Journal of International Economics 66 (2005) 249–266
Journal of INTERNATIONAL ECONOMICS
www.elsevier.com/locate/econbase

Monopolistic competition and trade, revisited: testing the model without testing for gravity[☆]

Peter Debaere[*]

Department of Economics, University of Texas, Austin, TX 78712-1173, United States

Received 11 November 2002; received in revised form 20 February 2004; accepted 24 February 2004

Abstract

In this paper, I show that the increasing similarity in GDPs among OECD country pairs leads to higher bilateral trade to GDP ratios. This finding provides some support for the prediction of Helpman [J. Jpn. Int. Econ. 1 (1987) 62], whose model explains intra-industry trade that is prevalent among developed countries. I also show that Helpman's prediction is rejected for non-OECD countries, among which intra-industry trade is not critical. This result contrasts with the findings of Hummels–Levinsohn [Q. J. Econ. 110 (1995) 799], which play an important role in the debate about whether or not New Trade Theory explains international trade patterns at the country level.

Keywords: Monopolistic competition and trade; New Trade Theory; Gravity

In the early 1980s, a new set of models gained prominence in international trade. Krugman (1979, 1980), Lancaster (1980), Helpman (1981) (and many other others) studied the far-reaching implications of monopolistic competition for international trade theory. To a large extent, this line of research that was part of 'New' Trade Theory was motivated by two stylized facts that the traditional theories of international trade à la Ricardo or Heckscher–Ohlin failed to explain. First, why does most of world trade flow

[☆] The title of the working paper version of this paper is "Testing 'New' Trade Theory without Testing for Gravity: Re-interpreting the Evidence".

[*] Tel.: +1 512 471 3211; fax: +1 512 471 3510.
E-mail address: debaere@eco.utexa.edu.

0022-1996/$ - see front matter
doi:10.1016/j.jinteco.2004.02.007

between developed countries that are similar in terms of endowments and technology? Second, why is it that a major fraction of trade consists of intra-industry trade in similar products?[1] Helpman and Krugman (1985) showed that a monopolistic competition model could explain both facts as long as firms produce differentiated products with increasing returns to scale technology, and as long as consumers have utility functions that reward diversity. In the present article, I revisit the empirical literature that has tested some of the implications of this monopolistic competition model for aggregate international trade patterns.

There is an extensive empirical literature on trade in differentiated products that in many instances preceded New Trade Theory. The early work by Verdoorn (1960), Balassa (1966) and Grubel and Lloyd (1975) documents the growing two-way intra-industry trade between developed countries.[2] This empirical work, however, mostly lacks an explicit link to theoretical models. Against this background, Helpman (1987) is an important contribution since Helpman explicitly derives testable implications from a monopolistic competition model in order to explain the increasing trade to GDP ratios among developed nations. In particular, Helpman predicts that countries exchange a larger fraction of output as they become more similar in terms of size and as their total size as a group increases, i.e. as they produce more varieties.

Helpman's prediction plays an important role in the empirical literature that tests some implications of monopolistic competition models for aggregate trade patterns with country-level data. Helpman (1987) provides graphical evidence to support his prediction for OECD countries that more similar countries trade more. Hummels and Levinsohn's (1995) econometric work confirms Helpman's finding. However, for non-OECD countries, for which intra-industry trade is thought not to matter, Hummels and Levinsohn cannot reject their test equation: The variables of interest have the predicted sign and the R^2 is reasonably high. Hummels and Levinsohn's finding raises the question why the theory that was designed to explain trade among developed countries finds some support among developing countries. This puzzling result has triggered much innovative work. It has led to a reassessment of the theoretical foundations of the gravity equation, see Deardorff (1998), and also to attempts to disentangle some of the alternative theories that are consistent with Helpman's test equation, see Feenstra et al. (2001) and Evenett and Keller (2002).[3]

However, despite the prominence of Helpman's prediction, until now, Hummels and Levinsohn's findings have not been re-visited. In this paper, I argue that an empirical re-investigation is warranted. Hummels and Levinsohn's test results do not account for countries' changing size in the world and, more importantly, their results are driven by the correlation between countries' volume of trade and their size, not by the similarity of their GDPs. Since we know from the gravity literature that the trade-size correlation is a feature of almost any theory and any data set, Hummels and Levinsohn's findings

[1] Davis (1995) shows that traditional models can generate intra-industry trade. Davis (1997) indicates how Hecksher–Ohlin–Vanek is theoretically consistent with massive trade flows among developed countries.

[2] For more recent evidence, see Helpman (1998).

[3] See also Baier and Bergstrand (2001).

are based on a very weak test of Helpman's theory. In this paper, I go back to Helpman's original formulation of the prediction that includes changing world shares and that is not affected by the trade-size correlation. I find that country-level data provides some qualitative support (signs corresponding to the predictions of the theory, a reasonable R^2) for Helpman's predictions of the aggregate trade patterns between countries where it is expected to, i.e. for developed countries. Trade among OECD countries as a fraction of their GDP is primarily explained by the increasing similarity in size and by the changing shares in the world economy. However, it is hard to interpret the empirical results for non-OECD countries, where one would not expect any support for Helpman's prediction to start with, as supporting the theory in any meaningful way: The changing size of non-OECD countries in the world drives the changing fraction of output that non-OECD countries trade; there is no significant, positive impact of country similarity.

It should be noted, however, that while my results question the ability of monopolistic competition models to explain international trade patterns between countries at the aggregate level especially among non-OECD countries, they do not disqualify New Trade Theory per se. For one, not all imperfect competition models of New Trade Theory use monopolistic competition as market structure. In addition, there is a flourishing empirical literature that relies on increasingly available firm-level data sets. This literature finds, also for non-OECD countries, some support for other features of New Trade Theory's models, such as for example for the fixed costs of production, the markups, etc. For an extensive discussion of this emerging literature, see Tybout (2001).

