Trade Flows, Multilateral Resistance, and Firm Heterogeneity*

Alberto Behar† Benjamin D. Nelson‡

Abstract

Anderson and van Wincoop ('AvW', 2003) showed the importance of 'multilateral resistance' general equilibrium effects in estimating the response of trade flows to trade costs. We integrate this into Helpman, Melitz and Rubinstein's (2008) extension of the AvW framework, which allows for firm heterogeneity, in order to quantify the different margins of adjustment. For bilateral trade cost changes, the general equilibrium effects are small. Surprisingly, most country-pairs reduce their trade after a multilateral fall in trade costs. The global trade response to lower costs is positive but amplified by firm entry and significantly dampened by multilateral resistance.

JEL Classifications: F10, F12, F14, F17

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

*The authors wish to thank Gordon Hanson and two anonymous referees for comments and suggestions which greatly improved this paper. The authors also wish to thank James Anderson, Jeffrey Bergstrand, Elhanan Helpman, Peter Neary and Adrian Wood, as well as participants at multiple seminars, for fruitful discussions. *Disclaimer*: The views expressed herein are those of the authors and should not be attributed to the IMF, its Executive Board, or its management, or to the Bank of England, the Monetary Policy Committee, or the Financial Policy Committee.

†International Monetary Fund. Email: ABehar@imf.org. Part of this work was conducted while at the Oxford Institute for Global Economic Development, Department of Economics, University of Oxford, and funding from the Economic and Social Research Council and the World Bank are gratefully acknowledged.

‡Bank of England. Email: benjamin.nelson@bankofengland.co.uk.

1 Introduction

How do changes in trade frictions affect trade flows? The answer to this question is important for understanding the welfare implications of trade liberalisations. This paper sheds light on this issue by examining comparative statics in a gravity model that integrates two forces, namely multilateral resistance ('MR') and firm heterogeneity, that were introduced in two important papers in the literature.

In the first paper, Anderson and van Wincoop (2003) ('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 demonstrate that traditional gravity equations capture the impact of only bilateral trade costs on trade flows but ignore the fact that regions operate in a multilateral world. As a result, traditional estimates fail to control for theoretically-motivated price terms, which aggregate both domestic and international trade costs and therefore capture MR. AvW show that bilateral trade flows depend on bilateral trade costs *relative to* MR. Failing to account for MR typically leads one to overstate the importance of changes in trade barriers on bilateral trade flows. Since then, further work has studied the general equilibrium interplay between trade costs, trade flows, and income.[1]

The second paper on which we draw is Helpman et al. (2008) ('HMR'). Heterogeneous firm productivity within a country means not all firms are productive enough to cover the fixed costs of exporting.[2] If fixed costs are high enough, no firms in a given country may find it profitable to export to a given destination. Hence, in the presence of fixed costs of trade, 'zeros' naturally arise in the trade data — a *country* selection effect.[3] 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.

However, in performing their counterfactual exercise, HMR abstracted from the general equilibrium effects of trade cost changes, and from the way in which these forces interact with the intensive and extensive margins that those authors distinguish. On the other hand, papers emphasizing MR have neglected the implications of firm heterogeneity. Our contribution computes general equilibrium comparative statics in a gravity model that marries the two strands of the literature and facilitates implementation in large cross-country datasets.

Our work is similar to Egger et al. (2011), who demonstrate biases that arise from neglecting the general equilibrium impact of changes in trade costs in the context of multiple Preferential Trade Agreements. In that study, estimation addresses the issues of endogeneity, country selection and MR, but only the last of these is also accounted for in comparative statics.[4]

We present our hybrid AvW-HMR gravity model in section 2, which applies AvW general equilibrium closure to the extensive and intensive margins of trade. In section 3, we study the theoretical comparative statics in a special case of our hybrid model – a frictionless initial equilibrium – which helps us to illustrate the mechanisms at work in the empirical comparative statics we subsequently compute. Here we show that comparative statics will in general comprise three terms: an intensive margin capturing the firm-level response to trade cost changes, an extensive margin capturing the effects of firm entry, and an adjustment term that captures the effects of MR. We show that changes in trade flows at the extensive margin are also dampened by MR in a general equilibrium setting. These comparative statics are useful for building intuition but are inappropriate for empirical implementation. In particular, the frictionless initial equilibrium is unrealistic and cannot generate the zeros and asymmetries observed in the trade data.

Hence, section 4 develops and empirically implements a Taylor approximation for capturing comparative static MR effects. We adapt Baier and Bergstrand's (2009) approach to a setting that includes firm heterogeneity and accounts for zeros in trade flows, centering the approximation at an initial equilibrium with positive asymmetric trade frictions. Compared to AvW, Baier and Bergstrand ('BB') show that the approximation error is

small for the vast majority of country-pairs and we believe the approach has a number of advantages over computational methods. First, it preserves the possibility of analyzing bilaterally asymmetric trade costs, which can explain asymmetric trade flows. In such settings, Bergstrand, Egger & Larch (2012) show that BB's flexible method is competitive with a computational approach (like AvW's) that assumes bilateral symmetry in trade costs. Second, this method does not require an estimate of the elasticity of substitution between product varieties, which has a large effect on general equilibrium comparative static outcomes (Bergstrand, Egger & Larch, 2012). Third, BB's approach is useful for gaining intuition about the effects at work when trade costs change. We show in a heterogeneous firms setting that multilateral resistance can be represented by three terms that are linear in trade costs, capturing world trade resistance, importer MR, and exporter MR. Having an analytically tractable expression makes the mechanics behind the comparative statics exercises transparent.

After replicating the estimation exercise in HMR, our counterfactual analysis decomposes overall trade flow elasticities into effects operating at the intensive margin, the extensive margin, and through multilateral resistance effects. Section 4 continues by illustrating the relative quantitative importance of each effect by means of three novel comparative static results.

First, we analyze the case of two countries changing their bilateral trade costs. We find that firm heterogeneity will tend to raise trade elasticities, while MR effects are typically small for most country pairs. Firm heterogeneity together with fixed costs of exporting imply the presence of an extensive margin of trade that amplifies country-level trade responses above those found in standard models. MR works against this, but the effects are muted. So, for practical empirical purposes, we find that ignoring MR is relatively innocuous for the majority of country pairs when only two countries change their trade costs in isolation. Furthermore, we show that bilateral liberalizations will tend to generate smaller trade responses at the extensive margin for larger exporting countries, as there is less scope for firm entry to take place for these exporters, compared to smaller exporters, in response to trade liberalisations. This means that smaller countries tend

to enjoy larger trade responses overall when trade costs change bilaterally. The average response indicates that around half the trade flow response operates through the extensive margin of trade.

