The Geography of Trade in Online Transactions: Evidence from eBay and MercadoLibre

Ali Hortaçsu
F. Asís Martínez-Jerez
Jason Douglas*

Abstract

We analyze geographic patterns of trade between individuals using transactions data from eBay and MercadoLibre, two large online auction sites. We find that distance continues to be an important deterrent to trade between geographically separated buyers and sellers, though to a lesser extent than has been observed in studies of non-Internet commerce between business counterparties. We also find a strong "home bias" towards trading with counterparties located in the same city. Further analyses suggest that location-specific goods, such as opera tickets; cultural factors; and the possibility of direct contract enforcement in case of breach may be the main reasons behind the same-city bias.

I.

An extensive literature in international economics analyzes the impact of distance on trade flows. Starting with Jan Tinbergen (1962), the stylized finding of a large number of papers estimating the "gravity equation" is that trade volume between two countries increases with the size of their economies and decreases with the distance that separates them. A subset of these papers also reports a significant "border effect": controlling for distance, trade between two regions is lower if the goods have to cross national borders (John McCallum, 1995, James E. Anderson and Eric van Wincoop, 2003). Moreover, when the home bias has been tested for U.S. intra-national trade flows, state limits seemed to have an effect on trade similar to that of the national borders (Holger C. Wolf, 2000, Russell Hillberry and David L. Hummels, 2003).

Anderson and van Wincoop (2004), in their recent review of this literature, point out transportation costs and tariffs/taxes as the main frictions contributing to both the decline of trade flow with distance and the border effect. They also discuss a growing number of papers on "informational frictions." Such informational frictions include search costs, which can impede geographically distant buyers and sellers from finding each other; communication barriers, which

hinder the efficiency of negotiations; and, more generally, "contracting costs," which are driven by the inability to monitor and discipline the misconduct of distant transacting parties. For instance, James E. Rauch and Vitor Trindade (2002) find that ethnic Chinese networks have increased bilateral trade flows between country pairs, and that in Southeast Asia the effect is larger for differentiated products than for homogenous goods. Pierre-Philippe Combes, Miren Lafourcade and Thierry Mayer (2005) report that firm and immigrant networks are important facilitators of intra-France trade. James E. Anderson and Douglas Marcouiller (2002) show that country-level indices of institutional quality are associated with trade flows. In an application to the trade of financial assets, Richard Portes and Helene Rey (2005) find a negative correlation between telephone traffic and bank presence and the distance effect on equity transactions. Finally, Hillberry and Hummels (2003, 2008) show that regional patterns of trade are determined by industry location and supply-chain optimization decisions. Agglomeration of industries in search of spillovers or national advantages (Glenn Ellison and Edward L. Glaeser, 1997) leads to a high volume of short-distance hauls of intermediate goods. Moreover, efficiency of hub-and-spoke distribution networks results in a high volume of intrastate shipments by wholesalers that are the recipients of intrastate trade. Pankaj Ghemawat (2001) argues that inattention to the non-physical dimensions of distance is at the root of many firms' international strategy failure.

The rise of the Internet naturally leads to the question of whether the institutional environment of online commerce alters the geography of trade flows. This paper analyzes geographic patterns of trade on two large online auction sites, eBay and MercadoLibre. eBay is the largest online auction site in the world, and our data is a representative sample of all eBay transactions (except eBay Motors) conducted within the 48 continental U.S. states. MercadoLibre is the largest online auction site in Latin America; we chose to study this site mainly to check the robustness of the results we obtained using eBay data, but also to understand whether additional geographic barriers to trade arise in the context of a less-developed set of economies.

Our setting is especially interesting because it allows us to observe commerce in its purest expression, as a transaction of end products between individual economic agents. Trading on online auction sites is largely independent of the geographic configuration of traditional distribution networks, whose impact on the geography of trade flows is emphasized by Hillberry

and Hummels (2003). Moreover, focusing on the trade of end products should isolate the patterns observed from the physical proximity chosen to optimize the supply chain in business-to-business commerce (Hillberry and Hummels, 2008). Finally, although our setting is unique, the phenomenon we analyze is neither rare nor irrelevant; more than 200 million eBay users worldwide listed over 600 million items on this site in the third quarter of 2006 alone.[1] For these reasons, our research can shed important light on our understanding of the geography of trade.

Furthermore, the online auction environment provides an exceptional opportunity to study the distance-dependence of trade, as the environment can be considered close to "frictionless" in certain important dimensions.[2] First, costs of computerized searches are practically non-existent and location-independent, as are costs of communication using email and a fairly uniform format and language. As for shipping costs, a flat shipping fee is quoted by most eBay sellers for transactions within the continental U.S., largely equalizing this margin across different seller locations. [3] Note also that, within the continental U.S., tariffs are nonexistent, and sales taxes, which are imposed by the states, should encourage out-of-state purchases as opposed to in-state purchases.

Our main result is the following: distance still has a negative effect on trade on the online auction sites eBay and MercadoLibre, though the effect is much smaller than has been observed in off-line trade. This effect is highly non-linear, with trading volume abnormally high within the same city. Once beyond the driving distance of the city limits, the effect of distance on trade is relatively small. As expected, the non-linearity of the distance effect is strongest for goods that have to be consumed in a specific location, such as opera tickets; however, it is evident in all categories of items. Further results suggest that "trust" may be a significant contributor to the distance effect: the "same-city" effect is much more pronounced in those categories where seller

[1] eBay, 8-K, January 24, 2007 (http://yahoo.brand.edgar-online.com/fetchFilingFrameset.aspx?dcn= 0000950134-07-001187&Type=HTML, accessed July 13, 2007).
[2] The "home bias" literature in finance can also be characterized as studying an environment that is similarly "frictionless." For example, Joshua D. Coval and Tobias Moskowitz (1999) find that mutual funds are likely to hold regionally-biased portfolios, and argues that monitoring costs may be an important factor.
[3] One may also consider the inconvenience caused by the time it takes to ship objects a long distance as an unobserved shipping cost. However, while one would expect shipping time to vary linearly with distance, as described later in the paper, we find a highly nonlinear pattern of distance-dependence that varies very little between 50 and 2000 kilometers.

reputation is lower. We also find some evidence that culture is a key factor in shaping the geography of trade: the same-city effect is strongest for local interest items such as sports memorabilia.

Our paper advances the literature on intra-national trade, helping to explain the factors behind an observed proximity bias that exists even after controlling for the most relevant causes previously identified—optimization of the supply chain, shipping costs, and search frictions. We show that even in the absence of search costs, information asymmetries, such as uncertainty regarding the reliability of a seller, may serve as an important barrier to trade[4]—and that proximity may serve as a substitute for trust.

Our paper also contributes to the literature on the impact of the Internet on the globalization of the economy, in which Caroline L. Freund and Diana Weinhold (2004), for instance, find that Internet connectivity is associated with increases in trade volume. We also complement the work of Manuel Blum and Avi Goldfarb (2006), who find that local tastes appear to be an important driver of digital-goods consumption. Some of our findings in physical-goods trade reinforce their conclusion regarding the importance of local tastes, though other findings point to factors such as trust as being another source of the observed home-bias on the Internet. Our results, along with Blum and Goldfarb's findings, may be interpreted as suggesting a potential limit to the Internet's ability to eliminate geographic barriers. Factors such as lack of trust and local tastes may still render geography an important factor in determining market boundaries and trade flows.

II. Theoretical Framework

To motivate the analysis we will use a simple auction model with an exogenously determined number of participants. The willingness of buyers to pay will depend on the characteristics of the good auctioned and the characteristics of the seller, including the buyer's geographic

[4] For example, Rauch and Trindade's (2002) ingenious study of how Chinese immigrant networks affect trade does not distinguish between a search cost story, in which trading partners cannot find each other, and an informational asymmetry story, in which trading partners do not trust each other.

proximity to the seller. Higher willingness to pay will lead to a higher probability of winning the auction. How distance among agents affects the willingness to pay will influence the probability of winning the auction, and therefore the number and aggregate value of transactions we should observe between agents from any pair of geographic locations.

