The Geography of Trade on eBay and MercadoLibre

(*Very* preliminary and incomplete. For seminar participants only, please do not distribute without authors' permission.)

September 19, 2005

Ali Hortacsu
University of Chicago and NBER

Asis Martinez-Jerez
Harvard Business School

Jason Douglas

Introduction

An extensive literature in international economics analyzes the impact of distance on trade flows. The stylized finding of a large numbers papers estimating the "gravity equation" is that trade volume among two countries increases with the size of their economies and decreases with the distance that separates the countries. A subset of these papers also report a very significant "border effect": controlling for distance, trade between two regions is much lower if the goods have to cross national borders. 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 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 hinder geographically distant buyers and seller from finding each other; communication barriers, which hinder efficiency of negotiations[1]; and, more generally, "contracting costs," driven by the inability to monitor and sanction misconduct of distant transacting parties.[2]

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 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 our results 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.

The online auction environment provides an interesting and unique opportunity to study the distance-dependence of trade, as the environment can be considered "frictionless" in certain important dimensions.[3] First, search costs are practically nonexistent on these markets, since the computerized search interface matches buyers almost instantly with the listings they might be interested in. Communication costs are also very low, given that most listings follow a fairly uniform format and language, and all transacting parties have access to e-mail for fast and reliable communication. 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. [4] Note also that, within the

[1] Portes and Rey (2002) find a negative correlation between telephone traffic and bank presence and the distance effect.

[2] Rauch and Trindade (1999) find that the percentage of Chinese immigrants in two countries is associated with the level of trade between these countries. Anderson and Marcouiller (2002) show that country-level indices of institutional quality are associated with trade flows.

[3] The "home bias" literature in finance can also be characterized as studying an environment that is similarly "frictionless." For example, Coval and Moskowitz (1999) find that mutual funds are likely to hold regionally-biased portfolios, arguing that monitoring costs may be an important factor.

[4] One may also consider the inconvenience caused by the time it takes one 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.

continental U.S., tariffs are nonexistent, and sales taxes, which are imposed by the states, should encourage out-of-state purchases as opposed to within state purchases.

Our main result is the following: even on eBay and MercadoLibre, distance has a strong negative effect on trade. This effect is highly non-linear, trading volume being abnormally high within the same city and abnormally low across country borders. Once beyond the driving distance of the city limits, the effect of distance on trade is relatively small between cities in the same state or country. Further results suggest that "trust" may be a significant contributor to the distance effect: we find that high ticket items (such as jewelry) have more difficulty crossing city limits, and that the "same city effect" is much more pronounced in those categories where the average seller reputation is lower. We also find some evidence that local tastes and item uniqueness have the opposite effect on the impact of distance on trade. The same city effect is strongest in location-specific items such as tickets or sports cards, and weakest in scarce items such as coins and collectibles.

To our knowledge, our paper is closest in spirit to the work of Blum and Goldfarb (2005), who analyze the web-surfing behavior of a panel of Internet users. They find that local culture appears to be an important driver of digital goods consumption – Korean Internet users are much more likely than non-Koreans to read online editions of Korean newspapers. Our findings reinforce their message regarding the importance of local tastes, but argue that this is not the only source of "home-bias" on the Internet. Our work also relates to Freund and Weinhold (2000), who analyze the impact of Internet penetration across countries on cross-country trade patterns, and find that Internet connectivity is associated with increases in trade volume. Our results maybe interpreted as pointing out a potential limit for the ability of the Internet to reduce geographic barriers – while the Internet may help reduce the distance barrier, as Freund and Weinhold's (2000) results show, factors such as trust and local tastes may be difficult to get around even with this technology. Especially, in the context of "trust," we contribute to the literature by showing that information about the *seller*, such as a reputation rating, seems to be more relevant than information about the *good*. A potential implication of this finding is that the growth of online marketplaces may be limited by their ability to develop exchange-like clearing mechanisms that dissociate the flow of goods from the identity of market participants. Simply reducing the information cost about the good traded will not allow the frictionless growth of the network of buyers and sellers.

Motivation of the Empirical Analysis

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 depends on the characteristics of the good auctioned and the characteristics of the seller, including the buyer's geographical proximity to the seller. Higher willingness to pay will lead to 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 of transactions we should observe between agents from any pair of geographical locations.