Second, we turn to multilateral changes in trade costs by analyzing the case where all countries reduce their international trade frictions. Just as for bilateral trade cost reductions, the extensive margin raises trade elasticities, but now the effects of MR are much more important. Because multilateral liberalizations imply small changes in relative trade costs for any given country pair, trade elasticities net of MR are much smaller. Surprisingly, we show analytically that bilateral elasticities can become negative and empirically that this is true for most country pairs. That is, multilateral trade liberalization will tend to redirect output across destinations, reducing exports to some locations but increasing trade with others. The pattern of responses is such that trade between smaller countries tends to become redirected to larger importers. The reason is that, while bilateral trade becomes more attractive, actual trade flows depend on the costs incurred in this trade relative to other destinations. Because bigger importers are less affected by MR, multilateral changes imply many changes in relative prices that tend to favour exporting to larger importers over smaller importers.[5]

Third, we study the response of world trade to multilateral changes in trade costs by aggregating the bilateral elasticities that properly capture the general equilibrium effects of MR. Our main finding here is that the dampening effect of MR dominates the amplification generated at the extensive margin. The elasticity net of general equilibrium considerations is around one third of the size of the aggregate trade elasticity that ignores relative price effects and less than 40% of those implied by standard models. In this empirical application, country entry makes no contribution to the increase in world trade and the role of firm entry is modest. The key results are summarized in Section 5.

2 The model

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

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

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

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

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

Thus solving for unit factor costs and substituting back into the gravity equation yields[7]

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

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

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

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

We define $Y^{I_j} \equiv \sum_{i \in I_j} Y_i$ as the total output of all importers from country j. It differs from AvW's notation only to the extent that fixed costs of exporting mean that country j firms will not export to all destinations. Using this notation, we similarly define $s_i^{I_j} \equiv Y_i / Y^{I_j}$, such that $s_i^{I_j}$ is the output share of country i as a fraction of the total output of all countries that import from j, I_j. We define the set J_i analogously – as the set of all exporters to country i. Following our convention, we then define $s_j^{J_i} \equiv Y_j / Y^{J_i}$ as the output of country j as a share of the total output of all exporters to country i, J_i.[8] Variable trade costs are weakly greater than unity between countries ($t_{ij} \geq 1$ for $i \neq j$) but are exactly unity for trade within countries ($t_{ij} = 1$ for $i = j$). We do not require that variable trade costs are bilaterally symmetric, which is important for capturing asymmetric trade flow patterns. Finally, the selection term V_{ij} would reduce to a constant in the absence of fixed costs of trade. In the presence of fixed costs, variation in the extensive margin of trade will contribute directly to variation in aggregate bilateral trade flows M_{ij} and indirectly because the MR terms are also a function of the extensive margin.

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

form, namely:

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

where f_{ij} are the fixed costs of exporting from country j to country i. Hence a consequence of requiring general equilibrium closure is that the extensive margin can be expressed in a convenient form akin to equation (1) for bilateral exports. Intuitively, higher fixed costs of trade reduce the cost cut-off, so fewer firms in j export to i. In addition, the cost level above which firms in j cannot export to i rises as the product of the two countries' outputs rises, as bilateral trade costs fall, or as MR rises. The fact that MR enters equation (4) just as it does for equation (1) highlights the fact that valid comparative statics on the extensive margin require accounting for general equilibrium effects, just as AvW emphasize for valid comparative statics at the aggregate level.

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

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

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

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

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

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

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

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

where $\delta \equiv (k-\sigma+1)/(\sigma-1)$. Equations (1) and (7) then make clear that the comparative static effects of trade costs will operate through three channels. First, changes in variable trade costs will affect bilateral trade through the intensive margin, captured by the t_{ij} term in (1). Second, to the extent that changes in trade costs also affect the cost of trading with alternative export destinations, the multilateral resistances in (1) will change, tending to oppose the direct effects of changes in trade costs through t_{ij}. Third, changes in variable trade costs will also affect the extensive margin of trade through V_{ij}. This final effect is also subject to competing forces. It comprises both direct effects of changes in trade costs through the t_{ij} term in the numerator of equation (4), which by (6) will affect the extensive margin in (7), together with indirect effects through the MR terms that enter (4). As in the aggregate trade flow equation (1), the MR effects operating through the extensive margin will tend to oppose the direct effects of trade costs through the extensive margin, providing some dampening effect. Overall, then, valid comparative statics must account for these two sources of effects of MR on trade flows.

The log of equation (1) gives the log-gravity equation:

$$m_{ij} = y_i + y_j - y^{I_j} - (\sigma - 1)\ln t_{ij} + w_{ij} + (\sigma - 1)\ln P_i \widehat{P}_j, \tag{8}$$

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

$$w_{ij} \equiv \ln\left(\exp \delta z_{ij} - 1\right), \quad z_{ij} \equiv \widetilde{z}_{ij} + \ln\left(P_i \widehat{P}_j\right)^{\sigma-1}. \tag{9}$$

We denote the extensive margin elasticity by $\varphi_{ij} \equiv \partial w_{ij}/\partial z_{ij} = \delta e^{\delta z_{ij}}/(e^{\delta z_{ij}} - 1)$, which captures the effects of changes in factors determining the latent variable z_{ij} on the proportion of firms exporting. Since we will be interested in the effects of changes in trade costs on countries of different sizes, it is useful to state the following result:

Lemma 1 *For given multilateral resistance, the derivative of the bilateral elasticity of trade at the extensive margin φ_{ij} with respect to country size s_h, $h = i, j$, satisfies*

$$\frac{\partial \varphi_{ij}}{\partial s_h} < 0, \quad h = i, j.$$

such that larger traders have smaller elasticities of bilateral trade at the extensive margin with respect to trade costs.

Proof. See appendix. ■

This result highlights one source of heterogeneity across countries' responses to changes in trade frictions, namely, the extensive margin of trade. The smaller the exporter, the greater the extent to which firm entry is encouraged following a trade liberalisation. Intuitively, smaller, more remote exporters have more scope for expanding the number of exporting firms than large, well-connected traders. So the effect is more muted for larger exporters. The extensive margin is not the only source of heterogeneity in countries' responses to changes in trade costs, however. As we show next, multilateral resistance also materially affects the response of trade flows to trade liberalisations.

3 Comparative statics in a frictionless world

Before we turn to the empirical implementation of our model, we study its properties in a simplified setting – around a frictionless initial equilibrium. At the end of the last section, we showed how country size drives heterogeneity in trade responses at the extensive margin. In this section, we illustrate how country size drives heterogeneous responses of trade flows through multilateral resistance. The frictionless case is useful to illustrate these general equilibrium effects in a way that abstracts from heterogeneity at the extensive margin.

