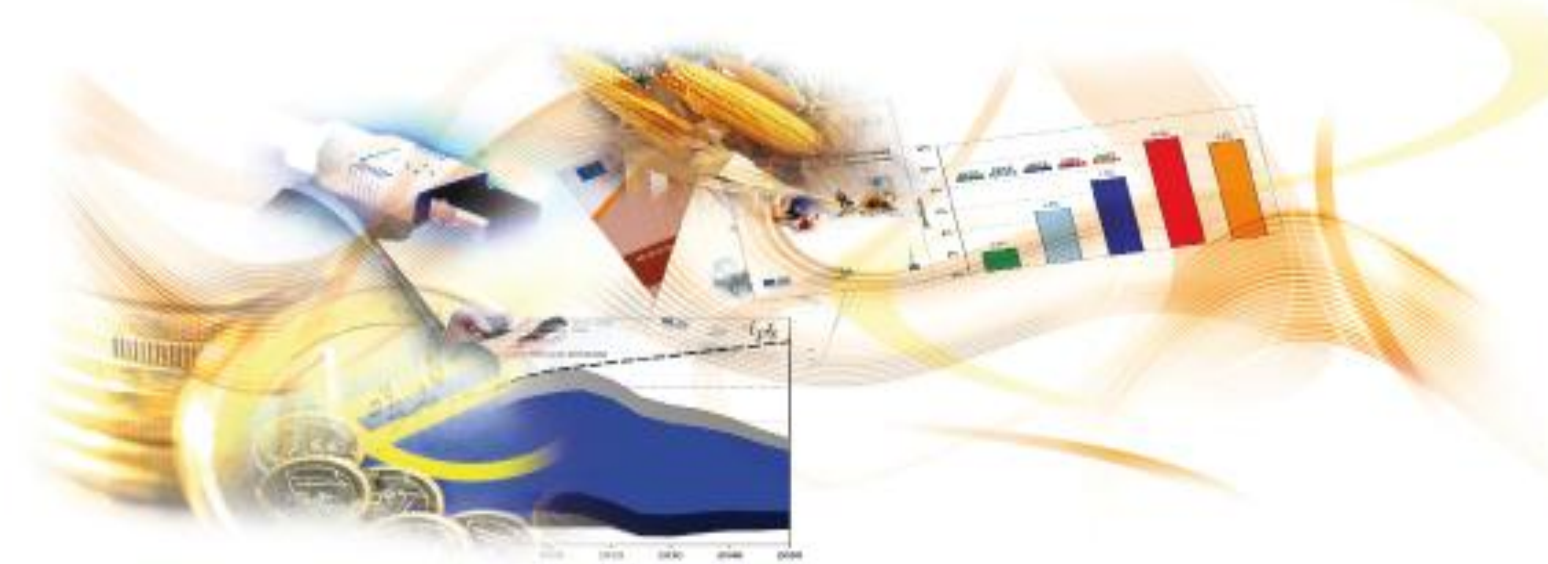
European
Commission



JRC TECHNICAL REPORTS

Institute for Prospective Technological Studies
Digital Economy Working Paper 2013/08

Does Online Trade live up to the Promise of a Borderless World? Evidence from the EU Digital Single Market

Bo Cowgill (University of California at Berkeley)
Cosmina Dorobantu (Oxford University)
Bertin Martens (JRC-IPTS, European Commission)

2013

Report EUR 26217 EN

Joint
Research
Centre

European Commission
Joint Research Centre
Institute for Prospective Technological Studies

Contact information
Address: Edificio Expo. C/ Inca Garcilaso, 3. E-41092 Seville (Spain)
E-mail: jrc-ipts-secretariat@ec.europa.eu
Tel.: +34 954488318
Fax: +34 954488300

http://ipts.jrc.ec.europa.eu/
http://www.jrc.ec.europa.eu/

This publication is a Working Paper by the Joint Research Centre of the European Commission. It results from the Digital Economy Research Programme at the JRC Institute for Prospective Technological Studies, which carries out economic research on information society and EU Digital Agenda policy issues, with a focus on Growth, Jobs and Innovation in the Single Market. The Digital Economy Research Programme is co-financed by the Directorate General Communications Networks, Content and Technology.

JRC84466

EUR 26217 EN

ISBN 978-92-79-33677-5 (pdf)

ISSN 1831-9424 (online)

doi:10.2788/32130

Luxembourg: Publications Office of the European Union, 2013

Printed in Spain

Preface

This report is based on proprietary e-commerce data collected by Google to which the JRC/IPTS has no access. Bo Cowgill has access to the data and, together with Cosmina Dorobantu, did earlier research work on international cross-border e-commerce using these Google data. This was presented at an IPTS seminar on cross-border e-commerce in December 2012 in Brussels. The IPTS subsequently asked both researchers for a specific analysis that focused on online border effects in the EU. The results of that analysis are presented in this report and were discussed at a seminar at the IPTS in Seville in June 2013.

The econometric analysis in Sections 2 and 3 of this report as well as the appendixes were produced for the JRC/IPTS by Bo Cowgill (UC Berkeley) and Cosmina Dorobantu (Oxford University) under Letters of Appointment nrs 153469-2013 and 153449-2013 respectively, which allocated a period of up to ten working days to each of them to carry out this assignment. The abstract, the introduction in Section 1 and the discussion in section 4 were written by Bertin Martens.

Abstract

An important EU Digital Single Market policy objective is to achieve an open and integrated market for online e-commerce in the EU, to make it easy for consumers to go outside their domestic market and shop online in other EU Member States. This study applies a standard gravity model of international trade to Google e-commerce data to estimate the prevalence of home bias in online shopping in the EU. It compares how much EU Member States trade domestically and with other Member States, and how much the EU trades with itself and with the rest of the world. The research confirms the findings of the (offline) international trade literature, according to which there is strong home bias. There is no unambiguous evidence about the strengths or weaknesses of the EU Digital Single Market. Strong intra-EU home bias suggests that online consumers have a tendency to stay in their home country market. Equally strong extra-EU home bias suggests that online consumers who do decide to shop abroad have a tendency to stay in the EU however, rather than going to a non-EU country. There are indications that online home bias is lower in a comparable cross-border trade setting in North America. Data and methodological limitations do not allow a more detailed analysis.

