Institute for Prospective Technological Studies

Digital Economy Working Paper 2012/1

The drivers and impediments for online cross-border trade in goods in the EU

Bertin Martens and Geomina Turlea[1]

July 2012

Abstract

There are no official statistics on international online trade in goods so far. This paper uses a consumer survey to construct a unique matrix of online B2C domestic and cross-border trade in goods between the 27 EU Member States. We compare online and offline trade patterns for similar goods. We find that the standard gravity model performs well in explaining online cross-border trade flows. The model confirms the strong reduction in geographical distance-related trade costs, compared to offline trade. However, the trade costs associated with crossing language barriers increase when moving from offline to online trade. Institutional variables such as the quality of legal institutions, online payments facilities and cost-efficiency of parcel delivery systems also play a role in cross-border trade. In a linguistically segmented market like the EU, online home market bias remains high. We conclude that there is some policy space available for regulators to boost online cross-border trade through improvements in legal and financial systems, and parcel delivery infrastructure.

JEL codes: F15, O52

[1] Institute for Prospective Technological Studies (IPTS), a research center of the European Commission in Sevilla (Spain). This paper is produced under the Digital Economy Research Programme. The authors can be reached at geomina.turlea //AT// ec europa eu and bertin.martens //AT// ec europa eu The views and opinions expressed in this paper are the authors' and do not necessarily reflects those of the European Commission, DG JRC or the IPTS.

1. Introduction

According to the European Commission (2012), ten years after the adoption of the EU E-Commerce Directive, e-commerce is still limited to less than 4% of total European trade, far below its full economic potential. The Commission's Digital Agenda for Europe aims to get 50% of all European citizens to buy online and 20% to engage in online cross-border transactions by 2015. The communication "Building trust in the Digital Single Market for e-commerce" (European Commission, 2012) proposed measures to boost online retail trade in the EU and overcome obstacles to ecommerce, including an insufficient number of online shops willing to sell across the border, inadequate payment and parcel delivery systems, and too many cases of abuse and disputes that are difficult to settle. These conclusions were mainly based on descriptive statistics derived from a consumer survey on online business-to-consumer (B2C) trade in goods (Civic Consulting, 2011). The current paper uses the same consumer survey data but submits them to a more in-depth econometric analysis to quantify the importance of the main drivers and impediments to online cross-border transactions by EU consumers.

The analytical tool that we use for this purpose is the "gravity model" of cross-border international trade, the standard workhorse for explaining international trade flows in the offline economy (Deardorff, 1984; Anderson & Van Wincoop, 2003). This model is rooted in the Newtonian idea that much of the observed patterns of international trade flows can be explained by the economic size of the trading partners and their physical distance. "Distance" can be more broadly interpreted as a catch-all variable and proxy for various sources of international trade costs that affect the relative price of domestic and imported goods. This may include physical transport costs, the cost associated with import tariffs and regulator barriers, and risks related to poor contract enforcement between different jurisdictions. In the traditional brick & mortar economy information retrieval is costly and requires physical transport, either to bring information about a good to potential customers or to bring the customer to the good to observe its characteristics. The rise of the internet and, more generally, digital communications technology, has led many observers to announce the "death of distance" (Cairncross, 1997; Blum & Goldfarb, 2006). It does not matter anymore where buyers and sellers are located since information is only a mouse click away and no longer related to physical distances. On the other hand, new sources of trade costs may emerge that are specific to online cross-border transactions. Low information costs and the resulting wider geographical reach may result in increased price competition among suppliers and increased variety of goods available, but also weaker interaction between consumers and suppliers. This reduces trust and increases risk perceptions, compared to offline trade. These factors may increase online trade costs.

This paper examines the importance of three sources of trade costs in online trade. First, it assesses to what extent the shift from ordinary offline trade to internet-enabled online trade has reduced the importance of geographical distance-related trade costs. While distance may no longer matter for information and purely digital products and services (Blum & Goldfarb, 2006), goods still need to be physically transported, and sometimes cross borders between different regulatory regimes, to reach the buyer. Consequently, only part of the distance-related trade costs is affected by the shift from analogue to digital information technology. Second, the paper examines the importance of a set of cultural and institutional factors, such as language and the quality of legal institutions, as determinants of online trade. The shift from offline to online trade may well increase their relative importance. Third, we assess the importance of two essential ingredients of online trading platforms, flexible online payments systems and cost-efficient parcel delivery systems. Though we focus on online digital trade, the money and the physical goods still need to be able to cross borders in order to complete a transaction. Finally, we examine how the shift from offline to online trade affects home bias or the "natural" preference for home markets.

The absence of official and comprehensive statistics on cross-border e-commerce trade flows in goods has so far prevented empirical work on this subject, or restricted it mostly to non-tangible cross-border transactions in pure information products that can be transported across the internet, basically with zero distance-related transport costs (Blum & Goldfarb, 2006). A new online consumer survey (Civic Consulting, 2011) has generated a unique dataset on the value of cross-border e-commerce in goods in the EU. We use these data to construct a bilateral online trade matrix for the 27 EU member states, plus some major non-EU trading partners. While these survey data have by no means the same comprehensive and exhaustive coverage as standard international trade data, they enable us to start looking into the drivers and impediments of online cross-border trade in the EU.

We find that that the importance of geographical distance-related trade costs is indeed reduced in online trade, compared to offline trade. On the other hand, socio-cultural variables such as language increase in importance and counterbalance the declining cost of distance. Moreover, other sources of trade costs gain in prominence for online transactions, in particular the quality of legal institutions, payments and parcel delivery systems. Overall, home bias turns out to be equally significant online and offline. This may be due to the fact that consumers (in a B2C online trade setting) are more sensitive to these new sources of trade costs than businesses (in a B2B offline trade setting) dealing with each other in more established offline relationships.

This study introduces several innovations. First, it uses a new and unique database on online cross-border trade in goods in the EU to go beyond previous research that focused on domestic online trade in goods (Hortacu et.al., 2009) and Blum & Goldfarb (2006) who focus on international online trade in pure information services. This paper is very similar to Lendl et.al. (2012) who also apply a gravity model to online cross-border trade in goods, using a propriety eBay database. Second, it builds on the work of McCallum (1995), Wolf (2000), Coughlin & Novy (2009) and Pacchioli (2011) who measure home bias and the border effect in offline trade. We apply this framework to online international trade in goods between the 27 EU member states.

The paper is structured as follows. The next section presents a brief literature overview of the extant literature on international online trade and the use of gravity models in international trade, including the role and interpretation of the distance variable in these models. Section 3 discusses the gravity model that we apply to this trade matrix. Section 4 explains the data sources for the model. The construction of the bilateral online trade matrix is explained in Annex 1. Section 5 presents the estimation results. Section 6 summarizes and presents some policy-related conclusions.