The thought experiment we perform is identical to one considered by AvW in their framework, which abstracts from firm heterogeneity. To draw a comparison with their results, consider a frictionless world in which variable trade costs $t_{ij} = 1$ for all i, j and in which all countries trade such that $I_j = J_i$. The support of a is such that all firms in country i export and that $V_{ij} = 1$ for all i, j at the initial equilibrium. In this world, importer and exporter multilateral resistances are symmetric and equal to unity, as in AvW's example. Using equations (4) and (5) in the equilibrium we describe, a change in

trade costs implies

$$\frac{dV_{ij}}{dt_{ij}} = -(\sigma - 1)\varphi \left(1 - \frac{dP_i}{dt_{ij}} - \frac{d\widehat{P}_j}{dt_{ij}}\right), \tag{10}$$

A rise in bilateral variable trade costs would reduce the number of firms exporting. But these effects are dampened to the extent that multilateral trade costs also rise, as captured by the third term. As is the case in standard gravity models at the intensive margin, a higher elasticity of substitution between products implies a large impact of trade costs on trade flows. In the case of equation (10) however, the effect here is at the extensive margin, or the number of firms exporting. This standard effect is multiplied by $\varphi = \delta(1 + \overline{k})$, which reflects the degree of heterogeneity across firms. When this is large, the extent of firm entry (or exit) following trade cost changes is also large. Using (10) in (1), a general expression for the effect of a change in trade costs on trade flows evaluated at the initial frictionless equilibrium is then given by:

$$\frac{d}{dt_{ij}}\left(M_{ij}\frac{Y}{Y_i Y_i}\right) = -(\sigma - 1)\,(1 + \varphi)\left(1 - \frac{dP_i}{dt_{ij}} - \frac{d\widehat{P}_j}{dt_{ij}}\right), \tag{11}$$

which we can use to understand the components of trade flow comparative statics. Just as the expression for comparative statics at the extensive margin in (10), equation (11) shows that comparative statics on overall trade flows between exporter j and importer i will comprise an intensive margin captured by $\sigma - 1$, some amplification due to firm exit at the extensive margin due to the factor $1 + \varphi > 1$, together with some dampening due to multilateral resistance $1 - dP_i/dt_{ij} - d\widehat{P}_j/dt_{ij}$. The last effect captures the general equilibrium repercussions of changes in trade frictions on average importer resistance and average exporter resistance. On the importer side, trade flows are dampened to the extent that cheaper varieties may now be available from elsewhere (i.e. from exporters other than j), while on the exporter side, dampening may occur because alternative export destinations may now be relatively more attractive (i.e. importers other than i).[9]

We can gain further insight by continuing to explore the MR adjustment term. Using equation (10), we totally differentiate the system of price indices given by (2) and (3)

with respect to variable trade costs. Evaluating at the frictionless equilibrium yields:

$$dP_i = \sum_j s_j dt_{ij} - \sum_j s_j d\widehat{P}_j, \tag{12}$$

$$d\widehat{P}_j = \sum_i s_i dt_{ij} - \sum_i s_i dP_i. \tag{13}$$

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

$$dP_i + d\widehat{P}_j = -\sum_l \sum_h s_l s_h dt_{lh} + \sum_l s_l dt_{lj} + \sum_h s_h dt_{ih}. \tag{14}$$

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

Consider first bilateral changes in trade costs. For a bilateral change, let $dt_{lh} = dt_{hl} = dt$ for $l, h = i, j$, or else let $dt_{lh} = 0$. Then using (11), GDP-weighted exports change according to

$$\frac{d}{dt}\left(M_{ij}\frac{Y}{Y_i Y_i}\right) = -(\sigma - 1)(1 + \varphi)\left(1 + 2s_i s_j - s_i - s_j\right). \tag{15}$$

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

Consider next the effect of a multilateral change in trade costs, as in AvW, such that

$dt_{ij} = dt > 0$ for all $i \neq j$, and $dt_{ii} = 0$ for all i. Using equation (11) and our expressions for the total differentials in (10) and (14), GDP-weighted exports change according to

$$\frac{d}{dt}\left(M_{ij}\frac{Y}{Y_i Y_i}\right) = -(\sigma - 1)(1 + \varphi)\left(-\sum_l s_l^2 + s_j + s_i\right). \tag{16}$$

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

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

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

trade when trade elasticities are generated, rather than all countries in the world.

4 Empirical comparative statics

For estimation purposes, the methodology developed in HMR deals with firm heterogeneity and country fixed effects control for importer and exporter multilateral resistance.[10] We replicate HMR's approach to estimation. For comparative statics however, we adapt Baier and Bergstrand's ('BB', 2009) 'Bonus vetus OLS' (BVOLS) approach method for approximating MR effects to a heterogeneous firms setting. In an homogeneous firms setting, BB accomplish this by taking a Taylor approximation for the non-linear MR terms around a world of symmetric but positive trade frictions. Their method allows us to expunge the endogenous components from the right hand side of the MR terms, such that approximate comparative statics are straightforward to implement.

AvW use a computational procedure to solve for MR explicitly and to compute comparative statics. Despite awareness of this procedure and the first-order consequences of MR, the approach of AvW has not been applied to cross-country data sets typically used in empirical applications of the gravity model.[11] This is largely due to the difficulty in implementation, which requires a customized program. Our application would be especially complex because the system we study includes asymmetric bilateral trade costs, firm selection, and different sets of active traders for different trading country pairs.

Like BB, our primary motive is to present a practical approach to computing comparative statics that has low barriers to implementation in empirical gravity work. It successfully corrects for the first-order inaccuracies associated with the general equilibrium effects of trade cost changes. Baier and Bergstrand (2009) show that their approximation is good for the majority of region pairs that they consider. In particular, comparative statics of economic integration based on the Taylor method are within 10% of the AvW approach for 83% of their pairs. Bergstrand et al. (2012) provide Monte Carlo evidence that the trade-flow comparative statics computed using a Taylor approach are close to those generated using nonlinear solvers. This is despite using a version of BVOLS for

estimation that, while better for comparative statics, starts with the handicap of being biased at the estimation stage. Furthermore, the accuracy of comparative statics improves as the number of countries increases. Therefore, using fixed effects rather than standard OLS with MR terms in estimation while using BVOLS for comparative statics in a large cross section of countries, as we do, would yield even better performance.

