some country,  $d \ln \lambda < 0$ , then the associated welfare gain is given by

$$d \ln W = -\frac{d \ln \lambda}{\varepsilon},$$

where  $d \ln W$  is the compensating variation associated with the shock expressed as a percentage of the income of the representative agent.

While the previous formula applies both to models with perfect and monopolistic competition, it relies on the assumption that all agents have Constant Elasticity of Substitution (CES) utility functions. This implies that the models with monopolistic competition considered by ACR necessarily feature constant markups, which de facto rules out any “pro-competitive” effects of trade. This is, potentially, a major limitation of our welfare formula. As [Helpman and Krugman \(1989\)](#) have noted, “the idea that international trade increases competition [...] goes back to Adam Smith, and it has long been one of the reasons that economists give for believing that the gains from trade and the costs from protection are larger than their own models seem to suggest.”

In this paper we drop the CES assumption and study how variable markups affect the gains from trade liberalization under monopolistic competition. Our main finding is that under standard assumptions on consumers’ demand and the distribution of firms’ productivity, which we describe in detail below, the welfare effect of a small trade shock is given by

$$d \ln W = -(1 - \eta) \frac{d \ln \lambda}{\varepsilon},$$

where  $\eta$  is a structural parameter that depends, among other things, on the elasticity of markups with respect to firm productivity. Under parameter restrictions commonly imposed in the existing literature,  $\eta$  is non-negative. Thus, perhaps surprisingly, gains from trade liberalization predicted by these new models are weakly *lower* than those predicted by models with constant markups.

Section 2 describes the basic environment in which this new formula applies. In order to isolate the contribution of variable markups to the gains from trade liberalization, we try to stay as close as possible from the models of monopolistic competition with

firm-level heterogeneity considered in ACR, while departing from CES utility functions. We start from a general demand system that encompasses the three main alternatives to CES utility functions considered in the trade literature: (i) separable, but non-CES utility functions, as in the pioneering work of [Krugman \(1979\)](#) and the more recent work of [Behrens and Murata \(2009\)](#), [Behrens, Mion, Murata, and Sudekum \(2009\)](#), [Saure \(2009\)](#), [Simonovska \(2009\)](#), [Dhingra and Morrow \(2012\)](#) and [Zhelobodko, Kokovin, Parenti, and Thisse \(2011\)](#); (ii) a quadratic, but non-separable utility function, as in [Ottaviano, Tabuchi, and Thisse \(2002\)](#) and [Melitz and Ottaviano \(2008\)](#); and (iii) a translog expenditure function, as in [Feenstra \(2003\)](#), [Bergin and Feenstra \(2009\)](#), [Feenstra and Weinstein \(2010\)](#), [Novy \(2010\)](#), and [Rodriguez-Lopez \(2010\)](#).<sup>1</sup> In addition, we assume that there are no fixed exporting costs, as in [Melitz and Ottaviano \(2008\)](#), and that there exists a finite choke price for all varieties, implying that consumers are not willing to spend an infinite amount to consume an extra variety of a differentiated good. This allows us to abstract from the welfare effects associated with the entry and exit of firms selling at that price.

Section 3 characterizes the trade equilibrium. We first describe how markups vary across firms as a function of their productivity. We show that firm-level markups only depend on the log-difference between the choke price and firms' marginal costs. When demand is log-concave, as assumed in [Krugman \(1979\)](#), markups are increasing with firm-level productivity, which implies incomplete pass-through of changes in marginal costs to prices. Log-concavity is one of the standard parameter restrictions alluded to before that leads to  $\eta \geq 0$  and lower gains from trade liberalization.

Under the assumption that the distribution of firm-level productivity is Pareto, we show that, in spite of variable markups, trade flows satisfy the same gravity equation as in models with CES utility functions and that aggregate profits are a constant share of revenues. Thus, conditional on the value of the trade elasticity, the macro-level predictions of models considered in this paper—namely, the predictions regarding the effects of changes in trade costs on wages and trade flows—are the same as in quantitative trade models with CES utility functions such as [Krugman \(1980\)](#), [Eaton and Kortum \(2002\)](#), [Anderson and Van Wincoop \(2003\)](#), and the versions of the [Melitz \(2003\)](#) model developed by [Chaney \(2008\)](#) and [Eaton, Kortum, and Kramarz \(2011\)](#). In our view, this provides an ideal benchmark to study how departures from CES utility functions may affect the welfare gains from trade liberalization. Since the macro-level behavior of new trade models considered in this paper is exactly controlled for, new gains (if any) may only reflect new

---

<sup>1</sup>Although the demand for differentiated goods in [Ottaviano, Tabuchi, and Thisse \(2002\)](#) and [Melitz and Ottaviano \(2008\)](#) is a special case of the demand system that we consider, it further requires the existence of an “outside good.” We discuss the robustness of our results to the introduction of such a good in Section 5.

micro-level considerations.

Section 4 explores the pro-competitive effects of trade, or lack thereof, in the economic environment described in Sections 2 and 3. We show that the gains from a small decline in trade costs are weakly *lower* than those predicted by ACR’s formula under parameter restrictions commonly imposed in the existing literature. This result builds on the point made above that under standard parameter restrictions, there is incomplete pass-through of changes in marginal costs from firms to consumers. As a result, foreign firms that become more productive because of lower trade costs tend to raise their markups, which tends to lower the welfare gains from trade.<sup>2</sup> Although domestic firms also tend to raise their markups on export markets, which tends to counteract the previous negative welfare effect, there is an overall labor reallocation away from undersupplied, high markup goods, which exacerbates the original distortion associated with variable markups. The net effect is a decrease in the gains from trade as captured by  $\eta$  in our new formula.

Section 5 explores the interaction between gains from new varieties, which we have abstracted from in our baseline analysis, and the existence of variable markups at the firm-level. To do so, we study, both analytically and through simulations, two extensions of our basic environment. The first one allows for fixed trade costs. In this case, marginal varieties are consumed in positive quantities, so entry and exit at the “cut-off” have non-negligible welfare effects. The second extension introduces an outside good, as in Melitz and Ottaviano (2008). This additional sector of production may lead to additional welfare effects through its impact on the size of the differentiated sector.

Our analysis is related to, and has implications for, a large number of theoretical and empirical papers in the international trade literature. On the empirical side, many authors have studied the relationship between international trade and firm-level markups; see e.g. Levinsohn (1993), Harrison (1994), Krishna and Mitra (1998), Konings, Van Cayseele, and Warzynski (2001), Chen, Imbs, and Scott (2009), Loecker and Warzynski (2012), and Goldberg, Loecker, Khandelwal, and Pavcnik (2012).<sup>3</sup> Methodologies, data sources, and conclusions vary, but a common feature of the aforementioned papers is their exclusive focus on domestic producers. As just discussed, however, our results highlight that the

---

<sup>2</sup>This result may seem to go against Krugman (1979)’s finding that under monopolistic competition, if demand functions are log-concave—which is one of our standard parameter restrictions—then larger markets are associated with lower markups. After all, an increase in market size is formally equivalent to an infinite decrease in trade costs starting from autarky. The apparent contradiction between the two sets of results is resolved by noticing that Krugman (1979)’s analysis of markups implicitly focuses on domestic firms, whereas our welfare analysis is crucially affected by the behavior of foreign markups.

<sup>3</sup>Among those, Goldberg, Loecker, Khandelwal, and Pavcnik (2012) is particularly relevant to the present analysis. Under general assumptions that strictly encompass ours, they estimate the elasticity of markups with respect to marginal costs for a large number of Indian firms. These estimates could provide one important piece of information to estimate  $\eta$ , which we plan to use in Section 4.3.

overall pro-competitive effects of trade may be very different from the effects on domestic producers. Under standard assumptions, a decrease in trade costs reduces the markups of domestic producers. Yet, because it also increases the markups of foreign producers, gains from trade liberalization are actually lower than those predicted by standard models with CES utility functions.

A recent empirical paper by [Feenstra and Weinstein \(2010\)](#) is closely related to our analysis. The authors use a translog demand system—which is one of the demand systems covered by our analysis—to measure the contribution of new varieties and variable markups on the change in the U.S. consumer price index between 1992 and 2005. They find that the contribution of these two margins is of the same order of magnitude as the contribution of new varieties estimated by [Broda and Weinstein \(2006\)](#) under the assumption of CES preferences. Our theoretical results show that in the translog case, the overall gains from a hypothetical decline in trade costs are exactly the same as under CES, which resonates well with [Feenstra and Weinstein \(2010\)](#)'s empirical findings. It should be clear, however, that the two exercises are distinct. [Feenstra and Weinstein \(2010\)](#) is a measurement exercise that uses observed trade data to infer changes in particular components of the U.S. price index. The present paper is a counterfactual exercise that focuses on the welfare effect of a particular shock, namely a change in variable trade costs. This approach has both advantages and limitations: it relies on the full structure of the model, including the distribution of firm-level productivity being Pareto, but it allows us to take all general equilibrium effects into account when computing the exact welfare changes caused by trade liberalization.<sup>4</sup>

On the theory side, the introduction of imperfect competition into trade models has been one of the major accomplishments in the international trade literature; see [Helpman and Krugman \(1985\)](#) and [Helpman and Krugman \(1989\)](#). One important message of this literature is that the consequences of changes in trade policy are very sensitive to the underlying market structure. Even in the case of an industry with a single monopolist, changes in trade costs have very different effects on consumer prices and welfare depending on whether the monopolist is a domestic or a foreign firm. As [Helpman and Krugman \(1989\)](#) note—and formalize in a model with one foreign monopolist—a “popular argument about tariffs is that they will be largely absorbed through a decline in foreign markups rather than passed onto consumers.” The flipside of that argument is that a decrease in trade costs may also be largely absorbed through an increase in foreign markups rather than passed onto consumers. In this regard, the main contribution of our paper is to show that under restrictive, but commonly used assumptions on market structure, de-

---

<sup>4</sup>In the translog case, our new welfare formula readily extends to the case of large changes in trade costs.

mand, and firm-level productivity, this channel dominates the pro-competitive effects on domestic firms often emphasized in academic and policy discussions.

Since there are many ways to introduce imperfect competition and variable markups, we have made our modeling choices based on two main considerations. First, we have tried to stay as close as possible to existing trade papers in this area, such as [Krugman \(1979\)](#), [Feenstra \(2003\)](#), and [Melitz and Ottaviano \(2008\)](#). Second, we have tried to stay as close as possible to ACR's assumptions in order to isolate how variable markups affect the mapping between trade openness, as measured by the share of domestic expenditure  $\lambda$ , and welfare. In recent work, [Edmond, Midrigan, and Xu \(2011\)](#) have used the model with two symmetric countries, Cournot competition, and CES utility functions developed by [Atkeson and Burstein \(2008\)](#) to study the magnitude of the gains from trade in economies with variable markups. When calibrating this model using data on manufacturing Taiwanese firms, they find gains from trade that are significantly larger than those predicted by ACR's formula, though the numbers vary depending on assumptions made on the correlation of productivity between Taiwanese and non-Taiwanese firms. We come back to the comparison between their results and ours in Section [4.2](#).