Let's assume that in a certain auction there are *S* seller types and *B* buyer types. We will assume that a type is defined by the individual's geographical location. However, there is no problem in extending the concept of type to include other agent characteristics, such as reputation. In the auction will participate N_s sellers of type *s* and N_b buyers of type *b.* When seller *j*, of type s_j, puts up an item for sale, the willingness to pay of buyer *i* of type b_i for the item, υ_{ij} will be:

$$\upsilon_{ij} = \gamma + \mu_{b_i,s_j} + \varepsilon_{ij}$$

where γ is some fundamental value of the item auctioned, μ_{b_i,s_j} is the valuation attached by buyer *i* on the item sold by seller *j,* and ε_{ij} is an IID disturbance. In its most simple interpretation, μ_{b_i,s_j} represents how the distance between the locations of seller *j* and buyer *i* affects the willingness to pay. For instance, it is intuitive to think that buyer *i* will reduce her valuation by the cost of hauling the item from seller *j*'s location to her location, or that she will factor in tax differentials among locations. It is possible to think without much effort that μ_{b_i,s_j} may be affected by the nature of the good being auctioned or the reputation of the seller. For instance, distance will have more impact on more valuable items or on less reputable sellers because the cost of a recourse action increases with remoteness.

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_{ij} follow the Type-I extreme value distribution we can express the probability that a buyer of type *b* wins an auction in which the good is sold by a seller of type *s* as

$$\Pr\{\text{type b wins auction of type s}\} = \frac{N_b \exp(\gamma + \mu_{b,s})}{\sum_{b'=1}^{B} N_{b'} \exp(\gamma + \mu_{b',s})} \qquad (1)$$

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 happen for 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 one buyer of this type. Observe that if the geographical 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 economies with a larger number of buyers will be more likely to buy the item.

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 of type *b,* N_b:

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

where

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

The total number of transactions between buyers of type *b* and sellers of type *s* will depend on the total number of goods auctioned by sellers of type *s* and the probability for buyers of type *b* to win those auctions. Multiplying equation (1) by T_s we obtain the expected number of transactions between buyers of type *b* and sellers of type *s*. Taking logs and including a disturbance term, $\nu_{b,s}$, we obtain the following expression of the Gravity equation:

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

where $T_{b,s}$ is the total number of transactions between buyers of type *b* and sellers of type *s*, T_s is the total number of goods auctioned by sellers of type *s,* and k_s is a constant term capturing the effect of the fundamental value of the good, γ, and the c_s.

The gravity equation suggests the following testable hypotheses: (1) The total number of transactions (we will repeat the analysis with the total dollar value of the transactions) between buyers of type *b* and sellers of type *s*, T_s is proportional to the size of the economy of the buyers, $\log N_b$, and of the sellers, $\log T_t$; (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 the 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.

eBay and MercadoLibre

We developed this study with the data of two leading online auction sites: eBay, arguably the firm that popularized the concept of online auctions, and the largest player in this industry; and MercadoLibre, 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 Internet commerce. According to the Forrester Technographics survey, in 2004, close to 30% of U.S. households had bid in an Internet auction. Therefore, our results will be indicative of how the Internet may affect the geography of trade.

eBay was founded by Pierre Omidyar in 1995 in San Jose. Since then it has continuously grown 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 a total of $34.2 billion worth of merchandise transacted.[5] MercadoLibre was founded by Marcos Galperin in 1999 in Buenos Aires, and in 2001 eBay became its largest shareholder when it bought 20% of the company as part of a strategic partnership agreement. With sites in nine Latin American countries[6] and more than 9.5 million items transacted in 2004 for an aggregate value of $425 million, MercadoLibre is the leading Latin American and the second largest world online auction site.[7]

The eBay data is the result of a stratified sampling of eBay listings taken between February and May 2004. From each of the 30 main categories of items in eBay, we extracted a daily random sample of 1000 listings. By the end of the data collection period we had over 3 million total listings, 41% of which resulted in an actual sale. For each sample 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. This allowed us to obtain the location of 27% of buyers in these transactions. This missing data problem skews our sample towards recording the locations of buyers who are more "experienced" traders on eBay, as they have to be both buyers and sellers in a short period of time.

