There Goes Gravity: How eBay Reduces Trade Costs*

Andreas Lendle[†]
Marcelo Olarreaga[‡]
Simon Schropp[§]
Pierre-Louis Vézina[¶]

June 2013

Abstract

We compare the impact of geographic distance, a standard measure of trade costs, on eBay and offline international trade flows. We consider the same set of 61 countries and the same basket of goods for both types of transactions. We find the effect of distance to be on average 65% smaller on eBay than offline. We argue this difference is due to a reduction of search costs; It increases when information frictions are high, i.e. when trade partners speak different languages and when corruption in the exporting country is high. Moreover, eBay-ratings technology, which reduces information frictions, further reduces the distance effect on eBay. We estimate the welfare gains from a reduction in offline distance-related frictions to the level prevailing online at 4% on average.

JEL CODES: F10, F13, L81.
Key Words: Trade costs, gravity, online trade, eBay.

*We are grateful to Richard Baldwin, Christine Barthelemy, Mathieu Crozet, Anne-Célia Disdier, Jonathan Eaton, Peter Egger, Phil Evans, Simon Evenett, Lionel Fontagné, Gordon Hanson, Torfinn Harding, Beata Javorcik, Thierry Magnac, Bertin Martens, Thierry Mayer, Hanne Melin, Marc Melitz, Peter Neary, Emanuel Ornelas, Cristian Ugarte, Tony Venables, Yoto Yotov, and seminar participants at Oxford, the Villars PEGGED workshop, the WTO's Trade and Development Workshop, the University of Neuchatel, CERGE-EI in Prague, the Paris Trade Seminar, the RIEF meeting at Bocconi University, ERWIT in Barcelona, the XIII Conference on International Economics in Granada, the GTAP conference in Geneva, the ETSG conference in Leuven, the LSE International Economics Workshop, the Canadian Economic Association Meeting in Montreal, and at the Tiger Forum in Toulouse for their constructive comments and suggestions. We also thank Daniel Bocian, Steve Bunnel, and Sarka Pribylova at eBay for their time, patience and efforts with our data requests, and eBay for funding. All errors are the sole responsibility of the authors.

[†]Graduate Institute, Geneva. email : andreas.lendle@graduateinstitute.ch
[‡]University of Geneva and CEPR. email: marcelo.olarreaga@unige.ch
[§]Sidley Austin LLP. email: sschropp@sidley.com
[¶]University of Oxford. email: pierre-louis.vezina@economics.ox.ac.uk

1 Introduction

In the 1990s advances in transportation and communication technologies led many commentators to believe that geographic distance between countries would soon no longer encumber international transactions (e.g. Cairncross 1997). Despite some anecdotal evidence in support of the "death of distance" (e.g. Friedman 2005), a large number of academic papers suggests that distance is "thriving", not "dying". Disdier and Head (2008), using a meta-analysis based on 1,000 gravity equations, found that the estimated coefficient on distance has been slightly on the rise since 1950. Chaney (2011) argues that the need for direct interactions between trading partners, resulting from information frictions first highlighted by Rauch (1999), explains why distance still matters for international trade today. Similarly, Allen (2011) estimates that information frictions account for 93% of the distance effect. This would suggest that, as suggested by Leamer (2007), advances in technology in recent decades have failed to reduce information frictions between countries. Is this the death knell for the "death of distance" hypothesis?

In this paper we breathe new life into the "death of distance" hypothesis. We argue that the right place to look may be in online markets which, as opposed to "offline" markets, make full use of technologies that can reduce information frictions. Indeed, as argued by Hortaçsu et al. (2009), Goldmanis et al. (2010), and Lieber and Syverson (2012), the main benefit of the internet as a trade facilitator is to reduce search costs, and it is reasonable to think of online marketplaces as "frictionless" in this regard. Exporters no longer need to make multiple phone calls, send faxes, write emails, attend trade fairs and networking events and make contacts. And while importers still incur some search costs, these are typically brought down to a simple internet search, uncorrelated with exporters' remoteness. If, as suggested by Chaney (2011) and Allen (2011), the distance effect captures mostly information frictions, it should be much smaller on eBay.

The heart of our paper is a dataset on cross-border transactions conducted over eBay, the world's largest online marketplace. This dataset allows us to compare the effect of distance on international trade online and offline. Our approach is similar to that of Hortaçsu et al. (2009) who, using a sample of within-US eBay transactions and gravity equations,

showed that the coefficient on distance on trade was much smaller online than offline. Yet international search costs may be very different from within-US costs as information and trust frictions are probably much more important. Their sample may thus not be fully appropriate to examine the impact of internet technology in international trade.[1] Furthermore, as noted by the authors, the products traded on eBay are mainly household durables, and thus their comparison with total offline trade patterns may be problematic.

Our dataset allows us to overcome these shortcomings and compare the distance effect on eBay and offline international trade considering the same set of countries and goods. It covers all eBay transactions, disaggregated into 40 product categories, between 61 countries (representative of 92% of total world trade) during 2004-2009. To create the best-possible comparison groups, we match eBay product categories to product descriptions from the 6-digit level HS classification to build a comparable basket of goods. We also drop from our eBay data all transactions that were concluded via auctions (60% of eBay traded value), as well as those sold by consumers, so that our eBay data reflect offline practises. This matching allows us to get as close as possible to comparing apples and apples, rather than apples and oranges, which could be a problem for the existing empirical literature comparing online and offline flows. Prima-facie evidence (Figure 1) indicates that the relationship between trade flows and distance is indeed more flat-sloped on eBay than offline.

To identify the effect of distance on trade we use the gravity framework, controlling for other standard gravity trade costs such as the absence of a common language, a common legal system, a shared border, a colonial history, or a free-trade agreement. We find the distance effect to be 65% smaller online than offline. This difference in distance coefficients is statistically significant at the 99% level, robust to using OLS or Poisson pseudo-maximum likelihood estimations as well as to various aggregations of goods online and offline. What's more, we find distance to have a bigger effect offline for each and every matched product category, confirming our result is not due to poor matching. We then show that the distance-effect difference is even higher when information frictions are high, i.e. in countries that do not share a language. This supports the prediction of Chaney (2011), namely that

[1] Hortaçsu et al. (2009) do provide international evidence using MercadoLibre, another online market, though it only covers 12 Latin American countries.

in a world where search costs are greatly reduced by technology, the role of distance in explaining trade flows is smaller.

To dig deeper into the effect of eBay technology on the distance elasticity of trade, we compare established sellers with top ratings, i.e. PowerSellers, with regular sellers. The idea is that the PowerSeller certification is one of the mechanisms by which eBay reduces search costs. We find that distance does matter significantly less for PowerSellers. In other words, when information frictions are reduced thanks to consumer ratings, distance matters less. These results provide further evidence that the drop in distance elasticity is explained by a drop in search costs.

Yet, as highlighted by Goolsbee (2000) and Lieber and Syverson (2012), the demographics of eBay users may be online-specific and not representative of the offline world. Based on a Forrester Technographics US survey, they suggest that higher-income, more educated, and younger consumers are more likely to do online purchasing. It could be argued that if offline shoppers were as rich, educated, and 'international' as eBay shoppers, the offline distance effect would be as low as on eBay. In our gravity regressions, we control for the direct effect of demographics on trade by including importer and exporter fixed effects that are specific to online and offline flows. As a further robustness check using interaction terms, we show that even in the extreme country with the highest internet penetration and education and income levels, where offline consumers are most likely to be similar to eBay's, eBay technology would still significantly reduce the distance effect.

While the importance of distance is 65% smaller on eBay on average, it still matters significantly. One explanation is that, while greatly reduced, search costs are still present on eBay. Yet, according to the literature, distance may capture not only search costs but also shipping costs (Feyrer 2009) or taste similarity, as argued by Blum and Goldfarb (2006) who showed that gravity holds in the case of website visits. Using data on shipping costs from eBay and USITC, we show that shipping costs are uncorrelated with distance (Figure 3). What's more, adding shipping costs to the eBay gravity model barely affects the distance coefficient. It thus seems unlikely that the significant distance effects captures shipping costs. Rather, it is probable that the distance effect on eBay trade captures taste differences on

top of the remaining information frictions.

We conclude with an estimate of the welfare gains that could be brought about by a widespread use of internet technologies in international trade. We use the formula proposed by Arkolakis, Costinot and Rodríguez-Clare (2012) to calculate the welfare gains that would result from a drop in offline search costs to the online level, as captured by the difference in distance effects. We find that in the average country, real income would increase by 4%. Small and open countries would gain most.

The reminder of the paper is organized as follows. In section 2 we provide some descriptive statistics regarding international trade flows on eBay and describe our goods-basket construction. Section 3 presents our empirical strategy and section 4 the results. Section 5 presents the trade gains from world flattening. Section 6 concludes.

2 International trade on eBay and offline: Building a comparable basket of goods

Our data covers all eBay trade flows between 61 developing and developed countries over the period 2004-2009. These 61 countries, identified in Figure 2, represent around 92% of global offline trade in 2007. Total cross-border flows on eBay were on average USD 6 billion per year over that period, representing a small fraction (0.06%) of world trade. The correlation between the logs of bilateral offline and eBay trade is 0.71, suggesting trade patterns are geographically similar online and offline. Since we want to compare online and offline trade flows as precisely as possible, we focus on the period 2004-2007 to abstract from unusual experiences during the Great Trade Collapse of 2008-2009 (Baldwin 2009).[2] We then average trade flows over the 2004-2007 period.

