



DEPARTMENT OF ECONOMICS
DISCUSSION PAPER SERIES

**TRADE FLOWS, MULTILATERAL RESISTANCE AND FIRM
HETEROGENEITY**

Alberto Behar and Benjamin D. Nelson

Number 440
July 2009

Manor Road Building, Oxford OX1 3UQ

TRADE FLOWS, MULTILATERAL RESISTANCE AND FIRM HETEROGENEITY¹

Alberto Behar² Benjamin D. Nelson³

Version: July 13, 2009

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: the extensive margin is also affected by multilateral resistance. If all countries reduce their trade frictions, the impact of multilateral resistance is so strong that bilateral trade falls in many cases. This is despite the larger trade elasticities implied by firm heterogeneity. For isolated bilateral changes in trade frictions, multilateral resistance effects are small for most countries, but are large when big importers are involved.

JEL Classifications: F10, F12, F14, F17

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

1. INTRODUCTION

This paper presents a gravity model that unites two strands of the recent literature in international trade: that stressing the importance of multilateral resistance (MR) and that stressing the importance of firm heterogeneity.⁴ We show that the comparative statics obtained with existing gravity models which account for firm heterogeneity but which ignore MR to be significantly misleading. Modelling firm heterogeneity leads to larger country-level trade elasticities, but MR can significantly dampen these responses. For example, even in the presence of firm heterogeneity, multilateral resistance makes most bilateral exports fall in response to lower global trade barriers and reduces the world-wide trade response by two thirds.

We use fixed effects to control for MR in estimation, together with Helpman, Melitz and Rubinstein's (2008) method to control for the implications of firm heterogeneity. But as argued by Anderson and van Wincoop (2003), comparative statics must account for the potential 'third country' effects of changes in trade costs which influence trade flows in a multilateral world. In this setting, it is bilateral trade

¹We are extremely grateful to James Anderson, Elhanan Helpman, Peter Neary, Adrian Wood and participants in seminars at Oxford University and Nottingham University for helpful comments and suggestions. All errors remain, however, our own. Funding from the Economic and Social Research Council (ESRC) and World Bank are gratefully acknowledged.

²CSAE and Department of Economics, University of Oxford. *Email:* ALBERTO.BEHAR@ECONOMICS.OX.AC.UK

³Nuffield College and Department of Economics, University of Oxford. *Email:* BEN.NELSON@NUFFIELD.OX.AC.UK

⁴The gravity equation models bilateral exports from one country to another as proportional to country-pair GDP, and as inversely proportional to trade frictions. Frictions typically include geographical features like distance or policy variables like free trade agreements. Eaton and Kortum (2002) also consider technology in a Ricardian trade model. Gravity models continue to be employed to estimate the elasticity of trade flows with respect to changes in trade frictions, the approach having been subject to a number of theoretical and empirical refinements. See, *inter alia*, Anderson (1979), Bergstrand (1985), Eaton & Kortum (2002) and Chaney (2008) in addition to the papers cited in the text.

barriers relative to multilateral trade barriers, or multilateral resistance, that determine bilateral trade flows.

In the presence of firm heterogeneity, bilateral exports depend on both exports per firm and the proportion of domestic firms that export. The latter is the extensive margin, which we show also takes a gravity-like form. That is, the proportion of firms that actively exports is a function of bilateral trade frictions relative to multilateral trade frictions, together with the product of the trading pair's GDPs.

To compute comparative statics, we develop Taylor approximated price indices which capture the effects of MR on trade responses, akin to Baier and Bergstrand (2009). We show that Baier and Bergstrand's method can be extended to include the effects of firm heterogeneity on multilateral resistance, while retaining the advantage of tractability.

We draw heavily on Anderson and Van Wincoop (2003), Baier and Bergstrand (2009), and Helpman, Melitz and Rubinstein (2008).⁵ The first and second of these papers do not account for firm heterogeneity in estimation, but do address the effects of MR on comparative statics, while the third paper accounts for firm heterogeneity in estimation, but does not address MR effects in comparative statics. We do both.

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, while empirically 'successful', capture the impact of only bilateral trade costs on trade flows, ignoring the fact that countries operate in a multilateral world. As a result, traditional estimates suffer from omitted variable bias since they fail to control for theoretically motivated price index terms, which aggregate both domestic and international trade costs, and therefore capture *multilateral resistance*. AvW show that bilateral trade flows depend on bilateral trade costs *relative to* multilateral resistance. Failing to account for MR effects typically leads one to overstate the importance of changes in trade barriers on bilateral trade flows.

External trade barriers have a large impact on the MR terms of small countries, which typically trade a large proportion of their output internationally. For example, a given increase in an exporter's external trade costs increases its MR by more the smaller is the country in question. This dampens the negative effect of an external border on bilateral trade flows by more: the increase in bilateral trade costs *relative to MR* resulting from the border is smaller. For comparative statics to be valid, modelling these general equilibrium effects due to MR is essential.

The second paper on which we draw is Helpman, Melitz and Rubinstein (2008) ('HMR'). Heterogeneous firm productivity within a country means not all firms engage in exporting in the presence of fixed costs of trade.⁶ The reason for this is that not all firms will have a productivity level high enough to generate profits sufficient to cover the fixed costs of exporting. It follows that, if fixed costs are high

⁵We also draw on the working paper version of the last paper, Helpman, Melitz and Rubinstein (2007).

⁶See Bernard, Jensen, Redding and Schott (2007) for an overview of firms in international trade.

enough, no firms in a given country may find it profitable to export. Hence in the presence of fixed costs of trade, ‘zeros’ naturally arise in the trade data — a *country* selection effect.⁷ 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.

Using bilateral country-level trade data, their remedy is to estimate the predicted probability that at least one firm will export from one country to another (i.e. that country-level exports are positive). They then use this estimate to construct *two* controls: one for the country selection effect (the selection of country pairs into trade) and another for the firm selection effect (the proportion of firms in a country that export, or the extensive margin). Together with fixed effects, which control for MR in estimation, including these two controls allows HMR to obtain consistent estimates of the coefficients in their gravity equation.

We replicate the estimation exercise in HMR and, following them, begin by performing partial equilibrium simulations which do not account for MR. For all our observations, traditional linear estimates bias downwards the effect of observable trade barriers on country-level trade flows. This difference, rather than the firm-level bias in the opposite direction highlighted by HMR, is arguably more relevant for policy. Larger countries have smaller elasticities at the extensive margin, and hence have lower overall ‘country-level’ responses.

Next, we analyse the general equilibrium effects of MR on trade responses using approximated price index terms. We show that Baier and Bergstrand’s (2009) method for obtaining Taylor approximated price indices can be extended to account for the impact of firm heterogeneity on MR. We conduct simulations in which (i) all countries reduce their trade frictions (multilateral changes) and (ii) only two countries do so (bilateral changes). Consistent with AvW’s ‘Implication 1’, after accounting for MR, larger countries have larger net elasticities of bilateral trade in response to multilateral changes in trade costs. This is because larger countries are less affected by MR. We show that once we account for firm entry into trade however, this is no longer necessarily true: bigger countries have lower elasticities at the extensive margin, which oppose the effects of MR on net trade elasticities.

Consistent with AvW, we find the effects of ignoring MR to be dramatic for multilateral changes in trade costs. After accounting for MR, bilateral trade responses are much lower. Most elasticities are negative, which means that the general equilibrium effects are so strong that many country pairs

⁷High 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.

reduce their bilateral trade after a multilateral reduction in trade frictions. For world trade as a whole, allowing for firm entry into exports generates a substantially larger trade elasticity. However, multilateral resistance reduces this response by two thirds. Compared to simple estimates therefore, the simulated impact net of MR effects is less than half as large.

We also consider bilateral changes in trade costs. For two countries liberalising trade in isolation, it is intuitive that the general equilibrium effects will be much smaller. For most country pairs, it is indeed safe to ignore MR in comparative statics, but we identify two types of country-pair for which this is not the case. First, for big country pairs, within the G7 perhaps, MR materially reduces the bilateral trade response. Second, for a small and remote exporter with few export destinations, a big importer can account for a “very” big proportion of export demand. Since price indices are a function of actively traded goods when firm and country selection are at work, MR effects are larger in each of these cases and so must be accounted for here as well.

The paper proceeds as follows. In the next section we present the model, before outlining our approach to comparative statics in section 3. In section 4, we describe and implement the estimation strategy of HMR. In section 5 we simulate comparative statics based on our estimates and using our MR approximation method. Section 6 concludes.

2. THEORY

2.1. The Model

There are J countries, $j = 1, \dots, J$. Within each country are monopolistically competitive firms which produce a continuum of differentiated products. Consumers have a ‘taste for variety’, embodied in CES preferences given by

$$u_j = \left[\int_{l \in B_j} x_j(l)^\alpha dl \right]^{\frac{1}{\alpha}}, \quad (1)$$

where $x(l)$ is consumption of variety l , contained in the set of varieties available in j , B_j . Let $\sigma \equiv 1/(1-\alpha)$ be the elasticity of substitution. With exogenous income in j of Y_j , firms face demand of

$$x_j(l) = \frac{Y_j}{P_j^{1-\sigma}} p_j(l)^{-\sigma}, \quad (2)$$

where $p_j(l)$ is the price of variety l in j and P_j is j ’s ideal price index, given by $P_j = \left[\int_{l \in B_j} p_j(l)^{1-\sigma} dl \right]^{\frac{1}{1-\sigma}}$. Note that it is defined over the set of available goods B_j , which constitutes the set of goods imported by j from active exporters to j .

Each country produces N_j varieties exogenously given, with one variety per firm. The unit cost of production by a firm with unit input requirement a is $c_j a$. a is firm specific as in Melitz (2003), while c_j reflects the exogenously given cost of factors of production in country j . Firms draw a independently from the identical distribution function $G(a)$ with support $[a_L, a_H]$, such that a_L is the lower bound on

possible unit input requirement draws, while a_H is the upper bound. We can identify each firm's unique variety l with its cost level a .

There are two types of cost of exporting. The first is an 'iceberg' variable trade cost $t_{ij} > 1$. The second is a fixed cost of exporting $f_{ij} > 0$, $f_{ii} = 0$. Taken together, a firm in j exporting to i producing q_{ij} units of output has a cost function given by

$$C_{ij}(a) = at_{ij}c_jq_{ij} + c_jf_{ij}. \quad (3)$$

Given demand and costs, each firm chooses price so as to maximise its profits. This gives the price and profit function for a firm exporting from j to i as

$$p_{ij}(a) = \frac{t_{ij}c_ja}{\alpha}, \quad (4)$$

$$\pi_{ij}(a) = (1 - \alpha) \left[\frac{t_{ij}c_ja}{\alpha P_i} \right]^{1-\sigma} Y_i - c_jf_{ij}. \quad (5)$$

Sales by firms in country j are only profitable in country i if $\pi_{ij}(a) > 0$. Hence we define a productivity cut-off a_{ij} by $\pi_{ij}(a_{ij}) = 0$, which is the cost level (or inverse productivity level) below which it is profitable to export. Firms with $a > a_{ij}$ do not generate profits high enough to cover the fixed costs of exporting f_{ij} . Using an exporting firm's profit function above then gives us the cut-off as

$$a_{ij} = \left[\frac{Y_i(1 - \alpha)}{f_{ij}c_j} \right]^{\frac{1}{\sigma-1}} \frac{\alpha P_i}{c_j t_{ij}}. \quad (6)$$

This gives us the *extensive margin* of trade. When a_{ij} is higher, the extensive margin is greater, implying a larger subset of firms exports. It rises as the income of the importing country rises, and as both fixed and variable costs of trade fall. Whenever $a_{ij} < a_H$, there will be firm selection into exporting. In particular, firms with the highest variable costs will choose not to export.

The total value of imports by country i from country j is given by $M_{ij} = \int_{a_L}^{a_{ij}} p_j q_i N_j dG(a)$. Substituting in for prices and quantities, we obtain

$$M_{ij} = \left[\frac{t_{ij}}{\alpha P_i} \right]^{1-\sigma} N_j Y_i \int_{a_L}^{a_{ij}} a^{1-\sigma} dG(a). \quad (7)$$

We then define $V_{ij} \equiv \int_{a_L}^{a_{ij}} a^{1-\sigma} dG(a)$ as a term capturing the firm selection effect. Note that as a_{ij} rises, indicating that the cost level above which firms find it unprofitable to export rises, V_{ij} rises. In other words, as this export cut-off rises, a larger set of firms export. Using this, we have bilateral exports from j to i given by

$$M_{ij} = \left[\frac{c_j t_{ij}}{\alpha P_i} \right]^{1-\sigma} N_j Y_i V_{ij}. \quad (8)$$

V_{ij} in equation (8), which is the same as in HMR, forms the basis for accounting for firm heterogeneity.

Omitting this term leads to erroneous estimates of the impact of trade barriers on firm level trade.

Unlike HMR, we wish to model P_i , explicitly incorporating firm heterogeneity. In particular, trade costs can be high enough to prohibit exports to some locations. This arises naturally in a model with fixed costs of exporting. In this setting, it becomes useful to define two types of set of countries. First, we define J_i as the set of all exporters to i ; if country j exports to i , we say that it is contained in the set J_i . Similarly, we define I_j as the set of all importers from j . So if i imports from j , it is contained in I_j . We model asymmetries in trade flows by allowing for $J_i \neq I_j$. This says that the set of exporters to country i is not necessarily equal to the set of importers from country j . In particular, we allow for $J_k \neq I_k$, which says that the set of exporters to country k is not necessarily equal to the set of importers from country k ; k might import from j , but not export back. (Note also that country i is contained in the set of all exporters to i , J_i , and that j is contained in the set of all importers from j , I_j . The sets therefore capture both domestic and international trade.)

An important implication of this is that the price indices are defined over these sets. More explicitly, country i 's price index is given by $P_i^{1-\sigma} = \sum_{j \in J_i} \int_{a_L}^{a_{ij}} p_{ij}(a)^{1-\sigma} dG(a)$. Using $p_{ij}(a) = c_j a t_{ij} / \alpha$ the price index can be written

$$P_i^{1-\sigma} = \sum_{j \in J_i} (c_j t_{ij} / \alpha)^{1-\sigma} N_j V_{ij}, \quad (9)$$

such that it is defined over the set of exporters to i . Since $P_i^{1-\sigma}$ aggregates trade costs, AvW interpret the price indices in the gravity equation as multilateral resistance terms, which we discuss further below. It is important to note however that firm heterogeneity and hence possible country selection into trade imply that multilateral resistance terms must be defined over the set of active traders.

2.2. General Equilibrium

The model has been described in partial equilibrium. To get closer to AvW's system, we assume trade balance in order to achieve general equilibrium closure. We will show that this allows us to write an AvW style gravity equation for bilateral exports. In addition, trade balance allows us to derive a gravity equation for the extensive margin. Both of these equations make explicit the role of price indices, or multilateral resistance, in general equilibrium.

