Trade Flows, Multilateral Resistance, and Firm Heterogeneity

Alberto Behar[*], Benjamin D Nelson[†]

February 3, 2012

Abstract

We present a gravity model that accounts for multilateral resistance, firm heterogeneity and country-selection into trade, while accommodating asymmetries in trade flows. A new equation for the proportion of exporting firms takes a gravity form, such that the extensive margin is also affected by multilateral resistance. We develop Taylor approximated multilateral resistance terms with which to capture the comparative static effects of changes in trade costs. For isolated bilateral changes in trade frictions, multilateral resistance effects are small for most countries. But if all countries reduce their trade frictions, the impact of multilateral resistance is so strong that bilateral trade falls in most cases, despite the larger trade elasticities implied by firm heterogeneity. As a consequence, the world-wide trade response, though positive, is much lower.

JEL Classifications: F10, F12, F14, F17

Key Words: Gravity models, multilateral resistance, firm heterogeneity.

1 Introduction

What is the impact of changes in trade frictions on trade flows? This paper sheds light on this issue by developing and estimating a gravity model that emphasizes two important forces: firm heterogeneity and multilateral resistance. This framework shows that comparative statics obtained under most existing gravity models are significantly misleading. Omitting firm heterogeneity misrepresents bilateral trade elasticities, whereas omitting multilateral resistance overstates bilateral trade elasticities. The two forces will interact differently depending on whether trade cost changes are bilateral or multilateral in nature and will significantly affect the likely response of aggregate trade flows to declining trade barriers.

Our framework places Helpman, Melitz & Rubinstein (2008) gravity model, which incorporates firm heterogeneity and allows trade responses to occur through firm entry/exit as well as firm-level export volumes, in an Anderson & van Wincoop (2003) general equilibrium setting. This setting allows us to account for multilateral resistance – or the importance of *relative* trade costs in determining trade flows. We demonstrate how multilateral resistance ('MR') works through both the intensive and extensive margins of trade and how the margins affect MR. We do this by adapting Baier & Bergstrand's (2009) method for approximating the effects of MR to the heterogeneous firms setting. Using this framework, this paper makes three contributions.

First, we analyze the case of two countries changing their bilateral trade costs. We find that firm heterogeneity will tend to raise trade elasticities, while MR effects are typically small for most country pairs. Firm heterogeneity together with fixed costs of exporting imply the presence of an extensive margin of trade that amplifies country-level trade responses above those found in standard models. MR works against this but the effects are muted. So, for practical empirical purposes, we find that ignoring MR is relatively innocuous for the majority of country pairs when only two countries change their trade costs. Furthermore, bilateral liberalizations will tend to generate smaller trade responses for larger countries because both the amplification effect of liberalization at the extensive margin will be smaller and dampening effect due to MR will be larger as country size grows.

Second, we turn to multilateral changes in trade costs by analyzing the case where all countries reduce their international trade frictions. Just as for bilateral trade cost reductions, the extensive margin raises trade elasticities, but now the effects of MR are much more important.

Because multilateral liberalizations imply small changes in relative trade costs for any given country pair, trade elasticities net of MR are much smaller. Surprisingly, we show analytically that bilateral elasticities can become negative and empirically that this is true for most country pairs. That is, multilateral trade liberalization will tend to redirect output across destinations, reducing exports to some locations but increasing trade with others. The pattern of responses is such that trade between smaller countries tends to become redirected to larger importers. The reason is that, while bilateral trade becomes more attractive, actual trade flows depend on the costs incurred in this trade relative to other destinations. Because bigger importers are less affected by MR, multilateral changes imply many changes in relative prices that tend to favour exporting to larger importers over smaller importers.

Third, we study the response of world trade to multilateral changes in trade costs by aggregating the bilateral elasticities that properly capture the general equilibrium effects of MR. Our main finding here is that the dampening effect of MR dominates the amplification generated at the extensive margin. The elasticity net of general equilibrium considerations is around one third of the size of the aggregate trade elasticity that ignores relative price effects and less than 40% of those implied by standard models. In our simulations country entry makes no contribution to the increase in world trade and the contribution of firm entry is modest.

Section 2 develops a new gravity model that captures the features of two recent papers in the international trade literature. In the first paper, Anderson and van Wincoop ('AvW'), solve the so-called 'border puzzle' – the implausibly large negative effect of the US-Canadian border on trade between US states and Canadian provinces highlighted by McCallum (1995). AvW do this by showing that traditional gravity equations capture the impact of only bilateral trade costs on trade flows but ignore the fact that regions operate in a multilateral world. As a result, traditional estimates fail to control for theoretically motivated price index terms, which aggregate both domestic and international trade costs, and therefore capture MR. AvW show that bilateral trade flows depend on bilateral trade costs *relative to* MR. Failing to account for MR typically leads one to overstate the importance of changes in trade barriers on bilateral trade flows.

The second paper on which we draw is Helpman, Melitz and Rubinstein (2008) ('HMR'). Heterogeneous firm productivity within a country means not all firms are productive enough to cover the fixed costs of exporting.[1] If fixed costs are high enough, no firms in a given country

---

[1] See Bernard, Jensen, Redding & Schott (2007) for an overview of firms in international trade.

may find it profitable to export to a given destination. Hence in the presence of fixed costs of trade, 'zeros' naturally arise in the trade data — a *country* selection effect.[2] In HMR's data, which we also use, the proportion of countries that do not trade with each other or trade in only one direction is around half of all observations.

HMR explore a further implication. With heterogeneous firm productivity and fixed costs of trade, a fall in variable trade costs makes exporting firms export more, but also induces new firms to export. These two effects are referred to as the *intensive* and the *extensive* margins respectively. HMR argue that failure to account for firm heterogeneity causes standard gravity estimation to conflate the impact of trade costs on these two margins by ignoring this *firm* selection effect.

In section 3, we study theoretical comparative statics in a special case of our hybrid AvW-HMR gravity model – a frictionless initial equilibrium, as studied in AvW – which helps us to illustrate the mechanisms at work in the empirical comparative statics we subsequently compute. Here we show that comparative statics will in general comprise three terms: an intensive margin capturing the firm-level response to trade cost changes, an extensive margin capturing the effects of firm entry, and an adjustment term that captures the effects of MR. We show that changes in trade flows at the extensive margin are also dampened by MR in a general equilibrium setting. These comparative statics are useful for building intuition but are inappropriate for empirical implementation. In particular, the frictionless initial equilibrium is unrealistic, and importantly it abstracts from zeros and asymmetries in the trade data that are known to be important empirically.

Hence in section 4 we develop a Taylor approximation for capturing comparative static MR effects. We centre the approximation at an initial equilibrium with positive, asymmetric trade frictions, which adapts Baier and Bergstrand's (2009) approach to a setting that includes firm heterogeneity. Compared to AvW, Baier and Bergstrand ('BB') show that the approximation error is small for the vast majority of country-pairs and we believe the approach has a number of advantages over computational methods. First, it preserves the possibility of analyzing bilaterally *asymmetric* trade costs. While AvW's approach provides an exact solution in the system they consider, it does so under the assumption of bilaterally *symmetric* trade costs. Bergstrand, Egger and Larch (2007) show that BB's flexible method can be superior to a

---

[2]High fixed costs of exporting from country j to country i do not imply that the absence of trade in the opposite direction, from i to j. These *asymmetric* trade flows are also a salient feature of the data that can be accommodated by modeling firm heterogeneity.

computational approach that assumes bilateral symmetry in trade costs despite the empirical presence of asymmetric trade costs and firm heterogeneity. Following HMR, we want to preserve asymmetric trade costs as a way of explaining the asymmetric trade flows one observes in the data. Second, this method does not require an estimate of the elasticity of substitution between product varieties, which has a large effect on comparative static outcomes (Bergstrand, Egger and Larch, 2011). Third, BB's approach is useful for gaining intuition about the effects at work when trade costs change. We show in a heterogeneous firms setting that multilateral resistance can be represented by three terms that are linear in trade costs, capturing world trade resistance, importer MR, and exporter MR. Having an analytically tractable expression makes the mechanics behind the comparative statics exercises transparent. Following this, our main message is that to employ a first order adjustment for MR is better than applying no adjustment at all.

This message complements recent work by Egger, Larch, Staub & Winkelmann (2011), who demonstrate biases that arise from neglecting the general equilibrium impact of changes in trade costs in the context of multiple Preferential Trade Agreements (PTAs). In that study, estimation addresses the issues of endogeneity, country selection[3] and MR, only the last of which is accounted for in comparative statics. MR comparative static effects of PTAs are analyzed using the Taylor method in Behar & Cirera-i Criville (2011), but without taking firm heterogeneity into account. In contrast, comparative statics in our study explicitly account for firm heterogeneity, general equilibrium and the interactions between the two. Both these features are considered by Behar, Manners & Nelson (2011) in their investigation of logistics quality and exports.

After replicating the estimation exercise in HMR[4], we isolate the effects of firm heterogeneity in the absence of MR. Thereafter, we implement our Taylor approximation empirically by studying in turn bilateral trade cost reductions, multilateral trade cost reductions, and the effect on aggregate world trade elasticities of multilateral trade cost reductions, accounting for both MR and firm heterogeneity. Hence section 4 contains our core analytical and empirical results. Section 5 concludes.

---

[3]The authors consider regressions that aim to control for firm heterogeneity using a related approach to HMR.

[4]HMR's method for accounting for the effects of firm heterogeneity uses a two-step regression procedure, which generates controls for country selection and the proportion of firms trading.

2 The model

The basis of our study is the gravity model proposed by Anderson and van Wincoop ('AvW', 2003) and extended by Helpman, Melitz and Rubinstein ('HMR', 2008) to include firm heterogeneity. The environment is one in which CES consumers within a set of endowment economies demand differentiated products produced by monopolistically competitive firms that differ according to their unit costs. The crucial feature of the HMR gravity set up is the presence of fixed costs of exporting. This generates selection into exporting by the lowest cost firms only. In some cases, trade costs can be so high that no firms in a given country export to a particular destination. This generates both firm selection and country selection into trade and the latter explains the presence of 'zeros' in bilateral trade data.

For brevity, we do not repeat the derivation of HMR's gravity equation here. We begin where HMR left off. Their gravity equation relates imports by country i from N_j firms in country j to variable 'iceberg' trade costs t_{ij}, the fixed total output of the importer Y_i, an index of inward multilateral resistance in the importer P_i, and a selection term V_{ij} which measures the proportion of firms in j that actively export to country i:

$$M_{ij} = \left(\frac{c_j t_{ij}}{\alpha P_i} \right)^{1-\sigma} N_j Y_i V_{ij}.$$

(cf. HMR equation (6)). In this equation, σ denotes the elasticity of substitution between varieties, $\alpha \equiv 1 - (1/\sigma)$ and c_j captures unit labour costs in exporter j.[5] We follow AvW in solving for the endogenous unit costs by closing the model through assuming trade balance for each country. This says that the total output of each exporter must equal the sum of imports across all importing destinations, $Y_j = \sum_{i \in I_j} M_{ij}$, where set I_j denotes the set of countries that import from country j. Using trade balance, HMR's gravity equation is:

$$Y_j = \left(\frac{c_j}{\alpha} \right)^{1-\sigma} N_j \sum_{i \in I_j} \left(\frac{t_{ij}}{P_i} \right)^{1-\sigma} Y_i V_{ij} \Leftrightarrow \left(\frac{c_j}{\alpha} \right)^{1-\sigma} N_j = \frac{Y_j}{\sum_{i \in I_j} \left(\frac{t_{ij}}{P_i} \right)^{1-\sigma} Y_i V_{ij}}$$

Thus solving for unit factor costs and substituting back into the gravity equation yields[6]

$$M_{ij} = \frac{Y_i Y_j}{Y^{I_j}} \left(\frac{t_{ij}}{P_i \widehat{P}_j} \right)^{1-\sigma} V_{ij}, \tag{1}$$

---

[5] In AvW, the equivalent term is their 'small p'.