2. Literature review

The gravity model is a well-known and often-used workhorse in international economics to explain the pattern of trade flows between countries. Tinbergen (1962) was the first to suggest that Newton's law of gravity could be applied to trade flows between economies. Deardorff (1984) notes the somewhat dubious underpinnings of the gravity model in economic theory but nevertheless concludes that it is very successful in predicting trade flows. Over the last two decades the gravity model has undergone substantial improvement and modifications, especially after Anderson & Van Wincoop (2003) provided solid theoretical underpinnings and a reformulation of the original Tinbergen model. Since then it has become widely used in trade economics and beyond. Distance is a key variable in the gravity model. It is basically a catch-all term that proxies various sources of cross-border trade costs that affect the relative price of domestic and imported goods. This may include transport costs, import tariffs, differences in technical standards and regulatory regimes between countries that induce additional trade costs, and risks related to poor institutional quality and weak contract enforcement across borders. The higher the trade costs, the less outward-oriented or the more home-biased trade patterns will be. McCallum (1995) applies the gravity model to trade between Canadian provinces and US states. He finds that although Canadian provinces are often closer to neighbouring US states than to neighbouring provinces, the US-Canada border still constitutes a significant source of trade costs and thus a barrier to trade. Wolf (2000) uses gravity to estimate home bias as an alternative measure of border costs in trade between US states. As language, culture, regulatory regimes and technical standards are pretty much similar across US states, at least much more so than between the US and other countries, one would expect the border effect or home bias to disappear in trade flows between US states. His research shows that home bias is indeed substantially lower in intra-US trade than in intra-OECD trade but remains a significant trade barrier. In the absence of regulatory differences, this may simply reflect a "natural" degree of consumer preference for local suppliers. Closeness of buyers and sellers may enhance the perception of trust, verification of product quality and easier settlement of disputes. Coughlin & Novy (2009) extend Wolf's research to cover both domestic trade between US states and international trade between US states and foreign trade partners. Somewhat surprisingly, they find that the domestic border between US states constitutes a larger trade barrier than crossing the international US border. Frankel & Wei (1993) apply the gravity model to trade between EU countries and conclude that trade costs at intra-EU borders are significant despite the fact that all these countries belong to a customs union. Pacchioli (2011) compares home bias in the US and in the EU internal market, as a proxy measure of the success of the EU's drive to complete the Single Market. She uses data on trade flows between EU member states and between states in the US. She finds a higher degree of home bias in the EU than in the US and concludes that there is still some way to go for the EU Single Market.

All the above work was conducted on traditional offline trade data. Empirical work on online trade flows has been very limited so far, mainly because of the absence of official statistics on online cross-border trade. Data are generated mainly by private companies involved in online e-commerce. Commercial interests stop them from publishing these data. A few empirical research papers have tried to circumvent this data gap and examine cross-border online operations from various indirect angles. Freund & Weinhold (2000) examine how the extent of internet penetration in countries affects their ordinary offline trade patterns. Blum & Goldfarb (2006) try to explain international internet click stream patterns using a gravity model. Their research focuses on online digital information products that can be transported across the internet at zero trade cost – anything but physical goods that need physical transport to reach the consumer. Still, they find that geographical distance plays a relevant role in these international purely digital transaction patterns. This finding is particularly true for digital products that depend on what they call "taste". Distance decreases the likelihood of a shared cultural context. For less taste and culture-dependent products distance has no statistically significant effect in their findings. This fits well with a comment by Grossman (1998) who notes that the usual order of magnitude of estimates of the distance elasticity in gravity models exaggerates the real cost of transport by at least an order of magnitude. He speculates that estimates of the distance elasticity must thus be due to other factors such as cultural differences and lack of familiarity between trading partners. Hortacu et.al. (2009) are the first to look at actual online transactions in physical goods. They take a sample of intra-US trade observations from eBay and cross-border trade from MercadoLibre to examine the importance of distance in these transactions. They conclude that distance still has an impact on trade, though less so in online than in offline transactions. Lendl et.al. (2012) uses eBay data on cross-border transactions between 62 countries for the period 2004-2007 to estimate a gravity model of online trade with several explanatory variables, including distance, transport costs, common language, border, legal regime or colonial background and quality of governance. They find that nearly all these factors generate less trade costs on eBay than in offline trade, except language and shipping costs – very much in line with our own findings.

The role of distance in international trade and the interpretation of the distance variable in gravity models of trade remains a difficult issue. Despite the decline in international transport costs and especially communication costs, the importance of distance does not appear to decline in gravity estimates over longer time periods. Disdier & Head (2008) perform a meta-analysis on 1467 estimates of the distance elasticity in gravity models. They conclude that the importance of distance decreased in international trade between 1870 and 1950 but since then it has been rising again. Berthelon & Freund (2003) show that this increase is not related to changes in the composition of trade. The importance of distance has increased for about 25 percent of all goods, mostly for homogenous goods sold on bulk exchange markets. Information on homogenous goods is easier to convey, mainly thanks to the fall in information costs. That results in a relative decline in the importance of information costs and a relative increase in the importance of distance-related transport costs. By contrast, differentiated products require more information and are thus relatively less sensitive to distance-related transport costs. Despite the fall in information costs, the ratio of distance to non-distance related trade costs for differentiated goods seems to have remained fairly constant. In contrast, Rauch (1999) argues that the fall in communication costs had a greater impact on differentiated goods than on homogenous goods.

3. The gravity model

In order to make sense out of the bilateral trade data generated by the consumer survey, we need an explanatory model for these trade flows. In line with previous research on online trade (Blum & Goldfarb, 2006; Hortaçu, 2010; Lendl et.al., 2012) we apply the well-known gravity model, the workhorse of international trade modelling. The gravity model explains the value of bilateral trade (Qij) between two countries i and j as a function of the product of the size of the two economies (proxied by GDP) and the distance (D) between them. This approach is essentially the equivalent of Newton's gravity theory from physics:

$$Qij = a\ (GDPi\ GDPj)\ /\ Dij \qquad (1)$$

The advantage of putting this model in log-log format is that the coefficients become elasticities. The value of b for instance is the percentage change in cross-border trade Qij induce by a one-percent change in GDPi.

$$\log Qij = \log a + b\ \log GDPi + c\ \log GDPj + d\ \log Dij \qquad (2)$$

Despite its strong empirical performance, this traditional version of the model had not roots in economic theory. Several authors (Deardorff (1998), Bergstrand (1990), Anderson (1979)) have tried to provide some economic foundations for the gravity model. One of the most authoritative developments was presented by Anderson & Van Wincoop (2003). They derive the gravity model starting from a consumer demand system based on a CES utility function. Consumers maximize utility by consuming goods produced in several regions, including their own, subject to a budget constraint. Prices differ between regions because of trade costs t_{ij}. From this they derive a reduced form gravity equation:

$$Q_{ij} = (GDP_i \, GDP_j \,/ \sum_i GPD_i\,)(t_{ii} \,/ P_i \, P_j\,)^{(1-s)} \quad (3)$$

$$\text{with } P_i = [\,\sum (\, b \; p_j \; t_{ij})^{(1-s)}\,]^{1/(1-s)} \quad (4)$$

where t = trade costs, b = distribution parameter and s = substitution elasticity between products from all countries.

The price index P is interpreted as a multilateral resistance term (MRT) since it depends on all bilateral resistances, not only between the trading country pair i and j. The intuition behind this multilateral term is that trade between countries i and j is not only a function of i,j specific factors but also affected by the presence of other countries. For instance, a price change in one country will affect trade flows not only with its direct trading partners but also between third countries. The impact of i on j is thus a combination of bilateral and multilateral effects. The main obstacle in the empirical application of this model is that the MRTs are not directly observable and need to be estimated using indirect methods. There are several options to do this:

(a) Starting with McCallum's (1995) path-breaking work on cross-border trade between Canadian provinces and US states, researchers have used "remoteness" as a measure of multilateral resistance, i.e. the weighted average distance between a country and all its trading partners. Anderson & Van Wincoop (2003) criticize this remoteness concept because it has no theoretical economic foundations in gravity theory. Still, it remains a relatively straightforward way to introduce MRT in gravity models, for instance Pacchioli (2011) uses remoteness measures to compare intra-EU and intra-US trade.

