An Anatomy of Online Trade: Evidence from eBay Exporters

Andreas Lendle[1]

August 2013

-- Preliminary draft --

Abstract

Online platforms such as eBay offer technologies that facilitate for small entrepreneurs to export. The mechanisms through which exporting is done vary considerably from traditional "offline" exporting activities. Many assumptions of trade models that try to understand the export behavior of firms, such as the presence of market-specific fixed costs to export, are unlikely to apply to eBay trade. I use a new firm-level dataset that covers domestic and international sales made through eBay by sellers based in five different countries. The data is comparable – if not more detailed, to firm-level datasets used in the trade literature. This allows me to compare exporting patterns of online sellers with those of "offline firms". I find that more eBay sellers export, and to many more markets. I show that Armenter & Koren's balls-and-bins model fairly well replicates eBay trade and much more so than traditional offline trade. I argue that this is evidence for the near absence of market-specific fixed costs. However, there is evidence that some barriers remain to become an exporter on eBay.

JEL CODES: F14, F17, L81

Keywords: Firm-level data, online trade, eBay, gravity equation, balls and bins model.

---

[1] Centre for Trade and Economic Integration at the Graduate Institute of International and Development Studies (HEID), Geneva and Sidley Austin LLP, Geneva. Email: andreas.lendle@graduateinstitute.ch

* The author would like to thank participants of the 2013 Villars "PEGGED" workshop and participants at HEID workshops for their constructive comments. I also thank eBay for providing me with data and giving valuable feedback, in particular Hanne Melin and the eBay "data team" in Prague, as well as the World Bank's "Exporter Dynamics Database" team that provided me with detailed "offline data". The author gratefully acknowledges financial support through the Swiss National Science Foundation's ProDoc scheme.

1 Introduction

"One of the most striking features of the micro data is that firm participation in international trade is exceedingly rare." (Bernard *et al.*, 2011)

97% of commercial US-based eBay sellers export.[2]

This paper analyzes a new firm-level dataset of traders on eBay in five countries. eBay is one of the world's largest online marketplaces and allows sellers to reach customers across the globe. The dataset provides new insights into the organization of firms in international trade and how e-commerce can change the way firms operate.

I first explain that selling and exporting on eBay functions very differently from what is normally assumed in models on the export behaviour by firms. I explain that the entry costs into selling on eBay are very small and that there are likely only very small entry costs for particular markets.

By analyzing the data and comparing it with findings from "offline" firm-level data, I then establish that there are indeed striking differences between eBay sellers and traditional ("offline") firms and exporters with regard to the extensive margin. It is a well-known fact from the empirical firm-level literature that most firms don't export, and those that do export reach customers in only a very small number of markets – which is usually explained by fixed costs that firms have to undertake to export. What I observe on eBay is entirely different. Practically all but the smallest sellers export, and most exporters sell to a fairly wide range of foreign markets. These results are even more striking given the fact that most eBay sellers are very small (in terms of annual sales) compared to "offline" exporters, and even the largest ones would still be considered "SMEs" in the "offline world".[3]

These results support my hypothesis that exporting on eBay entails very few, if any fixed costs, whether to export at all, or – even more so – to a particular market, at least when compared to traditional exporters. However, most firm-level trade models would then predict that all sellers sell everywhere, which is clearly not the case. I argue that even when assuming no fixed costs, actual data will still include many "zeroes", e.g., in seller-country combinations. This has to be taken into consideration when testing whether there exist fixed costs to export on eBay. Using Armenter & Koren's balls-and-bins model – which implies that there are no such fixed costs, I show that eBay data can be fairly well replicated using this model, and much better than offline data.

2 Literature review

There is a very wide range of empirical papers using firm-level sales and export data. This interest in the behavior of firms has been driven by theoretical advances in trade theory, but also because such data has become more easily available in recent years. Empirical papers are usually based on either of two types of data: (i) Census data or other data that covers all firms or a (possibly representative) selection of firms operating in a country, or (ii) customs data with export or import transactions that allow the identification of individual firms.

[2] 97% of sellers among all US-based eBay sellers with annual sales of at least USD 10'000 are exporting.
[3] One should note that some eBay sellers may in fact be larger when considering all their sales, including sales made outside of eBay. When comparing such sellers to "offline" exporters, one should also keep in mind that offline exporters are usually *wholesalers* or producers that sell large quantities to wholesalers. Platforms such as eBay allow for international *retail* sales to take place.

Data of the first type has shown that very few firms export, which possibly indicates that there are high barriers to export that only few firms are able to breach. For example, Bernard *et al.* (2007) showed that in 2000 only 4% of US-based firms were exporting. Similarly, Eaton *et al.* (2009 & 2011) showed that only 15% of French firms were exporting. Mayer & Ottaviano (2007) provided similar figures for a number of other European countries.

Data of the second type has shown that among exporters, there is a very high concentration of exports among the largest firms. For example, the 10% largest exporters usually account for 80% or more of all exports (see, for example, Mayer & Ottaviano (2007)). Another empirical fact that has been found among exporter datasets in many countries is that most exporters only sell to a small number of markets. For example, a recently developed dataset by the World Bank (2012), based on customs data of 44 countries, finds that the average number of export destinations across these countries is only 1.4 to 7.4, with the average across all countries being 2.9.[4] Other authors have found similar figures. For example, Bernard *et al.* (2007) found that US exporters sell on average only to around 3.5 markets and 64% of exporters only sell to one single market. The same authors have also shown that a very significant share of exporters sell only one type of products to one market.[5] Some authors have focused on the extensive margin of products being traded. One advantage is that such data is easily available because it can be constructed from published customs data. A notable example is Baldwin & Harrigan (2011), who showed that very large shares of products are not traded in any given year between the US and its trading partners – roughly 90% of country-product observations at the most detailed level are zero.

These facts are usually explained by the presence of fixed costs to enter the export market, to enter a particular destination, or the production and selling of a new product. In Melitz (2003), firms draw their productivity from a random distribution. Because they face fixed costs to then enter the domestic market, and higher fixed costs to enter the export market, only sufficiently productive firms will start producing (or selling), and within those only the more productive ones will also export. This fits with the empirical findings that exporters are larger and more productive. In extended versions of the Melitz model, authors assume fixed entry costs in each destination market, which may vary, but are always positive (e.g., Helpman, Melitz & Rubinstein (2008), Eaton *et al.* (2011), and Baldwin & Harrigan (2011)).