Our dataset also allows us to focus on the same goods traded online and offline. It covers all eBay transactions disaggregated into 40 product categories that we match with product codes at the 6-digit level of the HS classification using information on sub-categories from

[2]The Great Trade Collapse may have come with goods shifting, trade finance problems, or new protectionist pressures that may have affected online and offline trade differently.

the eBay website (see our matching table (8)). Since it is impossible to match some eBay categories to HS codes, we dropped those goods from our eBay basket of goods. It is also important to note that the selected HS categories all fall into the "final good" category of the WTO's Trade Policy Review classification, and are all classified as "consumer goods" in the BEC classification. This is important because we do not want our offline basket to include within-business transactions on intermediate inputs as the determinants of these flows may be very different from those of final-goods trade. To improve the matching between online and offline flows we only look at eBay exports by businesses, and we ignore all imports purchased via auctions, which are prevalent on eBay but quite uncommon offline.[3]

The matching of goods is crucial as it allows us to control for differences in trade costs due to the composition of trade.[4] For example, tickets to sport-events traded online are likely to be very sensitive to distance whereas exports of rare earths, which are produced in a few countries but consumed all over the world, are not likely to be very sensitive to distance. If tickets tend to be traded online and rare earths offline, differences in the impact of distance will be explained by the different goods, and not by information technology.

To verify whether our product matching is correct, we estimate the elasticity of substitution associated with our baskets of goods online and offline. This step is important as different elasticities of substitution could also be behind the difference in distance effects (see Archanskaia and Daudin 2012). Indeed the coefficient in front of each trade-cost variable in the gravity equation is a combination of the trade elasticity (i.e. the elasticity of trade with respect to trade costs) which depends on the elasticity of substitution, and the elasticity of total trade costs with respect to each trade cost variable. Thus, a smaller coefficient on distance for online flows could simply signal that the bundle of online products has a lower elasticity of substitution than the offline bundle.

To estimate these elasticities of substitution we assume that trade costs online and offline are Gamma distributed with shape parameter k^f, where f is the type of flow, but an identical

[3]The share of sales by consumers is 66% and the share of sales through auctions is 65%. Once we exclude both, we are left with 15% of total eBay's cross border flows. As we show in our robustness checks, results hold when including all flows.

[4]See Berthelon and Freund (2008) or Carrère et al. (2009) for a discussion of the impact of the composition of trade on the role of distance.

scale parameter. Using existing estimates of the elasticity of substitution for aggregate trade flows, we can back up consistent estimates of the elasticity of substitution in our online and offline basket of goods, using the fact that the variance of a gamma distributed variable is proportional to the mean by a factor equal to the scale parameter (that we estimated using aggregate offline data). For a detailed description of the methodology, see the appendix.[5]

Our estimates suggest that for an estimate of the aggregate elasticity of substitution of 5 (see Eaton and Kortum 2012), the online elasticity of substitution is 5.6. The offline elasticity of substitution for the corresponding goods is around 5.1. The online estimate is within the [3.6 ; 5.9] range estimated by Einav et al. (2012) using intra-US trade flows and identified with differences in sales tax across states.[6] The offline estimate is quite close to the Broda and Weinstein (2006) median estimate of 5.9 in our bundle of HS-6 digit goods. Moreover, while the online and offline elasticities of substitution are statistically different from zero at the 5 percent level, they are not statistically different from each other. This comforts us in our matching of online and offline products, and suggests that statistical differences in the estimated coefficients of the gravity equation will be due to the contribution of these variables to trade costs, rather than to differences in the elasticity of substitution.

Offline trade data and trade-cost variables come from the usual sources and are described in the Data Appendix.

3 The empirical model

To examine the impact of trade costs online and offline, our starting point is the gravity model. It suggests that bilateral trade between two countries is proportional to their economic mass and the multilateral resistance indices of the importer and the exporter,[7] and inversely proportional to trade costs between the two countries, often proxied by the

[5]In the appendix we actually estimate trade elasticities, i.e., the sensitivity of bilateral trade flows to changes in bilateral trade costs. Within our CES framework the elasticity of substitution is equal to 1 minus the trade elasticity.

[6]They are significantly lower than the estimates of De los Santos et al. (forthcoming) but these correspond to price elasticities of particular book varieties, and therefore we would expect them to be higher than our aggregated bundle of goods.

[7]The multilateral resistance terms are weighted averages of price indices in the importer's and exporter's trading partners.

geographic distance between them (see Anderson and Van Wincoop (2003) for an elegant derivation):

$$(1) \quad m_{ij} = \frac{y_i y_j}{y_w}\left(\frac{t_{ij}}{P_i \Pi_j}\right)^{\epsilon}$$

where m_{ij} are imports of country i from country j, y_i is total income in importing country i, y_j is total income in exporting country j, y_w is total world income, t_{ij} are trade costs between country i and country j, ϵ is the trade cost elasticity of bilateral imports,[8] and P_i and Π_j are the multilateral resistance terms in the importing (inward) and exporting (outward) country, respectively.[9]

We follow the literature and model bilateral trade costs (t_{ij}) as a function of geographic distance and other trade cost variables:

$$(2) \quad t_{ij} = D_{ij}^{\alpha_D} e^{NB_{ij}\alpha_{NB}} e^{NC_{ij}\alpha_{NC}} e^{NCL_{ij}\alpha_{NCL}} e^{NCLS_{ij}\alpha_{NCLS}} e^{NFTA_{ij}\alpha_{NFTA}}$$

where all αs are parameters, D_{ij} is the geographic distance between countries i and j, NB_{ij} is a dummy variable taking the value 1 when countries i and j do not share a border, NC_{ij} is a dummy variable taking the value 1 when countries i and j did not share a colonial link, NCL_{ij} is a dummy variable taking the value 1 when countries i and j do not share a common language, $NCLS_{ij}$ is a dummy variable taking the value 1 when countries i and j do not share a common legal system, and $NFTA_{ij}$ is a dummy variable taking the value 1 when countries i and j are not part of the same Free Trade Agreement.[10]

[8]Given by 1 - σ in Anderson and Van Wincoop (2003) where σ is the elasticity of substitution between different import sources in the importing country.

[9]The expressions for the inward and outward multilateral resistance terms are $P_i = \left[\sum_j (t_{ij}/\Pi_j)^{\epsilon} \frac{y_j}{y_w}\right]^{1/\epsilon}$ and $\Pi_j = \left[\sum_i (t_{ij}/P_i)^{\epsilon} \frac{y_i}{y_w}\right]^{1/\epsilon}$.

[10]Note that we measure the absence of common language, common legal system, colonial links or trade agreements, rather than their presence as in most of the literature. This has no consequences for the estimates, but it allows us to interpret these variables as trade costs (like distance) rather than as trade-enhancing variables.

We then substitute (2) into (1) and take logs on both sides to obtain:

$$
\begin{aligned}
\ln(m_{ij}) = \quad & \ln(y_i) + \ln(y_j) - \ln(y_w) + \beta_D \ln(D_{ij}) + \beta_{NB} NB_{ij} + \\
& \beta_{NC} NC_{ij} + \beta_{NCL} NCL_{ij} + \beta_{NCLS} NCLS_{ij} + \beta_{NFTA} NFTA_{ij} + \\
& -\epsilon \ln(P_i) - \epsilon \ln(\Pi_j)
\end{aligned} \tag{3}
$$

where all βs are parameters to be estimated and $\beta_k = \epsilon\alpha_k$, where k is the subscript indicating the different trade cost variables. Because we are interested in understanding the variation of different βs offline and online, and because P_i and Π_i are not observable (and difficult to estimate) we proceed as in much of the empirical literature and control for the multilateral resistance terms (and y_i and y_j) including importer i and exporter j fixed effects.

A stochastic fixed-effect version of equation (3) is our baseline specification to understand the importance of different trade costs offline and online. We estimate it separately for online and offline flows, but also append the offline and online data so that we can test whether coefficients are statistically different online and offline by introducing an eBay dummy that we interact with each of the trade cost variables. In both cases we allow for importer and exporter fixed effects to be different online and offline. This captures differences in prices for online and offline products, and can also correct for a selection of buyers and sellers into online and offline platforms. We use a least-square dummy-variable estimator, but also a Poisson estimator to control for heteroscedasticity (see Santos-Silva and Tenreyro 2006).[11]

4 Results and robustness checks

Table 1 provides the results of the estimation of (3) using distance as the only trade costs in columns (1) and (5). The elasticity of distance is 65% smaller online than offline. In columns (2) and (6) of Table 1 we provide the estimates of (3) including the other usual trade costs variables. When we introduce these additional trade costs, the coefficient on distance declines both online and offline. Still it remains around 65% smaller online, suggesting a flatter world

[11] Since some of our left-hand side variables were zeros (20% on eBay, less than 1% offline), we added a dollar to the import value before taking the logs and estimating the linear model.

on the eBay platform.

Some interesting patterns emerge regarding the other trade-cost variables. Common legal systems, trade agreements, colonial links and borders seem to matter much more offline. On the other hand the absence of a common language seem to matter more online than offline. We test for the statistical significance of these differences by appending the online and offline datasets and estimating the gravity equation including interactions of each trade costs with an eBay dummy which takes a value of one if the flow on the left-hand side is the eBay flow and zero if it is the offline flow. As explained above we also include importer-eBay and exporter-eBay fixed effects that control for any country-level differences between importers and exporters online and offline. As seen in Table 2, we find that the difference in the effect of distance is statistically significant. What's more, we find that the absence of colonial links and common legal systems also matter significantly less online. Hence technology may also reduce the distortions caused by historical legacies and problems of contract enforcement across different types of legal systems. We find no significant difference in the effect of free-trade agreements, borders, or language.