MercadoLibre gave us comprehensive statistics of the geographic patterns of trade for its different websites. Thus for any pair of buyer and seller locations (states/provinces), we have the number and amount of all the monthly transactions completed during the period August 2003 to July 2004. We got this information for each of the main 30 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.[8] The advantage of having all the transactions in the database eliminates any measurement error associated with the sampling procedure used with the eBay data, but comes at the price of not being able to get all the listing characteristics that might affect demand. Nevertheless, having the location information for all the buyers avoids the missing data problem of the eBay sample and allows the data obtained from MercadoLibre to act as a robustness check of our analyses.

---

[5] eBay Annual Report, 2004

[6] Argentina, Brazil, Chile, Colombia, Ecuador, Mexico, Peru, Uruguay and Venezuela. At the time of our study the Peru site was not yet operational.

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

[8] 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 as visited on September 12, 2005)

There are several reasons why these data sources are especially relevant not only for firms that try to expand their network of users but also for economists who try to understand the geography of trade, as they can help us to understand the sources of frictions that cause distance to have a negative impact in commercial activity. First of all, these marketplaces are pure intermediaries that facilitate trade in a variety of goods. 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 varieties, helping us to understand which product characteristics make trading more sensitive to distance. Second, 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. Third, shipping and handling fees are often quoted explicitly by seller on these websites, and, for a wide class of goods, these fees apply uniformly to buyers of differing locations (at least within the same country). Thus, shipping cost differentials across locations are effectively controlled for (except for variation 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 site, although encompassing a large variety, are mainly new and used household durables, and thus extrapolations to other goods categories is not possible. Second, a similar "representativeness" criticism may also be leveled against the demographic characteristics of the users of these websites. However, one may apply the same criticisms to the Internet trade in general. At a minimum, our analyses may shed some light on what are the frictions behind the distance effect for goods and agents currently involved in Internet commerce, and provide hints on how to overcome these frictions.

Gravity Equation

In this section we analyze whether physical distance between buyer and seller reduces the intensity of Internet trade or not. 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 the purchasing behavior of buyers when the seller is close or faraway.

In a regression framework based on equation (3), the vanishing of the distance effect will mean that variables that proxy for distance between buyer and seller will have no explanatory power when the dependent variables are measures of internet trade.

Table F 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 U.S. 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 even when they include a dummy for observations in which buyer and seller are located in the same state. Two-tier pricing by shipping companies, elimination of tax differentials, and the logistical possibility of an in-person visit to enforce the purchase agreement suggest the need to confirm that the distance effect is not

a spurious relationship created by two distinct types of observations: intrastate and interstate trades. Even if the relationship is not spurious, there is likely to be a discontinuity in the relationship. Thus the need for testing this specification.

When we run the regression with eBay data, we obtain results that have a sign consistent with prior studies, but that are smaller in magnitude. In all the specifications of the regression, the coefficients on distance are negative and significant while the coefficients on the same state dummy are also significant but positive. These signs are robust whether we measure trade by the number of transactions – the specification that we derived theoretically in the prior section – or by the dollar value of the trade, and whether we use the raw distance between the buyer and seller state or its log.

When we compare our results to the models of reference we observe that the effect of distance in Internet trade is much smaller that in the Census data. The coefficients of distance in the eBay regressions are roughly one tenth of the magnitude of the coefficients in Wolf (2000) and Hillberry and Hummels (2003). However, the coefficient on same state is of similar level – about ½ that of Wolf and the same level of Hillberry and Hummels. The combination of these two results seems to indicate that though eBay is fairly effective in mitigating the effect of distance in interstate commerce, a "home bias" still continues to exist.

The analyses using the MercadoLibre data support these findings. Table FA shows the results of the analyses. In the first three models, distance is measured by the distance between country capitals; in the last three, by the distance between provincial capitals. We observe a negative effect of distance that 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 that the same province effect. This difference may just be caused by the customs barriers, but it may also be due to an amplification of the frictions manifested in the "same province" effect.

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 of state borders raises the question of what causes this force and whether state lines are the critical distance point where a discontinuity occurs. In the next sections we will review some of the frictions that have been pointed out as responsible for the distance effect, such as shipping costs, trust, and taxes, to test whether these frictions are still present in Internet commerce and whether state lines are the only or even the most important point of discontinuity in the distance effect.

Frictions to Internet Trade

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. As we saw above, a puzzling pattern of internet trade is that frictions to trade seem to remain strong within the borders of a state but are mitigated

once commerce crosses state borders. For that reason, in this section we will give special consideration to the possible reasons behind the observed discontinuity.