Armenter & Koren (2010) provide, in my view, a very novel and innovative approach to explain these patterns of the extensive margin. They show that *scarcity* of the data can *also* explain some of the stylized facts that have been established by the empirical literature. Their main idea is that because the number of transactions that a seller undertakes, while not *exogenous*, is certainly *limited*, and *small* compared to the number of, for example, possible export markets. For example, the high share of firms that export to only one country is less surprising when one accounts for the number of export transactions that sellers undertake. Many sellers are small and have very few transactions, which means they will "naturally" only ship to few markets, or a small number of different products. They show that some of the stylized facts found in the literature can be replicated by assuming that trade flows are derived from a random distribution of a discrete number of trade flows (one pattern, among others, that cannot be explained by scarcity is the low share of firms exporting). While Armenter & Koren do not suggest that their balls-and-bins model is a plausible model that can fully explain trade

[4] One of the first papers making use of this new dataset is Cebeci *et al.* (2012). Another one, focusing on the importance of very large exporters, is Freund & Pierola (2012).

[5] I will show later that at least in those countries for which I have detailed "offline" data (which I use for comparison purposes), most of these exporters are also single-transaction exporters.

flows, they propose that empirical research takes the sparsity of data into account and focuses on explaining the difference between actual data and what the balls-and-bins model would predict.

The balls-and-bins model requires information about individual shipments. Armenter & Koren estimated shipping numbers per firm using aggregated data. In most firm-level datasets typically used in the literature (or in closely related papers such as Baldwin & Harrigan (2011), using detailed trade data at product level), such information is missing. There are therefore almost no papers that have provided additional evidence on the degree to which actual trade flows can be explained by the model. One notable exception is a recent paper by Shi (2011), who used a very detailed dataset on Chinese export transactions and found that the balls-and-bins model replicates many aspects of the Chinese data, similar to what Armenter & Koren found for the US.

In the world of online trade (such as on eBay), the assumptions behind the balls-and-bins model make a lot more sense than for offline trade because it can be assumed that market entry costs are very low and that transactions are in most cases made independently of each other. eBay trade is thus a prime example to test whether the balls-and-bins model holds. At a minimum, taking into account the sparsity of the data is very important because eBay data is sparse – many sellers make a single transaction in a year. To my knowledge, this paper is the first using detailed "firm-level" data for international online transactions that are comparable with datasets for offline firms and exporters. Two recent papers – Hortaçsu *et al.* (2009)[6] and Lendle *et al.* (2012) – use datasets for online transactions from MercadoLibre and eBay at an aggregated level. While they allow insights into the determinants of aggregated online trade, they do not analyze the behavior of individual sellers.

One paper that is somewhat similar in spirit is Hillberry & Hummels (2007), who use a detailed dataset on intra-national shipments of firms within the US that includes seller and buyer location by ZIP code. Similarly to what I will argue for eBay trade, the authors argue that assuming fixed entry costs "into ZIP codes" makes little sense. Nevertheless, they do find that the number of firms shipping to a state or ZIP code falls with distance, and most firms only ship to very close areas. They provide evidence that this home bias can be explained by local production networks: Many firms only face positive demand for their products by other firms located very closely to them, for whom they produce specialized inputs.

In the next section, I will explain the differences between eBay and offline trade and which empirical findings one should expect to find on eBay.

3 Trading on eBay versus offline

Trading on eBay is very different from traditional exporting. I will thus provide a short introduction how selling on eBay works. A prospective seller has to register with eBay. She then uploads items that she wants to sell (with a description, pictures etc.) on the eBay site[7] and specifies a price.[8] eBay may charge a small fee for listing an item[9] and a final value fee based on the total amount of the

[6] Hortaçsu *et al.*'s data is derived from a sample of individual eBay sellers, but they do not analyze the extensive margin of eBay trade.
[7] eBay has several different eBay sites in major markets. Products listed on one site (e.g., the US site at eBay.com) can also be listed on other sites. A French buyer could buy from a US-based seller on the US site, or the French site, or on a third country site. Typically, transactions are done on the site of either the buyer or seller country, if they have a site.
[8] An auction format can also be used. The data does not contain information on the sales format used. However, nowadays a large share of sales on eBay is conducted as "fixed price" sales.
[9] In some cases, listings are free of charge.

transaction. She also has to specify where the item can be shipped to. For example, one may limit it do the same city, same country, to all countries, or a specific list of countries, or all countries except for certain countries. While I do not have data on the choices made by sellers, the large share of sellers that ship to a wide range of markets indicates that allowing for global shipment to most or all destinations is very common. Notably, the seller will not be able to charge different prices for different countries.[10]

An important distinction to traditional trade is that if a seller allows for shipment to many countries, she will not know beforehand to which country she will be selling. Prospective buyers are browsing through the eBay platform and search for products they want to buy. They will be able to see where a seller is located and whether the seller offers delivery to their country. eBay also provides a sophisticated rating mechanism (for both buyers and sellers), which helps buyers to verify whether the seller is trustworthy. Here, one should note that buyers typically pay first before the product is shipped. They thus face a risk that the seller does not deliver the item, which eBay reduces through specific guarantee schemes and online resolution procedures.[11]. Once the buyer decides to buy a product, he makes the payment (e.g., by bank transfer, credit card or through eBay's PayPal payment system) and the seller ships it. Shipping fees are usually listed in the offer, and paid by the buyer, although some sellers offer "free shipping".[12]

Shipping normally occurs through the postal system or similar services.[13] When the shipment is international, then the seller may have to fill out a customs declaration form, and the buyer may be responsible for paying customs fees in his country. However, if the transaction value is low (which is often the case on eBay), most countries would allow for simplified procedures. That means that the item would often not be subject to any import duties or taxes (such as VAT).[14] On the other hand, if taxes and duties apply, then additional fees may be charged by the postal service or customs, which, in ad-valorem terms, can be very significant. Taking advantage of preferential duties in an RTA may be more difficult for such small shipments because additional documentation may be required. One would thus expect that RTAs are less beneficial for eBay trade. In contrast, a customs union may be very beneficial for eBay transactions because customs controls are removed.

In "traditional" trade, the process is different. A prospective seller of a product faces a number of – possibly very high – fixed costs. First of all, there is a fixed cost to enter the market. There is then a fixed cost to become an exporter, which involves becoming familiar with customs and shipping procedures of the exporting country. In addition, there may be destination-specific fixed costs, such as import procedures and having to deal with destination-specific regulations and possibly a different language.