Columns (3) and (7) of Table 1 add shipping costs to the set of explaining variables. Our eBay data also includes data on average bilateral ad-valorem shipping costs. While we do not have an equivalent for bilateral offline flows, in the case of US imports we do have data on freight and insurance costs from USITC. When plotting these costs against distance (see Figure 3) we find that for both online and offline flows, shipping costs are uncorrelated with distance, even though shipping costs seem to be much higher online, probably as there are less bulk-shipping scale economies for online shipments.[12] Still, we include a bilateral ad-valorem average of eBay shipping costs as a control both online and offline where it may also be a valid proxy for shipping costs. We find no significant effect for shipping costs,[13] and our distance elasticities are unaffected by this inclusion.[14]

Columns (4) and (8) provide the results using the Poisson pseudo-maximum likelihood

[12] Using data on all country pairs online gives a similar picture.

[13] This could be explained by endogeneity problems, as larger trade flows lead to economies of scale.

[14] This result also suggests that the death of distance online is not due to a reduction in shipping costs. Adding other controls such as bilateral average tariffs or trade-restrictiveness indices does not affect the results (not shown).

estimator which was suggested for gravity models by Santos Silva and Tenreyro (2006) to control for zero trade flows in the double log specification of the gravity equation and heteroscedasticity. Again we find that distance matters more offline. The estimated distance elasticity is around 45% smaller online.

To check whether our result might be driven by a composition effect within the online and offline bundles, we estimate gravity equations for each eBay category using the specification of column (2) of Table 1. The estimated coefficients, using both linear and Poisson pseudo maximum likelihood estimators, are summarized in Figure 4. For each product category, distance has a bigger effect offline. It is on average 60% smaller online. Pooling the product regressions together and estimating an average effect using importer-category and exporter-category fixed effects yields distance coefficients of -0.287 online and -1.167 offline, thus 75% smaller online.

In Table 3 we include the results of various robustness checks. As an important part of eBay trade is in used goods (25%) or occurs through auctions (65%) we replicate Table 1 disaggregating imports into used vs. new goods (this is done on a 2008 cross section because it is the only year for which we have the used versus new good information) and auctions vs. direct sales. We also report results when looking at all trade flows reported on comtrade, i.e. not just the eBay image, as well as all eBay trade flows and not only those that match offline products. Results are consistent across aggregations, suggesting that distance matters less than online across all types of eBay flows.[15]

The final two columns of Table 3 verify whether eBay seller reputation matters for the impact of distance on trade flows. Online platforms adopt mechanisms to overcome the incentives for opportunistic behavior in global markets where buyers and sellers do not necessarily meet repeatedly. The eBay PowerSeller status is one of these mechanisms.[16] It certifies that the seller has received 98% positive feedback, has been active for more than 90 days, has completed at least 100 transactions or transactions worth at least $3000 during

[15]We also run the same specification for sales by non-business exporters (e.g. consumers) and perhaps surprisingly found a similar distance elasticity as for business flows of around -0.5.

[16]Another important mechanism is the disclosure of information through photos and text. Lewis (2011) shows that they strongly influence auction prices on eBay motors as they help define the terms of the contract between sellers and buyers who cannot directly observed the goods they are buying.

the past year, and complies with eBay policies.[17] Seller reputation is in principle much more important than buyer reputation on eBay as transactions are usually of the "cash-in-advance" type where the buyer pays first and waits for the seller to send the goods.[18] The last two columns of Table 2 look at whether the impact of distance on trade flows is different for PowerSellers and non-PowerSellers. If the distance coefficient captures information frictions, and if the PowerSeller mechanism were to be effective, then we would expect a smaller distance coefficient for transactions undertaken by PowerSellers. As predicted, we find that distance affects non-PowerSellers more. We test for the statistical significance of the difference on the distance coefficient of PowerSellers by appending the PowerSeller and non-PowerSeller data and interacting each of the trade cost variables with a dummy indicating whether the flow involves PowerSeller or not. The only statistically-different coefficient at the 5% level is the one on distance as shown in Table 4. This reinforces the result that eBay technology reduces the distance coefficient by reducing information frictions.[19]

To examine whether eBay reduces search costs associated with product information as suggested by Rauch (1999), we use Broda and Weinstein's (2006) estimates of elasticity of substitution (sigma). The median of their HS-6 digit estimates measures the need for information, or the level of product differentiation, within each category. As substitution among import sources is smaller there is a stronger need for product information. In Table 5 we thus interact the distance coefficient with the sigmas associated with the HS codes associated with each eBay product category. The results are summarized in the middle panel of Figure 5. It shows that offline, the distance elasticity increases with the level of product

[17] See eBay's website for more details here: http://pages.ebay.com/sellerinformation/sellingresources/powerseller.html.

[18] Bohnet et al. (2005), using lab experiments, have argued that it is not only buyers learning about sellers but also sellers learning about other sellers that is an important ingredient of well-functioning markets prone to moral hazard. eBays feedback mechanism could thus facilitate exchanges through both these mechanisms. See Cabral and Hortaçsu (2010) for a recent analysis of the consequences of seller reputation on eBay.

[19] Hortaçsu et al. (2009) also tried to show that seller reputation lowered the distance elasticity of trade. However, their results were not conclusive. When estimating gravity equations for sales for bad and good sellers separately, they found that distance mattered slightly more, rather than less, for sellers with good reputations. In another specification, they computed a measure of bad-reputation at the State level based on the median-rating distribution and found that distance mattered more for bad-reputation States, thus conflicting with their previous result. This discrepancy is probably due to fact that when computing the bad-reputation State measures, the authors end up capturing any negative aspect of a State rather than strictly eBay-sellers' reputation, thus failing to identify the ratings effect on the distance coefficient.

differentiation, as was shown by Rauch (1999), whereas on eBay product differentiation barely matters for the distance effect. This suggests that eBay reduces the distance effect most when search costs are high. We also interact distance with measures of information frictions at the country level, namely corruption in the exporting country and language sharing. The logic is the same. If eBay reduces the distance effect by reducing information frictions, it should do so most when those frictions are high, and these are likely to be high when the seller is located in a country characterized by corruption problems or when the buyer speaks a language different than the one in the exporting country. Indeed, as Leamer (2007) wrote, "Physical distance may create and reinforce linguistic and cultural barriers that make it difficult to exchange thoughts between people located far from each other in the cultural landscape." The results are summarized in Figure 5. Offline, the distance elasticity is significantly higher when the exporting country suffers from corruption problems and when partners speak different languages. On eBay distance matters the same across country pairs. The main message is thus that eBay reduces the distance effect to the offline level prevailing in "clean" countries or between partners sharing a language. In other words, the reduction in the distance effect is indeed highest when information frictions are high.

4.1 Selection bias

As mentioned earlier, the difference in the effect of distance could be due to a selection of 'international' buyers rather than a 'technology' effect. While the appended model including importer-eBay and exporter-eBay fixed effects partly corrects for these selection effects, buyer and seller characteristics might also affect the impact of distance. For example, eBay buyers may tend to be richer and rich individuals may prefer purchasing goods from faraway countries. Ideally, we would like to observe individual characteristics of buyers online and offline, but we do not have access to that data. According to Lieber and Syverson (2012), higher-income and more educated individuals are more likely to do online purchasing. We thus check whether the distance effect is statistically different online and offline even in 'extreme' countries where offline consumers are most likely to be similar to online ones, i.e. in countries with high internet penetration, high income, low inequalities, and high education

levels. The idea is that in highly unequal societies with low internet penetration only a few privileged 'international' buyers have access to internet and buy on eBay. In these countries buyers on eBay and offline are likely to be most different. As reported in Figure 6 (drawn from Table 6), we find the biggest differences in distance effects in poorer, less educated, more unequal countries and in countries with low internet penetration, suggesting part of the difference may reflect a selection of 'international' buyers online. Still, we find that even for the most equal or most internet-penetrated countries, where the online and offline buyers are plausibly most similar, the distance effect is still statistically smaller online. This reinforces the idea that technology has a distance-reducing impact beyond importer selection.

5 Welfare gains

The reduction in search costs in online markets may thus be behind a significant reduction in the distance elasticity of trade. But would this reduction have any significant impact in terms of welfare gains? In other words, how richer would the world be in a hypothetical situation characterized by a widespread use of internet technologies in international trade? In order to estimate the welfare gains that would result from search costs being reduced to the level on online platforms, i.e. if distance mattered offline as little as online, we follow the General Equilibrium Trade Impact estimation procedure suggested by Head and Mayer (2013). We first need to calculate the change in intranational trade shares in each country using our gravity estimates. We can then compute the changes in real income following Arkolakis, Costinot and Rodríguez-Clare (2012). Indeed, according to their proposition 1, assuming that trade is balanced, that the ratio of profits to total income is constant, and that the import demand system is such that bilateral trade flows are given by a gravity specification consistent with the presence of a single production factor (labor), we can express the welfare change as:

$$(4) \quad \widehat{W}_i = \left[\frac{\widehat{m_{ii}}}{y_i}\right]^{1/\epsilon}$$

where, for any variable x, $\widehat{x} = x'/x$, and x' is the value of x after the shock. The change in intranational trade as a share of income is given by (see Proposition 2 in Arkolakis et al. 2012):

$$(5) \quad \frac{\widehat{m_{ii}}}{y_i} = \frac{\left(\widehat{w}_i \widehat{t_{ii}}\right)^{\epsilon}}{\sum_{j=1}^{n} \frac{m_{ij}}{y_i} \left(\widehat{w_j} \widehat{t_{ij}}\right)^{\epsilon}}$$

Hence, in order to calculate the change in welfare associated with a partial "death of distance" offline, we need an estimation of the change in trade costs ($\widehat{t_{ij}}$), as well as an estimation of the change in wages ($\widehat{w_j}$) in all n countries. The former can be obtained using the estimates of the distance coefficient online and offline:

$$(6) \quad \widehat{t_{ij}} = e^{\frac{1}{\epsilon}\left(\beta_D^{\text{online}} - \beta_D^{\text{offline}}\right) \ln D_{ij}}$$

We use the β_D coefficients reported in columns (4) and (8) of Table 1 which have been consistently estimated using importer and exporter fixed effects specific to online and offline flows and a Poisson estimator to control for heteroscedasticity. We can then easily compute $\widehat{t_{ij}}$ using an estimate of ϵ for aggregate trade flows from the existing literature. Eaton and Kortum (2012) suggest that the current best estimate sets $\epsilon = -4$. Note that this assumes that the elasticity of substitution for aggregate trade flows is identical to the elasticity of substitution for online flows and offline matched flows, as the β_D coefficients have been estimated using online and offline matched flows. We test this hypothesis in the Appendix.