[6] A full derivation is given in the technical appendix to this paper.

in which the indices of *inward multilateral resistance* P_i and *outward multilateral resistance* $\widehat{P}_j$ are defined respectively by

$$P_i^{1-\sigma} = \sum_{j \in J_i} \left(\frac{t_{ij}}{\widehat{P}_j} \right)^{1-\sigma} s_j^{J_i} V_{ij}, \tag{2}$$

$$\widehat{P}_j^{1-\sigma} = \sum_{i \in I_j} \left(\frac{t_{ij}}{P_i} \right)^{1-\sigma} s_i^{I_j} V_{ij}. \tag{3}$$

Some explanation of notation is in order. We define $Y^{I_j} \equiv \sum_{i \in I_j} Y_i$ as the total output of all importers from country j. It differs from AvW's notation only to the extent that fixed costs of exporting mean that country j firms will not export to all destinations. Using this notation, we similarly define $s_i^{I_j} \equiv Y_i / Y^{I_j}$, such that $s_i^{I_j}$ is the output share of country i as a fraction of the total output of all countries that import from j, I_j. We define the set J_i analogously – as the set of all exporters to country i. Following our convention, we then define $s_j^{J_i} \equiv Y_j / Y^{J_i}$ as the output of country j as a share of the total output of all exporters to country i, J_i. Variable trade costs are weakly greater than unity between countries ($t_{ij} \geq 1$ for $i \neq j$) but are exactly unity for trade within countries ($t_{ij} = 1$ for $i = j$). We do not require that variable trade costs are bilaterally symmetric, which is important for capturing asymmetric trade flow patterns. Finally, the selection term V_{ij} would reduce to a constant in the absence of fixed costs of trade. In the presence of fixed costs, variation in the extensive margin of trade will contribute directly to variation in aggregate bilateral trade flows M_{ij} and indirectly because the MR terms are also a function of the extensive margin.

How is the extensive margin of trade determined? In HMR's set up, there exists a unit cost level above which firms in j find it too costly to export to i, which we denote by a_{ij}. Imposing general equilibrium closure as we do allows us to write this cost cut-off as a function of outputs, trade costs, and MR terms. Interestingly, it takes a gravity-like form, namely:

$$a_{ij}^{\sigma-1} = \frac{1-\alpha}{N_j} \frac{Y_i Y_j}{Y^{I_j}} \frac{1}{f_{ij}} \left(\frac{t_{ij}}{P_i \widehat{P}_j} \right)^{1-\sigma}, \tag{4}$$

where f_{ij} are the fixed costs of exporting from country j to country i. Hence a consequence of requiring general equilibrium closure is that the extensive margin can be expressed in a convenient form akin to equation (1) for bilateral exports. Intuitively, higher fixed costs of trade reduce the cost cut-off, so fewer firms in j export to i. In addition, the cost level above

which firms in j cannot export to i rises as the product of the two countries' outputs rises, as bilateral trade costs fall, or as MR rises. The fact that MR enters equation (4) just as it does for equation (1) highlights the fact that valid comparative statics on the extensive margin require accounting for general equilibrium effects, just as AvW emphasize for valid comparative statics at the aggregate level.

Following HMR, if firm unit costs a are Pareto distributed with shape parameter k on the interval $a \in [a_L, a_H]$, then the extensive margin selection term V_{ij} entering equation (1) can be written

$$V_{ij} \equiv \max \left\{ \overline{k} \left[\left(\frac{a_{ij}}{a_L} \right)^{k-\sigma+1} - 1 \right], 0 \right\}, \tag{5}$$

where $\overline{k} \equiv k/(a_H^k - a_L^k)$. HMR relate the cost cut-off a_{ij} to a latent variable Z_{ij}, which is the ratio of the variable profits of the lowest cost firm a_L in j to the fixed costs of exporting. It is related to the cost cut-off according to

$$Z_{ij} = \left(\frac{a_{ij}}{a_L} \right)^{\sigma-1}. \tag{6}$$

In our general equilibrium setting, this latent variable is in turn related to MR through equation (4). In particular, we can write

$$Z_{ij} = \widetilde{Z}_{ij} \left(P_i \widehat{P}_j \right)^{\sigma-1}, \quad \widetilde{Z}_{ij} \equiv \frac{(1-\alpha)a_L^{1-\sigma}}{N_j} \frac{Y_i Y_j}{Y^{I_j}} \frac{t_{ij}^{1-\sigma}}{f_{ij}}.$$

In turn, this allows us to write the selection term in (5) as

$$V_{ij} = \max \left\{ \overline{k} \left[\widetilde{Z}_{ij}^{\delta} \left(P_i \widehat{P}_j \right)^{\delta(\sigma-1)} - 1 \right], 0 \right\}, \tag{7}$$

where $\delta \equiv (k - \sigma + 1)/(\sigma - 1)$. Equations (1) and (7) then make clear that the comparative static effects of trade costs will operate through three channels. First, changes in variable trade costs will affect bilateral trade through the intensive margin, captured by the t_{ij} term in (1). Second, to the extent that changes in trade costs also affect the cost of trading with alternative export destinations, the multilateral resistances in (1) will change, tending to oppose the direct effects of changes in trade costs through t_{ij}. Third, changes in variable trade costs will also affect the extensive margin of trade through V_{ij}. This final effect is also subject to competing forces. It comprises both direct effects of changes in trade costs through the t_{ij} term in the

numerator of equation (4), which by (6) will affect the extensive margin in (7), together with indirect effects through the MR terms that enter (4). As in the aggregate trade flow equation (1), the MR effects operating through the extensive margin will tend to oppose the direct effects of trade costs through the extensive margin, providing some dampening effect. Overall, then, valid comparative statics must account for these two sources of effects of MR on trade flows.

3   Comparative statics in a frictionless world

The main contribution of this study is to examine the comparative static effects of trade cost changes in the environment described above. Before we turn to the empirical implementation, we begin with a simplified case to build intuition. The thought experiment we perform is identical to one considered by AvW in their framework, which abstracts from firm heterogeneity. To draw a comparison with their results, consider a frictionless world in which variable trade costs $t_{ij} = 1$ for all i, j and in which all countries trade such that $I_j = J_i$. The support of a is such that all firms in country i export and that $V_{ij} = 1$ for all i, j at the initial equilibrium. In this world, importer and exporter multilateral resistances are symmetric and equal to unity, as in AvW's example. Using equations (4) and (5) in the equilibrium we describe, a change in trade costs implies

$$\frac{dV_{ij}}{dt_{ij}} = -(\sigma - 1)\varphi \left(1 - \frac{dP_i}{dt_{ij}} - \frac{d\widehat{P}_j}{dt_{ij}} \right), \tag{8}$$

where $\varphi \equiv \delta(1 + \overline{k})$. This illustrates the dependence of the extensive margin effect on the elasticity of substitution σ, the parameters governing firm heterogeneity captured in φ, and the net effect of MR. In particular, a rise in variable trade costs will reduce the extensive margin of exporting firms by a factor equal to one minus the effect of changes in MR, $1 - dP_i/dt_{ij} - d\widehat{P}_j/dt_{ij}$; times the (adjusted) elasticity of substitution between varieties, $\sigma - 1$; and mediated by a factor φ, which captures the extent of heterogeneity across firms. Thus, when we examine the role of the extensive margin in aggregate level comparative statics, we should observe terms net of MR effects, mediated by the extent of firm heterogeneity, and proportional to the elasticity of substitution.

Using (8) in (1), a general expression for the effect of a change in trade costs on trade flows evaluated at the initial frictionless equilibrium is then given by:

$$\frac{d}{dt_{ij}} \left(M_{ij} \frac{Y}{Y_i Y_i} \right) = -(\sigma - 1)(1 + \varphi) \left(1 - \frac{dP_i}{dt_{ij}} - \frac{d\widehat{P}_j}{dt_{ij}} \right), \tag{9}$$

9

which we can use to understand the components of trade flow comparative statics. Just as the expression for comparative statics at the extensive margin in (8), equation (9) shows that comparative statics on overall trade flows between exporter j and importer i will comprise an intensive margin captured by $\sigma - 1$, some amplification due to firm exit at the extensive margin due to the factor $1 + \varphi > 1$, together with some dampening due to multilateral resistance $1 - dP_i/dt_{ij} - d\widehat{P}_j/dt_{ij}$. The last effect captures the general equilibrium repercussions of changes in trade frictions on average importer resistance and average exporter resistance. On the importer side, trade flows are dampened to the extent that cheaper varieties may now be available from elsewhere (i.e. from exporters other than j), while on the exporter side, dampening may occur because alternative export destinations may now be relatively more attractive (i.e. importers other than i).[7]

We can gain further insight by continuing to explore the MR adjustment term. Using equation (8), we totally differentiate the system of price indices given by (2) and (3) with respect to variable trade costs. Evaluating at the frictionless equilibrium yields:

$$dP_i = \sum_j s_j dt_{ij} - \sum_j s_j d\widehat{P}_j, \tag{10}$$

$$d\widehat{P}_j = \sum_i s_i dt_{ij} - \sum_i s_i dP_i. \tag{11}$$

Combining these two expressions results in the total differential of the MR terms:

$$dP_i + d\widehat{P}_j = -\sum_l \sum_h s_l s_h dt_{lh} + \sum_l s_l dt_{lj} + \sum_h s_h dt_{ih}. \tag{12}$$

This equation summarizes the behavior of endogenously determined MR to exogenous changes in trade costs. It says that a rise in average trade costs across all import and export destinations tends to raise MR, stimulating bilateral trade between i and j (second and third terms), but that the average world trade resistance has also risen, tending to reduce bilateral trade (first term).

Consider first bilateral changes in trade costs. For a bilateral change, let $dt_{lh} = dt_{hl} = dt$

---

[7]While we emphasise the general equilibrium effect as operating through $\widehat{P}_j$, this exporter MR term may be thought of as capturing effects through factor costs, which are endogenous. In particular, a fall in $\widehat{P}_j$ corresponds to a rise in c_j. See the expression above equation (1) which we used to solve for and eliminate c_j.

for $l, h = i, j$, or else let $dt_{lh} = 0$. Then using (9), GDP-weighted exports change according to

$$\frac{d}{dt}\left(M_{ij}\frac{Y}{Y_iY_i}\right) = -(\sigma-1)(1+\varphi)\left(1+2s_is_j-s_i-s_j\right). \tag{13}$$

Note that firm heterogeneity provides an amplification factor through the term $(1+\varphi)$ when trade costs change, tending to raise the elasticity of trade with respect to trade costs relative to AvW's homogeneous firms case. We can also see from this expression that, whenever $s_i < 1/2$, larger countries experience smaller trade elasticities under bilateral change in trade costs. Moreover, the effects of MR will typically be small under bilateral changes in trade costs, as the term $1 + 2s_is_j - s_i - s_j$ will typically be close to unity. Intuitively, since only one set of trade costs changes under bilateral liberalization, relative price changes are captured fairly well by absolute price changes, such that only a small adjustment due to MR is required.

Consider next the effect of a multilateral change in trade costs, as in AvW, such that $dt_{ij} = dt > 0$ for all $i \neq j$, and $dt_{ii} = 0$ for all i. Using equation (9) and our expressions for the total differentials in (8) and (12), GDP-weighted exports change according to

$$\frac{d}{dt}\left(M_{ij}\frac{Y}{Y_iY_i}\right) = -(\sigma-1)(1+\varphi)\left(-\sum_l s_l^2 + s_j + s_i\right). \tag{14}$$

This corresponds precisely to AvW's system (see their equation (15)), with the addition of a term reflecting firm heterogeneity, φ. Compared to their model, the firm heterogeneity term means the trade elasticity will be larger by a factor $1+\varphi > 1$. But as in AvW's model, it is the case that MR dampens the effects of multilateral trade cost changes, and by more for smaller countries. As such, larger countries experience larger trade elasticities under multilateral changes in trade costs. The sign of the net effect of a trade liberalization in exports is ambiguous in this world, and the source of this ambiguity is the effect of MR through the term $-\sum_l s_l^2 + s_j + s_i$. Small country pairs (for which s_i+s_j is small) may well have negative bilateral elasticities. In principle, firm heterogeneity will magnify the impact of any negative elasticities arising from MR.

The analysis around the frictionless equilibrium has introduced the intuition for our main results. Firm heterogeneity introduces an additional factor $\varphi > 0$ into trade elasticities. For bilateral changes in trade costs, MR effects could be minor but, for multilateral changes, MR effects could be sufficiently strong to make some trade elasticities negative.

We want to measure these effects, but the problem with implementing these theoretical expressions empirically is that the world is obviously far from frictionless. To the contrary, trade

costs are significant (Anderson & van Wincoop 2004). Moreover, although the expressions above incorporate firm heterogeneity, we lose an important dimension of variation across countries because the φ term has no country-pair subscript. Finally, the expressions above do not allow the zeros in trade flows to play a role in determining likely trade responses. As is clear from (2) and (3), zeros in the trade data imply that the indices of MR should be adjusted to account for the sets of active traders. Concretely, this means that the output shares captured by the s_j terms in the expressions above should be computed with respect to the sets of countries with which exporters actually trade when trade elasticities are generated, rather than all countries in the world.