Before analyzing the different frictions, we need to establish that the same state effect is not artificially induced by the states with the largest area. If the effect of distance were completely linear, the size of the same state effect would be proportional to the average radius of transportation within the state. If we were to consider state-specific dummy variables, geographically large states would have larger coefficients simply because their borders span larger markets. One can argue that the same state effect observed before was caused by large amount of trade that occurred within the borders of the largest states. Table G Model 3 results suggest that this is not the case. The coefficient on the same state dummy remains positive, significant, and approximately at the same level when we include individual dummies for the intrastate trade of the five largest states. Moreover, in the case of California, the coefficient is negative and significant, maybe because the influence of Silicon Valley makes Californians more comfortable with internet commerce.

The most evident barrier to distance-less commerce is shipping costs, which are likely to increase with distance. Also, it is possible for transportation companies to have a two-tier pricing structure for interstate and intrastate transport, which gets reflected in the shipping cost quoted by sellers and would generate the same state effect. In the previous section we assumed that the flat shipping & handling cost usually quoted in eBay listings was the norm. In models 1 and 2 of Table G we relax this assumption by explicitly including the shipping cost in the regression framework. Shipping costs have the expected negative, and statistically significant, impact on trade activity, and they seem to be partially responsible for the distance effect, as the coefficient on the distance variable falls when shipping costs are included. However, this coefficient remains negative and significant, suggesting that shipping costs are only a partial explanation for the distance effect. Moreover, they do not seem to be behind the "home bias," as the coefficient of the same state dummy remains positive, significant, and at about the same level.

Previous research[9] has shown that online bidders commonly wait to place their bids until just before the auction expires, a strategy known as snipping. This strategy may be more difficult to implement for a specific auction if buyer and seller are in the same time zone. Less frequent trading across time zones may cause the distance effect, as states in different time zones tend to be further away, on average, than states in the same time zone. We test this conjecture in Table G, Model 2 by including a dummy variable that takes the value of 1 if buyer and seller are in the same time zone and 0 otherwise. The coefficient of this variable is not statistically significant and its inclusion has virtually no impact on the other variables of the model. Thus, differences in time zones do not seem behind the distance effect. eBay's proxy bidding[10] and third party services that provide snipping strategies (e.g., eSnipe.com) seem to have eliminated this potential friction.

[9] Roth and Ockenfels (2002), Bajari and Hortacsu (2003).

[10] When a buyer places a bid, she enters the maximum amount she would be willing to pay for the item. That maximum amount is kept confidential from other bidders and the seller. The eBay system compares

One of the difficulties of trading with a distant party is the difficulty of exerting any recourse if the other party does not fulfill 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 the trust of users in the online commerce platform, including Paypal, buyer protection, and a feedback system. Several papers in the empirical industrial organization literature show that bidders reduce their willingness to when they observe certain levels of negative feedback. 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 H, we interact distance and the same state dummy with dummy variables that identify bad sellers (98-99% rating) and very bad sellers (<98% rating). The significant coefficients in the interactions suggest that trust contributes to the effect of distance and the feedback system helps to mitigate, although not to completely eliminate, this effect. Furthermore, the impact of bad feedback is less visible within 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.

Finally, taxes are often associated with geographic patterns of trade. Goolsbee (1999) shows that internet purchases are partially driven by sales tax optimization. As a general rule, only intrastate internet trade is subject to sales tax. 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. Although it is difficult to conceive a positive coefficient on the same state dummy, one may argue that the value of the coefficient may be affected by the unduly influence of low or no sales tax states. Models 2 and 3 of Table H provide no evidence of sales taxes causing the distance effect. As expected, none of the interactions between the dummies identifying the tax regimes of the sellers and distance is significant. Moreover, when interacted with the same state dummy, coefficients do not monotonically increase as the sales tax rate falls, and even though the higher sales tax regime (absorbed in the intercept) has a negative coefficient relative to the other regimes, it may be caused by the weight of California, an exception to the same state effect as we saw above.

In summary, although shipping costs seem to deter internet commerce between distant buyer and seller, it only explains partially the effect of distance observed in online

---

the bid to those of the other bidders and places bids on the bidder's behalf, using only as much of her bid as is necessary to maintain her high bid position (or to meet the reserve price). The system will bid up to the bidder's maximum amount. (http://pages.ebay.com/help/buy/proxy-bidding.html) visited on September 17, 2005.

trading. Moreover, trust seems to be the only variable that has some impact on the same state effect. This finding is consistent with the possibility of a direct recourse in case of breach of contract, which provides a strong incentive for agents to keep trading relationships within limited radius. The question that arises is whether state limits are the relevant distance for intense commerce or if shorter radii, such as city limits or driving distance, are more important.