1. Explaining the volume of trade

Helpman (1987) studies why developed countries have been trading an increasing fraction of their output since World War II. He takes a structural equation from the model that he and Krugman developed in order to explain the increasing intra-industry trade between developed countries. In the model, scale economies and product differentiation induce specialization, so that each country produces a different subset of products. As countries have identical homothetic preferences, they consume each good proportional to their GDP. Therefore, with no trade impediments and balanced trade, countries should trade an increasing fraction of their GDP as they become more similar and as they grow faster than the rest of the world. The following equation expresses this idea for a group of countries A.[4]

$$\frac{VT_A}{Y_A} = e_A\left[1 - \sum_{i\in A}\left(\frac{Y_i}{Y_A}\right)^2\right] \tag{1}$$

[4] The derivation of the equation is found in the work of Helpman and Krugman (1985), pp. 166–167.

VT_A is the total volume of bilateral trade between all the countries i that are part of the group A. Y_i is the GDP of a country from group A and Y_A the joint GDP for the group. e_A is the share of total GDP of the group A in world GDP.

Eq. (1) contains two parts: the size of a group of countries in the world economy, e_A, and the size distribution within that group as measured by the similarity index in brackets. For a given set of countries, this index attains its maximum when countries are equal in size and its minimum when they are very unequal. The index has been the focus of the literature. It is an attractive measure, especially if one believes Helpman's hypothesis that the increasing fraction of output that developed countries trade among themselves reflects growing intra-industry trade. Since intra-industry trade should be primarily important among developed countries, the similarity index directly links the growing trade to GDP numbers for developed countries to a specific characteristic of these countries as a group, irrespective of their weight (e_A) in the world economy.[5]

Helpman (1987) provides some support for his prediction for 14 OECD countries for which intra-industry trade as measured by the Grubel–Lloyd index is typically high. Helpman calculates the total bilateral trade among them as a fraction of their joint GDP and relates it to their size similarity while assuming that the share of the OECD countries in the world economy, e_A, stays constant. He plots both series and finds a positive relation: More similarity in terms of country size is translated into higher trade to GDP ratios.

Hummels and Levinsohn (1995) provide an econometric test for country pairs of Helpman's prediction. They generate all (91) country pairs for Helpman's 14 countries during 1962–1977, which allows them to apply panel data techniques. The latter is an important innovation since bilateral trade flows are known to systematically depend on country-pair specific factors such as bilateral distance, tariffs, a common border, common language, etc. These non-random factors can be accounted for in a straightforward way by country-pair fixed effects in a panel. Hummels and Levinsohn also randomly choose 14 non-OECD countries for which intra-industry is not important and construct a similar panel. In doing so, they want to test whether the trade model finds empirical support in an environment where it is not appropriate.

A straightforward test equation of Helpman's prediction is Eq. (2a), which is a logarithmic transformation of Eq. (1) that is written as a fixed effect regression for any two countries i and j. (Throughout the paper, I use ij for the pair of countries i and j.) This specification has several advantages. It is directly consistent with the theory and it is also flexible, i.e. it allows us to estimate the coefficients for e_{ijt} and

[5] Eq. (1) hinges on product specialization, identical homothetic preferences and uniform international prices. Since these features are not unique to Helpman's model, a test of Eq. (1) can only tell whether the data is consistent with New Trade Theory or not. A Heckscher–Ohlin model with complete specialization can e.g. also generate Helpman's equation. The same is true for a perfect competition model with the Armington assumption and constant return to scale. Note, however, that the theory does not suggest in either of these cases that the equation should work in the OECD and not outside the OECD. In particular, one expects the opposite for the H–O model as complete specialization occurs with very dissimilar endowments.

sim_{ijt} (without imposing that they are the same or equal to one) and it allows us to separately assess the impact of e_{ijt} and sim_{ijt}. Moreover, the specification allows us to account for features of reality (such as transportation costs and bilateral trade frictions) that are known to matter and that are not included in the simple theoretical model. In particular, we have a multiplicative fixed effect that can capture bilateral transportation costs, language barriers or other bilateral trade frictions. This multiplicative fixed effect is consistent with a large body of empirical work.[6] The specification is ideal for the OECD countries. There is, however, for the group of non-OECD countries a concern, since there are about 7% zero bilateral trade observations. There is no perfect way to address this concern since there is a tradeoff between being true to the theoretical specification on the one hand and addressing the censoring issue on the other hand. I propose two alternatives. On the one hand, I take the logarithm of a small number in the case of a zero trade observation. The advantage of this approach is obviously that it leaves intact the logarithmic transformation of Helpman's equation and the advantages of this specification. In addition, this approach, at least to some extent, addresses the criticism that in many instances the reported zeros of trade statistics may in reality correspond to very small amounts of bilateral trade. I, however, also take another route that handles the zero observations better. I look for a specification in levels that will allow me to deal more accurately with censoring and run a Tobit regression or, alternatively, a linear fixed effect regression for the non-OECD countries. From a theoretical point of view, these level specifications will not be as good a translation of Eq. (1) as expression (2a) is. Hopefully we can, however, corroborate the estimates of Eq. (2a) with the level regressions.

There are two possible regressions in level that come to mind. Regression (2b) in levels is most closely related to Eq. (1). It is different from Eq. (2a), however, in that it estimates one coefficient for ($e_{ijt}.\text{sim}_{ijt}$). In other words, it imposes the same coefficient for e_{ijt} and sim_{ijt} and it does not allow us to separately evaluate the impact of country similarity, sim_{ijt}, that has been at the heart of the empirical debate. In addition, transportation costs that are not part of the theoretical setup but that are known to affect bilateral trade will be taken care of only additively through the additive fixed effect, which is a drawback.[7] Finally, there is regression (2c) that builds on Eq. (2b) and that separates out e_{ijt} from sim_{ijt}. While this regression in levels just like Eq. (2b) has the advantage of allowing for a Tobit estimation, it also addresses bilateral frictions in an additive way. Eq. (2c) should give us a sense of the separate impact of e_{ijt} and sim_{ijt} on VT_{ijt}/Y_{ijt}, even though, strictly speaking, the theoretical Eq. (1) dictates a multiplicative, not an additive

[6] Moreover, if one assumes as in Deardorff (1998) a Cobb Douglas utility function and iceberg transportation costs to theoretically justify the frequently used gravity equation with multiplicative transportation costs, one can derive a bilateral variant of Helpman's Eq. (1). In this variant, the term on the left-hand side, VT_{ij}/Y_{ij}, equals $t_{ij}.e_{ijt}.\text{sim}_{ijt}$. (Note that t_{ij} stands for the bilateral frictions (transportation costs) between countries i and j and as one can see, it enters multiplicatively.)