(b) Another way of measuring the importance of MRTs is to simply introduce a country dummy for the importer and exporter in the gravity equation. Country dummies capture omitted variables and country-specific (multilateral) effects. However, in a cross section data set that implies fixing the coefficient of GDP=1 to avoid perfect correlation between GDP and the country dummy. It requires a re-specification of the dependent variable to include the GDPs of the trading partners, in accordance with equation (3). The question is how realistic this constrained GDP coefficient is. Lendl et al (2012) use this method in the eBay study.

(c) A third option is to introduce MRTs for each explanatory variable (Baier & Bergstrand, 2009). This variable is not a dummy but calculated as the GDP-weighted average of the values for the relevant variables across all other trading partners. This method introduces an additional

constraint in the regression because the value of the coefficient of the original variable and the multilateral resistance term need to sum up to zero.
In this paper we apply and compare the results for all these specifications of the gravity equation.

Apart from the specification of the gravity model, Santos & Tenreyro (2006) question the estimation methods for gravity models. They argue that log-linearised gravity equations are potentially subject to biased estimation for two reasons: heteroskedasticity in the error term and zero values for some observations. The estimation of a stochastic version of the gravity equation with OLS assumes that the errors are normally distributed and not correlated with the explanatory variables. Heteroskedasticity tests show however that this is not always the case. Moreover, the logarithmic version of the gravity equation forces us to drop zeros and not available trade observations from the estimation because the logarithm of zero is not defined. OLS regressions simply drop all zero values for the dependent variable. This process discards information that may be useful for the regression. Santos & Tenreyro (2006) demonstrate that pseudo maximum likelihood (PML) is a less biased and more efficient estimator that avoids the problems of zero observations and heteroskedasticity. Here we apply both OLS and PML estimation methods and compare the results.

The interpretation of the coefficient for the distance variable in the gravity equation is not straightforward. Apart from transport costs directly linked to geographic distance, it may also include import tariffs, costs due to regulatory differences between countries, financial transaction costs, and information costs to bring the trading partners together in a transaction, etc. Since we are looking at intra-EU trade, there are no import tariffs on these cross-border transactions. Although VAT rates may vary across countries there is no VAT rate discrimination in function of the country of origin of the goods within the EU Single Market. VAT rates will therefore not affect bilateral trade patterns, only overall online demand for a good. The distance elasticity may also measure differences in regulatory barriers, combined with ways of getting around these barriers by switching from offline to online trade. The introduction of a legal governance quality variable may capture consumers' regulatory "regime switching" behaviour (see below).

Since goods still need to be physically transported to the consumer following an online transaction, we can assume that transport costs remain important in online trade. Online B2C trade usually implies transport of individual small parcels while offline B2B may benefit from economies of scale in large cargo consignments. Consequently, physical transport costs for goods bought online could actually be higher than offline. On the other hand, the higher number of intermediaries in offline trade (wholesalers, importers, etc) may add to offline trade costs. We have no data to compare online and offline trade costs between 27 EU member states and therefore limit the analysis to online trade costs only. We introduce as explicit parcel delivery cost variable in the gravity equation to test the importance of physical transport costs for online trade.

The gravity equation can also handle observations on domestic trade (country of origin and destination are identical, i=j). In that case, domestic distance is a measure of the size of a country. In line with the methodology applied by Pacchioli (2011), McCallum (1995) and Wolf (2000), we introduce a dummy variable for domestic trade observations in the gravity model. The coefficient on this dummy is an indicator of home bias, or the extent of consumer preference for domestic over foreign products. The home bias factor essentially measures the combined impact of all the variables that drive online (or offline) sales, including any omitted variables in the gravity equation such as "natural" preference for the home market. We calculate home bias only for online trade since we have no information on domestic sales for offline products. However, we can compare with home bias estimates for offline trade produced by other authors.

4. Data sources

While there is a wealth of statistics on ordinary offline cross-border trade, there are as yet no official statistics on online cross-border transactions. Companies involved in online trading have their own statistics on cross-border transactions but these offer only partial pictures, depending on the company's market position in different countries and product groups. In this paper we use a unique dataset from an online consumer survey in the 27 EU member states (Civic Consulting, 2011). The survey contains information on consumer online expenditure on goods, at home as well as in foreign countries. We use these data to construct a 27 x 27 bilateral online trade matrix for the EU27. We also construct an offline trade matrix between the same trading partners and for the same types of goods, so that we can compare online and offline trade patterns. See Annex 1 for a detailed explanation on the construction of the online bilateral trade matrix.

A critical issue in the construction of the online matrix is the extrapolation from survey level to population level. We use Eurostat data for the percentage of population that is connected to the internet (see table in Annex 3). However, there is a large difference between the Eurostat and the survey figures for the share of online consumers who actually buy online and buy online abroad. Since the Eurostat figures (43 and 10 percent of the population respectively) are lower than the survey figures (63 and 32 percent respectively), we stick to Eurostat to avoid overestimation. The survey figures would suggest that the EU Digital Agenda policy targets of getting 50% of all EU consumers to buy online and 20% actually shopping online abroad have already been reached in 2011; the Eurostat data suggest otherwise.

Based on the consumer survey, we estimate the total value of online traded goods in and between EU member states at about 241 billion €. Out of that total, 197 billion € (80%) is traded domestically. Only about 44 billion € (18%) crosses borders between EU member states, and another 6 billion € (2%) is imported from non-EU countries. Comparing the value of estimated online cross border trade (44 bln €) and observed offline intra-EU trade in the corresponding products categories (491 bln €) (Comext, 2010), we conclude that online trade represents about 8.7 percent of all cross-border trade in the EU. This indicates that online orders for the relevant categories of goods constitute a significant part of physical cross-border trade in goods.

The question arises to what extent the offline and online trade figures are actually comparable. On the one hand, offline and online trade involve the sale of identical consumer products: books, electronics, clothing, etc. These are final products and the trade volume is determined by consumer demand for these goods. However, the organisation of both supply chains is very different. Offline trade is mostly conducted business-to-business (B2B). Wholesalers export and import and use retailers as intermediaries before a good reaches the final consumer. By contrast, online trade is mostly B2C, with online wholesalers selling directly to final consumers. Differences in supply chains may, in turn, result in differences in the structure of the trade costs that underpin the two sets of trade flows. Wholesalers often have established relations with their foreign customers, with a fixed cost that can be amortized over many transactions. Transaction size is likely to be larger, again inducing economies of scale. Offline B2B cross-border trade figures would have to be augmented with retail gross price margins to produce a trade value figure that is comparable to direct B2C estimates. The above estimate of online B2C representing 8.7 percent of B2B cross-border trade should therefore be interpreted with caution.