## 2 Basic Environment

We consider a world economy comprising  $i = 1, \dots, n$  countries, one factor of production, labor, and a continuum of differentiated goods  $\omega \in \Omega$ . All individuals are endowed with one unit of labor, are perfectly mobile across the production of different goods, and are immobile across countries.  $L_i$  denotes the total endowment of labor and  $w_i$  denotes the wage in country  $i$ .

### 2.1 Consumers

The goal of our paper is to study the implications of trade models with monopolistic competition for the magnitude of the gains from trade in economies in which markups are variable. This requires departing from the assumption of CES utility functions. The existing trade literature has proposed three alternatives: (i) separable, but non-CES utility functions, as in the pioneering work of [Krugman \(1979\)](#); (ii) a quadratic, but non-separable utility function, as in [Ottaviano, Tabuchi, and Thisse \(2002\)](#); and (iii) a translog expenditure function, as in [Feenstra \(2003\)](#). In this paper, we consider a general demand

system for differentiated goods that encompasses all three.<sup>5</sup>

All consumers have the same preferences. If a consumer with income  $w$  faces a schedule of prices  $\mathbf{p} \equiv \{p_\omega\}_{\omega \in \Omega}$ , her Marshallian demand for any differentiated good  $\omega$  takes the form

$$\ln q_\omega(\mathbf{p}, w) = -\beta \ln p_\omega + \gamma \ln w + d(\ln p_\omega - \ln p^*(\mathbf{p}, w)), \quad (1)$$

where  $p^*(\mathbf{p}, w)$  is symmetric in all prices. Three properties of our demand system are worth emphasizing. First, the price elasticity  $-\beta + d'(\ln p_\omega - \ln p^*(\mathbf{p}, w))$  is allowed to vary with prices, which will generate variable markups under monopolistic competition. Second, other prices only affect the demand for good  $\omega$  through their effect on the aggregator  $p^*(\mathbf{p}, w)$ .<sup>6</sup> Third, the difference between the price elasticity and the cross-price elasticity, i.e. the elasticity with respect to  $p^*(\mathbf{p}, w)$ , is constant and equal to  $-\beta$ . This parameter will play a crucial role in our welfare analysis.

It is easy to check that the previous specification encompasses the case of separable utility functions considered in [Krugman \(1979\)](#). Using our notation, his model corresponds to a situation in which preferences are represented by a utility function,  $U = \int_{\omega \in \Omega} u(q_\omega) d\omega$  with  $u'(0) < \infty$ . The first-order conditions associated with utility maximization imply  $u'(q_\omega) = \lambda(\mathbf{p}, w) p_\omega$ , where  $\lambda(\mathbf{p}, w)$  is the Lagrangian multiplier of the budget constraint, i.e.,  $\int_{\omega \in \Omega} q_\omega p_\omega d\omega = w$ . Inverting the first-order conditions implies

$$\ln q_\omega(\mathbf{p}, w) = d(\ln p_\omega - \ln p^*(\mathbf{p}, w)),$$

where  $p^*(\mathbf{p}, w) \equiv 1/\lambda(\mathbf{p}, w)$  and  $d(x) \equiv u'^{-1}(e^x)$ . Thus the case of separable utility functions correspond to  $\beta = \gamma = 0$ . The mapping between our general demand system and the quadratic and translog cases can be established in a similar manner. We do so in the Appendix.

For future reference, note that the translog case corresponds to  $\beta = 1$  and  $\gamma = 1$ , whereas the quadratic case corresponds to  $\beta = -1$  and  $\gamma = 0$ , where the latter also assumes the existence of an outside good, something we study formally in Section 5. At this point, we only want to clarify the relationship between the demand for differentiated goods in this paper and in the existing literature. Note also that in the absence of an outside good, both the separable and translog cases satisfy  $\beta = \gamma$ .

Given existing parameter values in the literature, we impose the following restriction in our baseline analysis.

---

<sup>5</sup>As already mentioned in footnote 1, the quadratic utility function introduced by [Ottaviano, Tabuchi, and Thisse \(2002\)](#) assumes the existence of an “outside good,” which we will only introduce in Section 5.

<sup>6</sup>In this regard, our specification is less general than the Almost Ideal Demand System of [Deaton and Muellbauer \(1980\)](#), though it does not impose any functional form restriction on  $p^*(\mathbf{p}, w)$ .

**A1. [Existing Demand Systems]**  $\beta = \gamma \leq 1$ .

In addition, we impose the following restrictions on  $d(\cdot)$ .

**A2. [Choke Price]** For all  $x \geq 0$ ,  $d(x) = -\infty$ .

Assumption A2 implies that the aggregator  $p^*(\mathbf{p}, w)$  introduced in equation (1) acts as a choke price. The main benefit of Assumption A2 is that it allows us to study the selection of the most efficient firms into exports, while abstracting from fixed exporting costs. We come back to the importance of this assumption for our welfare analysis in Section 4. The main cost of Assumption A2 is that it excludes from our baseline analysis the case of separable preferences with utility function  $u$  satisfying Inada conditions, most notably the CES case. Section 5 discusses how relaxing A2 and introducing fixed costs affects our main results.

**A3. [Log-concavity]** For all  $x \leq 0$ ,  $d''(x) < 0$ .

Assumption A3 is equivalent to the assumption that demand functions are log-concave (in log-prices) for all differentiated goods. Though this is a fairly strong restriction, it is satisfied by the demand systems considered in Krugman (1979), Ottaviano, Tabuchi, and Thisse (2002), and Feenstra (2003).<sup>7</sup> It is also worth emphasizing that whereas Assumption A2 is necessary to derive the new welfare formula presented in the introduction, Assumptions A1 and A3 are not.<sup>8</sup> Assumptions A1 and A3 merely are sufficient conditions under which one can sign the new parameter  $\eta$  in that formula.

For future derivations, it is convenient to write the demand function in a way that makes explicit the symmetry assumption across goods as well as making explicit the way in which the choke-up price  $p^*(\mathbf{p}, w)$  affects the demand for all goods. Thus, we write  $q_\omega(\mathbf{p}, w) = q(p_\omega, p^*(\mathbf{p}, w), w)$ , with

$$\ln q(p_\omega, p^*, w) = -\beta \ln p_\omega + \gamma \ln w + d(\ln p_\omega - \ln p^*). \quad (2)$$

## 2.2 Firms

Firms compete under monopolistic competition with free entry. There is a large number of ex ante identical firms in each country  $i$  that have the option of hiring  $F_i > 0$  units of labor to enter the industry. We denote by  $N_i$  the measure of firms incurring this fixed entry cost in country  $i$ . After  $w_i F_i$  has been paid, production of any differentiated good

<sup>7</sup>In the case of separable preferences, this is equivalent to assuming that the elasticity of substitution between differentiated goods is decreasing with the level consumption.

<sup>8</sup>Though Assumption A2 requires demand functions to be log-concave locally around 0, it should be clear that it does not require them to be log-concave globally.

is subject to constant returns to scale. As in Melitz (2003), firm-level productivity  $z$  is the realization of a random variable  $Z_i$  drawn independently across firms from a distribution  $G_i$ . We assume that  $G_i$  is Pareto with the same shape parameter  $\theta$  around the world. We assume that  $\theta > \beta - 1$  to ensure that integrals are finite in subsequent sections.

**A4. [Pareto]** For all  $z \geq b_i$ ,  $G_i(z) \equiv \Pr(Z_i \leq z) = 1 - (b_i/z)^\theta$ , with  $\theta > \beta - 1$ .

Assumption A4 is crucial for our welfare analysis, so it is worth pausing to discuss its main implications. As demonstrated below, Assumption A4 implies that all the macro-level restrictions in ACR are satisfied. First, trade flows will satisfy the same gravity equation as in models with CES utility functions. This will allow us to calibrate our model and conduct counterfactual analysis in the exact same way as in ACR. Second, aggregate profits will be a constant share of revenues. This will allow us to ignore, as in ACR, the welfare impact of changes in the measure of entrants.<sup>9</sup>

Finally, we assume that international trade is subject to iceberg trade costs  $\tau_{ij} \geq 1$ , where we normalize  $\tau_{ii} = 1$ . Thus, for a firm with productivity  $z$  in country  $i$ , the constant cost of delivering one unit of the variety associated with that firm to country  $j$  is given by  $c_{ij}/z$ , where  $c_{ij} \equiv w_i \tau_{ij}$ . As mentioned earlier, there are no fixed exporting costs. Throughout our analysis, we assume that good markets are perfectly segmented across countries and that parallel trade is prohibited so that firms charge the optimal monopoly price in each market.

### 3 Trade Equilibrium

In this section we characterize the trade equilibrium for arbitrary values of trade costs. We proceed in two steps. We first study how the demand system introduced in Section 2 shapes firm-level variables. We then describe how firm-level decisions aggregate up to determine bilateral trade flows, and the measure of firms active in each market.

#### 3.1 Firm-level Variables

Consider the optimization problem of a firm producing good  $\omega$  in country  $i$  and selling it in a certain destination  $j$ . To simplify notation, and without risk of confusion, we drop indexes for now and denote by  $c \equiv c_{ij}/z$  the constant marginal cost of serving the market for a particular firm and by  $p^*$  and  $w$  the choke price and the wage in the destination country, respectively. Under monopolistic competition with segmented good markets

---

<sup>9</sup>The third macro-level restriction imposed in ACR is that trade in goods is balanced. Since there are no fixed costs of exporting in the present paper, it holds trivially as well.

and constant returns to scale, the firm chooses its market-specific price  $p$  in order to maximize profits in the market,

$$\pi(c, p^*, w) = \max_p \{(p - c) q(p, p^*, w)\},$$

taking  $p^*$  and  $w$  as given. The associated first-order condition is

$$\frac{p - c}{p} = -\frac{1}{\partial \ln q(p, p^*, w) / \partial \ln p'}$$

which states that monopoly markups are inversely related to the elasticity of demand.

**Firm-level markups.** We use  $m \equiv \ln(p/c)$  as our measure of firm-level markups. Marginal cost pricing corresponds to  $m = 0$ . Combining the previous expression with equation (2), we can express  $m$  as the implicit solution of

$$m - \ln \left( \frac{\beta - d'(m - v)}{\beta - 1 - d'(m - v)} \right) = 0, \quad (3)$$

where  $v \equiv \ln(p^*/c)$  can be thought of as a market-specific measure of the efficiency of the firm relative to other firms participating in that market, as summarized by the choke price prevailing in that market. Equation (3) will play a crucial role in our analysis. It implies that the choke price  $p^*$  is a sufficient statistic for all general equilibrium effects that may lead a firm to change its price in a particular market.

We assume that for any  $v > 0$ , there exists a unique solution  $\mu(v)$  of equation (3) in  $m$ , so that  $m = \mu(v)$ . Assumption A3 is a sufficient, but not necessary condition for existence and uniqueness. The properties of the markup function  $\mu(v)$  derives from the properties of  $d(\cdot)$ . Since  $\lim_{x \rightarrow 0} d(x) = -\infty$  by Assumption A2, we must also have  $\lim_{x \rightarrow 0} d'(x) = -\infty$ , which implies  $\mu(0) = 0$ . The least-efficient firm in a market has zero markup and marginal cost equal to the choke price in that market. Whether markups are monotonically increasing in productivity depends on the monotonicity of  $d'(\cdot)$ . As shown in the Appendix, if demand functions satisfy Assumption A3, then  $\mu'(v) > 0$  so that more efficient firms charge higher markups.