Assume trade balance for each country, such that $Y_j = \sum_{i \in I_j} M_{ij}$. We show that using this in (8) allows one to write (see Appendix A.1)

$$M_{ij} = \frac{Y_i Y_j}{Y^{I_j}} \left(\frac{t_{ij}}{P_i \hat{P}_j} \right)^{1-\sigma} V_{ij}, \quad (10)$$

where P_i is country i 's *importing* multilateral resistance term, \hat{P}_j is country j 's *exporting* multilateral resistance term, and $Y^{I_j} \equiv \sum_{h \in I_j} Y_h$ is the total output of the set of importers from j , I_j . In arriving

at this equation, using trade balance allows one to write these price indices as

$$P_i^{1-\sigma} = \sum_{j \in J_i} \left(\frac{t_{ij}}{\hat{P}_j} \right)^{1-\sigma} s_j^{J_i} V_{ij} R_{ij}, \quad (11)$$

$$\hat{P}_j^{1-\sigma} = \sum_{i \in I_j} \left(\frac{t_{ij}}{P_i} \right)^{1-\sigma} s_i^{I_j} V_{ij}. \quad (12)$$

We use the ‘hat’ to denote exporting multilateral resistance, in order to distinguish it from importing multilateral resistance.⁸ In these price index equations, the s_k^X terms represent country k ’s GDP as a share of the total output of all the countries in set X . That is, $s_i^{I_j} \equiv Y_i/Y^{I_j}$ is importer i ’s GDP as a share of the total output of all importers from j . Similarly we define $s_j^{J_i} \equiv Y_j/Y^{J_i}$, where $Y^{J_i} \equiv \sum_{k \in J_i} Y_k$, as exporter j ’s GDP as a share of the total output of all exporters to i .

In arriving at the above gravity equation, we have defined Y^{I_j} as the total output of all importers from country j . It is intuitive that as this quantity rises, it becomes relatively less attractive for firms in j to export to i , so that bilateral trade falls. $R_{ij} \equiv Y^{J_i}/Y^{I_j}$ is the output of exporters to i relative to that of importers from j . When it is high, relatively many exporters trade with i , indicating an increase in product variety, and hence a reduction in importer i ’s price index.

The inclusion of two price terms makes system (10)-(12) resemble that of AvW, with the crucial difference that it allows for firm heterogeneity and country selection.⁹ A corollary of this is that the multilateral resistance terms in AvW’s equation become asymmetric in our equation, the effects of which must be accounted for when performing comparative statics. Reductions in trade costs affect both the numerator and the denominator of the gravity equation. Because a reduction in t_{ij} affects the multilateral resistance terms, the resulting increase in bilateral trade will be smaller than in the absence of changes in multilateral resistance, all else being equal.

A further implication of imposing trade balance is that equation (6) for the extensive margin also takes a gravity-like form. In particular we show that (see Appendix A.2)

$$a_{ij}^{\sigma-1} = \frac{(1-\alpha)}{c_j N_j} \frac{Y_i Y_j}{Y^{I_j}} \frac{1}{f_{ij}} \left(\frac{t_{ij}}{P_i \hat{P}_j} \right)^{1-\sigma}, \quad (13)$$

which is a gravity equation for the cost cut-off defining the extent of the extensive margin. Just as for bilateral exports, it responds positively to the product of the trading countries’ GDPs, negatively to bilateral trade costs, and positively to multilateral resistance, captured by the $P_i \hat{P}_j$ term. Note also that the fixed trade cost f_{ij} enters equation (13), such that higher fixed costs reduce the cost level below which exporting is profitable. In this way, fixed costs affect the number of exporting firms, but not how much each exports. In other words, fixed costs affect the volume of bilateral exports, but only indirectly

⁸ Anderson (2009) refers to these as indices of sellers’ and buyers’ trade cost incidence respectively. This distinction is necessary given asymmetries in trade flows.

⁹ It also preserves the potential for asymmetries trade flows, which is precluded by the ‘decomposability assumption’ made by HMR in deriving a similar equation contained in their Appendix II. For further discussion, see footnote 12.

through their impact on a_{ij} , which determines V_{ij} . This is important for the identification strategy in the empirical section.

Further, (13) makes explicit the role of multilateral resistance on the extensive margin. Just as for bilateral trade flows, a multilateral increase in trade costs increases both the numerator and the denominator of (13); the effect of trade costs on the price indices in the denominator therefore acts to mitigate the direct effect in the numerator. Just as AvW show for bilateral trade flows, comparative statics on the extensive margin will be misleading where the latter effect is accounted for, but the former is not.

2.3. Towards Implementation

Imposing an assumption about the distribution of productivities grants us further analytical tractability. Following much of the recent trade literature, we impose a Pareto distribution on firm specific variable costs¹⁰, such that

$$\frac{1}{a} \sim \text{Pareto}(k) \quad a \in [a_L, a_H],$$

where a_L and a_H define the support of the distribution, consistent with above, and k is the shape parameter. In particular, this implies $G(a) = \frac{a^k - a_L^k}{a_H^k - a_L^k}$, and $g(a) = k \frac{a^k}{a_H^k - a_L^k} \frac{1}{a}$. Using this, we write the extensive margin V_{ij} as

$$V_{ij} = \max \left\{ \bar{k} \left[\left(\frac{a_{ij}}{a_L} \right)^{k-\sigma+1} - 1 \right], 0 \right\}, \quad (14)$$

such that whenever $a_{ij} < a_L$, or no firms in j generate profits sufficient to cover the fixed costs of exporting, $V_{ij} = 0$. From (10), this generates zero bilateral exports from j to i when $a_{ij} < a_L$. When will this scenario arise? Following HMR, one way to operationalise (14) is to consider the profits of the firm in j with the lowest variable costs a_L . If this firm does not find it profitable to export to i , then no firm in j will. Accordingly, for firm a_L , the ratio of variable profits to fixed export costs can be written as

$$Z_{ij} \equiv \frac{Y_i(1-\alpha) \left(\frac{c_j}{\alpha P_i} \right)^{1-\sigma} (t_{ij} a_L)^{1-\sigma}}{c_j f_{ij}}. \quad (15)$$

It then follows that

$$V_{ij} > 0 \quad \text{iff} \quad Z_{ij} > 1. \quad (16)$$

Z_{ij} is HMR's latent variable. It is unobserved, but can be estimated using a combination of the distributional assumption on variable costs a , and observable variables such as GDP and trade costs. We wish to highlight the role of multilateral resistance in this variable too. A further consequence of the trade

¹⁰See, *inter alia*, Chaney (2008) and Helpman, Melitz and Yeaple (2004), who use Pareto distributed firm productivity in theoretical and empirical work.

balance assumption is that HMR's Z_{ij} can be written as (see Appendix A.3)

$$Z_{ij} = \tilde{Z}_{ij} \left(P_i \hat{P}_j \right)^{\sigma-1}, \quad (17)$$

where

$$\tilde{Z}_{ij} \equiv \frac{Y_i Y_j}{Y^{I_j} N_j} \frac{(1-\alpha)(t_{ij} a_L)^{1-\sigma}}{c_j f_{ij}}. \quad (18)$$

That is, the latent variable can be decomposed into two components: multilateral resistance, and a component \tilde{Z}_{ij} independent of prices. Next, using (13), (17) and (18) allows us to relate the extensive margin cut off and the latent variable according to

$$Z_{ij} = \left(\frac{a_{ij}}{a_L} \right)^{\sigma-1}, \quad (19)$$

which gives (14) according to

$$V_{ij} = \max \left\{ \bar{k} \left[\tilde{Z}_{ij}^\delta \left(P_i \hat{P}_j \right)^{\delta(\sigma-1)} - 1 \right], 0 \right\}, \quad (20)$$

where $\delta \equiv \frac{k-\sigma+1}{\sigma-1}$. This again makes explicit the point that MR affects the extensive margin too; increases in trade costs will decrease \tilde{Z}_{ij} , but will increase $P_i \hat{P}_j$, mitigating the net effect on the latent variable Z_{ij} and hence on the extensive margin V_{ij} .

Taking logs of (10) yields the equation we work with for estimation and comparative statics. Specifying trade costs $t_{ij}^{\sigma-1} = D_{ij}^\gamma$, where D_{ij} is bilateral distance such that $d_{ij} = \ln D_{ij}$, gives

$$m_{ij} = \psi + \psi_i + \psi_j - \gamma d_{ij} + w_{ij} + \ln \left(P_i \hat{P}_j \right)^{\sigma-1}, \quad (21)$$

in which ψ is a constant, ψ_i and ψ_j are i and j specific constants respectively, and where w_{ij} is the ij specific component of $\ln V_{ij}$, given by $w_{ij} \equiv \ln \left\{ e^{\delta [\tilde{z}_{ij} + \ln(P_i \hat{P}_j)^{\sigma-1}]} - 1 \right\}$, which is the term containing the extensive margin. The last of these contains $z_{ij} = \ln Z_{ij}$, which since $Z_{ij} = \tilde{Z}_{ij} \left(P_i \hat{P}_j \right)^{\sigma-1}$ can be written

$$z_{ij} = \chi + \chi_i + \chi_j - \gamma d_{ij} - \ln f_{ij}, \quad (22)$$

in which χ is a constant, while χ_i and χ_j are i and j specific constants (including MR) respectively. The explicit consideration of MR will have no implications for estimation using HMR's method, but we wish to include the general equilibrium multilateral resistance effects explicit in (10) and (20) in our simulated comparative statics. It is to this issue that we now turn.

3. COMPARATIVE STATICS: THEORY

3.1. The bilateral trade elasticity

Consider a change in distance ∂d which includes (but is not limited to) a change in the bilateral distance between i and j , ∂d_{ij} . In the general case, the bilateral trade elasticity net of MR effects between importer i and exporter j is defined as $\xi_{ij} \equiv -\frac{\partial m_{ij}}{\partial d}$. From (21) it is given by

$$\xi_{ij} = \underbrace{\gamma}_{\text{Intensive margin}} + \underbrace{\varphi_{ij}\gamma}_{\text{Extensive margin}} - \underbrace{(1 + \varphi_{ij}) \frac{\partial \ln(P_i P_j)^{\sigma-1}}{\partial d}}_{\text{Multilateral Resistance}}, \quad (23)$$

where $\varphi_{ij} \equiv \frac{\partial w_{ij}}{\partial z_{ij}} = \frac{\delta e^{\delta z_{ij}}}{e^{\delta z_{ij}} - 1} > 0$. Equation (23) highlights the role of three effects which combine to form the net elasticity. The first is the firm-level effect, γ , or the intensive margin. This captures the impact of a change in trade costs on the level of exports per firm. The second is the extensive margin $\varphi_{ij}\gamma$. φ_{ij} provides a mapping from export profits of the most productive firm to the export profits of all firms in j given Pareto distributed productivities. γ gives the impact of a change in trade costs on the export profits of the most productive firm in j exporting to i . The product of φ_{ij} and γ then gives the impact of a change in trade costs on the change in export profits in country j that results, which in turn determines the impact on the number of firms in j exporting to i . We refer to the combination of the intensive and the extensive margin effects as the *gross elasticity*, given by

$$\xi_{ij}^{gross} = \gamma(1 + \varphi_{ij}). \quad (24)$$

The third term in (23) is the effect due to MR. Since increases in trade costs typically increase MR, the third term captures the dampening effect MR has on gross bilateral trade elasticities. Omission of this third term will therefore typically overstate the true net elasticity. In our model, which combines the effects of MR and firm heterogeneity, the MR elasticity $\frac{\partial \ln(P_i P_j)^{\sigma-1}}{\partial d}$ is multiplied by $(1 + \varphi_{ij})$ in determining the net bilateral trade elasticity. The term $(1 + \varphi_{ij})$ reflects the fact that MR affects both the intensive margin, as in AvW, and the extensive margin in the presence of firm heterogeneity.

One way to include MR effects is to construct the system of nonlinear price index equations and estimate the system following AvW. This method is computationally demanding however, especially when considering more than 150 countries, accommodating asymmetries in trade frictions, and allowing for the interdependence of the price indices and the extensive margin terms. As noted by AvW, asymmetric frictions make identification difficult, while Bergstrand, Egger and Larch (2007) have shown that, because of asymmetries, the AvW system solution can yield complex numbers. Furthermore, our application must accommodate country selection and use GDP shares that are specific to sets of active traders.

We therefore modify Baier and Bergstrand's (2009) (BB) method to account for the extensive margin. The method uses first order Taylor expansions to approximate the multilateral resistance terms. This

yields expressions for MR which contain exogenous variables which we later use to compute comparative statics. This approach has a number of advantages. First, as noted by BB, the parameters of the multilateral resistance terms are observable once approximated. Under an additional assumption we show that this extends to the case in which extensive margin effects due to firm heterogeneity are also considered. Second, we can use fixed effects for estimation, but use our Taylor approximations for comparative statics. Third, the Taylor expanded multilateral resistance terms offer a good intuition for the effects obtained from the comparative static exercises we subsequently consider. Against these advantages, of course, we should remember that what we obtain are linear approximations of non-linear price index equations. BB investigate the possible approximation errors associated with their method, and find them to be small for the majority of cases.¹¹

3.2. Approximation method

The aim of our approximation method is to yield an expression for the price index MR terms that contain only exogenous variables. In approximating the terms P_i and \hat{P}_j , we use a first order Taylor expansion around a world of symmetric trade frictions. The price indices contain extensive margin terms, which themselves contain the system of price indices. To disentangle these effects, we make the following assumption:

Decomposability Assumption: *The extensive margin terms V_{ij} entering the price indices are approximately $V_{ij} \simeq \tilde{Z}_{ij}^\delta \left(P_i \hat{P}_j\right)^{\delta(\sigma-1)}$.*¹²

Using this in the system of price index equations yields the following system for us to Taylor expand:

$$\hat{P}_j^{-k} = \sum_{i \in I_j} t_{ij}^{1-\sigma} s_i^{I_j} P_i^k \tilde{Z}_{ij}^\delta, \quad (25)$$

$$P_i^{-k} = \sum_{j \in J_i} t_{ij}^{1-\sigma} s_j^{J_i} \hat{P}_j^k \tilde{Z}_{ij}^\delta R_{ij}. \quad (26)$$

We show that this yields the following approximated MR term (see Appendix A.4.2)

$$\ln \left(P_i \hat{P}_j \right)^{\sigma-1} = \frac{1}{\delta+1} \times \left\{ \underbrace{- \sum_{l \in I_j} s_l^{I_j} \sum_{h \in J_l} s_h^{J_l} [\gamma d_{lh} - \delta \tilde{z}_{lh}]}_{\text{World Trade Resistance}} + \underbrace{\sum_{h \in J_i} s_h^{J_i} [\gamma d_{ih} - \delta \tilde{z}_{ih}]}_{\text{Importer's MR}} + \underbrace{\sum_{l \in I_j} s_l^{I_j} [\gamma d_{lj} - \delta \tilde{z}_{lj}]}_{\text{Exporter's MR}} \right\}, \quad (27)$$

¹¹They perform a Monte Carlo analysis in which only 8% of the approximation method comparative statics differ from AvW's results by more than 20%. They find that of these errors, all are accounted for by economically small countries located close to each other and to large trading partners.

¹²HMR make a similar assumption in their Appendix II. The crucial difference is that, in doing so, they assume that the ij component of V_{ij} is symmetric. Their assumption is that $V_{ij} = \left(\phi_{ij} \phi_i \hat{\phi}_j \right)^{1-\sigma}$, in which $\phi_{ij} = \phi_{ji}$, while ϕ_i and $\hat{\phi}_j$ 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 $\phi_{ij} \neq \phi_{ji}$. In our case, this implies it is possible that $\tilde{Z}_{ij} \neq \tilde{Z}_{ji}$.

which we can use for computing comparative statics. (27) extends Baier and Bergstrand's (2009) approximated MR term to the case of firm heterogeneity. The first modification to BB's expression is to aggregate trade costs over the relevant sets of traders, I_j and J_i . This reflects the asymmetry in trading relationships that arises under asymmetric trade costs and firm heterogeneity. The second modification is to include not only an intensive margin effect γd_{ij} in each component of the MR term, but also an extensive margin component $\delta \tilde{z}_{ij}$. Our approximated MR term shares with BB the advantage of yielding analytical tractability and a clear intuition for the comparative statics effects we subsequently compute.

