Distance decay in international trade patterns: a meta-analysis

Gert-Jan M. Linders
Department of Spatial Economics, Vrije Universiteit, De Boelelaan 1105, 1081 HV and Tinbergen Institute, Amsterdam, The Netherlands, Tel: +31 20 444 6106
Fax: +31 20 444 6004, Email: glinders@feweb.vu.nl.

Abstract
There is an irrevocable impact of distance on the patterns of international trade. The literature on bilateral trade identifies various factors that can explain the impact of geographic distance, ranging from transport costs to cultural familiarity. This paper performs a meta-analysis of existing empirical studies of bilateral trade, in order to contribute to our understanding of the distance barrier to trade. We focus on two key issues. First, we identify whether the impact of distance varies with time. Second, we investigate whether the estimated distance effect is affected by omitted variables bias with respect to other explanatory variables.

JEL codes: C19; F14; F15

Keywords: meta-analysis, bilateral trade, gravity model, distance decay

1. Introduction

Distance matters for trade. The impact of distance on bilateral trade reflects various distance-related trade costs. Besides transport costs of trading across distance, distance also leads to intangible trade costs, or transaction costs. Because of geographical information asymmetries, cultural familiarity decreases with distance, which raises the cost of trade. Rauch (2001) and Deardorff (2004) stress the key importance of transaction costs for understanding the relatively low intensity of foreign trade, also known as the 'mystery of missing trade' (Trefler, 1995). The empirical evidence on trade patterns suggests that border effects and geographical distance remain important sources of trade costs (Anderson and Van Wincoop, 2004). However, it has become common belief that advances in technology have vastly reduced the costs of transport and communication over the past decades (see Frankel, 1997, for an overview of the arguments). We will investigate whether there is evidence of such a development, as reflected in a downward trend of the distance decay in trade. This paper provides a meta-analysis of the distance effect in international trade to provide further insight into the determinants of the patterns of international trade.

A large number of primary studies has already investigated bilateral trade flows and identified a prominent role for geographical distance in accounting for the variation in trade flows. However, each of these primary studies has a somewhat different focus. Some studies focus on the impact of geographical distance. Others are motivated by a specific issue in international trade, ranging from the impacts of linguistic similarity, colonial ties or regional trade integration on trade, up to the effects of currency unions on trade and growth. Some studies focus on the investigation of the difference between domestic and foreign trade: so called border-effects; others intend to assess cross-country variation in bilateral trade. The methods and specifications used also vary widely. All of these studies, however, in some way consider the impact of transaction costs on international trade. Thus, they provide us with a valuable pool of information on the issue that we want to investigate.

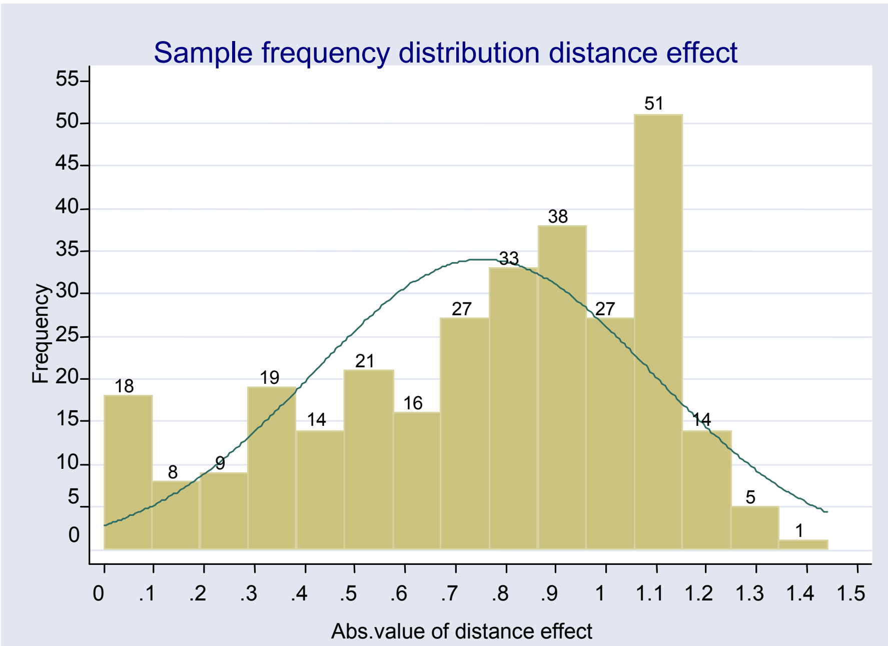
We need a way to systematically integrate the available evidence in the literature. The methods of meta-analysis are well suited to draw conclusions about the effect of geographical distance on trade flows, from the body of empirical studies performed until now. Glass (1976) stated that 'meta-analysis refers to the statistical analysis of a large collection of results from individual studies for the purpose of integrating the findings'. In other words, meta-analysis is a statistical analysis of existing empirical results on a specific topic. A meta-analysis starts from the observed cross-study variation in effect size estimates, and tries to account for this variation with differences in research design and specification. Meta-analysis has been widely used in social and experimental sciences, such as psychology, education and medicine. Recently, it has increasingly been applied in several fields of economics. Some references to meta-analysis performed in economics are Card and Krueger (1995), Smith and Huang (1995), and Görg and Strobl (2001). Given the widespread application of the gravity model to investigate the patterns of international trade, meta-analysis can be very useful as a quantitative review of existing evidence to acquire new insights (see, e.g. Stanley, 2001). Recently, Rose (2002) performed meta-analysis to summarize the current state of affairs regarding the impact of common currencies on international trade. We will use meta-analysis to statistically investigate variation in empirical evidence on distance decay that emerges from the literature on bilateral trade. Using meta-analysis, we can make statistically supported inferences about the impact of distance on trade and its development in time.

This paper is organized as follows. In section 3, we will describe the sample of studies reviewed for our purposes. In section 4, we discuss the meta-regression analysis. Finally, in section 5, we summarize the inferences that can be made.