4   Empirical comparative statics

For estimation purposes, the methodology developed in HMR deals with firm heterogeneity and country fixed effects control for importer and exporter multilateral resistance.[8] We replicate HMR's approach to estimation. For comparative statics however, we adapt Baier and Bergstrand's ('BB', 2009) method for approximating MR effects to a heterogeneous firms setting. In an homogeneous firms setting, BB accomplish this by taking a Taylor approximation for the non-linear MR terms around a world of symmetric but positive trade frictions. Their method allows us to expunge the endogenous components from the right hand side of the MR terms, such that approximate comparative statics are straightforward to implement.

AvW use a computational procedure to solve for MR explicitly and to compute comparative statics. Despite awareness of this procedure and the first order consequences of MR, the computational approach of AvW has not been widely adopted.[9] This is largely due to the difficulty in implementation, which requires a customized program. Our application would be especially complex because the system we study includes *asymmetric* bilateral trade costs, firm selection, and different sets of active traders for different trading country pairs. Like BB, our primary motive is to present an approach to computing comparative statics that has low barriers to im-

---

[8] The use of exporter and importer fixed effects in estimation is acknowledged by AvW, Baier & Bergstrand (2009) and Feenstra (2004) to be the most reliable estimation method in this context.

[9] There are isolated contributions but these are still somewhat limited in country coverage – for example modeling US and Canadian regions together with a limited set of additional countries or an aggregate for the rest of the world. Such studies typically conduct counterfactuals on binary variable like borders or free trade agreements (AvW; Anderson & Yotov (2010); Anderson & Yotov (2011)). To our knowledge, the most ambitious exercise is that of Egger et al. (2011), who on 128 countries allow for MR, entry and endogenous free trade agreements in estimation and account for MR in comparative statics. However, the comparative statics extensive margin is not explicitly dealt with, nor is its interaction with MR.

plementation in empirical gravity work. It successfully corrects for the first order inaccuracies associated with the general equilibrium effects of trade cost changes. Baier & Bergstrand (2009) show that their approximation is good for the majority of region pairs that they consider. In particular, comparative statics of economic integration based on the Taylor method are within 10% of the AvW approach for 83% of their pairs.[10]

Moreover, the method allows us to get closer to reality by dropping the assumption of *symmetric* bilateral trade costs, which AvW use in implementing their approach. The ability to account for asymmetric bilateral trade costs is a major advantage in accounting for asymmetric trade flows, which are a central characteristic of the data (HMR). Simulation analysis in Bergstrand, Egger & Larch (2007) concludes that the Taylor method performs better than the AvW approach when asymmetries are relevant, and is particularly important in a world of firm heterogeneity. Further, the approach does not require us to assume particular values for structural parameters like the elasticity of substitution, σ. Bergstrand, Egger & Larch (2011) show that comparative statics are very sensitive to the choice of this elasticity parameter and that the AvW approach is inaccurate when the assumed elasticity is different from the actual elasticity.[11] A final advantage of the BB method is its intuitive appeal. This, combined with the frictionless case just considered, allows us to be clearer about the mechanisms that lie behind the comparative statics we produce.

To apply BB's method to the case of firm heterogeneity, we make the following assumption for the purposes of comparative statics:[12]

Decomposability: *The extensive margin terms V_{ij} entering the indices of multilateral resistance $\{P_i, \widehat{P}_j\}$ are approximately $V_{ij} \simeq Z_{ij}^\delta \left(P_i \widehat{P}_j \right)^{\delta(\sigma-1)}$.*[13]

Using this, a Taylor expansion of importer and exporter multilateral resistances yields a

---

[10] They also *estimate* the parameters using OLS, so the approach to comparative statics would presumably perform even better when using fixed effects in estimation.

[11] Hertel, Hummels, Ivanic & Keeney (2007) discuss the wide range of σ estimates and the sensitivity of general equilibrium comparative statics to the value of σ.

[12] The assumption is not required for the purposes of estimation, so HMR's empirical procedure is not affected.

[13] HMR make a similar assumption in their Appendix II. They do so for V_{ij} in the main gravity equation while we do so only in the price index terms. Furthermore, they assume that the ij component of V_{ij} is symmetric. Their assumption is that $V_{ij} = \left(\phi_{ij} \phi_i \widehat{\phi}_j \right)^{1-\sigma}$, in which $\phi_{ij} = \phi_{ji}$, while ϕ_i and $\widehat{\phi}_j$ are importer and exporter specific effects. The assumption that $\phi_{ij} = \phi_{ji}$ precludes asymmetric trade flows, and for that reason is rejected by the authors. By contrast, we do not impose such symmetry. In particular, we allow for $\phi_{ij} \neq \phi_{ji}$ around our centre. In our case, this implies it is possible that $\widetilde{Z}_{ij} \neq \widetilde{Z}_{ji}$.

tractable expression for trade resistance that comprises three terms (see appendix):

$$(\delta + 1) \ln \left(P_i \widehat{P}_j \right)^{\sigma-1} \simeq -\overbrace{\sum_{l \in I_j} s_l^{I_j} \sum_{h \in J_l} s_h^{J_l} \left[(\sigma-1) \ln t_{lh} - \delta \widetilde{z}_{lh} \right]}^{\text{World Trade Resistance}}$$

$$+ \underbrace{\sum_{h \in J_i} s_h^{J_i} \left[(\sigma-1) \ln t_{ih} - \delta \widetilde{z}_{ih} \right]}_{\text{Importer's MR}} + \underbrace{\sum_{l \in I_j} s_l^{I_j} \left[(\sigma-1) \ln t_{lj} - \delta \widetilde{z}_{lj} \right]}_{\text{Exporter's MR}}, \quad (15)$$

in which $\widetilde{z}_{ij} \equiv \ln \widetilde{Z}_{ij}$. This expression has three terms which, by analogy with (12), can be straightforwardly interpreted. First, to the extent that world trade resistance is high, MR between the i-j trading pair will be low, tending to discourage international trade on average. Second, to the extent that importer i faces high trade costs in obtaining output from other exporters in the set J_i, exporter j will export more to i. Third, when exporter j finds it costly to export to other destinations in the set I_j, it will tend to export more to i instead, all else equal. Hence the expression above has a straightforward interpretation that makes clear the impact of *relative* trade costs on bilateral trade flows.

Moreover, our expression for MR incorporates the effects of firm heterogeneity in two respects. First, trade resistance occurring through the extensive margin is captured by the presence of the $\widetilde{z}_{ij}$ terms on the right hand side. Their presence makes clear, for example, that MR is affected by the intensive and extensive margins. Together with our discussion of the role of MR in the extensive margin in section 3, this indicates a two-way interaction between MR and the margins of trade. Second, the relevant average trade costs that constitute MR in a world of firm heterogeneity are taken over the sets of active traders. This is clear from the fact that the summation terms are taken over the sets J_i and I_j. This will have important implications for comparative statics as we will see below.

The log-gravity equation for the model is:

$$m_{ij} = y_i + y_j - y^{I_j} - (\sigma-1) \ln t_{ij} + w_{ij} + (\sigma-1) \ln P_i \widehat{P}_j, \quad (16)$$

where lower cases denote logs and where we follow HMR in capturing $\ln V_{ij}$ by

$$w_{ij} \equiv \ln \left\{ \exp \delta \left[\widetilde{z}_{ij} + \ln \left(P_i \widehat{P}_j \right)^{\sigma-1} \right] - 1 \right\}. \quad (17)$$

The estimation procedure we use is exactly as in HMR. It involves a two-step process whereby

a first stage probit regression is estimated for the probability that country j exports to country i. This probability is then used to construct (a) a control variable for the extensive margin, w_{ij}, and (b) the Inverse Mills Ratio as an additional control for country selection into trade. Both of these are used in the second stage regression, which takes the form of an otherwise standard gravity equation in which fixed effects control for MR in estimation. Since the procedure and data is exactly as in HMR, we provide an overview of the data, method and selected regression output in the appendix.

We turn instead to the implementation of our comparative statics, for which we need three things. First, we need an estimate of the 'firm level' intensive margin elasticity of trade flows to trade costs. We proxy trade costs with a number of observable variables and, for the elasticity, we take an estimate of the coefficient on log distance, γ. The estimates are $\hat{\gamma}_p = 0.66$ in the probit model and $\hat{\gamma} = 0.8$ in the 2nd stage regression.[14] Second, we need country-pair estimates of the extensive margin elasticities $\varphi_{ij} \equiv \partial w_{ij}/\partial z_{ij} = \delta e^{\delta z_{ij}}/(e^{\delta z_{ij}} - 1) > 0$, which we denote by $\widehat{\varphi}_{ij}$. The estimate of z_{ij} is a linear function of the the variables in the probit model and $\hat{\delta} = 0.72$ is estimated by non-linear least squares (NLS) in the second stage. The φ_{ij} terms clearly vary by country pair and they provide one source of cross country variation in trade elasticities. Finally, to make the adjustments for MR, we need country GDP shares defined relative to the relevant sets of active traders, $\{s_i^{J_i}, s_j^{I_j}\}$ for all i, j. These are easily generated from the matrix of observed trade flows.

4.1 Gross elasticities

Throughout our application it will be interesting to compare *gross elasticities*, which do not account for MR, with *net elasticities*, which do. Because gross elasticities ignore MR, they will be identical for both bilateral and multilateral changes in trade costs. The gross elasticity for country pair i-j is:

$$\xi_{ij}^{gross} = (\sigma - 1)(1 + \varphi_{ij}) \tag{18}$$

where $\xi_{ij}^{gross} \equiv -\partial m_{ij}/\partial t_{ij}|_{P_i, \widehat{P}_j}$. For the empirical analogue of this elasticity based on a 10% fall in distance, $\widehat{\xi}_{ij}^{gross} = \widehat{\gamma}(1 + \widehat{\varphi}_{ij})$, we reproduce the results from HMR in Table 1. The first row produces results from the NLS estimate derived from the Pareto assumption. The mean overall elasticity of 1.56 implies the bilateral country-level effect is due roughly

---

[14]We have made no theoretical distinction between the intensive margin γ and $\gamma_p = \frac{\partial \widehat{z}_{ij}^*}{\partial d_{ij}}$, which is the distance coefficient from the probit estimate. Our simulations take account of this small difference.

Functional Form	Mean	Median	Std. dev.	Max	Min
Pareto	1.56	1.47	0.29	3.78	1.28
Polynomial	1.85	1.85	0.52	3	1.14

Table 1: Summary statistics for estimates of gross country-level elasticity ξ_{ij}^{gross}

in equal parts to the intensive ($\hat{\gamma} = 0.8$) and the extensive margins, which is consistent with Behar, Manners & Nelson (2011). The estimate is higher than would be implied by OLS of 1.18 (see appendix) or by estimates attempting to deal with country selection alone (1.21 – see HMR). This is important because it implies one should allow for firm heterogeneity even if one is not interested in the composition into the margins.[15] Cross country variation in the extensive margin drives variation in the estimated gross elasticities, which have a standard deviation of 0.29. HMR also produce estimates derived from a polynomial function of z_{ij}.[16] We include these estimates for completeness and to show that the implications for bilateral elasticities are similar, albeit with a larger standard deviation.

These estimates consider only actively trading pairs, but, as is clear from the probit estimates, a fall in distance raises the predicted probability of a pair of countries trading. To translate this continuous effect on the probability into a comparative static simulation of binary country entry, a new trading pair is formed (expanding the set J_i) if the reduction in distance results in the most productive firm in j now being able to cover the fixed costs of exporting to i. Empirically, this means that the predicted value of z_{ij} after a fall in distance, $\widehat{\widetilde{z}}_{ij}^{\prime*}$, is positive when it previously was negative. Our analysis counts in how many cases pairs that were not previously trading now have $\widehat{\widetilde{z}}_{ij}^{\prime*} > 0$.

For the default 10% fall in distance, not a single instance of country entry exists. Generating a case of country entry required fall in log distance of 0.31, or 27%. This implies that the distribution of distance and other observed and unobserved trade costs is such that it would take very large changes in trade costs to generate new trading pairs. This is consistent with the observation in HMR that very little of the increase in world trade observed over time is due to the formation of new trading pairs. Similarly, we considered a rise in trade costs, estimating that a rise in log distance of 0.35 would have no effect on country entry and a rise in log distance of 0.4 would cause only one trading pair to stop trading. The results imply that we do not need

---

[15]This point on country-level comparative statics is distinct from the emphasis in HMR on the estimate of γ, which is lower when controlling for firm heterogeneity than when using OLS. Baranga (2009) and Belenkiy (2009) discuss the robustness of the HMR result.