In (27), MR is conveniently decomposed into three terms. The first of these captures world trade resistance, which averages the importing MR of all importers from j . When this world resistance term is higher, world trade in general is subject to higher trade frictions, reducing bilateral trade all else being equal. The second two terms in (27) are i 's importing MR and j 's exporting MR respectively. When either of these two terms is high, trading with other countries in the world trade system is subject to high trade costs, encouraging i and j to trade with each other instead. This clearly captures the idea that it is relative trade costs that matter in determining bilateral trade flows. For example, when $\sum_{h \in J_i} s_h^{J_i} [\gamma d_{ih} - \delta \tilde{z}_{ih}]$ is large, all exporters to i incur high trade costs in trading with i . Country i therefore incurs relatively small trade costs in importing from j , raising exports from j to i .

Before considering the net impact of changes in trade costs given in (23), we first examine the responses of the extensive margin and multilateral resistance in isolation. We then combine these effects with the intensive margin elasticity and consider the net impact of trade liberalisation.

3.3. Trade Costs and Multilateral Resistance

AvW show that, because small countries will in general trade a larger proportion of their output internationally, small countries' MR responses to multilateral changes in trade costs will be larger.¹³ This is also the case in our model. Denoting the elasticity of MR by $\frac{\partial \ln(P_i \hat{P}_j)^{\sigma-1}}{\partial d} \equiv \varepsilon_{ij}$, for Multilateral (M) changes $\partial d_{ij} = \partial d$ for all $i, j, i \neq j$, we show that for a change in the size of importer i , $\partial s_i^{J_i} = \partial s_i^{I_j} = \partial s_i$,

$$\frac{\partial \varepsilon_{ij}^M}{\partial s_i} < 0, \quad (28)$$

or that the elasticity of MR with respect to multilateral changes in trade costs *decreases* as importer i becomes larger relative to the set of active traders. We also show that the reverse is true for Bilateral (B) changes in trade costs in which $\partial d_{ij} \neq 0$ and $\partial d_{lk} = 0$ for all $k, l \neq i, j$. For a change in the size of importer i , $\partial s_i^{J_i} = \partial s_i^{J_j} = \partial s_i^{I_j} = \partial s_i$,

$$\frac{\partial \varepsilon_{ij}^B}{\partial s_i} > 0, \quad (29)$$

¹³This in part explains why calculations of the effects of the US/Canada border on Canadian provinces, which are small compared to US states, were particularly susceptible to the omission of MR in McCallum (1995).

such that the elasticity of MR with respect to bilateral changes in trade costs *increases* as importer i becomes larger relative to the set of active traders. What is the reason for the reversal of sign? When multilateral changes occur, a large fraction of a small country's total trade is affected, such that its price index falls by relatively more. Conversely, large countries' smaller proportion of internationally traded output implies that their MR terms are affected proportionately less. This pattern is reversed however when trade costs change only bilaterally. For a pair of large countries, each 'counts for a lot' in the other's MR term, since MR reflects weighted average trade costs in our approximation. For a pair of small countries, the reverse is true: they 'count for relatively little' in determining each other's MR. Then bilateral changes in trade costs have a bigger effect on the MR terms for larger country pairs. So while larger country pairs experience smaller effects through MR when trade costs change *multilaterally*, they experience larger effects through MR when trade costs change *bilaterally*. We summarise this in the following Lemma, which we use later:

LEMMA 1. *The elasticity of MR with respect to trade costs is*

- (a) *decreasing in importer size for multilateral changes in trade costs* $\frac{\partial \varepsilon_{ij}^M}{\partial s_i} < 0$;
- (b) *increasing in importer size for bilateral changes in trade costs* $\frac{\partial \varepsilon_{ij}^B}{\partial s_i} > 0$.

Proof. See Appendix A.5. ■

3.4. Trade costs and The Extensive Margin

How do we expect the elasticity of the extensive margin to vary by country size? Using $w_{ij} = \ln \left\{ e^{\delta [\tilde{z}_{ij} + \ln(P_i \hat{P}_j)^{\sigma-1}]} - 1 \right\}$, we write

$$\frac{\partial w_{ij}}{\partial d_{ij}} = -\varphi_{ij}\gamma + \varphi_{ij} \frac{\partial \ln(P_i \hat{P}_j)^{\sigma-1}}{\partial d_{ij}}, \quad (30)$$

where the first term $-\varphi_{ij}\gamma \equiv \frac{\partial \tilde{w}_{ij}}{\partial d_{ij}}$ is the gross effect of a fall in trade costs on the extensive margin, which yields the net extensive margin elasticity when the effect of MR through $\frac{\partial \ln(P_i \hat{P}_j)^{\sigma-1}}{\partial d_{ij}}$ is included. Holding constant the MR effect, we show that

LEMMA 2. *The gross extensive margin elasticity is larger in magnitude for smaller countries, such that*

$$\frac{\partial \tilde{w}_{ij}}{\partial d_{ij}} < 0, \quad \frac{\partial^2 \tilde{w}_{ij}}{\partial d_{ij} \partial Y_j} > 0.$$

Proof. See Appendix A.6. ■

Intuitively, since small, distant countries subject to high trade costs initially exhibit a very small range of exporting firms, reductions in trade costs which lead firms to enter international markets give rise to large proportional effects. In the extreme case, since the elasticity is given by $\frac{\partial w_{ij}}{\partial d_{ij}} = \frac{\partial V_{ij}}{\partial d_{ij}} \frac{d_{ij}}{V_{ij}}$ and

V_{ij} might be close to zero for small countries, the proportional response will be very large. This effect diminishes as trade costs fall, or the country gets larger, since as this occurs V_{ij} rises.

We use the results in Lemma 1 and Lemma 2 in what follows next, where we consider the overall effect of changes in trade costs on the elasticities of bilateral trade.

3.5. Bilateral changes in trade costs

It is intuitive that the general equilibrium effects captured by MR will be less material when considering changes in a small subset of countries. To investigate this, we consider the special case of two countries reducing their frictions, but nobody else doing so. For practical purposes, can one effectively ‘ignore’ MR when considering changes for sufficiently small countries? Denote the bilateral elasticity of exports from country j to country i when they reduce distance between each other by $\xi_{ij}^B \equiv -\frac{\partial m_{ij}}{\partial d_{ij}}$, in which $d_{ij} = d_{ji}$ and $\partial d_{ij} \neq 0$. Then

$$\xi_{ij}^B = \gamma (1 + \varphi_{ij}) \left(1 + s_i^{I_j} s_j^{J_i} + s_j^{I_j} s_i^{J_j} - s_j^{J_i} - s_i^{I_j} \right), \quad (31)$$

from which we establish

PROPOSITION 1. *For bilateral changes in trade costs, bilateral trade elasticities are decreasing in country size owing to*

- (i) *a larger multilateral resistance effect, which dampens the direct effect of changes in trade costs by more for larger countries, for a given extensive margin elasticity (Lemma 1 (b));*
- (ii) *smaller elasticities at the extensive margin for larger countries, for given multilateral resistance effects (Lemma 2).*

Proof. By Lemmata 1 (b) and 2 (see Appendix A.7). ■

COROLLARY 1. *For bilateral changes in trade costs, accounting for the effects of trade costs through multilateral resistance increases in importance as importer size increases relative to the set of active traders.*

The gross elasticity is $\xi_{ij}^{gross} = \gamma (1 + \varphi_{ij})$, as in (24), which ignores MR effects and is positive. The ratio of the net elasticity to the gross elasticity $\frac{\xi_{ij}^B}{\xi_{ij}^{gross}}$ gives a sense of the typical impact of MR. This ratio is given by

$$\frac{\xi_{ij}^B}{\xi_{ij}^{gross}} = 1 + s_i^{I_j} s_j^{J_i} + s_j^{I_j} s_i^{J_j} - s_j^{J_i} - s_i^{I_j}. \quad (32)$$

(32) is always positive, and captures in a simple way the dampening effect multilateral resistance has on bilateral trade elasticities. It has a maximum value of 1. The further below 1 is this ratio, the greater the extent to which MR affects the net elasticity. The ratio decreases further below unity as country size increases, reflecting Lemma 1 (b). Given that bigger countries also have lower elasticities at the

extensive margin (Lemma 2), we arrive at the above proposition. Here, the impact of MR goes in the same direction as the impact of firm heterogeneity when we consider bilateral trade cost changes. In particular, as well as having smaller elasticities at the extensive margin, larger countries ‘count for more’ in the MR terms of their respective trade partners. When these big trade partners reduce their trade barriers against each other, their MR terms fall by relatively more; this provides a larger MR dampening effect for larger countries, which reinforces their smaller extensive margin elasticities.

Country shares are typically small, so the value of the ratio $\frac{\xi_{ij}^B}{\xi_{ij}^{gross}}$ will usually be a value close to unity. This indicates that the impact of bilateral changes in trade costs on multilateral resistance will have a small impact on the net effect of bilateral changes in trade costs overall. One might infer therefore that it is only important to account for MR when considering two big countries. However, in the presence of firm heterogeneity, which causes ‘zeros’ and asymmetries in trade flows, it is the size of a country’s GDP relative to other active traders that matters for comparative statics. This is the intuition behind Corollary 1.

For example, consider a trade deal between a small developing country and the US. In this example, (a) $s_i^{I_j}$ is large; (b) $s_j^{J_i}$ is very small; and (c) $s_j^{I_i}$ is very small. (a) states that the importing country accounts for a large share of the combined output of all importers from j . This is the case if j exports to very few countries, of which i is the largest. (b) states that the share of country j in the total output of exporters to i is very small. This is the case, say, for a small developing country exporting to the US, which in turn imports from lots of other large countries. Finally, (c) states that the share of country j is small as a proportion of the total output of importers from i . Under these conditions, letting $s_j^{J_i}, s_j^{I_i} \approx 0$ in (32), the ratio of the net to the gross elasticity is

$$\frac{\xi_{ij}^B}{\xi_{ij}^{gross}} \approx 1 - s_i^{I_j}, \quad (33)$$

which clearly decreases as importer i accounts for more and more of the total output of importers from exporter j . For country pairs such as these, ignoring MR, even for bilateral changes in trade costs, may not be innocuous, the effect becoming bigger as the importer becomes larger.

3.6. Multilateral changes in trade costs

We now analyse the elasticity of bilateral exports from j to i given a multilateral change in trade costs $\xi_{ij}^M \equiv -\frac{\partial m_{ij}}{\partial d}$, in which $\partial d_{ij} = \partial d$ for all $i, j, i \neq j$. Intuitively, MR terms are likely to be crucial in determining correct comparative static effects under multilateral change in trade costs, as each country pair experiences large ‘third country’ changes. We show that the net export elasticity when all countries change their bilateral trade frictions is given by

$$\xi_{ij}^M = \gamma (1 + \varphi_{ij}) \left(- \sum_{l \in I_j} s_l^{I_j} s_j^{J_l} + s_i^{J_i} + s_j^{I_j} \right). \quad (34)$$

It can easily be verified that, were we to allow for falls in internal distance, $\xi_{ij}^M = 0$. This confirms that it is international trade costs *relative* to domestic trade costs that influence international trade flows. We then use (34) to state the following:

PROPOSITION 2. *For multilateral changes in trade costs*

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

$$\xi_{ij}^M \geq 0, \quad (35)$$

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 does not change ($\varphi_{ij} = 0$),*

$$\frac{\partial \xi_{ij}^M \big|_{\varphi_{ij}=0}}{\partial s_i} > 0, \quad (36)$$

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

(b) (ii) *if the extensive margin changes ($\varphi_{ij} > 0$),*

$$\frac{\partial \xi_{ij}^M}{\partial s_i} \geq 0, \quad (37)$$

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

Proof. See Appendix A.8. ■

As before, the gross elasticity is $\xi_{ij}^{gross} \equiv \gamma(1 + \varphi_{ij})$. Then to get a sense of the impact of MR, consider the ratio of the net to the gross elasticity, given by

$$\frac{\xi_{ij}^M}{\xi_{ij}^{gross}} = - \sum_{l \in I_j} s_l^{I_j} s_j^{J_l} + s_i^{J_i} + s_j^{I_j}. \quad (38)$$

In contrast to the bilateral case (32), the ratio in the multilateral case can be positive or negative for the typical country, consistent with Proposition 2, part (a). If positive, it is likely to be far from unity. The extra heterogeneity introduced by the extensive margin can affect the net elasticity in idiosyncratic ways, but it does not drive Proposition 2, part (a). Rather, the ambiguity of the sign in part (a) is driven by considering the impact of MR alone.¹⁴ The theoretical origin of this result is the ‘endowment economy’ nature of the model studied here; changes in trade costs serve to reallocate output from one

¹⁴Cf. AvW, their footnote number 15.

activity (e.g. domestic trade) to another (e.g. international trade). When viewed in this way, it seems perfectly natural for exports to one destination to be redirected towards another destination in response to a trade liberalisation, in order that incomes and expenditures be balanced. This ‘endowment economy effect’ gives rise to the possibility of negative export elasticities described above.

(b) part (i) of Proposition 2 repeats AvW’s ‘Implication 1’ using our Taylor approximation. It states that countries exporting to larger importers have larger elasticities of bilateral trade when multilateral trade liberalisation takes place. The result follows from Lemma 1 (a): larger countries typically trade a larger fraction of their output domestically for a given external tariff. An implication of this is that, for large countries, a smaller proportion of their total (i.e. domestic plus international) trade is affected by tariff changes. This reduces the size of the effect of trade cost changes on their MR terms. Because the dampening effect due to MR is smaller for larger countries, the effect of a multilateral tariff reduction on bilateral trade is bigger. This is stated in Proposition (2) (b) part (i) and can be seen by inspection of (38).

Proposition 2 (b) part (ii) however states that this theoretical relationship no longer holds when the extensive margin responds to multilateral trade liberalisations as well. Mechanically, Lemma 1 (a) and Lemma 2 work against each other in determining the effect of country size on bilateral trade elasticities given multilateral changes in trade costs. On the one hand, smaller countries experience larger direct effects at the extensive margin, as above, tending to increase the magnitude of their export elasticities. But compared to the bilateral case, small countries experience larger MR dampening effects when trade costs change multilaterally, owing to their increased dependence on international trade. These competing effects suggest a potentially non-monotonic relationship between country size and bilateral trade elasticities, and explain the ambiguity at the heart of Proposition 2 parts (a) and (b) (ii).

We turn next to an examination of the empirical implications of these theoretical results, beginning with estimation.

4. ESTIMATION

This section begins with a brief description of the methodology. It follows with an account of our replication of the HMR regression results. As in HMR, we illustrate the trade elasticity effects using the example of changes in distance.

4.1. Method

We follow a two stage procedure in which the first stage generates two controls for inclusion in the second. The first control is for country selection into trading, captured by the inverse Mills ratio. The second control is for firm selection, or the proportion of firms exporting to a particular destination. For the purposes of estimation, all multilateral resistance terms are captured in the constant and by country fixed effects. Accounting for multilateral resistance therefore has no implications for estimation

conditional on these fixed effects being included.