Besides the bilateral online trade estimates, we need a number of explanatory variables to populate the gravity model. Distance estimates were obtained from CEPII (2007). We use distance between capitals. Domestic distances are based on the greatest circle method. EU GDP figures are taken from Eurostat (2011). On top of these standard Newtonian gravity variables we add three types of explanatory variables:

(a) Cultural and institutional variables:

Contemporary applications of the gravity trade model routinely include shared language between trading partners as an explanatory variable, and in most cases this turns out to be significant. This is meant to capture the trade costs related to "cultural distance", signalled by Blum & Goldfarb (2006) and Grossman (1998). Language may be the most important measure of cultural distance, especially in a B2C trading environment where a shared language is essential. The relative importance of language may vary by type of good. It is likely to be more important for cross-border trade in books for instance, than for electronic goods that are more or less standardized across the world. Our dataset does not allow us to separate trade by type of good however.

To measure the role of institutions in online trade, we add to our explanatory variables an indicator of the relative quality of the legal system, taken from the World Bank (2011) dataset of global governance indicators. This will capture the differences in expected trade costs related to dispute settlement between importers and exporters in online trade. One peculiar aspect of online B2C is that consumers can choose the legal regime in which they carry out their online transactions. If they don't trust the domestic legal regime, they may "vote with their mouse" and move to a foreign legal regime, especially when "voting with your feet" is far more costly because of real transport and displacement costs.

(b) Quality of the online enabling environment:

It is important to identify possible trade costs linked to the specific organisational needs of online transactions in goods. Though they may be subsumed in the catch-all "distance" variable, it would be preferable to separate these specific issues. We introduce two explanatory variables explicitly related to the overall enabling environment for online trade in goods. Consumers need to have easy access to online means of cross-border payments to settle a transaction at the lowest possible transaction cost. We capture the maturity of online payment systems in two ways. First, the market share of cash payments on delivery is considered to be an indicator of the relative underdevelopment of payments systems, combined with an absence of trust in online payments and high transaction costs (the transport of money). Compared to credit or debit card payment systems, it is a costly and risky system as it involves the transport of large amounts of cash, and transporter and consumer need to be available at the same location and same point in time. Second, the market share of PayPal is taken as a proxy of the maturity of online payment systems whereby consumers trust a non-bank financial intermediary. It may however also point to deficiencies in the local banking system so that PayPal helps consumers to circumvent these deficiencies. Credit and debit cards are widely available in almost every country and supported by the banking system. We do not take the share of credit and debit cards as an indicator. These cards are very common in all EU countries and their share of transactions is highly negatively correlated with the previous two variables. In fact, cash-on-delivery and PayPal are also negatively correlated. To avoid multi-colinearity problems we use these variables in separate regressions. Both cash-on-delivery and PayPal indicators are obtained from the World Payments Report by CapGemini et.al. (2011).

Furthermore, an efficient parcel delivery system needs to be in place to physically ship the goods from their warehouses to the consumer and to minimize physical transport costs and delivery time. As argued above, the shift from offline to online trade only reduces the information cost component of trade costs, not the physical transport cost; on the contrary, because of diseconomies of scale in parcel delivery compared to bulk cargo, physical transport costs may actually increase. We capture this by introducing a parcel delivery cost indicator: the ratio of foreign to domestic parcel delivery costs, taken from Meschi et.al. (2011). We take foreign parcel delivery costs by country pair and direction of trade. Parcel transport costs are asymmetric for a given country pair.

5. Findings

We split the discussion in several parts. First, we examine the performance of the traditional gravity model for online and offline cross-border trade, using standard explanatory variables such as GDP, distance, shared language and governance. We apply different regression methods to check for possible biases in estimation methods. We also add more specific e-commerce related variables to the model, such as parcel delivery costs and the availability of online payments systems. Second, we compare these results with the outcomes of more sophisticated specifications of the gravity model, in line with Anderson & VanWincoop (2003) and Baier & Bergstrand (2009). Finally, we compare home bias in online and offline trading. The results of the various gravity model specifications and estimations can be found in Annex 4.

5.1. The traditional gravity model

We start by applying the traditional bilateral gravity model as in equation (2) to online as well as offline trade. Besides GDP and geographical distance, we include common language and quality of governance indicators, both for online and offline bilateral trade between the EU member states. The results are presented in Table 1 in Annex 4. As can be observed, there are some important changes in the coefficients when trade is switched from offline to online platforms. The most apparent and most expected change is that the coefficient of the distance variable is about three times as high for offline (elasticity of -1.328) then for online (elasticity of -0.450) trade. Lendl et.al. (2012) in their eBay study find a similar reduction in the distance coefficient. Distance matters far less for online trade, an empirical proof of the "death of distance" that is often associated with the rise of digital information technology and the internet and the corresponding decline in information costs. We have also run these regressions for pooled online and offline trade data and find that the difference between the online and offline coefficients is statistically significant.

As the importance of geographical distance, and the high analogue information costs that come with distance, decrease, other sources of information costs become more prominent in online trade, in particular language barriers. The coefficient for shared language between trading partners increases more than eightfold between offline (0.187) and online (1.540) trade. Lendl et.al (2012) find a fourfold increase in the importance of language in their eBay study[2]. In an offline B2B trade environment with established long-term relationships, economies of scale may facilitate the amortisation of translation costs, for instance by means of translated catalogues or hiring multilingual staff to deal with foreign clients. This is more difficult in a B2C online trading environment where consumers have direct exchanges with e-merchants.

Another minor change in coefficient values between columns (1) and (2) in Table 1 concerns the GDP coefficients, with a substantial increase for the importing country coefficient and an equally important decrease for the exporting country coefficient. We hesitate to attach much significance to this change as it may be due to the fact that the online trade data are derived from a B2C consumer survey that naturally emphasizes (and possibly introduces biases towards) demand side behaviour. If any economic interpretation could be given to this, it would probably mean that the supply side becomes relatively less important in online B2C trade. The drastic reduction in information costs opens up a much larger and more varied supply of goods to consumers, with more price and quality competition between suppliers. The online world is more of a buyers market, compared to the offline world.

Table 2 presents Pseudo-Poisson Maximum Likelihood (PPML) estimates for the same gravity model specification. The Stata routine for this estimation was prepared by Santos & Tenreyro (2006). It deals with potential heteroskedasticity problems and missing or zero bilateral trade observations in log-linear gravity models. We cannot directly compare the value of the coefficients in OLS and PPML because the dependent variable is formulated in a different way (logs in OLS, levels in PPML). Moreover, none of the estimated coefficients for offline trade are statistically significant in PPML. Putting aside this latter point, we can observe the same changes in the coefficients for distance and language between online and offline as observed in the OLS regressions.

Adding proxy indicators for the cost of cross-border parcel transport and the availability of efficient online payments systems (PayPal market share for an efficient system, and cash-on-delivery market share for inefficient systems) does not boost the explanatory powers of the traditional gravity model. The OLS coefficients in Table 1 are not significant, except for PayPal. However, in the PPML regressions in Table 2, parcel delivery costs and cash-on-delivery become very significant. This may be due to the increased information content of the PPML dataset because it does not discard the zero observations. The PPML regressions reinforce the perception that parcel delivery and online payments systems are significant drivers in online cross-border trade.