Same City Effect

The existence of a strong positive "same state" effect on 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 outside. 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 identify the point of discontinuity in the distance effect more precisely, we run a regression of the log of dollar value of trade between two cities on a series of dummy variables that take the value of 1 if the distance between buyer and seller is within a certain interval: a same city dummy and seller, and buyer city fixed effects. We graph the coefficients of the distance dummy variables in Figure A. It is remarkable to observe how the coefficients for all the distance intervals have similar levels and decrease smoothly as the distance increases, whereas the same city coefficient is about eight times the other coefficients. This result suggests that contrary to what one would expect, "driving distance" is relevant for commerce in the Internet. 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.

In the prior section we found that the same state effect was at least partially caused by direct enforcement ability and mitigated by reputation mechanisms. To explore what is behind the same city effect, we rerun the regression on distance dummies for each of the 30 main item categories in eBay. Table I 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 and baseball cards, which are very location-specific items. A person in Sacramento is less likely to buy a ticket to the Boston Lyric Opera than a Bostonian, and buyers in Minnesota are less likely to buy a Seattle Mariners card than a person in Seattle. Thus, it seems that cultural factors have an important role in causing the same city effect. However, we see at the bottom of the list items that also seem to have a strong cultural component, such as entertainment memorabilia, collectibles, coins, dolls, and bears. In contrast with the baseball cards, those items are less location specific. Moreover, these are items characterized by their uniqueness and for which buyers are more likely to expand geographically their search in order to buy them. Thus, same city effect observed for this items is smaller. Finally, a glance at Table I may suggest that as we go up the list the average value and size of the

items increases, consistent with shipping costs and trust being relevant forces behind the same city effect.

To further explore the impact of trust and shipping costs in shopping the same city effect, we regress the category specific coefficients of the same city dummies on the percentage of negative feedback in the median seller record as a measure of trust, and on the average weight and price of an item in the category. Table J 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 both the reputation measure and the average price in the category. An interpretation of these results is that as the likelihood of a breach of contract increases (the seller has a more negative reputation), or as the cost of such a breach (value of the item) increases, it is more important to have the possibility of a direct enforcement mechanism, measured by proximity in the same city effect. In contrast, the coefficient on weight is nonsignificant, albeit negative, suggesting that shipping costs, usually a function of weight, are not determinant of the same city effect.

In summary, the results in this section support the hypothesis that despite the facilitation of search provided by the Internet, the city limits or "driving distance" cause an "attraction field" that concentrates an excessive amount of trade within them. Cultural factors and the possibility of a direct enforcement action in case of breach of contract determine the existence/need of a local market even in the Internet space.

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 geographical distance is less of a deterrent to trade than it has been observed in studies of non-internet commerce, faraway buyers are still less likely to engage in a purchase agreement than closer ones. Furthermore, the geographical limits of a city, which we call driving distance, act as a field of attraction for transactions, so there is an abnormal concentration of commerce among buyers and sellers who are in the same city.

Further analyses suggest that cultural factors and the possibility of direct contract enforcement in case of breach are the main reasons behind the same city effect. The higher the likelihood of a breach (suggested by bad reputation of the seller), and the higher the cost of such a breach (more expensive items), the less likely a transaction between distant agents will take place. Also, very location-specific items such as tickets and 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 is too small to explain the bulk of the distance effect.

Our findings have two implications for online commerce platforms that want to extend their reach. First, they should continually innovate and perfect systems to increase the trust of market participants on their platforms. Features such as continuous monitoring of

listings, feedback systems, and buyer protection programs are of greater benefit more to distant agents who, in principle, are more defenseless than closer ones. Second, in their expansion drive, they should pay more attention to specialty items, such as collectibles or coins: the uniqueness of these items makes agents more willing to engage in distant trade, and in addition, buyers are more likely to have heterogeneous valuations and therefore prices are more likely to rise if markets are expanded geographically.

Future research could complement our findings by focusing on the impact of distance on prices at which items trade. Particularly interesting would be the analysis of sellers who expand their geographical 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.