Let's assume that in a certain auction there are L locations, and that at each location l, there are M_l sellers, indexed by $j=1,..,M_l$, and N_l buyers, indexed by $i=1,..,N_l$. The utility of buyer i in location b for a good sold by seller j in location s is given by

$$u_{ibjs} = \gamma + \mu_{b,s} + \varepsilon_{ibjs}$$

where γ is some fundamental value of the item auctioned, and $\mu_{b,s}$ is a factor that affects the utility of all buyers in location b for goods sold by sellers in location s, such as the distance and shipping costs between the two locations. $\mu_{b,s}$ may also be affected by the nature of the good being auctioned or the reputation of the seller, as the cost of a recourse action increases with distance. ε_{ibjs} is an IID random disturbance that is idiosyncratic to buyer i (in location b) and seller j (in location s).

If we assume that the auction mechanism is efficient—i.e., that it awards each good to the buyer with highest willingness to pay—and that ε_{ibjs} is IID across buyers and locations and follows a Type-I extreme value distribution, we can express the probability that a buyer from location b wins an auction in which the good is sold by seller j at location s as

$$(1) \quad \Pr\{\text{buyer from } b \text{ wins auction of seller } j \text{ at location } s\} = \frac{N_b \exp(\gamma + \mu_{b,s})}{\sum_{b'=1}^{B} N_{b'} \exp(\gamma + \mu_{b',s})}$$

following multinomial logit choice probabilities (the N_b terms reflect the population weighting of buyers across locations). The more positive the effect of the distance $\mu_{b,s}$, the more likely the largest valuation will be drawn by a buyer of type b. By the same token, the larger the number of buyers of type b, N_b, the more likely the highest valuation will occur in a buyer of this type.

Observe that if the geographic distance has no impact on a buyer's valuation, the probability of winning the auction depends exclusively on the number of buyers of each type, and the item will likely be sold to economies with a larger number of buyers.

If we take logs in equation (1) we obtain the following expression, which is linear in the effect of distance $\mu_{b,s}$ and the log of the number of buyers in location *b,* N_b:

$$(2) \qquad \log \Pr\{b,s\} = -\log c_s + \log N_b + \gamma + \mu_{b,s}$$

where

$$c_s = \sum_{b'=1}^{B} N_{b'} \exp(\gamma + \mu_{b',s})$$

Multiplying equation (1) by M_s we obtain the expected number of sales, $T_{b,s}$, by sellers in location *s* to buyers in location *b*. Taking logs and including a disturbance term, $v_{b,s}$, we obtain the following expression of the Gravity equation:

$$(3) \qquad \log T_{b,s} = k_s + \log M_s + \log N_b + \mu_{b,s} + \nu_{b,s}$$

where $T_{b,s}$ is the total number of sales to buyers in location *b* by sellers in location *s*; M_s is the total number of sellers in location *s*; and k_s is a constant term capturing the effect of the fundamental value of the good, γ, and c_s as defined above.

The gravity equation suggests the following testable hypotheses: (1) The total number of sales (we will repeat the analysis with the total dollar value of the sales) to buyers in location *b* by sellers in location *s*, $T_{b,s}$ is proportional to the size of the economy of the buyers, $\log N_b$, and of the sellers, $\log M_s$; (2) When all transactions are pooled in the analysis, the effect of distance on the intensity of trade, $\mu_{b,s}$, should be such that an increase in the distance between players should reduce the number of transactions; and (3) The impact of distance on the amount of trade will depend on the value of the item and the reputation of the seller.

In the following section we describe the data used in the empirical analysis.

III. Data Sources

We developed this study with data from two leading online auction sites: eBay, the firm that popularized and, arguably, invented the concept of online auctions, and the largest player in this industry; and MercadoLibre, the leader of online auctions in Latin America. Online auction sites are well-suited for our study not only because these firms are interested in minimizing the impact of distance to increase the size of their networks, but also because they are a good proxy for consumer-to-consumer Internet commerce. According to the Forrester Technographics survey, in 2004 close to 30 percent of U.S. households had bid in an Internet auction, and in the third quarter of 2006 eBay represented more than one fourth of U.S. Internet retail commerce.[5] Therefore, our results will be indicative of how geography and the Internet may affect commerce in its purest state, as end product transactions among individuals.

eBay was founded by Pierre Omidyar in 1995 in San Jose, California. Since then it has grown continuously to become the largest online auction site in the world. In 2004, more than 1.4 billion items were listed in eBay's marketplace, resulting in $34.2 billion worth of merchandise transactions.[6] MercadoLibre, founded by Marcos Galperín in 1999 in Buenos Aires, is currently the leading Latin American online auction site.[7] It operates in twelve Latin American countries[8] and in 2004 enabled the sale of more than 9.5 million items for an aggregate value of $425 million.

---

[5] In its 8-K of January 24, 2007, eBay reported that $14.4 billion worth of goods were traded in its marketplace during the third quarter of 2006, 51% of which were traded in the United States. (http://yahoo.brand.edgar-online.com/fetchFilingFrameset.aspx?dcn=0000950134-07-001187&Type= HTML, accessed July 13, 2007). According to the U.S. Census Bureau, e-commerce sales in the United States for the same period amounted to $27.5 billion (http://www.census.gov/mrts/www/data/html/ 06Q3.html, accessed July 13, 2007).

[6] eBay Annual Report, 2004.

[7] http://www.mercadolibre.com.ar/argentina/ml/p_loadhtml?as_menu=MPRESS&as_html_code=SML_05, accessed September 11, 2005.

[8] Argentina, Brazil, Chile, Colombia, Costa Rica, the Dominican Republic, Ecuador, Mexico, Panama, Peru, Uruguay, and Venezuela. At the time of our study the Costa Rica, Dominican Republic, Panama, and Peru sites were not yet operational.

The eBay data is the result of a stratified sampling of eBay listings we collected from the company site between February and May 2004. From each of the 27 main categories of items on eBay—excluding autos and real estate—we extracted a daily random sample. For each sampled listing, we obtained the description of the item being sold, the seller's location, the shipping and handling fee posted by the seller, and other listing characteristics that might affect demand (such as the seller's feedback rating, the insurance and payment methods allowed, listing time, etc.). Unfortunately, obtaining the buyer's location was less straightforward, since eBay does not report the location of the buyers explicitly. However, the buyer's location can be obtained if the buyer has previously sold an item on eBay, and if that item's listing is still recorded in the eBay database.[9] This allowed us to obtain the location of the buyer in 27 percent (or 266,588) of these transactions. This missing data problem skews our sample towards buyers who are more "experienced" traders on eBay, as they have to participate in trades as both buyers and sellers within a short period of time. We see this characteristic as a strength rather than a weakness, as it makes our sample more likely to reflect the conscious behavior of market participants who understand the impact of different transaction elements (as opposed to the potentially noisy decisions of occasional buyers who are not well versed in the workings of the community).

We collected the ex-ante shipping and handling fee declared by the seller. Whenever an auction did not have a cost of shipping associated with it, we deleted it from our sample. Thus, our sample does not include any transactions that were available only for local pickup. It also excludes all transactions for which the shipping cost was difficult to discern, i.e. when it was not explicitly disclosed by the seller.[10] We believe it is reasonable to assume that the items included in our analysis have flat shipping costs within the continental United States; although there is anecdotal evidence that shipping costs are sometimes negotiated ex-post between the seller and the winner of the auction, we do not have a way to measure the frequency of such ex-post negotiations or how they may affect the final cost.

[9] In our sample we identify any buyer who had listed an item in the 90 days prior to the day of the transaction or at any time after the transaction was consummated and before June 30, 2004, when we stopped collecting buyer location information.

[10] Listings that describe shipping costs as "Not specified" and those that instruct buyers to "Contact seller for S&H" were excluded from our sample.

The objective of the paper is to understand the impact of distance of internet trade. Some goods transacted in eBay can only be consumed in a particular city (e.g. tickets for a cultural event), and other goods are affinity goods, such as memorabilia from sports teams, which are sought mainly by people residing in certain locations. Including these goods in our sample could call into question whether our findings were driven by them or whether they reflected more general traits of Internet trading behaviors. Thus—with the exception of the section on the city-level effect—we focus our paper only on those goods that we can characterize as non-local. To do this, we classified all sub-categories of items (a total of some 23,000) into three groups: (1) not local, (2) somewhat local, and (3) definitely local, and included in our sample only those in group 1.[11] Table 1 presents some summary statistics on the final number of observations in our sample (123,333).[12]

MercadoLibre gave us comprehensive statistics on the geographic patterns of trade for its different websites. Thus, for any pair of buyer and seller locations (states/provinces), we have access to the number and amount of all the monthly transactions completed during the period from August 2003 to July 2004. This information was made available to us for each of the 30 main categories of items in MercadoLibre (they are similar but not exactly parallel to eBay's categories). For each pair of locations, there is one observation for auction transactions and one for fixed-price transactions.[13] Having all transactions in the database eliminates the problem of the missing buyer-location data and any measurement error associated with the sampling procedure used with the eBay data; however, it comes at the cost of not being able to obtain all the listing characteristics that might affect demand.