[7] Should there be a multiplicative component to transportation costs, which empirical work suggests there is, then the previous footnote indicates that an additive fixed effect specification of Eq. (1) as in Eqs. (2b) or (2c) will primarily relate transportation costs to Y_{ijt} instead of also to $e_{ijt}.\text{sim}_{ijt}$.

relationship between both variables. For the sake of completeness, I will in addition to estimating the regressions (2b) and (2c) for the non-OECD countries, report the results for the OECD countries.

$$\ln \mathrm{VT}_{ijt} - \ln Y_{ijt} = \alpha_{ij} + \beta_1 \ln e_{ijt} + \beta_2 \ln \mathrm{sim}_{ijt} + \varepsilon_{ijt},$$

$$\text{with } \mathrm{sim}_{ijt} = 1 - \left(Y_{it}/Y_{ijt}\right)^2 - \left(Y_{jt}/Y_{ijt}\right)^2 \tag{2a}$$

$$\mathrm{VT}_{ijt}/Y_{ijt} = \alpha_{ij} + \beta_1 \left(e_{it}\mathrm{sim}_{ijt}\right) + \varepsilon_{ijt}, \text{ with } \mathrm{sim}_{ijt} = 1 - \left(Y_{it}/Y_{ijt}\right)^2 - \left(Y_{jt}Y_{ijt}\right)^2 \tag{2b}$$

$$\mathrm{VT}_{ijt}/Y_{ijt} = \alpha_{ij} + \beta_1 e_{ijt} + \beta_2 \mathrm{sim}_{ijt} + \varepsilon_{ijt},$$

$$\text{with } \mathrm{sim}_{ijt} = 1 - \left(Y_{it}/Y_{ijt}\right)^2 - \left(Y_{jt}/Y_{ijt}\right)^2 \tag{2c}$$

It should be noted that Hummels and Levinsohn choose a specification that is different from mine. They maintain Helplman's assumption that e_{ij} is constant and that it is therefore part of the fixed effect. They bring the sum of the GDPs, Y_{ij}, to the right-hand side and multiply it with the similarity index as in regression (3). For the non-OECD countries, they run the regression in levels as in Eq. (4). As a yardstick to judge the extent to which the estimates fit Helpman's prediction, they note that by Helpman's logic, a positive coefficient on the product of the similarity index and the joint GDPs suggests that the monopolistic competition model works. They also find it important to obtain a reasonably high R^2. One could argue that the theory dictates that the β-coefficients should be one in all regressions. On the other hand, in the presence of distance or of various other trade frictions such as tariffs or quota and with the possibility of measurement error, it could be argued that requiring a positive coefficient, as they do, is a reasonable criterion.

$$\ln \mathrm{VT}_{ijt} = \alpha_{ij} + \beta \ln\left(\mathrm{sim}_{ijt} Y_{ijt}\right) + \varepsilon_{ijt} \tag{3}$$

$$\mathrm{VT}_{ijt} = \alpha_{ij} + \beta\left(\mathrm{sim}_{ijt} Y_{ijt}\right) + \varepsilon_{ijt} \tag{4}$$

In this paper, I argue that bringing Y_{ij} to the right-hand side in order to multiply it with the similarity index is not an innocuous transformation, especially not for the non-OECD countries. From the gravity literature, we know that there is a strong, positive correlation between a country's trade volume and its GDP that easily carries over to Helpman's variables VT_{ij} and Y_{ij}. If big countries on average trade more than small ones, there is a positive association between their bilateral trade and the sum of their GDP, so that the covariance Cov (VT_{ij}, Y_{ij}) is positive. Therefore, by multiplying the joint GDPs Y_{ij} and the similarity index sim_{ij} Hummels and Levinsohn allow size to impact on the estimation. In particular, as the proof in Appendix A indicates, a regression à la Hummels–Levinsohn yields a positive coefficient for the variable on the right-hand side (the similarity index of country pairs times the sum of their GDPs), even when there is no relation whatsoever between the trade to GDP ratio of country pairs and the extent to which these countries are similar. Moreover, due to the positive correlation between the volume of trade of countries and their GDPs, a reasonable R^2 is guaranteed even for non-OECD countries. In sum, a regression that employs a specification à la Hummels–Levinsohn is bound to deliver a positive coefficient and a relatively high R^2, the two criteria that Hummels and Levinsohn used as yardstick to judge the success of their test.

2. The data

I use the Feenstra et al. (1997) bilateral trade data, which recompile the UN data for the period 1970–1992. I work with the country aggregates. I collect the bilateral trade data for Helpman's 14 OECD countries and non-OECD countries that closely match the ones Hummels and Levinsohn chose.[8] In the first group, there are no zero bilateral trade observations and in the second group about 10% of the observations are zero.

The GDP numbers are from the *Penn World Tables* and the *IMF World Economic Outlook*. The first database converts all country values into dollars with PPP-values; in the latter, nominal dollar exchange rates are used. Using both data sets allows me to check the robustness of my results.[9] For each group of countries, I construct a balanced panel, making sure that there are no missing GDP data. The period runs from 1970 to 1989 since it allows for a maximum number of countries in both data sets. To check whether a difference in data sets drives my results, I replicate Hummels and Levinsohn's findings in Appendix B. My estimates are fairly similar. Note that I need time series data on world GDP for the implementation of the Eqs. (2a) and (2b) as I need to calculate a country pair's share of world GDP. I take the GDP numbers from 101 countries mentioned in Appendix C for which I have a full panel.