A step-by-step description of the procedure is provided in Appendix A.10. In the first stage, we estimate a probit model for the probability that j exports to i , denoted ρ_{ij} . At least one firm exports if the most productive firm can do so profitably. The most productive firm's profit is captured by the unobserved latent variable z_{ij} , modelled above. We let $z_{ij} = z(\chi_j, \chi_i, t_{ij}, f_{ij}) + (u_{ij} + v_{ij})$ where $z(\cdot)$ is log-linear. The χ 's capture country-specific effects, including multilateral resistance, while t_{ij} is an observed bilateral variable trade cost and f_{ij} is a bilateral observed fixed trade cost. In the stochastic setting these trade costs have unobserved components. In particular u_{ij} is unobserved variable trade costs and v_{ij} is unobserved fixed trade costs. The composite error term $u_{ij} + v_{ij} = \eta_{ij} \sim N[0, \sigma_\eta^2]$ where $\sigma_\eta^2 = \sigma_u^2 + \sigma_v^2$. The most productive firm in j exports to i when $z_{ij} > 0$. We cannot observe z_{ij} but infer that $z_{ij} > 0$ when we observe bilateral exports. Letting T_{ij} be unity when exports from j to i are observed and zero otherwise, we can write

$$\rho_{ij} = \Pr(T_{ij} = 1 | \text{Observables}). \quad (39)$$

Probit estimates can be used to generate predicted values for z_{ij} , \hat{z}_{ij}^* , where the “*” reflects the normalisation of the coefficient estimates by σ_η .¹⁵ It is well known that the predicted probability $\hat{\rho}_{ij}$ can be used to estimate the inverse Mills ratio $\hat{\eta}_{ij}^*$. This is a consistent estimate of the expected value of the unobserved trade frictions, given that trade takes place. Including $\hat{\eta}_{ij}^*$ in the second stage thus controls for the country selection effect. Furthermore, HMR show how \hat{z}_{ij}^* and $\hat{\eta}_{ij}^*$ can be used to account for firm selection. Define $\hat{x}_{ij} \equiv \hat{z}_{ij}^* + \hat{\eta}_{ij}^*$, which we call the *propensity to export*. This estimate of the latent variable z_{ij} is a function of both observable and (an estimate of) unobservable trade frictions. The propensity to export can enter our estimating equation in a number of ways. Following HMR, assuming Pareto distributed costs leads to the estimating equation

$$m_{ij} = \psi + \psi_i + \psi_j - \gamma d_{ij} + \ln(e^{\delta \hat{x}_{ij}} - 1) + \hat{\eta}_{ij}^* + \epsilon_{ij}, \quad (40)$$

which we refer to as the Non-Linear (‘NL’) specification because this requires a non-linear estimator. $\epsilon_{ij} \sim N[0, \sigma_e^2]$ is now orthogonal to the variables. The term $\ln(e^{\delta \hat{x}_{ij}} - 1)$ is the empirical analogue of w_{ij} from theory. The constant ψ includes $-y^{I_j}$, the invariant components of the MR terms and the constant parameters affecting firm selection. $\psi_i + \psi_j$ are country-specific fixed effects including GDP and MR terms.

4.2. Regression results

We begin with a reproduction of the coefficients in HMR, using data taken from Elhanan Helpman's web-site (see Appendix A.9). Table 1 presents the results. Column 1 presents the standard OLS regression

¹⁵The error terms are therefore distributed according to a unit normal distribution after this transformation.

results. Column 2 presents the first stage (probit) results. Column 3 presents the non-linear least squares specification (NL), which accounts for firm- and country-selection. For comparison with HMR, we have included the relevant page references at the bottom of the table, together with HMR's maximum likelihood estimates in column 4.¹⁶ In what follows we refer to estimates of γ , the intensive margin, as the *firm-level effect*. When this is combined with the estimates of the extensive margin, we refer to the result as the *country-level effect*. Thus the country level effect combines the impact of changes in trade costs through the extensive and intensive margins.

As explained in HMR, the OLS coefficient is more negative than the NL coefficient because linear estimates conflate effects at the intensive and extensive margin, and γ is only the firm-level effect in the theoretical setup. It can be biased in the opposite direction because of the omission of the inverse Mills ratio. However, the country selection effect is smaller than the firm selection effect for this particular dataset.

The variable indicating similarity of religion is used for identification in the second stage of the estimation procedure. Significance in the probit stage indicates religious similarity affects the fixed costs of exporting. Its non-significance and consequent exclusion from the second stage, conditional on inclusion of controls for country and firm selection, is required for identification. Not including these controls in a one-stage OLS regression does not produce a significant coefficient for religion in column 1.¹⁷ One could erroneously conclude that a fixed cost like religion does not matter for trade volumes. In fact, fixed costs affect the quantity of trade between two countries via the extensive margin.

5. COMPARATIVE STATICS: SIMULATIONS

We start with gross elasticities, which combine the intensive and extensive margins but exclude MR effects. We account subsequently for multilateral resistance in the case of a bilateral reduction in distance and conclude with the case of a multilateral reduction in distance.

5.1. Simulations without multilateral resistance effects: Gross elasticities

Our estimated gross elasticity at the country-level is¹⁸

$$\hat{\xi}_{ij}^{gross} = -\hat{\gamma} - \hat{\gamma}_p \hat{\varphi}_{ij}, \quad (41)$$

where $\hat{\gamma}_p$ is the distance elasticity obtained from the probit equation, given by $\gamma_p = \frac{\partial \hat{x}_{ij}}{\partial d_{ij}}$, and $\hat{\varphi}_{ij} = \frac{\hat{\partial} e^{\hat{\delta} \hat{x}_{ij}}}{e^{\hat{\delta} \hat{x}_{ij}} - 1}$ from our estimated values. The first term in (41) is the intensive margin while the second is the

¹⁶ We have not managed to replicate the NL specification in HMR (2008), but get close to the NL specification in the working paper version HMR (2007), despite the fact that, as in HMR (2008), we use non-linear least squares and that HMR (2007) use maximum likelihood.

¹⁷ It is possible for a fixed cost to show up as significant erroneously because of the omission of the aforementioned controls.

¹⁸ We do not allow for country entry into trade. This is consistent with our theoretical setup. One implication of this is we do not differentiate the inverse Mills ratio with respect to distance.

extensive margin. For all countries, the intensive margin is estimated to be $\hat{\gamma} = 0.799$. Probit estimates imply $\hat{\gamma}_p = 0.66$ for all countries. The theoretical analysis implies γ and γ_p should be equal, but the simulations in HMR implicitly allow them to differ. For consistent comparison with their results, our simulations also allow for this small difference. We calculate $\hat{\varphi}_{ij}$ using each country-pair's estimated propensity to export, \hat{x}_{ij} , together with the estimated $\hat{\delta}$. As noted in HMR, this is the source of cross-country variation in export elasticities in their application, which considers these gross elasticities.

HMR record the elasticities for a 10% fall in distance, which we reproduce in Table 2. The mean overall elasticity of 1.56 implies the country-level effect is due in roughly equal parts to the intensive and extensive margins. The minimum of 1.28 is still somewhat higher than $\hat{\gamma}$ and corresponds to the highest value of $\hat{\rho}$.¹⁹

In Figure 1, we map the elasticities for an infinitesimal change in distance against the propensity to export \hat{x}_{ij} . ‘Linear OLS’ gives conventional estimates of the country-level effect (1.176). ‘Firm-level’ elasticities capture the rise in exports at the intensive margin ($\hat{\gamma} = 0.799$) while ‘Country-level’ accounts for both the intensive and extensive margins, giving the gross country-level effect accounting for firm heterogeneity. Allowing for some movement at the extensive margin produces a larger country-level elasticity for all countries relative to OLS. Thus, in the absence of MR effects, our simulated results indicate larger countries have smaller country-level elasticities, driven by their smaller elasticities at the extensive margin.

Together with Figure 1, Table 2 shows that overall elasticities exceed those from OLS.²⁰ It is important to emphasise this because HMR’s point that the firm-level effect is overestimated by OLS does not imply the same of the country-level effect. In our case OLS provides an underestimate for all country pairs in the sample, conditional on their trading.²¹ The country-level effect is arguably more relevant for policy.

Figure 1 also shows that, the higher the predicted propensity to export \hat{x}_{ij} , the lower the absolute value of the elasticity. It follows that bigger countries will have lower overall country-level elasticities, supporting Lemma 2.²² These results are also consistent with those in HMR, where elasticities are lower for pairs of developed countries than for developing countries.²³ Furthermore, while the theoretical analysis indicated the potential for heterogeneity in elasticities, we have demonstrated empirically that it can be substantial.

¹⁹This raises the issue of what to do with predicted probabilities exceeding 0.9999999. A value of $\rho = 1$ would not allow for the proportion of firms to be identified and, akin to all firms exporting already, generates elasticities approaching $\hat{\gamma} = 0.799$. Alternatively, and seemingly preferred by HMR, the elasticities for $\rho > 0.9999999$ are set equal to those for $\rho = 0.9999999$. This means the proportion of firms is fixed based on this value of ρ and hence the minimum extensive margin is fixed, giving the minimum elasticity of 1.2832. This is what generates the substantial effect at the extensive margin for the highest values of ρ . For further discussion, see Baranga (2008).

²⁰As an alternative, HMR relax the assumption of Pareto-distributed productivity and estimate the firm-selection model with a polynomial function of \hat{x} . Doing so yields a mean elasticity of 1.85 and country-level elasticities that exceed the OLS figure for every observation. Results are available from the authors on request.

²¹Recent studies, such as Baranga (2009) and Belenkiy (2009), have questioned the robustness of the finding that estimates of γ are lower without firm- and country-selection controls. If the reverse held, the country effect would automatically be bigger.

²²By differentiating (41) with respect to \hat{x}_{ij} , it is clear that the elasticity will be lower (less negative) for higher values of \hat{x}_{ij} . Because \hat{x}_{ij} is a positive function of country size, the extensive margin is smaller for bigger countries. This illustrates Lemma 2 when Pareto distributed costs are assumed.

²³Results based on the alternative polynomial specification are also completely consistent with this.

5.2. Simulations with multilateral resistance effects: Net elasticities

We now turn to net elasticities $\hat{\xi}_{ij}^B$ and $\hat{\xi}_{ij}^M$, which do take multilateral resistance effects into account.

5.2.1. Bilateral changes in trade costs

Allowing for $\gamma \neq \gamma_p$ gives the empirical analogue of the net elasticity in (31) as

$$\xi_{ij}^B = \gamma + \varphi_{ij}\gamma_p - (1 + \varphi_{ij}) \frac{\gamma + \delta\gamma_p}{1 + \delta} \left\{ -s_i^{I_j} s_j^{J_i} - s_j^{I_j} s_i^{J_j} + s_j^{J_i} + s_i^{I_j} \right\}. \quad (42)$$

Figure 2 plots 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 ($\hat{\xi}_{ij}^B$) and gross ($\hat{\xi}_{ij}^{gross}$) elasticities are close together. Table 3 produces examples for illustration. The first two rows give the mean and median values for GDP shares and elasticities.²⁴ The typical elasticity after accounting for multilateral resistance is close to the gross elasticity. The table includes the ratio $\frac{\hat{\xi}_{ij}^B}{\hat{\xi}_{ij}^{gross}}$.²⁵ 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 $\frac{\hat{\xi}_{ij}^B}{\hat{\xi}_{ij}^{gross}} \simeq 0.9999$. By contrast, Japan and the USA comprised 45% of world GDP in our 1986 data. The world's two largest countries generate a ratio of $\frac{\hat{\xi}_{ij}^B}{\hat{\xi}_{ij}^{gross}} \simeq 0.79$. 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 1986 data, but they each had less than 2% of world GDP. Their ratio of 0.966 is still sufficiently close to unity to suggest it is only “very” big countries for which MR matters. Empirically, this is a stronger result than Corollary 1.

More generally, bigger countries have lower gross elasticities through the extensive margin *and* lower net-to-gross ratios through multilateral resistance, giving lower net elasticities overall (cf. Proposition 1). This pattern is further illustrated by simple correlation coefficients. There is a negative correlation of -0.29 between $\hat{\xi}_{ij}^{gross}$ and country-pair size (measured by summing their shares of world GDP). By contrast the correlation between $\hat{\xi}_{ij}^B$ and country size is -0.51, illustrating that the effect of MR reinforces that of the extensive margin, as suggested in Proposition 1. Figure 3 illustrates this negative relationship.

While the USA and Japan are the biggest in terms of world GDP shares, Table 4 shows lower net-to-gross ratios are generated by small (and distant) countries exporting to the largest countries, which we considered in theory above (cf. equation (33)). For example, the USA comprised 43% of the share of the GDP of all importers from French Guiana and the elasticity ratio is only $\frac{\hat{\xi}_{ij}^B}{\hat{\xi}_{ij}^{gross}} \simeq 0.57$. 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, multilateral resistance effects are material for (a) trade between the world's largest country

²⁴The GDP shares in Table 4 do not adjust for sets of trading partners.

²⁵This would be well approximated by the right hand side of equation (32), which strictly only applies when $\gamma = \gamma_p$.

pairs and (b) exports from small exporters with few export destinations to the world's largest countries.

5.2.2. Multilateral changes in trade costs

Allowing for $\gamma \neq \gamma_p$ gives the empirical analogue of the net elasticity in (34) as

$$\xi_{ij}^M = \gamma + \varphi_{ij}\gamma_p - (1 + \varphi_{ij}) \frac{\gamma + \delta\gamma_p}{1 + \delta} \left\{ 1 + \sum_{l \in I_j} s_l^{I_j} s_j^{J_l} - s_i^{J_i} - s_j^{I_j} \right\}. \quad (43)$$

Figure 4 demonstrates the dramatic effects of accounting for MR when *all* countries reduce frictions. Net elasticities $\hat{\xi}_{ij}^M$ are substantially lower than gross elasticities $\hat{\xi}_{ij}^{gross}$. The majority of the net elasticities are negative, which demonstrates that Proposition 2 (a) is not just a theoretical possibility, but 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 5 shows the median elasticity is -0.0188, while the mean elasticity is 0.0055. Because international frictions have fallen only relative to domestic frictions, one should expect a net effect close to zero.

Can we characterise the source of the MR effects and see where they are strongest? Table 5 hints at a positive relationship along the real line between the net-to-gross ratio $\frac{\hat{\xi}_{ij}^M}{\hat{\xi}_{ij}^{gross}}$ and the size of the trading pair.²⁶ 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 $\frac{\hat{\xi}_{ij}^M}{\hat{\xi}_{ij}^{gross}} \simeq 0.5$. The theoretical analysis introduced the opposing actions in the relationship between country size and the elasticity: bigger pairs have bigger multipliers (due to MR) but lower gross elasticities (due to the extensive margin). Recall the correlation between pair-size 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 5 demonstrates the overall positive relationship across all country-pairs. Thus, while weakened by the extensive margin effects, AvW's 'Implication 1' prevails in our data.

Overall, we have approximately 2,300 positive elasticities. This begs the question of where the positive elasticities are generated. In our theoretical section 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 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,

²⁶ As for the bilateral case, the multiplier is calculated according to $\frac{\xi_{ij}^M}{\xi_{ij}^{gross}}$. This would be well approximated by equation (38) when $\gamma = \gamma_p$.

they reduce their exports to most smaller countries to be able to expand their exports to a handful of big countries.

Aggregating elasticities We aggregate the bilateral elasticities by exporter, weighting each bilateral elasticity by the volume of exports. As should be the case, all countries see a net increase in trade even after accounting for MR. The mean elasticity is 0.29. One third of our countries have elasticities of less than 0.1. Canada (0.65) and Mexico (0.67) generate the biggest elasticities and we know what they have in common.