**Firm-level sales and profits.** In any given market, the price charged by a firm with marginal cost  $c$  and relative efficiency  $v$  is given by  $p(c, v) = ce^{\mu(v)}$ , where  $\mu(v)$  is the optimal markup given by equation (3). Given this pricing rule, the total sales faced by a firm with marginal cost  $c$  and relative efficiency  $v$  in a market with wage  $w$  and population  $L$ , are

equal to

$$x(c, v, w, L) \equiv Lw^\gamma \left( ce^{\mu(v)} \right)^{1-\beta} e^{d(\mu(v)-v)}. \quad (4)$$

In turn, the profits of a firm with marginal cost  $c$  and relative efficiency  $v$  selling in a market with wage  $w$  and population  $L$  are given by

$$\pi(c, v, w, L) \equiv \left( \frac{e^{\mu(v)} - 1}{e^{\mu(v)}} \right) x(c, v, w, L). \quad (5)$$

The relationship between profits and sales is the same as in models of monopolistic competition with CES utility functions. The only difference is that markups are now allowed to vary across firms.

### 3.2 Aggregate Variables

**Aggregate sales and profits.** Let  $X_{ij}$  denote the total sales by firms from country  $i$  in country  $j$ . Only firms with marginal cost  $c \leq p_j^*$  sell in country  $j$ . Thus there exists a productivity cut-off  $z_{ij}^* \equiv c_{ij}/p_j^*$  such that a firm from country  $i$  sells in country  $j$  if and only if its productivity  $z \geq z_{ij}^*$ . Accordingly, we can express bilateral trade flows between country  $i$  and  $j$  as

$$X_{ij} = N_i \int_{z_{ij}^*}^{\infty} x(c_{ij}/z, \ln(z/z_{ij}^*), w_j, L_j) dG_i(z).$$

Combining this expression with equation (4) and using our Pareto assumption A4, we get, after simplifications,

$$X_{ij} = \chi N_i b_i^\theta (w_i \tau_{ij})^{-\theta} L_j w_j^\gamma \left( p_j^* \right)^{1-\beta+\theta}. \quad (6)$$

where  $\chi \equiv \theta \int_0^\infty e^{-(1-\beta)(v-\mu(v))} e^{d(\mu(v)-v)} e^{-\theta v} dv > 0$ .<sup>10</sup>

Let  $\Pi_{ij}$  denote aggregate profits by firms from country  $i$  in country  $j$  (gross of fixed entry costs). This is given by

$$\Pi_{ij} = N_i \int_{z_{ij}^*}^{\infty} \pi(c_{ij}/z, \ln(z/z_{ij}^*), w_j, L_j) dG_i(z).$$

---

<sup>10</sup>Equation (6) implicitly assumes that the lower-bound of the Pareto distribution  $b_i$  is small enough so that the firm with minimum productivity  $b_i$  always prefers to stay out of the market,  $b_i < z_{ij}^*$ . This implies that the “extensive” margin of trade is always active in our paper.

Using equations (4) and (5), and again invoking our Pareto assumption A4, we get

$$\Pi_{ij} = \pi N_i \theta b_i^\theta (w_i \tau_{ij})^{-\theta} L_j w_j^\gamma (p_j^*)^{1-\beta+\theta}, \quad (7)$$

where  $\pi \equiv \theta \int_0^\infty (e^{\mu(v)} - 1) e^{-(1-\beta)v - \beta\mu(v)} e^{d(\mu(v)-v)} e^{-\theta v} dv > 0$ . For future reference, note that combined with (6), this implies that

$$\Pi_{ij} = \frac{\pi}{\chi} X_{ij}. \quad (8)$$

Thus, aggregate profits are a constant share of aggregate sales.

**Measure of Entrants.** Free entry requires that the sum of expected profits across all markets be equal to the entry costs, which can be expressed as

$$\sum_j \Pi_{ij} = w_i F_i N_i. \quad (9)$$

Labor market clearing and free entry, in turn, require that total sales across all markets be equal to the total wage bill,

$$\sum_j X_{ij} = w_i L_i. \quad (10)$$

Equations (8), (9) and (10) imply that the measure of entrants in each country is fully determined by country size  $L_i$  and the fixed cost of entry  $F_i$ ,

$$N_i = \frac{\pi L_i}{\chi F_i}. \quad (11)$$

This implies, in particular, that entry levels are invariant to changes in trade costs.

**Summary.** A trade equilibrium corresponds to price schedules,  $(p_1, \dots, p_n)$ , measures of entrants,  $(N_1, \dots, N_n)$ , and wages,  $(w_1, \dots, w_n)$ , for such that (i) prices set in country  $j$  by firms with productivity  $z$  located country  $i$  maximize their profits:

$$p_{ij}(z) = (w_i \tau_{ij} / z) e^{\mu(\ln(p_j^*(p_j, w_j) z / w_i \tau_{ij}))}, \text{ if } z \geq w_i \tau_{ij} / p_j^*(p_j, w_j), \quad (12)$$

and  $p_{ij}(z) \geq w_i \tau_{ij} / z$  if  $z < w_i \tau_{ij} / p_j^*(p_j, w_j)$ ; (ii) measures of entrants are consistent with free entry, equation (11); and (iii) wages are consistent with labor market clearing, equation (10), with aggregate sales  $X_{ij}$  determined by (6). Note that budget constraint in all countries imply that one of these  $n$  labor market conditions is redundant.

### 3.3 Discussion

In spite of the fact that the pricing behavior of firms, as summarized by equation (12), is very different in the present environment than in trade models with CES utility functions, all the macro-level restrictions imposed in ACR remain satisfied.

The fact that trade in goods is balanced trivially derives from labor market clearing in the absence of fixed costs of exporting. We have already pointed out that aggregate profits are a constant share of aggregate sales. To see that bilateral trade flows also satisfy the same macro-level restriction as in ACR, note that equation (6) implies

$$X_{ij} = \frac{N_i b_i^\theta (w_i \tau_{ij})^{-\theta} E_j}{\sum_k N_k b_k^\theta (w_k \tau_{kj})^{-\theta}}, \quad (13)$$

where  $E_j \equiv \sum_k X_{kj}$  denotes total expenditure in country  $j$ . This expression corresponds to what ACR refers to as a strong CES import demand system. In the rest of this paper we simply refer to equation (13) as a gravity equation.

Since the macro-level restrictions imposed in ACR remain satisfied, the macro-level predictions of models considered in this paper are exactly the same as in quantitative trade models using CES utility functions, such as Krugman (1980), Eaton and Kortum (2002), Anderson and Van Wincoop (2003), and Eaton, Kortum, and Kramarz (2011). As shown in the Appendix, once the models with variable markups considered in this paper are calibrated to match the trade elasticity  $\theta$  and the observed trade flows  $\{X_{ij}\}$ , they must predict the exact same changes in wages and trade flows for any change in variable trade costs. Yet, as we will see, differences in the behavior of firms at the micro-level opens up the possibility of new welfare implications.

Before we turn to our welfare analysis, it is worth emphasizing that there will be no gains from new varieties associated with trade liberalization in the present environment. Because  $N_i$  is unaffected by changes in trade costs, changes in the set of varieties available in a given importing country  $j$  may only come from selection effects, i.e. changes in the productivity cut-offs  $z_{ij}^*$ . But in the absence of fixed exporting costs, the creation or destruction of “cut-off” varieties has no welfare consequences either. Indeed, if “cut-off” varieties had some welfare benefit, they would be consumed in strictly positive amount.

Although the macro-level behavior of the models considered in this paper is the same as in ACR, the (assumed) lack of gains from new varieties stands in contrast to the trade models with CES utility functions and fixed exporting costs. We view the two classes of trade models with monopolistic competition and firm-level heterogeneity considered in this paper and in ACR as two useful polar cases. In ACR, as in Krugman (1980) and Melitz

(2003), there are potential gains from new varieties associated with trade liberalization, but CES utility functions rule out any variation in firm-level markups. Here instead, markups are allowed to vary, which opens up the possibility of pro-competitive effects of trade, but gains from new varieties are ruled out.

## 4 The Elusive Pro-Competitive Effects of Trade

In this section we explore the pro-competitive effects of trade, or lack thereof, in the economic environment described in Sections 2 and 3. We focus on a small change in trade costs from  $\tau \equiv \{\tau_{ij}\}$  to  $\tau' \equiv \{\tau_{ij} + d\tau_{ij}\}$ . In ACR, we have shown that under monopolistic competition with Pareto distributions of firm-level productivity and CES utility functions, the compensating variation associated with such a change—i.e., the net revenue of a planner who must compensate a representative agent in country  $j$ —is given by

$$d \ln W_j = -\frac{d \ln \lambda_{jj}}{\theta},$$

where the planner's revenue  $d \ln W_j$  is expressed as a percentage of the income of a representative agent in country  $j$ ;  $\theta$  is equal to the shape parameter of the Pareto distribution, like in the present paper; and  $d \ln \lambda_{jj}$  is the change in the share of domestic expenditure on domestic goods caused by the change from  $\tau$  to  $\tau'$ . By construction, the compensating variation  $d \ln W_j$  is positive if a change in trade costs leads to more trade,  $d \ln \lambda_{jj} < 0$ , because the planner would have to take money away from the consumer to bring her back to her original utility level. We now investigate how going from CES utility functions to the demand system described in equation (1) affects the above formula.

### 4.1 A New Welfare Formula

Let  $e_j \equiv e(\mathbf{p}_j, u_j)$  denote the expenditure function of a representative consumer in country  $j$ , with  $u_j$  being the utility level of such a consumer at the initial equilibrium. By Shephard's lemma, we know that  $de_j/dp_{\omega,j} = q(p_{\omega,j}, p_j^*, w_j) \equiv q_{\omega,j}$  for all  $\omega \in \Omega$ . Since all price changes associated with a change from  $\tau$  to  $\tau'$  are infinitesimal,<sup>11</sup> we can therefore