Bibliography (to be completed)

1. Anderson and Marcoullier

2. Anderson, James E., Eric van Wincoop, 2004, "Trade Costs," *Journal of Economic Literature* 42-3, 691-751.

3. Bajari and Hortacsu

4. Blum and Goldfarb

5. Coval, Joshua D, and Tobias J. Moskowitz, 1999, "The Geography of Investment: Informed Trading and Asset Prices," *Journal of Political Economy* 109, 811-841.

6. Freund and Weinhold

7. Goolsbee, A.

8. Roth and Ockenfels

9. Stulz, René, 1981, "On the Effects of Barriers to International Investment," *Journal of Finance* 36, 923-934.

10. Portes and Rey

11. Rauch and Trindade

Figure 1

Impact of Distance on Internet Trade

In this figure we graph the ratio between the coefficients in a trade regression of the different intervals of distance and the coefficient of very long distance (>4000 Km). We use a stratified sample of eBay listings with US buyers and sellers taken between February and May 2004. We run a regression of the log of dollar value of trade between two cities on a series of dummy variables that take the value of 1 if the distance between buyer and seller is within a certain interval: a same city dummy and seller, and buyer city fixed effects.

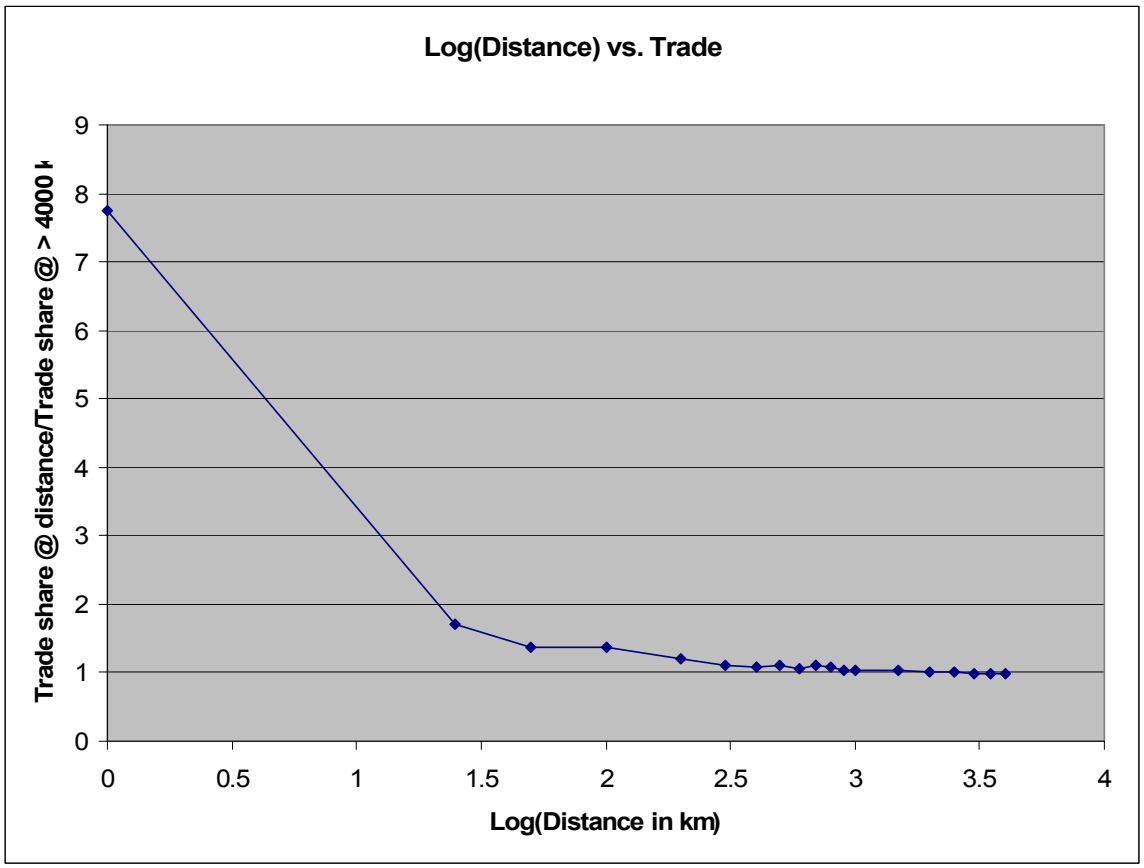

Table F

Impact of Distance on Internet Trade

In this table we regress measures of interstate trade on distance and economy size. We use a stratified sample of eBay listings with US buyers and sellers taken between February and May 2004. The dependent variable – intrastate trade – is measured in models 1-3 by the log of the number of transactions between state s (seller) and state b (buyer) while in models 4 and 5 it is measured by the log of the dollar value of these transactions. 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. The total number of transactions with state sellers or buyers proxies for the size of the economy. Models VI and VII are included for comparison purposes and present the results of Wolf (2000) and Hillberry and Hummels (2003). All coefficients are significant at the 1% level.

	Model I	**Model II**	**Model III**	**Model IV**	**Model V**	**Model VI**	**Model VII**
	EBay	**eBay**	**eBay**	**eBay**	**eBay**	**Wolf (2000)**	**Hillberry & Hummels (2003)**
	ln(# trans)	**ln(# trans)**	**ln(# trans)**	**ln($ sales)**	**ln($ sales)**	**1993 CFS**	**1997 CFS[a]**
DISTANCE_sb (1000km)	-0.096						
ln(DISTANCE_sb)		-0.127	-0.055	-0.10	-0.08	-1.00	-1.05
SAME_STATE			0.43	0.64	0.63	1.48	0.44
ln(T_s)	1.01	0.99	0.97	0.98	(seller f.e.)	1.02	(seller f.e.)
ln(T_b)	1.01	0.99	0.97	0.82	(buyer f.e.)	0.98	(buyer f.e.)
Observations	2297	2297	2297	2297	2297	2304	2304
Adj. R^2	0.98	0.98	0.98	0.91	0.93	0.84	0.91