Further, the approach does not require us to assume particular values for structural parameters like the elasticity of substitution, σ. Bergstrand et al. (2012) show that comparative statics are very sensitive to the choice of this elasticity parameter and that the AvW approach is inaccurate when the assumed elasticity is different from the actual elasticity.[12] A final advantage of the BB method is its intuitive appeal. This, combined with the frictionless case just considered, allows us to be clearer about the mechanisms that lie behind the comparative statics we produce.[13]

To apply BB's method to the case of firm heterogeneity, we make the following assumption for the purposes of comparative statics:[14]

Decomposability: *The extensive margin terms V_{ij} entering the indices of multilateral resistance $\{P_i, \widehat{P}_j\}$ are approximately $V_{ij} \simeq Z_{ij}^{\delta} \left(P_i \widehat{P}_j\right)^{\delta(\sigma-1)}$.*[15]

Using this, a Taylor expansion of importer and exporter multilateral resistances yields a tractable expression for trade resistance that comprises three terms:

$$(\delta+1)\ln\left(P_i\widehat{P}_j\right)^{\sigma-1} \simeq -\overbrace{\sum_{l\in I_j} s_l^{I_j} \sum_{h\in J_i} s_h^{J_i}\left[(\sigma-1)\ln t_{lh} - \delta\widetilde{z}_{lh}\right]}^{\text{World Trade Resistance}} + \underbrace{\sum_{h\in J_i} s_h^{J_i}\left[(\sigma-1)\ln t_{ih} - \delta\widetilde{z}_{ih}\right]}_{\text{Importer's MR}} + \underbrace{\sum_{l\in I_j} s_l^{I_j}\left[(\sigma-1)\ln t_{lj} - \delta\widetilde{z}_{lj}\right]}_{\text{Exporter's MR}}, \quad (17)$$

in which $\widetilde{z}_{ij} \equiv \ln \widetilde{Z}_{ij}$. This expression has three terms which, by analogy with (14), can be straightforwardly interpreted. First, to the extent that world trade resistance is high, MR between the i-j trading pair will be low, tending to discourage international trade on average. Second, to the extent that importer i faces high trade costs in obtaining output from other exporters in the set J_i, exporter j will export more to i. Third, when exporter

j finds it costly to export to other destinations in the set I_j, it will tend to export more to i instead, all else equal. Hence the expression above has a straightforward interpretation that makes clear the impact of *relative* trade costs on bilateral trade flows.

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

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

We turn instead to the implementation of our comparative statics, for which we need three things. First, we need an estimate of the 'firm level' intensive margin elasticity of trade flows to trade costs. We proxy trade costs with a number of observable variables and, for the elasticity, we take an estimate of the coefficient on log distance, γ. Our analytical results will be in terms of changes in variable trade costs, assuming fixed costs stay constant, and we will interpret the empirical illustration in this way. However, observable proxies included in the probit stage could affect both fixed and variable costs.

The empirical illustrations, which are computed for changes in distance, could therefore reflect changes in both costs. The estimates for the trade cost elasticities are $\hat{\gamma}_p = 0.66$ in the probit model and $\hat{\gamma} = 0.799$ in the 2nd stage regression.[16] Second, we need country-pair extensive margin elasticities $\varphi_{ij} = \delta e^{\delta z_{ij}}/(e^{\delta z_{ij}} - 1)$, the estimate of which we denote by $\widehat{\varphi}_{ij}$. In turn, these terms contain estimates of z_{ij}, a linear function of the the variables in the probit model, together with $\hat{\delta} = 0.72$. The φ_{ij} terms clearly vary by country pair and they provide one source of cross country variation in trade elasticities. Finally, to make the adjustments for MR, we need country GDP shares defined relative to the relevant sets of active traders, $\{s_i^{J_i}, s_j^{I_j}\}$ for all i, j. These are easily generated from the matrix of observed trade flows.

4.1 Gross elasticities

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

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

where $\xi_{ij}^{gross} \equiv -\partial m_{ij}/\partial t_{ij}|_{P_i, \widehat{P}_j}$. For the empirical analogue of this elasticity based on a 10% fall in distance, $\widehat{\xi}_{ij}^{gross} = \widehat{\gamma}(1+\widehat{\varphi}_{ij})$, we replicate the results from HMR in Table 1. For reference, the first row includes the linear estimate, while the second row produces results from the NLS estimate derived from the Pareto assumption made in section 2. The mean overall elasticity of 1.564 implies the bilateral country-level effect is due roughly in equal parts to the intensive and the extensive margins. The estimate is higher than would be implied by the linear OLS estimate or by estimates attempting to deal with country selection alone (1.21 – see HMR). This is important because it implies one should allow for firm heterogeneity even if one is not interested in the decomposition of the overall gross effect into different margins of adjustment.[17]

[Table 1 about here]

Cross country variation in the extensive margin elasticities drives variation in the estimated gross elasticities, which have a standard deviation of 0.289. The variation is generated from variations in $\widehat{\bar{z}}_{ij}^{*}$, which is monotonically related to the proportion of exporting firms. Small country pairs that are far apart will have a low probability of trading, which implies a low proportion of exporting firms and hence plenty of scope for firm entry after a reduction in trade costs, as shown earlier in Lemma 1.

This result is confirmed empirically in Table 1, where the correlation between ξ_{ij}^{gross} and $s_i + s_j$ is -0.286. One can analogously show that $\partial \xi_{ij}^{gross} / \partial \ln t_{ij} > 0$. For example, the correlation between the elasticity and log distance is 0.049. In addition to NLS estimates derived from the Pareto assumption, HMR also produce estimates derived from a polynomial function of z_{ij}. We include these estimates to show that the implications for bilateral elasticities are similar, albeit with a larger standard deviation.

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

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

pair to stop trading. The results imply that we do not need to consider changes in the sets of traders when studying the net elasticities for the baseline 10% change in trade costs that we do next.

4.2 Bilateral changes in trade costs

Allowing trade costs to change bilaterally results in the following:

Implication 1 *(Bilateral changes in trade costs) With Taylor approximated multilateral resistance terms, the elasticity of bilateral trade flows to bilateral (B) changes in variable trade costs is given by*

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

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

Proof. See appendix. ■

This is the empirical counterpart of the frictionless (15), which we implement by employing:

$$\xi_{ij}^{B} = \gamma + \varphi_{ij}\gamma_p - (1 + \varphi_{ij})\frac{\gamma + \delta\gamma_p}{1 + \delta}\left(-s_i^{I_j} s_j^{J_i} - s_j^{I_j} s_i^{J_i} + s_j^{J_i} + s_i^{I_j}\right), \tag{20}$$

using our estimated parameters. Table 2 shows this, together with an entry demonstrating MR in the linear case, which disregards firm heterogeneity, and an entry for the gross elasticity from section 4.1, which disregards MR. Because most countries are small, bilateral changes in trade costs have small MR implications. In the linear case, on average, the dampening effect of MR is only -0.033, so the average bilateral elasticity is close to that implied by the linear estimate of γ. If we allow for firm heterogeneity but no MR, the comparative static effect is substantially larger. The average MR effect is also small when we include firm heterogeneity, so net ($\widehat{\xi}_{ij}^{B}$) and gross ($\widehat{\xi}_{ij}^{gross}$) elasticities are close

together. As a result, the average amplifying effect of accounting for firm heterogeneity is much stronger than the dampening effect of MR when trade cost changes are bilateral.