JEL codes: F15, O52

Table of Contents

1. Introduction[1]

With the rise of the internet the cost of transmitting information has dramatically declined to almost nothing and any content is instantly accessible. This has led to the "death of distance" (Cairncross, 1997) hypothesis, the idea that distance has become irrelevant in anything that can be communicated across the internet. However, previous research has shown that this is not necessarily the case. Blum & Goldfarb (2006) showed that internet traffic density is still subject to the gravity law. Internet users do not venture far beyond their familiar cultural environment and geographical distance counts in the search for taste and culture sensitive services. The relevance of distance only diminishes for less taste-dependent services. Hortaçu et al (2009) demonstrate that distance still plays a role in online cross-border trade in goods, though less so than in offline cross-border trade. Lendle et al (2012) using e-Bay data, Gomez et.al (2013) using EU consumer survey data and Cowgill and Dorobantu (2013a) using Google e-commerce data, all come to similar conclusions. Paraphrasing Marc Twain, it seems that rumours about the death of online distance have been greatly exaggerated.

Establishing a borderless Single Market for offline trade between Member States was one of the main driving forces behind the creation of the EU: internal borders should not affect the free movement of goods and services, capital and labour. The Digital Agenda for Europe (European Commission, 2010) extended this objective to online trade. One important pillar of this policy agenda is the creation of a Digital Single Market, including for online e-commerce services. The Digital Agenda has set two ambitious policy targets in this respect: getting 50% of all EU citizens to buy online and 20% to engage in online cross-border transactions by 2015. The first target seems to be within reach as the percentage of online consumers stood at 45% in 2012. The second target however looks more difficult to attain. The percentage of cross-border online consumers reached 11% in 2012 and that number has barely grown in recent years (European Commission, 2013). In other words, despite the internet's promise of a seamless online market, borders still matter and online consumers prefer to shop at home. Home bias is possibly even more important in online than in offline trade (Gomez et.al, 2013). This prompted considerable research into online trade barriers and the launch of a new EU e-commerce policy package in early 2012 (European Commission, 2012).

Home bias is a well-documented phenomenon in offline international trade research. Cross-border trade can face obstacles created by government regulations and by private companies that seek to segment markets to achieve their profit-maximizing strategies. But beyond these man-made barriers, consumers seem to have an inherent preference for products that they can buy in their home market without going across the border. The ratio of cross-border to domestic consumption is lower than what can be explained in international trade models, even if one takes into account country-specific regulatory barriers. Taking into account this general finding from the international trade literature, the question is whether the ratio of cross-border to domestic e-commerce implicit in the two Digital Agenda policy targets can be achieved, given consumer preferences for home market transactions. In order to answer this question one needs to estimate the magnitude of home bias, or the so-called border effect, in online trade.

The present research report examines two types of border effects. The first is the border effect within the EU: How much more does an EU country trade domestically, within itself, compared to

[1] This section was written by Bertin Martens.

how much it trades cross-border with other EU countries? The second is the border effect between the EU and the rest of the world: How much higher is the volume of domestic trade within the EU between Member States of this regional trade block compared to the volume of external trade between an EU and a non-EU country? The findings of the (offline) international trade literature suggest that one can expect to find a positive border effect in our empirical research in both cases. A strong extra-EU border effect suggests that consumers have a preference for intra-EU over extra-EU trade. That would indicate that the EU Digital Single Market is successful in keeping internet users inside the EU as a regional home market. A strong intra-EU border effect would suggest that consumers still prefer to shop in their home country rather than in another EU country, despite the ease of crossing borders and overcoming distances on the internet. However, depending on the relative magnitude of the online and offline border effect, we might find some indications whether e-commerce has indeed achieved a more globalized world than traditional commerce.

The data for this research were collected by Google through its online advertising programme called AdWords. It includes a feature that tracks clicks on ads and enables advertisers to measure conversions into transactions. These data allow estimation of patterns and relative volumes of domestic and cross-border online trade as well as home bias, or the border effect, by type of online product and over a period of at least a few years. It also shows how home bias is affected by the nature of the products sold online.

The empirical analysis does indeed find a strongly positive border effect, both for intra-EU and for extra-EU online trade. On the one hand, the strong intra-EU border effect suggests that online shoppers still have a strong preference to stay in their own country, rather than going out to another EU country and despite the fact that the internet makes it much easier to find shopping opportunities across the border. On the other hand, the strong extra-EU border effect shows that once EU online shoppers decide to go abroad they have a strong preference to stay inside the EU rather than venturing outside to non-EU countries. These findings do not lead to an unambiguous conclusion with regard to the success or failure of the Digital Single Market. It seems to be successful in keeping online consumers within the EU but less successful in getting them out of their home market and into other EU countries. This may offer one possible explanation why the EU Digital Agenda is underperforming with regard to the online cross-border indicator. From an economic point of view, success should ultimately be measured in terms of consumer welfare improvements, not just trade volumes. That will require further research.

Findings from research carried out on similar data for the North American online market (US states –Canadian provinces) suggest that home bias is lower in that region. Consumers seem to feel more at ease going outside their home market in that North American setting. While this analysis again controls for many specific factors, it is not possible to control for every possible structural difference between the EU and North America. We should therefor interpret this finding with caution.

It is even more difficult to compare online and offline border effects. Differences in the composition of online and offline trade as well as difference in the structure of the supply chains (B2B offline versus B2C online) prevent a straightforward comparison. We can only point out that the observed online border effects are substantially higher than those observed in the offline trade literature. That does not constitute conclusive evidence however. Further research will be required that should control for difference in composition and supply chains in order to arrive at more conclusive evidence whether online trade is indeed more globalised or not than offline trade.

This report is structured as follows. Section 2 explains the trade model that we use to estimate the drivers of online cross-border trade. Section 3 discusses the results from the empirical estimation of the model, both for intra-EU and for extra-EU online trade, and for trade by type of product. Section 4 discusses the results, compares them with other research and attempts to draw some policy conclusions.