The estimation of $\widehat{w_j}$ requires solving the general equilibrium wages of all countries in our sample. The change in wages in all other countries are implicitly given by (see Arkolakis et al. 2012):

$$(7) \quad \widehat{w_j} = \sum_{i'=1}^{n} \frac{m_{i'j} \widehat{w_{i'}} \left(\widehat{w_j} \widehat{t_{i'j}}\right)^{\epsilon}}{y_j \sum_{j'=1}^{n} m_{i'j'}/y_i' \left(\widehat{w_{j'}} \widehat{t_{i'j'}}\right)^{\epsilon}}$$

We solve the n non-linear equations for the changes in wages ($\widehat{w_j}$) using the Stata code provided by Head and Mayer (2013). Substituting these and the estimates of the changes in trade costs in equation (6) into (5) and the result into (4) yields the changes in real income following a drop in the distance effect offline to the level prevailing online.

The welfare-gains gains per country are given in Table 7. The increase in real income associated with a reduction in distance-related search costs for all trade flows is on average equal to 4%, ranging from over 43% for Taiwan to 0.86% for the US. Hence, our results suggest that potential gains from the reduction in search costs brought about by online platforms are quite large. We then try to explain the variance in welfare gains across countries. We regress welfare gains on GDP, GDP per capita, remoteness, and openness. We find remoteness and openness to be statistically significant. As shown in Figure 7, the largest welfare gains would occur in the most open and remote countries.

6 Concluding Remarks

In his review of Friedman (2005), Leamer (2007) argues that advances in technology in recent decades have failed to reduce information frictions between countries. Humans are still like animals and cannot trust each other unless in the same physical space. Geography thus creates special relationships between buyers and sellers who reside in the same neighborhoods and this explains why the distance effect on international commerce is "possibly the only important finding that has fully withstood the scrutiny of time and the onslaught of economic technique.

Using a dataset on eBay cross-border transactions and comparable offline trade flows, we estimated a distance effect on trade flows about 65% smaller online than offline. We argued this difference in distance effects was due to online technologies that reduce search costs associated with geographic distance. Importantly, the welfare gains from the reduction in distance related trade costs are large. In a hypothetical world where information frictions offline were reduced to the level prevailing online, real income would increase by 4% in the average country.

References

Allen, Treb (2011). Information Frictions in Trade, Job-Market paper, Yale University.

Anderson, James and Eric van Wincoop (2003). Gravity with gravitas: a solution to the border puzzle, American Economic Review 93, 170-92.

Arkolakis, Costas, Arnaud Costinot and Andrés Rodríguez-Clare (2012). New Trade Models, Same Old Gains?, American Economic Review, forthcoming.

Archanskaia, Elizaveta and Guillaume Daudin (2012). Heterogeneity and the Distance Puzzle, FREIT Working Paper 448

Baldwin, Richard (2009). The great trade collapse. VoxEU.org Publication.

Berthelon, Matías and Caroline Freund (2008). On the Conservation of Distance in International Trade. Journal of International Economics 75(2), 310-310.

Blum, Bernardo, and Avi Goldfarb (2006). Does the internet defy the law of gravity? Journal of International Economics 70, 384-405.

Bohnet, Iris & Heike Harmgart & Steffen Huck & Jean-Robert Tyran, (2005). "Learning Trust," Journal of the European Economic Association, MIT Press, vol. 3(2-3), pages 322-329, 04/05.

Broda, Christian and David Weinstein (2006). Globalization and the Gains from Variety, The Quarterly Journal of Economics 121(2), 541-585.

Cabral, L. and A. Hortaçsu (2010). The dynamics of seller reputation: evidence from eBay. The Journal of Industrial Economics, 58(1), 54-78.

Cairncross, Frances (1997). The Death of Distance. Cambridge: Harvard Business School Press.

Carrère, Céline, Jaime de Melo and John Wilson (2009). The Distance Effect and the Regionalization of the Trade of Developing Countries. CEPR Discussion Paper 7458.

Chakravarti, Laha, and Roy (1967). Handbook of Methods of Applied Statistics, Volume I, John Wiley and Sons, 392-394.

Chaney, Thomas (2011). "The Gravity Equation in International Trade: an Explanation", mimeo.

De los Santos, Babur, Ali Hortaçsu and Matthijs Wildenbeest (forthcoming). Testing models of consumer search using data on web browsing and purchasing behavior. American Economic Review.

Disdier, Anne-Célia and Keith Head (2008). The Puzzling Persistence of the Distance Effect on Bilateral Trade. Review of Economics and Statistics 90(1), 37-48.

Eaton, Jonathan and Samuel Kortum (2012). Putting Ricardo to work. Journal of Economic Perspectives 26(2), 65-90.

Einav, Liran, Dan Knoepfle, Jonathan Levin and Neel Sundaresan (2012). Sales taxes and internet commerce. NBER Working paper 18018, National Bureau of Economic Research, Boston.

Feyrer, James (2009). Distance, Trade, and Income : The 1967 to 1975 Closing of the Suez Canal as a Natural Experiment. NBER Working Paper 15557, National Bureau of Economic Research, Boston.

Friedman, Thomas L. (2005). The world is flat. Farrar, Straus & Giroux.

Goldmanis, M., Hortaçsu, A., Syverson, C., and Emre, O (2010). E-Commerce and the Market Structure of Retail Industries. Economic Journal 120, 651-682.

Goolsbee, Austan (2000). In a world without borders: the impact of taxes on internet commerce. Quarterly Journal of Economics 115(2), 561-576.

Head, Keith and Thierry Mayer (2013). Gravity Equations: Workhorse, Toolkit, and Cookbook, CEPR Discussion Paper 9322

Hortaçsu, Ali, A. Martinez Jerez, and Jason Douglas (2009). The geography of trade in online transactions: Evidence from eBay and MercadoLibre", American Economic Journal: Microeconomics 1(1), 53-74.

Kaufmann, Daniel, Kraay, Aart, and Mastruzzi, Massimo, (2010) "The worldwide governance indicators : methodology and analytical issues," Policy Research Working Paper Series 5430, The World Bank.

Leamer, E.E. (2007), A Flat World, a Level Playing Field, a Small World After All, or None of the Above? A Review of Thomas L Friedman's The World is Flat." Journal of Economic Literature, 45(1): 83-126.

Lewis, Gregory (2011). Asymmetric information, adverse selection and online discolsure: The case of eBay motors. American Economic Review 101(4), 1535-1546.

Lieber, Ethan and Chad Syverson (2012) "Online vs. Offline Competition," Oxford Handbook of the Digital Economy, Martin Peitz and Joel Waldfogel (eds.).

Rauch, James E. (1999). Networks versus markets in international trade. Journal of International Economics 48(1), 7-35.

Santos Silva, J. M. C. and Silvana Tenreyro (2006). The Log of Gravity. The Review of Economics and Statistics 88(4), 641-658

Data Appendix

Below we discuss variable construction and data sources for all variables used in the empirical sections. The appendix Table provides descriptive statistics for each variable.

- Distance (D): Distance between two countries based on bilateral distances between the largest cities of those two countries, those inter-city distances being weighted by the share of the city in the overall country's population. Source: CEPII Distances database.

- Shipping cost (T): Ad-valorem shipping costs as a share of product price (logged). Source: eBay.

- No Border (NB): dummy variable indicating whether the two partners share a border. Takes the value 1 when the two partners do not share a border. Source: CEPII Distances database.

- No Colony (NC): dummy variable indicating whether the two countries have ever had a colonial link. It takes the value 1 when the two trading partners do not share a colonial link. Source: CEPII Distances database.

- No Common Language (NCL): dummy variable indicating whether the two countries share a common official language. It takes the value 1 when the two trading partners do not share a common language. Source: CEPII Distances database.

- No Common Legal System ($NCLS$): dummy variable indicating whether the two countries have the same legal origin. It takes the value 1 when the two partners do not share a legal origin. Source: CEPII Gravity database.

- No FTA ($NFTA$): dummy variable indicating whether the two countries have a free-trade agreement declared at the WTO. It takes the value 1 when the two partners do not have a free-trade agreement. Source: WTO.

- Corruption (C): Negative of control-of-corruption which captures perceptions of the extent to which public power is exercised for private gain, including both petty and

grand forms of corruption, as well as "capture" of the state by elites and private interests. Source: Kaufmann et al. (2010).

- Trade elasticity (sigma): Elasticity of substitution within HS-6 product categories. Source: Broda and Weinstein (2006).

- Internet penetration (@): Number of internet users over population. Source: World Bank World Development Indicators.

- Gini (Gini): Gini coefficient of income inequality. Source: World Bank World Development Indicators.

- Education (Educ): Gross enrolment ratio. Tertiary (ISCED 5 and 6). Total is the total enrollment in tertiary education (ISCED 5 and 6), regardless of age, expressed as a percentage of the total population of the five-year age group following on from secondary school leaving. Source: World Bank World Development Indicators.

- Remoteness (Remote): Inverse of a weighted sum of trading partners' GDPs, where the weights are inverted distances. Source: CEPII and World Bank World Development Indicators.

- Openness (Open): Sum of imports and exports over GDP. Source: World Bank World Development Indicators.

- GNI per capita (GNIPC): Gross National Income per capita, current $. Source: World Bank World Development Indicators.