<sup>11</sup>In principle, price changes may not be infinitesimal because of the creation of "new" goods or the destruction of "old" ones. This may happen for two reasons: (i) a change in the number of entrants  $N$  or (ii) a change in the productivity cut-off  $z^*$ . Since the number of entrants is independent of trade costs, as argued above, (i) is never an issue. Since the price of goods at the productivity cut-off is equal to the choke price, (ii) is never an issue either. This would not be true under CES utility functions and fixed exporting costs. In this case, changes in productivity cut-offs are associated with non-infinitesimal changes in prices since goods at the margin go from a finite (selling) price to an (infinite) reservation price, or vice versa. We

express the associated change in expenditure as

$$de_j = \sum_i \int_{\omega \in \Omega_{ij}} [q_{\omega,j} dp_{\omega,j}] d\omega,$$

where  $\Omega_{ij}$  is the set of goods produced in country  $i$  and exported in country  $j$  and  $dp_{\omega,j}$  is the change in the price of good  $\omega$  in country  $j$  caused by the change from  $\tau$  to  $\tau'$ . The previous expression can be rearranged in logs as

$$d \ln e_j = \sum_i \int_{\omega \in \Omega_{ij}} [\lambda_{\omega,j} d \ln p_{\omega,j}] d\omega, \quad (14)$$

where  $\lambda_{\omega,j} \equiv p_{\omega,j} q_{\omega,j} / e_j$  is the share of expenditure on good  $\omega$  in country  $j$  in the initial equilibrium. Using equation (4), equation (12), and the fact that firms from country  $i$  only sell in country  $j$  if  $z \geq z_{ij}^*$ , we obtain

$$d \ln e_j = \sum_i \int_{z_{ij}^*}^{\infty} \lambda_{ij}(z) [d \ln c_{ij} + d \ln m_{ij}(z)] dG_i(z), \quad (15)$$

where

$$\lambda_{ij}(z) \equiv \frac{N_i e^{-(1-\beta)(\ln(z/z_{ij}^*) - \mu(\ln(z/z_{ij}^*)))} e^{d(\mu(\ln(z/z_{ij}^*) - \ln(z/z_{ij}^*))}}}{\sum_k N_k \int_{z_{kj}^*}^{\infty} e^{-(1-\beta)(\ln(z'/z_{kj}^*) - \mu(\ln(z'/z_{kj}^*)))} e^{d(\mu(\ln(z'/z_{kj}^*) - \ln(z'/z_{kj}^*))}} dG_k(z')}. \quad (16)$$

Equation (15) states that the percentage change in expenditure is equal to a weighted sum of the percentage change in prices, with the percentage changes in prices themselves being the sum of the percentage change in marginal costs,  $d \ln c_{ij}$ , and markups,  $d \ln m_{ij}(z) \equiv -\mu'(\ln(z/z_{ij}^*)) d \ln z_{ij}^*$ . Combining equations (15) and (16) with Assumption A4 and equation (13), we get, after simplifications,

$$d \ln e_j = \sum_i \lambda_{ij} d \ln c_{ij} - \rho \sum_i \lambda_{ij} d \ln z_{ij}^*$$

where  $\lambda_{ij} \equiv X_{ij} / E_j$  is the share of expenditure on goods from country  $i$  in country  $j$  and  $\rho$  is a weighted average of the markup elasticities  $\mu'(v)$  across all firms,

$$\rho = \int_0^{\infty} \mu'(v) \frac{e^{-(1-\beta)(v-\mu(v))} e^{d(\mu(v)-v)} e^{-\theta v}}{\int_0^{\infty} e^{-(1-\beta)(v'-\mu(v'))} e^{d(\mu(v')-v')} e^{-\theta v'} dv'} dv,$$

---

come back to this issue in detail in Section 5.

with weights given by the share of total expenditure on firms with relative efficiency  $v$ .<sup>12</sup> Finally, using the definition of the productivity cut-off  $z_{ij}^* \equiv c_{ij}/p_j^*$ , we can rearrange the expression above as

$$d \ln e_j = \underbrace{\sum_i \lambda_{ij} d \ln c_{ij}}_{\text{Change in marginal costs}} + \underbrace{(-\rho) \sum_i \lambda_{ij} d \ln c_{ij}}_{\text{Direct markup effect}} + \underbrace{\rho d \ln p_j^*}_{\text{GE markup effect}}. \quad (17)$$

To fix ideas, consider a “good” trade shock,  $\sum_i \lambda_{ij} d \ln c_{ij} < 0$ . If markups were constant, the only effect of such a shock would be given by the first term on the RHS of (17). Here, the fact that firms adjust their markups in response to a trade shock leads to two additional terms. The second term on the RHS of (17) is a direct effect. *Ceteris paribus*, a decrease in trade costs makes exporting firms relatively more productive, which leads to changes in markups, by equation (3). Interestingly, under Assumption A3, this direct effect tends to *lower* gains from trade liberalization. The reason is simple. There is incomplete pass-through of changes in marginal costs from firms to consumers: firms that become more productive because of lower trade costs tend to raise their markups ( $\mu'(v) > 0$ ), leading to lower welfare gains ( $\rho > 0$ ). The third term on the RHS of (17) is a general equilibrium effect. It captures the change in markups caused by changes in the choke price  $p_j^*$ . This is the channel emphasized, for instance, by Melitz and Ottaviano (2008) and Corcos, del Gatto, Mion, and Ottaviano (2011). If trade liberalization leads to a decline in the choke price, reflecting a more intense level of competition, then  $\rho > 0$  implies a decline in markups and *higher* gains from trade liberalization.

In order to compare the direct and general equilibrium markup effects, we now need to compare the change in marginal costs,  $\sum_i \lambda_{ij} d \ln c_{ij}$ , to the change in the choke price,  $d \ln p_j^*$ . We can do so by using the labor market clearing condition (10). Totally differentiating  $\sum_i X_{ij} = w_j L_j$ , using equation (6), we obtain

$$d \ln p_j^* = \frac{\theta}{1 - \beta + \theta} \sum_i \lambda_{ij} d \ln c_{ij} + \frac{1 - \gamma}{1 - \beta + \theta} d \ln w_j. \quad (18)$$

Plugging equation (18) into equation (17), we finally get

$$d \ln e_j = \left[ 1 - \rho \left( \frac{1 - \beta}{1 - \beta + \theta} \right) \right] \left[ \sum_i \lambda_{ij} d \ln c_{ij} \right] + \rho \left( \frac{1 - \gamma}{1 - \beta + \theta} \right) d \ln w_j. \quad (19)$$

At this point, we can follow the exact same strategy as in ACR. By differentiating the

<sup>12</sup>Assumption A4 guarantees that the average markup elasticities are constant across source countries  $i$ . Without Pareto distributions, these elasticities—like trade elasticities—would vary across countries.

gravity equation (13), one can show that  $\sum_i \lambda_{ij} d \ln c_{ij}$  is equal to  $d \ln \lambda_{jj} / \theta + d \ln w_j$ . Combining this observation with equation (19) and using the fact that  $\beta = \gamma$  under Assumption A1—which guarantees that demand functions are homogeneous of degree zero in all prices—we obtain

$$d \ln e_j = \left[ 1 - \rho \left( \frac{1 - \beta}{1 - \beta + \theta} \right) \right] \frac{d \ln \lambda_{jj}}{\theta} + d \ln w_j.$$

Under free entry, we know that income per capita in country  $j$  is equal to the wage  $w_j$ . Accordingly, the compensating variation associated with a small change in trade costs from  $\tau \equiv \{\tau_{ij}\}$  to  $\tau' \equiv \{\tau_{ij} + d\tau_{ij}\}$  is given by

$$d \ln W_j = d \ln w_j - d \ln e_j = \left[ 1 - \rho \left( \frac{1 - \beta}{1 - \beta + \theta} \right) \right] \frac{d \ln \lambda_{jj}}{\theta}.$$

We are now ready to state the main result of our paper.

**Proposition 1** *Suppose that Assumptions A1-A4 hold. Then the compensating variation associated with a small trade shock in country  $j$  is given by*

$$d \ln W_j = - (1 - \eta) \frac{d \ln \lambda_{jj}}{\theta}, \text{ with } \eta \equiv \rho \left( \frac{1 - \beta}{1 - \beta + \theta} \right) \in [0, 1].$$

Despite the fact that markups are allowed to vary at the firm-level, we see that welfare analysis can still be conducted using only a few sufficient statistics. In particular, like in ACR, the share of expenditure on domestic goods,  $\lambda_{jj}$ , is the only endogenous variable whose changes needs to be observed in order to evaluate the welfare consequences of changes in trade costs.

Despite the fact that the models analyzed in this paper satisfy the same macro-level restrictions as in ACR, we also see that different predictions at the micro-level—namely the variation in markups across firms—lead to different welfare conclusions. In addition to the trade elasticity  $\theta$ , the compensating variation now also depends on an extra structural parameter  $\eta$ , whose interpretation we discuss in detail in the next section.

Finally, we see that under Assumptions A1-A4, the new structural parameter  $\eta$  is non-negative. Thus an increase in openness to trade,  $d \ln \lambda_{jj} < 0$ , is associated with welfare gains that are lower than  $-d \ln \lambda_{jj} / \theta$ . Since bilateral trade flows satisfy the gravity equation (13) and the measure of entrants is independent of trade costs, one can further check that the value of  $d \ln \lambda_{jj} / \theta$  associated with a trade shock is the same as in ACR; see Appendix. We conclude that under standard restrictions on consumers' demand and the

distribution of firms' productivity, gains from trade liberalization are weakly *lower* than those predicted by the models with constant markups considered in ACR.

## 4.2 Discussion

In order to understand why variable markups lower the gains from trade under Assumptions A1-A4, it is useful to step back and analyze how they affect different components of the changes in real income per capita  $d \ln W_j$ .

By definition, revenue per capita in country  $j$  is equal to  $r_j \equiv \frac{1}{L_j} \sum_i \int_{\omega \in \Omega_{ji}} p_{\omega,i} q_{\omega,i} L_i d\omega$ . Differentiating the previous expression—and using the fact that  $N_j$  is fixed and that “cut-off” varieties have zero sales as we did in the previous section—we get

$$d \ln r_j = \sum_i \int_{\omega \in \Omega_{ji}} \left[ \phi_{\omega,i} (d \ln p_{\omega,i} + d \ln q_{\omega,i}) \right] d\omega,$$

where  $\phi_{\omega,i} \equiv p_{\omega,i} q_{\omega,i} L_i / r_j L_j$  is the share of revenues in country  $j$  associated with sales of good  $\omega$  in country  $i$ . Subtracting the change in expenditure  $d \ln e_j$  in equation (14) from the previous expression, we can express the change in real income per capita as

$$d \ln W_j = \sum_{i \neq j} \left\{ \int_{\omega \in \Omega_{ji}} \left[ \phi_{\omega,i} d \ln p_{\omega,i} \right] d\omega - \int_{\omega \in \Omega_{ij}} \left[ \lambda_{\omega,j} d \ln p_{\omega,j} \right] d\omega \right\} + \sum_i \int_{\omega \in \Omega_{ji}} \left[ \phi_{\omega,i} d \ln q_{\omega,i} \right] d\omega,$$

where we have used the fact that  $\phi_{\omega,j} = p_{\omega,j} q_{\omega,j} / r_j = p_{\omega,j} q_{\omega,j} / e_j$  because of trade balance in country  $j$ . The first term in that decomposition is a standard terms-of-trade effect: real income in country  $j$  increases if the price of its exports increases and decreases if the price of imports increase. The second term captures both a direct productivity effect and the underlying distortion caused by monopolistic competition. As we show in the Appendix, it can be split as follows:

$$\sum_i \int_{\omega \in \Omega_{ji}} \left[ \phi_{\omega,i} d \ln q_{\omega,i} \right] d\omega = - \left[ \sum_i \int_{\omega \in \Omega_{ji}} \left[ \phi_{\omega,i} d \ln \tau_{ji} \right] d\omega \right] + cov \left( e^{\mu_{\omega,i}}, \frac{dL_{\omega,i}}{L_j} \right),$$

where  $L_{\omega,i}$  denotes total employment associated with the production of good  $\omega$  for country  $i$  in country  $j$ .  $\sum_i \int_{\omega \in \Omega_{ji}} \left[ \phi_{\omega,i} d \ln \tau_{ji} \right] d\omega$ , represents the productivity effect associated with a change in the trade costs, whereas

$$cov \left( e^{\mu_{\omega,i}}, \frac{dL_{\omega,i}}{L_j} \right) = \sum_i \int_{\omega \in \Omega_{ji}} \left[ e^{\mu_{\omega,i}} d \left( L_{\omega,i} / L_j \right) \right] d\omega$$

reflects the impact of trade costs on the distortion caused by variable markups.<sup>13</sup>