Since there is potentially an endogeneity problem for my regressions (see next section), I want to instrument for e_{ijt} and sim_{ij}. To do so, I instrument for a country's GDP with its endowments and use these instruments (instead of the regular GDP numbers) to construct e_{ijt} and sim_{ij}. As endowments, I take a country's total population and its capital stock data from the *Penn World Tables.* I do this for the OECD, the non-OECD and as well as for the 101 countries mentioned before.[10]

Figs. 1 and 2 suggest how bringing the sum of the GDPs to the right of the test equation as in the work of Hummels and Levinsohn (1995) may affect the outcome of the test. Fig. 1 plots the bilateral trade vs. the similarity index times GDP during 1970–1989 for two sets of country pairs, the US and Canada vs. South Korea and

[8] The countries are Canada, Belgium, the US, the Netherlands, Ireland, Italy, Austria, the UK, Japan, Sweden, Switzerland, France, Denmark and Germany. Thailand, South Korea, Brazil, Greece, Cameroon, Columbia, Nigeria, Pakistan, Paraguay, Norway, Peru, the Philippines. Congo and Ivory Coast are the only two countries from the Hummels and Levinsohn data set that are not included.

[9] Both data are being used in the literature. One can make an argument in support of either nominal exchange rates or PPP corrections.

[10] Since only population data are available for all 101 countries, I can only use population data (and not the capital stock) to instrument for *GDP* and thus for e_{ij}. Note that the capital stocks are available for the OECD countries and for most of the non-OECD countries. When a country's capital stock is not available, I approximate it by multiplying its population with the capital–labor ratio of a country with a similar per capita GDP.

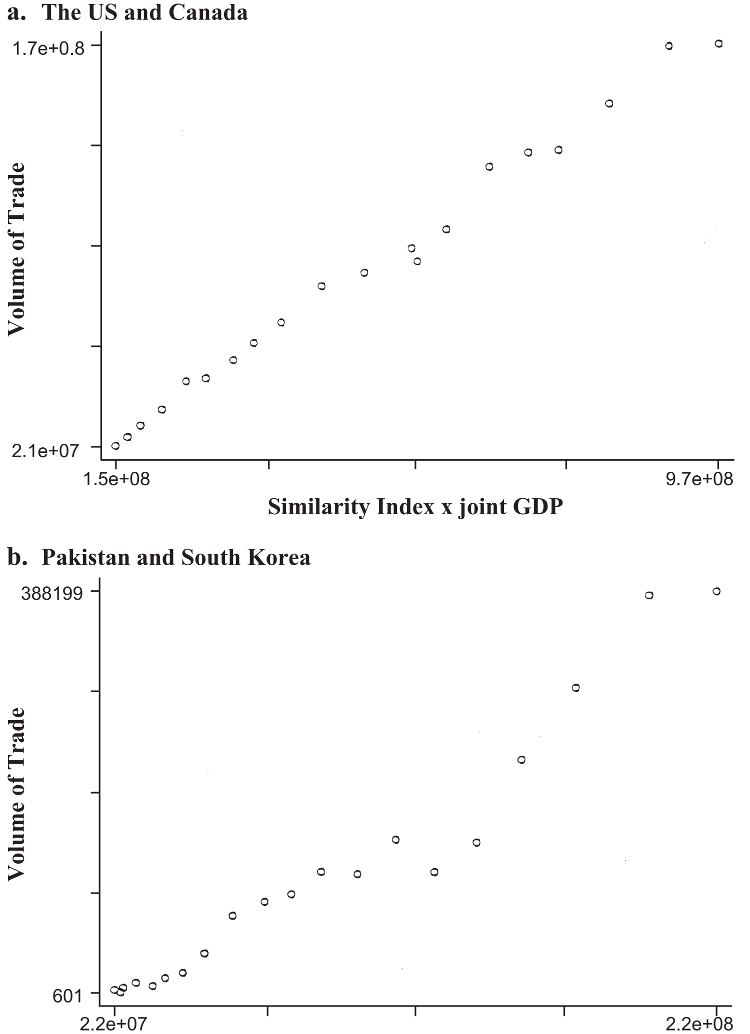

Fig. 1. Bilateral trade vs. similarity index times GDP.

Pakistan. Fig. 2 does the same for bilateral trade as a ratio of joint GDP on the one hand and the similarity index on the other hand. (I present the *Penn World Table* data. The *IMF* data are similar.) For both groups of countries, there is a positive relation between the volume of trade and the sum of the GDPs times the similarity index. However, there is a different relation between the trade to GDP ratio and the similarity index. It is positive among the OECD countries and negative among the non-OECD countries. One should be careful when interpreting such graphs, since a trend could be driving them. Still, Figs. 1 and 2 suggest the overriding impact of the correlation between the volume of trade and GDP. There can be a positive relation between trade