[16]Egger et al (2011) also adopt a variation of the polynomial approach in estimation (but not comparative statics).

to consider changes in the sets of traders when studying the net elasticities for the baseline 10% change in trade costs that we do next, beginning with bilateral changes.

4.2 Bilateral changes in trade costs

Allowing trade costs to change bilaterally results in the following:

Proposition 1 *(Bilateral changes in trade costs) The elasticity of bilateral trade flows to bilateral (B) changes in variable trade costs is given by*

$$\xi_{ij}^B = (\sigma - 1)(1 + \varphi_{ij}) \left(1 + s_i^{I_j} s_j^{J_i} + s_j^{I_j} s_i^{J_j} - s_j^{J_i} - s_i^{I_j} \right) \tag{19}$$

where $\xi_{ij}^B \equiv -dm_{ij}/d \ln t_{ij}$ for country pair i-j. Bilateral trade elasticities are always decreasing in country size for bilateral changes in trade costs.

For bilateral changes in trade costs, the elasticity of trade flows is empirically always positive. Moreover, the impact of MR will typically be small, because the adjustment made to the elasticity due to general trade frictions will *typically* be close to unity. There will be exceptions however. Consider the impact of a trade deal between a small developing country exporter and the US. In this case, $s_i^{I_j}$ is large, while $s_j^{I_j}$, $s_j^{J_i}$ and $s_j^{I_i}$ are very small. The bilateral elasticity will be approximately $(\sigma - 1)(1 + \varphi_{ij}) \left(1 - s_i^{I_j} \right)$ such that the adjustment due to MR is large. For country pairs such as these, ignoring MR effects even for bilateral changes need not be innocuous. For the majority of country pairs however, the effects of MR under bilateral changes in trade costs will be small. This is highly intuitive given the skewness in the distribution of country incomes such that most countries are 'small' relative to the mean. Since the essence of MR is to account for GDP-weighted changes in relative trade costs, bilateral changes between two typical countries have small general equilibrium effects. A corollary is that, when trade costs change only bilaterally, the relative cost of importing from or exporting to alternative destinations also changes. Then we can predict that the impact of the bilateral reduction in trade frictions will be to stimulate bilateral trade.

Since both the MR effect and the elasticity at the extensive margin are decreasing in country size, the impact of size on the trade elasticities is unambiguous in the case of isolated changes in trade costs. Small country pairs liberalizing trade bilaterally will tend to experience larger elasticities.

Group		GDP Share		Elasticity		Ratio
		Exporter	Importer	Gross ($\widehat{\xi}_{ij}^{gross}$)	Net ($\widehat{\xi}_{ij}^{B}$)	$\widehat{\xi}_{ij}^{B}/\widehat{\xi}_{ij}^{gross}$
Mean		1.269%	1.200%	1.56	1.53	0.97
Median		0.186%	0.112%	1.47	1.45	0.99
Mauritania	Togo	0.004%	0.005%	2.04	2.04	0.9999
USA	Japan	29.23%	15.74%	1.28	1.02	0.79
Mexico	Spain	1.721%	1.717%	1.29	1.25	0.97

Table 2: Elasticity for a bilateral reduction in trade frictions ξ_{ij}^{B}

We implement the empirical counterpart of the net elasticity in equation (19) for distance, $\xi_{ij}^{B} = \gamma + \varphi_{ij}\gamma_p - \left(1 + \varphi_{ij}\right)\frac{\gamma + \delta\gamma_p}{1+\delta}\left(-s_i^{I_j}s_j^{J_i} - s_j^{I_j}s_i^{J_j} + s_j^{J_i} + s_i^{I_j}\right)$,[17] using our estimated parameters and plotting the results in Figure 1. The chart shows the net elasticities generated in our sample against the gross elasticities. The overwhelming majority of observations form a pattern indistinguishable from a 45 degree line. This shows bilateral changes in trade costs have small MR implications, and so net ($\widehat{\xi}_{ij}^{B}$) and gross ($\widehat{\xi}_{ij}^{gross}$) elasticities are close together. Table 2 produces examples for illustration. The first two rows give the mean and median values for GDP shares and elasticities. The typical elasticity after accounting for MR is close to the gross elasticity. The table includes the ratio $\widehat{\xi}_{ij}^{B}/\widehat{\xi}_{ij}^{gross} \approx 1 + s_i^{I_j}s_j^{J_i} + s_j^{I_j}s_i^{J_j} - s_j^{J_i} - s_i^{I_j}$.[18] The mean value of this ratio is 0.97 and the median value is 0.99, which confirms that the typical ratio is close to unity, indicating a small material effect due to MR in the case of isolated bilateral changes.

Tiny countries like Mauritania and Togo have a ratio of $\widehat{\xi}_{ij}^{B}/\widehat{\xi}_{ij}^{gross} \simeq 0.9999$. By contrast, Japan and the USA comprised 45% of world GDP in the 1986 data that HMR study and generate a ratio of $\widehat{\xi}_{ij}^{B}/\widehat{\xi}_{ij}^{gross} \simeq 0.79$. This suggests that MR effects are large for big countries, even for bilateral changes in trade costs. Mexico and Spain were the 10th and 11th biggest countries in the world in our sample, but they each had less than 2% of world GDP. Their ratio of 0.97 is still sufficiently close to unity to suggest it is only "very" big countries for which MR matters for bilateral liberalizations. A related implication of MR is that trade agreements signed between developing countries will have a larger effect than trade agreements signed between developing and developed countries (Behar and Cirera-i-Criville, 2011).

More generally, bigger countries have lower gross elasticities through the extensive margin *and* lower net-to-gross ratios through MR, giving lower net elasticities overall. This pattern

---

[17]If the estimates in the probit ($\hat{\gamma}_p$) and second stage ($\hat{\gamma}$) are the same, then this simplifies to $\xi_{ij}^{B} = \gamma\left(1 + \varphi_{ij}\right)\left\{1 + s_i^{I_j}s_j^{J_i} + s_j^{I_j}s_i^{J_j} - s_j^{J_i} - s_i^{I_j}\right\}$
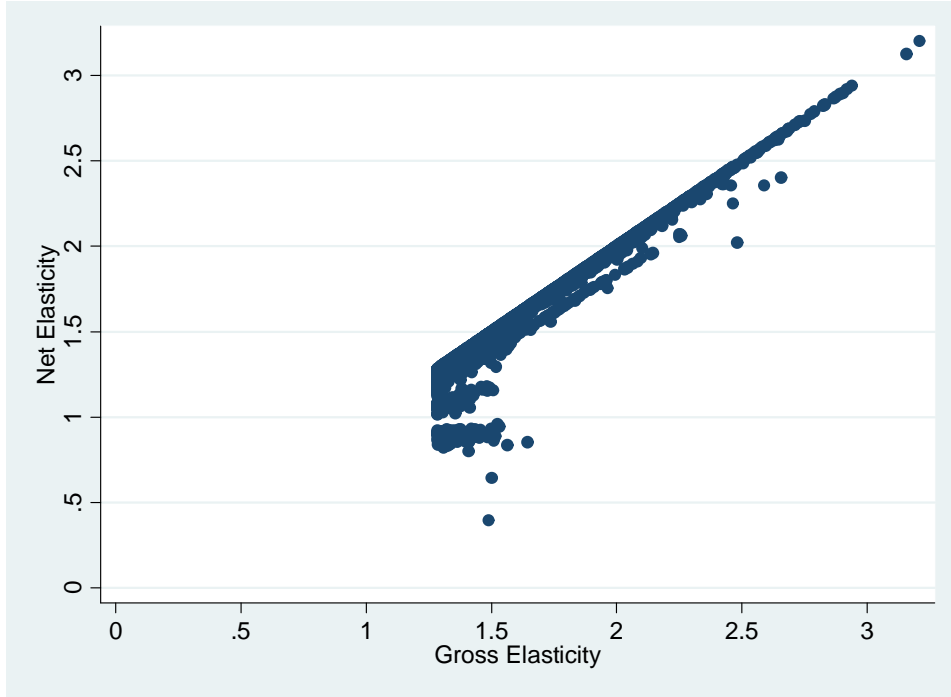
[18]The equality is exact for $\gamma = \gamma_p$.

Figure 1: Effects of MR for bilateral changes in distance: x-axis plots gross elasticity $\widehat{\xi}_{ij}^{gross}$; y-axis plots net elasticity $\widehat{\xi}_{ij}^{B}$.

is further illustrated by simple correlation coefficients. There is a negative correlation of -0.29 between $\widehat{\xi}_{ij}^{gross}$ and country-pair size (measured by summing their shares of world GDP). The correlation between $\widehat{\xi}_{ij}^{B}$ and country size is -0.51, illustrating that the effect of MR reinforces that of the extensive margin, as suggested in Proposition 1. Figure 2 illustrates this negative relationship.

While the USA and Japan are the biggest in terms of world GDP shares, Table 3 shows lower net-to-gross ratios are generated by small (and distant) countries exporting to the largest countries, which we considered in theory above. For example, the USA comprised 43% of the share of the GDP of all importers from French Guiana and the elasticity ratio is only $\widehat{\xi}_{ij}^{B}/\widehat{\xi}_{ij}^{gross} \simeq 0.57$. Japan makes up 72% of importers from Bhutan and the net-to-gross ratio is only 0.26. Thus, for bilateral changes in trade costs, MR effects are material for (a) trade between the world's largest country pairs and (b) exports from small exporters with few export destinations to the world's largest countries. Otherwise they are small.

4.3   Multilateral changes in trade costs

We turn next to changes in trade costs that are multilateral in nature. We find that:
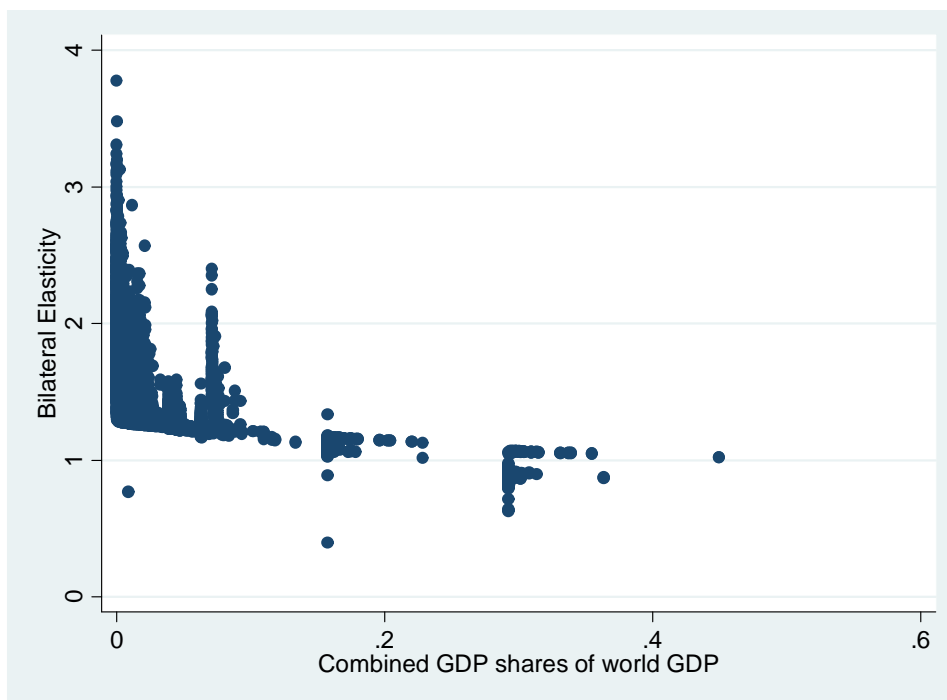
Figure 2: Net elasticity $(\widehat{\xi}_{ij}^B)$ and country-pair GDP share when bilateral distance falls in isolation.

| Countries | | GDP shares* | | | Ratio |
Exporter	Importer	Combined	Exporter	Importer	$\widehat{\xi}_{ij}^B/\widehat{\xi}_{ij}^{gross}$
Bhutan	Japan	15.75%	72.34%	0.001%	0.26
Eq. Guinea	USA	29.23%	56.10%	42.85%	0.43
Kiribati	USA	29.23%	46.78%	51.92%	0.52
Solomon Isl.	USA	29.23%	45.38%	53.58%	0.54
French Guiana	USA	29.24%	42.76%	56.73%	0.57

* 'Combined' shares do not adjust for active traders; 'exporter' and 'importer' shares do.