In order to help the comparison between this decomposition and the one considered in the previous section, let us use equation (4), equation (12), and the fact that firms from country  $i$  only sell in country  $j$  if  $z \geq z_{ij}^*$  to write

$$d \ln W_j = \sum_{i \neq j} \left\{ \left( \frac{E_i}{E_j} \right) \lambda_{ji} [d \ln c_{ji} + d \ln m_{ji}] - \lambda_{ij} [d \ln c_{ij} + d \ln m_{ij}] \right\} \\ - \sum_i \left( \frac{E_i}{E_j} \right) \lambda_{ji} d \ln \tau_{ji} + cov \left( e^{\mu_{\omega,i}}, \frac{dL_{\omega,i}}{L_j} \right),$$

where we use the notation  $d \ln m_{ij} \equiv \int_{z_{ij}^*}^{\infty} \frac{\lambda_{ij}(z)}{\lambda_{ij}} d \ln m_{ij}(z) dG_i(z)$ . The previous expression can be rearranged as

$$d \ln W_j = d \ln w_j - \sum_i \lambda_{ij} d \ln c_{ij} - \sum_{i \neq j} \left[ \lambda_{ij} d \ln m_{ij} - \left( \frac{E_i}{E_j} \right) \lambda_{ji} d \ln m_{ji} \right] + cov \left( e^{\mu_{\omega,i}}, \frac{dL_{\omega,i}}{L_j} \right). \quad (20)$$

Compared to a model with perfect competition, we see that variable markups affect the consequences of changes in trade costs in two ways. First, they affect terms of trade effects:  $\sum_{i \neq j} \left\{ \lambda_{ij} d \ln m_{ij} - \left( \frac{E_i}{E_j} \right) \lambda_{ji} d \ln m_{ji} \right\}$  is in general different from zero. Second, they create a new source of gains or losses from trade liberalization depending on whether the labor reallocation caused by changes in trade costs is positively or negatively correlated with markups. If trade liberalization triggers reallocation of factors of production towards goods that are under supplied, i.e. those with higher markups, then welfare gains from trade tend to be higher than under perfect competition. This corresponds to the case in which  $cov \left( e^{\mu_{\omega,i}}, \frac{dL_{\omega,i}}{L_j} \right)$  is strictly positive. Otherwise, welfare gains from trade are lower than under perfect competition.

The idea that gains from international trade may be higher or lower in the presence of distortions is an old one in the field; see e.g. [Bhagwati \(1971\)](#). A key contribution of the present paper is to provide a theoretical framework in which these effects can be signed and quantified using only a few sufficient statistics. By comparing equations (17) and (20), we see that the combined effects of variable markups on terms-of-trade effects and distortions is equal to  $-\rho \left[ \sum_i \lambda_{ij} d \ln c_{ij} - d \ln p_j^* \right]$ , which reduces to  $\eta d \ln \lambda_{jj} / \theta$ .<sup>14</sup> As discussed earlier, since  $\eta$  is positive under standard parameter restrictions, this implies

<sup>13</sup>It should be clear that this new decomposition is not specific to the particular market structure considered in this paper; see [Basu and Fernald \(2002\)](#) for a general discussion.

<sup>14</sup>The previous identity also explains why domestic markups appear to matter in the analysis of Section 4.1. Since labor supply is inelastic, domestic markups per se are a transfer from consumers towards

that more openness,  $d \ln \lambda_{jj} < 0$ , tends to be associated with lower welfare gains than under perfect competition (or monopolistic competition with constant markups).

To understand why variable markups lead to lower gains from trade, consider a symmetric economy in which iceberg trade costs uniformly decrease by  $d \ln \tau < 0$ . In this case, there are no terms-of-trade effects. Thus the new term in our welfare formula,  $\eta d \ln \lambda_{jj} / \theta$ , only reflects the impact of changes on trade costs on the severity of the distortion. Under Assumption A1 ( $\beta \leq 1$ ), general equilibrium effects are small in the sense that the response of the choke price  $p^*$  required to bring back trade balance after a change in trade costs is less than proportional to the shock,  $(\sum_{i \neq j} \lambda_{ij}) d \ln \tau - d \ln p_j^* < 0$  by equation (18). Under Assumption A3 (log-concavity), this tends to increase the relative demand for low-productivity/high-price goods ( $d^2 \ln q(p_\omega, p^*, w) / d \ln p_\omega d \ln p^* = -d''(\ln p_\omega - \ln p^*) > 0$ ) and, in turn, to trigger a reallocation of labor towards these goods. Since Assumption A3 also implies that low-productivity goods have lower markups, this tends to generate a negative correlation between markups and the labor reallocation,  $cov\left(e^{\mu_{\omega,i}}, \frac{dL_{\omega,i}}{L_j}\right) \leq 0$ , hence the lower gains from trade liberalization in Proposition 1 compared to ACR.

The logic of our welfare results stands in contrast to the results of [Edmond, Midrigan, and Xu \(2011\)](#) who emphasize how—in a symmetric environment that abstracts from terms-of-trade considerations—trade liberalization may affect the degree of misallocation in an economy through its effects on the distribution of markups. In our paper, the distribution of markups is invariant to changes in trade costs; see Appendix for details.<sup>15</sup> Yet, the degree of misallocation in the economy *does* vary with changes in trade costs. As equation (20) illustrates, what matters for welfare is the correlation between markups and changes in the labor allocation, not the changes in the dispersion (or other moments) of the distribution of markups. Under standard restrictions on consumers' demand and the producers. In our model, however, free entry implies:

$$\sum_{i \neq j} \left( \frac{w_i L_i}{w_j L_j} \right) \lambda_{ji} d \ln m_{ji} + cov\left(e^{\mu_\omega}, \frac{dL_\omega}{L_j}\right) = -\lambda_{jj} d \ln m_{jj}.$$

Thus, computing the changes in domestic markups allow us to account both for the change in terms-of-trade caused by changes in markups on the export markets as well as changes in the distortion.

<sup>15</sup>This reflects the countervailing effects of a productivity change, in general, and a change in trade costs, in particular, on markups. Under standard assumptions, productivity gains tend to raise the markups of incumbent firms, but also lead to the entry of less efficient firms, which charge lower markups. Under Pareto, the second effect exactly offsets the first one. A similar compositional effect is at play in [Bernard, Eaton, Jensen, and Kortum \(2003\)](#). In their model, Bertrand competition leads to variable markups at the firm-level, but distributional assumptions similar to ours make the distribution of markups invariant to changes in trade costs. [de Blas and Russ \(2010\)](#) and [Holmes, Hsu, and Lee \(2010\)](#) study, among other things, how departures from this distributional assumptions affect markups.

distribution of firms' productivity, we have shown that this new channel tends to lower rather than increase the gains from trade liberalization. In the next section, we take a first stab at quantifying this effect.

### 4.3 Quantitative Results

In ongoing empirical work we use disaggregated U.S. trade data to estimate  $\eta$ , and in turn, to quantify the importance of the pro-competitive effects of trade in practice. For now, we can get some sense of how variable markups may affect the gains from trade liberalization by exploring plausible bounds for  $\eta$ .

Without an outside good, our general demand system encompasses two possible sets of microtheoretical foundations: (i) additively separable utility functions, which corresponds to  $\beta = 0$ ; and (ii) translog expenditure functions, which corresponds to  $\beta = 1$ . In the latter case  $\eta = 0$ , so the gains are the same as in ACR. So we focus on the former case, which implies  $\beta = 0$  and hence  $\eta = \rho / (1 + \theta)$ .

Suppose first that demand is log-concave (Assumption A3). Then one can check that  $0 < \mu'(v) < 1$  (see Appendix), which implies that firms with higher productivity charge higher markups and lower prices. In turn, we must have  $0 < \rho < 1$  and hence

$$0 < \eta < 1 / (1 + \theta).$$

The first part of this inequality implies that the gains from trade liberalization are lower than in ACR, while the second part of the inequality implies that  $1 / (1 + \theta)$  is an upper bound to the adjustment of the gains from trade liberalization due to variable markups. Since trade flows satisfy a gravity equation with trade elasticity  $\theta$ , we follow ACR and set  $\theta$  between 5 and 10; see [Anderson and Van Wincoop \(2004\)](#). This implies that, under these particular assumptions, the (downward) adjustment to the gains from trade liberalization coming from variable markups can be at most 17%.

We can get more precise results by considering a reasonable value for  $\rho$  from the empirical literature. [Goldberg, Loecker, Khandelwal, and Pavcnik \(2012\)](#) estimate markups and use prices to back out the implied marginal costs across manufacturing firms in India. They find a positive correlation between productivity and markups, in line with the previous assumption that demand is log concave and so that  $\rho > 0$ . When running a cross-sectional regression of (log) prices on (log) marginal cost ( $m$  on  $-\ln z$ ), they find a "pass-through" coefficient of 0.35. For a given firm in our model, the pass through coefficient is  $1 - \mu'(v)$ : this is the elasticity of the price with respect to marginal cost for a firm with relative efficiency  $v$ . Accordingly,  $\rho$  is equal to one minus the average pass-through

coefficient across firms, weighted by their shares of domestic sales. If we use [Goldberg, Loecker, Khandelwal, and Pavcnik \(2012\)](#)'s estimate as a proxy for that average, then we can set  $\rho = 0.65$ . For  $\theta \in [5, 10]$ , this yields  $\eta \in [0.6, 0.11]$ , implying that the decrease in the gains from trade liberalization due to variable markups is between 6 and 11%.<sup>16</sup>

## 5 Robustness

In our baseline analysis, we have abstracted from the welfare gains from new varieties. The objective of this final section is to explore the interaction between such gains and the existence of variable markups at the firm-level. To do so, we consider two departures from the basic environment presented in Section 2: (i) fixed trade costs, as in ACR, and (ii) an outside good, as in [Melitz and Ottaviano \(2008\)](#). With fixed costs the marginal varieties are consumed in positive quantities, so entry and exit at the “cut-off” have implications for welfare. With an outside good, trade may lead to reallocation of resources and expenditure, and these forces will affect welfare through the set of goods consumed. To deal with both issues, we need to impose more structure on our demand system. We follow the previous literature by assuming separable utility functions or translog expenditure functions in the presence of fixed costs, as in [Krugman \(1979\)](#) and [Feenstra \(2003\)](#), and quadratic, but non-separable utility function in the presence of an outside good, as in [Ottaviano, Tabuchi, and Thisse \(2002\)](#).

### 5.1 Fixed Trade Costs

The basic environment is the same as in Section 2, except for the fact that after receiving their random productivity draw, firms must incur fixed trade costs in order to sell in each market. Specifically, the fixed cost for any country to sell in country  $j$  is  $w_j f_j$ .