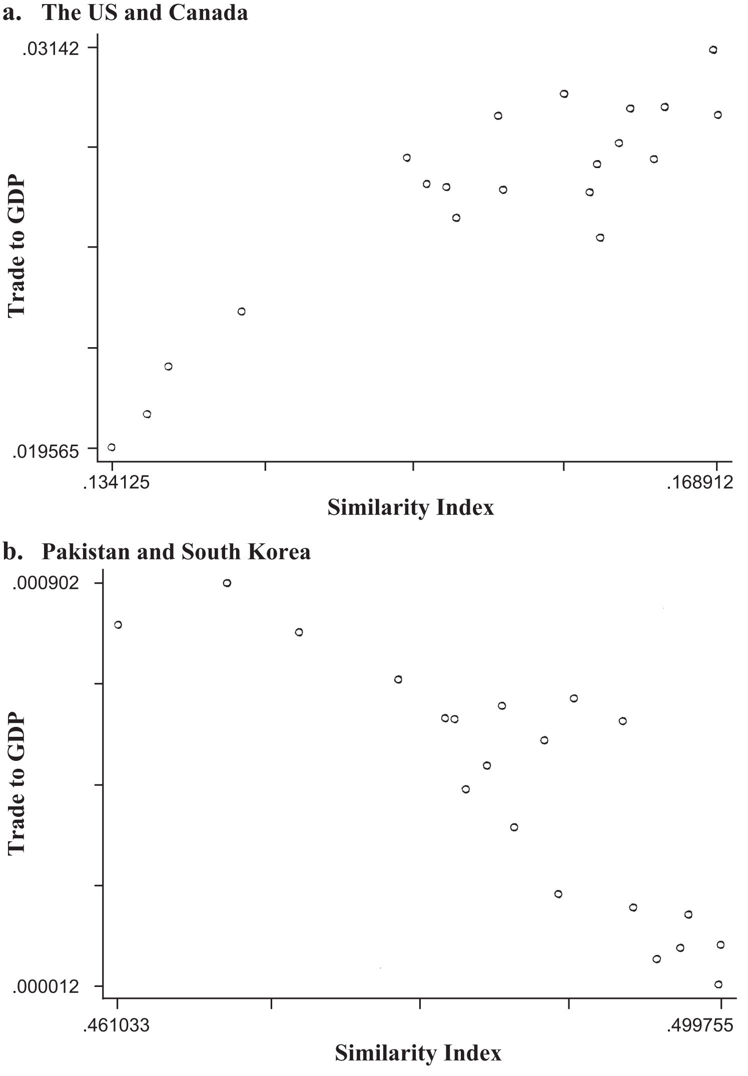

Fig. 2. Bilateral trade to GDP ratio vs. similarity index.

and the similarity index times the sum of the GDPs whether or not there is a positive relation between dispersion and trade to GDP ratios.

3. Empirical results and discussion

Tables 1, 2 and 3 present the estimation results respectively for regressions (2a), (2b) and (2c). The columns to the left report the results for the group of OECD countries and the columns to the right those for the group of non-OECD countries. Since a trade shock

Table 1
Eq. (2a) in logs with varying shares panel data

Fixed effects regression with time-specific effects

Dependent variable	OECD				Non-OECD			
	lnVT/Y	lnVT/Y	lnVT/Y	lnVT/Y	lnVT/Y	lnVT/Y	lnVT/Y	lnVT/Y
	Penn	IMF	Penn	IMF	Penn	IMF	Penn	IMF
ln(similarity)	1.57	0.89	–	–	−0.96	0.4	–	–
S.E.	0.11*	0.056*			0.99	0.23		
ln(world share)	1.3	0.47	–	–	1.98	0.99	–	–
S.E.	0.13*	0.05*			0.95*	0.22*		
ln(similarity) IV	–	–	0.25	0.66	–	–	−2.3	−0.5
S.E.			0.06*	0.06*			0.6*	0.26*
ln(world share) IV	–	–	0.62	0.98	–	–	5.4	3.1
S.E.			0.25*	0.22*			2.9	1.2*
R^2	0.61	0.45	0.56	0.41	0.02	0.14	0.03	0.14
Observ.	1820	1820	1820	1820	1320	1320	1320	1320

Standard errors under the estimated coefficients.
* Significant at 95%.

not only affects the dependent variable in the regression, but also all the regressors that are a function of countries' GDP (and thus, of their trade), there is potentially an endogeneity problem. As discussed, I therefore instrument for the measures e_{ij} and sim_{ij} as follows: I instrument for countries' GDPs with their endowments and then use these instruments (instead of the actual GDP numbers) to construct both the share and the similarity measure. Each time, I report the regular estimates in the first rows before turning to the instrumental variable estimates in the lower rows. Note that the regression also includes year-specific effects, since there are two oil shocks during the period that is studied.

As one can see for the logarithmic specification in Table 1, a clear pattern emerges. The logarithmic specification that is clearly the preferred one for the OECD countries explains, as the block of columns to the left indicates, between 60% and 40% of the variation in the data for both data sets. As the estimates in row two indicate, I obtain a positive coefficient for both the similarity index and for the share of the country pairs in the world economy. Both estimated coefficients are significant at the 95% level. Moreover, using instrumental variables (reported in the fourth row) does not alter the picture. As Helpman (1987) predicted, more country similarity makes developed countries trade a bigger fraction of their output and so does an increase in their share of the world economy. The estimates in Table 1 also illustrate that the estimated coefficients are statistically different from their theoretical value of 1 in almost all cases—a feature that was clear already in the work of Hummels and Levinsohn (1995) who like Helpman (1987) made the sign rather than the exact magnitude of the estimated coefficients the yardstick for judging empirical support. As noted before, one could argue that maybe because of the simplicity of the model (that does not fully account for all possible trade frictions) and because of the presence of measurement error, one should probably not expect a coefficient that is indistinguishable from unity. Note also that the coefficients on the share and similarity variable are statistically different in all cases.

Table 2
Eq. (2b) in levels with varying shares panel data

Dependent variable	Fixed effects regression with time-specific effects				Tobit regression with time and country-specific effects			
	OECD				Non-OECD			
	VT/Y	VT/Y	VT/Y	VT/Y	VT/Y	VT/Y	VT/Y	VT/Y
	Penn	IMF	Penn	IMF	Penn	IMF	Penn	IMF
Similarity*share	−0.029	0.59	–	–	0.07	0.17	–	–
S.E.	0.03*	0.9			0.002*	0.015*		
Similarity*share IV	–	–	−0.38	−0.14	–	–	0.012	0.029
S.E.	–	–	0.07*	0.05*	–	–	0.004*	0.006*
R^2	0.29	0.22	0.27	0.21	−0.08	0.08	−0.08	0.07
Observ.	1820	1820	1820	1820	1320	1320	1320	1320

For the set of non-OECD countries that are reported in the second block of columns in Table 1, I obtain different results that contrast with the systematic support for the estimation equation among the OECD countries. As noted, the logarithmic specification is attractive since it puts any multiplicative, pair-specific and non-random factors such as distance, tariffs, cultural barriers or border in the fixed effect. It also allows us to separately evaluate the impact of the share and the similarity index that is at the heart of the debate. The drawback of this specification, however, is that it does not account for the 7% zero observations all that well—the logarithm of a small number is included if there is no bilateral trade. For the non-OECD countries, I estimate a positive coefficient for a country pair's share in the world economy. The coefficient is significant at the 95% level, except in one case (Penn World data with IV). As for the coefficient on the similarity index, it is either insignificant or it has a sign that contradicts Helpman's prediction. For reference, the coefficients for either a country pair's share in the world economy or its size similarity never equal 1.[11] In addition, the coefficients for the similarity and the share index are clearly different.