2. Empirical specification[2]

We use a gravity model, the standard workhorse model in international trade. The gravity model links the observed volume of bilateral trade between country pairs with a series of explanatory variables. Originally, the gravity model of trade started from the assumption that the volume of trade between two countries was determined by the respective size of the two economies and the geographical distance between them – just like Newton's gravity model in physics explains attraction between two planets. Over time, many other explanatory variables have been added to the gravity model. Apart from distance the "standard" model now includes variables such as shared language, border, religion, legal regime, colonial history and currency zone, and a common land-locked or sea-locked status.

We estimate the gravity model on online bilateral trade flow data, in order to gauge the contribution of all these factors to facilitating or impeding that trade. Controlling for and neutralising the impact of all these factors enables us to isolate the effect of administrative borders between countries on trade. This "border effect" tells us something about the preference of consumers to shop at home or abroad. As such, it gives us an indication of how fragmented or integrated a regional market like the EU is.

Models are always an abstraction of reality; they do not capture every detail of that reality. The gravity model and the explanatory variables that we have added to it do not capture everything that may affect cross-border e-commerce. For instance, it does not capture the difficulty of legal recourse across different jurisdictions in cross-border transactions compared to domestic transactions that take place within a single jurisdiction. The model is sufficiently powerful however to explain a very large part of the observed variation in cross-border e-commerce (as indicated by the high adjusted R^2 in the regressions).

The gravity equation is estimated both at the aggregate economy-wide level and at the NAICS2 industry level. See Appendix 2 for an explanation of the break-down by industry.

Data on the volume of domestic and cross-border online trade comes from Google, and we provide a detailed description of it in Appendix 1. The data set covers the amount of domestic trade (a country's trade with itself) for each of the 27 EU countries, the exports and imports taking place between the 27 EU countries, and the exports and imports of each of the 27 EU countries and 191 other countries outside of the EU.

To estimate the two border effects we would ideally use our entire data set in a single regression including two dummy variables: one for a country's trade with itself, and one for trade between two EU countries. In this setting, the excluded group, against which the two border effects are measured, is trade between EU countries and the rest of the world.

Estimating this equation is not realistically possible. Ever since Anderson and Van Wincoop (2003)'s seminal work highlighting the importance of accounting for multilateral resistance terms in a gravity setting, empirical trade economists have been advocating for the inclusion of importer and exporter country fixed effects. There are two advantages arising from the inclusion of fixed effects. First, a gravity specification which employs importer and exporter fixed effects is consistent with the theoretical model used to derive it. Second, a specification with fixed effects allows the

[2] This section was written by Bo Cowgill and Cosmina Dorobantu.

researcher to control for all unobservable characteristics that have an effect on a country's overall level of exports or imports.

For our data set, the inclusion of importer and exporter fixed effects in addition to the two border effects dummy variables whose coefficients we want to estimate creates perfect multi-collinearity. Given this, we have two options: either to exclude the fixed effects, or to estimate two separate regressions, the first of which allows us to measure the border effect within the EU, while the second allows us to examine the border effect between the EU and the rest of the world. We choose the second option, due to the biases inherent in a gravity specification that lacks importer and exporter fixed effects.

We consider only the 27 EU countries when estimating the first regression, which measures the border effect within the EU. The data set, therefore, consists of a 27 × 27 matrix, reflecting the inclusion of each country's domestic trade, as well as the exports and imports between that country and its 26 EU trading partners. In terms of explanatory variables, we use the same controls that Helpman, Melitz and Rubinstein (2008) employed in their study. The gravity equation we estimate takes the following form:

$$\ln x_{ij} = \beta_0 + \beta_1 \ln(\text{Distance})_{ij} + \beta_2 \text{ Land border}_{ij} + \beta_3 \text{ Island}_{ij} + \beta_4 \text{ Landlock}_{ij} + \beta_5 \text{ Legal}_{ij}$$
$$+ \beta_6 \text{ Language}_{ij} + \beta_7 \text{ Colony}_{ij} + \beta_8 \text{ Currency}_{ij} + \beta_9 \text{ Home}_{ij} + \beta_{10} \text{ Religion}_{ij} + \lambda_i + x_i + \varepsilon_{ij}$$

where:

- xij: online conversions (exports) from country i to country j
- Distance: great circle distance, in km, between the capitals of i and j
- Land border: 1 if two countries share a land border, 0 otherwise
- Island: 1 if both countries are islands, 0 otherwise
- Landlock: 1 if both countries have no coastline, 0 otherwise
- Legal: 1 if both countries' legal systems share the same origin, 0 otherwise
- Language: 1 if both countries share an official language, 0 otherwise
- Colony: 1 if the two countries were in a colonial relationship, 0 otherwise
- Currency: 1 if both countries use the same currency, 0 otherwise
- Home: 1 for an EU country's trade with itself, 0 otherwise
- Religion: (Protestants in i × Protestants in j) + (Catholics in i × Catholics in j) + + (Muslims in i × Muslims in j)
- λi: exporter fixed effect
- xj: importer fixed effect
- εij: error term

For the second regression we estimate, which measures the border effect between the EU and the rest of the world, we include all exports and imports between the EU countries (a 26 × 27 matrix), as well as all exports and imports between each EU country and each non-EU country (a (2×27) × 191 matrix) . Note that in this second regression, we exclude from our data set the domestic trade volumes. The gravity equation we estimate takes the following form:

$$\ln x_{ij} = \beta_0 + \beta_1 \ln(\text{Distance})_{ij} + \beta_2 \text{ Land border}_{ij} + \beta_3 \text{ Island}_{ij} + \beta_4 \text{ Landlock}_{ij} + \beta_5 \text{ Legal}_{ij}$$
$$+ \beta_6 \text{ Language}_{ij} + \beta_7 \text{ Colony}_{ij} + \beta_8 \text{ Currency}_{ij} + \beta_9 \text{ EU}_{ij} + \beta_{10} \text{ Religion}_{ij} + \lambda_i + x_i + \varepsilon_{ij}$$

where the dependent and all independent variables are defined as above, with the exception of:

- EU: 1 for trade between two EU countries, 0 otherwise

We use OLS to estimate the two gravity equations specified above. In the academic literature, OLS has been criticized for failing to account for zero trade flows. In our first regression, intended to measure the magnitude of border effects within the EU, all trade flows are greater than zero. In this setting, where there are no zero trade flows, OLS is a suitable estimator. In our second regression, intended to measure the magnitude of the border effect between the EU and the rest of the world, around a third of the trade flows are zeros. OLS is less suitable for this second regression, as a selection bias is introduced by omitting the zero trade flows.