2. Descriptive statistics

To compare the distance decay estimates across the literature, we collected a sample of studies from the literature. An Econlit search on the keywords ("trade" or "distance") and "gravity" provided a list of papers on the gravity model that should be sufficiently representative of the literature concerned. We extended the list with some additional gravity studies identified from references in the literature (this is sometimes referred to as the "snowballing" technique). To ensure comparability, we only consider papers that are concerned with aggregate bilateral merchandize trade. Therefore, we do not include papers that focus on sectors or specific industries. Only papers in English are taken into account. This generated a list of more than 200 studies, 19 of which were randomly sampled for consideration in the meta-analysis. In total, 290 estimates of the distance decay parameter were collected from the primary studies.[1] A histogram of the observations collected from the sampled studies (Figure 1) shows the distribution of estimates. A normal density curve based on the mean and standard deviation in the sample has been plotted as well. A concentration of estimates above the mean emerges, indicating that the distribution is skewed.

Figure 1. Histogram of distance decay estimates.



[1] It turns out that 8 observations in our sample, based on a semi-log specification of the gravity equation, are not comparable to the double-log elasticities. We omit these from the sample, together with some observations that were explicitly identified as unreliable in the primary studies. This leaves a sample of 290 observations.

Table 1 below illustrates the sample of studies.The table indicates that there is substantial variation of the distance effect across the literature. The average distance elasticity in the sample is about 0.78 (defined positively), which places it somewhat higher than the value of 0.6 claimed by Leamer and Levinsohn (1995) in their review of empirical evidence. The standard deviation is quite large.

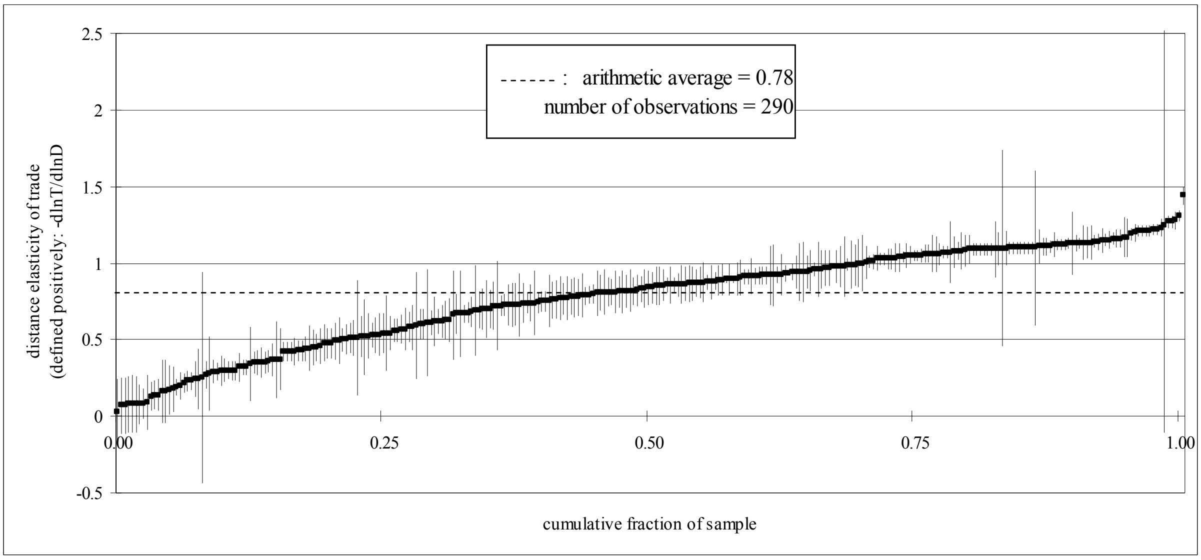
Table 1. Overview of the primary studies.

Study	Study-mean distance effect	Standard deviation around study mean	Number of estimates
1. Abraham et al. (1997)	0.46	0.21	4
2. Anderson and Marcouiller (2002)	1.11	0.02	8
3 .Van Bergeijk and Oldersma (1990)	0.99	0.09	4
4. Bougheas et al. (1999)	0.57	0.21	10
5. Dell'Ariccia (1999)	0.20	0.11	20
6. Djankov and Freund (2002)	0.84	0.32	12
7. Eichengreen and Irwin (1998)	0.54	0.26	31
8. Endoh (2000)	0.79	0.15	25
9. Frankel and Rose (2002)	1.11	0.07	5
10. Freund and Weinhold (2000)	0.84	0.21	49
11. Frankel, Stein and Wei (1995)	0.52	0.08	20
12. Hejazi and Safarian (2001)	0.70	0.23	18
13. Jakab et al. (2001)	0.94	0.05	6
14. Nilsson (2002)	0.64	0.17	14
15. Oguledo and MacPhee (1994)	0.51	0.25	3
16. Raballand (2003)	1.21	0.15	5
17. Rose (2000)	1.07	0.19	52
18. Rose (2004)	1.21	0.08	6
19. Wall (1999)	0.94	-	1
Total	0.78	0.32	290

Note: the distance decay elasticity of trade has been defined positively.

The question arises whether the non-normal distribution in figure 1 indicates that the sample is not sufficiently representative, or that the underlying population effect is not homogeneous. The extent of variation in the distance effect is effectively illustrated graphically in figure 2. All individual estimates are ordered according to their size. The vertical lines indicate the 95 percent confidence interval around the estimate that is denoted by a solid square. The overall mean distance effect in the sample is plotted as a broken horizontal line. The extent of heterogeneity of the estimates is shown clearly by the fact that the mean estimate does not enter into most confidence intervals of the individual estimates. This suggests that the variation in distance decay across the sample is not only due to the sampling error with which each observation is estimated. The pattern points out that the estimates of distance decay are heterogeneous.

Figure 2. Forest plot of individual estimates and confidence intervals.



A formal test of heterogeneity that is standard in meta-analysis in the medical field, can be applied to our meta-sample as well. The test compares the individual estimates to the pooled effect size. This fixed effect average is the weighted average effect size, where the weights are equal to the inverse of the sampling variance of the individual estimate.
The pooled effect size is given in equation (1):

$$\overline{D} = \frac{\sum_{i=1}^{n} w_i D_i}{\sum_{i=1}^{n} w_i} \tag{1}$$

in which D_i stands for the estimated distance effect size in primary estimation i, and w_i is the weight attached to the observation (Sutton et al.2000). We focus on the pooled effect size using random effects to capture variation between estimates. This allows for variance in the underlying (true) distance decay of each estimate to capture heterogeneity. The random effect thus defines the weight w_i as:

$$w_i = \frac{1}{\tau^2 + v_i} \tag{2}$$

Here, the variance of estimate i equals the estimation variance v_i and an additive term to capture the residual random effect variance τ^2 in the underlying effect.