In Table 2, I report the Tobit estimates of regression (2b) and in Table 3 those for regression (2c).[12] The primary reason for estimating Eqs. (2b) and (2c) is to address issues related to zero bilateral trade observations in the group of non-OECD countries. Let us first analyze these.

The Tobit regression (2b) estimates the coefficient on the joint product of the similarity index and the share of both countries in the world economy. We obtain a positive and significant coefficient for the non-OECD countries across the board, with or without using instrumental variables. (The estimated coefficient is significantly different from one.) One should be reluctant to interpret this outcome in terms of support for Helpman's prediction. When we additively separate the share and the

[11] Note that the estimates for the linear probability regressions based on Eq. (2b) that correspond to the Tobit estimates reported in Tables 2 and 3 are qualitatively similar. The R^2 for these regressions (instead of the Pseude R^2 of the Tobit regressions) is more directly comparable to the R^2 for the linear fixed effect regressions for the group of OECD countries. The R^2 for both regressions is relatively low; it ranges between 0.04 and 0.14.

[12] The country pair fixed effect were estimated as separate dummies.

Table 3
Eq. (2c) in levels with varying shares panel data

Dependent variable	Fixed effects regression with time-specific effects				Tobit regression with time and country-specific effects			
	OECD				Non-OECD			
	VT/Y	VT/Y	VT/Y	VT/Y	VT/Y	VT/Y	VT/Y	VT/Y
	Penn	IMF	Penn	IMF	Penn	IMF	Penn	IMF
Similarity	−0.06	0.001	–	–	4E−05	0.0001	–	–
S.E.	0.01*	0.002			0.0003	0.00003*		
Share of world	0.013	0.016	–	–	0.046	0.06	–	–
S.E.	0.004*	0.02*			0.005*	0.007*		
Similarity IV	–	–	−0.6	−0.38	–	–	−0.0002	−0.0003
S.E.			0.04*	0.03*			0.0003	0.0004
Share of world IV	–	–	0.009	0.01	–	–	0.07	0.09
S.E.			0.002*	0.002*			0.01*	0.01*
R^2/Pseudeo R^2	0.28	0.22	0.34	0.26	−0.08	0.07	−0.08	0.07
Observ.	1820	1820	1820	1820	1320	1320	1320	1320

Standard errors under the estimated coefficients.
* Significant at 95%.

similarity index as in regression (2c), we obtain a positive impact of the share variable (with and without instrumental variables) and a coefficient on the similarity index that is either insignificant or that has a negative sign. This suggests that only the shares are responsible for the positive coefficient in regression (2b). Even though specification (2b) is stricter than Eq. (2a) (It imposes the same coefficient on similarity and share, which, judging by the estimates of Eq. (2a) is not the case), and even though it together with specification (2c) does not account for multiplicative transportation costs, the estimates of the level regression (2b) and (2c) corroborate the results of the logarithmic specification (2a): Country shares affect trade to GDP ratios positively and the index of country similarity does not seem to play a significant role for non-OECD countries.

Let's now turn to the estimates of Eq. (2b) and (2c) for the OECD countries in the left panel of respectively Table 2 and 3. As noted, for this group of countries zero bilateral trade observations are not a concern and the logarithmic specification is clearly the preferred specification, since it does not impose conditions that do not hold. The estimates may be of interest, however, to assess the results for the group of non-OECD countries. Across the board (with and without instrumental variables) we obtain for regression (2b) a negative coefficient for the product of the similarity index and the share in Table 2. Table 3 suggests that this outcome is primarily driven by the negative impact of the share variable (which has a negative coefficient) and not by the index of country similarity (which has a positive sign). As mentioned before, compared to the logarithmic specification (2a) the assumptions underlying regressions (2b) and (2c) are different (multiplicative vs. additive fixed effect, separate vs. same coefficient for share and similarity index in (2b), multiplicative vs. additive relation between share and similarity index in (2c)).

Together with the logarithmic estimates, the level estimates confirm, on the one hand, what our findings in Table 1 had already suggested and what we know from Hummels and Levinsohn (1995): It is not the case that the literal formulation of Helpman's specification (with coefficients that are the same and equal to one) is borne out by the data, which is why we are primarily interested in signs not magnitudes. On the other hand, the additive level estimates for OECD and non-OECD countries, if anything, confirm the logarithmic estimates that country similarity is clearly a factor for the OECD, whereas it is not for the non-OECD countries. While it would have been nice if the estimates for regression (2c) (and hence) (2b) corroborated for the OECD the positive sign of the share coefficient from the logarithmic specification, the negative coefficient in regressions (2b) and (2c) should not question the positive contribution of country share to trade to GDP ratios for the OECD as suggested by Helpman's specification. On the one hand, the negative coefficients are either based on an additive specification that does not allow for multiplicative transportation costs or on a specification (in the case of 2b) that imposes restrictions (same coefficients for share and similarity) that do not hold. On the other hand, the preferred specification with multiplicative fixed effects that does not impose equal coefficients clearly shows the share variable with the right sign. Finally, note that the R^2 of the level regressions (2b) and (2c) range between 0.22 and 0.34 for the OECD countries, which is significantly higher than the R^2 that we obtain for the non-OECD countries from the linear regressions that correspond to the Tobit regressions: These range between 0.04 and 0.14. Based on the estimation results that I present, it is hard to argue that data for non-OECD countries systematically supports the aggregate predictions of the international trade patterns of one of New Trade Theory's most visible models. While the increased share of non-OECD country pairs seems to account for increasing trade to GDP ratios, the role of country similarity that has been at the heart of the debate is insignificant or works in a way that contradicts Helpman's prediction. I should also point out that the R^2 of the regressions for non-OECD countries is much lower than for OECD countries. In other words, measured by Hummels and Levinsohn's own criteria (signs corresponding to the theory and a reasonably high R^2), there is support at the aggregate level for Helpman's prediction among OECD countries, which is where one would expect support based on the theory. For the non-OECD countries, on the other hand, it is hard to argue that non-OECD countries support Helpman's prediction, which contrasts with Hummels and Levinsohn's basic point.