Shipping costs per transaction – in ad-valorem terms - can be fairly high on eBay due to the small size of transactions and the lack of economies of scale in shipping. However, these costs are rather uniform across foreign destinations, at least for small parcels. Offline, shipping costs are likely to vary more with distance, but are smaller in ad-valorem terms because of bulk shipping. Apart from costs, shipping *quality* can be a major impediment to eBay transactions. If there is, for example, a high risk

[10] The seller would typically charge a higher shipping fee for foreign countries, which should reflect the higher shipping costs. Charging higher (or lower) than actual shipping fees would allow for some price discrimination between countries.

[11] Sellers also face risks. Anecdotal evidence from eBay sellers shows that they are very concerned about receiving a bad rating from buyers, which could severely harm their future business on eBay.

[12] I have data on the shipping fee whenever it is separately listed.

[13] I do not consider products that may be traded electronically. Sellers may also offer bulky items that would have to be picked up.

[14] A similar advantage applies to domestic out-of-state orders in the US, for which often no sales tax is charged.

that an incoming parcel gets lost, then this would make many consumers reluctant to buy from a foreign eBay seller and it would also make eBay sellers reluctant to sell for fear of bad buyer ratings due to slow or no delivery.[15] This can explain why eBay purchases from many countries are fairly small, but because low postal service quality affects *all* shipments, it is not going to affect the distribution of purchases across potential sellers in the exporting country.[16]

Eventually, a transaction only occurs when a seller and a buyer match. One could therefore think of a fixed cost being attached to this. Part of this is a search cost – it takes time and effort to find a trading partner. I expect this cost to be fairly low. On eBay, it is the buyer searching for a seller, not vice versa (offline, it can go both ways). This search takes place on the eBay platform, which allows for a full search across all listed items. Language may however have an effect because eBay operates on different sites in different languages. While it is possible to search across different eBay sites, product descriptions in a foreign language can make it more difficult for a buyer to identify the right product and thus match with the seller. Such search costs also exist in the offline world, and are likely much higher. Another element of the buyer-seller matching is *risk*. Both buyer and seller face a risk that the other party does not comply (e.g., does not pay, does not deliver, delivers a product of low quality, late delivery, etc.). In the offline world, a range of mechanisms exist to reduce such risks (e.g., letter of credits). On eBay, there are mechanisms such as guarantee schemes, online resolution procedures or a feedback forum with a rating system that allows buyers and sellers to rate each other. But some risks remain for both buyers and sellers. Because a successful transaction can reveal information about the other party, this risk should be lower the more transactions are made between the parties. Thus, risk imposes a fixed cost at the buyer-seller level. Because of these two elements of fixed costs to match (search cost and risk), one should expect repeated transactions between sellers and buyers. Unfortunately, neither offline nor eBay data includes information on buyers.[17] As I will explain below, it is an important assumption that can explain why eBay trade patterns (and even more so offline trade patterns) deviate from the balls-and-bins model, under which it is assumed that transactions are fully independent from each other.

Another element of costs are transaction costs. On eBay, sellers pay a percentage fee of the sales value to eBay.[18] Due to this fee structure, there are fewer scale economies on eBay than offline.

In Table 1, below, I summarize the importance of these different fixed costs (and other costs) for offline and eBay trade and the related empirical regularity that we typically see for offline trade and the one that we would expect to see on eBay. For example, if entry costs into becoming an eBay seller are indeed small, then we should expect to see sellers with very few sales. Of course, unlike for exporters versus non-exporters, we do not observe the firms or sellers that do not overcome entry costs into the market. If the cost to become an exporter is high, then one would expect to see few firms exporting, which can be observed clearly in "offline" firm-level datasets.[19] If the hypothesis is correct

[15] Similarly, sellers would find it difficult to export through eBay if the postal service of the exporting country is of low quality, which can be an impediment for the use of such channels for developing country exporters. Since I focus on eBay exports from five developed countries with fairly good postal services, this is less likely to be an issue.

[16] One avenue for further research is to match postal service quality indicators with eBay trade flows.

[17] One notable exception is a new dataset used by Eaton *et al.* (2012). The authors match Colombian export transactions with US import transactions. The figures shown in their paper indicate that repeated transactions between buyers and sellers are common, as one would expect.

[18] In addition, fees for listing may apply. The percentage fee is lower for some product categories and for large sellers. See http://pages.ebay.com/help/sell/fees.html and http://pages.ebay.com/help/sell/storefees.html for the US eBay site.

[19] Note here that one should distinguish between direct and indirect exporting. Many offline firms may see parts of their production exported, but do not engage themselves in exporting (see my discussion on this in Section 6).

that the cost to become an exporter on eBay is low, then we should see a high share of eBay sellers to be exporting. Similarly, I expect that there is practically no additional cost to enter an additional foreign market, and thus eBay sellers, contingent on being exporters, should sell to many different markets. The magnitude of costs offline and online is certainly highly stylized and not necessarily true for all firms or all sectors.

Table 1: Costs online and offline

Type of cost	"Traditional" trade		eBay trade	
	Magnitude	**Expected empirical regularity**	**Magnitude**	**Expected empirical regularity**
Entry cost	High	Large minimum size of firms & small number of firms	Low	Small minimum size of sellers and large number of sellers
Cost to become an exporter	High	Few firms export, or only indirectly	Low	Most sellers export
Destination-specific entry costs	High	Exporters sell to few countries & large minimum sales per country	Low	Exporters sell to many countries and small minimum sales per country
Cost to establish a buyer-seller relationship	High, and increasing with distance (?)	Few and lasting buyer-seller relationships	Low, and independent of distance (?)	Many buyers per seller & many sellers per buyer
Transport (shipping) costs	Low, with variation across destinations	Trade falls with distance	High, but low variation across most (foreign) destinations	Trade is hardly affected by distance
Transaction costs	Low, and falling with firm size	No upper limit for firm size	High, and independent of size	Upper limit for firm size

Many of the predictions listed above can be empirically tested – the key testable hypotheses are summarized in

Table 2, below. However, the expected empirical regularities are different depending on whether Bay trade data is “dense” or “sparse” (the same applies for an analysis of offline trade data). As Armenter & Koren (2010) have shown, the “sparsity effect” has major implications for the interpretation of empirical patterns found in the offline data.