[a] Excluding wholesale

Notes: Wolf (2000) and Hillberry-Hummels (2003) use the Commodity Flow Survey of the U.S. Census, which covers a representative sample of shipment 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 FA
Impact of Distance on Internet Trade

In this table we analyze the impact of distance on international and interprovince trade through the Internet. The sample includes all the transactions completed in the MercadoLibre sites during the period August 2003 to July 2004. The dependent variable is measured the log of the dollar value of the transactions between country/province s (seller) and country/province b (buyer). For models 1-3 the geography unit is the country while in models 4-6 is the province. 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 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.

	Model I	**Model II**	**Model III**	**Model IV**	**Model V**	**Model VI**
ln(DISTANCE_sb)	-3.536	-0.696	-0.194	-1.898	-0.372	-0.372
	(0.000)	(0.269)	(0.676)	(0.000)	(0.000)	(0.000)
SAME PROVINCE					-0.007	
					(0.993)	
SAME COUNTRY		10.769			5.900	5.901
		(0.000)			(0.000)	(0.000)
Seller fixed effects	country	Country	country	province	Province	province
Buyer fixed effects	country	Country	country	province	Province	province
Observations	79	79	69	6966	6966	6962
Adj. R^2	0.48	0.69	0.57	0.41	0.66	0.66

Table G
Impact of Distance on Internet Trade
The Role of Shipping Costs, Time Zone, and Large States

In this table we test whether the effect of distance on interstate trade is caused by shipping costs, differences in time zone among states, or by the influence of large states in the regressions. We use a stratified sample of eBay listings with US buyer and seller taken between February and May 2004. The dependent variable is the log of the dollar value 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 average 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 buyer and seller and in states with the same time zone and 0 otherwise.

	Model I Shipping Rate	**Model II Time Zone**	**Model III Large States**
ln(DISTANCE_AB)	-0.06***	-0.07***	-0.15***
	(0.02)	(0.01)	(0.04)
SAME STATE	0.84***	0.83***	0.72***
	(0.05)	(0.05)	(0.06)
SHIPPING COST (%)	-0.03***	-0.03***	
	(0.01)	(0.01)	
SAME TIME ZONE		0.04	
		(0.03)	
SAME STATE CA			-0.80***
			(0.08)
SAME STATE NY			0.27***
			(0.10)
SAME STATE FL			1.80***
			(0.13)
SAME STATE TX			0.14
			(0.13)
SAME STATE MT			4.57***
			(1.02)
Seller state fixed effects	Yes	Yes	Yes
Buyer state fixed effects	Yes	Yes	Yes
Observations			
Adj. R^2			

Note: State abbreviations: CA (California); NY (New York); FL (Florida); TX (Texas); MT (Montana).