[11] For instance, single-disc DVD players were classified as not local, because they are of interest to potential buyers throughout the United States. On the other hand, college-related collectibles were classified as somewhat local, because even though colleges can recruit their students from all over the country, student bodies and alumni populations tend to be geographically concentrated. Finally, Idaho collectibles were classified as definitely local because of the high likelihood of co-location of buyers and sellers for items related to this state.

[12] To test the robustness of these results, we also ran the analyses in the next two sections with all types of items—local, somewhat local, and not-local—as well as with items for which the seller did not specify a shipping cost. The results were essentially the same. For the items without a specified shipping cost we input the average shipping cost of an item in its category.

[13] 86% of MercadoLibre's traffic is fixed price (http://www.mercadolibre.com.ar/argentina/ml/p_loadhtml?as_menu=MPRESS&as_html_code=SML_05, accessed September 12, 2005).

There are several reasons why these data sources are especially valuable for economists who try to understand the geography of trade. First of all, they depict the purest possible form of commerce: end product transactions between individual economic agents whose geographic patterns are dominated neither by the concentration of industries in search of spillovers or natural advantages, nor by the endowment of big suppliers of unique, branded goods. Second, these marketplaces are pure intermediaries that facilitate trade in a variety of goods by heterogeneous agents. Thus, short of running a comprehensive census, focusing on these sites allows us to examine geographic trading patterns for a relatively large cross-section of product and agent varieties, and helps us to understand which factors make trading more sensitive to distance. Third, one may argue that the main benefit of the Internet as a trade facilitator is to reduce search costs, and it is reasonable to think of these marketplaces as being essentially "frictionless" in this regard. Fourth, shipping and handling fees are often quoted explicitly by the seller on these websites, and, for a wide class of goods—the ones considered in this study—these fees apply uniformly to buyers of differing locations (at least within the same country). Thus, we can effectively control for shipping cost differentials across locations (with the exception of variations in the time-of-arrival dimension).

These features render eBay and MercadoLibre close approximations of the "unified marketplace" view of the Internet, though several caveats are in order. First, the products that are bought and sold through these sites, although encompassing a large variety, are mainly new and used household durables, and thus extrapolations to other categories of goods is not possible. Second, a similar "representativeness" criticism may be leveled against the demographic characteristics of the users of these websites or the Internet in general, qualifying any extrapolations to the off-line world.

IV. Results from Analyses Aggregated at the State/Province Level

In this section we analyze whether physical distance between buyer and seller reduces the intensity of Internet trade. If the online auction sites are able to eliminate the frictions that have been traditionally attributed to the distance effect, we should observe no difference in buyers'

purchasing behavior when the seller is close or far away. Specifically, in a regression framework based on equation (3), the variables that proxy for distance between buyer and seller should have no explanatory power when the dependent variables are measures of Internet trade.

Table 2 presents our first test of the gravity equation with the eBay sample. We include as benchmarks the results of Wolf (2000) and Hillberry and Hummels (2003), who tested the impact of distance on interstate commerce in the United States using data from the Commodity Flow Survey of the U.S. Census. Both of these studies find a negative and significant effect of the distance variable. These studies also find a very significant "home-state bias" effect, as seen in the large coefficient estimate on a dummy for same-state transactions, which suggests there are costs to trading across state borders that are not purely distance-dependent.[14]

In the regression with eBay data, we obtain results that have a sign consistent with prior studies. When we compare our results to the models of reference, we observe that the effect of distance in Internet trade is much smaller than in the Census data. The coefficients of distance in the eBay regressions are roughly one-tenth to one-twentieth the magnitude of the coefficients in Wolf (2000) and Hillberry and Hummels (2003), implying that trade falls 10 percent each time the distance doubles. However, the coefficient on same state is similar to what was found by Wolf (2000) and Hillberry and Hummels (2003), implying that intrastate commerce is 1.8 to 3 times higher than the amount that would be justified by other factors. The combination of these two results seems to indicate that although eBay is fairly effective in mitigating the effect of distance on interstate commerce, a "home bias" persists.[15]

[14] Specifically, Hillberry and Hummels (2003) show that the home-state bias found by Wolf (2000) diminishes once shipments from wholesalers are excluded from the CFS data set (wholesalers tend to ship in-state more than manufacturers). However, the same-state coefficient continues to be economically and statistically significant.

[15] We realize that the economic reality of the markets we analyze is very different from that used by Wolf and Hillberry and Hummels. Both of the latter studies use the Commodity Flow Survey of the U.S. Census, which covers a representative sample of shipments from U.S. mining, manufacturing, and wholesale establishments. By contrast, we use items traded in eBay, which are usually consumer goods. Thus, not only are the goods different in nature, but so are the parties who trade. In Wolf and in Hillberry and Hummels, the buyers and sellers are professionals acting on behalf of a corporation; in our paper, they are individuals acting in their own interests. As we focus on the transactions of end products between individuals, our study is especially relevant to understanding consumer preferences about trading with distant counterparts. Our results cannot be attributed to buyer and seller location choices based on the optimization of a supply chain.

The analyses using the MercadoLibre data support these findings. Table 3 shows the results of these analyses. In the first model, distance is measured by the distance between country capitals; in the second, by the distance between provincial capitals. We observe a negative distance effect that is somewhat higher than that of the eBay sample. This effect is attenuated by a very strong same-country effect and a relatively strong same-province effect. It is interesting to note from this analysis that the same-country effect seems to be much stronger (by six times) than the same-province effect. This difference may simply be caused by customs barriers, but it may also be due to an amplification of the frictions manifested in the "same-province" effect.

The "home-state bias" documented in the literature has been questioned on the basis of the omission of relative price indices (Anderson and van Wincoop, 2003) and mismeasurement of the intrastate distance (Keith Head and Thierry Mayer, 2002). Following Hillberry and Hummels (2003) we use the seller- and buyer-region fixed effect to address the omitted price-index critique. To reduce the measurement error bias, we use Wolf's (2000) measure of intrastate distance.[16] Moreover, the small magnitude of the distance effect suggests that our estimates of the same-state effect are not likely to be driven by mismeasurement of the intrastate distance (Head and Mayer, 2002). Finally, the latter part of the paper, which focuses on city pair distances, is not subject to this criticism.

We analyzed the robustness of our results in several ways. First, we considered whether the distance and same-state coefficients were actually the result of a relative density of buyers and sellers in certain geographic locations. Essentially, this raised the question of whether we would observe the same coefficients if buyers and sellers in our sample were matched randomly. To test this possibility, we randomly matched the same number of buyers and sellers as we included in our sample and ran the regression in model II of Table 2, repeating this procedure 1000 times. We found the same-state coefficient to be statistically significant in 56 instances, the number we would expect to find by chance alone. In addition, the number of same-state transactions we

[16] We measure distance as the great-circle distance between state capitals. For intrastate distances we use Wolf's (2000) formula, which utilizes the (population-weighted) distance between the two most populous cities within a state: $D_{ii,t} = 2 * \left[1 - \frac{P_{i,1}}{P_{i,1} + P_{i,2}}\right] * D_{i,12}$ where $P_{i,1}$ and $P_{i,2}$ denote the population of the largest and the second-largest city within the state respectively, and $D_{i,12}$ their distance from each other.

found ranged from 10,800 to 11,400, compared to 20,770 in the actual sample. In addition, the average coefficient on distance was -0.00083, which was statistically significant on 32 occasions and negative in 551 of the regressions. All these factors suggest that our results are due not to random pairing but to conscious choices on the part of buyers and sellers.