[Table 2 about here]

It is intuitive for the average MR effect to be small since the average country is small. Since the essence of MR is to account for GDP-weighted changes in relative trade costs, bilateral changes between two typical countries have small general equilibrium effects. However, there are some exceptions. Table 3 illustrates this with specific country examples and includes the ratio $\widehat{\xi}_{ij}^{B}/\widehat{\xi}_{ij}^{gross} \approx 1 + s_i^{I_j} s_j^{J_i} + s_j^{I_j} s_i^{J_i} - s_j^{J_i} - s_i^{I_j}$. Tiny pairs like Mauritania and Togo have $\widehat{\xi}_{ij}^{B}/\widehat{\xi}_{ij}^{gross} \simeq 0.9999$. By contrast, Japan and the USA comprised 45% of world GDP in the 1986 data and generate a ratio of 0.79. This suggests that MR effects are not insignificant for all countries, even for bilateral changes in trade costs. Even though Mexico and Spain were the 10th and 11th biggest countries in the world, they each had less than 2% of world GDP, such that the $\widehat{\xi}_{ij}^{B}/\widehat{\xi}_{ij}^{gross}$ ratio for this pair is still sufficiently close to unity to suggest it is only "very" big countries for which MR matters for bilateral liberalizations.

[Table 3 about here]

However, our adjusting for actual trading partners implies lower net-to-gross ratios are also generated by small (and remote) exporters if they have a dominant importer. An extreme example in Table 3 is Japan, which makes up almost three-quarters of importer GDP from Bhutan, so that the net-to-gross ratio is only 0.26 for exports from the latter to the former. The table also contains some other examples. Thus, for bilateral changes in trade costs, MR effects are material for (a) trade between the world's largest country pairs and (b) exports from small exporters with few export destinations to the world's largest countries. But given the skewness in the distribution of country incomes, most country pairs are small, and are therefore not materially affected by MR.

Since both the MR effect and the elasticity at the extensive margin are decreasing in country size, the impact of size on the trade elasticities is unambiguous in the case of

bilateral changes in trade costs. Small country pairs will tend to experience larger elasticities, as demonstrated in Implication 1. Figure 1 illustrates this negative relationship in terms of country GDP shares $s_i + s_j$ and the correlation of -0.506 is stronger than was the case for $\widehat{\xi}_{ij}^{gross}$.

[Figure 1 about here]

4.3 Multilateral changes in trade costs

When all countries reduce trade costs, we find that:

Implication 2 *(Multilateral changes in trade costs) With Taylor approximated multilateral resistance terms, the elasticity of bilateral trade flows to multilateral changes (M) in variable trade costs is given by*

$$\xi_{ij}^{M} = (\sigma - 1)(1 + \varphi_{ij}) \left(-\sum_{l \in I_j} s_l^{I_j} s_l^{J_i} + s_i^{J_i} + s_j^{I_j} \right) \tag{21}$$

where $\xi_{ij}^{M} \equiv -\sum_l \sum_h dm_{ij}/d\ln t_{lh}$*, in which* $d\ln t_{lh} = d\ln t$ *for all* $l \neq h$ *and* $d\ln t_{lh} = 0$ *for* $l = h$*. Then:*

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

$$\xi_{ij}^{M} \gtrless 0, \tag{22}$$

such that the sign of the elasticity of bilateral trade is ambiguous;

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

$$\frac{\partial \xi_{ij}^{M} \big|_{\varphi_{ij}=0}}{\partial s_h} > 0, \quad h = i, j \tag{23}$$

when $s_i^{I_j} + s_i^{J_i} < 1$ *if* $h = i$*, and when* $s_j^{I_j} + s_j^{J_i} < 1$ *if* $h = j$*, such that larger countries larger firm-level responses to multilateral trade liberalizations;*

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

$$\frac{\partial \xi_{ij}^{M}}{\partial s_h} \gtrless 0, \quad h = i, j \tag{24}$$

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

Proof. See appendix. ■

The expression in (21) is the empirical counterpart to (16) in the frictionless case. In contrast to the case of bilateral changes in trade costs, the elasticity of bilateral trade flows with respect to multilateral changes in trade costs can be either positive or negative in theory. If it is positive, it is likely to be far below the gross elasticity that one would compute ignoring MR. The theoretical origin of this result lies in the endowment economy studied here and in AvW. Changes in trade costs serve to reallocate output from one destination to another so that multilateral trade cost changes can mean that exports become redirected from one importer to another even if the bilateral friction between all country pairs falls. A fall in bilateral trade can be explained by the overall bilateral cost rising *relative* to the cost of importing from or exporting to alternative destinations.[18]

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

Part (b) (ii) of the Implication arises because Lemma 1 acts against part (b) (i) of the Implication when the extensive margin is in operation. Larger countries have smaller gross elasticities due to the extensive margin being less responsive, but also have smaller MR effects, so the theoretical relationship between country size and elasticity breaks

down. Therefore, the sign of the elasticity and its correlation with country size is an empirical matter, which requires implementation using

$$\xi_{ij}^{M} = \gamma + \varphi_{ij}\gamma_p - (1+\varphi_{ij})\frac{\gamma+\delta\gamma_p}{1+\delta}\left(1+\sum_{l\in I_j} s_l^{I_j} s_l^{J_i} - s_i^{J_i} - s_j^{I_j}\right). \tag{25}$$

Table 4 demonstrates the dramatic effects of accounting for MR when *all* countries reduce distance. Because international frictions have fallen relative only to domestic frictions, one might expect an average effect close to zero. Indeed, accounting for MR almost completely cancels the trade effect calculated in the absence of MR – whether in the linear case or when combined with firm heterogeneity. For example, the average net elasticity $\widehat{\xi}_{ij}^{M}$ is only 0.006, which is substantially lower than the average gross elasticity $\widehat{\xi}_{ij}^{gross}$. Moreover, there are only about 2,300 positive elasticities out of approximately 11,000 trade flows, so Implication 2 (a) is relevant empirically.[19]

[Table 4 about here]

The theoretical analysis introduced the opposing forces in the relationship between country size and the elasticity: bigger pairs have bigger MR multipliers but lower gross elasticities due to the extensive margin. For multilateral changes, the correlation between $\widehat{\xi}_{ij}^{M}$ and $s_i + s_j$ is 0.490. This positive correlation, which is illustrated in Figure 2, implies that the effect due to MR is stronger than that due to the extensive margin in the multilateral case.

[Figure 2 about here]

The general equilibrium effects we are describing are so strong that most country pairs reduce their bilateral trade after reductions in their frictions and trade is redirected elsewhere. 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, each exporter must have *at least one* import destination with which it has a positive export response. Empirically, this happens for every

exporter in our sample, but G7 importers are responsible for half of the positive elasticities. Putting it starkly, most countries trade a lot more with the G7 and less with almost everyone else. After a reduction in trade frictions, they reduce their exports to most smaller countries to be able to expand their exports to a handful of big countries.