We test for heterogeneity between estimates in the final sample. Starting from the pooled estimate and the individual estimates, the test variable (Q) directly follows (see Sutton et al., 2000):

$$Q = \sum_{i=1}^{n} w_i \left(D_i - \overline{D}\right)^2 \tag{3}$$

The assumption that each individual estimate reflects the same underlying effect size is rejected if the Q-statistic exceeds the critical value of a $\chi^2(n-1)$ distribution, for (n-1) degrees of freedom. The Q-test is significant with a p-value of 0.000. The Q-value found gives a strong indication of heterogeneity in the primary effect sizes. Consequently, the hypothesis of homogeneity is rejected. This confirms the need for a meta-regression analysis to find the determinants of the variation in the distance effect.

3. Meta-regression analysis

Meta-regression analysis extends the fixed and random effects approach to account for heterogeneity between estimates. A meta-regression model assumes that all variation between primary estimates, apart from the share attributable to sampling error, can be explained by covariates that reflect differences in estimation, data type or specification between studies. Since this model assumes there is no residual random effects variance between estimates, it is also known as fixed-effects regression. The meta-regression model can be estimated by a standard OLS or WLS approach. Alternatively, a mixed effect model (random-effects regression) allows for random effects variance in the error term, after controlling for study characteristics. The mixed model is generally preferred in meta-regression analysis, because the covariates cannot be expected to capture all heterogeneity between the primary estimates.[2]

Table 2 presents an overview of the different categories and groups of moderator (regressor) variables that we distinguish in the meta-regressions. Most of the moderator variables that we distinguish are dummy variables, indentifying specific groups of distance decay estimates. The exception to that is the time trend variable. The first category of covariates captures the estimation and general specification characteristics of the gravity equation estimates of distance decay. Secondly, the meta-regression incorporates moderators that indicate the presence of several control variables in the primary specification. Finally, spatial and temporal specifics of the estimates can be controled for in the meta-specification.

The choice of moderator variables is foremost motivated by theory. For example, econometric considerations suggest that estimates of the distance effect may vary depending on the use of instrumental variables to correct for endogeneity problems in the primary-study gravity equation. Also, Tobit regression estimates explicitly account for the censored nature of bilateral trade (zero is the lower limit of trade). In some other cases, there is no a-priori reason to expect a systematic difference in mean effects, but rather it is interesting to find out whether an effect exists. The most prominent example is the time trend effect in distance decay. Theoretical considerations may explain a trend in either direction. The meta-regression analysis can serve to help clarify the empirical question thus posed.

---

[2] See Sutton et al. (2000) for an overview of regression models for meta-analysis.

Table 2. Summary of moderator variables in the meta-regression analysis.

Categories of moderator variables	Sub-categories of moderator variables	Moderator variable name	Description	Hypothesized effect on |distance parameter|	No. of observations for moderator-dummy=1
Estimation and specification characteristics	Estimators	GLS dummy	1 if GLS-type estimator; 0 otherwise	?	40
	(OLS is base group: 215 observations)	Tobit dummy	1 if censored or outlier consistent; 0 otherwise	+	42
		ML/GLM estimators dummy	1 if ML, GLM; 0 otherwise	?	4
	Instrumental variables	IV dummy	1 if instrumental variables used; 0 otherwise	+/-	21
	Operationalization dependent variable	Export dummy	1 if export is dependent variable gravity equation	0	75
	(bilateral trade is base group: 155 observations)	Import dummy	1 if import is dependent variable gravity equation	0	45
		Transformed bilateral trade dummy	1 if (1+bilateral trade) is dependent variable	?	26
	Model specification	Equilibrium dummy	1 if based on theoretically derived gravity equation; 0 if empirical gravity equation	+/-	42
		Zero flows dummy	1 if zero flows included in sample; 0 otherwise	+	43
		Lagged trade dummy	1 if lagged trade regressor included; 0 otherwise	-	21
	Data type	Panel data dummy	1 if panel data; 0 if cross-section data	?	135
Primary study control variables		RIA dummy	1 if FTA/regional group included; 0 otherwise	-	249
		Language/colonial links dummy	1 if language/colonial ties included; 0 otherwise	+	198
		Contiguity dummy	1 if adjacency included; 0 otherwise	-	251
Spatial and temporal characteristics	Spatial categorization	U.S. trade patterns dummy	1 if gravity model applied to US trade patterns	?	20
		European trade patterns dummy	1 if gravity model applied to European trade patterns	?	49
		USSR trade patterns dummy	1 if gravity model applied to intra-USSR trade before 1991	-	4
	Temporal categorization	Great Depression dummy	1 if estimate applies to a year between 1930-1939	-	4
		post-WW2 reconstruction dummy	1 if estimate applies to a year between 1946-1950	+	9
		Time trend	0 (for 1904) – 95 (for 1999)	?	-

4.1 Basic empirical results

Table 3 presents the meta-regression results for a set of general estimation and specification moderator variables. The distance decay elasticity has been defined positively, to facilitate the interpretation of the impact on the size of the distance effect. The first output column shows results for a weighted least squares regression, with the inverse estimation standard error as weight. This is an example of fixed-effects regression, attributing all the variation in the distance effect (apart from estimation variance) to the moderator covariates. Because the estimation variances differ across estimates (see figure 2), the use of weights is necessary in the regression to correct for heteroskedasticity. The other output columns are estimated by random-effects regression, using the inverse estimation variance as weight.

Table 3. Meta-regression of distance decay: estimation and specification characteristics.