We also tested these results against the possibility that they were driven by the idiosyncratic characteristics of certain segments of our sample. Specifically, we included category fixed-effects in the regression models and ran the analyses using the data exclusively from one category in order to verify that the results were not driven by a certain type of item. We also classified the auctions in our sample as a function of the buyer's experience with eBay transactions and as function of the seller's reputation and repeated our analyses with each of these segments. Table 4 presents the results of running Model IV—the one we use as a base for the analyses in the next section—for the different sub-samples. The overarching conclusion is that the distance and the same-state effect are significantly present in all the segments of our sample.[17]

Finally, we inquired as to whether the absence of trade between certain pairs of states (in one direction or both) biased our estimations. First, we filled in the missing observations by defining the dependent variable as the logarithm of trade—in the number or value of transactions—plus one. Second, we ran our models using the Poisson specification described in João M.C. Santos Silva and Silvana Tenreyro (2006). In both cases our findings were essentially equivalent to the ones described here.

In summary, we find that the distance effect is present in both our samples. It seems that the Internet reduces but is unable to completely eliminate the frictions that cause the impact of distance. The powerful constraint posed by state borders raises the questions of a) what causes this force and b) whether state lines are the critical distance point at which a discontinuity occurs. We will address these questions in the sections that follow.

[17] We ran all the analyses in Table 2 and in Tables 5 and 6 for each of the sub-samples used in Table 4. The results were essentially the same.

V. What Drives the "Same-State" Effect on eBay?

In this section we analyze whether the distance effect observed in the prior section is explained by the same frictions that affect non-Internet trade, such as shipping costs, time zones, trust, or sales taxes. We will give special consideration to the possible reasons for the observed discontinuity at the state border.

The most evident trade friction is the cost of shipping, which is likely to increase with distance. Also, it is possible for transportation companies to have a two-tiered pricing structure for interstate and intrastate transport, a disparity which would generate the same-state effect. By focusing on transactions with flat shipping and handling rates, we potentially eliminate this friction in our analysis. However, the cost of shipping may still impact the decision to trade with a distant seller if the buyer considers the cost of a potential return. If we assume that the rates quoted by seller are proportional to the cost of shipping the item, the possibility of an in-person item return will be more valuable the higher the cost quoted in the listing. In models II, III, and IV of Table 5 we test this assumption by explicitly including the shipping cost in the regression.[18] Shipping cost has the predicted negative, and statistically significant, impact on trade activity; however, the same state coefficient remains unchanged, suggesting that shipping cost is not an explanation for the same-state effect. It does not seem to be explain the distance effect either, as the coefficient of the distance variable remains negative, significant, and at about the same level.

Another possibility is that the idiosyncratic culture of the Internet is responsible for geographic patterns of trade. For instance, previous research[19] has shown that online bidders commonly wait to place their bids until just before the auction expires, a strategy known as sniping. This strategy may be somewhat more difficult to implement for a specific auction if the buyer and seller are not in the same time zone (especially if the auction ends late at night or during the buyer's work hours). To test this possibility, we include a dummy variable in the model indicating whether the buyer and seller are in the same time zone; we find that the coefficient is not statistically

[18] Shipping costs as a percent of the item's final price are calculated for individual transactions from the data reported in the listing and the median is calculated for each state pair.
[19] See Alvin E. Roth and Axel Ockenfels (2002), Patrick Bajari and Ali Hortaçsu (2004).

significant and has virtually no impact on the other variables of the model (Table 5, Model III). These results lead us to discard time-zone difference as the cause of the physical distance effect.

The critique that overstatement of intrastate distances may artificially cause the "home bias" effect[20] could have greater basis in larger states (provided state size correlates with the measure of distance). This effect is potentially magnified by Wolf's (2000) finding that the share of shipments within a state to its total shipments is higher for larger states. To test whether the same-state effect we observed was caused by large amounts of trade within the borders of the largest states, we include in Table 5, Model IV individual dummies for the intrastate trade of five large states. The coefficient on the same-state dummy remains positive, significant, and approximately at the same level as before, while the same-state coefficient of the large states is negative and significant, except for Montana. Moreover, in the case of California, the sum of the same-state and the same-state-California coefficients is basically zero, perhaps because the influence of Silicon Valley makes Californians more comfortable with Internet commerce, or perhaps because the rest of the country is especially interested in items coming from California, a state known for setting trends in many cultural (and fashion) dimensions.

One of the risks of trading with a distant party is the difficulty of exerting any recourse if the other party does not fulfill his or her obligation. In the same way that the letter of credit was designed to address this issue in international trade, eBay has developed several features that aim to increase user trust in the online commerce platform. These include PayPal, the electronic payment system; buyer protection; and a feedback system. Several papers in the empirical industrial organization literature show that bidders respond to certain levels of negative feedback with a reduced willingness to pay.[21] If reputation is effective in mitigating the concerns of buyers for whom distance makes it difficult to exercise any recourse, we should observe that negative feedback generates a higher reduction in the willingness to pay of distant buyers. In Table 6, we interact distance and the same-state dummy with dummy variables that indicate

[20] Assuming a linear effect of distance on trade, an imputed intrastate distance larger than actual distance will yield smaller predicted intrastate commerce. The difference between the actual commerce and the imputed volume will be picked up by the same-state dummy (Hillberry and Hummels, 2003, and Wolf, 2000).
[21] See Bajari and Hortaçsu (2004) for a survey of these results.

whether the median seller's feedback rating for that state pair is below certain thresholds.[22] The significant coefficients in the interactions suggest that trust contributes to the effect of distance and that the feedback system helps to mitigate, but does not completely eliminate, this effect. Furthermore, the impact of negative feedback is less visible within the same state (the positive coefficient on the interaction term suggests that sellers with bad reputations are more likely to find a buyer in the same state), which is consistent with an interpretation of the same-state effect that attributes the higher intensity of intrastate commerce to an increased possibility of direct enforcement of the trade agreement, either by returning the good in person or by compelling the seller to deliver on his or her promise.

Finally, taxes are often associated with geographic patterns of trade. Austan Goolsbee (2000) shows that Internet purchases are partially driven by sales-tax optimization. In general, sales taxes are only collected if buyer and seller are located in the same state.[23] Thus, differences in the tax regime of the seller state should have no impact on trade except in the sense that when seller and buyer are in the same state they will be less likely to engage in trade, and even less likely if they are located in a state with a high sales tax. This would normally suggest a negative coefficient on the same-state dummy, contrary to the evidence above. However, if we were to have individual same-state dummies for each state, it would be reasonable to expect a positive coefficient on the low-sales-tax-states' dummies. One may argue that the value of the coefficient may be affected by the undue influence of low or no sales-tax states, though it is difficult to conceive how this could result in a positive coefficient on the same-state dummy. Models II and III of Table 6 provide no evidence of sales tax causing the distance effect. As expected, none of the interactions between the dummies identifying the sellers' tax regimes and distance are significant. However, when interacted with the same-state dummy, coefficients—although not consistently significant—increase as the sales-tax rate falls, and the highest coefficient corresponds to the states without sales tax.

[22] The variable BAD SELLERS indicates whether the median seller rating is between 98.2-99.3%, and the variable VERY BAD SELLERS indicates whether the median seller rating is below 98.2%. In our transaction data, 75% of sellers have better than 99.3% positive rating and 90% of sellers have better than 98.2% positive rating.

[23] The seller will not be responsible for collecting sales taxes if it has no physical presence, or nexus, in the state of the buyer. In these cases, the buyer is obligated to report and pay the use tax—which is basically equal to the sales tax—in his or her state of residence. However, given the administrative complexity and widespread ignorance of this obligation, a vast majority of interstate Internet buyers do not pay sales tax.

Another possible motivation for buyers to prefer sellers in their own vicinity is the desire for immediate gratification—to enjoy the purchase as soon as the transaction is completed. If this impatience were driving the geographic patterns of trade, we would observe a significantly stronger distance effect on the buy-it-now transactions than on the regular auctions. We would expect that buyers who were more sensitive to the delay in receiving the good would be less likely to engage in auctions that required waiting until the closing of the bidding process, and would choose either the buy-it-now option or other channels of trade.[24] To investigate the merit of this hypothesis, we re-estimated our basic gravity equation specification by interacting the same-state and (log) distance variables with the percent of the transactions volume that was achieved through buy-it-now sales between each state pair. We found that neither interaction variable affected the same-state or distance coefficients in an economically or statistically significant manner.[25]

In summary, although shipping cost seems to deter Internet commerce between distant buyers and sellers, it does not explain the effect of distance observed in online trading. Moreover, trust seems to be the only variable that has some reliable impact on the same-state effect. This finding is consistent with the possibility of direct recourse in case of breach of contract, which provides a strong incentive for agents to keep trading relationships within a limited radius. The question then arises of whether state limits are the relevant distance for intense commerce or whether shorter radii, such as city limits or driving distance, are more important. We address this concern in the following section.