In order to justify why we nevertheless opt for OLS as our estimator, we note that two alternative methodologies have been proposed in the academic literature to address zero trade flows. The first one, presented in Helpman, Melitz and Rubinstein (2008), consists of a two stage estimation, the first of which is a Probit equation describing trade flows. While this methodology would be ideal for the second regression we run (intended to capture the border effect between the EU and the rest of the world), it cannot be used for our first regression. Within the EU, all trade flows in our data set are greater than zero, implying that the probability of trade between two countries is always 1. Running a Probit, as a first stage estimation, therefore, is not possible. Since the results of the two regressions we estimate need to be comparable, we rule out the methodology proposed by Helpman, Melitz and Rubinstein (2008), which can only be employed to estimate the second of our gravity equations. The second methodology, presented in Santos and Tenreyro (2008), proposes a Poisson pseudo maximum likelihood estimator (PPML). A simple RESET test, however, shows that for our data set, the OLS estimator does a better job than the PPML estimator at addressing potential heteroskedasticity problems.

3. Results[3]

We estimate the regressions separately for each of the four years included in our data set. Table 1 presents the results we obtain when running the first regression, which is meant to measure the border effect within the EU. We discuss the coefficients for each variable separately.

Border effects

International borders, even within the EU, impact online trade flows. The coefficient on the home dummy variable is large, positive, and statistically significant. It ranges from a minimum of 3.255 in 2010 to a maximum of 4.018 in 2009. Taking the exponent of the coefficients reported for the Home dummy variable gives the average border effect. In 2010, the year for which we obtain the lowest border effect, the exponent of 3.255 is approximately equal to 25.9. The highest home dummy coefficient of 4.018 is obtained for 2010, translating into a border effect of approximately 55. This indicates that within the EU, domestic online trade is between 26 and 55 times higher than international online trade. For comparison purposes, in a North American setting with online trade inside and between US states and Canadian provinces, Cowgill and Dorobantu (2013b) find that domestic online trade is only 5.8 times higher than international online trade.

Table 1: Determinants of intra-EU online trade

	2008	2009	2010	2011
Distance	-0.448***	-0.435***	-0.631***	-0.621***
	(0.120)	(0.132)	(0.131)	(0.147)
Land border	1.015***	1.047***	1.190***	1.260***
	(0.195)	(0.210)	(0.207)	(0.239)
Island	-0.823	-1.510**	-1.227	-1.386
	(0.653)	(0.715)	(0.766)	(0.904)
Landlock	0.381	0.322	0.466	0.667*
	(0.337)	(0.406)	(0.362)	(0.348)
Legal	0.327***	0.444***	0.475***	0.317***
	(0.097)	(0.109)	(0.116)	(0.115)
Language	0.671***	0.281	0.743**	0.825***
	(0.248)	(0.303)	(0.310)	(0.315)
Colony	-0.136	0.267	0.100	0.210
	(0.283)	(0.379)	(0.345)	(0.359)
Currency	0.354**	0.420**	0.473***	0.258
	(0.156)	(0.176)	(0.169)	(0.181)
Home	3.901***	4.018***	3.255***	3.447***
	(0.428)	(0.494)	(0.461)	(0.516)
Religion	0.379*	0.351	0.365	0.534**
	(0.229)	(0.237)	(0.261)	(0.269)
Obs.	729	729	729	729
Adjusted R^2	0.884	0.853	0.876	0.864

Notes: exporter and importer fixed effects. Estimated by OLS. Dependent variable is in logs. Robust standard errors (clustering by country pair). *** significant at 1%, ** significant at 5%, * significant at 10%.

[3] This section was written by Bo Cowgill and Cosmina Dorobantu.

We do not really have an explanation for the considerable variation in the value of the coefficient across these four years. Four observations in time are not enough to perform a meaningful statistical analysis. We can only point out that this was a period with considerable turbulence in the world and in the EU economy in the wake of the financial crisis.

Distance

Table 1 indicates that distance has a negative and significant influence on online trade. The elasticity of distance ranges between a minimum of -0.435 in 2009 and a maximum of -0.631 in 2010. These elasticities imply that all else being equal, a 1% increase in the distance between two countries reduces the volume of trade between them by approximately 0.44% in 2009 and 0.63% in 2010. While the magnitude of the coefficients we estimate for the EU is similar to the one obtained in gravity studies analyzing worldwide online trade flows, it is larger than the magnitude of coefficients obtained in gravity studies focusing solely on online trade within North America. For online exchanges happening within the United States and Canada, researchers find the elasticity of distance to be around -0.1.

Language

Sharing a common official language has a positive and statistically significant effect on online trade in all years except for 2009. The magnitude of the coefficients is fairly large, ranging between 0.671 in 2008 to 0.825 in 2011. Taking the exponent of these coefficients allows us to conclude that all else being equal, two countries that have a common official language trade roughly two times more online than countries that do not.

Table 2: Determinants of extra-EU online trade

	2008	2009	2010	2011
Distance	-0.437***	-0.407***	-0.722***	-0.705***
	(0.061)	(0.067)	(0.064)	(0.061)
Land border	0.876***	0.938***	0.764***	0.897***
	(0.164)	(0.164)	(0.155)	(0.175)
Island	-0.006	0.018	-0.012	0.014
	(0.088)	(0.093)	(0.114)	(0.116)
Landlock	0.152	0.164	0.228**	0.139
	(0.100)	(0.100)	(0.095)	(0.092)
Legal	0.155***	0.131***	0.143***	0.150***
	(0.036)	(0.039)	(0.040)	(0.040)
Language	0.612***	0.990***	1.015***	1.101***
	(0.078)	(0.090)	(0.097)	(0.091)
Colony	0.501***	0.630***	0.656***	0.343***
	(0.105)	(0.116)	(0.127)	(0.121)
Currency	0.218**	0.158	0.456***	0.153*
	(0.099)	(0.104)	(0.100)	(0.093)
EU	2.847***	4.083***	4.152***	3.432***
	(0.268)	(0.267)	(0.225)	(0.224)
Religion	0.460***	0.197**	0.123	0.245***
	(0.091)	(0.097)	(0.098)	(0.094)
Obs.	7,452	7,748	7,768	7,841
Adjusted R^2	0.884	0.882	0.877	0.891