Some studies, for example Wilson, Mann and Otsuki (2005), use gravity models to simulate the effects of multilateral reductions in trade barriers on global trade. We calculate world-wide elasticities, which are reported in Table 6. The homogenous parameter γ implies the intensive margin effect is 0.799 for the whole world. Allowing for the extensive margin of trade to operate and aggregating across all bilateral elasticities, the gross world trade elasticity $\hat{\xi}^{gross}$ is 1.29. This is close to the minimum of the distribution because the world's largest economies have elasticities at the minimum or very close to it.²⁷ With reference to equation (34), allowing for the extensive margin multiplies the gravity model parameter by more than 1.5. Including MR effects however, aggregating $\hat{\xi}^M$ yields a world-wide net trade elasticity of 0.47. This means that, for the world as a whole, the net-to-gross ratio $\frac{\hat{\xi}^M}{\hat{\xi}^{gross}}$ is approximately one third. 0.47 is also less than half of the elasticity implied by OLS, which is 1.17. Hence, existing gravity-based simulations of multilateral reductions in trade barriers can be seriously misleading.

In sum, Table 6 confirms that multilateral resistance 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.

6. CONCLUSION

We have presented a gravity model that accounts explicitly for firm selection into exports and for general equilibrium price effects acting through multilateral resistance terms. We have shown that multilateral resistance affects the extensive margin of trade, which takes a gravity form similar to the equation for bilateral trade. While using fixed effects for estimation, a Taylor approximation along the lines of Baier & Bergstrand (2009) can be used to allow for multilateral resistance effects when conducting comparative statics. We have shown that the approach can be extended to include the effects of firm heterogeneity on multilateral resistance.

We have emphasised that overestimates of the firm-level distance coefficient do not imply an overes-

²⁷This result is therefore susceptible to the truncation of \hat{p} - see footnote 19.

timate of the overall country-level effect in the presence of firm heterogeneity. Our data show that, for all countries that already trade, failing to account for firm heterogeneity underestimates the effects of a fall in trade frictions. The intensive margin makes up approximately half the effect on average. The extensive margin elasticity is smaller the larger is the country-pair involved.

Isolated bilateral reductions in trade frictions by two countries are typically subject to small MR effects. For large countries however, even bilateral changes in trade costs generate material MR effects. In addition, the simultaneous modelling of MR and heterogeneity identifies some small exporters — those with few trading partners and who export to a large importer — as additional candidates for which MR effects are important under bilateral changes in trade costs. Overall, bigger country pairs have lower net elasticities when bilateral trade costs change in isolation.

In contrast to the bilateral case, our simulations for multilateral changes in trade costs show that MR wipes out a large proportion of the comparative static effect found when MR is ignored. MR reduces the responsiveness of world trade by two thirds. Despite the effect at the extensive margin, which raises trade elasticities, we calculate a world elasticity of 0.47. This is less than half of that typically implied by the ‘naive’ OLS coefficient. At the country-pair level, bilateral trade often falls as exports are redirected from some destinations to others. The MR effect dampens elasticities by less for larger countries, but the extensive margin effect is bigger for smaller traders. Net country-level elasticities are therefore potentially non-monotonic in country size.

The Taylor approach provides a tractable way to generate MR effects for the purposes of comparative statics. It is, however, an approximation. In Baier and Bergstrand’s (2009) sample of 88 countries, Monte Carlo simulations revealed only 8% of simulation results differed from the AvW system by 20% or more. The biggest inaccuracy was a 38% deviation from the ‘true’ value. Baier and Bergstrand note that the largest inaccuracies involved countries in the European Economic Area (EEA) which were relatively small and relatively close to their largest trading partners. Extending beyond the EEA in our case, some North African countries may be in the same category for our 1986 data, while some small East Asian countries may now be susceptible because of the emergence of China and India. Bhutan, for example, borders both. Baier and Bergstrand propose an iterative method which allows the Taylor approximation to converge on the results produced by the AvW system. This certainly is an approach that can be taken when it comes to implementation. But our main message is that to employ a first-order adjustment is better than to use no adjustment for MR at all.

The number of negative bilateral trade responses we recorded have not allowed for the possibility that some countries would no longer trade. We have not considered the formation of new export relationships between two countries with zero trade. While HMR argue that very little of the expansion of world trade seen over the last few decades is attributable to new trading pairs, another fruitful area for research would be the use of the empirical framework to examine this possibility.

REFERENCES

- Anderson, J. (1979), ‘A theoretical foundation for the gravity equation’, *American Economic Review* pp. 106–116.
- Anderson, J. (2009), ‘Gravity, Productivity and the Pattern of Production and Trade’, *mimeo* .
- Anderson, J. and Van Wincoop, E. (2003), ‘Gravity with gravitas: a solution to the border puzzle’, *American Economic Review* **93**(1), 170–192.
- Baier, S. and Bergstrand, J. (2009), ‘Bonus vetus OLS: A simple method for approximating international trade-cost effects using the gravity equation’, *Journal of International Economics* .
- Baranga, T. (2009), ‘Unreported Trade Flows and Gravity Equation Estimation’, *mimeo* .
- Belenkiy (2009), ‘Robustness of the Extensive Margin in the Helpman, Melitz and Rubinstein (HMR) Model’, *mimeo* .
- Bergstrand, J. (1985), ‘The gravity equation in international trade: some microeconomic foundations and empirical evidence’, *Review of Economics and Statistics* pp. 474–481.
- Bergstrand, J., Egger, P. and Larch, M. (2007), ‘Gravity Redux: Structural Estimation of Gravity Equations with Asymmetric Bilateral Trade Costs’, *mimeo* .
- Bernard, A., Jensen, J., Redding, S. and Schott, P. (2007), ‘Firms in international trade’, *Journal of Economic Perspectives* **21**(3), 105–130.
- Carayol (2006), ‘RE: “Expression too long” error with nl command’, *Accessed via Statalist* .
- Chaney, T. (2008), ‘Distorted gravity: The intensive and extensive margins of international trade’, *American Economic Review* **98**(4), 1707–1721.
- Eaton, J. and Kortum, S. (2002), ‘Technology, geography, and trade’, *Econometrica* pp. 1741–1779.
- Helpman, E., Melitz, M. and Rubinstein, Y. (2007), ‘Estimating Trade Flows: Trading Partners and Trading Volumes’, *NBER Working Paper 12927* .
- Helpman, E., Melitz, M. and Rubinstein, Y. (2008), ‘Estimating Trade Flows: Trading Partners and Trading Volumes’, *Quarterly Journal of Economics* **123**(2), 441–487.
- Helpman, E., Melitz, M. and Yeaple, S. (2004), ‘Export versus FDI with heterogeneous firms’, *American Economic Review* **94**(1), 300–316.
- McCallum, J. (1995), ‘National borders matter: Canada-US regional trade patterns’, *American Economic Review* pp. 615–623.

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

APPENDIX A: PROOFS AND DERIVATIONS

A.1. Deriving the gravity equation

Using bilateral trade balance in (8) gives

$$Y_j = \left(\frac{c_j}{\alpha}\right)^{1-\sigma} N_j \sum_{i \in I_j} \left(\frac{t_{ij}}{P_i}\right)^{1-\sigma} Y_i V_{ij}, \quad (44)$$

$$\Rightarrow \left(\frac{c_j}{\alpha}\right)^{1-\sigma} N_j = \frac{Y_j}{\sum_{i \in I_j} \left(\frac{t_{ij}}{P_i}\right)^{1-\sigma} Y_i V_{ij}}. \quad (45)$$

Use in (8) to obtain

$$M_{ij} = Y_i Y_j \frac{1}{P_i^{1-\sigma}} \frac{t_{ij}^{1-\sigma}}{\sum_{i \in I_j} \left(\frac{t_{ij}}{P_i}\right)^{1-\sigma} Y_i V_{ij}} V_{ij}. \quad (46)$$

Then define Y^{I_j} as the total GDP of all importers from j , or $Y^{I_j} \equiv \sum_{i \in I_j} Y_i$. Dividing top and bottom by this gives

$$M_{ij} = \frac{Y_i Y_j}{Y^{I_j}} \frac{1}{P_i^{1-\sigma}} \frac{t_{ij}^{1-\sigma}}{\sum_{i \in I_j} \left(\frac{t_{ij}}{P_i}\right)^{1-\sigma} \frac{Y_i}{Y^{I_j}} V_{ij}^{\sigma-1}} V_{ij}. \quad (47)$$

Further, define $s_i^{I_j}$ as the share of i ’s GDP in the set of all importers from j , or $s_i^{I_j} \equiv \frac{Y_i}{Y^{I_j}} = \frac{Y_i}{\sum_{i \in I_j} Y_i}$.

Then it follows that $\sum_{i \in I_j} s_i^{I_j} = \sum_{i \in I_j} \frac{Y_i}{\sum_{i \in I_j} Y_i} = 1$. We write the gravity equation as

$$M_{ij} = \frac{Y_i Y_j}{Y^{I_j}} \frac{1}{P_i^{1-\sigma}} \frac{t_{ij}^{1-\sigma}}{\sum_{h \in I_j} \left(\frac{t_{hj}}{P_h}\right)^{1-\sigma} s_h^{I_j} V_{hj}} V_{ij}. \quad (48)$$

Finally, define the exporter’s MR term $\hat{P}_j^{1-\sigma} \equiv \sum_{h \in I_j} \left(\frac{t_{hj}}{P_h}\right)^{1-\sigma} s_h^{I_j} V_{hj}$. Note that, intuitively, it is defined over the set of importers from j , I_j . Using this in our earlier equation

$$\left(\frac{c_j}{\alpha}\right)^{1-\sigma} N_j = \frac{Y_j}{Y^{I_j}} \frac{1}{\hat{P}_j^{1-\sigma}} \quad (49)$$

$$\Rightarrow \hat{P}_j^{1-\sigma} = \frac{Y_j}{Y^{I_j} \left(\frac{c_j}{\alpha}\right)^{1-\sigma} N_j}, \quad (50)$$

which expresses the level of the exporter's MR in terms of exogenous variables. In (49), perform

$$\left(\frac{c_j}{\alpha}\right)^{1-\sigma} N_j = \frac{Y_j}{Y^{I_j}} \frac{1}{\widehat{P}_j^{1-\sigma}} \frac{\sum_{j \in J_i} Y_j}{\sum_{j \in J_i} Y_j} \quad (51)$$

$$= \frac{Y_j}{\sum_{j \in J_i} Y_j} \frac{1}{\widehat{P}_j^{1-\sigma}} \frac{\sum_{j \in J_i} Y_j}{\sum_{i \in I_j} Y_i}, \quad (52)$$

Then define $s_j^{J_i}$ is the share of j 's GDP in the set of all exporters to i , such that $s_j^{J_i} \equiv \frac{Y_j}{\sum_{j \in J_i} Y_j}$. We then have the convenient property that $\sum_{j \in J_i} s_j^{J_i} \equiv \sum_{j \in J_i} \frac{Y_j}{\sum_{j \in J_i} Y_j} = 1$. We are left with the object $\sum_{j \in J_i} Y_j / \sum_{i \in I_j} Y_i$. It is the ratio of the total output of all exporters to i to that of all importers from j . It is defined over the sets I_j and J_i , and as such is an “ ij ” variable. Define this ratio as $R_{ij} \equiv \frac{\sum_{j \in J_i} Y_j}{\sum_{i \in I_j} Y_i}$ and use it to write

$$\left(\frac{c_j}{\alpha}\right)^{1-\sigma} N_j = \frac{s_j^{J_i}}{\widehat{P}_j^{1-\sigma}} R_{ij}, \quad (53)$$

and the price index for the importer as

$$P_i^{1-\sigma} = \sum_{j \in J_i} (c_j t_{ij} / \alpha)^{1-\sigma} N_j V_{ij} \quad (54)$$

$$= \sum_{j \in J_i} \left(\frac{t_{ij}}{\widehat{P}_j} \right)^{1-\sigma} s_j^{J_i} V_{ij} R_{ij}. \quad (55)$$

The multilateral resistance terms are therefore given by

$$\widehat{P}_j^{1-\sigma} = \sum_{i \in I_j} \left(\frac{t_{ij}}{P_i} \right)^{1-\sigma} s_i^{I_j} V_{ij}, \quad (56)$$

$$P_i^{1-\sigma} = \sum_{j \in J_i} \left(\frac{t_{ij}}{\widehat{P}_j} \right)^{1-\sigma} s_j^{J_i} V_{ij} R_{ij}. \quad (57)$$

The gravity equation then looks like that in the text (10),

$$M_{ij} = \frac{Y_i Y_j}{Y^{I_j}} \left(\frac{t_{ij}}{P_i \widehat{P}_j} \right)^{1-\sigma} V_{ij}. \quad (58)$$

A.2. Deriving the gravity equation for the extensive margin

Recall equation (6) in the text gives the extensive margin as

$$a_{ij} = \left[\frac{Y_i (1 - \alpha)}{f_{ij} c_j} \right]^{\frac{1}{\sigma-1}} \frac{\alpha P_i}{c_j t_{ij}}, \quad (59)$$

and, after assuming trade balance, equation (50) gives

$$\hat{P}_j = \left[\frac{Y_j}{Y^{I_j} \left(\frac{c_j}{\alpha}\right)^{1-\sigma} N_j} \right]^{\frac{1}{1-\sigma}}. \quad (60)$$

Using this in (6) allows us to write

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

which can be written as

$$a_{ij}^{\sigma-1} = \left[\frac{(1-\alpha)}{c_j N_j} \right] \frac{Y_i Y_j}{Y^{I_j}} \frac{1}{f_{ij}} \frac{t_{ij}^{1-\sigma}}{\left(P_i \hat{P}_j\right)^{1-\sigma}}, \quad (61)$$

in the text.

A.3. Deriving the latent variable

In the text, we define the ratio of the variable profits of the firm with lowest costs to the fixed costs of exporting as

$$Z_{ij} \equiv \frac{Y_i (1-\alpha) \left(\frac{c_j}{\alpha \hat{P}_i}\right)^{1-\sigma} (t_{ij} a_L)^{1-\sigma}}{c_j f_{ij}}, \quad (62)$$

which, using equation (49) giving $\left(\frac{c_j}{\alpha}\right)^{1-\sigma} = \frac{Y_j}{Y^{I_j}} \frac{1}{N_j \hat{P}_j^{1-\sigma}}$, can be written

$$\begin{aligned} Z_{ij} &= \frac{Y_i Y_j (1-\alpha) \frac{1}{N_j} (t_{ij} a_L)^{1-\sigma}}{Y^{I_j} c_j f_{ij}} \left(P_i \hat{P}_j\right)^{\sigma-1} \\ &= \tilde{Z}_{ij} \left(P_i \hat{P}_j\right)^{\sigma-1}, \end{aligned}$$

used in the text.