The governance variable (quality of legal institutions in the importing country) follows a similar pattern. It has the expected sign but is statistically not very significant in the OLS version, but becomes much larger and very significant in the PPML version. The negative sign implies that an increase in the quality of legal institutions at home results in less cross-border trade. The

[2] The increase may be lower for eBay-specific trade because eBay has many language versions of its trading platform,

interpretation could be that consumers may not be able to vote with their feet and leave a country to carry out an offline transaction elsewhere if they don't appreciate the legal environment at home; however, they can vote with their mouse and shift their online transactions to a more reliable legal setting abroad if they are not satisfied with the environment at home. One would expect the quality legal governance to be a significant factor in online transactions because consumers are likely to have weaker social relations with their online suppliers, compared to the stronger social interaction, on-the-spot verification and less anonymity in offline transactions. More anonymous transactions require a higher degree of trust and reliance on third-party enforcement by the legal system in case a deal fails.

As the PPML regressions produce more robust results for the parcel delivery costs variable in the traditional gravity model, we can use it to check whether the distance variable still matters for online trade. The cost of transporting information declines to virtually zero in a digital economy online trading platform. Consequently, geographical distance only matters to the extent that online trade still requires distance-related transport costs to deliver the goods to the consumer. In Table 1 column (4) and Table 2 column (4) we drop the distance variable and only leave parcel delivery costs as an explanatory variable, a proxy variable for distance. In the OLS regression in Table 1 it slightly decreases the explanatory powers of the gravity model. In the PPML regression in Table 2 dropping geographical distance and leaving in transport costs does not change the log-likelihood of the PPML regression, compared to the version with distance. We can therefore conclude that geographical distance as such does not matter anymore in online trade and that it can be replaced by the cost of transporting the goods.

In Table 1 equations (1) and (2) are estimated without domestic trade since we have no domestic trade observations for comparable product categories in offline trade. In equation (7) we introduce domestic online trade, to examine whether that would change the estimates in equation (1) that covers cross-border trade only. The introduction of a dummy variable for domestic online trade allows us to quantify the well-known phenomenon of home bias in international trade: the "inherent" preference to buy on the domestic market rather than abroad. We find a domestic dummy coefficient value of 3.19, resulting in a border effect in EU online markets of 24 (exp(3.19)=24). This online home bias estimate is at the higher end of available estimates for overall offline trade. Since we do not have domestic trade values for offline trade, we can only compare our home bias in online trade with estimates for offline trade from the available

---

probably more so than the average e-retailer website.

economic literature. For example, Pacchioli (2011) compares home bias in offline trade for EU member states and US states. Depending on the specification of the gravity model, she finds border effects in the EU between 7.4 and 24: EU member states are between 7 and 24 times more likely to buy at home than in any other EU member state, considerably higher than in the US where border effects are estimated to be between 2.6 and 7. One can question whether our finding for online trade, based on a limited number of online traded consumer products, is comparable with the home bias values found for overall goods trade patterns, including consumer goods as well as intermediates and primary products. Ideally, the comparison would have to be made for a similar product composition. More research will be required to get a better understanding of the magnitude and sources of home bias in online trade, compared to offline trade. Understanding this mechanism is important for policy makers who want to boost online cross-border trade.

5.2. Enhanced versions of the gravity model

In Tables 3 to 7 we test the results for several enhancements of the traditional gravity model. These improvements are derived from the theoretical economic underpinnings of the gravity model. The question is whether they also improve the estimation results. We first reformulate the dependent variable by including GDP in it, in line with equation (3). This implies that in the gravity model, the coefficients for importer and exporter country GDP are fixed at 1. The gravity model now runs entirely on the explanatory powers of distance, language and governance variables. Tables 3 (OLS) and 4 (PPML) show that this still generates statistically significant coefficients for distance and language in many cases. However, the explanatory power of this version of the gravity model is very poor, as indicated by the very low R-squared in the OLS regressions. The log-likelihood of the PPML regressions is also significantly lower than in Table 2. The regressions in Table 5 combine the fixed GDP coefficients with country fixed effects (represented by country dummies) to capture hidden variables. The results are statistically significant for distance and language but legal governance performs poorly. Overall R-squared and log-likelihood are again significantly below those in Tables 1 and 2. In Table 6 we test for the introduction of multilateral resistance variables for each explanatory variable, in line with the gravity model proposed by Baier & Bergstrand (2009). This requires imposing an additional constraint on the regression coefficients (apart from putting the GDP coefficient equal to 1) whereby, for each explanatory variable, the sum of the coefficient of the variable and their multilateral resistance term equals 0. These constraints severely handicap the regression process.

The coefficients for distance and language turn out to be significant, except for distance in the offline OLS regression. Still, the results are somewhat puzzling, for instance the increase in the value of the distance coefficient when moving from offline to online trade in the PPML regression. The log-likelihood of the PPML regressions is also very low. We conclude that these alternative and theoretically improved versions of the gravity model perform poorly compared to the traditional gravity model. Constrained GDP coefficients and multilateral resistance terms severally handicap the regressions and do not improve the predictive powers of the model.

Finally, in Table 7 we introduce the theoretically unfounded but widely used remoteness-variable and a proxy for multilateral effects. Remoteness is calculated as the GDP-weighted average of all distances between an importer or an exporter and all its partner countries. We introduce remoteness in the traditional gravity model (first two columns of table 7), in the enhance gravity model with GDP coefficients = 1 (columns 3 and 4) and in a PPML version of the traditional gravity model (last two columns). Remoteness turns out to have some significance as an explanatory variable for the importing country in these three models. However, it is not significant for the exporting country. When it is statistically significant for the exporter (online in the PPML model) its value is zero. Remoteness can therefore be discarded as a relevant explanatory variable for online trade, except perhaps for the importing country. This could be expected since distance (or remoteness) plays a less important role in online trade.

6. Summary and conclusions

We could paraphrase Marc Twain and say that "rumours about the death of distance are greatly exaggerated". Nevertheless, there is some truth in this rumour. First, the results show that importance of geographical distance is strongly reduced in online trade, compared to offline trade, due to a drastic reduction in information costs in the digital economy that enables consumers to scan a much wider territory to satisfy their wishes and place their buying orders. However, this is compensated by an increase in the trade costs associated with crossing linguistic borders. The change in coefficient values for distance and language is confirmed across these variations of the basic model. We can even entirely discard geographical distance as a self-standing explanatory variable and replace it with parcel delivery costs, a good proxy for distance-related transport costs. Second, the quality of legal governance, the cost of parcel delivery and the efficiency of online payments systems play an important role in explaining cross-border online trade in the EU. This leaves an important policy margin for regulators to boost cross-border online trade and indicates that EU policies that aim to increase competition in parcel delivery and online payment systems are a step in the right direction. Third, the results provide a preliminary indication that home bias is not significantly different in online markets and traditional offline trade. Despite the fact that reduced information widen the market for consumers and facilitate buying abroad, consumer still have a strong tendency to buy at home. Language barriers certainly play an important role here, but other as yet unobserved variables may also be part of the explanation. In particular, we have not been able yet to examine the role of consumer perceptions for buying abroad, compared to the perceived risks of buying at home. Further research is required in this domain.