Moving beyond bilateral trade responses, we aggregate the bilateral elasticities by exporter, weighting each bilateral elasticity by the volume of exports. As should be the case after aggregating across all export destinations, all countries see an aggregate increase in exports even after accounting for MR. The mean aggregate elasticity is only 0.29, while one-third of our countries have elasticities of less than 0.1, a small fraction of the elasticity implied by OLS. Researchers who are interested in the comparative static impacts of trade costs for a particular country are being particularly mislead by estimates based on gravity model coefficients that ignore general equilibrium.

Some studies use gravity models to simulate the effects of multilateral reductions in trade barriers on global trade (eg Wilson et al. (2005)). It is interesting to calculate similar world-wide elasticities, which we report in Table 5. Ignoring both MR and firm heterogeneity, the implied worldwide impact would be the linear estimate of γ reported in the first row of the table. Using the same estimates but allowing for MR yields a much lower global impact, as shown in the second row. In the third row, we allow for the extensive margin of trade to operate, such that aggregating across all bilateral elasticities yields a world elasticity $\widehat{\xi}^{gross}$ of 1.291. This is close to the minimum bilateral elasticity presented in Table 1 because the world's biggest economies, which have a large weight in the aggregate elasticity, tend to have the lowest individual elasticities. For global elasticities (unlike bilateral elasticities), the intensive margin accounts for over 60% of the trade increase and the gross elasticity is about 10% higher than that implied by OLS. Recall that there was no effect on country entry after a 10% fall in distance.

[Table 5 about here]

Once we account for MR effects, however, world-wide elasticities are much lower. As reported in the forth row of Table 5, aggregation yields a world-wide net trade elasticity $\widehat{\xi}^{M}$ of 0.467, which is barely a third of the gross elasticity, and less than 40% of the

elasticity implied by OLS. We draw the same conclusions from an analogous empirical implementation using instead the polynomial estimates. Therefore, existing gravity-based simulations of the global trade implications of multilateral reductions in trade barriers are seriously misleading.

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

5 Concluding remarks

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

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

can ignore MR for practical purposes.

By contrast, large effects due to MR emerge after multilateral changes in trade costs, for example a comprehensive new multilateral trade agreement, are considered. The theoretical sign of the bilateral trade response is ambiguous and we find empirically that most bilateral trade elasticities are negative. Because large countries are less affected by MR, exports get reoriented from small importers to large importers.[20]

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

Notes

[1] Anderson and Yotov (2011) perform an analysis of the effects on the terms of trade – and hence real income – of changes in trade costs arising from free trade agreements. Behrens et al. (2009) show the endogenous responses of wages and productivity, due to Melitz (2003), matter for counterfactual trade flow analysis. They find that the estimated counterfactual border effect is around half the size of the 'pure' border effect ignoring these changes.

[2]See Bernard et al. (2007) for an overview of firms in international trade.

[3]High fixed costs of exporting from country j to country i do not imply 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.

[4]The authors consider regressions that aim to control for but do not separately identify firm heterogeneity.

[5]This is related to the 'constructed home bias' – or the disproportionate share of domestic trade – of Anderson and Yotov (2010). Because a larger share of large countries' trade is domestic rather than international, changes in trade costs have a smaller effect on large countries' price indices. This home bias effect is distinct from the 'home market effect' – whereby changes in trade costs shift the location of production – as studied in, inter alia, Krugman (1980), Helpman and Krugman (1985), Davis (1998) and Hanson and Xiang (2004).

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

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

[8]Note that, as long as the sets of active traders $\{J_i, I_j\}$ are fixed – as for the comparative statics we consider (see below) – GDP shares will also be constant. We therefore abstract from effects that trade cost changes might have on aggregate output shares, $\{s_j^{J_i}, s_i^{I_j}\}$. See Baier and Bergstrand (2009) and AvW for similar set-ups. Anderson and Yotov (2011) compute terms of trade effects arising from trade liberalisation, which would have implications for real income and therefore world GDP shares to the extent that these effects are asymmetric.

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

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

[11]To our knowledge, the most ambitious exercise is that of Egger et al. (2011), who allow for MR comparative statics and other factors on a dataset of 128 countries. Other contributions are still limited to modeling US and Canadian regions together with a limited set of additional countries or an aggregate for the rest of the world (AvW; Anderson and Yotov (2010); Anderson and Yotov (2011); Behrens et al. (2009)). All these studies conduct counterfactuals on binary variables like borders or free trade agreements and not the broader class of continuous trade cost proxies. Behar et al. (forthcoming) apply the methods developed here to the case of logistics.

[12]Hertel et al. (2007) discuss the wide range of σ estimates and the sensitivity of general equilibirum comparative statics to the value of σ.

[13]Note that, following Baier & Bergstrand (2009) and most dynamic macroeconomic models, we will consider a first-order Taylor approximation to the multilateral resistance terms. A first-order approximation will allow us to obtain analytical solutions but rules out interactions between distance and other trade costs in the MR terms.

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

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

we allow for $\phi_{ij} \neq \phi_{ji}$ around our centre. In our case, this implies it is possible that $\widetilde{Z}_{ij} \neq \widetilde{Z}_{ji}$.

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

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

[18]In parallel, the worldwide reduction in trade frictions is associated with a rise in factor costs, which reduce demand. To see the interplay between MR and factor costs, consider what happens to trade between i and j when trade barriers between j and a third country $h \neq i$ fall, but the bilateral barrier between i and j remains unchanged. This manifests itself as a fall in j's exporter MR. At the same time, higher demand from the third country bids up j's factor costs. The higher factor costs make j's goods more expensive in i and, given the barriers between i and j remain constant, this lowers demand from i. So, the fall in exporter MR makes supplying goods to i less attractive in relative terms, while the rise in factor costs reduces demand by i. This general equilibrium channel can be strong enough to reduce bilateral trade between i and j even when the i-j trade barrier is not constant and actually falls along with other 'third country' trade frictions, as in a multilateral liberalisation.

[19]Similarly, Egger et al (2011) find that the signing of multiple preferential trade agreements leads many signatories to experience some falls in bilateral trade as a result of third country effects. This illustrates how the general equilibrium concepts in this literature are related to the traditional notion of trade diversion Viner (1950). Moreover, the structural gravity model provides a very explicit interpretation of trade diversion, and one can show analytically that trade between the signatory to an agreement and all third parties falls. Furthermore, Behar and Criville (2011) show how country size influences the effect of PTAs through MR.

[20]In a counterfactual analysis that lies in between the bilateral and multilateral cases

studied here, Behar et al. (forthcoming) study the effects of a reduction in exporter-specific trade costs. A unilateral improvement in logistics quality increases exports by more in larger countries, chiefly because they are less affected by multilateral resistance.

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

Functional Form	*Mean*	*Median*	*Std. dev.*	*Maximum*	*Minimum*	Correlation with size
Linear	1.176	1.176	-	-	-	-
Pareto	1.564	1.468	0.289	3.777	1.283	-0.286
Polynomial	1.853	1.846	0.518	2.995	1.141	-0.348

Table 2: Decomposition of trade flow elasticity for a bilateral reduction in trade frictions.