Table 3: Lowest bilateral elasticities.

Proposition 2 (*Multilateral changes in trade costs*) *The elasticity of bilateral trade flows to multilateral changes in variable trade costs is given by*

$$\xi_{ij}^M = (\sigma - 1)(1 + \varphi_{ij})\left(-\sum_{l \in I_j} s_l^{I_j} s_j^{J_l} + s_i^{J_i} + s_j^{I_j}\right) \tag{20}$$

where $\xi_{ij}^M \equiv -dm_{ij}/d\ln t$, $d\ln t_{ij} = d\ln t$ for all i, j, $i \neq j$. Then:

(a) *after accounting for effects through multilateral resistance, it is the case that*

$$\xi_{ij}^M \gtrless 0, \tag{21}$$

such that the sign of the elasticity of bilateral trade is ambiguous in theory. In particular, for some country pairs, it could be negative;

(b) (i) *if the extensive margin is held constant ($\varphi_{ij} = 0$),*

$$\frac{\partial \xi_{ij}^M\big|_{\varphi_{ij}=0}}{\partial s_i} > 0, \tag{22}$$

such that countries exporting to larger importers have larger firm-level responses to multilateral trade liberalizations;

(b) (ii) *if the extensive margin is allowed to change ($\varphi_{ij} > 0$),*

$$\frac{\partial \xi_{ij}^M}{\partial s_i} \gtrless 0, \tag{23}$$

such that the relationship between importer size and the country-level bilateral export elasticity is ambiguous in theory.

In contrast to the case of bilateral changes in trade costs, the elasticity of bilateral trade flows with respect to multilateral changes in trade costs can be either positive or negative in theory. If it is positive, it is likely to be far below the gross elasticity that one would compute ignoring MR, given by $(\sigma-1)(1+\varphi_{ij})$. The theoretical origin of this result lies in the endowment economy nature of the environment studied here, as in AvW's set up. Changes in trade costs serve to reallocate output from one activity (e.g. domestic trade) to another (e.g. international trade) so that multilateral trade cost changes can mean that exports become redirected from

one destination to another even if the bilateral friction between a country pair falls. A fall in bilateral trade can be explained by the overall bilateral cost rising *relative* to the cost of importing from or exporting to alternative destinations.[19]

Part (b) (i) of Proposition 1 replicates AvW's 'Implication 1' in the environment in which we Taylor approximate MR. It states that countries exporting to larger importers have larger elasticities of bilateral trade when multilateral trade liberalization takes place. The reason is that larger countries typically trade a smaller fraction of their output internationally, instead trading proportionately more domestically. This means that large countries are less affected by MR, which in turn means that the dampening effect of MR on trade elasticities is smaller for large countries. This remains true in our case as long as the extensive margin does not respond to trade liberalizations ($\varphi_{ij} = 0$).

Part (b) (ii) of the Proposition shows what happens when the extensive margin can respond to multilateral trade cost changes too. Now the relationship between country size and elasticity breaks down, and this is due to the way in which the elasticity at the extensive margin varies with country size. To see this, note that the elasticity of the extensive margin can be written $\partial w_{ij}/\partial \ln t_{ij} = -(\sigma-1)\varphi_{ij} + \varphi_{ij}\left[\partial \ln\left(P_i\widehat{P}_j\right)^{\sigma-1}/\partial \ln t_{ij}\right]$, where the first term is the gross effect of a fall in trade costs on the extensive margin, which yields the net extensive margin elasticity once MR is adjusted for. Taking the first term in isolation, it can be shown that $\frac{\partial \varphi_{ij}}{\partial Y_j} > 0$, such that for given MR, the elasticity of the extensive margin with respect to trade costs is *decreasing* in the size of the exporter. From equation (4), large exporters tend also to have high cost cutoffs, all else equal. Therefore a reduction in trade costs has a small proportional effect on the number of exporting firms. Conversely, the extensive margin is very sensitive to trade cost changes for small countries. Hence as country size changes, there are two opposing effects at work. First, larger countries are less subject to MR, so will tend to have larger trade elasticities, all else equal. But larger countries will also be less sensitive to trade costs at the extensive margin, tending to reduce their trade elasticities. The net effect is ambiguous, as detailed in Part (b) (ii) of Proposition 1.

---

[19]In parallel, the worldwide reduction in trade frictions is associated with a rise in factor costs, which reduce demand. To see the interplay between MR and factor costs, consider what happens to trade between i and j when trade barriers between j and a third country $h \neq i$ fall, but the bilateral barrier between i and j remains unchanged. This manifests itself as a fall in j's exporter MR. At the same time, higher demand from the third country bids up j's factor costs. The higher factor costs make j's goods more expensive in i and, given the barriers between i and j remain constant, this lowers demand from i. So, the fall in exporter MR makes supplying goods to i less attractive in relative terms, while the rise in factor costs reduces demand by i. This general equilibrium channel can be strong enough to reduce bilateral trade between i and j even when the i-j trade barrier is not constant and actually falls along with other 'third country' trade frictions, as in a multilateral liberalisation.
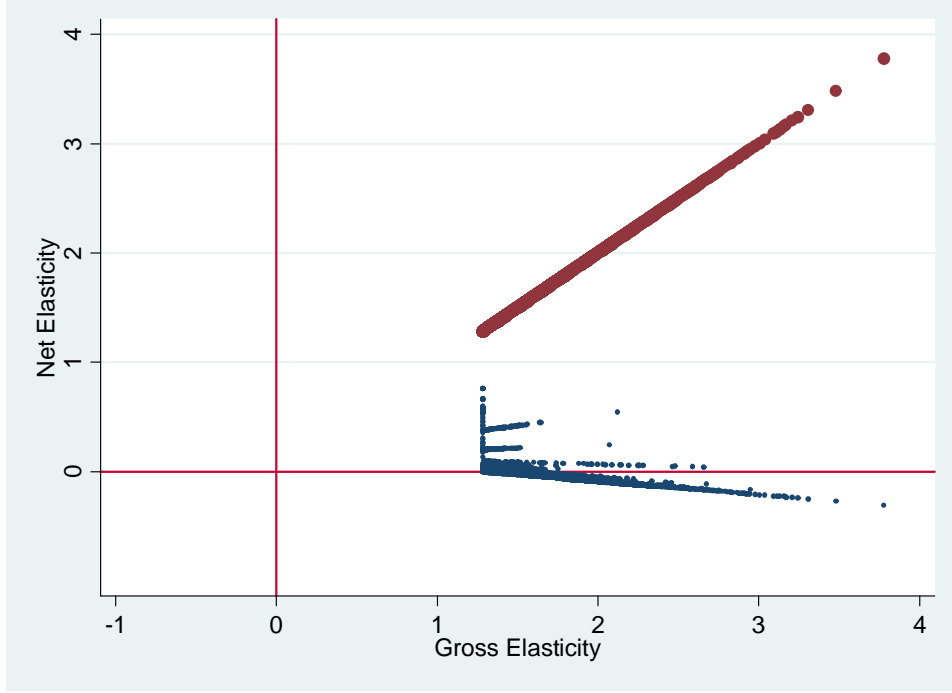
Figure 3: Effects of MR when all countries reduce frictions. Line with positive slope plots $\widehat{\xi}_{ij}^{gross} = \widehat{\xi}_{ij}^{gross}$ for reference. X-axis is $\widehat{\xi}_{ij}^{gross}$; y-axis is $\widehat{\xi}_{ij}^{M}$.

Figure 3 demonstrates the dramatic effects of accounting for MR when *all* countries reduce distance.[20] Estimates of net elasticities $\widehat{\xi}_{ij}^{M}$ are substantially lower than gross elasticities $\widehat{\xi}_{ij}^{gross}$. The majority of the net elasticities are negative, which demonstrates that Proposition 2 (a) is relevant empirically. This implies that the general equilibrium effects are so strong that most country pairs reduce their bilateral trade after reductions in their frictions, and trade is redirected elsewhere. Consistent with this, Table 4 shows the median elasticity is -0.019, while the mean elasticity is 0.006. Because international frictions have fallen only relative to domestic frictions, one should expect a net effect close to zero.

Can we characterize the source of the MR effects and see where they are strongest? Table 4 hints at a positive relationship along the real line between the net-to-gross ratio $\widehat{\xi}_{ij}^{M} / \widehat{\xi}_{ij}^{gross} \approx -\sum_{l \in I_j} s_l^{I_j} s_j^{J_l} + s_i^{J_i} + s_j^{I_j}$ and the size of the trading pair. Mauritania and Togo have a negative net-to-gross ratio, Mexico and Spain are just above zero, while the USA and Japan have a value of of $\widehat{\xi}_{ij}^{M} / \widehat{\xi}_{ij}^{gross} \simeq 0.5$. The theoretical analysis introduced the opposing forces in the relationship between country size and the elasticity: bigger pairs have bigger MR multipliers but lower gross elasticities due to the extensive margin. Recall the correlation between pair-size

---

[20] We compute this as $\xi_{ij}^{M} = \gamma + \varphi_{ij}\gamma_p - (1 + \varphi_{ij}) \frac{\gamma + \delta\gamma_p}{1+\delta} \left\{ 1 + \sum_{l \in I_j} s_l^{I_j} s_j^{J_l} - s_i^{J_i} - s_j^{I_j} \right\}$, i.e. allowing for $\gamma \neq \gamma_p$ between the probit and second stage regressions.

Group		GDP share		Elasticity		Ratio
		Exporter	Importer	Gross $\left(\widehat{\xi}_{ij}^{gross}\right)$	Net $\left(\widehat{\xi}_{ij}^{M}\right)$	$\widehat{\xi}_{ij}^{M}/\widehat{\xi}_{ij}^{gross}$
Mean		1.269%	1.200%	1.56	0.006	0.011
Median		0.186%	0.112%	1.47	-0.019	-0.013
Mauritania	Togo	0.004%	0.005%	2.04	-0.094	-0.047
USA	Japan	29.23%	15.74%	1.28	0.657	0.512
Mexico	Spain	1.721%	1.717%	1.29	0.020	0.015

Table 4: Elasticity for a multilateral reduction in trade frictions ξ_{ij}^{M}
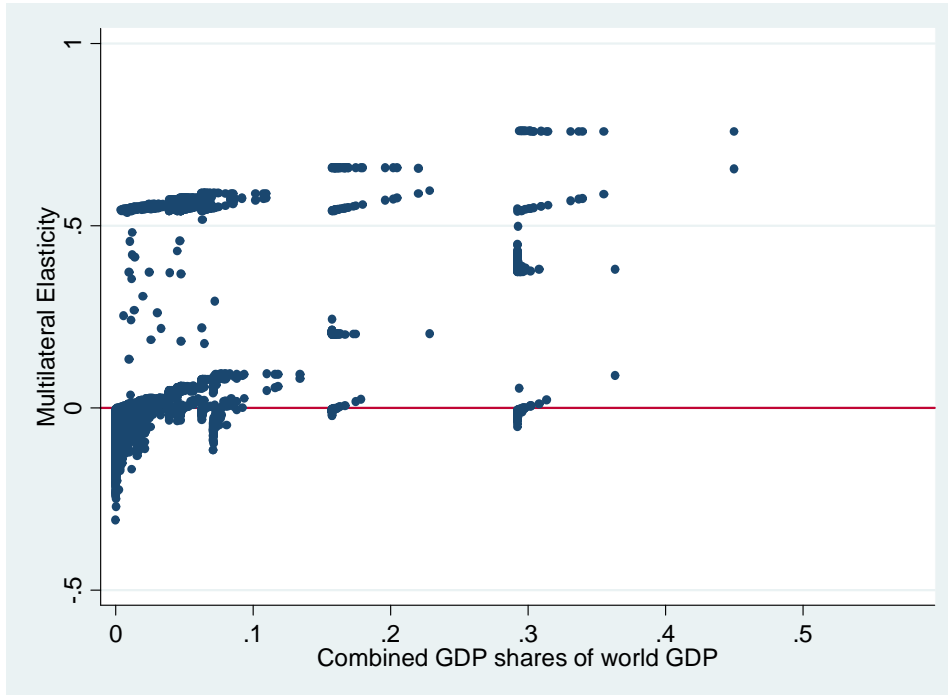

Figure 4: Net elasticity $\widehat{\xi}_{ij}^{M}$ and country-pair GDP share when all countries reduce frictions.

and the gross elasticity is -0.29. For multilateral changes, the correlation between the net-to-gross ratio and country-pair size is 0.49. Figure 3 demonstrates the overall positive relationship across all country-pairs. Thus, while weakened by the extensive margin effects, country size and bilateral trade elasticities are positively related for multilateral changes to trade costs in our data set.