The total volume of consumer online expenditure is likely to increase over time as more consumers become more confident with online shopping and move a larger share of their shopping online. An important limit on that growth potential is the composition of the online shopping basket. The consumer survey shows that this is heavily biased towards a limited number of goods such as electronics, clothing, music/film and a few other items. The online shopping basket differs considerably from the overall consumer goods basket, probably because other types of goods do not lend themselves so easily to online trade. Further research is also needed to explain the composition and restrictions on the online consumer basket and explore ways to widen the range of goods that can be traded online. Even if the total volume of online shopping still has very considerable growth potential, the gravity model indicates that the ratio of domestic to foreign online shopping may not change that much because it is held back by linguistic fragmentation in the EU market. Since only 36 out of 702 EU27 country pairs share a common language, online retailers who want to expand their business abroad are strongly advised to have a range of language versions of their websites. However, it is difficult to see how language could become a policy variable; there is probably little that EU policy makers can do in that respect.

Another conclusion from this extensive comparison of different versions and estimation techniques for the gravity model is that the traditional gravity model, although not well-founded in economic theory, performs well compared to its theoretically improved cousins, especially when appropriate PPML estimation techniques are used. The enhanced versions of the gravity perform poorly almost across the board. The core gravity model variables (GDP and distance), combined with a few variables that relate to the specific sources of trade costs in an online B2C environment (language, governance and delivery infrastructure indicators) explains most of the observed online trade patterns. Adding more or different regulatory indicators is probably not going to affect much the relative importance of these core variables. These findings can contribute to the debate on trends in international trade costs, building on the arguments explored by Berthelon & Freund (2008) and Disdier & Head (2008).

A final word of caution. This analysis is based on a single EU consumer survey data set that offers some unique insights into the value and direction of online cross-border trade between EU countries. Obviously, these data do not have the same validity as the far more comprehensive and detailed international offline trade in goods statistics that have accumulated over the years. They offer a first insight but more effort will have to go into the construction of more comprehensive and reliable online cross-border trade data sets that will enable a more detailed and rigorous testing of the drivers and impediments to online cross-border trade. This would have to include more details on product-specific cross-border trade, transport costs, prices and information costs.

Bibliography

Anderson, J. (1979) "A theoretical foundation for the gravity equation", American Economic Review, vol 69, pp 106-116.

Anderson, J. and Van Wincoop, E. (2003) "Gravity with gravitas: a solution to the border puzzle", American Economic Review, vol 93:1, pp170-192.

Anderson, J and Van Wincoop, E. (2004) "Trade costs", Journal of Economic Literature, vol 42:3, pp 691-751

Berthelon, M. and Freund, C. (2004) "On the conservation of distance in international trade", World Bank working paper 3293.

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

Brown, J. and Goolsbee A. (2000) "Does the Internet Make Markets More Competitive?," NBER Working Papers nr 7996

Brynjolfsson, E, Smith, M and Hu, Y. (2003) "Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers," Management Science, Vol 49, No. 11

Brynjolfsson, E., A. Dick and M. Smith (2010) "A nearly perfect market", Quantitative Marketing and Economics, Springer, vol. 8(1), pages 1-33, March 2010

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

CapGemini, Royal Bank of Scotland, and EFMA (2011) "World Payments Report", available at http://gbm.rbs.com/docs/gbm/insight/gts/perspectives/WPR_2011.pdf

Goolsbee, A. and Chevalier, J. (2002) "Measuring prices and price competition online: Amazon and Barnes and Noble", NBER working paper nr 9085, July 2002.

Civic Consulting (2011) "Consumer market study on the functioning of e-commerce and internet marketing and selling techniques in the retail of goods". Available at http://ec.europa.eu/consumers/consumer_research/market_studies/e_commerce_study_en.htm

Coase, R. (1937) "The nature of the firm", Economica, vol 4(4), pp 386-405.

Coughlin, C. and D. Novy (2009) "Is the international border effect larger than the domestic border effect: evidence from US trade", Research paper series nr 2009/29, Leverhulme Centre, Nottingham University

CEPII (2010) "Gravity data set", available at http://www.cepii.fr/anglaisgraph/bdd/gravity.htm

Deardorff, A. (1995) "Determinants of bilateral trade: does gravity work in a neo-classical world", in J. Frenkel ed "Regionalisation of the world economy", Chicago University Press.

Disdier, A. and K. Head (2008) "The puzzling persistence of the distance effect on bilateral trade", The Review of Economics and Statistics, MIT Press, vol. 90(1), pages 37-48, 09

eBay (2012) "Enabling Traders to Enter and Grow on the Global Stage", available at http://www.ebaymainstreet.com/sites/default/files/eBay_Enabling-Traders-to-Enter-and-Grow-on-the-Global-Stage.pdf

Eurobarometer (2011) "Special report on Consumer empowerment", report nr 342, available at http://ec.europa.eu/consumers/consumer_empowerment/docs/report_eurobarometer_342_en.pdf

European Commission (2000) "Directive on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market (Directive on electronic commerce)", Directive 2000/31/EC of 8 June 2000 available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32000L0031:EN:HTML

European Commission (2011) Communication from the Commission "Towards Single Market Act", COM(2010) 608, available at: http://ec.europa.eu/internal_market/smact/docs/single-market-act_en.pdf

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

Frankel, J. and Wei Shang-Jin (1995) "Trading blocks and the Americas", Journal of development economics, vol 47:1, pp 61-96

Freund, C. and Weinhold, D. (2000) "The effect of the internet on international trade", Journal of international economics, vol 62:1, pp 171-189

Garciano, L., Kaplan, S. (2000) "The effect of business-to-business ecommerce on transaction costs", NBER working paper nr 8017, November 2000.

Goolsbee, A. and Klenow, P. (2006) "Valuing Consumer Products by the Time Spent Using Them: An Application to the Internet," American Economic Review vol. 96(2), pages 108-113, May

Grossman, G. (1998) "Comment" pp 29-31 in J.A. Frenkel ed "The regionalisation of the world economy", Chicago university press.

Hortacu, A., Martinez-Jerez F. and Douglas, J. (2009) "The geography of trade in online transactions: evidence from eBay and Mercado Libre", American Economic Journal: Microeconomics, vol 1, pp 53-74.

Lendle, Andreas, Marcelo Olarreaga, Simon Schropp and Pierre-Louis Vezina (2012), "There goes gravity: how eBay reduces trade costs", CEPR discussion paper, London

McCallum, J. (1995) "National borders matter: Canada-US regional trade patterns", American Economic Review, vol 85, pp 615-23.

Meschi, M., Irving, T and Gillespie, M. (2011) "Intra-community cross-border parcel delivery", FTI Consulting, London.

Milgrom, P. and Roberts (1995) "Complementarities and fit strategy, structure, and organizational change in manufacturing," Journal of Accounting and Economics, Elsevier, vol. 19(2-3), pages 179-208, April.

North, D. (1994) "Integrating institutional change and technical change in economic history: a transaction cost approach", Journal of institutional and theoretical economics, vol 150(4), pp 609-24.

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

Rauch, J. (1999) "Networks versus Markets in International Trade" Journal of International Economics, vol 48, 7-35

Tinbergen J (1962), Shaping the World Economy, The Twentieth Century Fund, New York.

Williamson, O. (1985) "The economic institutions of capitalism", Free Press, New York.