Table H
Impact of Distance on Internet 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 with US buyer and seller taken between February and May 2004. The dependent variable is the log of the dollar value 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 seller has a rating between 98% and 99% and 0 otherwise. VERY BAD SELLER is a dummy variable that takes the value of 1if the seller has a rating below 98% and 0 otherwise. (TAX==X%) are dummy variables to account for the level of state sales taxes; state rates are rounded up to the numbers included; states with sales tax rate equal to or higher than 7% (e.g. CA) are captured by the intercept.

	Model I Seller feedback	**Model II Sales Taxes**	**Model III Feedback & Taxes**
ln(DISTANCE)	-0.09***	-0.10**	-0.12***
	(0.01)	(0.05)	(0.05)
SAME STATE	0.42***	-0.01	-0.07
	(0.06)	(0.18)	(0.18)
ln(DISTANCE)*BAD_SELLER	-0.01**		-0.01**
	(0.005)		(0.004)
ln(DISTANCE)*VERY BAD SELLER	-0.02**		-0.02**
	(0.01)		(0.01)
SAME STATE * BAD SELLER	0.60***		0.34***
	(0.07)		(0.11)
SAME STATE * VERY BAD SELLER	0.68***		0.40
	(0.24)		(0.25)
LN(DISTANCE) * (TAX==6%)		0.02	0.03
		(0.06)	(0.06)
LN(DISTANCE) * (TAX==5%)		-0.04	-0.02
		(0.06)	(0.06)
LN(DISTANCE) * (TAX==4%)		0.004	0.02
		(0.06)	(0.06)
LN(DISTANCE) * (TAX==3%)		0.13	0.15
		(0.09)	(0.09)
LN(DISTANCE) * (TAX==0%)		-0.10	-0.08
		(0.08)	(0.08)
SAME STATE * (TAX==6%)		0.84***	0.64***
		(0.22)	(0.24)
SAME STATE * (TAX==5%)		0.61***	0.64***
		(0.20)	(0.20)
SAME STATE * (TAX==4%)		0.98***	0.83***
		(0.21)	(0.22)
SAME STATE * (TAX==3%)		1.07***	1.11***
		(0.33)	(0.33)
SAME STATE * (TAX==0%)		0.62**	0.54
		(0.35)	(0.35)
Seller state fixed effects			
Buyer state fixed effects			
Observations			
Adj R^2	0.94	0.94	0.94

Table I
Impact of Distance on Trade Patterns of Different Types of Goods

In this table we rank the coefficients of the same city dummy variables in regressions of measures of intrastate trade on distance and economy size by category of good traded. We use a stratified sample of eBay listings with US buyer and seller taken between February and May 2004. We run the regression for each of the 30 main categories of goods in eBay. The dependent variable is the log of the dollar value of the 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_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
Tickets	3.049
Sports Mem., Cards & Fan Shops	1.571
Jewelry & Watches	1.111
Consumer Electronics	1.085
Home & Garden	0.954
Business & Industrial	0.913
Cameras & Photo	0.83
Travel/Luggage	0.788
Computers & Networking	0.75
Toys & Hobbies	0.706
Antiques	0.684
Pottery & Glass	0.663
Clothing, Shoes & Accessories	0.658
Video Games	0.639
Stamps	0.635
Sporting Goods	0.629
Musical Instruments	0.613
Gift Certificates	0.584
DVDs & Movies	0.532
Music	0.526
Art	0.499
Entertainment Memorabilia	0.448
Books	0.446
Collectibles	0.439
Health & Beauty	0.438
Coins	0.375
Everything Else	0.347
Dolls & Bears	0.288
Crafts	0.201

Table J
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 intrastate trade on distance and economy size by category of good traded. We use a stratified sample of eBay listings with US buyer and seller taken between February and May 2004. We exclude from the regression the categories with extreme same city coefficients. E[Weight] and E[Price] are the average weight and price respectively of the goods sold in the category. Seller's reputation is measured by the median percentage of negative feedback received by sellers in the category.

	Dependent Variable: Coefficient on SAME_CITY
E[Weight] in Category	-0.021
	(0.014)
%Negatives in Median Seller's Record	1.112
	(0.307)***
E[Price] in Category	0.0047
	(0.0012)***
Observations	27
R^2	0.53