Consider a firm with marginal cost  $c$  and relative efficiency  $v \equiv \ln(p^*/c)$ . Fixed costs do not affect firm-level markups, which remains a function of  $v$  alone, but they affect

---

<sup>16</sup>[Goldberg, Loecker, Khandelwal, and Pavcnik \(2012\)](#) also run a regression of (log) price on (log) marginal cost with firm fixed effects, so that the identification comes from the variation over time in their panel. This yields a lower pass-through coefficient of 0.2, which would imply  $\rho = 0.8$  and  $\eta \in [0.72, 0.13]$ . [Goldberg, Loecker, Khandelwal, and Pavcnik \(2012\)](#) have data for multiple products, so they are careful to include appropriate fixed effects to get within product identification of the pass-through coefficient. In the cross-section regression they control for product-year fixed effects while in the time-series regression they control for firm-product fixed effects.

firm-level profits. We now have:

$$\pi(c, v, w, L) \equiv \left( \frac{e^{\mu(v)} - 1}{e^{\mu(v)}} \right) L w^\beta \left( c e^{\mu(v)} \right)^{1-\beta} e^{d(\mu(v)-v)} - w f,$$

where we have used the fact that  $\beta = \gamma$  under Assumption A1. Accordingly, a firm will enter a given market  $j$  if and only if  $v \geq v_j^*$ , with  $v_j^*$  implicitly defined by

$$\left( \frac{e^{\mu(v_j^*)} - 1}{e^{\mu(v_j^*)}} \right) \left( e^{\mu(v_j^*)-v_j^*} \right)^{1-\beta} e^{d(\mu(v_j^*)-v_j^*)} = \frac{f_j}{L_j} \left( \frac{w_j}{p_j^*} \right)^{1-\beta}. \quad (21)$$

If  $f_j = 0$  then  $v_j^* = 0$ . Thus firms only enter market  $j$  if their marginal cost is below the choke price in that market, as in Section 3. In general, however, we need to keep track of a new equilibrium variable,  $v_j^* \equiv v_j(\ln p_j^*)$ , where  $v_j(\cdot)$  is implicitly determined by (21). Given  $v_j^*$  and  $p_j^*$ , productivity cut-offs can be computed as  $z_{ij}^* = c_{ij} e^{v_j^*} / p_j^*$ .

The gravity equation (13) and the free entry condition (??) are the same as before, while the labor market clearing condition now takes into account the resources associated with the fixed trade costs,

$$\sum_j (\lambda_{ij} w_j L_j - N_i \sum_k w_k f_k (1 - G_i(z_{ik}^*))) = w_i L_i - w_i f_i \sum_l N_l (1 - G_l(z_{li}^*)).$$

Finally, prices are exactly as above,  $p_{ij}(z, p_j^*, c_{ij}) = (c_{ij}/z) e^{\mu(\ln(p_j^* z / c_{ij}))}$ , if  $z \geq z_{ij}^*$ .

If  $\beta < 1$ , which is the case under separable utility function, then equation (21) implies that  $v_j^*$  is affected by changes in  $p_j^*$  and hence is no longer invariant to changes in trade costs. This implies that the markup distribution and entry levels will now be affected by changes in trade costs. If  $\beta = 1$ , which is the case under translog expenditure functions,  $v_j^*$  is no longer equal to zero, but it remains unaffected by trade costs. Accordingly, both the markup distribution and entry levels will remain constant. This is the case that we study first.

### 5.1.1 Translog Expenditure Function

In the translog case, the expenditure function can be expressed in closed form as a function of good prices,  $p_j$ , as well as the number of entrants,  $N \equiv (N_1, \dots, N_n)$ , and the pro-

ductivity cut-offs,  $z_j^* \equiv (z_{1j}^*, \dots, z_{1j}^*)$ , that affects the set of available varieties:

$$e(p_j, N, z_j^*, u_j) = \left[ \sum_i N_i \int_{z_{ij}^*} p_{ij}(z) q(p_{ij}(z), \tilde{p}_j) dG_i(z) \right] \times u_j, \quad (22)$$

where:

$$q(p, \tilde{p}) = \frac{\xi}{p} \ln(\tilde{p}/p),$$

$$\tilde{p}_j = \frac{1 + \xi \sum_i N_i \int_{z_{ij}^*} \ln p_{ij}(z) dG_i(z)}{\sum_i N_i (1 - G_i(z_{ij}^*))}.$$

Totally differentiating equation (22) yields

$$d \ln e_j = - \sum_i \lambda_{ij} \varphi_{ij}^N d \ln N_i + \sum \lambda_{ij} \varphi_{ij}^Z d \ln z_{ij}^* \\ + \sum \lambda_{ij} d \ln c_{ij} + \left( -\rho(v_j^*) \right) \sum \lambda_{ij} d \ln c_{ij} + \rho(v_j^*) \sum \lambda_{ij} d \ln p_j^*.$$

where:

$$\varphi_{ij}^N \equiv \int_{z_{ij}^*} N_i p_{ij}(z) q(p_{ij}(z), \tilde{p}_j) dG_i(z) + \frac{\partial \ln \tilde{p}_j}{\partial \ln N_i} \int_{z_{ij}^*} p_{ij}(z) \frac{\partial q(p_{ij}(z), \tilde{p}_j)}{d \ln \tilde{p}_j} dG_i(z),$$

$$\varphi_{ij}^Z \equiv -p_{ij}(z_{ij}^*) q(p_{ij}(z_{ij}^*), \tilde{p}_j) g_i(z_{ij}^*) z_{ij}^* + \frac{\partial \ln \tilde{p}_j}{\partial \ln z_{ij}^*} \int_{z_{ij}^*} p_{ij}(z) \frac{\partial q(p_{ij}(z), \tilde{p}_j)}{d \ln \tilde{p}_j} dG_i(z),$$

$$\rho(v_j^*) \equiv \int_{v_j^*}^{\infty} \mu'(v) \frac{e^{-(1-\beta)(v-\mu(v))} e^{d(\mu(v)-v)} e^{-\theta v}}{\int_0^{\infty} e^{-(1-\beta)(v'-\mu(v'))} e^{d(\mu(v')-v')} e^{-\theta v'} d v'} d v.$$

$\varphi_{ij}^N$  and  $\varphi_{ij}^Z$  measure the utility gains from new varieties in country  $j$  associated with changes in the measure of entrants in country  $i$  and changes in the productivity cut-offs, respectively. Like in Section 4.1,  $\rho(v_j^*)$  measures the average pass-through in country  $j$ , which now depends on the level of fixed costs. Following the exact same steps as in this section, one can show that

$$d \ln e_j = - \sum_i \lambda_{ij} \varphi_{ij}^N d \ln N_i + \sum \lambda_{ij} \varphi_{ij}^Z d \ln z_{ij}^* + \frac{d \ln \lambda_{jj}}{\theta}.$$

Thus, gains from trade liberalization will be larger than in the case without fixed costs if and only if there are gains from new varieties,  $-\sum_i \lambda_{ij} \varphi_{ij}^N d \ln N_i + \sum \lambda_{ij} \varphi_{ij}^Z d \ln z_{ij}^* > 0$ .

However, one can show that  $d \ln N_i = 0$  for all  $i$ , because profits are a constant share of total revenues, and that  $\varphi_{ij}^Z = 0$  for all  $i$ , because average sales are constant across destinations. These two features of our model directly derive from Assumption A4 and imply that  $d \ln e_j = -d \ln \lambda_{jj} / \theta$ . Under the assumption of translog expenditure functions, whether or not there are fixed costs of exporting, gains from trade liberalization are the same as in ACR.

### 5.1.2 Separable Utility Function

We now turn to the case of separable utility functions. Since we consider an environment with fixed costs, we no longer need to assume the existence of a choke price in order to generate the selection of the most productive firms into export. Thus we drop Assumption A2. Here, in line with the discussion of Section ??, one can simply think of  $p^*(\mathbf{p}, w)$  as the inverse of the Lagrange multiplier associated with the budget constraint of the representative agent or, equivalently, the Lagrange multiplier associated with the constraint in the expenditure minimization problem.

Without loss of generality, we normalize the subutility function of the representative consumer so that  $u(0) = 0$ . The expenditure function can then be written as

$$e(\mathbf{p}_j, \mathbf{N}, \mathbf{z}_j^*, u_j) = \min_q \sum_i N_i \int_{z_{ij}^*}^{\infty} p_{ij}(z) q_i(z) dG_i(z)$$

subject to

$$\sum_i N_i \int_{z_{ij}^*}^{\infty} u(q_i(z)) dG_i(z) \geq u_j.$$

Differentiating we then get the same decomposition as in the translog case:

$$\begin{aligned} d \ln e_j = & - \sum_i \lambda_{ij} \tilde{\varphi}_j^N d \ln N_i + \sum \lambda_{ij} \tilde{\varphi}_j^Z d \ln z_{ij}^* \\ & + \sum \lambda_{ij} d \ln c_{ij} + \left( -\rho(v_j^*) \right) \sum \lambda_{ij} d \ln c_{ij} + \rho(v_j^*) \sum \lambda_{ij} d \ln p_j^*. \end{aligned}$$

The only difference between this new expression and the one derived in the translog case

comes from the gains from new varieties. We now have:

$$\begin{aligned}\tilde{\varphi}_j^N &\equiv \int_{v_j^*}^{\infty} \left[ \frac{1}{\varepsilon^u(q(v))} - 1 \right] \frac{e^{\mu(v)-(1+\theta)v} q(v)}{\int_{v_j^*}^{\infty} e^{\mu(v')-(1+\theta)v'} q(v') dv'} dv, \\ \tilde{\varphi}_j^Z &\equiv \left[ \frac{1}{\varepsilon^u(q_j^*)} - 1 \right] \frac{e^{\mu(v_j^*)-(1+\theta)v_j^*} q(v_j^*)}{\int_{v_j^*}^{\infty} e^{\mu(v)-(1+\theta)v} q(v) dv}.\end{aligned}$$

where  $\varepsilon^u(q) \equiv u'(q)q/u(q)$  denotes the elasticity of the subutility function, which determines the magnitude of the gains from new varieties. The CES case considered in ACR corresponds to  $\varepsilon^u(q) = (\sigma - 1) / \sigma$  and  $\rho = 0$ . In this situation, the first two terms are equal to zero, which immediately leads to ACR's formula.

