Economics Bulletin

Volume 36, Issue 3

Distance-sensitivity of German imports: First evidence from firm-product level data

Joachim Wagner
Leuphana University Lueneburg

Abstract

This paper uses a tailor-made new data set of 3,376,598 observations for German imports at the firm-product-country of origin level to estimate a gravity equation and to investigate the link between the amount of firms' imports and the distance to countries of origin. It is shown that, in line with stylized facts based on aggregate data, the quantity of imports declines significantly with distance within a firm for a given product.

I thank an anonymous referee for helpful comments. All computations were done at the Research Data Centre of the Federal Statistical Office in Wiesbaden. I thank Melanie Scheller for preparing the transaction level data for imports and for checking the output of my do-files for the violation of privacy. The micro data used are strictly confidential but not exclusive; see http://www.forschungsdatenzentrum.de/datenzugang.asp for information on how to access the data. To facilitate replications the Stata do-file used is available from the author on request.
Citation: Joachim Wagner, (2016) "Distance-sensitivity of German imports: First evidence from firm-product level data", *Economics Bulletin*, Volume 36, Issue 3, pages 1275-1279
Contact: Joachim Wagner - wagner@leuphana.de.
Submitted: April 06, 2016. **Published:** July 08, 2016.

1. Motivation

The negative effect of distance on international trade qualifies as a stylized fact. As Leamer (2007, p. 110) put it, "(t)here is very little that we economists fully understand about global trade but there is one thing that we know – commerce declines dramatically with distance." This distance effect is due to trade costs that tend to increase with the distance between the country of origin and the country of destination, including transportation costs (freight costs and time costs), information costs, language barriers, different currencies, legal and regulatory costs, and cultural distance.

Empirical estimates of the distance-sensitivity of trade usually are based on estimated gravity equations[1] that regress the (log of) trade between two countries on the (log of) distance between them plus a number of other variables that control for the size of the countries, the level of per capita income, whether a country is landlocked, and whether the countries share a common language or a common border, among others. Gravity equations are usually estimated with aggregate data at the country or sectoral level. Trade costs, however, that are behind the negative distance effect, can be expected to vary to a large amount between heterogeneous firms engaged in trade. For example, firms differ with regard to intercultural competence of their managers, knowledge of foreign languages of employees and experience of doing business with various countries. Furthermore, as Anderson and van Wincorp (2004, p. 747) point out, trade costs also vary widely across products lines. Therefore, trade costs related to distance can be expected to vary not only between an importing country and different countries of origin, but also between the different products traded between two countries and between firms engaged in these transactions. Estimates of the distance-sensitivity of imports, therefore, should control for these firm specific and product specific effects by including firm-product fixed effects. Identification of the regression coefficient of the distance variable then comes from the within-firm-product variation of import values across countries of origin.

For exports, there are two studies that apply this empirical strategy. Bastos and Silva (2010) report results for Portugal that show that export quantities decrease with distance when identification comes from the within-firm-product variation of export quantities across destination. Wagner (2016) shows that this is the case for exports by German firms, too. To the best of my knowledge, comparable estimates of the distance-sensitivity of imports are missing. This paper contributes to the literature by reporting such estimates for Germany, one of the most important actors on the world market for goods.

The rest of the paper is organized as follows. Section 2 discusses the data and measurement issues. Section 3 presents the results of the empirical investigation. Section 4 concludes.

2. Data and measurement issues

The empirical investigation uses a new tailor-made data set that combines high quality import data from official statistics with data on the distance between Germany and countries of origin of imports plus other information for characteristics of these countries.

Data on imports are based on either the reports by German firms on transactions with firms from countries that are members of the European Union (EU) or on transaction-level

[1] A discussion of the gravity model and its relation to theoretical models of trade is far beyond the scope of this letter; see Head and Mayer (2014) for a comprehensive treatment. For a meta-analysis of 1,467 distance effects estimated in 103 papers see Disdier and Head (2008).

data collected by the customs on trade with countries outside the EU. The raw data relate to one transaction of a German firm with a firm located outside Germany at a time. For a given year, the sum over all import transactions is identical to the figures published by the Federal Statistical Office for total imports of Germany.