Notes: exporter and importer fixed effects. Estimated by OLS. Dependent variable is in logs. Robust standard errors (clustering by country pair). *** significant at 1%, ** significant at 5%, * significant at 10%

Table 2 reports the results we obtain by running our second regression, meant to capture the border effect between the EU and the rest of the world. The border effect ranges from a minimum of 2.847 in 2008 to a maximum of 4.152 in 2010. The exponent of 2.847 is roughly equal to 17.2, indicating that online trade within the EU is at least 17.2 times higher than online trade between EU countries and non-EU countries. For 2010, the exponent reaches a value of 59.3. This coefficient seems to imply that the (Digital) Single Market encourages trade between the EU member states in the virtual world. The results for the other two variables of interest, distance and language, are similar to ones reported in Table 1.

Industry level results

As a final exercise to get more insight into the border effect within the EU, we classify our data at the NAICS 2-digit level. The methodology we follow is described in detail in Appendix 2. We report here the results that we obtain by running the first of our two regressions for each services sector. While we provide the results obtained by running our second regression at the industry-level, as well, we believe they are not conclusive due to the high proportion of zeros in the data. For many combinations of countries and industries, there were zero trade flows in our data.[4]

In table 3, the first four columns report the elasticity of distance that we estimate for each year and industry considered, as well as the standard errors. Columns 5 to 8 report the coefficients that we estimate for the Home dummy variable and the associated standard errors. These coefficients tell us how much more an EU country trades with itself compared to how much it trades with other EU countries. Finally, columns 9 to 12 report the coefficients estimated for the third variable of interest, namely the dummy variable taking the value of 1 when two EU countries share the same language.[5]

A few items stand out in this final table. First, language effects appear in natural places. In line with the findings of Blum and Goldfarb (2006), they are largest in culture and taste-sensitive industries like Education, Accommodation and Food, and the Arts, Entertainment and Recreation industries. Second, the home bias appears to be strongest among regulated industries such as Health care and Finance and Insurance. For Finance and Insurance, for example, the coefficient on distance is insignificant, indicating that this is a service that travels well. The border effect, however, is very large. This suggests that although this service can be easily provided over large distances, it does not cross borders very well. Finally, the distance effects seem to be largest for the Cultural, and Accommodation and Food industries.

[4] For some industries, over 90% of the data set consists of zeros.

[5] Table 3 presents an abridged version of the results. A full version is also available.

Table 3: Online trade estimates by product category

	Distance				Home bias				Language			
	2008	2009	2010	2011	2008	2009	2010	2011	2008	2009	2010	2011
Wholesale trade	-0.68***	-0.80***	-0.43**	-0.09	3.41***	2.91***	4.53***	5.00***	0.90	0.73	0.89	1.34***
	(0.24)	(0.19)	(0.21)	(0.18)	(0.93)	(0.83)	(0.87)	(0.65)	(0.64)	(0.61)	(0.60)	(0.41)
Retail trade	-0.27*	-0.37**	-0.61***	-0.68***	4.22***	4.33***	3.65***	4.09***	1.26***	0.64*	1.00***	0.81**
	(0.16)	(0.16)	(0.17)	(0.19)	(0.52)	(0.58)	(0.53)	(0.65)	(0.32)	(0.37)	(0.35)	(0.41)
Transportation, warehousing	-0.35	-0.84***	-0.89***	-0.59***	5.24***	3.74***	4.29***	4.74***	-0.43	0.60	0.47	0.43
	(0.29)	(0.18)	(0.20)	(0.17)	(0.97)	(0.80)	(0.89)	(0.72)	(0.53)	(0.45)	(0.56)	(0.44)
Utilities	-0.04	-0.22	-0.51**	-0.58***	6.49***	4.17***	2.99***	4.94***	1.19	2.13***	2.95***	1.21**
	(0.27)	(0.17)	(0.25)	(0.21)	(1.08)	(1.02)	(0.87)	(0.79)	(0.81)	(0.74)	(0.67)	(0.50)
Info and cultural industries	-0.47***	-0.47***	-0.55***	-0.61***	4.24***	4.03***	3.20***	3.73***	0.24	0.37	0.82**	0.69*
	(0.16)	(0.15)	(0.13)	(0.14)	(0.60)	(0.66)	(0.60)	(0.63)	(0.32)	(0.40)	(0.41)	(0.39)
Finance and insurance	-0.13	-0.07	-0.25	-0.36	5.46***	5.66***	5.41***	4.77***	1.45***	1.55***	1.34***	1.65***
	(0.27)	(0.22)	(0.24)	(0.24)	(0.88)	(0.70)	(0.72)	(0.75)	(0.52)	(0.47)	(0.49)	(0.43)
Real Estate	-0.24	-0.45**	-0.59***	-0.65***	4.53***	4.98***	4.30***	5.14***	1.66***	0.89**	1.22**	0.37
	(0.25)	(0.20)	(0.19)	(0.18)	(0.95)	(0.72)	(0.83)	(0.75)	(0.62)	(0.44)	(0.54)	(0.41)
Prof., sci. and tech services	-0.01	-0.05	-0.33**	-0.15	3.75***	4.17***	3.86***	4.51***	0.85**	0.43	0.80**	0.64**
	(0.10)	(0.12)	(0.17)	(0.15)	(0.55)	(0.53)	(0.53)	(0.53)	(0.36)	(0.36)	(0.37)	(0.32)
Administrative and support	-0.54***	-0.67***	-0.62***	-0.51***	3.00***	3.43***	2.93***	3.11***	1.18***	0.94***	0.98***	1.24***
	(0.20)	(0.16)	(0.14)	(0.16)	(0.69)	(0.67)	(0.65)	(0.67)	(0.35)	(0.34)	(0.33)	(0.40)
Education services	-0.65**	-0.35*	-0.13	-0.22	3.95***	3.90***	5.30***	4.50***	1.18**	1.75***	0.40	0.95
	(0.27)	(0.21)	(0.23)	(0.21)	(0.87)	(0.72)	(0.84)	(0.85)	(0.58)	(0.49)	(0.51)	(0.62)
Health care	-0.85***	-0.80***	-0.50*	-0.67***	3.76***	4.41***	4.53***	4.73***	1.14*	0.25	1.25*	1.48***
	(0.28)	(0.27)	(0.26)	(0.19)	(0.84)	(0.87)	(0.93)	(0.82)	(0.65)	(0.50)	(0.68)	(0.51)
Arts, entert., recreation	-0.50***	-0.66***	-0.53***	-0.82***	2.51***	2.37***	3.41***	2.83***	0.98***	1.15***	1.32***	1.47***
	(0.17)	(0.22)	(0.17)	(0.22)	(0.60)	(0.64)	(0.61)	(0.66)	(0.37)	(0.39)	(0.34)	(0.43)
Accommodation and food	-1.58***	-1.89***	-0.49	-1.03***	1.71	1.49	5.10***	5.71***	2.29***	1.81**	1.63*	1.47*
	(0.33)	(0.47)	(0.42)	(0.27)	(1.22)	(1.33)	(1.40)	(1.17)	(0.78)	(0.83)	(0.92)	(0.84)
Other services	-0.19	-0.26	-0.19	-0.40**	3.73***	4.81***	5.47***	3.99***	1.61**	0.90	0.93	2.00***
	(0.19)	(0.24)	(0.22)	(0.18)	(1.00)	(0.91)	(0.95)	(0.76)	(0.68)	(0.62)	(0.68)	(0.50)
Public Administration	-0.05	-0.19	-0.58	-0.64	3.94***	4.96***	5.20***	4.52***	1.31**	1.52***	0.86	1.14**
	(0.20)	(0.26)	(0.42)	(0.40)	(0.82)	(0.77)	(0.96)	(1.12)	(0.58)	(0.48)	(0.56)	(0.55)