Note finally that I have presented in the Tables 1, 2 and 3 the estimates for the most general specification of Helpman's equation that includes both the similarity of country pairs in addition to country pairs' changing share in the world economy. This formulation is less restrictive than the specification that Helpman and Hummels and Levinsohn took as a starting point for their analysis. These authors mainly, if not exclusively, focused on the country similarity, while ignoring changing world shares or while assuming that these stay constant. Note, however, that the same pattern (support for Helpman's prediction among OECD countries with the predicted positive role of country similarity vs. no support among non-OECD countries and no positive impact of country similarity) is found for these other, more restrictive specifications.[13]

4. Conclusion

Helpman (1987) and Hummels and Levinsohn (1995) have played a prominent role in the empirical trade literature that studied some of the implications of monopolistic competition models for the patterns of international trade. Helpman provided the first test for OECD countries that was based on a structural equation that was directly taken from the theory. He provided some evidence that related country similarity to increasing trade to GDP ratios. Hummels and Levinsohn expanded Helpman's results and brought to the fore a paradox that for many years has been central to discussions about the empirical support for New Trade Theory. They questioned Helpman's test of the monopolistic competition model because it could not differentiate OECD trade (that should be predominantly trade in differentiated products) from non-OECD trade (for which differentiated products should not be so prevalent). Hummels and Levinsohn's finding inspired much innovative research.

In the present article I have revisited the empirical support for Helpman's prediction, while drawing on the lessons from the gravity literature that exposes the strong correlation between trade and GDP. I find that increased trade to GDP ratios for a group of OECD countries are positively related to their share in the world economy and to their similarity in terms of size, which is consistent with the theory. On the other hand, I show how Hummels and Levinsohn's results for non-OECD countries are effectively driven by the correlation between country size and trade volume. I argue that it is hard to find systematic support for Helpman's prediction among non-OECD countries. While the trade to GDP ratios for these countries are positively related to their share in the world economy, there is no meaningful way to relate these to their similarity in terms of size.

Note finally that while the reported results question the ability of Helpman and Krugman's monopolistic competition model to explain international trade patterns between countries at the aggregate level especially among non-OECD countries, they do not disqualify New Trade Theory per se for the study of non-OECD countries. For one, even though monopolistic competition models have played a prominent role in New Trade Theory, not all its imperfect competition models use monopolistic competition as

[13] If one estimates regressions (2a) and (2c) while setting the share-coefficient equal to zero, one estimates a coefficient for the similarity index whose sign is consistent for the OECD countries with what Helpman predicts. For non-OECD countries, this is not the case. The R^2 is also reasonably high for the OECD and markedly lower for non-OECD countries. (The estimated coefficients for the similarity index mostly differ from 1). Only in one case (with IV and IMF data) does one obtain a positive and significant coefficient on similarity for non-OECD countries. In all other instances is the coefficient for similarity insignificant or is the estimated sign not conform to the theory. Cross-sectional estimates reveal a similar pattern of support (to capture country pair specific effects, I include bilateral distance).

market structure. In addition, there is a flourishing empirical literature that studies (and finds support for) other features of New Trade Theory with more disaggregate, firm-level data.

Acknowledgement

I thank Alan Deardorff, Ufuk Demiroglu, Robert Feenstra, Li Gan, Gene Grossman, Jim Harrigan, Wolfgang Keller, Subal Kumbhakar, Keith Maskus, Bob Stern and Jim Tybout for their suggestions and support. Presentations at the fall NBER international trade and investment meeting in Stanford, the empirical trade conference at Rutgers University and the Midwest International Economics Meetings at the University of Illinois at Urbana Champane and the ASSA meetings in Boston proved helpful as well and so were the suggestions by two anonymous referees of this journal. All remaining errors are mine.

Appendix A

I show in a simple example how a Hummels–Levinsohn specification yields a positive coefficient, even when there is by assumption no relation whatsoever between trade to GDP ratios and country similarity.[14] Consider the following regression. For notational simplicity I replace the similarity index sim_{ij} by x_k, the volume of bilateral trade VT_{ij} by y_k and the joint GDP Y_{ij} by z_k.

$$\frac{y_k}{z_k} = \alpha + \beta x_k + \varepsilon k \tag{A1}$$

I assume that there is a positive correlation between y and z, $\mathrm{Cov}(y_k,z_k)>0$, which corresponds to the trade-size correlation. Assume furthermore that

$$E(\varepsilon|x_k) = 0$$

$$\text{And } \underline{x} > 0 \left(\underline{x} = \sum_t x_t/T \right)$$

[14] I thank Ufuk Demiroglu for pointing me towards this example.

Note that the first is a standard assumption in an OLS context and the second one is automatically satisfied in the context of Helpman's specification since x_k lies in the interval [0, 1/2] for country pairs.

Now consider the slope coefficient in a regression of y_k on $x_k z_k$,

$$\begin{aligned} b &= \Sigma(x_k z_k - \underline{x}_k \underline{z}_k)(y_k - \underline{y}_k) / \left[\Sigma(x_k z_k - \underline{x}_k \underline{z}_k)^2\right] \\ &= [\Sigma(x_k z_k - \underline{x}_k \underline{z}_k)/T(y_k - \underline{y}_k)/T] / \left[\Sigma(x_k z_k - \underline{x}_k \underline{z}_k)^2/T^2\right] \\ &= Cov(y_k, x_k z_k)/Var(x_k z_k) \end{aligned}$$

Assume now that x is independent of y and z. In that case, the estimated coefficient b will be positive, even though $\beta=0$ in Eq. (A1) by assumption. In order to show this, I focus on Cov(y_k, $x_k z_k$) because Var($x_k z_k$) is positive.

$$Cov(y_k, x_k z_k) = E[(x_k z_k - E(x_k z_k))(y_k - E(y_k))]$$

$$= E[x_k z_k(y_k - E(y_k))] - E[E(x_k z_k)(y_k - E(y_k))]$$

Use the fact that x is independent of y and z:

$$= E(x_k)E[z_k(y_k - E(y_k))] - E(x_k z_k)E(y_k - E(y_k))$$

Since the last term is zero, I obtain

$$= E(x_k)E[z_k(y_k - E(y_k))] = E(x_k)Cov(z_k, y_k)$$

Since both terms are positive by assumption, Cov(y_k, $x_k z_k$)>0. In other words, we estimate a positive b in a Hummels–Levinsohn regression, even though there is by assumption no relation whatsoever between country similarity and trade to GDP, i.e $\beta=0$.