	(1) WLS	(2) mixed effect (ME) model	(3) Lagged trade control (ME)	(4) drop RE-estimates (ME)	(5) Preferred model (ME)
Panel data dummy	0.205	0.107	0.091	0.095	0.099
	[0.043]**	[0.040]**	[0.038]*	[0.037]**	[0.034]**
Export dummy	-0.073	-0.014	-0.017	-0.061	
	[0.049]	[0.042]	[0.040]	[0.040]	
Import dummy	-0.113	-0.021	-0.022	-0.071	
	[0.099]	[0.067]	[0.065]	[0.063]	
Transformed bilateral	-0.571	-0.605	-0.393	-0.418	-0.329
trade dummy	[0.116]**	[0.120]**	[0.123]**	[0.117]**	[0.098]**
GLS dummy	-0.062	-0.379	-0.398	-0.239	-0.284
	[0.044]	[0.061]**	[0.059]**	[0.070]**	[0.054]**
Tobit dummy	0.024	0.080	0.068	0.049	
	[0.066]	[0.088]	[0.084]	[0.080]	
ML/GLM estimators	0.128	0.269	0.261	0.258	0.261
dummy	[0.149]	[0.160]	[0.154]	[0.148]	[0.148]
IV dummy	-0.202	0.026	0.075	-0.064	
	[0.078]**	[0.081]	[0.078]	[0.082]	
Zero flows dummy	0.181	0.253	0.247	0.283	0.261
	[0.067]**	[0.087]**	[0.083]**	[0.080]**	[0.074]**
Equilibrium dummy	0.090	0.028	0.027	0.009	
	[0.092]	[0.066]	[0.064]	[0.062]	
Lagged trade dummy			-0.383	-0.368	-0.374
			[0.077]**	[0.074]**	[0.073]**
Constant	0.824	0.792	0.811	0.832	0.807
	[0.040]**	[0.030]**	[0.029]**	[0.028]**	[0.023]**
Observations	290	290	290	279	279
Adjusted R-squared	0.28	-	-	-	-

Standard errors in brackets; * significant at 5%, ** significant at 1%. Dependent variable: distance elasticity (defined positively). Semi-log elasticities excluded for lack of comparability. WLS regression uses the inverse of primary estimate standard errors as weights. Mixed effect models are indicated by the acronym ME and use the inverse of primary estimate variances as weights. RE estimates that fail Hausman test have been excluded in specification (4) and (5).

As already indicated, the literature on meta analysis argues that the use of mixed models is generally preferred to fixed-effects meta regression. Still, the parameter estimates of the WLS meta-regression model and the mixed model are qualitatively comparable for most moderator variables. The theoretical expectations concerning the use of Tobit and the inclusion of zero flows are partly confirmed. The distance effect is substantially larger in samples that include zero flows. This suggests that censoring of the data causes a bias in the parameter estimate if not accounted for. However, the use of estimation techniques that explicitly deal with censored trade, captured by the Tobit dummy, does not have an additional significant impact, although the sign is as expected (see table 2). Compared to OLS estimates, GLS and IV estimates appear to be systematically lower. The WLS meta regression attributes the effect to the use of IV, whereas the mixed effect models show a significant effect for the GLS dummy instead. Because GLS includes several IV estimators, it may be difficult to disentagle the effect (see table 2).

With regard to model specification, recent literature has stressed the importance of deriving a reduced form gravity equation on the basis of theory, rather than using the traditional 'empirical' gravity equation (for an overview, see Anderson and Van Wincoop, 2004). In particular, as empirical gravity equations do not control for country-specific equilibrium effects, such as relative price or relative distance (captured, for example, in catch-all country-specific fixed effects), the estimates suffer from potential omitted variables bias. The Equilibrium dummy is included to reflect the importance of this type of omitted variables bias. The parameter estimate is never significantly different from zero, though. At least for the distance decay parameter, the bias from using the traditional empirical gravity equation does not seem to be important. Conform expectation, gravity equations that include lagged trade have a substantially and statistically lower distance decay elasticity (see column 3 to 5). As noted by Eichengreen and Irwin (1998), the parameters in such a gravity model become short-term effects (or direct impact effects) of the variable in question. For distance decay this means that, compared to a third country at closer distance, if a country at some distance from the importer opens up to trade, the direct effect of distance on their bilateral trade is measured by the distance decay elasticity, but the long-run equilibrium effect operates through lagged trade.

Table 3 also includes moderator dummies for several operationalization categories, as well as a dummy for data type. Theoretical considerations do not give a prior for the effect of data type. The expectation is that the operationalization of trade should in general not affect the parameter estimate of the distance effect. The results in table 3 confirm that the choice of import or export, rather than bilateral trade, has no statistical effect. However, gravity models based on transformed bilateral trade (see table 2) have a significantly lower distance parameter. Adding the moderator for lagged trade partly captures this effect (see output column 3), since a relatively high share of the 'lagged trade' observations coincide with the transformed dependent variable specification. Half of the effect remains, though, and is highly significant statistically. A theoretical explanation is that the transformed dependent variable leads to a gravity model that approximates a semi-log specification for small trade flows, while maintaining the double-log form. This implies that the distance parameter strictly speaking is not a double-log elasticity at low values of trade. Because of the different treatment of low trade flows, the estimated distance parameter might be expected to differ systematically. Without an apparent theoretical explanation, distance decay is significantly higher in panel data studies.

Output columns 4 and 5 move towards a preferred meta model of the impact of general estimation and specification characteristics on the size of distance decay. First, a number of observations that are based on random effects estimation in a panel setting (RE) are deleted from the meta sample. These observations failed the Hausman specification test of RE model consistency. This does not affect the model outcomes qualitatively. The final regression-output column presents a mixed-effect model that includes the moderators that showed significant effects.

4.2 The 'death of distance': re-birth after meta-regression?

The development over time of distance decay in trade is particularly interesting. Technological advances in transport and more recently in ICT, together with the advent of globalization, have led to the view that distance is rapdily becoming less important for economic activity. This has given rise to metaphors, such as the 'global village' and the 'death of distance' to describe the new economy of the late 20th and 21st century. This trend may show up in different ways with respect to international trade. First, trade has grown more rapidly than world output, especially in the second half of the last century (see Baier and Bergstrand, 2001, for a discussion of possible explanations). For our purpose, it is of interest how the decline in transport and other distance related costs would affect the distance elasticity of trade. If distance becomes less important to trade, one would expect that the distance decay parameter had declined in magnitude over time.