	Intensive margin	*Mean extensive margin*	*Country entry*	*Mean MR effect*	*Mean total elasticity*
Linear MR	1.176	-	-	-0.033	1.143
Firm heterogeneity	0.799	0.765	Negligible	-	1.564
Combination	0.799	0.765	Negligible	-0.037	1.526

Intensive margin is OLS estimate of γ in the linear model and second stage NLS estimate of γ otherwise.

Extensive margin for each pair is $\widehat{\varphi}_{ij}\hat{\gamma}_p$. Country entry based on pairs where $\widehat{\tilde{z}}^*_{ij} > 0$ and $\widehat{\tilde{z}}^*_{ij} \leq 0$.

MR is $-\left(1+\widehat{\varphi}_{ij}\right)\frac{\widehat{\gamma}+\hat{\delta}\hat{\gamma}_p}{1+\hat{\delta}}\left(-s_i^{I_j}s_j^{J_i}-s_j^{I_j}s_i^{J_i}+s_j^{J_i}+s_i^{I_j}\right)$ or linear analogue for each pair.

Total is from eqn (20) or linear analogue.

Table 3: Bilateral elasticities for specific country pairs.

Countries		*GDP shares*		Gross	Net	*Ratio*
Exporter	*Importer*	*Exporter*	*Importer*	$\widehat{\xi}_{ij}^{gross}$	$\widehat{\xi}_{ij}^{B}$	$\widehat{\xi}_{ij}^{B}/\widehat{\xi}_{ij}^{gross}$
Bhutan	Japan	0.001%	72.34%	1.49	0.39	0.26
Eq. Guinea	USA	56.10%	42.85%	1.50	0.64	0.43
Kiribati	USA	46.78%	51.92%	1.65	0.85	0.52
Solomon Isl.	USA	45.38%	53.58%	1.56	0.84	0.54
French Guiana	USA	42.76%	56.73%	1.41	0.80	0.57
USA	Japan	29.23%	15.74%	1.28	1.02	0.79
Mexico	Spain	1.721%	1.717%	1.29	1.25	0.97
Mauritania	Togo	0.004%	0.005%	2.04	2.04	0.9999

Table 4: Decomposition of trade flow elasticity for a multilateral reduction in trade frictions

	Intensive margin	*Mean extensive margin*	*Country entry*	*Mean MR effect*	*Mean total effect*
Linear MR	1.176	-	-	-1.162	0.014
Firm heterogeneity	0.799	0.765	Negligible	-	1.564
Combination	0.799	0.765	Negligible	-1.558	0.006

Intensive margin is OLS estimate of γ in the linear model and second stage NLS estimate of γ otherwise.

Extensive margin for each pair is $\widehat{\varphi}_{ij}\hat{\gamma}_p$. Country entry based on pairs where $\widehat{\tilde{z}}_{ij}^{\prime *} > 0$ and $\widehat{\tilde{z}}_{ij}^{*} \leq 0$.

MR is $-\left(1+\varphi_{ij}\right)\frac{\gamma+\delta\gamma_p}{1+\delta}\left\{1+\sum_{l\in I_j} s_l^{I_j} s_l^{J_i} - s_i^{J_i} - s_j^{I_j}\right\}$ or linear analogue for each pair.

Total is from eqn (25) or linear equivalent.

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

	Intensive margin	*Extensive margin*	*Country entry*	*MR effect*	*Total effect*
Linear (no MR or heterog.)	1.176	-	-	-	1.176
Linear (MR but no heterog.)	1.176		-	-1.112	0.064
Firm heterogeneity (heterog. but no MR)	0.799	0.492	Negligible	-	1.291
Combination (heterog. & MR)	0.799	0.492	Negligible	-0.824	0.467
Combination (heterog. & MR - poly.)	0.862	0.425	Negligible	-0.826	0.462

Entries are trade-weighted aggregates of bilateral elasticities calculated in table 4 (or polynomial equivalents).

Figure 1: Net elasticity ($\widehat{\xi}_{ij}^{B}$) and country-pair GDP share when bilateral distance falls in isolation.