Overall, we have approximately 2,300 positive elasticities out of about 11,000 positive trade flows. This begs the question of where the positive elasticities are generated. Above we discussed how negative bilateral responses are the outcome of redirection from some export destinations to others. Given the endowment economy model studied here, negative export elasticities with some destinations should have offsetting positive elasticities with others. For consistency with theory therefore, it is necessary for each exporter to have *at least one* import destination with

which it has a positive export response. This happens for every country in our sample. G7 importers are responsible for half of the positive elasticities, because most nations increase exports to these destinations. The results reflect major *trade redirection*. Putting it starkly, most countries trade a lot with the G7 and a little with everyone else. After a reduction in trade frictions, they reduce their exports to most smaller countries to be able to expand their exports to a handful of big countries.

4.4 World-wide elasticities

Finally, we aggregate the bilateral elasticities by exporter, weighting each bilateral elasticity by the volume of exports. As should be the case, after aggregating across all export destinations, all countries see an aggregate increase in trade even after accounting for MR. The mean aggregate elasticity is 0.29 and one third of our countries have elasticities of less than 0.1, which are a fraction of the elasticity implied by standard approaches. Canada (0.65) and Mexico (0.67) generate the biggest elasticities - and we know what they have in common.

Some studies (e.g. Wilson, Mann & Otsuki (2005)) use gravity models to simulate the effects of multilateral reductions in trade barriers on global trade. It is interesting to calculate similar world-wide elasticities, which we report in Table 5. For reference, we include the OLS estimate, which ignores both MR and firm heterogeneity. Focussing on the Pareto estimates, the intensive margin effect given by the estimate of γ is 0.799 for the whole world. Allowing for the extensive margin of trade to operate and aggregating across all bilateral elasticities yields a gross world trade elasticity $\widehat{\xi}^{gross}$ of 1.29. This is close to the minimum of the distribution of bilateral elasticities discussed above because the world's largest economies, which dominate the aggregate elasticity, tend to have the lowest individual elasticities. For global elasticities (unlike bilateral elasticities: cf table 1), the implications are that the intensive margin accounts for over 60% of the trade increase and that the gross elasticity is only about 10% higher than that implied by OLS. Recall that there was no effect on country entry of a 10% fall in distance.

Once we account for MR effects, however, aggregation yields a world-wide net trade elasticity $\widehat{\xi}^{M}$ of 0.47. Hence, for the world as a whole, the net-to-gross ratio $\widehat{\xi}^{M}/\widehat{\xi}^{gross}$ is approximately one third and the net elasticity is less than 40% of the elasticity implied by OLS. We draw the same conclusions from an analogous empirical implementation using the polynomial estimates. Therefore, existing gravity-based simulations of multilateral reductions in trade barriers can be seriously misleading.

Functional Form	Elasticities						Ratios		
	OLS	Intensive	Extensive	Country	*Gross*	Net	*Gross: intensive.*	*Net: gross.*	Net: OLS
Pareto	1.176	0.799	0.492	0	1.291	0.467	1.616	0.362	0.397
Polynomial	1.176	0.862	0.425	0	1.287	0.462	1.493	0.359	0.392

Table 5: Impact of firm heterogeneity and MR on world-wide elasticities

In sum, Table 5 confirms that MR dramatically reduces the responsiveness of world trade to a multilateral reduction in trade frictions. However, this masks the reorientation taking place bilaterally. We saw that most bilateral trade elasticities are negative, despite the larger gross elasticities generated by effects at the extensive margin. This dramatic impact is pervasive but greater amongst smaller trading pairs. This is consistent with the findings of AvW and shows the redirection of exports away from most export destinations towards the G7.

The analyses in this section have been of multilateral and bilateral reductions in bilateral (or dyadic) trade frictions. Behar, Manners & Nelson (2011) study the impact of a rise in country-specific (or monadic) logistics quality on exports. Echoing the intuition for the multilateral case here, bigger countries experience a larger trade response to an improvement in their logistics because they are less affected by MR. This effect outweighs the smaller response by bigger countries at the extensive margin. A further implication of MR in that study is that the regression coefficient on a monadic variable applies to a country of average size. Because most countries are smaller than average, MR makes the trade response much smaller for most countries.

5   Concluding remarks

We have presented a gravity model for which we have computed comparative statics that account for the effects of both firm heterogeneity and multilateral resistance. To do so, we imposed general equilibrium closure on Helpman, Melitz and Rubinstein's (2008) gravity model, rendering it comparable to Anderson and van Wincoop's (2003) system while preserving asymmetries in trade frictions and zeroes in the trade data. Adapting Baier and Bergstrand's approximate multilateral resistance (MR) terms to the case of firm heterogeneity allowed us to compare the effects of bilateral and multilateral trade liberalizations in a context that accounts for both the extensive margin of trade and the importance of relative trade costs in determining equilibrium

responses.

In general, the presence of firm heterogeneity tends to amplify the magnitude of trade cost changes on trade flows. Therefore, even if one is not interested in the margins of trade, they need to be considered when estimating the overall trade response. These effects will typically be dampened by MR however. Except for the largest countries, the dampening due to MR effects is small for bilateral changes in trade costs. One implication is that most analyses of trade agreements between two countries (and by extension a handful of countries) can ignore MR for practical purposes.

By contrast, large effects due to MR emerge after multilateral changes in trade costs, for example a comprehensive new multilateral trade agreement, are considered. The theoretical sign of the bilateral trade response is ambiguous and we find empirically that most bilateral trade elasticities are negative. Because large countries are less affected by MR, exports get reoriented from small importers to large importers.

A further implication of this is that the world trade response to a multilateral liberalization is positive but less than 40% of that implied by standard approaches that ignore MR. Properly accounting for MR may contribute, for example, to explanations of why dramatic worldwide transport technology improvements did not make a commensurately large contribution to trade increases over time (Behar & Venables 2011).

References

Anderson, J. & van Wincoop, E. (2003), 'Gravity with gravitas: a solution to the border puzzle', *The American Economic Review* **93**(1), 170–192.

Anderson, J. & van Wincoop, E. (2004), 'Trade costs', *Journal of Economic Literature* **42**(3), 691–751.

Anderson, J. & Yotov, Y. (2010), 'The changing incidence of geography', *The American Economic Review* **100**(5), 2157–2186.

Anderson, J. & Yotov, Y. (2011), 'Terms of trade and global effciency effects of free trade agreements, 1990-2002', *mimeo* .

Baier, S. & Bergstrand, J. (2009), 'Bonus vetus OLS: a simple method for approximating international trade-cost effects using the gravity equation', *Journal of International Economics* **77**(1), 77–85.

Baranga, T. (2009), 'Unreported trade flows and gravity equation estimation', *mimeo.* .

Behar, A. & Cirera-i Criville, L. (2011), 'Does it matter who you sign with? comparing the impacts of north-south and south-south trade agreements on bilateral trade', *World Bank Policy Research Working Paper* .

Behar, A., Manners, P. & Nelson, B. (2011), 'Exports and international logistics', *World Bank Policy Research Working Paper* .

Behar, A. & Venables, A. (2011), *Transport costs and international trade*, in Handbook of Transport Economics, Edward Elgar, p. 97.

Belenkiy, M. (2009), 'The extensive margin in the industry trade: Estimation, significance and implications', *mimeo.* .

Bergstrand, J., Egger, P. & Larch, M. (2007), 'Gravity redux: structural estimation of gravity equations with asymmetric bilateral trade costs', *mimeo* .

Bergstrand, J., Egger, P. & Larch, M. (2011), 'Gravity redux: estimation of gravity equation coefficients, elasticities of substitution, and general equilibrium comparative statics under asymmetric bilateral trade costs', *mimeo* .

Bernard, A., Jensen, J., Redding, S. & Schott, P. (2007), 'Firms in international trade', *Journal of Economic Perspectives* **21**(3), 105–130.

Egger, P., Larch, M., Staub, K. & Winkelmann, R. (2011), 'The trade effects of endogenous preferential trade agreements', *American Economic Journal: Economic Policy* **3**(3), 113–143.

Feenstra, R. (2004), *Advanced international trade: theory and evidence*, Princeton Univ Pr.

Helpman, E., Melitz, M. & Rubinstein, Y. (2008), 'Estimating trade flows: trading partners and trading volumes', *The Quarterly Journal of Economics* **123**(2), 441–487.

Hertel, T., Hummels, D., Ivanic, M. & Keeney, R. (2007), 'How confident can we be of CGE-based assessments of free trade agreements?', *Economic Modelling* **24**(4), 611–635.

McCallum, J. (1995), 'National borders matter: Canada-US regional trade patterns', *The American Economic Review* **85**(3), 615–623.

Wilson, J., Mann, C. & Otsuki, T. (2005), 'Assessing the benefits of trade facilitation: a global perspective', *The World Economy* **28**(6), 841–871.

A   Technical Appendix

A.1   Theory

A.1.1   Derivation of gravity equation

From HMR's equation:

$$M_{ij} = \left(\frac{c_j t_{ij}}{\alpha P_i} \right)^{1-\sigma} N_j Y_i V_{ij},$$

assume trade balance, which gives

$$Y_j = \left(\frac{c_j}{\alpha} \right)^{1-\sigma} N_j \sum_{i \in I_j} \left(\frac{t_{ij}}{P_i} \right)^{1-\sigma} Y_i V_{ij} \Leftrightarrow \left(\frac{c_j}{\alpha} \right)^{1-\sigma} N_j = \frac{Y_j}{\sum\limits_{i \in I_j} \left(\frac{t_{ij}}{P_i} \right)^{1-\sigma} Y_i V_{ij}}.$$

Substitute for $(c_j/\alpha)^{1-\sigma} N_j$ in HMR's gravity equation to yield

$$M_{ij} = \frac{Y_j Y_i}{\sum\limits_{i \in I_j} \left(\frac{t_{ij}}{P_i} \right)^{1-\sigma} Y_i V_{ij}} \left(\frac{t_{ij}}{P_i} \right)^{1-\sigma} V_{ij}.$$

Divide and multiply by $Y^{I_j} \equiv \sum_{i \in I_j} Y_i$, then we have the gravity equation a la AvW, but which adjusts for the extensive margin of trade and sets of active traders, given by:

$$M_{ij} = \frac{Y_j Y_i}{Y^{I_j}} \left(\frac{t_{ij}}{P_i \widehat{P}_j} \right)^{1-\sigma} V_{ij},$$

where

$$\widehat{P}_j^{1-\sigma} \equiv \sum_{i \in I_j} \left(\frac{t_{ij}}{P_i} \right)^{1-\sigma} s_i^{I_j} V_{ij}$$

is exporter multilateral resistance in which $s_i^{I_j} \equiv Y_i / Y^{I_j}$. From the consumer price index, we have also that importer multilateral resistance is:

$$P_i^{1-\sigma} = \sum_{j \in J_i} \left(\frac{c_j t_{ij}}{\alpha} \right)^{1-\sigma} N_j V_{ij} = \sum_{j \in J_i} \left(\frac{t_{ij}}{\widehat{P}_j} \right)^{1-\sigma} s_j^{J_i} r_{ij} V_{ij},$$

where $s_j^{J_i} \equiv Y_j / Y^{J_i}$ and $r_{ij} \equiv Y^{J_i} / Y^{I_j}$. The latter term r_{ij} captures the relative size of all exporters to i relative to all importers from j. For the purposes of the comparative statics we consider empirically, this term is a constant if the two sets J_i and I_j do not change. As we discuss in the body of the paper, for the changes in distance that we consider, these sets of traders are indeed constant. For our comparative statics then, we can subsume the r_{ij} into the constant in the gravity equation and use

$$P_i^{1-\sigma} = \sum_{j \in J_i} \left(\frac{t_{ij}}{\widehat{P}_j} \right)^{1-\sigma} s_j^{J_i} V_{ij}$$

to complete the general equilibrium system.