To explore whether gains from new varieties affect our welfare predictions in any systematic way we now turn to a simple numerical example. Our goal is to compute the overall gains from trade when fixed export costs are present and compare these gains to the ones that arise in the model with constant markups as in ACR and the model with variable markups but no fixed export costs (as in Section 4). To allow for all these possibilities within the same framework we consider a generalized CES utility function,

$$u(q) = (q + a)^{(\sigma-1)/\sigma} - a^{(\sigma-1)/\sigma},$$

with  $\sigma > 1$  and  $a > 0$ . This utility function reduces to the simple CES case if  $a = 0$  but features variable markups whenever  $a > 0$ . Thus, if we set  $a = 0$  the model collapses to the one in ACR whereas if we set fixed exporting costs to zero the model collapses to the one in Section 4. \*\*\* To Be Completed \*\*\*

## 5.2 Outside Good

We now come back to the case without fixed costs. The basic environment is the same as in Section 2, except for the existence of an outside good. As in [Melitz and Ottaviano \(2008\)](#), the outside good is freely traded, produced one-to-one from labor in all countries, and used as the numeraire. In addition, we assume that the preferences of the representative agent over the outside good and the differentiated goods can be represented by a

quadratic utility function:

$$U(q_0, \mathbf{q}) = q_0 + \kappa_1 \sum N_i \int_{z_{ij}^*}^{\infty} q_i(z) dG_i(z) - \frac{\kappa_2}{2} \sum N_i \int_{z_{ij}^*}^{\infty} (q_i(z))^2 dG_i(z) - \frac{\kappa_3}{2} \left( \sum N_i \int_{z_{ij}^*}^{\infty} q_i(z) dG_i(z) \right)^2,$$

where  $q_0$  is the consumption of the outside good and  $\kappa_1, \kappa_2, \kappa_3 > 0$  are preference parameters. Thus, substituting for the consumption of the outside good, the consumer's minimization problem can be expressed as

$$e_j(N, \mathbf{z}_j^*, \mathbf{c}, u_j) = \min_{q_0, \mathbf{q}} \sum_i N_i \int_{z_{ij}^*}^{\infty} p(z, p_j^*, c_{ij}) q_i(z) dG_i(z) - \left( \kappa_1 \sum N_i \int_{z_{ij}^*}^{\infty} q_i(z) dG_i(z) - \frac{\kappa_2}{2} \sum N_i \int_{z_{ij}^*}^{\infty} (q_i(z))^2 dG_i(z) - \frac{\kappa_3}{2} \left( \sum N_i \int_{z_{ij}^*}^{\infty} q_i(z) dG_i(z) \right)^2 \right).$$

Differentiating the previous expression, one can show that:

$$d \ln W_j = \sum_i \lambda_{ij}^d \hat{\varphi}_j^N d \ln N_i - (1 - \rho) \sum_i \lambda_{ij}^d d \ln c_{ij} - \rho d \ln p_j^*,$$

where  $\lambda_{ij}^d$  denotes the share of expenditure on differentiated goods from country  $i$  in country  $j$  and  $\hat{\varphi}_j^N \equiv 1 - \kappa'_1/p_j^* + \kappa'_2 + \kappa'_3 M_j$  denotes the gains from new varieties associated with changes in the measure of entrants in country  $i$ , which now depends on the choke price,  $p_j^*$ , as well as the measure of goods available in country  $j$ ,  $M_j \equiv \sum_i N_i (b_i/z_{ij}^*)^\theta$ .

Let  $\lambda_j^d \equiv \sum_i \lambda_{ij}^d$  denote the share of expenditure on differentiated goods in country  $j$  and let  $E_j^d$  denote total expenditure on differentiated goods in country  $j$ . Using the gravity equation—which still holds in the differentiated good sector—one can further simplify the previous equation as

$$d \ln e_j = \sum_i \lambda_{ij}^d \hat{\varphi}_j^N d \ln N_i - \left[ \left( \frac{1 - \rho}{\theta} \right) \right] \sum_i \lambda_{ij}^d d \ln (N_i/N_j) - \lambda_j^d \frac{\rho}{2 + \theta} d \ln (E_j^d/N_j) - \lambda_j^d (1 - \eta) \frac{d \ln \lambda_{jj}}{\theta},$$

where  $\eta$  is defined in the exact same way as in Proposition 1 and  $\lambda_{jj} \equiv \lambda_{jj}^d/\lambda_j^d$  now denotes, without risk of confusion, the share of expenditure on domestic goods in the differentiated sector.

To gain further intuition, consider the case in which countries are symmetric. In this situation, we must have  $d \ln (N_i/N_j) = 0$  for all  $i$ . We must also have total revenue in the differentiated good sector being equal to total expenditure in that sector country-by-country. Since the measure of entrants is proportional to total revenues, this implies  $d \ln (E_j^d/N_j) = 0$ . Combining these two observations with the previous expression, we get

$$d \ln W_j = \lambda_j^d \times \left( \hat{\varphi}_j^N d \ln E_j^d - (1 - \eta) \frac{d \ln \lambda_{jj}}{\theta} \right).$$

Besides the obvious fact that the change in expenditure now needs to be scaled down by the share of expenditure in the differentiated sector, we see that the only difference compared to the expression derived in the absence of an outside good is the term  $-\hat{\varphi}_j^N d \ln E_j^d$ . This reflects the fact that total expenditure in the differentiated good is no longer pinned down by the exogenous labor supply in country  $j$ . The exact same issue arises in a model of monopolistic competition with constant markups; see [Krugman \(1980\)](#). This channel may either increase or lower the gains from trade liberalization depending on the degree of substitutability between the outside good and differentiated goods. If it is low enough, then a decrease in trade costs in the differentiated good sector will be associated with a decrease in expenditure in that sector,  $d \ln E_j^d < 0$ , and hence lower gains from trade liberalization than those predicted by [Proposition 1](#).

## 6 Concluding Remarks

We have studied the pro-competitive effects of international trade, or lack thereof, in models with monopolistic competition, firm-level heterogeneity, and variable markups. Under standard restrictions on consumers' demand and the distribution of firms' productivity, we have shown that gains from trade liberalization are weakly lower than those predicted by the models with constant markups considered in [Arkolakis, Costinot, and Rodríguez-Clare \(2012\)](#). Our preferred estimates suggest that the decrease in the gains from trade liberalization due to variable markups is between 6 and 11%.

# A Proofs

## A.1 Existing Demand Systems

**Linear demand.** A quadratic utility function as in [Ottaviano, Tabuchi, and Thisse \(2002\)](#) implies the following linear demand for differentiated goods:

$$q(p_\omega, p^*, w) = \frac{\kappa_1}{\kappa_2 + \kappa_3 M(p^*)} - \frac{p_\omega}{\kappa_2} + \frac{\kappa_3}{\kappa_2} \frac{\int_{p_\omega \leq p^*} p_\omega d\omega}{\kappa_1 + \kappa_3 M(p^*)}, \quad (23)$$

where  $M(p^*) \equiv \int_{p_\omega \leq p^*} d\omega$  and  $p^*$  is implicitly defined as the solution of

$$p^* = \frac{\kappa_1 \kappa_2}{\kappa_2 + \kappa_3 M(p^*)} + \frac{\kappa_3}{\kappa_2 + \kappa_3 M(p^*)} \int_{p_\omega \leq p^*} p_\omega d\omega.$$

Demand in (23) can be rewritten as

$$\ln q(p_\omega, p^*, w) = \ln p_\omega - \ln \kappa_2 + \ln (\exp(-[\ln p_\omega - \ln p^*]) - 1).$$

In terms of the demand system introduced in Section 2, this implies  $\beta = -1$ ,  $\gamma = 0$  and  $d(x) = -\ln \kappa_2 + \ln(e^{-x} - 1)$ .

**Translog demand.** A translog expenditure function as in [Feenstra \(2003\)](#) implies the following demand for differentiated goods:

$$q(p_\omega, p^*, w) = \frac{w}{p_\omega} \zeta \ln(p^*/p_\omega),$$

where  $p^*$  is implicitly defined as the solution of

$$\ln p^* = \frac{\frac{1}{\zeta} + \int_{p_\omega \leq p^*} \ln p_\omega d\omega}{M(p^*)},$$

with  $M(p^*) \equiv \int_{p_\omega \leq p^*} d\omega$ . In terms of the demand system introduced in Section 2, this implies  $\beta = \gamma = 1$ , and  $d(x) = \ln \zeta + \ln(-x)$ .

## A.2 Monotonicity of Markups

To see this, let  $f(m, v) \equiv m - \ln \left( \frac{\beta - d'(m-v)}{\beta - 1 - d'(m-v)} \right)$ . Differentiating with respect to  $m$  and  $v$ , we obtain

$$\begin{aligned} f_m(m, v) &= 1 + \left[ \frac{1}{\beta - d'(m-v)} - \frac{1}{\beta - 1 - d'(m-v)} \right] d''(m-v), \\ f_v(m, v) &= - \left[ \frac{1}{\beta - d'(m-v)} - \frac{1}{\beta - 1 - d'(m-v)} \right] d''(m-v). \end{aligned}$$

Note that  $\beta - d'(m-v) > \beta - 1 - d'(m-v) > 0$ , where the last equality follows from the condition that at firms' chosen prices the elasticity of demand must be higher than one. But then  $d''(m-v) < 0$  implies  $f_m(m, v) > 0$  and  $f_v(m, v) < 0$ . By the implicit function theorem, equation (3) therefore implies  $\mu'(v) = -f_v(m, v)/f_m(m, v) > 0$ . For future purposes, also note that  $-f_v(m, v)/f_m(m, v) < 1$ , hence  $\mu'(v) < 1$ .

## A.3 Equivalence of Counterfactual Changes in Trade Flows

By equations (13) and (10), we know that

$$\begin{aligned} \lambda_{ij} &= \frac{N_i b_i^{-\theta} (w_i \tau_{ij})^{-\theta}}{\sum_k N_k b_k^{-\theta} (w_k \tau_{kj})^{-\theta}}, \\ \sum_j \lambda_{ij} w_j L_j &= w_i L_i. \end{aligned}$$

Consider a counterfactual change in variable trade costs from  $\tau \equiv \{\tau_{ij}\}$  to  $\tau' \equiv \{\tau'_{ij}\}$ . Let  $\hat{v} \equiv v'/v$  denote the change in any variable  $v$  between the initial and the counterfactual equilibrium. Since  $N_i$  fixed for all  $i$ , using the exact same argument as in the proof of Proposition 2 in ACR, one can show that

$$\hat{\lambda}_{jj} = \frac{1}{\sum_{i=1}^n \lambda_{ij} (\hat{w}_i \hat{\tau}_{ij})^{-\theta}},$$

where  $\hat{w}_j = 1$  by choice of numeraire, and  $\{\hat{w}_i\}_{i \neq j}$  are implicitly given by the solution of

$$\hat{w}_i = \sum_{j'=1}^n \frac{\lambda_{ij'} \hat{w}_{j'} Y_{j'} (\hat{w}_i \hat{\tau}_{ij'})^{-\theta}}{Y_i \sum_{i'=1}^n \lambda_{i'j'} (\hat{w}_{i'} \hat{\tau}_{i'j'})^{-\theta}}.$$

By Proposition 2 in ACR, this implies that conditional on trade flows and expenditures in the initial equilibrium and an estimate of the trade elasticity, counterfactual changes in trade flows are the same as in ACR.