Appendix: Estimates of trade elasticities online and offline

In this section we estimate the trade elasticities of the online and offline baskets. This is important because the estimated effect of distance on trade flows (β_D) is a combination of the trade elasticity (ϵ) and the impact of distance on trade costs (α_D). The lower estimated impact online could be therefore explained by a lower trade elasticity for the online bundle, rather than a smaller impact of distance on trade costs.

Our identification strategy is as follows. We assume that the log of aggregate (offline) trade costs, eBay trade costs, and eBay-image trade costs are all gamma distributed with the same scale parameter θ, but with a different shape parameter (k_f) that is specific to each flow. More formally, we assume that trade costs $\ln(t_{ij}^f)$ of each flow f are generated by a gamma distribution with scale parameter θ and shape parameter k^f:

$$(8) \quad \ln(t_{ij}^f) \sim \frac{1}{\theta^{(k^f)}\Gamma(k^f)}(\ln(t_{ij}^f))^{k^f-1}e^{\frac{-\ln(t_{ij}^f)}{\theta}}$$

We do not observe trade costs directly, but the product of the trade elasticity with the log of trade costs for each type of flow. This product can be consistently estimated using equation (3) for each type of flow. To backup the scale parameter (θ) that we assume is common to the three type of trade costs, we need an estimate of the trade elasticity for aggregate trade flows that we borrow from the existing literature. Eaton and Kortum (2012) suggest a consensus estimate of -4. Dividing by -4 the estimated product of the trade elasticity and the log of trade costs for aggregate flows, we then have the empirical distribution of aggregate trade costs. To obtain the scale parameter θ, we then use a convenient property of gamma distributions which implies that the variance-to-mean ratio is equal to the scale parameter.[20]

To estimate the trade elasticities for eBay and eBay-image flows, we use a second convenient property of gamma distributions, i.e. if $\ln(t_{ij}^f) \sim \text{Gamma}(\theta, k^f)$, then $\epsilon^f \ln(t_{ij}^f) \sim$

[20] The mean of a gamma distribution is equal to $k^f\theta$ and its variance is given by $k^f\theta^2$.

Gamma($\epsilon^f\theta, k^f$). Combining this with the first property mentioned above, we have that the variance-to-mean ratio of the the estimated product of $\epsilon^f \ln(t_{ij}^f)$ for eBay and eBay-image is equal to $\epsilon^f\theta$. Thus,

$$(9) \quad \epsilon^f = \frac{\text{var}\left[-\epsilon^f \ln(t_{ij}^f)\right]}{\text{mean}\left[-\epsilon \ln(t_{ij})\right]\theta}$$

Using 9 we have have estimates of ϵ^f for eBay and eBay-image trade flows that we can compare to the estimate of -4 on aggregate trade flows that we borrow from the existing literature.

Before proceeding to empirically testing our method, we check that the empirical distribution of $\ln(t_{ij})$ for aggregate trade flows fits a gamma distribution. Figure 8 shows the kernel density estimate of this empirical distribution. A single gamma distribution clearly does not fit our empirical estimates of $\ln(t_{ij})$, as the distribution is bimodal. This is not new and is explained by the fact that there at moderate trade costs (those that would occur over distances between continents) we have very few observations. A simple way of solving this is to assume that the bimodal distribution of trade costs is due to the combination of two unimodal gamma distributions for trade costs within a continent, and across continents. Using Figure 8 we then assume that the split between the two distributions occur when the log of trade costs is around 1.75.

We then estimate θ for each of these two gamma distributions of aggregate trade flows by taking the variance-to-mean ratios.[21] This yields $\theta = 0.013$ for (log) trade costs below 1.75 and $\theta = 0.004$ for (log) trade costs above 1.75.

Using these θ estimates and equation (9), we can estimate the trade elasticities for eBay and eBay-image flows. For low trade costs, the trade elasticity of eBay flows equals -4.6 and of image flows -4.1. For high trade costs, the trade elasticity of eBay flows equals -6.6 and of image flows -4.2. Thus the trade elasticity for image flows is almost identical to the trade elasticity on aggregate trade flows. For eBay flows the trade elasticity if anything is larger

[21] Note that this assumes that the trade elasticity is invariant to trade costs which is consistent with existing estimates using aggregate trade costs.

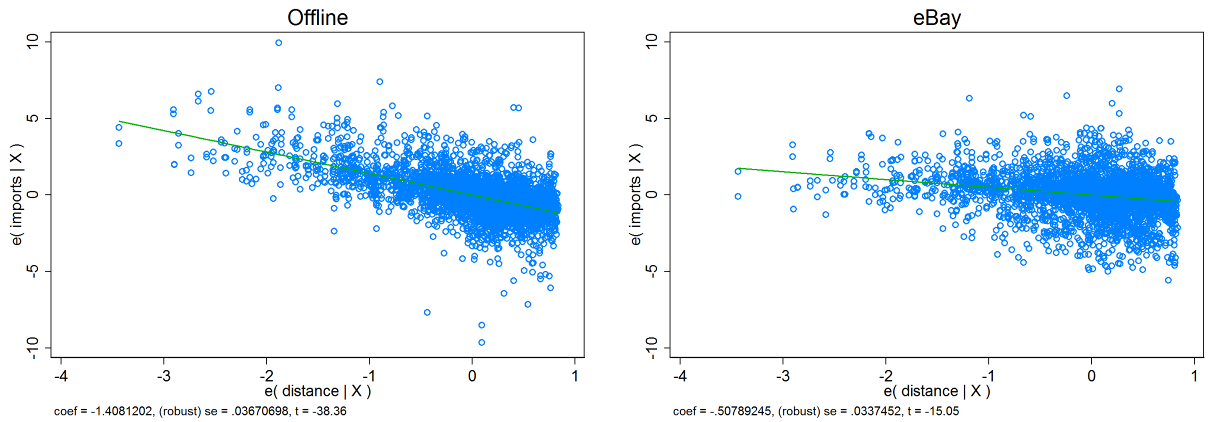
than for aggregate trade flows or eBay-image, which suggests that a lower trade elasticity for eBay trade flows cannot explain why we observe a lower estimated impact of trade costs variables such as distance on online trade flows.

To check that the trade elasticity estimates are statistically different from zero, but not statistically different from each other, we construct bootstrapped standard errors taking into account the sampling error as well as the error associated with the offline aggregate elasticity-of-substitution estimates. For the latter, we assume that ϵ is normally distributed with mean -4 and a variance equal to 1, which yields a 95% confidence interval of $[2; 6]$, which practically covers all of the estimates surveyed in Eaton and Kortum (2012). We allow for 500 repetitions and allow two-way clustering of standard errors within each importing and exporting country. This procedures yields a standard error equal to 1 for both estimated trade elasticities of image flows (i.e. for large and small trade costs). The standard errors for eBay flows are equal to 1.7 for large trade costs and 11 in the case of small trade costs. This suggests that none of the trade elasticity estimates is statistically different from -4, or from each other.

Finally, we can test the assumption that $\ln(t_{ij})$ of aggregate trade flows is gamma distributed with a scale parameter $\theta = 0.013$ for values of $\ln(t_{ij})$ below 1.75, and gamma distributed with a scale parameter $\theta = 0.004$ for values of $\ln(t_{ij})$ above 1.75, using a Kolmogorov-Smirnov test of equality-of-distributions.[22] The values of the two Kolmogorov-Smirnov statistic (D) are close to zero (with p-values of 0.8 for high trade costs, and 0.3 for low trade costs) and therefore we cannot reject at the 5% level the null hypothesis that $\ln(t_{ij})$ is gamma distributed with a scale parameter $\theta = 0.013$ for values of $\ln(t_{ij})$ below 1.75, and gamma distributed with a scale parameter $\theta = 0.004$ for values of $\ln(t_{ij})$ above 1.75.

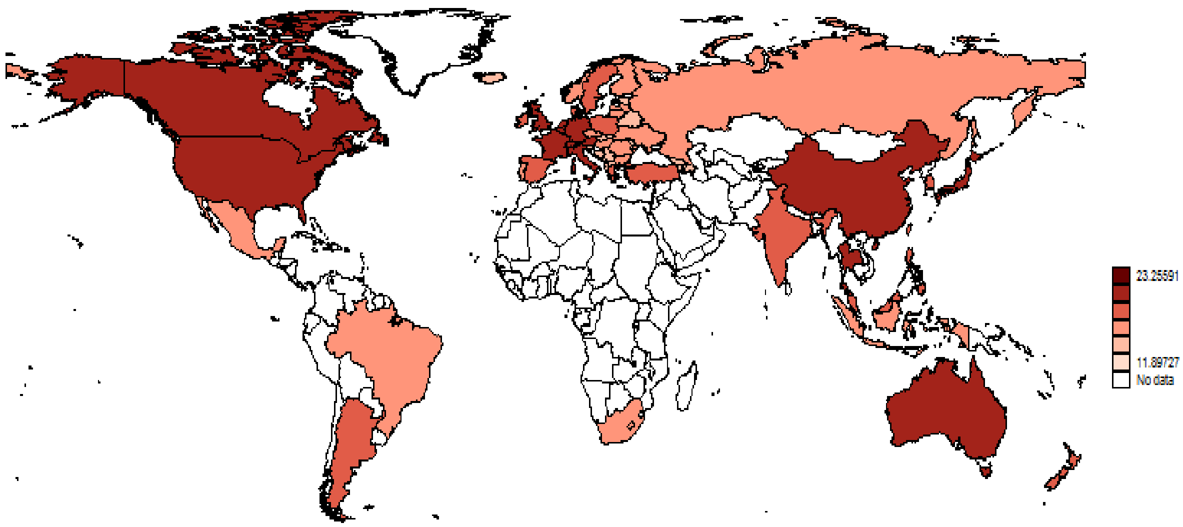
[22]For a discussion of the Kolmogorov-Smirnov test see Chakravarti, Laha, and Roy (1967).