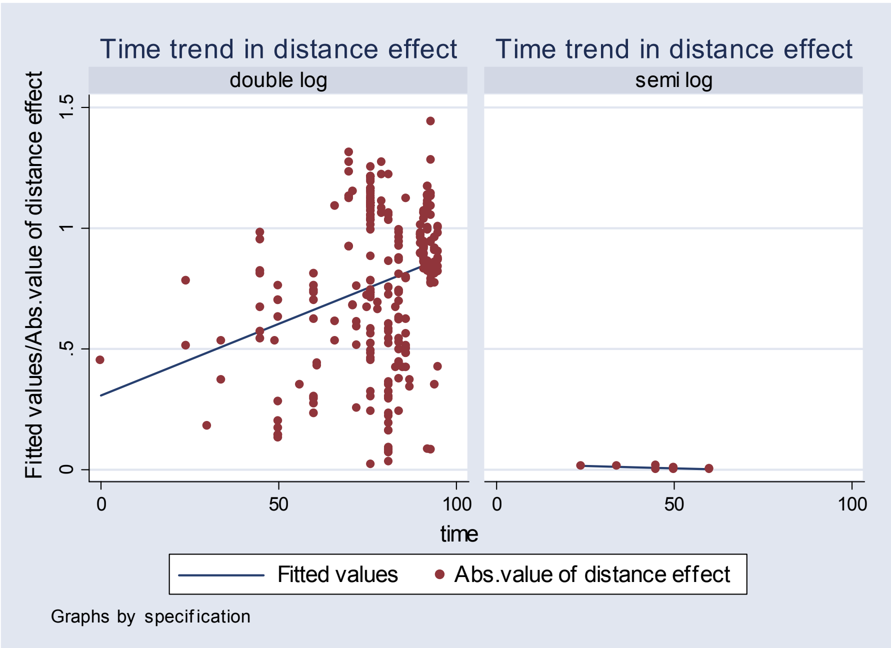
Several primary studies, however, have identified that the trend in the distance effect is not falling (e.g., Brun et al.,2002; see Frankel, 1997 for an overview). This has been identified as a puzzle, given the apparent fall in transport and communication costs (Leamer and Levinsohn, 1995; Wonnacott, 1998). Various explanations for this paradox have been suggested. Frankel (1997) argues that less tangible transaction costs associated with distance may be more important than transport costs. Alternatively, the lack of a clear trend in the distance effect has been attributed to the possibility that technological progress in transportation and communication has been "distance neutral" (Eichengreen and Irwin, 1998; Boisso and Ferrantino, 1997; Anderson and Van Wincoop, 2003). If transport costs fall proportionately irrespective of distance, trade would increase proportionately as well, leaving the distance parameter unaffected. This argument has been formally derived in Buch et al. (2004). However, as shown by Grossman (1998), starting form the assumption that transport costs increase with distance, a proportionate fall in transport costs at all distances will lower the relative import (c.i.f.) price of a product for long distance trade relative to short distance trade. This is due to the fact that transport costs are higher for long distance trade to begin with. Hence, if technological progress has reduced distance sensitive trade costs over the past century, we would generally expect the distance decay parameter to decline over time for two reasons.[3] First, when distance costs fall as a fraction of total costs, the rise in relative price as a result of a percentage increase in distance falls as well. The reaction in trade to a percentage

[3] Distance neutral technological progress only leaves the distance coefficient unaffected if imports are separable in f.o.b. prices and transport costs, as for example in Bröcker and Rohweder (1990).

increase in distance (the distant effect) will become less as a result (cf. Rauch, 1999). Second, transport costs will fall most for distant trade. Thus, the relative price of distant trade will fall vis-à-vis proximate trade, which reduces distance decay in trade. The fall in distance decay complements the overall rise in trade that follows from the decline of transport costs at all distances. That aspect of falling trade costs is indeed captured by the constant term in the gravity equation, as argued by Buch et al. (2004).

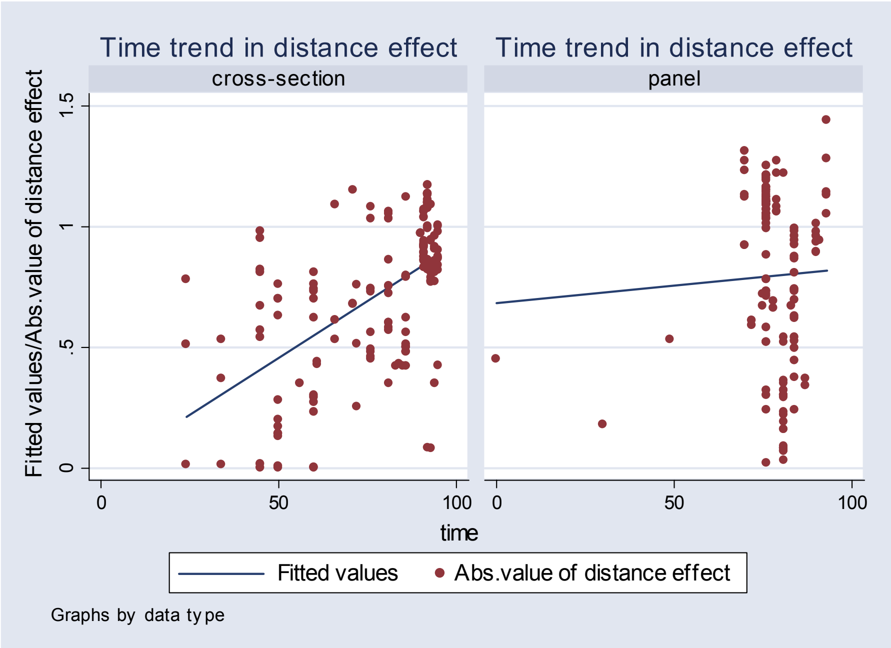
Meta-analysis can help to confront these hypotheses systematically with the empirical evidence across the literature. In the meta-sample, observations span from 1904 to 1999. Figures 3 and 4 provide a first impression of the development of distance decay over time. They present scatter plots of the observed distance parameters against the year of observation (0 stands for 1904).

Figure 3. Distance decay over time.