A.1.2   Taylor approximation

Following Baier and Bergstrand, we take a first-order Taylor approximation of the MR indices. Our model contains an additional effect due to the impact of MR on the extensive margin: the price indices contain extensive margin terms, which themselves contain a system of price indices. In order to extract the exogenous components of the system, we make the further assumption that the extensive margin

terms that appear in the price index equations (but not the main gravity equation) can be decomposed according to $V_{ij} \simeq \widetilde{Z}_{ij}^{\delta} \left(P_i \widehat{P}_j \right)^{\delta(\sigma-1)}$, or $w_{ij} \simeq \delta \widetilde{z}_{ij} + \ln \left(P_i \widehat{P}_j \right)^{\delta(\sigma-1)}$. This approximated extensive margin value approaches that computed under the assumption of a Pareto distribution for δz_{ij} sufficiently large, since $w_{ij} = \ln \left\{ e^{\delta z_{ij}} - 1 \right\} \simeq \delta z_{ij} = \delta \widetilde{z}_{ij} + \ln \left(P_i \widehat{P}_j \right)^{\delta(\sigma-1)}$ when this is the case. Using this, and defining $W_{ij} \equiv \widetilde{Z}_{ij}^{\delta}$ yields price indices of

$$\widehat{P}_j^{-k} = \sum_{i \in I_j} t_{ij}^{1-\sigma} s_i^{I_j} P_i^k W_{ij} \tag{1}$$

$$P_i^{-k} = \sum_{j \in J_i} t_{ij}^{1-\sigma} s_j^{J_i} \widehat{P}_j^k W_{ij}, \tag{2}$$

where we have used that $(\delta+1)(\sigma-1) = (\frac{k-\sigma+1}{\sigma-1}+1)(\sigma-1) = k$. The terms (1) and (2) are those which we Taylor approximate. We do this around a symmetric centre in which trade frictions are positive but symmetric and sufficiently low for all countries to trade. Then $t_{ij} = t > 1$ for all i, j. Similarly, $W_{ij} = W$, $P_i = P$, $I_j = J_i$ and $\widehat{P}_j = \widehat{P}$. This yields the system

$$\widehat{P}^{-k} = t^{1-\sigma} P^k W \sum_i s_i = t^{1-\sigma} P^k W,$$

$$P^{-k} = t^{1-\sigma} \widehat{P}^k W \sum_j s_j = t^{1-\sigma} \widehat{P}^k W,$$

which has a solution at $P = \widehat{P}$. Any observed variations in t_{ij} are pertubations from this centre which can be approximated using Taylor's method. These can be asymmetric in the sense that $t_{ij} \neq t_{ji}$. To do this, we use

$$e^{-k \ln \widehat{P}_j} = \sum_{i \in I_j} e^{\ln s_i^{I_j}} e^{(1-\sigma) \ln t_{ij}} e^{k \ln P_i} e^{\ln W_{ij}}, \tag{3}$$

$$e^{-k \ln P_i} = \sum_{j \in J_i} e^{\ln s_j^{J_i}} e^{(1-\sigma) \ln t_{ij}} e^{k \ln \widehat{P}_j} e^{\ln W_{ij}}, \tag{4}$$

Then expanding the LHS of (4) around $\ln P$ and the RHS around $\ln t$, $\ln W$, and $\ln \widehat{P}$ gives

$$P^{-k} - k P^{-k} [\ln P_i - \ln P] = \sum_{j \in J_i} s_j^{J_i} t^{1-\sigma} W \widehat{P}^k + \sum_{j \in J_i} s_j^{J_i} (1-\sigma) t^{1-\sigma} W \widehat{P}^k [\ln t_{ij} - \ln t]$$

$$+ \sum_{j \in J_i} s_j^{J_i} t^{1-\sigma} W \widehat{P}^k [\ln W_{ij} - \ln W] + \sum_{j \in J_i} s_j^{J_i} k t^{1-\sigma} W \widehat{P}^k [\ln \widehat{P}_j - \ln \widehat{P}].$$

Divide through by $P^{-k} = t^{1-\sigma} W \widehat{P}^k$ and use that $\sum_{j \in J_i} s_j^{J_i} = 1$ to give

$$-k[\ln P_i - \ln P] = \sum_{j \in J_i} s_j^{J_i} (1-\sigma)[\ln t_{ij} - \ln t]$$

$$+ \sum_{j \in J_i} s_j^{J_i} [\ln W_{ij} - \ln W] + \sum_{j \in J_i} s_j^{J_i} k [\ln \widehat{P}_j - \ln \widehat{P}]$$

$$= \sum_{j \in J_i} s_j^{J_i} (1-\sigma) \ln t_{ij} - (1-\sigma) \ln t$$

$$+ \sum_{j \in J_i} s_j^{J_i} \ln W_{ij} - \ln W + \sum_{j \in J_i} s_j^{J_i} k \ln \widehat{P}_j - k \ln \widehat{P}.$$

Use $-k \ln P = (1-\sigma) \ln t + \ln W + k \ln \widehat{P}$ to give

$$\ln P_i = \sum_{j \in J_i} s_j^{J_i} \left(\frac{\sigma-1}{k} \right) \ln t_{ij} - \frac{1}{k} \sum_{j \in J_i} s_j^{J_i} \ln W_{ij} - \sum_{j \in J_i} s_j^{J_i} \ln \widehat{P}_j.$$

2

Writing the approximated price indices in terms using subscript h as the generic exports and l as the generic importer, we have

$$\ln P_i = \sum_{h\in J_i} s_h^{J_i}\left(\frac{\sigma-1}{k}\right)\ln t_{ih} - \frac{1}{k}\sum_{h\in J_i} s_h^{J_i}\ln W_{ih} - \sum_{h\in J_i} s_h^{J_i}\ln\widehat{P}_h,$$

$$\ln\widehat{P}_j = \sum_{l\in I_j} s_l^{I_j}\left(\frac{\sigma-1}{k}\right)\ln t_{lj} - \frac{1}{k}\sum_{l\in I_j} s_l^{I_j}\ln W_{lj} - \sum_{l\in I_j} s_l^{I_j}\ln P_l.$$

Substituting the first line into the second gives

$$\ln\widehat{P}_j = \frac{1}{k}\sum_{l\in I_j} s_l^{I_j}\left[(\sigma-1)\ln t_{lj} - \ln W_{lj}\right]$$

$$-\sum_{l\in I_j} s_l^{I_j}\left[\frac{1}{k}\sum_{h\in J_l} s_h^{J_l}\left[(\sigma-1)\ln t_{lh} - \ln W_{lh}\right] - \sum_{h\in J_l} s_h^{J_l}\ln\widehat{P}_h\right],$$

such that adding $\ln P_i + \ln\widehat{P}_j$ yields

$$\ln P_i + \ln\widehat{P}_j = -\frac{1}{k}\sum_{l\in I_j} s_l^{I_j}\sum_{h\in J_l} s_h^{J_l}\left[(\sigma-1)\ln t_{lh} - \ln W_{lh}\right]$$

$$+\frac{1}{k}\sum_{h\in J_i} s_h^{J_i}\left[(\sigma-1)\ln t_{ih} - \ln W_{ih}\right] + \frac{1}{k}\sum_{l\in I_j} s_l^{I_j}\left[(\sigma-1)\ln t_{lj} - \ln W_{lj}\right]$$

Using $(\sigma-1)\ln t_{ij} = \gamma d_{ij}$ and $\ln W_{ij} = \delta\widetilde{z}_{ij}$, our MR approximation is then given by

$$\ln P_i\widehat{P}_j = -\frac{1}{k}\sum_{l\in I_j} s_l^{I_j}\sum_{h\in J_l} s_h^{J_l}\left[\gamma d_{lh} - \delta\widetilde{z}_{lh}\right] + \frac{1}{k}\sum_{h\in J_i} s_h^{J_i}\left[\gamma d_{ih} - \delta\widetilde{z}_{ih}\right] + \frac{1}{k}\sum_{l\in I_j} s_l^{I_j}\left[\gamma d_{lj} - \delta\widetilde{z}_{lj}\right].$$

Our comparative statics require the term $\frac{\partial\ln(P_iP_j)^{\sigma-1}}{\partial d} = (\sigma-1)\frac{\partial\ln(P_iP_j)}{\partial d}$, so we write

$$(\sigma-1)\ln P_i\widehat{P}_j$$

$$= \frac{(\sigma-1)}{k}\left\{-\sum_{l\in I_j} s_l^{I_j}\sum_{h\in J_l} s_h^{J_l}\left[\gamma d_{lh} - \delta\widetilde{z}_{lh}\right] + \sum_{h\in J_i} s_h^{J_i}\left[\gamma d_{ih} - \delta\widetilde{z}_{ih}\right] + \sum_{l\in I_j} s_l^{I_j}\left[\gamma d_{lj} - \delta\widetilde{z}_{lj}\right]\right\},$$

and use $\delta = \frac{k-\sigma+1}{\sigma-1}$, such that $(\delta+1)(\sigma-1) = k$, finally yielding the approximation:

$$(\sigma-1)\ln P_i\widehat{P}_j$$

$$= \frac{1}{\delta+1}\left\{-\sum_{l\in I_j} s_l^{I_j}\sum_{h\in J_l} s_h^{J_l}\left[\gamma d_{lh} - \delta\widetilde{z}_{lh}\right] + \sum_{h\in J_i} s_h^{J_i}\left[\gamma d_{ih} - \delta\widetilde{z}_{ih}\right] + \sum_{l\in I_j} s_l^{I_j}\left[\gamma d_{lj} - \delta\widetilde{z}_{lj}\right]\right\}.$$

A.1.3   Proof of Proposition 1: Bilateral changes

Proof. The trade elasticity for a bilateral change is given by

$$\xi_{ij}^B = (\sigma-1) + \varphi_{ij}(\sigma-1) - \left(1+\varphi_{ij}\right)\frac{\partial\ln\left(P_i\widehat{P}_j\right)^{\sigma-1}}{\partial d_{ij,ji}},$$

$$\frac{\partial\ln\left(P_i\widehat{P}_j\right)^{\sigma-1}}{\partial d_{ij,ji}} = (\sigma-1)\left\{-s_i^{I_j}s_j^{J_i} - s_j^{I_j}s_i^{J_j} + s_j^{J_i} + s_i^{I_j}\right\}, \tag{5}$$

$$\Rightarrow \xi_{ij}^B = (\sigma-1)\left(1+\varphi_{ij}\right)\left\{1 + s_i^{I_j}s_j^{J_i} + s_j^{I_j}s_i^{J_j} - s_j^{J_i} - s_i^{I_j}\right\}.$$

3

or $\xi_{ij}^B = \gamma \left(1 + \varphi_{ij}\right) \left\{1 + s_i^{I_j} s_j^{J_i} + s_j^{I_j} s_i^{J_j} - s_j^{J_i} - s_i^{I_j}\right\}$ if $\gamma = \sigma - 1$. For a change in the size of importer i such that $\partial s_i^{J_i} = \partial s_i^{I_j} = \partial s_i$,

$$\frac{\partial \xi_{ij}^B}{\partial s_i} = \frac{\partial\left(\gamma\varphi_{ij}\right)}{\partial s_i} \left\{1 + s_i^{I_j} s_j^{J_i} + s_j^{I_j} s_i^{J_j} - s_j^{J_i} - s_i^{I_j}\right\} + \gamma\left(1 + \varphi_{ij}\right)\left\{s_j^{J_i} + s_j^{I_j} - 1\right\}.$$

Use that $\frac{\partial \widetilde{w}_{ij}}{\partial d_{ij}} \equiv -\gamma\varphi_{ij}$ so $\frac{\partial\left(\gamma\varphi_{ij}\right)}{\partial s_i} = -\frac{\partial^2 \widetilde{w}_{ij}}{\partial d_{ij} \partial s_i} = -\frac{\partial^2 \widetilde{w}_{ij}}{\partial d_{ij} \partial Y_i} \frac{\partial Y_i}{\partial s_i} < 0$ to argue

$$\frac{\partial \xi_{ij}^B}{\partial s_i} < 0 \text{ if } s_j^{J_i} + s_j^{I_j} < 1$$

$\blacksquare$

A.1.4 Proof of Proposition 2: Multilateral Changes

Proof. The trade elasticity for a multilateral change is given by

$$\xi_{ij}^M = (\sigma - 1) + \varphi_{ij}(\sigma - 1) - \left(1 + \varphi_{ij}\right)\frac{\partial \ln\left(P_i \widehat{P}_j\right)^{\sigma-1}}{\partial d},$$

$$\frac{\partial \ln\left(P_i \widehat{P}_j\right)^{\sigma-1}}{\partial d} = (\sigma - 1)\left\{1 + \sum_{l \in I_j} s_l^{I_j} s_j^{J_l} - s_i^{J_i} - s_j^{I_j}\right\},$$

$$\Rightarrow \xi_{ij}^M = (\sigma - 1) + \varphi_{ij}(\sigma - 1) - \left(1 + \varphi_{ij}\right)(\sigma - 1)\left\{1 + \sum_{l \in I_j} s_l^{I_j} s_j^{J_l} - s_i^{J_i} - s_j^{I_j}\right\}.$$

This yields

$$\xi_{ij}^M = (\sigma - 1)\left(1 + \varphi_{ij}\right)\left(-\sum_{l \in I_j} s_l^{I_j} s_j^{J_l} + s_i^{J_i} + s_j^{I_j}\right),$$

or $\xi_{ij}^M = \gamma\left(1 + \varphi_{ij}\right)\left(-\sum_{l \in I_j} s_l^{I_j} s_j^{J_l} + s_i^{J_i} + s_j^{I_j}\right)$ if $\gamma = \sigma - 1$. Then:

(a) Ignoring MR gives the gross elasticity as $\gamma\left(1 + \varphi_{ij}\right) > 0$ but including the MR term $\left(-\sum_{l \in I_j} s_l^{I_j} s_j^{J_l} + s_i^{J_i} + s_j^{I_j}\right)$ implies $\xi_{ij}^M \gtrless 0$ since $\left(-\sum_{l \in I_j} s_l^{I_j} s_j^{J_l} + s_i^{J_i} + s_j^{I_j}\right) \gtrless 0$.