Appendix B. Replicating Hummels and Levinsohn fixed effect regression:

$$\ln \mathrm{VT}_{ijt} = \alpha_{ij} + \beta \ln\left(\mathrm{sim}_{ijt} Y_{ijt}\right) + \varepsilon_{ijt} \text{ for OECD}$$

$$\mathrm{VT}_{ijt} = \alpha_{ij} + \beta\left(\mathrm{sim}_{ijt} Y_{ijt}\right) + \varepsilon_{ijt} \text{ for non-OECD}$$

OECD (in logs)	HL (1995)	IMF	PennW
Estimated b coefficient	1.405	1.217	1.361
t-Statistic	110.8	154.5	124.2
R^2	0.865	0.932	0.899
Obs.	2002	1820	1820
Non-OECD (in levels)	**HL (1995)**	**IMF**	**PennW**
Estimated b coefficient	0.00157	0.00268	0.0013
t-Statistic	24.4	23.25	20.6
R^2	0.304	0.305	0.25
Obs.	1456	1320	1320

Appendix C. The list of countries in the data set

South Africa	Jamaica	Australia
Algeria	Trinidad–Tobago	New Zealand
Morocco	Panama	Fiji
Sudan	Surinam	
Tunisia	Israel	
Egypt	Japan	
Cameroon	Iran	
Gabon	Iraq	
Angola	Jordan	
Gambia	Kuwait	
Ghana	Oman	
Guinea	Saudi Arabia	
Madagascar	Turkey	
Malawi	Myanmar	
Mali	Hong Kong	
Mauritania	India	
Mauritius	Indonesia	
Mozambique	Korea	
Nigeria	Malaysia	
Guinea–Bissau	Pakistan	
Senegal	Philippines	
Sierra Leone	Singapore	
Somalia	Thailand	
Zimbabwe	Taiwan	
Togo	China	
Zambia	Belgium–Luxembourg	
Canada	Denmark	
USA	France	
Argentina	Germany	
Bolivia	Greece	
Brazil	Ireland	
Chile	Italy	
Colombia	Netherlands	
Ecuador	Portugal	
Mexico	Spain	
Paraguay	United Kingdom	
Peru	Austria	
Uruguay	Finland	
Venezuela	Iceland	
Costa Rica	Norway	
El Salvador	Sweden	
Guatemala	Switzerland	
Honduras	Malta	
Nicaragua	Czechoslovakia	
Barbados	Hungary	
Dominican RP	Poland	
Haiti	Romania	

References

Baier, S., Bergstrand, J., 2001. The growth of world trade: tariffs, transport costs, and income similarity. Journal of International Economics 53, 1–27.

Balassa, B., 1966. Tariff reduction and trade in manufactures among the industrial countries. American Economic Review 56, 466–473.

Davis, D., 1995. Intra-industry trade: a Heckscher–Ohlin–Ricardo approach. Journal of International Economics 39, 201–225.

Davis, D., 1997. Critical evidence on comparative advantage, North–North trade in a multilateral world. Journal of Political Economy 105, 1051–1160.

Deardorff, A., 1998. Determinants of bilateral trade: does gravity work in a neoclassical world? In: Frankel, J.A. (Ed.), The Regionalization of the World Economy. University of Chicago Press, Chicago, pp. 7–28.

Evenett, S., Keller, W., 2002. On theories explaining the success of the gravity equation. Journal of Political Economy 110, 281–316.

Feenstra, R., Lipsey, R., Bowen, H.P., 1997. World trade flows, 1970–1992, with production and tariff data. National Bureau of Economic Research Working Paper No. 5910, CD-ROM.

Feenstra, R., Markusen, J., Rose, A., 2001. Using the gravity equation to differentiate between alternative theories of trade. Canadian Journal of Economics 34, 430–447.

Grubel, H.G., Lloyd, P.J., 1975. Intra-Industry Trade: The Theory and Measurement of International Trade in Differentiated Products. Wiley, New York.

Helpman, E., 1981. International trade in the presence of product differentiation, economies of scale and monopolistic competition: a Chamberlin–Heckscher–Ohlin approach. Journal of International Economics 11, 305–340.

Helpman, E., 1987. Imperfect competition and international trade: evidence from fourteen industrial countries. Journal of the Japanese and International Economies 1, 62–81.

Helpman, E., 1998. The structure of foreign trade. National Bureau of Economic Research Working Paper No. 6752.

Helpman, E., Krugman, P., 1985. Market Structure and Foreign Trade. Increasing Returns, Imperfect Competition and the International Economy. MIT Press, Cambridge.

Hummels, D., Levinsohn, J., 1995. Monopolistic competition and international trade: reconsidering the evidence. Quarterly Journal of Economics 110, 799–836.

Krugman, P., 1979. Increasing returns to scale, monopolistic competition and international trade. Journal of International Economics 9, 469–479.

Krugman, P., 1980. Scale economies, product differentiation and the pattern of trade. American Economic Review 70 (5), 950–959.

Lancaster, K., 1980. Intra-industry trade under perfect monopolistic competition. Journal of International Economics 10, 151–175.

Tybout, J., 2001. Plant- and firm-level evidence on 'new' trade theories. Mimeo, Pennsylvania State University Working Paper.

Verdoorn, P.J., 1960. Intra-block trade of benelux. In: Robinson, E.A. (Ed.), Economic Consequences of the Size of Nations. Macmillan, London.