At first sight, distance decay appears to have increased over time. We focus on the double-log estimates in the left panel of figure 3 (for completeness, the semi-log estimates are presented in the right hand panel). The sample consists of both cross-section estimates and panel (pooled) estimates of distance decay. The pooled estimates apply to the range of years in the panel considered. We have assigned the estimate to the median year in the panel, as a proxy to compare different panels and cross-section estimates. Figure 4 shows that the general trend over time is positive for both sub-samples of estimates. However, most panels cluster substantially in later years of the time span. As a result, the time trend is more difficult to identify statistically in the panel sample.

Figure 4. Distance decay over time: panel versus cross-section estimates.



The descriptive analysis at least suggests that a negative trend in distance decay is not obvious. Further assessment requires multivariate analysis, controlling for other characteristics of the observations in the meta sample. Table 4 below presents meta-regression results that extend the preferred model of table 3 with a time trend variable. The evidence from the meta-analysis does not offer support for a 'death of distance' hypothesis. In fact, the importance of distance sensitive costs for trade even seems to increase somewhat over time. The time trend in the distance decay parameter is positive and significant in most meta-specifications in the table. The trend implies that, all else equal, the distance elasticity increases with 0.03 to 0.04 over the course of a decade. In other words, the representative decay in trade as a result of a 1% increase in bilateral distance has increased from 0.52% in 1904 (the value of the constant term, or base group of estimates, in the first model) to 0.81% in 1999, a rise of around 0.29 percentage points. This suggests that intangible costs of distance, relating to unfamiliarity with distant markets and cultures are still important and did not decrease over time, relative to other trade costs (e.g., Frankel, 1997). Furthermore, the sustained or rising importance of distance for trade could also be due to the fact that multilateral liberalization of trade has increased the relative importance of natural trade barriers imposed by distance, compared to policy induced trade costs. As a consequence, the distance effect might rise over time.[4] The common wisdom of falling (relative) transportation costs has been challenged as well. Hummels (1999), for example, provides evidence that long distance shipping has not become cheaper as a percentage of value added. All in all, the results suggest that distance has not become less important for trade over time.

[4] I am indebted to both Albert de Vaal and Erik Verhoef for suggesting this interpretation.

Table 4. Meta-regression of distance decay: time trend.

	(1)	(2)	(3)	(4)
	Mixed effect model	WLS	OLS	Spatial/temporal controls (ME)
Panel data dummy	0.115 [0.034]**	0.250 [0.033]**	0.108 [0.034]**	0.302 [0.028]**
Transformed bilateral trade dummy	-0.183 [0.115]	-0.127 [0.141]	-0.181 [0.116]	-0.095 [0.083]
GLS dummy	-0.248 [0.056]**	-0.117 [0.030]**	-0.238 [0.056]**	-0.008 [0.045]
ML/GLM estimators dummy	0.263 [0.147]	0.131 [0.146]	0.270 [0.127]*	0.103 [0.109]
Zero flows dummy	0.244 [0.073]**	0.150 [0.062]*	0.250 [0.075]**	0.129 [0.052]*
Lagged trade dummy	-0.397 [0.073]**	-0.450 [0.114]**	-0.392 [0.073]**	-0.416 [0.052]**
Time trend	0.003 [0.001]*	0.002 [0.002]	0.004 [0.001]*	0.004 [0.001]**
U.S. trade dummy				-0.407 [0.048]**
European trade dummy				-0.583 [0.041]**
USSR trade dummy				-0.401 [0.087]**
Great Depression dummy				-0.147 [0.115]
post-WW2 reconstruction dummy				0.266 [0.078]**
Constant	0.521 [0.123]**	0.618 [0.134]**	0.502 [0.124]**	0.448 [0.095]**
Observations	279	279	279	279
Adjusted R-squared	-	0.31	0.28	-

Standard errors in brackets; * significant at 5%; ** significant at 1%. Dependent variable: distance elasticity (defined positively). Semi-log elasticities excluded for lack of comparability. WLS regression uses the inverse of primary estimate standard errors as weights. Mixed effect models are indicated by the acronym ME and use the inverse of primary estimate variances as weights. RE estimates that fail Hausman test have been excluded.

Though random-effects meta regression is preferred, both WLS and OLS fixed-effects regression give a similar picture (see output colums 2 and 3). Remarkably, OLS results are very similar to random-effects regression, whereas WLS parameter-estimates differ more, both qualitatively and quantitatively. The model displayed in output column 4 adds several control variables to test whether the time trend (as well as the other moderators) is robust.[5] The moderators identify specific spatial or temporal characteristics of various groups of estimates. Several studies focus on European or U.S. trade patterns, which may differ systematically from global patterns. The remaining controls intend to capture findings reported in the primary studies with respect to distance decay. Trade between regions and republics in the former USSR, before the fall of communism, was determined more by central planning than by natural factors such as distance. Therefore, distance decay in that region was smaller before 1991 than thereafter. Trade in the period of the Great Depression was also characterized by policy more than by economic fundamentals. Heavy general protectionism,

[5] See table 2 for a description.

combined with specific preferences towards colonies and allies, reduced the effect of distance on trade in exchange for the effect of commercial and monetary initiatives (Eichengreen and Irwin, 1995). On the same note, the late 1940s had to cope with the breakdown of infrastructure, which impeded international trade, especially across longer distances. The distance effect may therefore have been more pronounced in that period. Especially the last three controls may affect the estimated time trend, because the effects are situated in early or late periods of the meta sample. The results show that the effect of the added controls is as expected. Moreover, both European and U.S. trade patterns are less sensitive to distance than average. The time trend is robust to these control variables, and becomes statistically more significant. The moderators on the operationalization of trade and on remaining estimators that were left from table 3 are no longer statistically significant at 10%.

So far, the meta-analysis has indicated that the variation in the distance effect across the literature is systematically related to differences in research design and estimation, and that the distance effect shows an increasing trend over time. However, a large part of variation in the distance parameter has not been captured by the moderator variables included so far. The fit of the model is not very high, at an adjusted R-squared of around 0.30. The next section extends the meta-model specification to include differences in primary study control variables.