(b) (i) If the extensive margin does not operate, the elasticity is $\left.\xi_{ij}^M\right|_{\varphi_{ij}=0} = \gamma\left(-\sum_{l \in I_j} s_l^{I_j} s_j^{J_l} + s_i^{J_i} + s_j^{I_j}\right)$, such that $\frac{\partial \left.\xi_{ij}^M\right|_{\varphi_{ij}=0}}{\partial s_i} = \gamma\left(1 - s_j^{J_i}\right)$, which is positive if $s_j^{J_i} < 1$.

(b) (ii) If the extensive margin operates, the sign of $\frac{\partial \xi_{ij}^M}{\partial s_i}$ is subject to two competing effects. Using $\frac{\partial\left(\gamma\varphi_{ij}\right)}{\partial s_i} = -\frac{\partial^2 \widetilde{w}_{ij}}{\partial d_{ij} \partial s_i} = -\frac{\partial^2 \widetilde{w}_{ij}}{\partial d_{ij} \partial Y_i}\frac{\partial Y_i}{\partial s_i} < 0$, this can be seen from

$$\frac{\partial \xi_{ij}^M}{\partial s_i} = \underbrace{-\frac{\partial^2 \widetilde{w}_{ij}}{\partial d_{ij} \partial s_i}}_{-\text{ve}}\underbrace{\left(-\sum_{l \in I_j} s_l^{I_j} s_j^{J_l} + s_i^{J_i} + s_j^{I_j}\right)}_{+\text{ve or } -\text{ve}} + \underbrace{\gamma\left(1 + \varphi_{ij}\right)\left(1 - s_j^{J_i}\right)}_{+\text{ve}}.$$

The MR effect from (b) (i) increases the elasticity for larger countries, but larger countries experience smaller direct effects of trade costs on the extensive margin. This second effect can go the opposite way if $-\sum_{l \in I_j} s_l^{I_j} s_j^{J_l} + s_i^{J_i} + s_j^{I_j} < 0$, yielding a positive elasticity overall, but whether $-\sum_{l \in I_j} s_l^{I_j} s_j^{J_l} + s_i^{J_i} + s_j^{I_j} < 0$ is ambiguous. The response of the overall elasticity to changes in country size is therefore ambiguous so that $\frac{\partial \xi_{ij}^M}{\partial s_i} \lessgtr 0$. $\blacksquare$

A.2 Empirical implementation

A.2.1 Description of variables

The dependent variable is the log of exports in 1986 measured in constant (2000) US dollars. The main variables used are distance: the log of distance in km between the importer and exporter; border: a binary variable indicating whether the country pair shares a common physical boundary; island: a binary variable indicating whether at least one country is an island (in HMR, this is described as an indicator of whether both countries are an island); landlocked: a binary variable indicating whether at least one country is landlocked (in HMR, this is described as an indicator of whether both countries are landlocked); legal: a binary variable indicating whether or not the country pair share the same legal origin; language: a binary variable indicating whether the country pair share a common language; colonial ties: a binary variable indicating whether one country every colonised the other; FTA: a binary variable indicating whether or not the country pair formed a regional trade agreement; religion: a variable constructed by HMR indicating how similar the religious composition is in the country pair;[1] exporter and importer dummies to capture fixed effects including multilateral resistance.

To construct GDP Shares, GDP data are sourced from the World Bank's World Development Indicators WDI). Where necessary, WDI data were combined to meet the country definitions in the trade data (for example combining Belgium and Luxembourg). Where possible, missing observations were supplemented with data from the United Nations Common Database (UNCDB). Otherwise, a GDP of 0.1 was inputted manually. The denominator is constructed based on the sum of the individual countries' GDPs. Shares are based on the subset of countries the importer imports from or the exporter exports to. Except in the example of a small exporter with few trading partners exporting to a large exporter, we have achieved similar results simply by using the share of world GDP, where world GDP is the value provided by the WDI.

A.2.2 Description of empirical methodology

Here, we describe the methodology used for estimation and the comparative statics exercises.

Estimation

1. Estimate a probit model for the probability of positive exports from j to i. The first stage includes the bilateral variables listed in the data appendix together with importer and exporter dummies. The set of bilateral variables must include a variable which will be omitted from the second stage. Theoretically, such a variable should affect the fixed costs of exporting, but not variable costs. HMR use religion for this purpose.[2] The probit generates the estimate of $\hat{\gamma}_p = 0.66$ in column 2 below and noted in the text.

2. Generate predicted probabilities ($\hat{\rho}$) of positive exports. Many of these are practically indistinguishable from unity such that many values of the Inverse Mills Ratios are not distinguished from each other. The approach in HMR converts all $\hat{\rho} > 0.9999999$ to values of $\hat{\rho} = 0.9999999$. For a further discussion of truncation issues, see Baranga (2009).

3. Predict the Inverse Mills Ratio $\widehat{\bar{\eta}}_{ij}^* = \frac{\phi(\widehat{z}_{ij}^*)}{\Phi(\widehat{z}_{ij}^*)}$, where $\phi(.)$ is the the standard normal density function and $\Phi(.)$ is the standard normal distribution function. Generate $\widehat{z}_{ij}^*$ as a linear function of the probit variables and $\widehat{\bar{z}}_{ij}^* = \widehat{z}_{ij}^* + \widehat{\bar{\eta}}_{ij}^*$.

4. Estimate the second stage including the predicted controls for firm selection $\log(e^{\delta \widehat{\bar{z}}_{ij}^*} - 1)$, derived from the assumption of Pareto distributed firm productivity, and country-selection $\widehat{\bar{\eta}}_{ij}^*$. Helpman, Melitz & Rubinstein (2007) use maximum likelihood while the published version (Helpman, Melitz & Rubinstein (2008)) use non-linear least squares. We follow HMR08. Because of the large number of dummies (over 300), we follow Carayol (2006) and use the function evaluator program version of the nl command. This generates estimates of $\widehat{\delta} = 0.72$ and $\gamma = 0.8$ in column 3 and noted in the text. HMR also develop a semi-parametric controls for firm selection using a polynomial in $\widehat{\bar{z}}_{ij}^*$; results for this are presented in column 4.

---

[1] We use the version employed by HMR in their paper. An updated version has subsequently become available on Elhanan Helpman's website but does not affect any of our results.

[2] Some countries export to everyone else or import from everyone else. The fixed effects are thus perfect predictors and cannot be included in the probit. HMR omit such variables from the study. However, an alternative option might be to generate predicted probabilities equal to 0.9999999, as done in step 2 for a number of countries, and continue from there.

5. Generate predicted values for trade $\hat{m}$ including a predicted value $\omega = \log(e^{\hat{\delta}\hat{\widetilde{z}}^*} - 1)$ or alternatively using the regression coefficients for the polynomial in $\hat{\widetilde{z}}^*$.

Simulations   Generate an alternative measure of the variable of interest. In the case of distance, the new measure is 10% lower than the original.

1. Generate a new latent variable using estimated coefficients together with the new distance measure.

2. Generate the new predicted probabilities $\hat{\rho}'$, setting values above 0.9999999 equal to 0.9999999 as before.

3. Keep the originally estimated Inverse Mills Ratio $\hat{\widehat{\eta}}^*_{ij}$. The original estimate of unobserved trade frictions, conditional on the same countries trading, is still the prediction based on the original values.

4. Generate new predictions $\hat{\widehat{z}}^{*\prime}$, using the same value for $\hat{\widehat{\eta}}^*_{ij}$ but the new predicted probabilities in $\Phi^{-1}(\hat{\rho})$. Use the estimate $\hat{\delta}$ and $\hat{\widehat{z}}^{*\prime}$ to generate an alternative prediction of the non-linear term $\omega' = \log(e^{\hat{\delta}\hat{\widehat{z}}^{*\prime}} - 1)$ or alternatively using the polynomial in $\hat{\widehat{z}}^{*\prime}$.

5. Generate the alternative predicted values for trade $\hat{m}'$ based on the new distance values and ω'.

6. Calculate the gross elasticity $\xi = \frac{\hat{m}' - \hat{m}}{|\log 0.9|}$. For country-pairs with $\hat{\rho}' > 0.9999999$, assign the elasticity generated by a value of $\hat{\rho}' = 0.9999999$. By doing this, HMR are taking a probability of 0.9999999 to imply a certain proportion strictly less than unity of firms exporting.

7. Calculate the net elasticities, which account for MR, applying the appropriate country shares and formulae derived in the text.

A.2.3   Regression results

The first column gives the results from standard OLS estimates ignoring firm heterogeneity. Column 2 gives first-stage probit estimates and column 3 gives the second-stage regression, which includes controls for country selection (the Inverse Mills Ratio $\hat{\widehat{\eta}}^*_{ij}$) and the proportion of firms exporting (represented by the non-linear parameter δ). In all cases, multilateral resistance is controlled for in estimation by the use of exporter and importer fixed effects. For completeness, column 4 includes results based on the semi-paramteric polynomial approximation for the proportion of firms exporting.

Regression Results[3]	(1) OLS	(2) Probit	(3) NL	(4) Polynomial
Distance $(-\gamma)$	-1.176***	-0.660***	-0.799***	-0.862***
	(0.0269)	(0.0206)	(0.0352)	(0.038)
Border	0.458***	-0.382***	0.832***	0.786***
	(0.118)	(0.0941)	(0.124)	0.149
Island	-0.391***	-0.345***	-0.17	-0.2
	(0.107)	(0.0751)	(0.097)	0.12
Landlocked	-0.561***	-0.181	-0.447***	-0.482—
	(0.164)	(0.105)	(0.128)	0.135
Legal	0.486***	0.0964**	0.388***	0.385***
	(0.043)	(0.0297)	(0.039)	0.044
Language	0.176**	0.284***	0.023	0.045
	(0.0556)	(0.0381)	(0.059)	0.066
Colonial ties	1.299***	0.325	1.003***	1.038***
	(0.132)	(0.292)	(0.091)	0.092
Currency Union	1.364***	0.492***	1.024***	1.106***
	(0.216)	(0.132)	(0.269)	0.249
FTA	0.759***	1.985***	0.378**	0.457***
	(0.162)	(0.314)	(0.144)	0.108
Religion	0.102	0.261***		
	(0.09)	(0.061)		
$\widehat{\overline{\eta}}^{*}$			0.392***	1.131***
			(0.049)	0.137
$\widehat{\delta}$			0.716***	
			(0.049)	
$\widehat{z}^{*}$				3.602***
				0.348
$\widehat{z}^{*2}$				-0.782
				0.112
$\widehat{z}^{*3}$				0.064***
				0.011
Constant	3.168***	3.813***	517.2	4.197***
	(0.788)	(0.371)	(764.8)	0.517
Observations $[R^2]$	11,146 [0.709]	24,649	11,146 [0.718]	11,146 [0.721]

---

[3]Significance at 0.001 & 0.01 levels denoted with '***' and '**' respectively. Standard errors in parentheses (based on 50 bootstrap replications in columns 3 and 4).

References

Baranga, T. (2009), 'Unreported trade flows and gravity equation estimation', *mimeo.* .

Carayol, T. (2006), '"expression too long" error with nl command', *Statalist* .

Helpman, E., Melitz, M. & Rubinstein, Y. (2007), 'Estimating trade flows: trading partners and trading volumes', *mimeo.* .

Helpman, E., Melitz, M. & Rubinstein, Y. (2008), 'Estimating trade flows: trading partners and trading volumes', *The Quarterly Journal of Economics* **123**(2), 441–487.