Notes: exporter and importer fixed effects. Estimated by OLS. Dependent variable is in logs. Robust standard errors (clustering by country pair). *** significant at 1%, ** significant at 5%, * significant at 10%.

4. Discussion[6]

a) As expected and in line with the findings from the international trade literature, the empirical analysis does indeed find a strongly positive border effect, both for intra-EU and for extra-EU online trade. Over the period 2008-2011, online consumers in the EU are between 25 and 55 times more likely to buy in their own country than in another EU country. They are between 17 and 59 times more likely to buy in the EU than outside the EU. The lower values in this range are comparable to the findings of Gomez et.al (2013) who estimate a home bias factor of 16 for online trade inside the EU in 2011 – to our knowledge the only other study that examines online home bias in the EU.

b) These findings do not lead to an unambiguous conclusion with regard to the strengths or weaknesses of the Digital Single Market. On the one hand, the strong intra-EU border effect suggests that online shoppers still have a strong preference to stay in their own country, rather than going out to another EU country – despite the fact that the internet makes it much easier to find shopping opportunities across the border. On the other hand, the strong extra-EU border effect shows that once EU online shoppers decide to go abroad they have a strong preference to stay inside the EU rather than venturing outside to non-EU countries. The Digital Single Market seems to be successful in keeping online consumers within the EU but less successful in getting them out of their home market and into other EU countries. This may offer one possible explanation why the EU Digital Agenda is underperforming with regard to the online cross-border indicator. From an economic point of view, success should ultimately be measured in terms of consumer welfare improvements, not just trade volumes. That will require further research.

c) Note that these border effect estimations control for and neutralise the impact of differences between countries in terms of distance between them, language, legal systems and other country-specific factors. Besides, we should refrain from a direct comparison between the intra- and extra-EU border effects, given the methodological problems that were pointed out above. Ideally, we would have run a single gravity model combining the intra- and extra-EU border effects that would have given us a straight answer to this question. However, that was not feasible because of methodological constraints.

d) Findings from research carried out on similar data for the North American online market (US states –Canadian provinces) suggest that home bias is lower in that region. Consumers seem to feel more at ease going outside their home market in that North American setting. While this analysis again controls for many specific factors, it is not possible to control for every possible structural difference between the EU and North America. We should therefor interpret this finding with caution.

e) If online home bias would be lower than offline, then e-commerce would have lived up at least partially to the promise of a borderless world (Cairncross, 1997) and achieved a more globalized world than traditional commerce. Unfortunately, a direct comparison of online and offline border effects is fraught with problems. Differences in the composition of online and offline trade as well as differences in the structure of the supply chains (B2B offline versus B2C online) prevent a straightforward comparison. Several researchers have therefore limited the comparison to differences in the importance of geographical distance. The findings of Hortaçu

[6] This section was written by Bertin Martens.

et al (2009), Lendle et al (2012) and Gomez et al (2013) confirm that the importance of distance is considerably reduced in online trade compared to offline trade, for a similarly composed trade basket. This seems to support the "death of distance" hypothesis. The present study does not compare the distance effect for online and offline trade. Reduction in distance costs can only be squared with an increase in home bias if other variables can be brought into the picture that undo and reverse the reduction in distance-related trade costs. Gomez et al (2013) point out that language could be such a factor, especially in a linguistically and culturally segmented market like the EU. While distance related trade costs diminish substantially, language related trade costs increase when moving from offline B2B to online B2C trade. Lendle et al (2012) come to similar conclusions.

f) We can only point out that the observed online border effects in this study are substantially higher than those observed elsewhere in the offline trade literature. For example, Pacchioli (2011) estimates home bias in offline trade and finds that EU countries are between 7 and 24 times more likely to trade at home than with one another, considerably higher than US states which are only 2.6 to 7 times more likely to trade at home. That does not constitute conclusive evidence however that home bias is actually higher online. Further research will be required that should control for differences in composition and supply chains in order to arrive at more conclusive evidence whether online trade is indeed more or less globalised than offline trade.

g) This is one of the first studies to examine digital trade patterns within the EU. The initial set of results is promising, and we hope it will motivate other researchers to dedicate their time to this topic. We suggest two specific ways to move the research further. First, the presence of many zero bilateral trade flow observations in the data set, both at the aggregate and industry level, may distort the gravity estimation and must be addressed. Second, it is important to find a more robust way to compare results from online data to traditional, offline trade figures.