4.3 Distance decay and omitted variables bias

Apart from variation across time and space and according to estimation characteristics, heterogeneity in distance decay may also reflect differences in the specification of the gravity equation explaining bilateral trade. We have already paid attention to primary study specification issues, when we considered the effect of a lagged trade regressor and equilibrium versus empirical gravity equations. Table 5 extends the meta-regression with moderators that reflect whether primary studies have included regressors that tend to be correlated with geographical distance. The argument is that distance effect estimates may suffer from omitted variables bias, if the gravity equation does not control for other explanatory characteristics that are correlated to distance. In this respect, the most promising control variables often included in the gravity literature are dummies for contiguity, common language and colonial ties, spatial regions and preferential trade agreements (see table 2).

Countries that share borders are geographically closer than two average non-bordering countries. As a result, distance and adjacency are correlated. This may lead to omitted variable bias in the distance effect. Generally, the effect of common borders on trade is estimated positive. This reflects both measurement error in the distance variable (the distance of trade between bordering countries is often relatively overestimated), as well as stronger historical links and trade concentration along borders. We expect that the distance estimate will be lower once contiguity has been controlled for. This is corroborated by the results in table 5, although the effect only becomes significant in the mixed model once the spatial and temporal controls have been included.

Table 5. Meta-regression of distance decay: omitted variables heterogeneity.

	(1)	(2)	(3)	(4)	(5)
	WLS	Mixed effect	ME: Time trend	WLS: Time trend	ME: controls
RIA dummy	-0.184 [0.061]**	-0.068 [0.052]	-0.079 [0.052]	-0.182 [0.060]**	-0.150 [0.038]**
Language/colonial links dummy	0.382 [0.058]**	0.204 [0.043]**	0.174 [0.046]**	0.430 [0.061]**	0.265 [0.034]**
Contiguity dummy	-0.287 [0.066]**	-0.090 [0.046]	-0.073 [0.047]	-0.327 [0.067]**	-0.314 [0.037]**
Lagged trade dummy	-0.424 [0.091]**	-0.359 [0.064]**	-0.344 [0.064]**	-0.450 [0.091]**	-0.478 [0.047]**
Time trend			0.002 [0.001]	-0.003 [0.001]*	-0.000 [0.001]
U.S. trade dummy					-0.292 [0.047]**
European trade dummy					-0.515 [0.037]**
USSR trade dummy					-0.576 [0.098]**
Great Depression dummy					-0.362 [0.119]**
Reconstruction dummy					0.219 [0.081]**
Constant	1.052 [0.053]**	0.825 [0.053]**	0.661 [0.105]**	1.316 [0.121]**	1.152 [0.090]**
Observations	279	279	279	279	279
Adjusted R-squared	0.26			0.27	

Standard errors in brackets; * significant at 5%; ** significant at 1%. Dependent variable: distance elasticity (defined positively). Semi-log elasticities excluded for lack of comparability. WLS regression uses the inverse of primary estimate standard errors as weights. Mixed effect models are indicated by the acronym ME and use the inverse of primary estimate variances as weights. RE estimates that fail Hausman test have been excluded.

Language and colonial ties reduce trade costs. Countries that share these cultural and historical links tend to trade disproportionately despite generally large distances. The correlation with distance is a source of bias to the distance decay parameter. Therefore, we expect distance decay to rise once these historical links have been accounted for. Because observations on both control variables tend to go together in the meta-sample, we have grouped them in one moderator dummy. The results in table 5 confirm the expectations. The effect on distance decay is substantial and statistically significant at 1%.

Membership in common preferential trade blocs is generally concentrated in space. Therefore, the distance effect captures the effect of regional integration if not explicitly identified in the primary specifcation. Once controlled for this source of omitted variables bias, the distance decay is expected to fall. We combined both formal PTA/FTAs and informal regional groups into one moderator. Regional group dummies often reflect regional concentration of trade that is not captured by standard gravity variables. Because regional group membership is defined by some notion of geographical proximity, the same general argument holds as for FTAs. The regression results support the expectation that distance decay falls once regional integration is accounted for in the gravity equation. Once again, the extended meta model in output column 5 gives the most explicit result. It is noteworthy that the time trend is less pronounced in the mixed models of table 5, lacking statistical significance, while the trend turns negative and significant for the WLS fixed-effect regression.

It is interesting to see the outcomes for a more complete meta-model, as the overall fit of the model is again rather low at 0.27 (for WLS). Table 6 shows a complete meta model that summarizes the effect of different types of characteristics to account for variation in the distance effect. The complete model significantly improves the overall fit. Around 70% of the variation in the distance effect can be explained (OLS). The time trend is once again firmly positive in both OLS and mixed effect models, although the conclusions for WLS regression remain opposite. Most variables of influence that were identified before are robust to the model extension. All are highly significant in the preferred mixed effect model. As an exception, the distance effect no longer significantly depends on the inclusion of zero-trade flows. This provides support for the statement that zero flows do not pose severe problems for gravity analysis in practice (Frankel, 1997; Baldwin, 1994). The results confirm that the operationalization and remaining estimator dummies can be dropped once the time trend has been included. Moreover, the GLS dummy looses significance in the preferred specification.

Table 6. Meta-regression of distance decay: complete model.

Category of moderators	Moderator variable	(1)	(2)	(3)	(4)
		WLS	Mixed effect	OLS	Preferred ME
Data type	Panel data dummy	0.198 [0.022]**	0.246 [0.027]**	0.248 [0.028]**	0.242 [0.027]**
Operationalization	Transformed bilateral trade dummy	-0.030 [0.089]	0.122 [0.081]	0.114 [0.084]	
Estimators	GLS dummy	-0.065 [0.018]**	0.109 [0.044]*	0.114 [0.047]*	0.084 [0.043]
	ML/GLM estimators dummy	0.082 [0.083]	0.113 [0.098]	0.154 [0.083]	
Model specification	Zero flows dummy	0.021 [0.043]	0.013 [0.054]	0.017 [0.057]	
	Lagged trade dummy	-0.460 [0.065]**	-0.449 [0.047]**	-0.439 [0.048]**	-0.396 [0.041]**
Temporal categorization	Time trend	-0.003 [0.001]**	0.003 [0.001]**	0.004 [0.001]**	0.002 [0.001]**
Primary specification control variables	RIA dummy	-0.060 [0.037]	-0.074 [0.034]*	-0.076 [0.035]*	-0.086 [0.033]**
	Language/colonial links dummy	0.227 [0.040]**	0.180 [0.033]**	0.173 [0.034]**	0.167 [0.033]**
	Contiguity dummy	-0.333 [0.045]**	-0.289 [0.037]**	-0.278 [0.037]**	-0.273 [0.035]**
	Constant	1.251 [0.094]**	0.741 [0.094]**	0.722 [0.097]**	0.844 [0.084]**
Observations		279	279	279	279
Adjusted R-squared		0.78		0.70	