A.4. Comparative statics

A.4.1. Gross and Net elasticities

Taking logs of (10) gives

$$m_{ij} = y_i + y_j - y^{I_j} - (\sigma-1) \ln t_{ij} + w_{ij} + (\sigma-1) \ln P_i \hat{P}_j, \quad (63)$$

where $w_{ij} \equiv \ln \{e^{\delta z_{ij}} - 1\} = \ln \left\{e^{\delta [\tilde{z}_{ij} + (\sigma-1) \ln P_i \hat{P}_j]} - 1\right\}$ is the log of the extensive margin containing $\tilde{z}_{ij} = \ln \tilde{Z}_{ij}$ given by

$$\tilde{z}_{ij} = \chi + \chi_i + \chi_j - (\sigma-1) \ln t_{ij} - \ln f_{ij}.$$

We specify trade costs t_{ij} as $t_{ij}^{\sigma-1} = D_{ij}^\gamma$, such that $(\sigma - 1) \ln t_{ij} = \gamma d_{ij}$, where $d_{ij} = \ln D_{ij}$. This yields

$$m_{ij} = y_i + y_j - y^{I_j} - \gamma d_{ij} + w_{ij} + (\sigma - 1) \ln P_i \hat{P}_j \quad (64)$$

$$\tilde{z}_{ij} = \chi + \chi_i + \chi_j - \gamma d_{ij} - \ln f_{ij}. \quad (65)$$

Then for a change in trade costs ∂d , the change in log imports by i from j is

$$\frac{\partial m_{ij}}{\partial d} = -\gamma + \varphi_{ij} \left[-\gamma + \frac{\partial \ln (P_i P_j)^{\sigma-1}}{\partial d} \right] + \frac{\partial \ln (P_i P_j)^{\sigma-1}}{\partial d}, \quad (66)$$

which gives an elasticity $\xi_{ij}^{net} \equiv -\frac{\partial m_{ij}}{\partial d}$ of

$$\xi_{ij}^{net} = \gamma + \varphi_{ij} \gamma - (1 + \varphi_{ij}) \frac{\partial \ln (P_i P_j)^{\sigma-1}}{\partial d}, \quad (67)$$

where $\varphi_{ij} \equiv \frac{\delta e^{\delta z_{ij}}}{e^{\delta z_{ij}} - 1}$. We refer to the elasticity omitting the last term, which is the effect due to MR, as the gross elasticity, $\xi_{ij} = \gamma(1 + \varphi_{ij})$. Including MR effects however gives the net elasticity, equation (67).

A.4.2. Multilateral resistance approximation

To compute net elasticities, we need to compute $\frac{\partial \ln (P_i P_j)^{\sigma-1}}{\partial d}$. Our method is to take a first-order Taylor approximation around a symmetric centre in which all countries trade, following Baier and Bergstrand. Our model contains an additional effect due to the impact of MR on the extensive margin made clear in equation (20). The price indices therefore 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 appears in the price index equations can be decomposed according to $V_{ij} \simeq \tilde{Z}_{ij}^\delta (P_i \hat{P}_j)^{\delta(\sigma-1)}$, or $w_{ij} \simeq \delta \tilde{z}_{ij} + \ln (P_i \hat{P}_j)^{\delta(\sigma-1)}$. This approximated extensive margin value approaches that computed under the assumption of a Pareto distribution for δz_{ij} sufficiently large, since $w_{ij} = \ln \{e^{\delta z_{ij}} - 1\} \simeq \delta z_{ij} = \delta \tilde{z}_{ij} + \ln (P_i \hat{P}_j)^{\delta(\sigma-1)}$ when this is the case. Using this, and defining $W_{ij} \equiv \tilde{Z}_{ij}^\delta$ yields price indices of

$$\hat{P}_j^{1-\sigma} = \sum_{i \in I_j} \left(\frac{t_{ij}}{P_i} \right)^{1-\sigma} s_i^{I_j} (P_i \hat{P}_j)^{\delta(\sigma-1)} W_{ij} \quad (68)$$

$$\Rightarrow \hat{P}_j^{-k} = \sum_{i \in I_j} t_{ij}^{1-\sigma} s_i^{I_j} P_i^k W_{ij} \quad (69)$$

$$P_i^{1-\sigma} = \sum_{j \in J_i} \left(\frac{t_{ij}}{\hat{P}_j} \right)^{1-\sigma} s_j^{J_i} (P_i \hat{P}_j)^{\delta(\sigma-1)} W_{ij} R_{ij} \quad (70)$$

$$\Rightarrow P_i^{-k} = \sum_{j \in J_i} t_{ij}^{1-\sigma} s_j^{J_i} \hat{P}_j^k W_{ij} R_{ij}, \quad (71)$$

where we have used that $(\delta + 1)(\sigma - 1) = (\frac{k-\sigma+1}{\sigma-1} + 1)(\sigma - 1) = k$. The terms (69) and (71) are those which we Taylor approximate. We do so as follows. Consider a symmetric world in which all countries trade. Then $t_{ij} = t$, $W_{ij} = W$, $P_i = P$, $R_{ij} = 1$, $I_j = J_i$ and $\hat{P}_j = \hat{P}$. This yields the system

$$\begin{aligned}\hat{P}^{-k} &= t^{1-\sigma} P^k W \sum_i s_i = t^{1-\sigma} P^k W, \\ P^{-k} &= t^{1-\sigma} \hat{P}^k W \sum_j s_j = t^{1-\sigma} \hat{P}^k W,\end{aligned}$$

which has a solution at $P = \hat{P}$. This is the centre around which we take our approximation. To do this, we use

$$e^{-k \ln \hat{P}_j} = \sum_{i \in I_j} e^{\ln s_i^{J_j}} e^{(1-\sigma) \ln t_{ij}} e^{k \ln P_i} e^{\ln W_{ij}}, \quad (72)$$

$$e^{-k \ln P_i} = \sum_{j \in J_i} e^{\ln s_j^{J_i}} e^{(1-\sigma) \ln t_{ij}} e^{k \ln \hat{P}_j} e^{\ln W'_{ij}}, \quad (73)$$

where $W'_{ij} \equiv W_{ij} R_{ij}$. Then expanding the LHS of (73) around $\ln P$ and the RHS around $\ln t$, $\ln W'$, and $\ln \hat{P}$ gives

$$\begin{aligned}P^{-k} - kP^{-k}[\ln P_i - \ln P] &= \sum_{j \in J_i} s_j^{J_i} t^{1-\sigma} W' \hat{P}^k + \sum_{j \in J_i} s_j^{J_i} (1-\sigma) t^{1-\sigma} W' \hat{P}^k [\ln t_{ij} - \ln t] \\ &\quad + \sum_{j \in J_i} s_j^{J_i} t^{1-\sigma} W' \hat{P}^k [\ln W'_{ij} - \ln W'] + \sum_{j \in J_i} s_j^{J_i} k t^{1-\sigma} W' \hat{P}^k [\ln \hat{P}_j - \ln \hat{P}].\end{aligned}$$

Divide through by $P^{-k} = t^{1-\sigma} W' \hat{P}^k$ and use that $\sum_{j \in J_i} s_j^{J_i} = 1$ to give

$$\begin{aligned}-k[\ln P_i - \ln P] &= \sum_{j \in J_i} s_j^{J_i} (1-\sigma) [\ln t_{ij} - \ln t] \\ &\quad + \sum_{j \in J_i} s_j^{J_i} [\ln W'_{ij} - \ln W'] + \sum_{j \in J_i} s_j^{J_i} k [\ln \hat{P}_j - \ln \hat{P}],\end{aligned}$$

simplifying to

$$\begin{aligned}-k \ln P_i + k \ln P &= \sum_{j \in J_i} s_j^{J_i} (1-\sigma) \ln t_{ij} - (1-\sigma) \ln t \\ &\quad + \sum_{j \in J_i} s_j^{J_i} \ln W'_{ij} - \ln W' + \sum_{j \in J_i} s_j^{J_i} k \ln \hat{P}_j - k \ln \hat{P},\end{aligned}$$

and use that $-k \ln P = (1-\sigma) \ln t + \ln W' + k \ln \hat{P}$ to give

$$\ln P_i = \sum_{j \in J_i} s_j^{J_i} \left(\frac{\sigma-1}{k} \right) \ln t_{ij} - \frac{1}{k} \sum_{j \in J_i} s_j^{J_i} \ln W_{ij} R_{ij} - \sum_{j \in J_i} s_j^{J_i} \ln \hat{P}_j. \quad (74)$$

The term $\frac{1}{k} \sum_{j \in J_i} s_j^{J_i} \ln R_{ij}$ is an i -specific constant which is invariant to the changes in trade costs that we consider, so that, for our comparative statics purposes, we can ignore it. Writing the approximated price indices in terms using subscript h as the generic exports and l as the generic importer, we have

$$\ln P_i = \sum_{h \in J_i} s_h^{J_i} \left(\frac{\sigma - 1}{k} \right) \ln t_{ih} - \frac{1}{k} \sum_{h \in J_i} s_h^{J_i} \ln W_{ih} - \sum_{h \in J_i} s_h^{J_i} \ln \hat{P}_h, \quad (75)$$

$$\ln \hat{P}_j = \sum_{l \in I_j} s_l^{I_j} \left(\frac{\sigma - 1}{k} \right) \ln t_{lj} - \frac{1}{k} \sum_{l \in I_j} s_l^{I_j} \ln W_{lj} - \sum_{l \in I_j} s_l^{I_j} \ln P_l. \quad (76)$$

Substituting the first line into the second gives

$$\begin{aligned} \ln \hat{P}_j &= \frac{1}{k} \sum_{l \in I_j} s_l^{I_j} [(\sigma - 1) \ln t_{lj} - \ln W_{lj}] \\ &\quad - \sum_{l \in I_j} s_l^{I_j} \left[\frac{1}{k} \sum_{h \in J_l} s_h^{J_l} [(\sigma - 1) \ln t_{lh} - \ln W_{lh}] - \sum_{h \in J_l} s_h^{J_l} \ln \hat{P}_h \right], \end{aligned}$$

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

$$\begin{aligned} \ln P_i + \ln \hat{P}_j &= \frac{1}{k} \sum_{h \in J_i} s_h^{J_i} [(\sigma - 1) \ln t_{ih} - \ln W_{ih}] - \sum_{h \in J_i} s_h^{J_i} \ln \hat{P}_h \\ &\quad + \frac{1}{k} \sum_{l \in I_j} s_l^{I_j} [(\sigma - 1) \ln t_{lj} - \ln W_{lj}] \\ &\quad - \sum_{l \in I_j} s_l^{I_j} \left[\frac{1}{k} \sum_{h \in J_l} s_h^{J_l} [(\sigma - 1) \ln t_{lh} - \ln W_{lh}] - \sum_{h \in J_l} s_h^{J_l} \ln \hat{P}_h \right] \\ &= -\frac{1}{k} \sum_{l \in I_j} s_l^{I_j} \sum_{h \in J_l} s_h^{J_l} [(\sigma - 1) \ln t_{lh} - \ln W_{lh}] + \frac{1}{k} \sum_{h \in J_i} s_h^{J_i} [(\sigma - 1) \ln t_{ih} - \ln W_{ih}] \\ &\quad + \frac{1}{k} \sum_{l \in I_j} s_l^{I_j} [(\sigma - 1) \ln t_{lj} - \ln W_{lj}] - \sum_{h \in J_i} s_h^{J_i} \ln \hat{P}_h + \sum_{l \in I_j} s_l^{I_j} \sum_{h \in J_l} s_h^{J_l} \ln \hat{P}_h. \end{aligned}$$

Then use (50) to argue that, with fixed output, costs and firm numbers, $\hat{P}_j = \eta_j$ where η_j is an exporter specific constant. The terms $-\sum_{h \in J_i} s_h^{J_i} \ln \hat{P}_h + \sum_{l \in I_j} s_l^{I_j} \sum_{h \in J_l} s_h^{J_l} \ln \hat{P}_h$ are then given by $-\sum_{h \in J_i} s_h^{J_i} \ln \eta_h + \sum_{l \in I_j} s_l^{I_j} \sum_{h \in J_l} s_h^{J_l} \ln \eta_h$ and are irrelevant for the purposes of comparative statics. Using $(\sigma - 1) \ln t_{ij} = \gamma d_{ij}$ and $\ln W_{ij} = \delta \tilde{z}_{ij}$, our MR approximation is then given by

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

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

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

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

$$(\sigma-1) \ln P_i \hat{P}_j = \frac{1}{\delta+1} \left\{ - \sum_{l \in I_j} s_l^{I_j} \sum_{h \in J_l} s_h^{J_l} [\gamma d_{lh} - \delta \tilde{z}_{lh}] + \sum_{h \in J_i} s_h^{J_i} [\gamma d_{ih} - \delta \tilde{z}_{ih}] + \sum_{l \in I_j} s_l^{I_j} [\gamma d_{lj} - \delta \tilde{z}_{lj}] \right\}. \quad (77)$$

The right hand side of this expression contains only exogenous variables, since both γd_{ij} and $\delta \tilde{z}_{ij}$ are free of price index terms. We use this expression for our comparative statics.

A.5. Proof of Lemma 1: MR elasticities

Proof. Denote the elasticity of this MR with respect to a change in trade costs, by $\varepsilon_{ij} = \frac{\partial \ln(P_i \hat{P}_j)^{\sigma-1}}{\partial d}$.

Consider two cases

1. $\partial d_{ij} = \partial d$ for all i, j , for $i \neq j$. This is a *Multilateral* (M) change in trade costs. Then

$$\begin{aligned} \varepsilon_{ij}^M &= \frac{1}{\delta+1} \left\{ - \sum_{l \in I_j} s_l^{I_j} \sum_{h \in J_l, h \neq j} s_h^{J_l} [\gamma + \delta \gamma] + \sum_{h \in J_i, h \neq i} s_h^{J_i} [\gamma + \delta \gamma] + \sum_{l \in I_j, l \neq j} s_l^{I_j} [\gamma + \delta \gamma] \right\} \\ &= \gamma \left\{ 1 + \sum_{l \in I_j} s_l^{I_j} s_j^{J_l} - s_i^{J_i} - s_j^{I_j} \right\}. \end{aligned}$$

Then for a change in the size of importer i , $\partial s_i^{J_i} = \partial s_i^{I_j} = \partial s_i$,

$$\frac{\partial \varepsilon_{ij}^M}{\partial s_i} = \gamma \left\{ s_j^{J_i} - 1 \right\} < 0, \quad \text{if } s_j^{J_i} < 1.$$

Thus as s_i increases from 0 to 1, or as country i trades more with itself but less internationally, the elasticity of its price index, or importing MR term, with respect to *multilateral* changes in trade costs falls. AvW's 'Implication 1' result therefore holds under our approximation: larger importers have smaller MR elasticities with respect to multilateral changes in trade costs.

2. $\partial d_{ij} = \partial d_{ji} > 0$, and $\partial d_{ij} = 0$ for all country pairs except ij . This is a *Bilateral* (B) change in trade costs. Then we can write, for given exporter MR terms, that

$$\begin{aligned} \varepsilon_{ij}^B &= \frac{1}{\delta+1} \left\{ - \left(s_i^{I_j} s_j^{J_i} [\gamma + \delta \gamma] + s_j^{I_j} s_i^{J_j} [\gamma + \delta \gamma] \right) + s_j^{J_i} [\gamma + \delta \gamma] + s_i^{I_j} [\gamma + \delta \gamma] \right\} \\ &= \gamma \left\{ -s_i^{I_j} s_j^{J_i} - s_j^{I_j} s_i^{J_j} + s_j^{J_i} + s_i^{I_j} \right\}. \end{aligned}$$

Then for a change in the size of importer i , $\partial s_i^{J_i} = \partial s_i^{I_j} = \partial s_i$,

$$\frac{\partial \varepsilon_{ij}^B}{\partial s_i} = \gamma \left\{ 1 - s_j^{J_i} - s_j^{I_j} \right\} > 0, \quad \text{if } s_j^{J_i} + s_j^{I_j} < 1.$$

This is in contrast to the multilateral case in (a). It shows that, for bilateral changes in trade costs, larger importers have larger price index elasticities all else equal. Intuitively,

when trade is liberalised bilaterally between two large countries, they ‘count for more’ in each others’ MR terms. That is, bigger countries are bigger determinants of multilateral resistance. Hence when countries such as these liberalise trade bilaterally, they experience larger general equilibrium effects through MR than smaller countries.