In a trade model using a monopolistic competition framework and heterogeneous producers (e.g., as in Eaton et al, 2011), but no fixed costs to sell to a market, even the least efficient active producer would be able to sell to any market. If trade flows could be infinitesimally small (i.e., the observed data would be “dense”), we would see all sellers exporting everywhere. If data is sparse, then we may not observe a trade flow, even if we should expect a positive trade flow. Let’s say we assume for a given seller’s product there is positive demand from all countries and that there is no fixed cost to sell to a market. If all sellers would sell very large numbers of items during the period that we observe, then the data would show positive trade flows for all seller-destination combinations. However, because eBay sellers list (and sell) a *discrete* (and often very small) number of items per year and because we observe only one (or a few years) of eBay trade, many of these potential trade flows cannot be observed.

Table 2: Predictions when eBay trade is "dense" or "sparse"

Assumption	Expected empirical regularity when eBay data is dense	Expected empirical regularity when eBay data is sparse
No fixed cost to export	All sellers export	Not all sellers export. Probability to be an exporter increases with number of transactions.
No destination-specific entry costs	All sellers export to all countries	Some (many?) seller-destination bins are zero. Number of destinations increases with number of transactions.
	Gravity: Extensive margin (sellers per destination) does not matter.	Gravity: Extensive margin matters more the "sparser" trade is.

For some dimensions of the eBay data, sparsity is indeed less important, e.g., for product-destination combinations.[20] But for seller-destination combinations it is: The number of active sellers in a given year can be more than one million, and there are around 200 possible destination countries to which we observe any exports.[21] A country would see tens of millions in annual eBay imports if all eBay sellers exported there, exceeding the overall "offline" imports of some of the smallest destination countries.

Armenter & Koren's balls-and-bins model provides a benchmark against which one can compare eBay trade patterns. Because their model takes the sparsity of the data directly into account and because its assumptions imply that there are no fixed costs to export, or to sell to a specific destination, finding that eBay trade can be predicted by the model would be strong evidence that my assumption is correct. Similarly, if one assumes that trade is not "random" (as one does for offline trade), the balls-and-bins model should *not* do well in replicating actual data.

As in Armenter & Koren (2010), I assume that the size of eBay sellers (in terms of the number of items that they manage to sell) is exogenous. One may assume that the number of sales that a seller makes is directly derived from its productivity, which is distributed in some exogenous way among sellers (as for example in Melitz (2003)). Because eBay sellers normally do not produce what they sell, productivity could be defined as a seller's capability to purchase items in the market for a value that is below the market price that can be achieved on eBay (over one period). The more productive a seller is, the more items she will be able to acquire at a cost low enough to be sold with a non-negative profit on eBay. eBay sales are roughly distributed according to Zipf's law, which has also been found for offline firms (see Section 5). This suggests one could assume a Pareto-distributed productivity. Productivity can be very low - some sellers will only find it profitable to sell one single item.[22] Thus, the minimum productivity to enter the market is obviously extremely small. The data suggests that there could be an upper limit to productivity and thus seller size because so far there are only a few sellers with sales of more than USD 10 million.

I also assume that country demand (as a share of overall eBay transactions) is exogenous. I therefore do not attempt to explain what drives the overall size of the eBay market. But what I can show is how, given a certain number of transactions sold by a seller, or purchased by a country, such transactions

[20] The reason is that the eBay data is aggregated across only 37 product categories. Thus, relatively few zeroes can be observed. The same can be found in offline data at highly aggregated product categories (HS2), but many zeroes are found at disaggregated levels (see, for example, Baldwin & Harrigan, 2011).

[21] The five exporting countries for which I have data show at least one export transaction to almost all destinations.

[22] With an even lower productivity, sellers would enter the market (i.e., register on eBay and list an item) and then not be able to sell anything.

will be distributed across destinations or sellers and how they would be if trading on eBay entailed no country-specific fixed costs.

In the remainder of this paper, I will show that eBay patterns do indeed follow to a large extent what one would expect in the absence of fixed costs. I will also provide explanations for deviations from the predicted outcome and numerous comparisons with offline data. I can also show that – at least compared to eBay trade - the balls-and-bins model does *not* replicate offline data very well, which shows that trading on eBay is indeed different.

The next section describes the dataset, followed by a few descriptive statistics. I will then focus on the exporter status of eBay sellers before analyzing the extensive margin of eBay exporters.

4 Data sources

I use a new firm-level dataset that covers all domestic and international sales conducted by eBay sellers based in five different countries (US, Germany, France, UK and Australia). [23] The data was provided by eBay. It contains data for all sellers (Australia, France) or all but the smallest sellers (other countries)[24] for the period 2006 to 2011 (US: only until 2010; Australia: until 2012). For each seller, the dataset provides for the number and value of transactions by destination country (including domestic sales), by product category, and by eBay site used.[25] The data also contains the number of shipped items (some transactions include more than one item), and the shipping fee.

The data is anonymous - all eBay sellers are only identified by a number. One "seller" may have different seller profiles, and each seller may offer and sell on different eBay sites. However, its registration code is unique, and the data is aggregated at that level. For the main part of my analysis, in which I focus on the exporter status and exports to individual countries, I mainly use data for the latest available year (2010 for US; 2012 for Australia; 2011 for other countries). Gravity regressions are done using all available years.

eBay transactions are divided into around 40 different product categories. After dropping those categories that cannot be considered as goods (such as tickets and travel), or for which there is a negligible number of observations, we are left with 37 unique product categories.[26] When comparing eBay trade flows with offline trade flows in the gravity regression, I use only a subset of the offline

[23] Note that the location of a seller can be different from the eBay site used. I refer to sellers from a country as sellers being based in that country, independent of which site is used for their transactions.
[24] The Australian and French data includes all sellers, including the smallest ones. The data for DE and UK roughly includes all sellers that had at least 25 transactions in that year, and some of the sellers with fewer transactions. For the US dataset only, the data includes all sellers that had at least 25 sales transactions in the *previous* year, which, by accident, was chosen instead of the more appropriate value for the same year when the data was retrieved. This means that the US data is not necessarily representative for these extremely small sellers.
[25] Each of the five countries has its own eBay site. Most sales transactions are conducted via the site of the country (e.g., eBay.com for the US), while a small fraction of sales is made on other sites. However, this varies a lot and for some countries the share of trade conducted via other sites can be quite high. Also, international transactions are often conducted on the site of the country in which the buyer is located. I do not exploit this dimension of the data in this version of the paper. However, it appears that such sales are mainly language-driven. For example, sales by US sellers made through the Austrian eBay site would show a fairly large share of Austrian buyers, but also of German buyers.
[26] These categories are listed in Table 3.

data that includes only products that are similar to the items traded on eBay. To do that, I use a correspondence table between the 37 eBay categories and the Harmonized System.[27]

I have used a wide range of "offline" firm-level data for comparison. In most cases, I use aggregated figures that are taken from the literature, including from a new World Bank database (see World Bank, 2012). I received detailed transaction-level customs data for two countries (Malawi and Peru) that are covered by this database, which I analyze in more detail. The Australian Bureau of Statistics also provides data on the number of exporters and number of export transactions per destination country, which I have used as well (see ABS, 2012).