Figure 1
The importance of distance with and without search costs



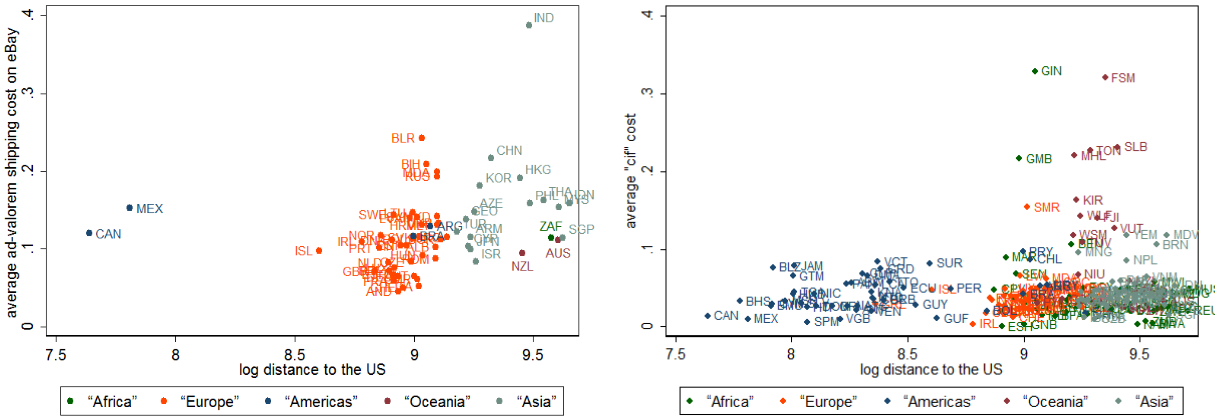
Note: Offline bilateral trade data is from UN Comtrade for 61 countries which represent more than 92% of world trade and is restricted to the set of goods which are traded on the eBay platform. eBay bilateral trade data is from eBay for the same set of countries.

Figure 2
Country coverage



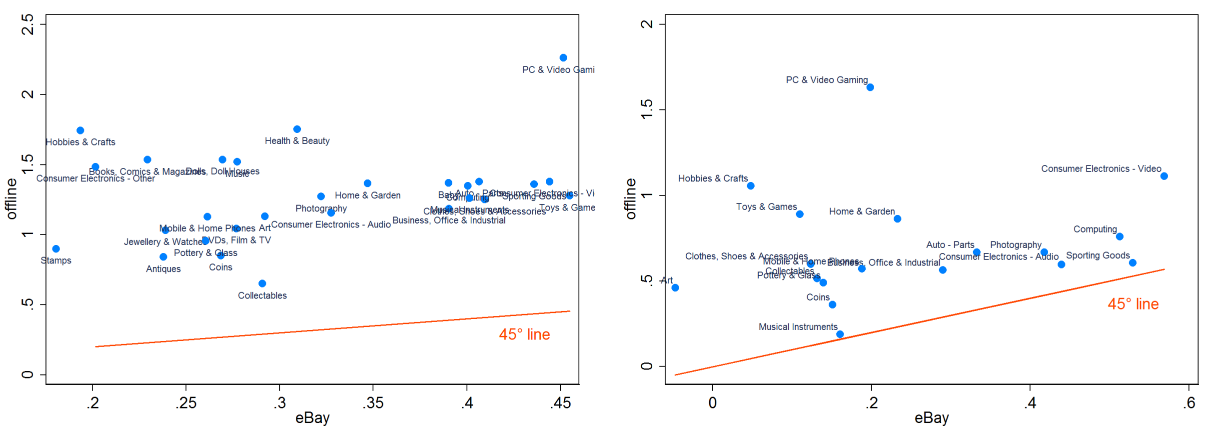
Note: The intensity of the red color signals the value of the log of eBay exports

Figure 3
Distance and shipping costs offline and online



Sources: USITC and eBay

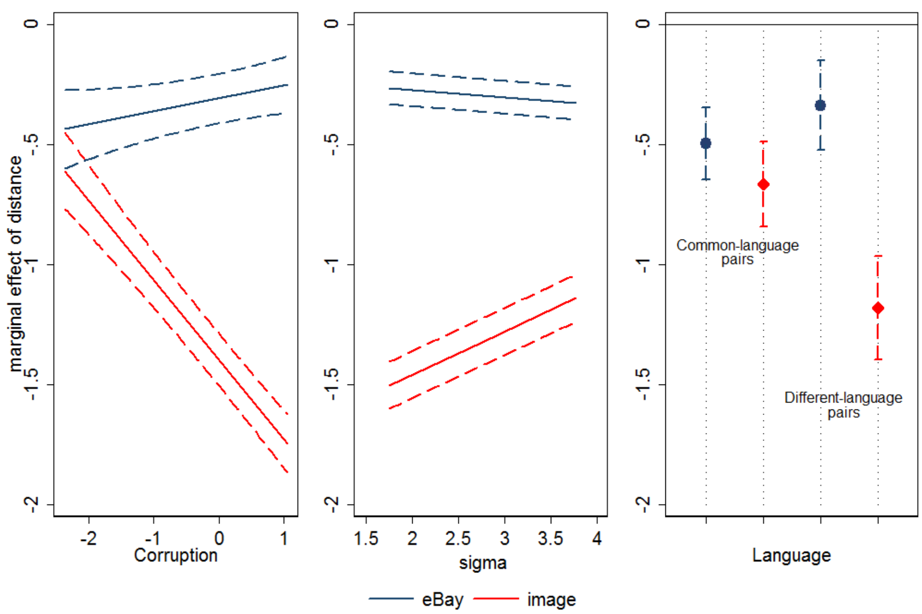
Figure 4
Distance coefficient by eBay category



Note: The left panel reports estimates using an OLS estimator and the right panel reports estimates using a poisson estimator. Each distance coefficient is estimated in a separate regression with a specification identical to the one reported in column (2) of Table 1.

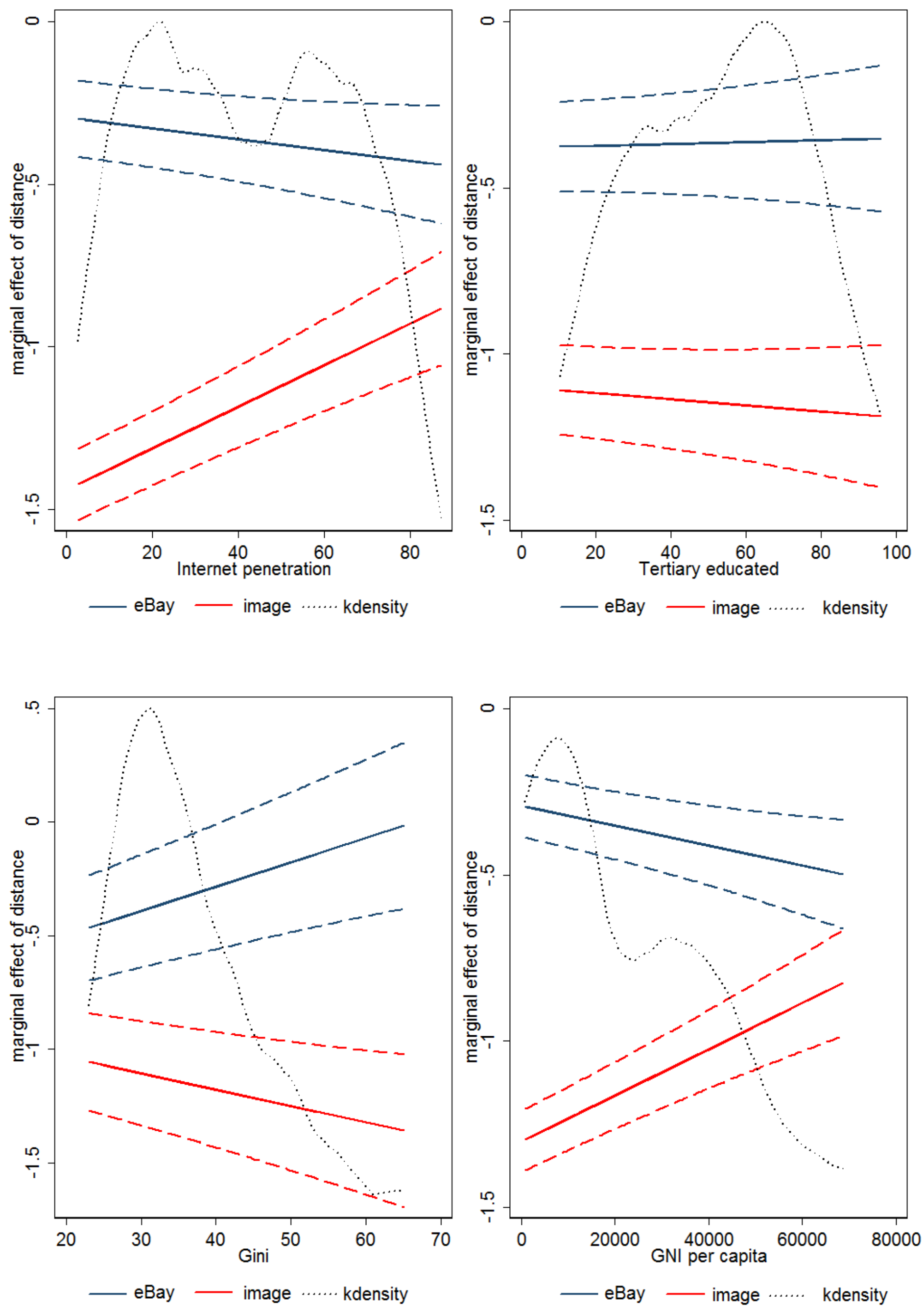
Figure 5

Distance coefficients by exporter corruption, elasticity of substitution, and language sharing