The key explanatory variables used in this gravity model of cross-border e-commerce (distance, language, home bias) do a good job in explaining the observed patterns of trade flows at the country and industry level. Admittedly, they are not policy instruments and offer little prospect for policy makers to try to achieve a seamless online Digital Single Market in the EU. Policy instruments are situated at the more disaggregated level of individual online transactions and the obstacles that consumers may face when they want to shop outside their home market. These may be related to trust, geographical access restrictions, legal issues, online payment systems, parcel delivery, etc. More research will be required to explore the relative importance of all these potential obstacles and ways to overcome them.

Appendix 1: Data Description[7]

Google data

The data we use come from Google's online advertising program, called “AdWords.” One of the features that Google offers to AdWords advertisers is a service called “conversion tracking.” This feature anonymously tracks users viewing or clicking an ad through the checkout process. The feature enables AdWords advertisers to measure the cost effectiveness of Google’s ads -- both in terms of dollars (or other currencies) as well as transaction counts. The resulting measures of transaction sizes and volumes are stored on Google’s servers, so that clients can view them alongside advertising costs when they browse their advertising accounts.

In this study, we use data from the aforementioned “conversion tracking” feature. To our knowledge, this is the most comprehensive data set used to date to examine digital trade-related questions. However, like all data sets, the Google data has limitations. These are minor compared to the limitations faced by other studies that looked at online trade. Nonetheless, we clearly spell them out below:

- Counts only: The conversion tracking code reliably generates counts of transactions and other conversions. We do not use the transaction values associated with these counts. The counts represent the number of times users reach pages where the advertisers placed the code.
- AdWords ads only: Our sample is necessarily limited to sellers who are advertisers with Google -- including its advertising platform that sends ads to other publishers' sites. Google partners with millions of online businesses around the world. Because of its high quality targeting technology, Google’s advertising is attractive to many online businesses. Nonetheless, some online conversions will not be tracked in this dataset because the sellers were not registered with Google -- or because users visited the seller’s site without going through a Google ad. For example, if a user types www.ebay.com into their browser and purchases an object directly from www.ebay.com, that transaction will not be recorded in the data set, because it happened outside of Google's advertising program.
- 30 day limit: The conversion tracking feature generates a record if a user clicks on an AdWords ad and reaches the page where the code is placed within 30 days of the initial click on the ad. If the user reaches the page after 30 days from the original click, the conversion will not be recorded.
- Optional tool: The conversion tracking tool is optional, so advertisers are not required to enable it. However, the adoption rate is high, because it is the easiest way for advertisers to measure the conversions coming through their ads.

The location data associated with our transactions come from two sources. For buyers, we use estimates based on IP (Internet Protocol) address. For sellers, we use the self-reported address data required of businesses who sign up to be Google advertisers.

We believe that this information allows us to place most buyers and sellers in the correct region. Nonetheless, it is possible that we are attributing a small part of the buyers or sellers to the wrong location. Our IP addresses might indicate an incorrect physical location if the user is accessing the

---

[7] This is a reproduction of the data section presented in Cowgill and Dorobantu (2013a). Appendix 1 and 2 were written by Bo Cowgill and Cosmina Dorobantu.

internet using a virtual private network (VPN) connection or if the user is, for whatever reason, using software to mask his or her actual IP address. We believe this is a small enough percentage of the total internet users to make our location identification on the buyers' side reliable.

On the seller side, we have access to two address fields: a general mailing address and a billing address associated with the advertiser's credit card. We use the former, but a data check reveals that the general mailing address reported by the sellers very rarely differs from the billing address.

Representativeness and biases

We cannot assess the statistical representativeness of this database in the absence of information on the entire universe of online transactions in the world and/or in the countries we study. That information does not exist. All we can claim is that this is the most extensive database used to date.

Concerns might arise about possible build-in home bias in Google's ad targeting algorithms. However, online advertisers can select in which geographical area they want their ads to pop up. If advertisers know that their product sells only locally, they can restrict ads to local areas. If they have a more global reach they can advertise more widely. Consequently, geo-targeting is not Google algorithm's choice but the advertisers' own choice. This is similar to offline shops that decide where to advertise: only in their shop window to attract local passers-by, or in a wider geographical area using regional newspapers, or globally using other means, etc. Likewise, consumers can decide where to look for the goods they intend to buy, either locally or further away. This study measures total home bias which is a combination of retailers' and consumers' home bias. There is no attempt in this study, or indeed in any other border effect analysis that the authors are aware of, to separate the two sources of home bias. This may be a subject for future studies.

Coverage

Google announced in 2007 that AdWords had more than one million online advertising accounts. This is the last publicly available figure. This indicates that six years ago, there were already more than one million businesses opting for Google's advertising platform. Google also announced publicly that no client accounts for more than 10% of the company's revenues, indicating that its advertising program is not dominated by one large business.

Our data covers all conversion counts recorded over a period of four years, from 2008 to 2011. While we cannot provide precise figures, the overall number of recorded conversion counts is far in excess of 10 billion. The data are available at the daily level, although we aggregate it to the yearly level to generate our main results. All except 6 countries are featured in our data. A trade embargo between the United States and Burma, Cuba, Iran, North Korean, Sudan and Syria keep these countries out of the seller side of our data. As a US-based company, Google cannot generate sales from these countries.

Other Variables

The only other variable that requires explanation is our distance measure, d_{ij}. We use the CEPII Distances database, which reports a measure for distance between the capital of country i and the capital of country j in kilometers, calculated using the Great Circle Distance Formula. We used information reflected in CIA's World Factbook to construct all the other variables.