## A.4 Welfare Decomposition

In the main text we have argued that

$$\sum_i \int_{\omega \in \Omega_{ji}} [\phi_{\omega,i} d \ln q_{\omega,i}] d\omega = \sum_i \int_{\omega \in \Omega_{ji}} [\phi_{\omega,i} d \ln \tau_{ji}] d\omega + \text{cov} \left( e^{\mu_{\omega,i}}, \frac{dL_{\omega,i}}{L_j} \right). \quad (24)$$

We now establish this result formally. First note that  $q_{\omega,i} L_i = z_{\omega} L_{\omega,i} / \tau_{ji}$ , where  $z_{\omega}$  is the productivity of the firm producing good  $\omega$  in country  $j$  and  $L_{\omega,i}$  is the amount of labor used for exports towards country  $i$ . Thus we have

$$\sum_i \int_{\omega \in \Omega_{ji}} [\phi_{\omega,i} d \ln q_{\omega,i}] d\omega = - \sum_i \int_{\omega \in \Omega_{ji}} [\phi_{\omega,i} d \ln \tau_{ji}] d\omega + \sum_i \int_{\omega \in \Omega_{ji}} [\phi_{\omega,i} d \ln L_{\omega,i}] d\omega.$$

Since  $\phi_{\omega,i} = p_{\omega,i} q_{\omega,i} L_i / r_j = (e^{\mu_{\omega,i}} \tau_{ji} w_j / z_{\omega}) (z_{\omega} L_{\omega,i} / \tau_{ji}) / w_j L_j = e^{\mu_{\omega,i}} L_{\omega,i} / L_j$ , we can rearrange the previous expression as

$$\sum_i \int_{\omega \in \Omega_{ji}} [\phi_{\omega,i} d \ln q_{\omega,i}] d\omega = - \sum_i \int_{\omega \in \Omega_{ji}} [\phi_{\omega,i} d \ln \tau_{ji}] d\omega + \sum_i \int_{\omega \in \Omega_{ji}} \left[ e^{\mu_{\omega,i}} d \left( \frac{L_{\omega,i}}{L_j} \right) \right] d\omega. \quad (25)$$

Since labor market clearing in country  $j$  implies  $\sum_i \int_{\omega \in \Omega_{ji}} L_{\omega,i} d\omega = L_j$ , we know that  $\sum_i \int_{\omega \in \Omega_{ji}} d(L_{\omega,i} / L_j) d\omega = 0$ . Thus if we denote by  $e^{\mu} \equiv \sum_i \int_{\omega \in \Omega_{ji}} e^{\mu_{\omega,i}} d\omega$ , we must have

$$\sum_i \int_{\omega \in \Omega_{ji}} [e^{\mu_{\omega,i}} d(L_{\omega,i} / L_j)] d\omega = \sum_i \int_{\omega \in \Omega_{ji}} [(e^{\mu_{\omega,i}} - e^{\mu}) [d(L_{\omega,i} / L_j) - 0]] d\omega. \quad (26)$$

Let  $L_{\omega}$  denote total employment associated with the production of good  $\omega$  in country  $j$ . Using this notation and combining equations (25) and (26), we obtain equation (24).

## A.5 Distribution of Markups

Let  $M_{ij}(m; \tau)$  denote the distribution of markups set by firms from country  $i$  in country  $j$  in a trade equilibrium if trade costs are equal to  $\tau \equiv \{\tau_{ij}\}$ . Since firm-level markups only depend on the relative efficiency of firms, we can express

$$M_{ij}(m; \tau) = \Pr \{ \mu(v) \leq m | v \geq 0 \},$$

where the distribution of  $v$  depends, in principle, on the identity of both the exporting and the importing country. Recall that  $v \equiv \ln(p^*/c)$  and  $c = c_{ij}/z$ . Thus for a firm with productivity  $z$  located in  $i$  and selling in  $j$ , we have  $v = \ln p_j^* - \ln c_{ij} + \ln z = \ln z - \ln z_{ij}^*$ .

Combining this observation with Bayes' rule, we can rearrange the expression above as

$$M_{ij}(\mu; \tau) = \frac{\Pr \left\{ \mu(\ln z_{ij}^* - \ln z) \leq m, \ln z_{ij}^* \leq \ln z \right\}}{\Pr \left\{ \ln z_{ij}^* \leq \ln z \right\}}.$$

Under Assumption A3,  $\mu(\cdot)$  is strictly monotone, and under Assumption A4, firm-level productivity  $z$  is drawn from a Pareto distribution. Thus the previous equation implies

$$M_{ij}(m; \tau) = \frac{\int_{\ln z_{ij}^*}^{\ln z_{ij}^* - \mu^{-1}(m)} \tilde{g}_i(u) du}{\left( b_i / z_{ij}^* \right)^\theta},$$

where  $\tilde{g}_i(u) \equiv \theta b_i^\theta e^{-\theta u}$  is the density of  $u \equiv \ln z$ . This can be further simplified into

$$M_{ij}(m; \tau) = 1 - e^{\theta \mu^{-1}(m)}.$$

Since the function  $\mu(\cdot)$  is identical across countries and independent of  $\tau$ , by equation (3), this establishes that for any exporter  $i$  and any importer  $j$ , the distribution of markups  $M_{ij}(\cdot; \tau)$  is independent of the identity of the exporter  $i$ , the identity of the importer  $j$ , and the level of trade costs  $\tau$ . As a result, the overall distribution of markups in any country  $j$  is also invariant to changes in trade costs.

## References

- ANDERSON, J. E., AND E. VAN WINCOOP (2003): "Gravity with Gravitas: A Solution to the Border Puzzle," *The American Economic Review*, 93(1), 170–192.
- (2004): "Trade Costs," *Journal of Economic Literature*, 42(3), 691–751.
- ARKOLAKIS, C., A. COSTINOT, AND A. RODRÍGUEZ-CLARE (2012): "New Trade Models, Same Old Gains?," *American Economic Review*, 102(1), 94–130.
- ATKESON, A., AND A. BURSTEIN (2008): "Pricing to Market, Trade Costs, and International Relative Prices," *American Economic Review*, 98(5), 1998–2031.
- BASU, S., AND J. G. FERNALD (2002): "Aggregate Productivity and Aggregate Technology," *European Economic Review*, 46, 963–991.
- BEHRENS, K., G. MION, Y. MURATA, AND J. SUDEKUM (2009): "Trade, Wages, and Productivity," *CEP Discussion Papers*, 0942.
- BEHRENS, K., AND Y. MURATA (2009): "Globalization and Individual Gains from Trade," *CEPR discussion paper*, 7438.
- BERGIN, P. R., AND R. FEENSTRA (2009): "Pass-Through of Exchange Rates and Competition Between Floaters and Fixers," *Journal of Money Credit and Bankin*, 41(1), 35–70.
- BERNARD, A. B., J. EATON, J. B. JENSEN, AND S. KORTUM (2003): "Plants and Productivity in International Trade," *American Economic Review*, 93(4), 1268–1290.
- BHAGWATI, J. N. (1971): *The Generalized Theory of Distortions and Welfare*, vol. Trade, Balance of Payments, and Growth: Papers in International Economics in Honor of Charles P. Kindleberger. Amsterdam: North-Holland.
- BRODA, C., AND D. WEINSTEIN (2006): "Globalization and the Gains from Variety," *Quarterly Journal of Economics*, 121(2), 541–585.
- CHANEY, T. (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *The American Economic Review*, 98(4), 1707–1721.
- CHEN, N., J. IMBS, AND A. SCOTT (2009): "The Dynamics of Trade and Competition," *Journal of International Economics*, 77(1), 50–62.
- CORCOS, G., M. DEL GATTO, G. MION, AND G. OTTAVIANO (2011): "Productivity and Firm Selection: Quantifying the "New" Gains From Trade," *Economic Journal*, forthcoming.
- DE BLAS, B., AND K. N. RUSS (2010): "Teams of Rivals: Endogenous Markups in a Ricardian World," *NBER Working Paper*, 16587.
- DEATON, A., AND J. MUELLBAUER (1980): "An Almost Ideal Demand System," *American Economic Review*, 70(3), 312–326.

- DHINGRA, S., AND J. MORROW (2012): "The Impact of Integration on Productivity and Welfare Distortions Under Monopolistic Competition," *mimeo*, LSE.
- EATON, J., AND S. KORTUM (2002): "Technology, Geography and Trade," *Econometrica*, 70(5), 1741–1779.
- EATON, J., S. KORTUM, AND F. KRAMARZ (2011): "An Anatomy of International Trade: Evidence from French Firms," *Econometrica*, 79(5), 1453–1498.
- EDMOND, C., V. MIDRIGAN, AND D. XU (2011): "Competition, Markups and the Gains from Trade," *mimeo*, Duke and NYU.
- FEENSTRA, R. C. (2003): "A Homothetic Utility Function for Monopolistic Competition Models, Without Constant Price Elasticity," *Economics Letters*, 78(1), 79–86.
- FEENSTRA, R. C., AND D. WEINSTEIN (2010): "Globalization, Markups, and the U.S. Price Level," *NBER working paper*, 15749.
- GOLDBERG, P., J. D. LOECKER, A. KHANDELWAL, AND N. PAVCNIK (2012): "Prices, Markups and Trade Reform," *mimeo*.
- HARRISON, A. E. (1994): "Productivity, Imperfect Competition and Trade Reform. Theory and Evidence," *Journal of International Economics*, 36(1-2), 53–73.
- HELPMAN, E., AND P. KRUGMAN (1985): *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy*. MIT Press, Cambridge, Massachusetts.
- HELPMAN, E., AND P. R. KRUGMAN (1989): *Trade Policy and Market Structure*. MIT Press, Cambridge, Massachusetts.
- HOLMES, T. J., W.-T. HSU, AND S. LEE (2010): "Plants and Productivity in Regional Agglomeration," *mimeo*, University of Minnesota.
- KONINGS, J., P. VAN CAYSEELE, AND F. WARZYNSKI (2001): "The Dynamics of Industrial Mark-Ups in Two Small Open Economies: Does National Competition Policy Matter?," *International Journal of Industrial Organization*, 19(5), 841–859.
- KRISHNA, P., AND D. MITRA (1998): "Trade Liberalization, Market Discipline and Productivity Growth: New Evidence from India," *Journal of Development Economics*, 56(2), 447–462.
- KRUGMAN, P. (1979): "Increasing Returns Monopolistic Competition and International Trade," *Journal of International Economics*, 9(4), 469–479.
- KRUGMAN, P. (1980): "Scale Economies, Product Differentiation, and the Pattern of Trade," *The American Economic Review*, 70(5), 950–959.
- LEVINSOHN, J. (1993): "Testing the Imports-as-Market-Discipline Hypothesis," *Journal of International Economics*, 35(1), 1–22.

- LOECKER, J. D., AND F. WARZYNSKI (2012): "Markups and Firm-Level Export Status," *forthcoming American Economic Review*.
- MELITZ, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1695–1725.
- MELITZ, M. J., AND G. I. P. OTTAVIANO (2008): "Market Size, Trade, and Productivity," *The Review of Economic Studies*, 75(1), 295–316.
- NOVY, D. (2010): "International Trade Without CES: Estimating Translog Gravity," *CEP Discussion paper*, 1031.
- OTTAVIANO, G. I., T. TABUCHI, AND J.-F. THISSE (2002): "Agglomeration and Trade Revisited," *International Economic Review*, 43(2), 409–436.
- RODRIGUEZ-LOPEZ, J.-A. (2010): "Prices and Exchange Rates: A Theory of Disconnect," Manuscript, University of California Irvine.
- SAURE, P. (2009): "Bounded Love of Variety and Patterns of Trade," *mimeo*, Swiss National Bank.
- SIMONOVSKA, I. (2009): "Income Differences and Prices of Tradables," Manuscript, University of California, Davis.
- ZHELOBODKO, E., S. KOKOVIN, M. PARENTI, AND J.-F. THISSE (2011): "Monopolistic Competition in General Equilibrium: Beyond the CES," *mimeo*.