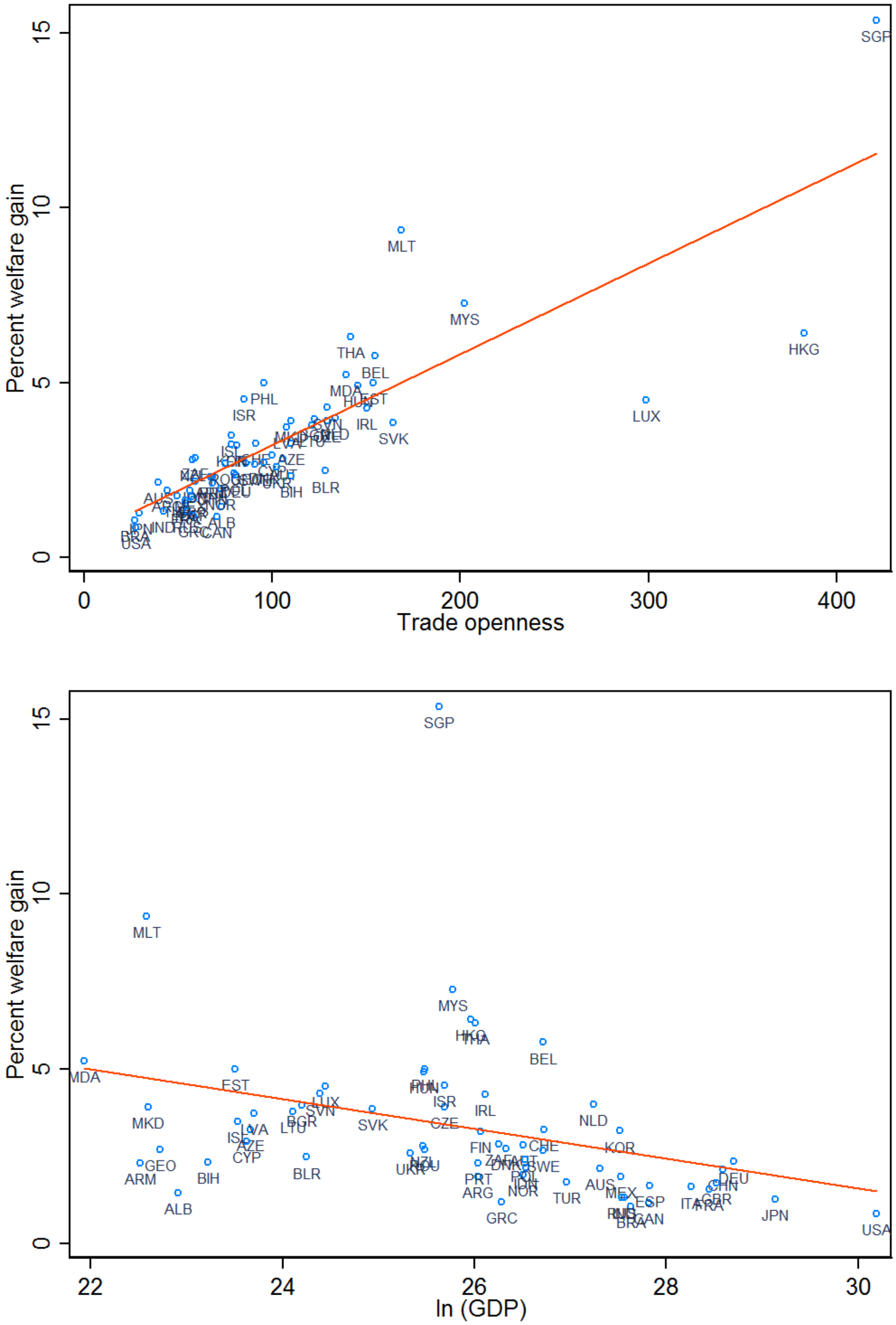
Note: The figures are based on regressions in Table 5. The dashed lines are the 95% confidence interval.

Figure 6
The role of consumer self-selection



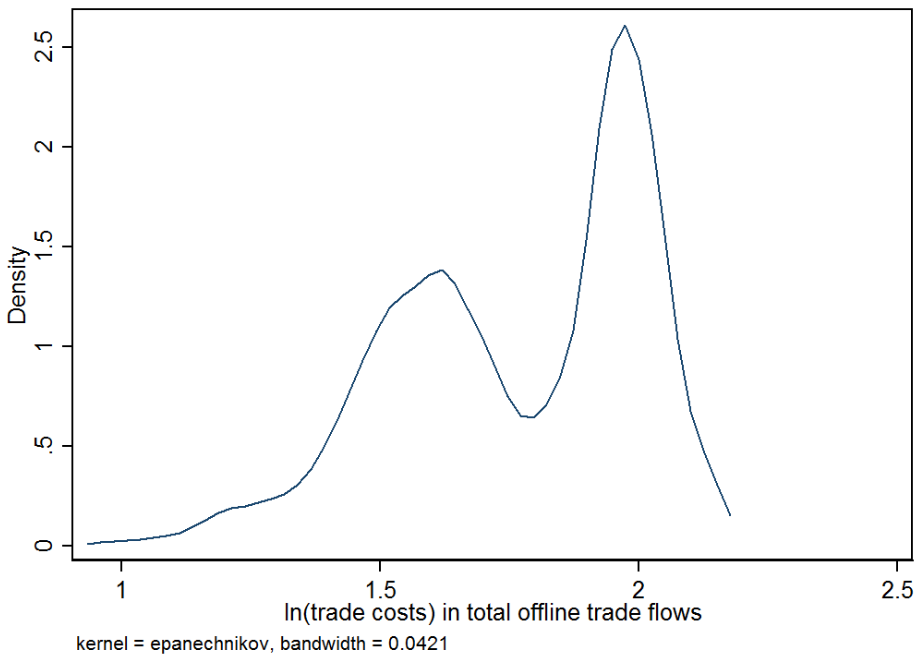
Note: The figures are based on regressions in Table 6. The dotted lines give the kernel density estimate of the x axis variable. The dashed lines are the 95% confidence interval.

Figure 7
Welfare gains from lower distance-related search costs



Note: Added-variable plots obtained after regressing welfare gains on GDP, GDP per capita, remoteness and openness.

Figure 8
Kernel density estimate of aggregate $\ln(t_{ij})$



Note: The $\ln(t_{ij})$ for aggregate trade flows was consistently estimates using equation (3) on aggregate trade data and the consensus trade elasticity parameter -4 (Eaton and Kortum 2012).

Table 1
Gravity, online and offline

	(1) eBay	(2) eBay	(3) eBay	(4) eBay	(5) offline	(6) offline	(7) offline	(8) offline
Distance	-0.508***	-0.351***	-0.338***	-0.369***	-1.408***	-1.134***	-1.129***	-0.663***
	(0.0337)	(0.0465)	(0.0463)	(0.107)	(0.0367)	(0.0501)	-0.0503	(0.0456)
No common legal sys.		-0.241***	-0.194***	-0.254**		-0.572***	-0.571***	-0.379***
		(0.0592)	(0.0583)	(0.103)		(0.0543)	-0.0546	(0.0591)
No colony		0.0624	0.0301	-0.285**		-0.462***	-0.492***	0.0300
		(0.142)	(0.143)	(0.128)		(0.166)	-0.166	(0.109)
No common language		-0.412***	-0.418***	-0.960***		-0.195*	-0.193*	0.218**
		(0.0946)	(0.0944)	(0.141)		(0.110)	-0.11	(0.0956)
No border		-0.137	-0.125	-0.750***		-0.318**	-0.287*	-0.285***
		(0.132)	(0.130)	(0.143)		(0.151)	-0.152	(0.0895)
No FTA		-0.193**	-0.225**	-0.295*		-0.318***	-0.311***	-0.430***
		(0.0880)	(0.0885)	(0.174)		(0.0905)	-0.0914	(0.0830)
Shipping costs			-0.000237				-0.00178*	
			(0.000892)				-0.000941	
Observations	3,660	3,660	3,636	3660	3,660	3,660	3636	3660
R-squared	0.895	0.896	0.898		0.882	0.889	0.888	

Note: All regressions are estimated using an importer and exporter fixed effect linear model, except for columns (4) and (8) which use a poisson pseudo maximum likelihood estimator. The figures in brackets are pair-clustered standard errors, and *, **, and *** stand for statistical significance at the 10, 5, and 1 percent level respectively.

Table 2

Testing differences in gravity coefficients

	Distance	No common legal system	No colony	No common language	No border	No FTA
Gravity coefficient	-1.134***	-0.572***	-0.462***	-0.195*	-0.318**	-0.318***
	(0.0501)	(0.0543)	(0.166)	(0.110)	(0.151)	(0.0905)
Interaction with eBay dummy	0.783***	0.332***	0.524**	-0.217	0.181	0.125
	(0.0662)	(0.0804)	(0.226)	(0.142)	(0.207)	(0.130)

Note: The dependant variable is log imports. Regression estimated using importer-eBay and exporter-eBay fixed effect linear model. 7,320 Observations. R-squared: 0.908. The figures in brackets are pair-clustered standard errors, and *, **, and *** stand for statistical significance at the 10, 5, and 1 percent level respectively.