The record of the transaction usually includes a firm identifier (tax registration number) of the trading firm. Using this identifier information at the transaction level can be aggregated at the level of the trading firm to generate year-firm-product-value--country data. These data show who trades how much of which good with customers from which country in a given year. Products are distinguished here according to the Harmonized System at 6-digit level (HS6).This paper uses data for the reporting year 2011.

The dependent variable in the empirical models is the value of imports of a (HS6) good by a firm from a country of origin. The import value is regressed on the distance to the country of origin and on a set of control variables.

Data on *distance to import origin* between Germany and the countries of origin of imports are taken from the CEPII's *GeoDist* database (Mayer and Zignago 2011). The "distw" – measure is used that calculates the distance between two countries based on bilateral distances between the biggest cities of those two countries, those inter-city distances being weighted by the share of the city in the overall country's population (see Mayer and Zignago (2011, p. 11) for details).

The empirical models include a number of standard gravity variables as control variables (see Bustos and Silva 2010, and Wagner 2016).

Market size is proxied by the Gross Domestic Product (GDP) of the country of origin, measured in Millions of US-Dollar in current prices. Information is taken from the World Bank World Development Indicators database (see http://data.worldbank.org/indicator/NY.GDP.MKTP.CD).

GDP per capita is measured in current prices and U. S. dollars. Data are from the International Monetary Fund's World Economic Outlook Data Base, April 2012 edition (see https://www.imf.org/external/pubs/ft/weo/2012/01/weodata/index.aspx).

Landlocked is a dummy-variables that takes the value 1 if a country has no direct access to the sea. Information is taken from the CEPII's *GeoDist* database (Mayer and Zignago 2011).

Furthermore, two groups of trade partner countries are distinguished, namely countries that are *members of the European Union* (EU) and Non-EU countries. This controls for the different level of policy-made trade barriers that are absent for intra-EU trade.

3. Results

The empirical investigation uses 3,376,598 observations for the value of imports at the firm-product-country of origin level for the reporting year 2011. For trade with the EU-countries the number of observations is 1,332,507. Data are from 48,121 firms that imported 4,909 products (recorded at the HS6-level) in 795,228 firm-product combinations. Data for imports from non-EU countries are for 2,044,091 observations from 123,852 firms that traded 4,908 different goods in 1,432,235 firm-product combinations. A table with descriptive statistics for all variables is available on request.

To investigate the link between import quantity and the distance to the countries of origin separate regression models are estimated for observations on imports from other EU-countries and from non-EU countries with the log of import quantity as the endogenous variable and the log of distance to the country of origin as exogenous variable. Following Bustos and Silva (2010) and Wagner (2016), the model includes a set of control variables for

further characteristics of the country of origin (described in detail above), namely the log of GDP, the log of GDP per capita, and a dummy variable indicating whether the country is landlocked. To control for the role of firm and goods specific factors fixed effects at the firm-good level are included in the empirical models, too. Here, identification of the regression coefficient of the distance variable comes from the within-firm-product variation of import values across countries of origin. Results are reported in Table I.

The regression coefficient of log(distance) is an estimate of the elasticity of the quantity of firms' import with respect to distance to countries of origin. According to the results, therefore, an increase in distance by one percent leads to a decrease of import quantity by 0.271 percent in intra-EU trade and by 0.081 percent in extra-EU trade, controlling for the size (measured by GDP) and the wealth (measured by GDP per capita) of the country of origin and for the status of a landlocked country.[2] This means that doubling the distance is related to a decrease in imports by 27 percent inside the EU and by 8 percent outside the EU. Note that these estimates are considerably smaller than comparable estimates for the distance-sensitivity of exports reported in Wagner (2016), where the respective figures are 42 percent for exports inside the EU and 25 percent for extra-EU exports.