Standard errors in brackets; * significant at 5%; ** significant at 1%. Dependent variable: distance elasticity (defined positively). Semi-log elasticities excluded for lack of comparability. WLS regression uses the inverse of primary estimate standard errors as weights. Mixed effect models are indicated by the acronym ME and use the inverse of primary estimate variances as weights. Potentially biased RE estimates are excluded, as before. Spatial and temporal control moderators have been included as before, but their parameters are not reported here for the sake of brevity.

4. Conclusions

This paper has illustrated that the empirical evidence about the effect of geographical distance on aggregate trade flows is heterogeneous. The distance effect on trade varies according to the estimation and specification characteristics of the primary studies. Meta-regression analysis shows that most variation in distance decay across the literature can be explained by a few characteristics of the primary specification and estimation.

In short, when focusing on statistically significant effects, the following patterns have been identified. Panel studies generally produce a larger distance decay parameter. Models that include a lagged trade regressor as explanatory variable result in lower distance elasticities. This reflects that the elasticity in these models is a short run, rather a than long run effect. The omission of several bilateral variables from the gravity model leads to substantial omitted variables bias in the distance parameter. These variables are: membership in a common trade bloc or regional grouping, common language or colonial ties and contiguity of countries. On the other hand, the use of empirical gravity specifications, rather than consistent theoretical foundations, does not appear to bias the estimated distance effect. Neither does the omission of zero flows from the regression sample systematically affect the estimated distance coefficient in gravity models of aggregate bilateral trade.

Most estimates of the trend in distance decay indicate that the effect of distance on trade has increased over time. This suggests that distance has not become less important for trade costs, despite the popular notions of the 'death of distance' and declining transport and communication costs. This finding supports the view that distance costs are multidimensional and that less tangible, informal distance barriers remain important.

Reference List

Anderson, J.E. and E. Van Wincoop (2003): 'Gravity with Gravitas: A Solution to the Border Puzzle', *American Economic Review*, 93 (1), pp. 170-192.

Anderson, J.E. and E. Van Wincoop (2004): 'Trade Costs', *Journal of Economic Literature*, 42 (September), pp. 691-751.

Boisso, D. and M. Ferrantino (1997): 'Economic Distance, Cultural Distance, and Openness in International Trade:Empirical Puzzles', *Journal of Economic Integration*, 12 (4), pp. 456-484.

Brun, J.-F., C. Carrere, P. Guillaumont and J.d. Melo (2002): 'Has Distance Died? Evidence from a panel gravity model', *CEPR Discussion Paper*, 3500.

Buch, C.M., J. Kleinert and F. Toubal (2004): 'The Distance Puzzle: On the Interpretation of the Distance Coefficient in Gravity Equations', *Economics Letters (June)*, 83 (3), pp. 293-98.

Card, D. and A.B. Krueger (1995): 'Time-series minimum-wage studies: a meta-analysis', *American Economic Review*, 85, pp. 238-243.

Deardorff, A.V. (2004): 'Local Comparative Advantage: Trade Costs and the Pattern of Trade', *Research Seminar in International Economics Discussion Paper, University of Michigan*, 500.

Eichengreen, B. and D.A. Irwin (1995): 'Trade blocs, currency blocs and the reorientation of world trade in the 1930s', *Journal of International Economics*, 38, pp. 1-24.

Eichengreen, B. and D.A. Irwin (1998): 'The Role of History in Bilateral Trade Flows', in: J.A. Frankel (ed.), *The Regionalization of the World Economy*, pp. 33-57, Chicago and London: The University of Chicago Press.

Frankel, J.A. (1997): *Regional Trading Blocs in the World Economic System*, Washington D.C.: Institute for International Economics.

Glass, G.V. (1976): 'Primary, secondary and meta-analysis of research', *Educational Researcher*, 5, pp. 3-8.

Görg, H. and E. Strobl (2001): 'Multinational companies and productivity spillovers: a meta-analysis', *Economic Journal*, 111, pp. 723-740.

Hummels, D. (1999): 'Have International Transportation Costs Declined?', *Mimeo, University of Chicago*.

Leamer, E.E. and J. Levinsohn (1995): 'International trade theory: the evidence', in: G. Grossman and K. Rogoff (eds.), *The Handbook of International Economics, Vol. III* North-Holland: Elsevier.

Rauch, J.E. (1999): 'Networks versus markets in international trade', *Journal of International Economics*, 48, pp. 7-35.

Rauch, J.E. (2001): 'Business and Social Networks in International Trade', *Journal of Economic Literature*, 39, pp. 1177-1203.

Rose, A.K. (2002): *The effect of common currencies on international trade: a meta-analysis,* unpublished manuscript, forthcoming in: Monetary Unions and Hard Pegs: Effects on Trade, Financial Development and Stability, Alexander, V., Mélitz, J., and Von Furstenberg, G. M.

Smith, V.K. and J.-C. Huang (1995): 'Can markets value air quality? A meta-analysis of hedonic property value models', *Journal of Political Economy*, 103, pp. 209-225.

Stanley, T.D. (2001): 'Wheat from Chaff: Meta-analysis as Quantitative Literature Review', *Journal of Economic Perspectives*, 15 (3), pp. 131-150.

Sutton, A.J., K.R. Abrams, D.R. Jones, T.A. Sheldon and F. Song (2000): *Methods for Meta-analysis in Medical Research*, Wiley Series in Probability and Statistics, Chichester etc.: John Wiley & sons.

Trefler, D. (1995): 'The case of missing trade and other mysteries', *American Economic Review*, 85, pp. 1029-1046.