Table 3

Robustness checks: Gravity for different types of eBay flows

	(1) eBay total	(2) offline total	(3) New goods	(4) Used goods	(5) Auctions	(6) Non-auctions	(7) PowerSellers	(8) Non-PowerSellers
Distance	-0.440***	-1.308***	-0.530***	-0.340***	-0.116***	-0.228***	-0.323***	-0.429***
	(0.0321)	(0.0511)	(0.0448)	(0.0427)	(0.0358)	(0.0497)	(0.0396)	(0.0426)
No common legal sys.	-0.158***	-0.571***	-0.208***	-0.261***	-0.104**	-0.346***	-0.158***	-0.237***
	(0.0394)	(0.0565)	(0.0536)	(0.0554)	(0.0433)	(0.0590)	(0.0510)	(0.0521)
No colony	-0.285***	-0.279	-0.175	0.0395	-0.0311	0.212	-0.359***	-0.207*
	(0.0934)	(0.173)	(0.150)	(0.133)	(0.136)	(0.157)	(0.128)	(0.116)
No common language	-0.387***	-0.326***	-0.308***	-0.451***	-0.259**	-0.556***	-0.305***	-0.313***
	(0.0663)	(0.111)	(0.0775)	(0.0900)	(0.113)	(0.125)	(0.0752)	(0.0885)
No border	-0.242***	-0.136	-0.162	-0.144	-0.447***	-0.388**	-0.296**	-0.347***
	(0.0915)	(0.155)	(0.126)	(0.122)	(0.129)	(0.155)	(0.117)	(0.109)
No FTA	-0.100*	-0.230**	0.0437	-0.180**	0.0378	-0.0912	-0.233***	-0.0429
	(0.0580)	(0.0955)	(0.0805)	(0.0812)	(0.0518)	(0.0761)	(0.0788)	(0.0787)
Observations	3,660	3,660	3,660	3,660	3,660	3,660	3,660	3,660
R-squared	0.943	0.856	0.900	0.900	0.630	0.810	0.911	0.903

hline

Note: All regressions are estimated using an importer and exporter fixed effect linear model. The figures in brackets are pair-clustered standard errors, and *, **, and *** stand for statistical significance at the 10, 5, and 1 percent level respectively.

Table 4

Testing differences in gravity coefficients for PowerSellers

	Distance	No common legal system	No colony	No common language	No border	No FTA
Gravity coefficient	-0.429***	-0.237***	-0.207*	-0.313***	-0.347***	-0.0429
	(0.0426)	(0.0521)	(0.116)	(0.0885)	(0.109)	(0.0787)
Interaction with PowerSeller dummy	0.106**	0.0792	-0.152	0.00772	0.0515	-0.191**
	(0.0458)	(0.0613)	(0.118)	(0.0839)	(0.120)	(0.0952)

Note: The dependant variable is log of eBay imports. Regression estimated using importer-PS and exporter-PS fixed effect linear model. Observations: 7,320. R-squared: 0.915. The figures in brackets are pair-clustered standard errors, and *, **, and *** stand for statistical significance at the 10, 5, and 1 percent level respectively.

Table 5

eBay reduces the distance elasticity most when search costs are highest

	eBay	image	eBay	image	eBay	image
Distance	-0.210***	-1.816***	-0.495***	-0.665***	-0.308***	-1.394***
	-0.0125	-0.0276	-0.0764	-0.0906	-0.0522	-0.0548
No common legal sys.	-0.315***	-0.788***	-0.259***	-0.513***	-0.261***	-0.445***
	-0.0161	-0.0359	-0.06	-0.0533	-0.059	-0.0543
No colony	-0.0304	-0.844***	0.0417	-0.395**	0.0536	-0.408**
	-0.042	-0.082	-0.142	-0.167	-0.14	-0.16
No common language	-0.723***	-0.197***	-1.726***	4.068***	-0.391***	-0.330***
	-0.0323	-0.0648	-0.6	-0.729	-0.0923	-0.106
No border	-0.556***	-0.708***	-0.0762	-0.515***	-0.149	-0.243*
	-0.0369	-0.0694	-0.135	-0.156	-0.129	-0.137
no FTA	0.00314	-0.229***	-0.201**	-0.292***	-0.186**	-0.360***
	-0.02	-0.0462	-0.0873	-0.0892	-0.0875	-0.0874
Distance $\times$ σ	-0.0314***	0.180***				
	-0.00118	-0.00208				
Distance $\times$ No common language			0.158**	-0.514***		
			-0.0699	-0.0824		
Distance $\times$ corruption					0.0539*	-0.332***
					-0.0281	-0.0265
Observations	76,410	76,410	3,660	3,660	3,660	3,660
R-squared	0.066	0.15	0.896	0.891	0.896	0.895

Note: All regressions are estimated using an importer and exporter fixed effect linear model. The figures in brackets are pair-clustered standard errors, and *, **, and *** stand for statistical significance at the 10, 5, and 1 percent level respectively.

Table 6
Selection effects (not for publication)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ebay	ebay	ebay	ebay	offline	offline	offline	offline
Distance	-0.380***	-0.295***	-0.713***	-0.292***	-1.099***	-1.441***	-0.890***	-1.303***
	(0.0841)	(0.0723)	(0.139)	(0.0577)	(0.0837)	(0.0676)	(0.127)	(0.0569)
No common legal sys.	-0.240***	-0.237***	-0.243***	-0.236***	-0.562***	-0.526***	-0.571***	-0.524***
	(0.0614)	(0.0589)	(0.0584)	(0.0594)	(0.0566)	(0.0546)	(0.0538)	(0.0553)
No colony	0.0668	0.0541	0.0661	0.0720	-0.430**	-0.439***	-0.465***	-0.437**
	(0.150)	(0.142)	(0.140)	(0.143)	(0.176)	(0.164)	(0.165)	(0.167)
No common language	-0.443***	-0.435***	-0.404***	-0.433***	-0.251**	-0.274**	-0.201*	-0.265**
	(0.103)	(0.0941)	(0.0926)	(0.0943)	(0.124)	(0.110)	(0.109)	(0.110)
No border	-0.0728	-0.119	-0.136	-0.161	-0.408**	-0.260*	-0.318**	-0.277*
	(0.136)	(0.130)	(0.130)	(0.132)	(0.157)	(0.145)	(0.149)	(0.149)
no FTA	-0.210**	-0.182**	-0.198**	-0.197**	-0.302***	-0.329***	-0.314***	-0.314***
	(0.0928)	(0.0879)	(0.0871)	(0.0876)	(0.0962)	(0.0899)	(0.0896)	(0.0902)
Distance × Education	0.000292				-0.000930			
	(0.00129)				(0.00122)			
Distance × Gini		-0.00168				0.00641***		
		(0.00109)				(0.00104)		
Distance × Internet			0.0107***				-0.00722**	
			(0.00387)				(0.00340)	
Distance × GNI per capita				-3.00e-06**				6.95e-06***
				(1.26e-06)				(1.20e-06)
Observations	3,360	3,600	3,660	3,540	3,360	3,600	3,660	3,540
R-squared	0.894	0.897	0.897	0.898	0.888	0.890	0.889	0.890

Note: All regressions are estimated using an importer and exporter fixed effect linear model. The figures in brackets are pair-clustered standard errors, and *, **, and *** stand for statistical significance at the 10, 5, and 1 percent level respectively.

Table 7
The welfare gains from lower search costs

country	welfare gain (%)	country	welfare gain (%)	country	welfare gain (%)
USA	0.86	BIH	2.33	SVK	3.84
BRA	1.06	DEU	2.34	CZE	3.90
CAN	1.16	POL	2.39	MKD	3.91
GRC	1.19	BLR	2.47	BGR	3.96
JPN	1.27	UKR	2.58	NLD	3.98
IND	1.31	SWE	2.67	IRL	4.27
RUS	1.31	ROU	2.69	SVN	4.28
ALB	1.45	GEO	2.70	LUX	4.50
FRA	1.55	DNK	2.72	ISR	4.53
ITA	1.61	NZL	2.80	HUN	4.90
ESP	1.65	AUT	2.83	PHL	4.98
GBR	1.73	ZAF	2.85	EST	4.99
TUR	1.75	CYP	2.92	MDA	5.23
MEX	1.91	FIN	3.20	BEL	5.75
ARG	1.92	KOR	3.22	THA	6.31
NOR	1.97	AZE	3.25	HKG	6.42
CHN	2.11	CHE	3.27	MYS	7.26
AUS	2.14	ISL	3.48	MLT	9.36
IDN	2.18	LVA	3.71	SGP	15.36
PRT	2.29	LTU	3.76	TWN	43.27
ARM	2.31				

Note: Welfare gains estimated using formula in Arkolakis et al. (2012).

Table 8

Matching eBay categories and HS codes

eBay category	HS codes
Antiques	9701 9702 9703 9706
Baby	6111 6209 9501 9502 9503 490300 871500 940490
Books, Comics and Magazines	4901 4902 970200
Business, Office and Industrial	85 8201 8304 8459 8460 8461 8462 8463 8465 8517 9018 481960 852020 940600
Auto - Parts	8803 8714 8708
Clothes, Shoes and Accessories	61 62 64 65 420221 420222 420229 630900
Coins	490700 7118
Collectables	9705
Computing	8471
Consumer Electronics - Other	853110 8507 8506
Dolls, Doll Houses	9502
Hobbies and Crafts	4203 4414 4817 340600 701610 731910 960330 960340
Home and Garden	57 3922 6303 8450 9103 9105 9405 230910 392610 481960 830300 850910 851610 851660 940310 940330 940340
Jewellery and Watches	7101 7102 7103 7113 7114 7116 7117 9101 9102
DVDs, Film and TV	3706 852490
Music	8524
Photography	9006
Pottery and Glass	6911 7013
Sporting Goods	9506
Sports Memorabilia	9705
Stamps	490700 970400
Toys and Games	9501 9502 9503 9505 9506
Musical Instruments	92
Mobile and Home Phones	8517
PC and Video Gaming	950410 950430
Consumer Electronics - Audio	8518 8519 8521 8528 9007
Art	9701 9702 9703
Health and Beauty	3301 3303 3304 3305 3306 330730 330790

Note: When aggregated into a basket, HS categories that fit into many eBay categories are added only once.