Wolf, H. (2000) "Intranational home bias in trade", Review of economics and statistics, vol 84:4, pp 555-563

World Bank (2012) "Governance indicators", database available at http://data.worldbank.org/data-catalog/worldwide-governance-indicators

ANNEXES

ANNEX 1: THE CONSTRUCTION OF THE ONLINE BILATERAL TRADE MATRIX

For the purpose of this paper we use a unique dataset from an online survey of 29.100 consumers in the 27 member states of the EU, carried out by Civic Consulting (2011) on behalf of the European Commission. The advantage of consumer survey data is that they offer a more comprehensive picture, not affected by specific market and product biases of company data. The disadvantage of this survey is that it generates no details on products and prices and only provides an overall expenditure pattern. Eurostat household and firm level survey data also offer some insights into consumer online spending and company online sales. However, they provide no information on the geographical direction of cross-border online trade; consequently, they could not be used for the purpose of this research. Another major advantage of the consumer survey is that it covers online trade in goods, a subject on which there is little information to date. At the same time, the fact that it covers goods only constitutes a major limitation. Online trade in services is probably more important than online goods trade.

The survey also makes an explicit distinction between online expenditures on domestic and foreign websites. One may question if consumers are consistently able to distinguish between these categories. The dot com or country extensions of web addresses are not always a good indicator of the actual geographical location of the supplier. Some major e-merchants have physical supply networks that are unrelated to the website addresses. That opens up several possibilities for the definition of cross-border online trade in goods. The simplest definition is an online transaction that triggers a flow of goods crossing one or more national borders. However, these may not necessarily be the same border(s) as between the country of residence of the buyer and the country of origin of the website of the seller – depending on where the warehouse of the seller is located. Another possible definition is a transaction that triggers a financial transfer across national borders, independently of the underlying physical transaction. In this study we stick to the simple definition and assume that the physical delivery and the financial transaction follow the same geographical pattern.

and in the underlying consumer survey data it is assumed that cross-border online trade means that the good crosses at least one state border.

Questionnaires were administered through computer-assisted web interviews[3]. As such, they cover consumers who have internet access and could potentially carry out e-commerce transactions, though not all of them actually do so. Overall, 85% of the respondents in the sample are online shoppers, i.e. they bought at least one product online in the twelve months prior to the survey date. This compares to an

[3] Except for Malta and Cyprus where interviews were administered by telephone.

EU average of 52% of respondents who have Internet access at home and who do online shopping, according to a recent Eurobarometer (2011) survey. Online panels may over-represent online shoppers, though there are no a-priori reasons to give more credibility to either of these surveys. It is possible that online panels are biased towards consumers who feel more at ease with computers and the internet. Oversampling is not problematic when the data are used to analyze the patterns of domestic and foreign online shopping. We should however bear the potential risks of oversampling in mind when we extrapolate the survey data to population levels to estimate the extent and total value of online shopping. We use the more conservative Eurobarometer (2011) data to extrapolate, to avoid overestimation. Panel size was approximately 1000 per country, with some variation according to country size. Country survey panels were built to ensure that all key demographic groups (e.g. gender, age, region, household size, occupation) are represented. The sample distribution for gender and age is close to the national figures.

The survey questionnaire contains information about the number of domestic and foreign online transactions over the last 12 months, the countries where cross-border transactions where made, and the amount of money spent on domestic and foreign transactions. We use this information to construct a first matrix of the sample-level value of online transactions among the 27 EU member states. Theoretically, the matrix could contain up to 27 x 27 = 729 trade observations. In practice, some cells are empty when no cross-border transactions are reported for particular pairs of countries. The survey also contains information on cross-border transaction with non-EU countries: the US, China, Norway, Iceland and Switzerland, and the residual category Rest of the World. The diagonal line of that matrix contains the value of domestic online transactions. For the non-diagonal cells (cross-border trade), we use survey information on the total amount spent per consumer on cross-border transactions over the last 12 months, and the countries where this spending took place. We calculate average spent for each consumer per cross-border transaction and apply the same average to all transactions, assuming that all cross-border transactions for a given consumer have the same value. This is admittedly a simplification but survey data do not allow us to be more specific.

Since all country survey samples have a more or less similar size of 1000 respondents, the sample-level trade matrix is not representative of the population-level trade pattern in the EU. Ten percent of respondents in Malta doing a transaction with the rest of the EU have not the same economic weight as ten percent of German respondents doing a cross-border transaction. We therefore construct a second trade matrix at population level, using the ratio of sample (online) population to total (online) population as a multiplier. Total online population figures are taken from the Eurostat (see Annex 3)

ANNEX 2: FREQUENCY DISTRIBUTION BY TYPE OF GOOD IN THE SAMPLE SURVEY

Electronics	19%
Cloth & shoes	17%
Books	10%
Music/video	6%
Cosmetics	6%
Software	6%
Electrical	6%
Toys	5%
Sports eq.	4%
Car parts	3%
Furniture	2%
Tools	2%
Medicines	2%
Other	12%
TOTAL	100%

Source: Civic Consulting (2011) and own calculations

ANNEX 3: PERCENTAGE OF POPULATION BUYING ONLINE AND ABROAD

Country	Population ('000)	%online	%online that buys	%online that buys abroad	%pop that buys online /Survey	%pop that buys online abroad /Survey	%pop that buys online /Eurostat	%pop that buys online abroad /Eurostat
Austria	8.372	80%	96%	77%	76%	62%	44	32
Belgium	10.827	83%	84%	51%	69%	43%	43	24
Bulgaria	7.576	51%	72%	34%	37%	17%	7	3
Cyprus	801	58%	57%	56%	33%	32%	21	18
Czech	10.512	73%	96%	24%	70%	18%	30	5
Denmark	5.547	91%	95%	55%	86%	50%	70	28
Estonia	1.340	77%	72%	26%	56%	20%	21	10
Finland	5.350	89%	94%	52%	84%	47%	62	28
France	64.709	80%	94%	40%	75%	32%	53	14
Germany	81.757	83%	97%	41%	81%	34%	64	9
Greece	11.125	53%	91%	57%	48%	30%	18	7
Hungary	10.013	70%	69%	15%	48%	10%	22	4
Ireland	4.468	77%	89%	74%	68%	57%	43	22
Italy	60.397	57%	86%	50%	49%	28%	15	5
Latvia	2.249	72%	70%	38%	50%	27%	20	8
Lithuania	3.329	65%	74%	32%	48%	21%	16	5
Luxemburg	502	91%	82%	75%	75%	69%	65	56
Malta	416	69%	59%	58%	41%	40%	45	38
Netherlands	16.577	92%	89%	29%	82%	27%	69	14
Poland	38.164	65%	95%	27%	62%	18%	30	2
Portugal	10.637	58%	81%	43%	47%	25%	18	7
Romania	21.466	44%	79%	19%	35%	9%	6	1
Slovakia	5.424	78%	96%	47%	75%	36%	31	11
Slovenia	2.054	69%	79%	36%	54%	25%	37	11
Spain	47.150	69%	84%	46%	58%	32%	27	9
Sweden	9.348	94%	95%	43%	89%	40%	71	16
UK	62.042	87%	97%	45%	84%	39%	71	10
TOTAL %		74%	85%	44%	**63%**	**32%**	43%	**10%**
TOTAL #	502.152	370.069						

Source: Eurostat for population and percentage online, online buyers and buyers abroad from the Civic Consulting (2011) consumer survey and from Eurostat.