[24] The main driver of a buyer's choice of the buy-it-now method is probably price certainty. We would expect impatience to have a second-order effect on this choice, if any. If the impatience effect exists, we should observe it in the cross-section of the buy-it-now transactions.

[25] The buy-it-now results also indicate that our findings are not caused by "shill-bidding," a strategy employed by eBay sellers to raise their selling price by using a different screen name (alias) to bid on their own items. This was an unlikely explanation because shill-bidding could only be responsible for a proximity bias if (1) shilling sellers won an extraordinary number of auctions, which would come at great expense to them; (2) the alternative aliases sellers used for shilling were registered in the same state, a choice that might reveal their plot; and (3) in the case of the eBay data, sellers also used their shilling alias for selling, which would make it more difficult to build a reputation.

VI. Results from Analyses Aggregated at the City Level

The existence of a strong positive "same-state" or "same-province" effect on the intensity of Internet commerce suggests the existence of a sort of "trading gravity field" within which the intensity of transactional activity is much greater than without. Whether state borders or other milestones determine the reach of the attraction field is an empirical question. In principle, we may hypothesize that the city limits, the county line, or a specific travel length are justifiable alternatives to the state border. To further explore this question we exploit the fact that our eBay dataset contains the location of all buyers and sellers in the sample.[26] For the analyses in this section we also use sub-categories or items classified as local or somewhat local as we aim to identify the characteristics that determine the locality of a good.

To identify the point of discontinuity in the distance effect more precisely, we test the gravity equation aggregating the data at the city level and decomposing the distance variable into a series of dummy variables that take the value of 1 if the distance between buyer and seller is within a certain interval. We graph the coefficients of the distance dummy variables in Figure 1. Remarkably, the coefficients for all the distance intervals have similar levels and decrease smoothly as the distance increases, whereas the same-city coefficient is more than six times the other coefficients. This result suggests that, contrary to what one would expect, "driving distance" is relevant for Internet commerce. In theory, the Internet would enable markets to extend their reach almost limitlessly, and even if that benefit is partially observed for all other levels of distance, the city limits seem to represent an important barrier to trade.

Hillberry and Hummels (2008) find a similar result for the flow of commodities in the United States. The cause behind their finding, the co-location of participants in the supply chain to exploit natural advantages and spillovers (Ellison and Glaeser, 1997), does not seem plausible in our setting. In the previous section we found that the same-state effect was at least partially caused by direct enforcement ability and mitigated by reputation mechanisms. To explore the causes of the same-city effect, we rerun the regression on distance dummies for each of the 27

[26] Unfortunately, the MercadoLibre dataset was provided to us aggregated at the province level.

main item categories in eBay. Table 7 presents a ranked list of the coefficients of the same-city dummy for the different categories. An inspection of this list suggests several hypotheses. First, we observe at the top of the list tickets, which need to be used in a specific location. Next to tickets we find sports memorabilia, items that are likely to be owned and sought after by fans residing in the location of a particular team. For instance, a person in Sacramento is less likely to buy a Seattle Mariners baseball card than a person in Seattle. Thus, it seems that cultural factors—of which sports fandom is an example—have an important role in causing the same-city effect. However, we see that the items on which the city effect is smallest also seem to have a strong cultural component: entertainment memorabilia, art, collectibles, books, and dolls and bears. In contrast with baseball cards, these items are probably of interest to consumers nationwide, and their uniqueness makes buyers more likely to expand their search geographically in order to buy them.

We then regress the category-specific coefficients of the same-city dummies on the percentage of negative feedback in the average seller record and on the average price, average shipping cost, and average weight of an item in that category. Table 8 presents the results of the regression. Despite the potential attenuation bias caused by an estimated dependent variable, we observe a positive and significant effect of the reputation measure. The other variables do not enter significantly into this regression. One interpretation of these results is that as the likelihood of a breach of contract increases (the seller has a more negative reputation), it is more important to have the possibility of a direct enforcement mechanism, measured by proximity in the same-city effect. In contrast, the coefficients on shipping costs and weight are insignificant, suggesting that freight is not the main determinant of the same-city effect.

In summary, the results in this section support the hypothesis that despite the ease of searching that the Internet provides, the city limits or "driving distance" cause an "attraction field" that results in an excessive concentration of trade within them. In addition to goods that need to be locally consumed, cultural factors and the possibility of a direct enforcement action in case of breach of contract may determine the existence/need of a "local" market, even on the Internet.

VII. Conclusion

In this paper we analyze whether online auction sites have been able to create a virtual market in which the physical distance between buyer and seller becomes irrelevant. Using transactions data from two Internet auction platforms (eBay and MercadoLibre), we find that even though geographic distance is less of a deterrent to trade than it has been observed to be in studies of non-Internet commerce, more distant buyers are still less likely to engage in a purchase agreement than closer ones. Furthermore, there is an abnormally large concentration of commerce among buyers and sellers within the same city limits.

As expected, the non-linearity of the distance effect is strongest for goods that have to be used in a specific location, such as opera tickets; however, it is evident in all categories of items. Further analyses suggest that cultural factors and the possibility of direct contract enforcement in case of breach are the main reasons behind the distance effect. The higher the likelihood of a breach (suggested by poor seller reputation), the less likely a transaction between distant agents will take place. Also, items of local interest such as baseball cards tend to be traded in local markets. Shipping costs, at least for intrastate U.S. trade, lightly deter distant trade, but their influence can not explain the bulk of the proximity effect.

Given our focus on end-product transactions between individuals, this study is especially relevant to understanding consumer preferences about trading with distant counterparts. Our results cannot be attributed to buyer and seller location choices based on the optimization of a supply chain. In this sense, for firms designing strategies for geographic expansion, our study stresses the importance of paying careful attention to the non-physical dimensions of distance.

Our findings have implications for online commerce platforms that want to extend their reach. Our results suggest that they should continually innovate and perfect systems to increase the trust of market participants. Features such as continuous monitoring of listings, feedback systems, and buyer-protection programs are of greater benefit to more distant agents who, in principle, have fewer options to ensure fair trade than closer ones.

Future research could complement our findings by focusing on the impact of distance on the prices at which items trade. Of particular interest would be the analysis of sellers who expand their geographic reach by listing the same item in different sites and/or list their items in different languages. Another revealing study would analyze how distant and close buyers differ in their bidding behavior throughout an auction.

*Corresponding authors: Ali Hortaçsu, Department of Economics, University of Chicago, and NBER, 1126 E. 59th St. Chicago, IL 60637, Tel: (773) 702 5841, E-mail: hortacsu@uchicago.edu; F. Asís Martínez-Jerez, Accounting & Management Unit, Harvard Business School, Soldiers Field, Boston, MA 02163-9986, Tel: (617) 495-6844, E-mail: amartinezjerez@hbs.edu. We thank seminar participants at the FTC Internet Auctions Roundtable, Harvard Business School, INFORMS Annual Meeting, International Industrial Organization Conference, Northwestern University, Real Colegio Complutense, Universidad de Alicante, University of California at Berkeley, University of Chicago, and at the University of Southern California for valuable comments. We also thank Hays Golden, Gabriel Rosenhouse, and Katherine Miller for excellent research assistance. Financial support from the NET Institute (www.netinst.org) is gratefully acknowledged. Hortaçsu additionally acknowledges financial support from the John M. Olin Foundation and the NSF (SES-0449625). Martínez-Jerez acknowledges the financial support of the Harvard Business School Division of Research.

References

1. Anderson, James E. and Douglas Marcoullier. 2002. "Insecurity and the Pattern of Trade: An Empirical Investigation." *Review of Economics and Statistics*, 84(2): 342–52.

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

3. Anderson, James E. and Eric van Wincoop. 2004. "Trade Costs." *Journal of Economic Literature,* 42(3): 691–751.

4. Bajari, Patrick and Ali Hortaçsu. 2004. "Economic Insights from Internet Auctions." *Journal of Economic Literature*, 42(2): 457–486.

5. Blum, Manuel and Avi Goldfarb. 2006. "Does the Internet Defy the Law of Gravity?" *Journal of International Economics,* 70(2): 384-405 .