Note that a major difference between the structure of the eBay data and that of most firm-level datasets used in the offline literature is the fact that my data contains the number of transactions made. This makes the data much more suitable for an analysis with the balls-and-bins model. The only countries for which such offline data is available are Malawi, Peru and (with limitations) Australia.

5 Descriptive statistics

In this section, I provide a brief overview of the data, based on 2010 data for the US and 2011 data for other countries.[28] The total number of sellers for which I have data is around 4 million.[29] However, the coverage of small sellers varies across countries, which means that the overall number of sellers is even higher. The coverage for large sellers is fairly complete for all countries. Sellers large enough to be considered "commercial sellers", which I define as those with annual sales of USD 10'000 or more, are very numerous, with up to 174'000 per country. When compared with the overall population of the country, the UK and Australia have by far the highest "density" of such sellers, with more than 1'000 such sellers per one million inhabitants. A small share of very large sellers exist, with a small number of sellers having annual sales of USD 10 million or more. This is a remarkable turnover for eBay sellers, but certainly not a lot compared to sales of many "offline" firms.

A significant share of eBay sales are exported, i.e., they are shipped to an eBay user whose location is in a different country. Nevertheless, most sales are still domestic – there is thus a clear "home bias", which can also be found in the "offline" world. The share of sales that are exported (by number of transactions) varies across countries and ranges from 6% to 13%. In the sections below, I will take a closer look into the distribution of exports across destination markets and how this compares to offline trade. Overall, the main export destinations for eBay sellers are, not surprisingly, countries that are large, close and use the same language – which indicates that eBay sales may be similarly distributed as traditional export sales. For example, the main destination countries for US sellers are Canada, Australia and the UK. Overall, across the five eBay markets, one can find exports to practically all countries and customs territories around the world – the entire dataset contains 218 different destinations.

<u>Distribution of seller size</u>

The distribution of seller size, whether measured by number of transactions or in annual sales, is very unequal and roughly follows Zipf's law – but only when I exclude the smallest sellers (roughly those

[27] This is the same approach that was already used in Lendle *et al.* (2012). See there for details. Note that gravity findings hold even when using all offline trade.
[28] Note that due to the sensitivity of some of the data, I can only provide limited aggregated figures of sales and number of sellers.
[29] Note that the actual number of sellers is even higher because I do not have data for most of the smallest sellers, except for Australia and France.

with 100 or less transactions). A similar result has been found for the whole sample of US "offline firms.[30] On eBay, there are more very small sellers compared to what one would expect from a Zipf distribution.

Using some other measures of inequality, one can show that eBay sales are indeed less concentrated among sellers than sales or exports among offline firms. For example, when comparing the sales (export) share of the 5% largest eBay sellers (exporters) with the export share of the largest 5% of offline exporters, I find a much stronger concentration offline than online.[31]

Shipping costs and unit values

The data includes information on shipping costs. Generally, speaking, shipping costs are fairly high in ad-valorem terms, ranging from an average of 9% to 15% across the five countries.[32] Average shipping fees hardly vary *within* foreign destination countries, but are significantly higher (by about 20-100%, depending on the country) than for domestic shipments. The most remarkable difference is in the US, where foreign shipping costs about twice as much on average, with hardly any difference between the most important and closest destination (Canada) and other countries. I run a gravity regression (excluding domestic trade) to look at the impact of distance on shipping costs and found a positive, but fairly small impact. A 10% increase in distance increases shipping costs by only 0.3-0.7%.[33]

Interestingly, I find (for international trade) that unit values slightly decrease with distance, with a distance coefficient of around -0.1. This is in contrast to the findings of Baldwin & Harrigan (2011) for US offline exports. However, my result may easily be driven by a composition effect, for which I cannot fully control because the product categories on eBay are very broadly defined.[34]

Export status

As already mentioned above, the eBay seller-level data shows some remarkable differences to "offline" firm-level literature. The first difference is that almost all sellers export. I consider a seller as an exporter if she shipped at least one item abroad in a given year. The share of sellers exporting varies by seller size. A wide majority of sellers that had at least 25 transactions on eBay in a year is exporting. For fairly large sellers, that number increases to 96%-98%.

The share of exporters among offline firms is much smaller, ranging from 2% to 67% (see Table 8 for an overview). These figures vary widely across and within countries, depending on the source (e.g., results based on datasets that include only large firms tend to have higher exporter shares), but are in all cases much lower than for eBay sellers, despite the fact that these firms are larger than eBay sellers (e.g., in terms of annual sales). One particularly useful example is the comparison for Australia: The share of "offline firms" exporting refers to all registered businesses in Australia – only 2% export.

[30] See, for example, Axtell (2001).

[31] I have compared the eBay shares with data provided by the World Bank (2012) for exporters in a wide range of countries.

[32] It should be noted that the data contains the shipping costs provided by the seller in the listing. Actual shipping costs can be different. For example, a seller may offer "free shipping", which means that shipping costs are included in the purchase price and not observable. Actual shipping costs could also be lower than what is stated by the seller.

[33] This regression and the one explained in the next paragraph were done using importer-year-product and exporter-year-product fixed effects. Results are fairly robust when using other fixed effects (e.g., for importer, exporter, year and product category).

[34] Baldwin & Harrigan's result is based on almost 5'000 different product categories, whereas the eBay data is aggregated across 37 categories only.

Many of these businesses are certainly "too small" to export – but that is also true for eBay sellers. If I take all Australian eBay sellers, many of which only sold a single item in 2012, I still find that around 14% export. When ignoring these tiny sellers, the share of sellers exporting goes up: 77% of "commercial sellers" (annual sales of USD 10'000 or more) are exporting.