■

A.6. Proof of Lemma 2: Extensive margin elasticity

Proof. From the text, we have that

$$\frac{\partial w_{ij}}{\partial d_{ij}} = \frac{\delta e^{\delta z_{ij}}}{e^{\delta z_{ij}} - 1} \left[\frac{\partial \tilde{z}_{ij}}{\partial d_{ij}} + \frac{\partial \ln \left(P_i \hat{P}_j \right)^{(\sigma-1)}}{\partial d_{ij}} \right] \quad (78)$$

$$= -\varphi_{ij}\gamma + \varphi_{ij} \frac{\partial \ln \left(P_i \hat{P}_j \right)^{(\sigma-1)}}{\partial d_{ij}}, \quad (79)$$

$$\varphi_{ij} \equiv \frac{\delta e^{\delta z_{ij}}}{e^{\delta z_{ij}} - 1}. \quad (80)$$

The first term is the gross extensive margin elasticity, which we denote $\frac{\partial \tilde{w}_{ij}}{\partial d_{ij}} \equiv -\varphi_{ij}\gamma$. Now consider the derivative of the gross elasticity with respect to s_j , holding constant the MR effect. We have

$$\begin{aligned} \left. \frac{\partial^2 \tilde{w}_{ij}}{\partial d_{ij} \partial Y_j} \right|_{P_i, \hat{P}_j} &= -\gamma \frac{\partial}{\partial Y_j} \left(\frac{\delta e^{\delta z_{ij}}}{e^{\delta z_{ij}} - 1} \right) \\ &= -\gamma \times -\frac{\delta^2 e^{\delta z_{ij}}}{(e^{\delta z_{ij}} - 1)^2} > 0, \\ \left. \frac{\partial^2 \tilde{w}_{ij}}{\partial d_{ij} \partial Y_j} \right|_{P_i, \hat{P}_j} &= -\gamma \times -\frac{\delta^2 e^{\delta z_{ij}}}{(e^{\delta z_{ij}} - 1)^2} \times \left. \frac{\partial z_{ij}}{\partial Y_i} \right|_{P_i, \hat{P}_j} \\ &= -\gamma \times -\frac{\delta^2 e^{\delta z_{ij}}}{(e^{\delta z_{ij}} - 1)^2} > 0, \end{aligned}$$

since $\left. \frac{\partial z_{ij}}{\partial Y_j} \right|_{P_i, \hat{P}_j}, \left. \frac{\partial z_{ij}}{\partial Y_i} \right|_{P_i, \hat{P}_j} = 1$. We then write

$$\frac{\partial \tilde{w}_{ij}}{\partial d_{ij}} < 0, \quad \left. \frac{\partial^2 \tilde{w}_{ij}}{\partial d_{ij} \partial Y_j} \right|_{P_i, \hat{P}_j} > 0, \quad (81)$$

such that larger exporters have gross extensive margin elasticities of smaller absolute size. ■

A.7. Proof of Proposition 1: Bilateral changes

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

$$\xi_{ij}^B = \gamma + \varphi_{ij}\gamma - (1 + \varphi_{ij}) \frac{\partial \ln \left(P_i \widehat{P}_j \right)^{\sigma-1}}{\partial d_{ij,ji}}, \quad (82)$$

$$\frac{\partial \ln \left(P_i \widehat{P}_j \right)^{\sigma-1}}{\partial d_{ij,ji}} = \gamma \left\{ -s_i^{I_j} s_j^{J_i} - s_j^{I_j} s_i^{J_j} + s_j^{J_i} + s_i^{I_j} \right\}, \quad (83)$$

$$\Rightarrow \xi_{ij}^B = \gamma + \varphi_{ij}\gamma - (1 + \varphi_{ij}) \gamma \left\{ -s_i^{I_j} s_j^{J_i} - s_j^{I_j} s_i^{J_j} + s_j^{J_i} + s_i^{I_j} \right\}. \quad (84)$$

This yields

$$\xi_{ij}^B = \gamma (1 + \varphi_{ij}) \left\{ 1 + s_i^{I_j} s_j^{J_i} + s_j^{I_j} s_i^{J_j} - s_j^{J_i} - s_i^{I_j} \right\}. \quad (85)$$

Then consider a change in the size of importer i such that $\partial s_i^{J_i} = \partial s_i^{I_j} = \partial s_i$. Then

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

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

$$\frac{\partial \xi_{ij}^B}{\partial s_i} < 0,$$

when $s_j^{J_i} + s_j^{I_j} < 1$. Thus, Lemmata 1 and 2 combine to establish Proposition 1. ■

A.8. Proof of Proposition 2: Multilateral Changes

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

$$\xi_{ij}^M = \gamma + \varphi_{ij}\gamma - (1 + \varphi_{ij}) \frac{\partial \ln \left(P_i \widehat{P}_j \right)^{\sigma-1}}{\partial d}, \quad (86)$$

$$\begin{aligned} \frac{\partial \ln \left(P_i \widehat{P}_j \right)^{\sigma-1}}{\partial d} &= \gamma \left\{ 1 + \sum_{l \in I_j} s_l^{I_j} s_j^{J_l} - s_i^{J_i} - s_j^{I_j} \right\}, \\ \Rightarrow \xi_{ij}^M &= \gamma + \varphi_{ij}\gamma - (1 + \varphi_{ij}) \gamma \left\{ 1 + \sum_{l \in I_j} s_l^{I_j} s_j^{J_l} - s_i^{J_i} - s_j^{I_j} \right\}. \end{aligned} \quad (87)$$

This yields

$$\xi_{ij}^M = \gamma (1 + \varphi_{ij}) \left(- \sum_{l \in I_j} s_l^{I_j} s_j^{J_l} + s_i^{J_i} + s_j^{I_j} \right). \quad (88)$$

Now consider the following:

(a) Ignoring MR gives the gross elasticity as $\gamma (1 + \varphi_{ij}) > 0$. 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

$$\xi_{ij}^M \geq 0, \quad (89)$$

since $\left(-\sum_{l \in I_j} s_l^{I_j} s_j^{J_l} + s_i^{J_i} + s_j^{I_j}\right) \geq 0$.

(b) (i) If the extensive margin does not operate, the elasticity is

$$\xi_{ij}^M \Big|_{\varphi_{ij}=0} = \gamma \left(-\sum_{l \in I_j} s_l^{I_j} s_j^{J_l} + s_i^{J_i} + s_j^{I_j} \right), \quad (90)$$

such that

$$\frac{\partial \xi_{ij}^M \Big|_{\varphi_{ij}=0}}{\partial s_i} = \gamma (1 - s_j^{J_i}),$$

which is positive if $s_j^{J_i} < 1$. Then

$$\frac{\partial \xi_{ij}^M \Big|_{\varphi_{ij}=0}}{\partial s_i} > 0. \quad (91)$$

(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 $\frac{\partial(\gamma \varphi_{ij})}{\partial s_i} = -\frac{\partial^2 \tilde{w}_{ij}}{\partial d_{ij} \partial s_i} = -\frac{\partial^2 \tilde{w}_{ij}}{\partial d_{ij} \partial Y_i} \frac{\partial Y_i}{\partial s_i} < 0$, this can be seen from

$$\frac{\partial \xi_{ij}^M}{\partial s_i} = \underbrace{-\frac{\partial^2 \tilde{w}_{ij}}{\partial d_{ij} \partial s_i}}_{-ve} \underbrace{\left(-\sum_{l \in I_j} s_l^{I_j} s_j^{J_l} + s_i^{J_i} + s_j^{I_j} \right)}_{+ve \text{ or } -ve} + \underbrace{\gamma (1 + \varphi_{ij}) (1 - s_j^{J_i})}_{+ve}. \quad (92)$$

The MR effect from (b) (i) increases the elasticity for larger countries, but, by Lemma 2, 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, giving

$$\frac{\partial \xi_{ij}^M}{\partial s_i} \leq 0. \quad (93)$$

■

A.9. 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 language²⁸
- colonial ties: a binary variable indicating whether one country ever 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 (% Protestants in j multiplied by % Protestants in i + % Catholics in j multiplied by % Catholics in i + % Muslims in j multiplied by % Muslims in i).
- exporter and importer dummies to capture fixed effects

To construct GDP Shares, GDP data are sourced from the World Bank’s World Development Indicators. 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 (for the former Soviet and Yugoslav Republics for example) were supplemented with data from the United Nations Common Database (UNCDB). Otherwise, a GDP of 0.1 was inputted manually. The denominator is manually 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.10. Description of empirical methodology

Here, we describe the methodology used for estimation and the comparative statics exercises. Our empirics were run on Stata 10.0 MP.

Here, we describe the methodology used for estimation and the comparative statics exercises. Our empirics were run on Stata 10.0 MP.

A.10.1. 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 so that identification does not rely on assumptions of joint normality in the errors in the first and second stage. Theoretically, such a variable should affect the fixed costs of exporting, but not variable costs. For the full sample, HMR use religion for this purpose. Some countries export to

²⁸HMR do not explain the construction of this variable, so it is not clear how this is defined.

everyone else or import from everyone else. The fixed effects are thus perfect predictors and cannot be included in the probit; they therefore do not feature in the subsequent stages.²⁹

2. Generate predicted probabilities ($\hat{\rho}$) of positive exports from country j to i . Many of these are close to unity. Values predicted to exceed 0.9999999 can be indistinguishable from unity in some statistical packages. The Inverse Mills Ratio would be undefined and the terms capturing firm heterogeneity (see point 4) would be unidentified. The approach in HMR converts all $\hat{\rho} > 0.9999999$ to values of $\hat{\rho} = 0.9999999$.
3. Predict the Inverse Mills Ratio $\hat{\eta}_{ij}^* = \frac{\phi(\hat{z}_{ij}^*)}{\Phi(\hat{z}_{ij}^*)}$, where $\phi(\cdot)$ is the standard normal density function and $\Phi(\cdot)$ is the standard normal distribution function. Generate $\hat{x}_{ij} = \hat{z}_{ij}^* + \hat{\eta}_{ij}^*$.
4. Estimate the second stage including the predicted controls for firm selection $\log(e^{\hat{\delta}\hat{x}_{ij}} - 1)$ and country-selection $\hat{\eta}_{ij}^*$. HMR07 use maximum likelihood while HMR08 use non-linear least squares. We follow HMR08. In principle, the default interactive version of the nl command should be sufficient. However, the large number of dummies (over 300) generates an error message. Following Carayol (2006), we use the function evaluator program version of the nl command.
5. Generate predicted values for trade \hat{m} including a predicted value $\hat{\omega} = \log(e^{\hat{\delta}\hat{x}} - 1)$.

A.10.2. 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 $\hat{\eta}_{ij}^*$. 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{x}' , using the same value for $\hat{\eta}_{ij}^*$ but the new predicted probabilities in $\Phi^{-1}(\hat{\rho})$. Use the estimate $\hat{\delta}$ and \hat{x}' to generate an alternative prediction of the non-linear term $\hat{\omega}' = \log(e^{\hat{\delta}\hat{x}'} - 1)$.
5. Generate the alternative predicted values for trade \hat{m}' based on the new distance values and $\hat{\omega}'$.
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 of firms exporting.

²⁹HMR 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.

7. Calculate the net elasticities applying the appropriate country shares adjusting for trading partners and adjustment formulae given in equations () or (). A good approximation can be achieved by applying the multiplier given in () and () and/or using unadjusted country shares.

	1	2	3	4
	OLS	Probit	NL	NL (HMR07)
Dependent variable: Bilateral Exports				
Distance (- γ)	-1.176*** <i>0.0269</i>	-0.660*** <i>0.0206</i>	-0.799*** <i>0.0352</i>	-0.801
border	0.458*** <i>0.118</i>	-0.382*** <i>0.0941</i>	0.832*** <i>0.124</i>	0.831
island	-0.391*** <i>0.107</i>	-0.345*** <i>0.0751</i>	-0.17 <i>0.0972</i>	-0.171
landlocked	-0.561*** <i>0.164</i>	-0.181 <i>0.105</i>	-0.447*** <i>0.128</i>	-0.448
legal	0.486*** <i>0.043</i>	0.0964** <i>0.0297</i>	0.388*** <i>0.0393</i>	0.388
language	0.176** <i>0.0556</i>	0.284*** <i>0.0381</i>	0.0232 <i>0.0593</i>	0.024
colonial ties	1.299*** <i>0.132</i>	0.325 <i>0.292</i>	1.003*** <i>0.0906</i>	1.003
currency union	1.364*** <i>0.216</i>	0.492*** <i>0.132</i>	1.024*** <i>0.269</i>	1.026
FTA	0.759*** <i>0.162</i>	1.985*** <i>0.314</i>	0.378** <i>0.144</i>	0.386
religion	0.102 <i>0.09</i>	0.261*** <i>0.061</i>		
η			0.392*** <i>0.0491</i>	0.399
δ			0.716*** <i>0.0494</i>	0.716
constant	3.168*** <i>0.788</i>	3.813*** <i>0.371</i>	517.2 <i>764.8</i>	
Reference	HMR08:471	HMR07:37	HMR07:37	
Observations	11146	24649	11146	11146
R ²	0.709		0.718	0.709

Table 1: Regression results. Significance at 0.001 & 0.01 levels denoted with *** & **. Standard errors in italics (based on 50 bootstrap replications in column 3).

mean	median	sd	max	min
1.5635	1.4682	0.2886	3.7771	1.2832

Table 2: Summary statistics for elasticity estimates of gross country-level elasticity (ξ^{gross}).

Group		GDP Share		Elasticity		Ratio (ξ^B / ξ^{gross})
		Exporter	Importer	Gross (ξ^{gross})	Bilateral (ξ^B)	
mean		1.269%	1.200%	1.5635	1.5262	0.9734
median		0.186%	0.112%	1.4682	1.4505	0.9930
Mauritania	Togo	0.004%	0.005%	2.0354	2.0352	0.9999
USA	Japan	29.23%	15.74%	1.2832	1.0181	0.7934
Mexico	Spain	1.721%	1.717%	1.2929	1.2489	0.9660

Table 3: Elasticity for a bilateral reduction in trade frictions. Gross elasticity ignores multilateral resistance while bilateral elasticity includes it. The Ratio divides bilateral elasticity by gross elasticity.

Countries		GDP shares			Ratio (ξ^B / ξ^{gross})
Exporter	Importer	Combined	Exporter	Importer	
Bhutan	Japan	15.75%	72.337%	0.001%	0.2636
Eq. Guinea	USA	29.23%	56.101%	42.845%	0.4284
Kiribati	USA	29.23%	46.777%	51.918%	0.5192
Solomon Isl.	USA	29.23%	45.384%	53.577%	0.5358
French Guiana	USA	29.24%	42.757%	56.732%	0.5673

Table 4: Lowest bilateral elasticities. Combined shares do not adjust for trading partners while exporter and importer shares do.