6. Combes, Pierre-Philippe, Miren Lafourcade and Thierry Mayer. 2005. "The Trade-Creating Effects of Business and Social Networks: Evidence from France." *Journal of International Economics*, 66(1): 1–29.

7. Coval, Joshua D. and Tobias J. Moskowitz. 1999. "The Geography of Investment: Informed Trading and Asset Prices." *Journal of Political Economy,* 109(4): 811–841.

8. Ellison, Glenn and Edward L. Glaeser. 1997. "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach." *Journal of Political Economy,* 105(5): 889–927.

9. Freund, Caroline L. and Diana Weinhold. 2004. "On the Effect of the Internet on International Trade." *Journal of International Economics*, 62(1): 171–189.

10. Ghemawat, Pankaj. 2001. "Distance Still Matters: The Hard Reality of Global Expansion." *Harvard Business Review,* 79(8): 137–147.

11. Goolsbee, Austan. 2000. "In a World Without Borders: The Impact of Taxes on Internet Commerce." *Quarterly Journal of Economics*, 115(2): 561–576.

12. Head, Keith and Thierry Mayer. 2002. “Illusory Border Effects: Distance Mismeasurement Inflates Estimates of Home Bias in Trade.” Centre d’Etudes Prospectives et d’Informations Internationales (CEPII) Working Paper 2002–01.

13. Hillberry, Russell and David L. Hummels. 2003. “Intranational Home Bias: Some Explanations.” *Review of Economics and Statistics,* 85(4): 1089–92.

14. Hillberry, Russell and David L. Hummels. 2008. “Trade Responses to Geographic Frictions: A Decomposition Using Micro-Data.” *European Economic Review,* 52(3): 527-550.

15. McCallum, John. 1995. “National Borders Matter: Canada-U.S. Regional Trade Patterns.” *American Economic Review,* 85(3): 615–623.

16. Portes, Richard and Helene Rey. 2005. “The Determinants of Cross-Border Equity Flows.” *Journal of International Economics*, 65(2): 269–296.

17. Rauch, James E. and Vitor Trindade. 2002. “Ethnic Chinese Networks in International Trade.” *Review of Economics and Statistics*, 84(1): 116–130.

18. Roth, Alvin E. and Axel Ockenfels. 2002. “Last-Minute Bidding and the Rules for Ending Second-Price Auctions: Evidence from eBay and Amazon Auctions on the Internet.” *American Economic Review*, 92(4): 1093–1103.

19. Santos Silva, João M.C., and Silvana Tenreyro. 2006. “The Log of Gravity.” *Review of Economics and Statistics,* 88(4): 641–658.

20. Tinbergen, Jan. 1962. *Shaping the World Economy: Suggestions for an International Economic Policy*. New York: The Twentieth Century Fund.

21. Wolf, Holger C. 2000. “Intra-National Home Bias in Trade.” *Review of Economics and Statistics*, 82(4): 555–63.

Figure 1

Impact of Distance on Internet Trade

In this figure we graph the results of a regression in which we ran specification (3) at the city-level. We used a flexible specification for distance by constructing indicator variables that take the value of 1 if the distance between buyer and seller is within a certain interval; the figure plots the regression coefficients for these distance indicators. We set the log of distance to be zero if the buyer and seller and located the same city. We use a stratified sample of eBay listings with U.S. buyers and sellers taken between February and May 2004. Distance between cities is measured as the great-circle distance.

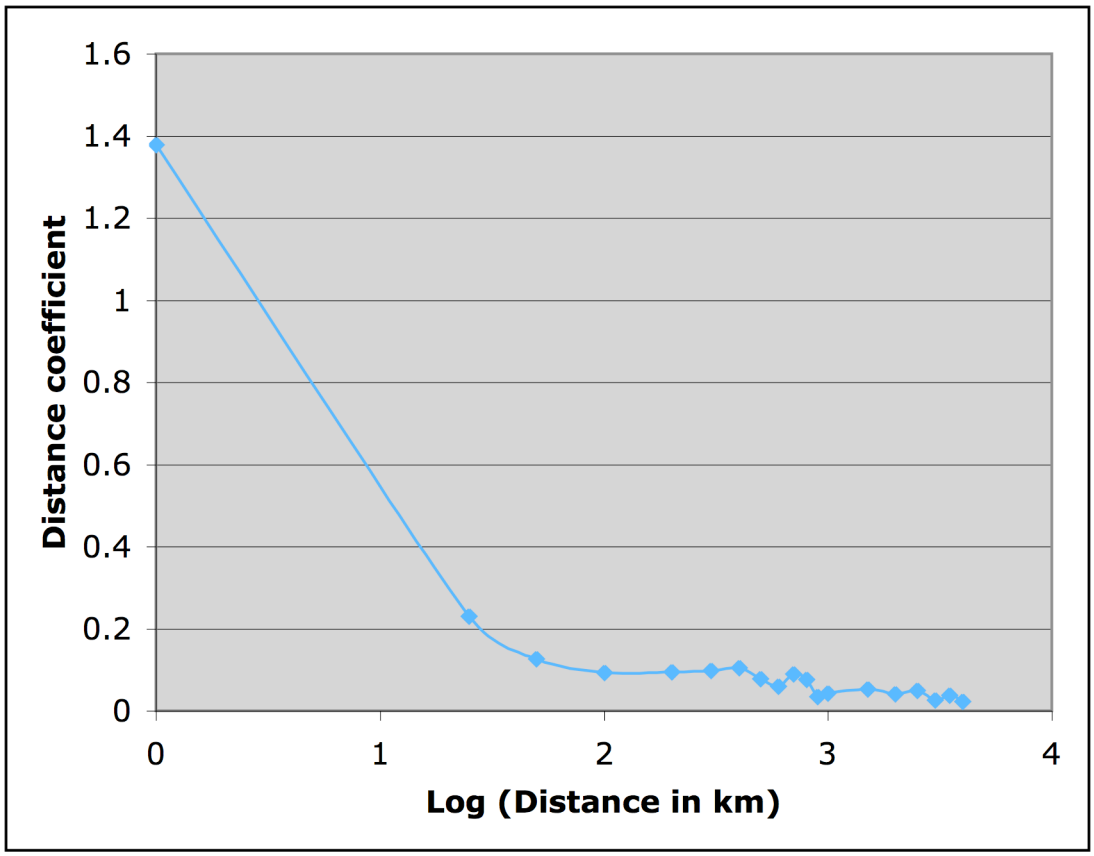

Table 1

Summary Statistics

Panel A. eBay Sample

In this table we summarize the composition and price distribution of our eBay sample. We use a stratified sample of eBay listings taken between February and May 2004 that involve U.S. buyers and sellers. We include in the sample only those sub-categories of items that could be clearly identified as non-local.

			Price Quartiles ($)		
Category	Number of Transactions	Average Price ($)	Q1	Median	Q3
Collectibles	2,760	48.84	4.99	9.99	24.33
Everything Else	3,199	16.71	3.51	7.50	15.00
Toys & Hobbies	3,829	21.75	3.99	9.25	19.99
Dolls & Bears	3,886	25.79	5.57	10.50	23.00
Stamps	2,822	21.80	3.00	6.00	13.50
Books	3,884	14.08	2.99	5.52	10.50
Jewelry & Watches	3,908	175.13	4.99	13.99	102.50
Cons. Electronics	3,517	49.01	3.50	10.50	44.95
Sporting Goods	2,145	48.27	8.02	18.25	42.99
Art	1,285	50.63	6.00	13.99	46.00
Musical Instr.	6,089	117.38	11.50	31.05	102.50
Cameras & Photo	5,423	84.56	10.00	24.99	76.00
Pottery & Glass	2,894	25.93	7.50	12.39	26.00
Video Games	9,116	17.60	4.99	9.99	19.50
Travel/Luggage	1,568	52.13	9.99	19.99	50.00
Coins	7,857	83.29	6.50	14.50	41.50
DVDs & Movies	5,952	15.88	3.99	7.95	12.79
Music	7,476	8.25	2.10	5.00	9.51
Clothing, Shoes, Accessories	4,873	23.85	4.99	9.99	20.01
Home & Garden	4,292	27.79	5.50	10.99	25.03
Business & Indust.	4,222	78.54	7.00	14.50	34.00
Crafts	4,058	10.86	2.95	5.03	10.50
Antiques	1,558	69.49	9.95	18.17	47.12
Health & Beauty	8,230	14.81	4.00	8.99	16.49
Entertainment Memorabilia	2,218	20.63	4.99	9.99	16.99
Computers & Networking	4,820	109.46	9.99	24.50	69.89
Sports Mem., Cards & Fan Shop	11,230	19.22	1.75	4.80	13.50
TOTAL	123,111	44.23	4.25	9.99	24.95

Table 1

Summary Statistics

Panel B. MercadoLibre.com Sample

In this table we summarize the main characteristics of the MercadoLibre.com sample by country. The sample includes all transactions completed through the MercadoLibre web sites during the period of August 2003 to July 2004. For any pair of buyer and seller locations (states/provinces), the firm provided the number and amount of all the monthly transactions completed during the period.