Export destinations

Another striking difference between online and offline data is the fact that online sellers export to a very high number of different countries. This is demonstrated in Figure 10, which plots the number of different export destinations against the average value of the total exports of a seller, for both the five countries covered by out eBay data, as well as 44 countries from a recent World Bank dataset (World Bank, 2012). While none of the five "eBay countries" is included in the World Bank data, the 44 countries include a wide mixture of small and large, developing and developed countries. The comparison shows that – not surprisingly – eBay exporters export much less (in terms of annual export value) than their offline counterparts, but – certainly somewhat surprisingly – export to many more different countries. For example, eBay sellers with annual exports of at least USD 1'000 export on average to between 7 and 21 different countries, while offline sellers only export to less than five different countries in all but one of the 44 countries for which such data is currently available. Data for offline exporters in the US, Australia and France also shows that such exporters reach very few markets on average.

Other characteristics of eBay sellers

There are many other interesting characteristics of eBay sellers that can be found in the data, which provides scope for future research. For example, one can analyze survival patterns of eBay sellers.[35] Using the product dimension, one can also observe an interesting product specialization pattern of eBay sellers – large sellers concentrate on fewer product categories. I do not further analyze these characteristics in this paper.

6 Exporter status of eBay sellers

Introduction

In this section, I will compare eBay with offline sellers with regard to their export status. It is a widely cited stylized fact in the (offline) firm-level literature that relatively few firms export, and those that do usually export only to very few markets. A higher share of large firms export, and larger exporters sell to more countries – but the number of destination countries remains small for all but the largest exporters. This suggests that – at least offline - exporting entails costs that only larger and more productive firms can afford or for which they will find it worth investing.[36]

These stylized facts based on offline data should be read with a grain of salt. The share of firms exporting or selling to a large range of countries, especially when one focuses on manufacturers, seems surprisingly low. Although this may be caused by a variety of trade barriers that generate high fixed costs for entering foreign markets, it could also simply be related to the fact that trade data does not reflect value chains in today's interconnected industries. Many firms may supply inputs to other

[35] This can be done for Australia and France, where I have data for all sellers. For other countries, one has to limit the analysis to larger sellers because many small sellers are not included in the data. The data also contains the eBay registration date, which can be used for a survival analysis.

[36] Such costs could for example include finding foreign customers, setting up a distribution network, dealing with international shipments, different technical and other regulations, etc.

firms in the same country, or in a very small number of other countries, but the final product assembled by these firms may be "consumed" in a large number of countries. Even producers of final products may not sell directly to final consumers in all countries. Instead, products could be delivered through wholesalers, and they could be exported to "hubs" in large countries, and then re-exported to smaller countries. Thus, the "value added" produced by a supplier in country may in fact be "consumed" in 100 or more countries, but there is no record in the data that this supplier ever exported anything.[37] Of course, even if there is a lot of indirect and thus unobserved exporting, this does not imply that models assuming a high fixed cost to export are wrong, but one may then re-interpret these costs rather as costs to *export directly* instead of exporting through an intermediary that may use scale economies to reduce the fixed costs to export.

The eBay data gives an insight into trade relationships that are different from offline firm level literature: Since trade on eBay is typically in final products which are sold directly to final consumers, I observe a more complete set of global trade in the sense that there are no "unobserved" indirect exports through other firms.

Another reason why the number of exporters (or export destinations, as I will analyze below) in offline data is very low could be related to the size of the firms and the number of shipments. I will show this for eBay data, and it has been shown for offline firms by Armenter & Koren (2010). The fact that data is sparse, meaning that the number of firms and the number of products they produce or shipments that they make is much smaller than the possible combinations of firms and countries, or firm-country-product combinations, can explain many patterns of offline and online trade.

<u>Export status of offline firms versus eBay sellers</u>

Bernard *et al.* (2007) showed that a remarkably small proportion of US firms engage in international trade: Out of 5.5 million firms operating in the US in 2000, only 4% exported. Bernard *et al.*'s data sample includes *all* US firms–for many of which one would not expect any export activities, such as small-scale retailers. However, even when focusing solely on firms active in manufacturing, mining and agriculture, the authors find that only 15% were exporters. The share of firms exporting varies across sectors and ranges from 5% in the printing industry to 38% for computers and electrical equipment. Similar evidence can be found for other countries.

European firms tend to export more frequently, which is not surprising since their home markets are smaller and foreign markets are close and well integrated. Nevertheless, Mayer *et al.* (2007) show that "only" 28% of UK firms export. Figures are higher for Germany (59%), France (67%), and Italy (74%), although it should be noted that the data samples used to generate the figures for these countries exclude small firms; as a result, figures including all firms are likely to be significantly lower.[38] For example, a more complete study of French firms, which has been widely cited in the trade literature, shows that only 15% of French manufacturers export (Eaton *et al.*, 2009). The French data also reveals that exporting is almost exclusively performed by large firms: only three percent of the smallest 10% of French firms (measured by total sales) export, while 65% of the largest 10% of French firms export (see Table 8 for an overview of the offline data and comparison with eBay).

How do these figures compare with eBay? I find that eBay sellers are different from offline exporters in two respects: A remarkably high share of eBay sellers is engaged in cross-border sales. For

[37] The automobile industry is a good example. While one could find a BMW being sold in almost any country, the records would show that none of the possibly hundreds of firms acting as a supplier in the value chain exports to all of these countries.
[38] See Mayer & Ottaviano, 2007. The European data contains only firms with more than 20 employees (France, Germany) or more than 11 employees (Italy). In the case of France, these are 43% of all firms.

example, out of those US-based sellers that I would define as commercial sellers (annual sales above USD 10'000) I find that 97% export - a share that is significantly higher than what one can find in any study of offline trade. When using data for all eBay sellers covered by the dataset,[39] the share of sellers exporting is smaller (68%), but still remarkably high compared to offline firms, especially when one takes into consideration that this includes many users with negligible sales. For other eBay countries, I find similar patterns, except for Australia, where the share of exporters is high, but consistently lower than for the other four countries.

> ***Empirical result 1:*** *Almost all commercial eBay sellers export, and even significant shares of tiny eBay sellers ship abroad.*

Bernard *et al.* (2007) showed that exporters are larger, more productive, pay higher salaries and employ higher-skilled workers (see their Table 3). I cannot replicate these results since I do not have a lot of data on seller characteristics, such as capital, number of employees or wages. However, my data also shows that exporters are much larger. Taking all US eBay exporters, one can find that exporters are 3.3 times larger (in terms of overall sales). Focusing on "commercial sellers" (those with annual sales above USD 10'000), this exporter premium drops to 47%, and there is no statistically significant exporter premium within the group of very large sellers (with annual sales above USD 100'000). In contrast, Bernard *et al.* found an exporter premium of 148%.