Group		GDP Share		Elasticity		Ratio (ξ^M / ξ^{gross})
		Exporter	Importer	Gross (ξ^{gross})	Multilateral (ξ^M)	
mean		1.269%	1.200%	1.5635	0.0055	0.0111
median		0.186%	0.112%	1.4682	-0.0188	-0.0130
Mauritania	Togo	0.004%	0.005%	2.0354	-0.0937	-0.0460
USA	Japan	29.23%	15.74%	1.2832	0.6569	0.5119
Mexico	Spain	1.721%	1.717%	1.2929	0.0196	0.0151

Table 5: Elasticity for a multilateral reduction in trade frictions. Gross elasticity ignores multilateral resistance while Multilateral elasticity includes it. The Ratio divides multilateral elasticity by gross elasticity.

Elasticities			Ratios	
Firm	Gross $\xi^{gross} = \gamma(1+\phi)$	Net $\xi^M \approx \gamma(1+\phi)(1-MR)$	Extensive-to-intensive margin $\xi^{gross} / \gamma \approx (1+\phi)$	Net-to-gross elasticity $\xi^M / \xi^{gross} \approx (1-MR)$
0.798	1.291	0.467	1.618	0.362

Table 6: Impact of firm heterogeneity and MR on world-wide elasticities. Firm-level elasticity is intensive margin given by gravity model parameter. Gross elasticity includes extensive margin and allows for firm entry. Net elasticity accounts for multilateral resistance captured by price indices. The extensive margin ratio is the amplification effect of allowing for firm entry in comparative statics. The ratio of net-to-gross elasticities summarizes the dampening effect of accounting for changes in multilateral resistance on the world-wide trade elasticity.

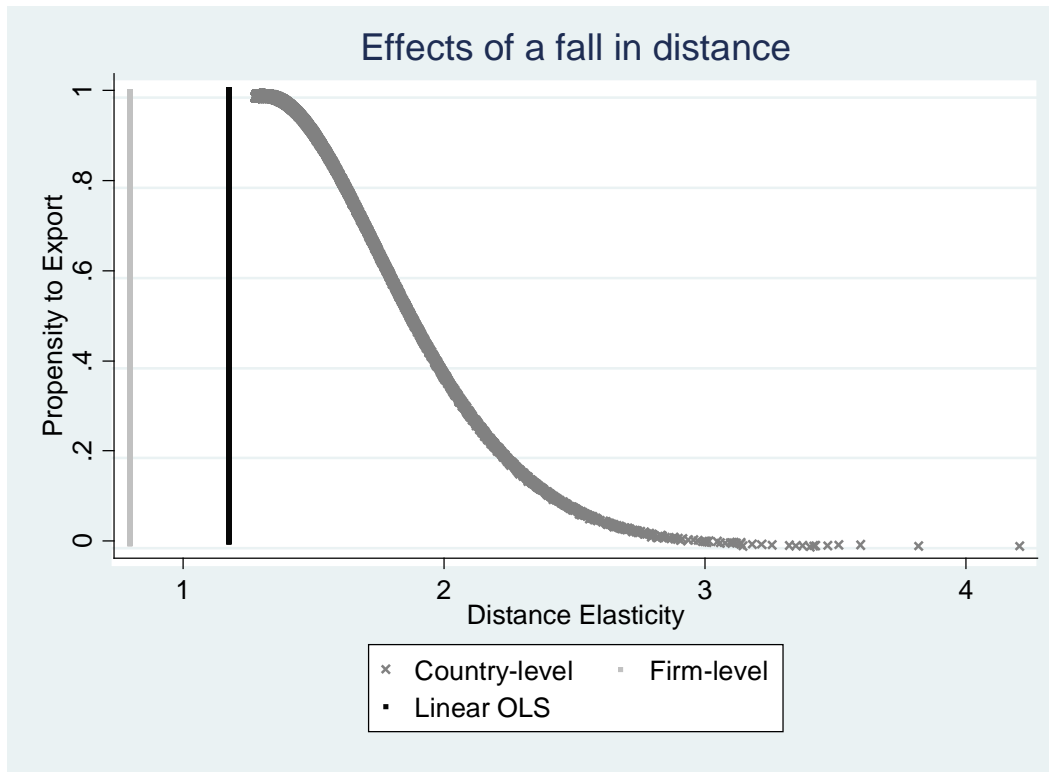


Figure 1: Elasticity estimates for 11,146 country pairs that trade, based on NL estimates. Our focus is the x-axis, which gives the absolute value of the elasticity of trade with respect to distance. 'Linear OLS' estimate gives conventional estimates of the country-level effect (1.176). The other two estimates come from methods accounting for firm heterogeneity. 'Firm-level' elasticities capture the rise in exports at the intensive margin ($\gamma=0.799$). The 'country-level' or gross elasticity (ξ^{gross}) accounts for both the intensive and extensive margins and has a mean of 1.5635 (cf. Table 2).

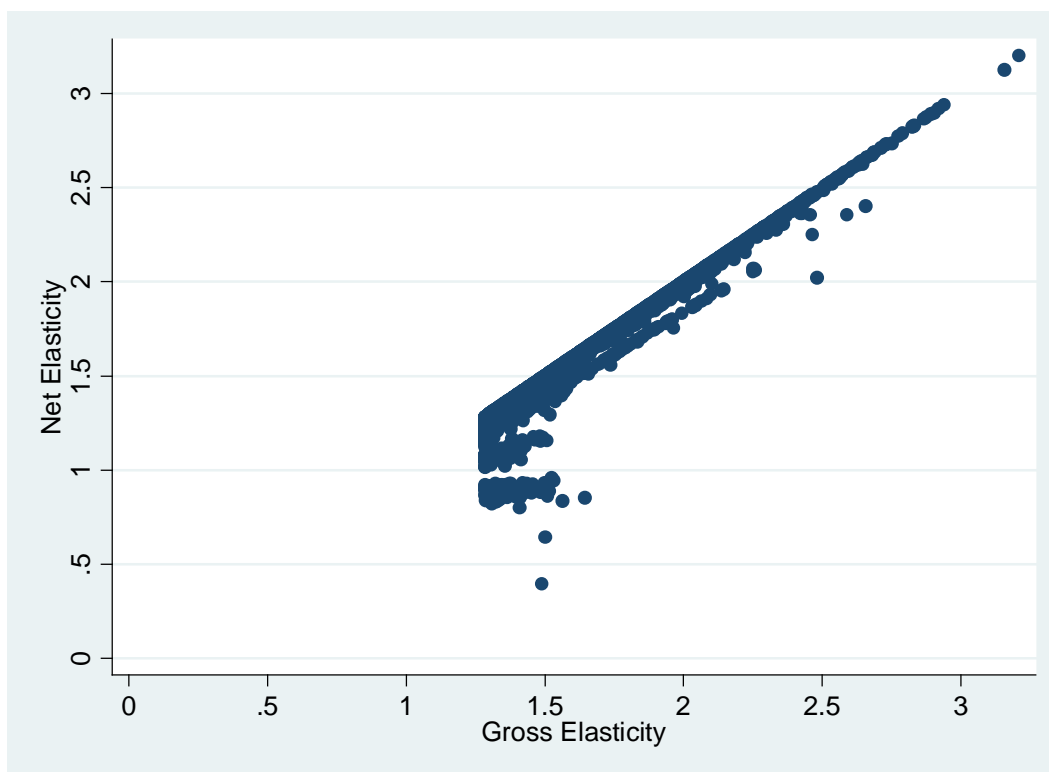


Figure 2: Effects of MR for bilateral changes in distance: x-axis is before accounting for MR (ξ^{gross}); y-axis is after accounting for MR (ξ^{B}).

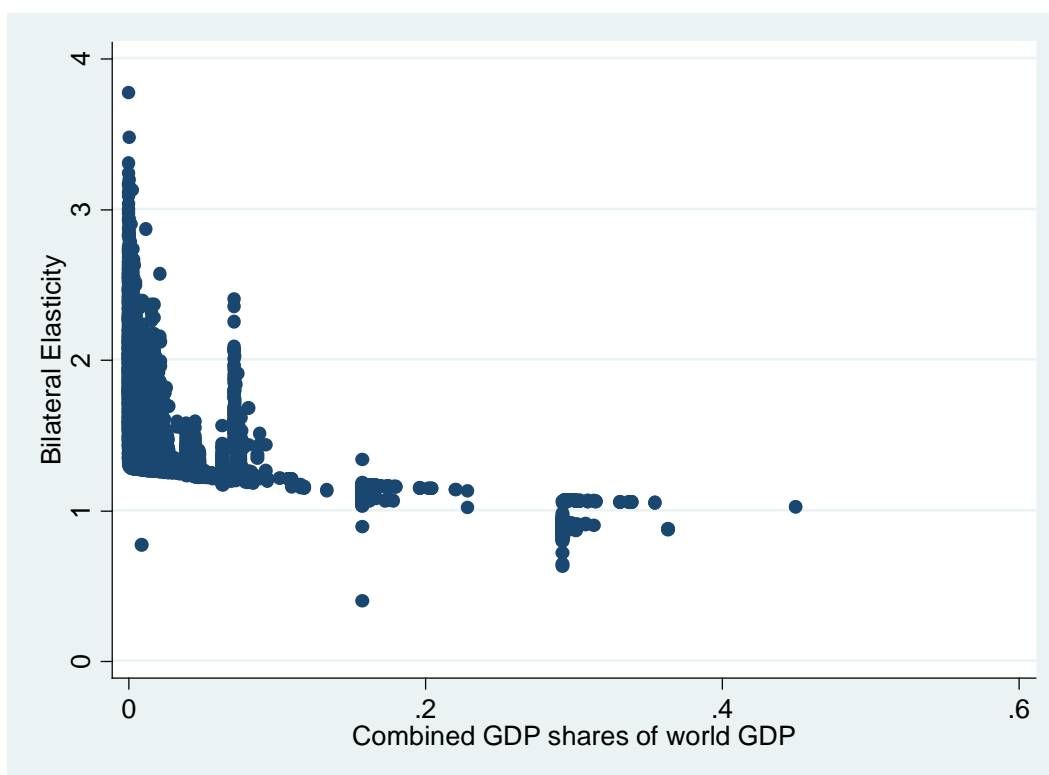


Figure 3: Net elasticity (ξ^{B}) and country-pair size when two countries reduce their frictions

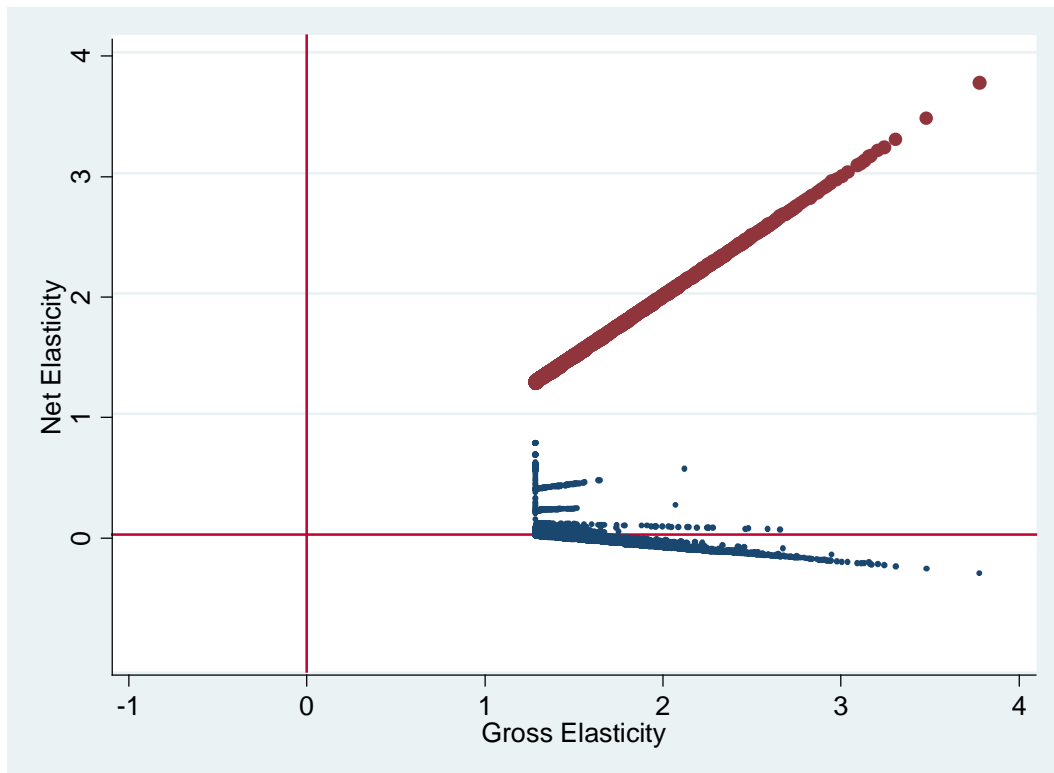


Figure 4: Effects of MR when all countries reduce frictions (line with positive slope is for reference). x-axis is ξ^{gross} ; y-axis is ξ^{M} .

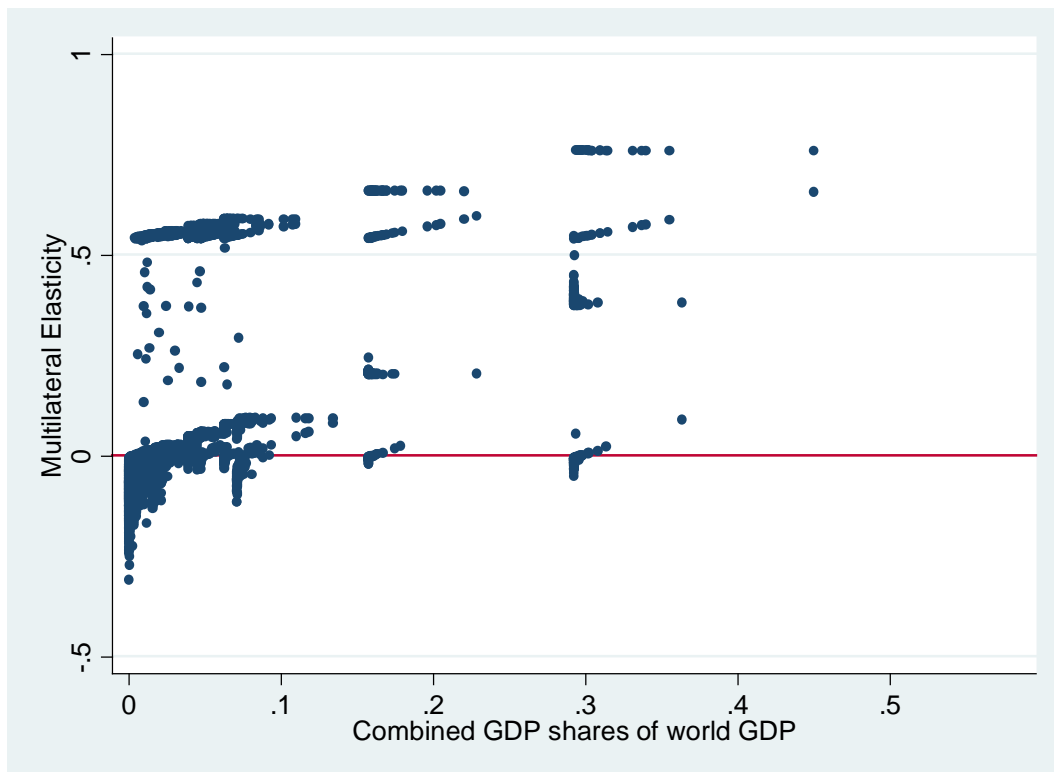


Figure 5: Net elasticity (ξ^{M}) and country size when all countries reduce frictions