Country	Number of Transactions	Average Price (US$)	Fixed-Price Transactions (Percent)	International Transactions (Percent)
Argentina	628,736	83.6	79.5%	0.8%
Brazil	2,004,677	55.2	85.9%	0.1%
Chile	77,003	148.7	84.8%	0.8%
Colombia	65,491	88.4	92.9%	0.6%
Ecuador	17,501	118.5	95.5%	0.1%
Mexico	258,052	159.4	83.1%	0.3%
Uruguay	31,403	80.7	92.2%	0.3%
Venezuela	174,118	103.3	88.1%	0.4%
TOTAL	3,256,981	75.0	84.8%	0.3%

Table 2

Impact of Distance on eBay Trade

In this table we regress measures of interstate trade on distance and economy size. We use a stratified sample of eBay listings taken between February and May 2004 that involve U.S. buyers and sellers. We include in the sample only those sub-categories of items that could be clearly identified as non-local. The dependent variable—interstate trade—is measured either by the log of the number of transactions (models I-II) or by the log of the dollar volume of commerce (III-IV) between state s (seller) and state b (buyer). We measure distance as the great-circle distance between state capitals. For intrastate distances we use Wolf's (2000) formula, which utilizes the (population-weighted) distance between the two most populous cities within a state. SAME_STATE is a dummy variable that takes the value of 1 if buyer and seller are located in the same state and 0 otherwise. Ln(T_s) is the natural logarithm of the total number of transactions with a seller from state s. Ln(T_b) is the natural logarithm of the total number of transactions with a buyer from state b. The total number of transactions performed by state sellers or buyers proxies for the size of the economy. The results of Wolf (2000) and Hillberry and Hummels (2003) are reproduced in columns 5 and 6 for comparison purposes. Aside from t-statistics for Wolf, standard errors are in parentheses.

	eBay					
	ln(number transactions)			ln($ sales)		
	Model I	Model II	Model III	Model IV	1993 CFS (Wolf, 2000)	1997 CFS[a] (Hillberry & Hummels, 2003)
ln(DISTANCE_sb)	-0.10‡	-0.05‡	-0.07‡	-0.07†	-1.00‡	-1.05‡
	(0.010)	(0.011)	(0.013)	(0.031)	(43.32)	(0.02)
SAME_STATE		0.60‡	0.56‡	1.03‡	1.48‡	0.44‡
		(0.097)	(0.093)	(0.179)	(11.53)	(0.10)
ln(T_s)	0.96‡	0.96‡	(seller f.e.)	(seller f.e.)	1.02‡	(seller f.e.)
	(0.008)	(0.008)			(62.04)	
ln(T_b)	0.95‡	0.95‡	(buyer f.e.)	(buyer f.e.)	0.98‡	(buyer f.e.)
	(0.015)	(0.015)			(59.33)	
Observations	2181	2181	2181	2181	2137	2304
Adjusted R^2	0.93	0.93	0.94	0.78	0.84	0.91

*, †, ‡ significant at the 10%, 5%, and 1% level respectively

[a] Excluding shipments by wholesalers

Notes: Wolf (2000) and Hillberry-Hummels (2003) use the Commodity Flow Survey of the U.S. Census, which covers a representative sample of shipments from U.S. mining, manufacturing, and wholesale establishments.
Wolf (2000) uses driving distances obtained from Rand-McNally.
Hillberry and Hummels (2003) use actual shipping distances collected by the Commodity Flow Survey.

Table 3

Impact of Distance on MercadoLibre Trade

In this table we analyze the impact of distance on international and interprovince trade conducted via the Internet. The sample includes all transactions completed through the MercadoLibre web sites during the period of August 2003 to July 2004. The dependent variable is the log of the dollar value of the transactions between country/province s (seller) and country/province b (buyer). We measure distance as the great-circle distance between country or province capitals. For intracountry or intraprovince distances we use Wolf's (2000) formula, which utilizes the (population-weighted) distance between the two most populous cities within a state. SAME COUNTRY is a dummy variable that takes the value of 1 if buyer and seller are located in the same country and 0 otherwise. SAME PROVINCE is a dummy variable that takes the value of 1 if buyer and seller are located in the same province and 0 otherwise. Standard errors are in parentheses.

	Model I Country Level	**Model II Province Level**
Ln(DISTANCE_sb)	-0.546	-0.382‡
	(0.518)	(0.030)
SAME PROVINCE		1.011‡
		(0.156)
SAME COUNTRY	10.814‡	6.068‡
	(1.757)	(0.080)
Seller fixed effects	Country	Province
Buyer fixed effects	Country	Province
Observations	79	7175
Adjusted R^2	0.69	0.69

*, †, ‡ significant at the 10%, 5%, and 1% level respectively

Table 4

Impact of Distance on eBay Trade: Sensitivity to Sample Composition

In this table we analyze the sensitivity of the impact of distance on eBay trade to the sample composition. We use a stratified sample of eBay listings taken between February and May 2004 that involve U.S. buyers and sellers. We include in the sample only those sub-categories of items that could be clearly identified as non-local. To control for category heterogeneity we run our base specification by aggregating transactions between location pairs at the category level and using category fixed-effects. We also run the specification using data exclusively from one category (once for video games and once for jewelry, respectively). To control for heterogeneous buyer experiences, we classify the auctions in our sample as a function of the number of feedback reports received by the buyer, regardless of their sign. We classify auctions as having new or sophisticated buyers as a function of where they fall in the distribution of feedback reports (below or above the median). To control for heterogeneous seller reputations, we classify the auctions in our sample as a function of the percentage of negative feedback reports received by the seller. We then classify auctions as having good or bad sellers as a function of where they fall in the distribution of negative feedback reports (below or above the median). We use Model IV of Table 2 as reference. The dependent variable—interstate trade—is measured by the log of the volume of commerce between state s (seller) and state b (buyer). We measure distance as the great-circle distance between state capitals. For intrastate distances we use Wolf's (2000) formula, which utilizes the (population-weighted) distance between the two most populous cities within a state. SAME_STATE is a dummy variable that takes the value of 1 if buyer and seller are located in the same state and 0 otherwise. Standard errors are in parentheses.

	Base Case	Category Fixed Effects	Video Games	Jewelry	New Buyers	Sophisti-cated Buyers	Good Sellers	Bad Sellers
ln(DISTANCE_sb)	-0.07‡	-0.03‡	-0.004	-0.04	-0.06‡	-0.08‡	-0.08‡	-0.06‡
	(0.013)	(0.006)	(0.023)	(0.026)	(0.014)	(0.034)	(0.013)	(0.015)
SAME_STATE	0.56‡	0.41‡	0.30†	0.39‡	0.63‡	0.40‡	0.44‡	0.56‡
	(0.093)	(0.029)	(0.15)	(0.14)	(0.11)	(0.09)	(0.09)	(0.11)
Seller state fixed effects	yes	yes	yes	Yes	Yes	Yes	Yes	Yes
Buyer state fixed effects	yes	yes	yes	Yes	Yes	Yes	Yes	Yes
Category fixed effects	no	yes	No	No	No	No	No	No
Observations	2,181	27,599	1,288	975	2,061	2,077	2,083	2,010
Adjusted R^2	0.78	0.59	0.69	0.61	0.91	0.90	0.90	0.90

*, †, ‡ significant at the 10%, 5%, and 1% level respectively

Table 5

Impact of Distance on eBay Trade:

The Role of Shipping Costs and Large States

In this table we test whether the effect of distance on interstate trade is caused by shipping costs, by differences in time zone among states, or by the influence of large states in the regressions. We use a stratified sample of eBay listings taken between February and May 2004 that involve U.S. buyers and sellers. We include in the sample only those sub-categories of items that could be clearly identified as non-local. The dependent variable is the log of the number of transactions between state s (seller) and state b (buyer). We measure distance as the great-circle distance between state capitals. For intrastate distances we use Wolf's (2000) formula, which utilizes the (population-weighted) distance between the two most populous cities within a state. SAME_STATE is a dummy variable that takes the value of 1 if buyer and seller are located in the same state and 0 otherwise. SHIPPING COST is the median transportation cost for shipments from state s to state b in percentage. SAME TIME ZONE is a dummy variable that takes the value of 1 if the buyer and seller are in states with the same time zone and 0 otherwise. Standard errors are in parentheses.