These results can be interpreted in two ways: First of all, it appears that exporting is much easier for eBay sellers. For an eBay seller who is shipping items by mail, fixed costs attached with selling abroad are probably fairly small, especially when compared to offline firms. One may argue that eBay sellers are not *comparable* with offline firms, especially manufacturers, but the fact that almost all of those that one may define as "commercial sellers" can reach foreign customers is remarkable, especially when one considers that these sellers are very small compared to their offline counterparts. Among large sellers on eBay, all but the largest have sales of not more than a few million USD per year. A second – and not necessarily contradicting – explanation for these differences is that I observe all exports on eBay because of the nature of the sales and the data: Sales are (usually) made from a final seller to a final consumer. I can observe all these transactions (and nothing else). Offline firm-level data usually does not show indirect exports from one manufacturer through an intermediary (e.g., an assembler or a wholesaler) to a final consumer in another country. Thus, while the results clearly suggest that there are very low fixed costs to export online, such fixed costs may also be lower offline than the low export shares in offline data suggest.

<u>Are there fixed costs to export on eBay?</u>

How can one test whether there are fixed costs to export on eBay? In the presence of fixed costs, one would expect to observe that smaller sellers (or firms) are exporting less frequently. If all firms were identical, one would expect only sellers above a certain threshold size to be selling abroad. Obviously, sellers vary widely in terms of products they offer, or the individual "fixed costs" they may face to export. We would therefore expect a gradual increase in the share of firms exporting along indicators such as total sales.

This is indeed exactly what one can observe. In Figure 2, I show the average export status for deciles of eBay sellers, and also for French offline firms (which is the only country for which I have found such offline data). We can see that larger sellers (or offline firms) are more likely to export, and also

[39] Note that the US dataset does not cover the smallest eBay sellers (roughly those with less than 25 transactions per year). See also the notes to Figure 3.

that all but the largest French offline firms are less likely to export than eBay sellers, even though they are much larger.[40]

Empirical result 2: *Large sellers or firms are more likely to export – both on eBay and offline.*

However, this alone cannot be a clear indication of fixed costs to export because such a pattern in the data could also be observed if there were no fixed costs to export. The reason is simple: Many eBay sellers (but also "offline" firms) are so small that they may not export simply *by chance*. I am therefore using an approach similar to Armenter & Koren (2010) to test whether what we observe in the data can be interpreted as an indication for fixed costs to export, or whether it is simply what one would expect to see even without any fixed costs.

Let us assume that there are no fixed costs to export, i.e., shipping an individual transaction abroad is as easy as shipping it to a seller in the same country. The probability that an individual purchase is made from a foreign buyer should then equal the share of demand from foreign buyers. As in Armenter & Koren's (2010) "balls-and-bins" model, transactions can be thought of as balls that are randomly thrown into bins, and the width of bins is proportional to the probability that a bin is hit. Since for the moment I am only interested in whether a seller is exporting, there are essentially only two "bins" – domestic and foreign consumers.

This is obviously only a thought experiment. Doing this for offline firms may seem odd because offline exporters certainly do not randomly ship items (i.e., throwing balls) to certain countries or consumers. Instead, they may advertise in certain markets, or otherwise actively recruit customers. Their transactions are also more likely to be correlated with each other – the same customer may purchase several times from the same buyer.

It does however seem a more plausible model for eBay because eBay sellers are not necessarily actively marketing items to certain consumers. In theory, they may do nothing but uploading their items to the eBay platform and then wait for somebody to buy.[41] As long as they offer international shipment, the buyer could theoretically come from anywhere. For eBay, one should intuitively think of buyers picking sellers randomly – the buyers are "throwing balls" at them. This is shown schematically in Figure 1. It shows a number of buyers and sellers, each from four different countries. Each "ball" is a purchase, and the number of balls per buyer country is assumed to be exogenous. If all buyers make random purchases, then the seller (and thus country) from which a transaction is made will only depend on the width of a seller's bin, which is proportional to the number of items that each seller is listing on eBay. In the US data (the same applies to data for the four other countries), we only observe sales by US-based sellers. What we observe is thus the balls thrown on US sellers, which will be only a certain share of all transactions, and equal to the US' share in eBay listings. I do not know what this share is, but this does not matter as long as all transactions are assumed to be independent of each other. The distribution of purchases made from the US should therefore follow the same allocation across buyers and buyer countries as the sales from any other selling country, or the aggregate sales from all countries.

Focusing for now on the export status of a firm, I assume that there is a certain probability p_{exp} that an individual transaction will be shipped abroad – provided that the seller allows for international

[40] Note that the differences between eBay countries are mainly driven by different coverage of the data. The AU and FR dataset covers all sellers, including the smallest ones. The US dataset also covers some of the smallest sellers, whereas the DE and UK dataset covers very few sellers with less than 25 transactions.
[41] Of course, in practice an eBay seller could also do location-specific marketing, e.g., through Google Ads.

shipment. With t independent transactions, the expected export status for a seller can then be simply calculated as follows:

$$p(\text{export\_}status = 1) = 1 - (1 - p_{exp})^t$$

With

$$\lim_{t \to \infty} E(\text{export}_{status} = 1\,) = 1$$

With transactions assumed to be independent, the probability that a transaction is exported can be taken from the overall share of items that are exported.[42] If 20% of eBay sales from a country were shipped abroad (as a share of transactions), the probability that a seller with t sales does *not* export would thus simply be 0.8^t, and hence $p(\text{export\_}status = 1) = 1 - 0.8.^t$.

I can now calculate the expected export status for a seller with a given number of transactions and compare this with what we observe in the actual data. We should expect to see that many small sellers don't export, but almost all large sellers should export.

Large non-exporters (by number of transactions) on eBay are indeed rare: For example, 96% of sellers with more than 100 transactions are exporting – but we should expect almost 100% (99.9999105%). There are only two US sellers with more than 10'000 transactions in 2010 that did not ship abroad. But they are not as rare as one would expect given the "chance" of an item shipped abroad – the chance that 10'000 sales are all domestic would be almost zero and it would be extremely unlikely that one would observe this in a sample of less than one million sellers. However, for relatively small sellers (i.e., those with few transactions), we observe fewer exporters than one would expect from the "balls-and-bins" model.