4. Concluding remarks

The bottom line, then, is that the link between the quantity of imports and the distance to countries of origin is negative within firm-product import flows to Germany. This empirical result is based on a gravity-type empirical model. As Anderson and van Wincoop (2003) point out in their seminal paper on gravity models, after controlling for size, trade between two countries is decreasing in their bilateral trade barrier relative to the *average* barrier of the two countries with all their trade partners. They refer to the theoretically appropriate average trade barrier as *multilateral resistance* (MR). Neglecting this issue leads to biased estimates of the coefficients of the variables in the gravity equation. Given that MR is unobserved, usually importer and exporter country – year fixed effects are included in the estimated equations as proxy variables for MR in order to obtain unbiased estimates (see e.g. Behar and Nelson (2014), p. 543). With the cross section firm level data from Germany used in my study this approach is not feasible.

It might be the case that firms from more remote foreign countries supply less of a given good to German importers because these foreign firms are closer to other potential markets. The reason for the decline of the quantity of imports with distance within a firm for a given product then is not only the larger distance between Germany and the country of origin, but the smaller distance between the country of origin and other potential markets, too. The central result of the paper, however, that the quantity of imports declines significantly with distance within a firm for a given product is still valid.

This new evidence is in line with stylized facts based on aggregate data and with evidence for exports based on firm-product-destination data by Bastos and Silva (2010) for Portugal and by Wagner (2016) for Germany. Distance is not dead at all when it comes to international trade in goods.

[2] Given that the control variables serve as controls only in the empirical models we do not comment on the results for the estimated regression coefficients.

5. References

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

Anderson, James E. and Eric van Wincoop (2004), Trade Costs. *Journal of Economic Literature* XLII (3), 691-751.

Bastos, Paulo and Joana Silva (2010), The quality of a firm's exports: Where you export to matters. *Journal of International Economics* 82 (2), 99–111.

Behar, Alberto and Benjamin D. Nelson (2014), Trade Flows, Multilateral Resistance, and Firm Heterogeneity. *Review of Economics and Statistics* 96 (3), 538-549.

Disdier, Anne-Célia and Keith Head (2008). The puzzling persistence of the distance effect on bilateral trade. *Review of Economics and Statistics* 90 (1), 37-48.

Head, Keith and Thierry Mayer (2014), Gravity Equations: Workhorse, Toolkit, and Cookbook. Handbook of International Economics, Vol. 4, 131-195.

Leamer, Edward E. (2007), A Flat World, a Level Playing Field, a Small World After All, or None of the Above? A Review of Thomas L. Friedman's *The World is Flat. Journal of Economic Literature* XLV (1), 83-126.

Mayer, Thierry and Soledad Zignago (2011), Notes on CEPII's distance measures: The GeoDist database. CEPII Document de Travail No 2011-25, December.

Wagner, Joachim (2016), Distance-sensitivity of German exports: Fist evidence from firm-product level data. *Applied Economics Letters* (in press).

Table I: Quantity of firms' imports and characteristics of countries of origin: Regression results

Characteristics of country of origin		Dependent variable: log (quantity of imports) EU-countries	Non-EU countries
log (distance)	ß	-0.271	-0.081
	p	0.000	0.003
log (GDP)	ß	0.249	0.325
	p	0.000	0.000
log (GDP per capita)	ß	-0.219	-0.077
	p	0.000	0.000
Landlocked (Dummy variable)	ß	0.134	0.705
	p	0.002	0.000
Constant	ß	-4.541	-10.015
	p	0.000	0.000
Fixed effects: Firms*Goods		yes	yes
R-squared		0.432	0.802
Number of observations		1,332,507	2,044,091

Note: OLS regressions. ß is the estimated regression coefficient, p is the prob-value (based on heteroscedasticity-consistent standard errors clustered at the level of the importing firm). For a detailed definition of the variables see text.