Annex 4: Regression results

Table 1: The traditional gravity model, OLS estimates

LH=logCBT	Online	Offline	Online				
logGDP_M	1.144***	0.922***	1.145***	1.151***	1.134***	1.135***	1.129***
logGDP_X	0.685***	0.959***	0.679***	0.703***	0.832***	0.934***	0.670***
log_distance	-0.450***	-1.328***	-0.473***		-0.518***	-0.453***	-0.444***
comlanguage	1.540***	0.187	1.544***	1.906***	1.185***	1.202***	1.554***
loggovern	-0.039	-0.049*	-0.036	-0.045	-0.035	-0.005	-0.039
logpostcst			0.003	-0.004			
paypal					0.018***		
cash						-0.008	
home bias							3.190***
_cons	-14.368***	-8.166***	-14.256***	-17.710***	-16.085***	-17.308***	-14.042***
N	583	701	582	582	363	363	610
r2	0.689	0.812	0.689	0.675	0.745	0.741	0.743
F	296.359	426.255	245.773	247.150	209.134	200.756	388.336
ll	-1.0e+03	-1.1e+03	-1.0e+03	-1.0e+03	-588.384	-591.539	-1.1e+3

legend: * p<0.05; ** p<0.01; *** p<0.001

Variables:
LogGDP M = log GDP of the importing country
LogGDP X = log GDP of the exporting country
Comlanguage = common language dummy
Loggovern = difference of logs of quality of legal governance indicators in M and X(World Bank)
Logparcelcst = log of parcel cost = (parcelcost X->M)/(domestic parcelcost M)
Paypal = log market share of PayPal in transactions in importing country
Cash = log of market share of cash-on-delivery transactions in importing country

Table 2: The traditional gravity model, PPML estimates

LH=CBT	Online	Offline	Online			
logGDP_M	1.334***	0.710	1.330***	1.317***	1.378***	1.377***
logGDP_X	0.844***	0.668	0.836***	0.822***	1.088***	1.029***
log_distance	0.056***	-0.598	0.094***		0.103***	0.112***
comlanguage	1.530***	0.767	1.469***	1.376***	1.727***	1.726***
loggovern	-0.433***	0.016	-0.432***	-0.441***	-0.491***	-0.491***
logpostalcst			-0.007***	-0.006***		
paypal					-0.006***	
cash						-0.001
_cons	-24.979***	-20.195	-24.865***	-23.873***	-29.201***	-28.590***
N	702	702	701	701	390	390
r2						
F						
ll	-1.6e+04	-5.756	-1.5e+04	-1.5e+04	-1.1e+04	-1.1e+04

legend: * p<0.05; ** p<0.01; *** p<0.001

Table 3: The improved gravity model (GDP coefficient = 1, OLS estimates)

LH=logCBT_GDP	Online	Offline	Online			
log_distance	-0.427***	-1.304***	-0.383***		-0.551***	-0.515***
comlanguage	1.519***	0.153	1.510***	1.813***	1.162***	1.160***
loggovern	0.040	-0.054*	0.030	0.022	-0.002	0.021
logpostalcst			-0.007	-0.013*		
paypal					0.010*	
cash						-0.006
_cons	-0.261	6.590***	-0.341	-2.865***	0.097	0.218
N	583	701	582	582	363	363
r2	0.098	0.369	0.102	0.079	0.147	0.142
F	25.610	97.810	18.965	22.409	20.790	20.181
ll	-1.1e+03	-1.1e+03	-1.1e+03	-1.1e+03	-595.912	-597.050

legend: * p<0.05; ** p<0.01; *** p<0.001

Table 4: The improved gravity model (GDP coefficient = 1, PPML estimates)

LH=CBT_GDP	Online	Offline	Online			
logdistance	-0.662***	-1.257***	-0.494***		-0.515	-0.498
comlanguage	0.763*	-0.412	0.682*	1.108***	1.030	1.217*
loggovern	-0.038	-0.026	-0.073	-0.081	-0.062	-0.095
logpostcst			-0.034**	-0.045***		
paypal					-0.001	
cash						0.011
_cons	2.441**	7.100***	2.266*	-0.838**	0.808	0.414
N	702	702	701	701	390	390
r2						
F						
ll	-256.207	-387.179	-250.058	-255.453	-89.712	-89.104

legend: * p<0.05; ** p<0.01; *** p<0.001

Table 5: The improved gravity model (country fixed effects, OLS on left, PPML on right)

LH variable	logCBT_GDP		CBT_GDP	
	Online	Offline	Online	Offline
Logdistance	-0.740***	-1.349***	-0.921***	-1.410***
comlanguage	1.315***	0.657*	1.315***	1.208***
loggovern	0.078	0.102*	0.196	-0.030
_cons	2.022**	7.758***	3.508***	8.043***
Statistics				
N	583	701	702	702
r2	0.593	0.597		
F	18.579	17.876		
ll	-834.948	-930.256	-185.044	-271.934

legend: * p<0.05; ** p<0.01; *** p<0.001

PS: countries and fixed effects coefficients not shown

Table 6: Improved gravity with multilateral resistance terms for each variable (OLS and PPML)

LH variable	---- logCBT_GDP ----		------ CBT_GDP ------	
	Online	Offline	Online	Offline
logdistance	-0.065***	-0.005		
MRdistance	0.065***	0.005		
comlanguage	2.053***	1.354***		
MRlanguage	-2.053***	-1.354***		
loggovern	0.030	-0.060*		
MRloggovern	-0.030	0.060*		
_cons	-3.550***	-2.641***		
logdistance			-0.048***	0.030**
MRdistance			0.048***	-0.030**
comlanguage			1.527***	0.967**
MRlanguage			-1.527***	-0.967**
loggovern			0.041	-0.008
MRloggovern			-0.041	0.008
_cons			-2.384***	-1.178***
Statistics				
N	583	701	702	702
r2				
F	35.535	8.524		
ll	-1.0e+03	-1.2e+03	-260.411	-493.772

legend: * p<0.05; ** p<0.01; *** p<0.001

Table 7: Traditional gravity with "remoteness" as a proxy for multilateral resistance (OLS and PPML)

LH variable	---- LH=log_CBT -----		---- LH=logCBT_GDP ----		------- LH=CBT -------	
	Online	Offline	Online	Offline	Online	Offline
logGDP_M	1.126***	0.924***				
logGDP_X	0.688***	0.963***				
logdistance	-0.525***	-1.317***	-0.507***	-1.280***		
comlanguage	1.432***	0.183	1.381***	0.165		
loggovern	-0.073*	-0.047*	-0.005	-0.048*		
REMij	0.519**	-0.046	0.588***	-0.128		
REMji	-0.000	-0.000	-0.000*	-0.000		
_cons	-14.353***	-8.174***	-0.415	6.702***		
logGDP_M					1.321***	0.721
logGDP_X					0.857***	0.678
log_distance					0.071***	-0.586
comlanguage					1.368***	0.765
loggovern					-0.360***	0.039
REMij					-0.437***	-0.291
REMji					-0.000***	-0.000
_cons					-24.091***	-20.042
Statistics						
N	583	701	583	701	702	702
r2	0.695	0.812	0.122	0.370		
F	226.688	304.466	18.155	61.946		
ll	-1.0e+03	-1.1e+03	-1.1e+03	-1.1e+03	-1.5e+04	-5.753

legend: * p<0.05; ** p<0.01; *** p<0.001