	Model I Baseline	Model II Shipping Rate	Model III Time Zone	Model IV Large States
ln(DISTANCE_sb)	-0.07‡	-0.07‡	-0.07‡	-0.07‡
	(0.013)	(0.012)	(0.015)	(0.012)
SAME STATE	0.56‡	0.56‡	0.57‡	0.56‡
	(0.093)	(0.092)	(0.094)	(0.090)
SHIPPING COST (Percent)		-0.13*	-0.13*	-0.13*
		(0.074)	(0.074)	(0.074)
SAME TIME ZONE			0.002	
			(0.021)	
SAME STATE CA				-0.53‡
				(0.089)
SAME STATE NY				-0.37†
				(0.091)
SAME STATE FL				-0.31†
				(0.093)
SAME STATE TX				-0.38‡
				(0.092)
SAME STATE MT				1.71‡
				(0.150)
Seller state fixed effects	yes	yes	yes	Yes
Buyer state fixed effects	yes	yes	yes	yes
Observations	2180	2153	2153	2153
Adjusted R^2	0.78	0.94	0.94	0.94

*, †, ‡ significant at the 10%, 5%, and 1% level respectively

Table 6

Impact of Distance on eBay Trade: The Role of Trust and Taxes

In this table we test whether the effect of distance on interstate trade is caused by taxes or trust. We use a stratified sample of eBay listings taken between February and May 2004 that involve U.S. buyers and sellers. We include in the sample only those sub-categories of items that could be clearly identified as non-local. The dependent variable is the log of the number of transactions between state s (seller) and state b (buyer). We measure distance as the great-circle distance between state capitals. For intrastate distances we use Wolf's (2000) formula, which utilizes the (population-weighted) distance between the two most populous cities within a state. SAME_STATE is a dummy variable that takes the value of 1 if buyer and seller are located in the same state and 0 otherwise. BAD SELLER is a dummy variable that takes the value of 1 if the median seller rating for that state pair is between 98.2% and 99.3% and 0 otherwise. VERY BAD SELLER is a dummy variable that takes the value of 1 if the median seller rating is below 98.2% and 0 otherwise. (SELLER TAX==X%) are dummy variables to account for the level of seller-state sales taxes; state rates are rounded up to the numbers included; states without sales tax are captured by the intercept. Standard errors are in parentheses.

	Model I	**Model II**	**Model III**
	Seller feedback	**Sales Tax**	**Feedback & Taxes**
ln(DISTANCE)	-0.07‡	-0.07*	-0.06
	(0.012)	(0.039)	(0.040)
SAME STATE	0.50‡	0.029	0.066
	(0.093)	(0.139)	(0.141)
SHIPPING COST (Percent)	-0.13*	-0.12*	-0.12*
	(0.073)	(0.073)	(0.073)
ln(DISTANCE)*BAD SELLER	-0.03‡		-0.03‡
	(0.009)		(0.009)
ln(DISTANCE)*VERY BAD SELLER	-0.04†		-0.04†
	(0.016)		(0.016)
SAME STATE * BAD SELLER	0.63‡		0.73‡
	(0.22)		(0.280)
SAME STATE * VERY BAD SELLER	0.93‡		1.24‡
	(0.14)		(0.171)
LN(DISTANCE) * (SELLER TAX==6%)		-0.05	-0.07
		(0.049)	(0.050)
LN(DISTANCE) * (SELLER TAX==5%)		-0.01	-0.02
		(0.045)	(0.046)
LN(DISTANCE) * (SELLER TAX==4%)		0.02	0.01
		(0.045)	(0.045)
LN(DISTANCE) * (SELLER TAX==3%)		-0.01	-0.02
		(0.066)	(0.066)
LN(DISTANCE) * (SELLER TAX==0%)		0.06	0.057
		(0.060)	(0.061)
SAME STATE * (SELLER TAX==6%)		0.40*	0.04
		(0.237)	(0.211)
SAME STATE * (SELLER TAX==5%)		0.44†	0.40†
		(0.189)	(0.190)
SAME STATE * (SELLER TAX==4%)		0.44‡	0.40‡
		(0.167)	(0.168)
SAME STATE * (SELLER TAX==3%)		0.84	0.63
		(0.541)	(0.559)
SAME STATE * (SELLER TAX==0%)		1.14‡	1.11†
		(0.446)	(0.439)
Seller state fixed effects	Yes	Yes	Yes
Buyer state fixed effects	Yes	Yes	Yes
Observations	2153	2153	2153
Adjusted R^2	0.94	0.94	0.94

*, †, ‡ significant at the 10%, 5%, and 1% level respectively

Table 7

Impact of Distance on Trade Patterns of Different Types of Goods

In this table we test whether the effect of distance on interstate trade is caused by taxes or trust. We use a stratified sample of eBay listings taken between February and May 2004 that involve U.S. buyers and sellers. In this table we rank the coefficients of the same-city dummy variables in regressions of measures of intercity trade on distance and economy size by category of good traded. We run the regression for each of the 27 main categories of goods on eBay (aside from real estate and autos). The dependent variable is the log of the dollar value of the transactions between city s (seller) and city b (buyer). We use seller- and buyer-city fixed effects to control for economy size. SAME_CITY is a dummy variable that takes the value of 1 if buyer and seller are located in the same state and 0 otherwise.

Category	SAME_CITY coefficient	Standard error
Tickets	1.80‡	0.10
Sports Mem, Cards & Fan Shop	0.78‡	0.09
Travel/Luggage	0.53‡	0.11
Cameras & Photo	0.52‡	0.11
Clothing, Shoes & Accessories	0.48‡	0.09
Jewelry & Watches	0.48‡	0.11
Video Games	0.46‡	0.09
Pottery & Glass	0.44‡	0.11
Home & Garden	0.42‡	0.09
Toys & Hobbies	0.41‡	0.10
Business & Industrial	0.39‡	0.12
Consumer Electronics	0.38‡	0.12
Sporting Goods	0.38‡	0.09
DVDs & Movies	0.37‡	0.10
Music	0.36‡	0.11
Computers & Networking	0.30‡	0.09
Health & Beauty	0.27‡	0.09
Musical Instruments	0.25‡	0.09
Antiques	0.24*	0.14
Coins	0.23†	0.12
Crafts	0.23†	0.11
Everything Else	0.19*	0.11
Stamps	0.18	0.20
Dolls & Bears	0.16	0.12
Books	0.15*	0.09
Collectibles	0.14	0.10
Art	0.09	0.15
Gift Certificates	-0.01	0.22
Entertainment Memorabilia	-0.01	0.15

*, †, ‡ significant at the 10%, 5%, and 1% level respectively

Table 8

Impact of Distance on Trade Patterns of Different Types of Goods

In this table we regress the impact of distance on trade on characteristics of the goods traded and the reputation of their sellers. The dependent variable is the coefficient of the same-city dummy variables from regressions of measures of intercity trade on distance and economy size by category of good traded. We exclude from this regression the Tickets and Sports Memorabilia categories. Avg. Weight is the average weight of the goods sold in the category, based on our estimation of item weights from 50 randomly sampled listings from each category. Seller's reputation is measured by the average percentage of negative feedback received by sellers in the category. Standard errors are in parentheses.

Dependent Variable: Coefficient on SAME_CITY	
Avg. Weight in Category	-0.001
	(0.007)
Avg. (Percent) Shipping Cost in Category	-0.001
	(0.001)
Percent Negatives in Average Seller's Record	0.170†
	(0.073)
Avg. Price in Category	0.001
	(0.001)
Observations	27
Adj. R^2	0.17

*, †, ‡ significant at the 10%, 5%, and 1% level respectively