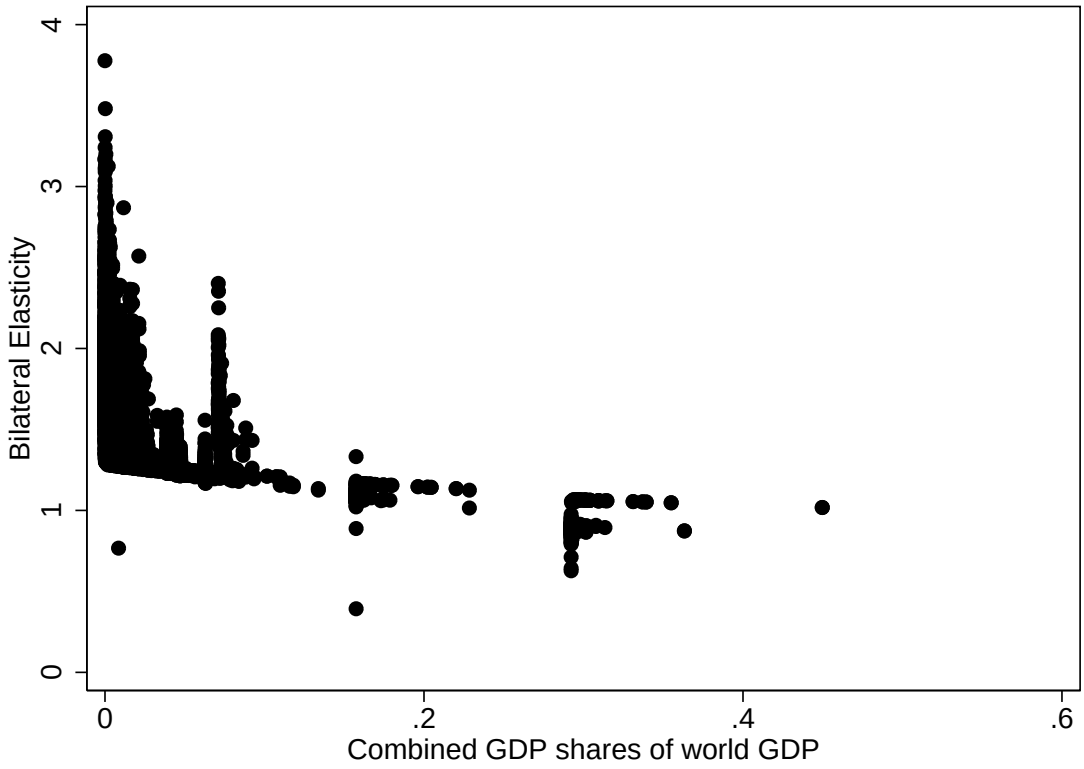
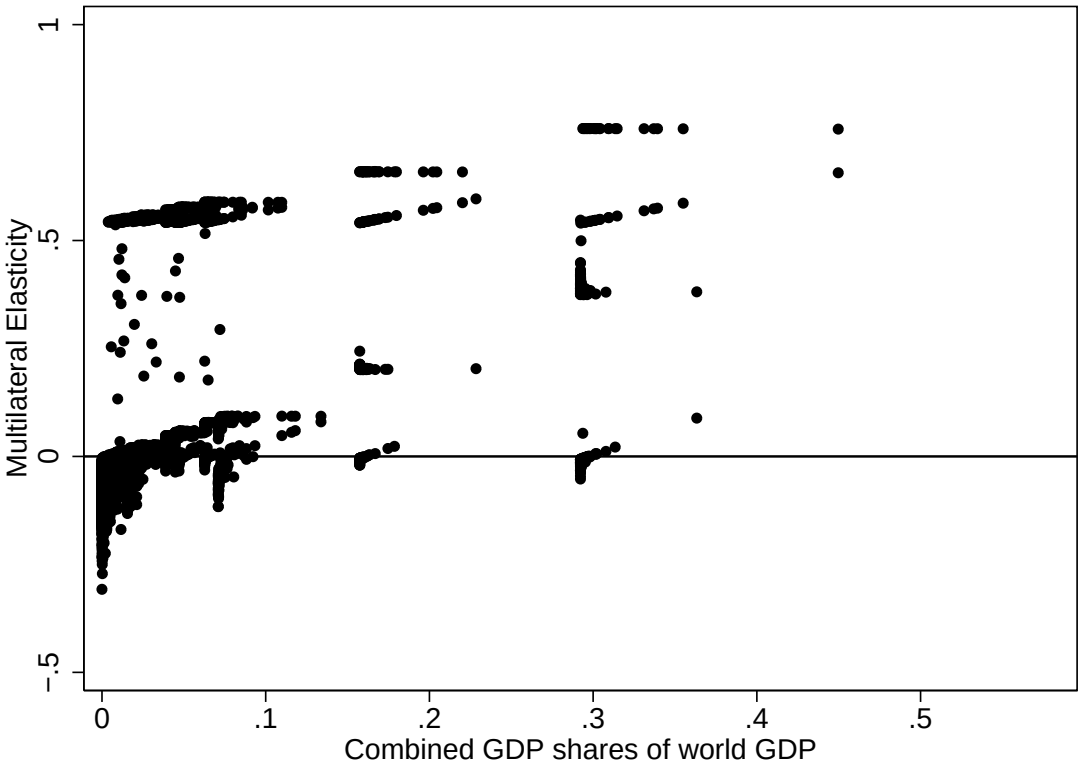


Figure 2: Net elasticity ($\widehat{\xi}_{ij}^{M}$) and country-pair GDP share when all countries reduce frictions.



References

James E. Anderson and Eric van Wincoop. Gravity with gravitas: A solution to the border puzzle. *American Economic Review*, 93(1):170–192, March 2003.

James E. Anderson and Eric van Wincoop. Trade Costs. *Journal of Economic Literature*, 42(3):691–751, September 2004.

James E. Anderson and Yoto V. Yotov. The changing incidence of geography. *American Economic Review*, 100(5):2157–2186, December 2010.

James E. Anderson and Yoto V. Yotov. Terms of trade and global effciency effects of free trade agreements, 1990-2002. *NBER working paper No. 17003*, April 2011.

Scott L. Baier and Jeffrey H. Bergstrand. Bonus vetus OLS: A simple method for approximating international trade-cost effects using the gravity equation. *Journal of International Economics*, 77(1):77–85, February 2009.

Thomas Baranga. Unreported trade flows and gravity equation estimation. *mimeo*, 2009.

Alberto Behar and Laia Cirera i Criville. Does it matter who you sign with? *World Bank Policy Research Working Paper No. 5626*, April 2011.

Alberto Behar and Anthony J. Venables. Transport costs and international trade. In A. de Palma, R. Lindsey, E. Quinet, and R. Vickerman, editors, *Handbook of Transport Economics*, chapter 5, page 97. Edward Elgar, 2011.

Alberto Behar, Benjamin D. Nelson, and Phil Manners. Exports and international logistics. *Oxford Bulletin of Economics and Statistics*, forthcoming.

Kristian Behrens, Giordano Mion, Yasusada Murata, and Jens Suedekum. Trade, wages, and productivity. *CEPR discussion paper No. 7369*, July 2009.

Maxim Belenkiy. The extensive margin and the industry trade: estimation, significance, and implications. *mimeo*, 2009.

Jeffrey H. Bergstrand, Peter Egger, and Mario Larch. Gravity redux: Estimation of gravity-equation coefficients, elasticities of substitution, and general equilibrium comparative statics under asymmetric bilateral trade costs. *Journal of International Economics*, 89(1):110–121, January 2012.

Andrew B. Bernard, J. Bradford Jensen, Stephen J. Redding, and Peter K. Schott. Firms in international trade. *Journal of Economic Perspectives*, 21(3):105–130, Summer 2007.

Donald R. Davis. The home market, trade, and industrial structure. *American Economic Review*, 88(5):1264–1276, December 1998.

Peter Egger, Mario Larch, Kevin E. Staub, and Rainer Winkelmann. The trade effects of endogenous preferential trade agreements. *American Economic Journal: Economic Policy*, 3(3):113–143, August 2011.

Robert Feenstra. *Advanced international trade*. Princeton University Press, 2004.

Gordon H. Hanson and Chong Xiang. The home-market effect and bilateral trade patterns. *American Economic Review*, 94(4):1108–1129, September 2004.

Elhanan Helpman and Paul R. Krugman. *Market structure and foreign trade*. MIT Press, 1985.

Elhanan Helpman, Mark J. Melitz, and Yona Rubinstein. Estimating trade flows: Trading partners and trading volumes. *Quarterly Journal of Economics*, 123(2):441–487, May 2008.

Thomas Hertel, David Hummels, Maros Ivanic, and Roman Keeney. How confident can we be of cge-based assessments of free trade agreements? *Economic Modelling*, 24(4): 611–635, July 2007.

Paul R. Krugman. Scale economies, product differentiation, and the pattern of trade. *American Economic Review*, 70(5):950–959, December 1980.

John McCallum. National borders matter: Canada-us regional trade patterns. *American Economic Review*, 85(3):615–623, June 1995.

Mark J. Melitz. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6):1695–1725, November 2003.

Jacob Viner. *The Customs Union Issue*. Carnegie Endowment for International Peace, 1950.

John S. Wilson, Catherine L. Mann, and Tsunehiro Otsuki. Assessing the benefits of trade facilitation: a global perspective. *The World Economy*, 28(6):841–871, 2005.