Appendix 2: Methodology to break down data by NAICS2 Sectors[8]

Google has developed advanced text classification software to automatically categorize ads running on its online platforms. Google's verticals are structured like NAICS sectors: There are 27 top level categories, which are similar to the 2-digit NAICS sectors -- these are the broadest classifications available. Under these top level verticals, there are 241 second and third level categories -- similar to the 4-digit and 6-digit NAICS classifications. The aforementioned algorithms assign each search to the relevant vertical and sub-vertical. For example, a search for [car tires] would be classified under the third level category 'Vehicle Tires,' the second level category 'Auto Parts' and the top level vertical 'Automotive.'

We map Google's verticals to NAICS sectors using a classification scheme designed by Google Chief Economist Hal Varian in Choi and Varian (2012). A few examples of the mapping are provided below:

Google Vertical		NAICS Sector	
ID	Title	ID	Title
47	Automotive	441	Motor vehicle and parts dealers
5	Computers and Electronics	443	Electronics and appliance stores
1868	Apparel	448	Clothing and clothing accessories stores
...	...	...	...

Where Choi and Varian (2012) did not provide a NAICS/vertical mapping, we selected and assigned an encoding.

We successfully assigned a 2-digit NAICS classification to over 99% of the conversions in our data set. 0.2% of our data remain unclassified. Of the classified conversions, the vast majority (98% +) fell in the service sectors.[9] Only 1.5% of them fall within the Agriculture, Mining, Construction and Manufacturing sectors.

The underrepresentation of Manufacturing in our data set warrants further discussion. Google does have a vertical for “Industries,” the company's equivalent of what economists would call “Manufacturing.” Less than 0.2% of conversions in our data set are classified as Industries. Google also has verticals such as 'Computers and Electronics.' We can classify this vertical as NAICS 334, Computer and Electronic Product Manufacturing, or as NAICS 443, Electronics and Appliance Stores. The first classification would fall under Manufacturing, the second one would fall under Services. Following Choi and Varian (2012), we classify 'Computers and Electronics' within Services, under Retail Trade. They select this classification because retailers, rather than manufacturers, are Google's usual advertising clients.

[8] This is a reproduction of the methodology to break down data by NAICS2 sectors presented in Cowgill and Dorobantu (2013b).

[9] The only service sector that we cannot map any conversions to is 'Management of Companies and Enterprises.

Bibliography

Anderson, James E., and Eric van Wincoop (2003) "Gravity with Gravitas: A Solution to the Border Puzzle." American Economic Review 93.1 (2003): 170-92.

Blum, Bernard and Avi Goldfarb (2006) "Does the internet defy the law of gravity", Journal of international economics, vol 70 (2006), pp 384-405

Cairncross, Frances (1997) "The death of distance; how the communications revolution will change our lives", London, Orion Business Books

Choi, Hyunyong, and Hal Varian (2012) "Predicting the Present with Google Trends." The Economic Record 88:1 (2012): 2-9.

Cowgill, Bo and Cosmina Dorobantu (2013a) "Worldwide Gravity in Online Commerce", Working Paper.

Cowgill, Bo and Cosmina Dorobantu (2013b) "The US-Canada Border Effect in Online Commerce", Working Paper.

European Commission (2010) "A Digital Agenda for Europe", available at http://ec.europa.eu/digital-agenda/digital-agenda-europe

European Commission (2012) "A coherent framework to build trust in the Digital single market for e-commerce and online services", available at http://ec.europa.eu/internal_market/ecommerce/communication_2012_en.htm

European Commission (2013) "Digital Agenda Scoreboard", 2013 update, available at http://ec.europa.eu/digital-agenda/en/scoreboard

Gomez, Estrella, Bertin Martens and Geomina Turlea (2013) "The drivers and impediments to online cross-border trade in the EU", JRC/IPTS Digital Economy working paper 2013-02

Helpman, Elhanan, Marc Melitz, and Yona Rubinstein (2008) "Estimating Trade Flows: Trading Partners and Trading Volumes" The Quarterly Journal of Economics 123.2 (2008): 441-87.

Hortacsu, Ali, F. Asis Martinez-Jerez, and Jason Douglas (2009) "The Geography of Trade in Online Transactions: Evidence from eBay and MercadoLibre." American Economic Journal: Microeconomics 1.1 (2009): 53-74.

Lendle, Andreas, Marcelo Olarreaga, Simon Schropp, and Pierre-Louis Vezina (2012) "There Goes Gravity: How eBay Reduces Trade Costs", Centre for Economic Policy Research Discussion Paper No. 9094.

Pacchioli, Consuelo (2011) "Is the EU internal market suffering from an integration deficit? Estimating the home bias effect", CEPS working document nr 348, May 2011.

European Commission
EUR 26217 – Joint Research Centre – Institute for Prospective Technological Studies

Title: Does online trade live up to the promise of a borderless world? Evidence from the EU Digital Single Market

Authors: Bo Cowgill, Cosmina Dorobantu, Bertin Martens

Luxembourg: Publications Office of the European Union

2013 – 19 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series – ISSN 1831-9424 (online)

ISBN 978-92-79-33677-5 (pdf)

doi:10.2788/32130

Abstract

An important EU Digital Single Market policy objective is to achieve an open and integrated market for online e-commerce in the EU, to make it easy for consumers to go outside their domestic market and shop online in other EU Member States. This study applies a standard gravity model of international trade to Google e-commerce data to estimate the prevalence of home bias in online shopping in the EU. It compares how much EU Member States trade domestically and with other Member States, and how much the EU trades with itself and with the rest of the world. The research confirms the findings of the (offline) international trade literature, according to which there is strong home bias. There is no unambiguous evidence about the strengths or weaknesses of the EU Digital Single Market. Strong intra-EU home bias suggests that online consumers have a tendency to stay in their home country market. Equally strong extra-EU home bias suggests that online consumers who do decide to shop abroad have a tendency to stay in the EU however, rather than going to a non-EU country. There are indications that online home bias is lower in a comparable cross-border trade setting in North America. Data and methodological limitations do not allow a more detailed analysis.

LB-NA-26217-EN-N

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new standards, methods and tools, and sharing and transferring its know-how to the Member States and international community.

Key policy areas include: environment and climate change; energy and transport; agriculture and food security; health and consumer protection; information society and digital agenda; safety and security including nuclear; all supported through a cross-cutting and multi-disciplinary approach.

Publications Office

ISBN 978-92-79-33677-5

9 789279 336775