Overall, I find that 68% of US eBay sellers export,[43] whereas the random simulation suggests that we should expect 89% of sellers exporting. Figure 3 shows these figures for all five countries, as well as for US offline firms (manufacturers only). Data for the latter is taken from simulations by Armenter & Koren (2010).

One should first note that the actual share of exporters and also the share of items exported vary across eBay countries. The main reason for these differences lies in data coverage. As I have explained above (see footnote 24), only the Australian and French datasets cover all sellers, including all of the smallest ones. The lower share of exporters for Australia and France should therefore not be seen as an indicator that eBay sellers in these countries find it more difficult to export. What we see is that for all countries, the actual share of sellers exporting is lower than the share we would expect from simulating the "balls-and-bins" model. Figure 4 shows this in more detail. It shows the actual average export status by number of transactions for all five countries for which I have data. One can see that the share of exporters is lower than one would expect. The overall share of transactions exported is 6-13% across the five countries covered in the data, whereas the data for small sellers would be more

[42] As in Armenter & Koren, I derive the export probability from the overall share of items exported. Of course, if there are fixed costs to export, then the overall share of items exported will be biased downwards because the observed export probability is a weighted average of export shares of large sellers (who have overcome the fixed costs to exports and thus face the "real" share of foreign demand) and small sellers who do not export. This means that in the presence of fixed costs to export, one is using an export probability that is too small. The balls-and-bins model may thus under-predict the share of firms that should export under the assumption of no fixed costs. However, at least for eBay, the share of transactions exported is not much larger for large sellers than for the aggregate, mainly because large sellers make up for most of the sales.

[43] See the notes to Figure 3, explaining the differences between this number and some figures directly obtained from eBay.

closely replicated with an export probability of individual transactions of only about 1-3%. To illustrate this, the chart also shows the expected average export status for export probabilities 1% and 3%. However, for large numbers of transactions (> 1'000), almost all eBay sellers in all five countries export, as one would expect (see Figure 5). One exception is Australia, where we can observe a fair number of very large sellers that do not export at all.

> ***Empirical result 3:*** *Most eBay sellers with large numbers of transactions export. For sellers with few transactions, the number of exporters is much lower than predicted by the "balls-and-bins" model.*

How can one explain these differences with the "balls-and-bins" model? Are these differences between "expected" and actual exporter shares an indication that we need to reject the assumption that exporting on eBay entails no fixed costs? Not necessarily, because there are some possible explanations for this other than fixed costs:

- Some sellers may sell products for which there is no demand from abroad because they are too bulky or for other reasons, and this may be particularly the case for small sellers. In the US, the product category with the highest export share (by transaction) is stamps, and the one with the lowest is "home and garden" items. However, even for stamps I find that the share of exporters within those with few transactions is smaller than what one would expect if the export decision was purely random. I have repeated the simulation results for each product category, and while there are categories in which the export status resembles more the one we would expect if sales were purely random, there are very few sectors in which the actual share of sellers exporting is as high as or higher than the expected one. Note however that some items may simply have to be sold locally because it makes no sense to ship it (e.g., an old sofa). Unfortunately, the data does not allow to quantify this. I can observe that certain items are not exported, but I can not observe the buyer and seller location within the US. Whether there is an "export barrier" for such countries, or simply a "distance barrier" remains open. Ideally, one had data for eBay shipments at ZIP code level to verify this. Hillberry and Hummels (2007) showed with such data for US sales by "offline firms" that there is indeed a huge "home bias" for sales, but they showed that this is due to distance, and not due to borders (in their case state borders). However, if one aggregated such data at state level, one could draw the wrong conclusion that US state borders act as trade barriers. Similarly, one may find on eBay that items that are hardly exported are also not sold across a large distance within the US.[44]
- Small sellers (in terms of transactions sold) export smaller shares of their sales. The largest 10-20% of sellers export higher shares.[45] One may therefore assume that the likelihood of an item of a small seller being picked from a foreign buyer may be smaller than the overall average, which can explain the deviation from the "balls-and-bins" model. As I have stated above, this may be due to the type of products they sell. But another reason could be that there are indeed fixed costs for sellers to overcome. One way to *quantify* the fixed costs to export is to assume that an eBay seller – on average - only exports once she can expect to make a certain number of export transactions. Armenter & Koren (2010) propose the following: They define all firms with an expected number of export transactions being below a certain threshold k^* to be non-exporters. Intuitively, this means a firm is only expected to export once

[44] This could also be very relevant for the interpretation of "offline" data. It may well be that many of the non-exporters that one observes in firm-level data only sell products across short distances. What prevents them from exporting is thus not the country border, but the distance to the border.

[45] Because they account for most of the sales, their export share is close to the overall share of transactions exported (see footnote 42, above).

it can expect to reach a certain number of export transactions. On eBay, sellers would only offer international shipment after they have sold a number of items domestically. For US offline exporters, Armenter & Koren found that only for a threshold value of $k^* = 24$ will the resulting expected share of firms exporting be the same as the actual share. With an estimated[46] average value per transaction of USD 36'000, this means that US offline firms are only expected to export ones the expected export value reaches USD 850'000, which they interpret as an indicator that there are substantial export barriers in traditional offline trade. I replicated this for eBay, and I find much lower threshold values. The expected share of sellers exporting falls to the actual share for values of k^* between 1 and 5. The numbers vary across countries because of the different data coverage for small sellers. If I only include sellers with at least 25 transactions ("regular sellers"), for which I have an almost complete coverage in all five countries, the estimated value for k^* is between 3 and 5, i.e., much lower than for offline trade. In absolute terms, the difference is even more pronounced because eBay transactions are tiny compared to offline transactions. The average transaction value on eBay is fairly small (below USD 100), as compared to the (estimated) average transaction value in US offline data (USD 36'000).[47] In other words, while Armenter & Koren's estimate suggests that only exports of USD 850'000 make offline firms overcome the fixed costs to export, the equivalent figure for eBay is somewhere around USD 100-200.[48]

- Another explanation would be that transactions are not independent because there are matching costs between buyers and sellers, which can be linked to both search costs and risks (see Section 3 above). Buyers therefore make repeated orders from the same seller - not because they pick the seller again "by accident", but because they are satisfied with her products and delivery service and thus the perceived risk to order from her is lower. The buyer may also have incurred some search costs to find the right seller for the products he is interested in, which also explains repeated purchases.[49] Hence, "balls" can be serially correlated, which violates one of the key assumptions of the balls-and-bins model. However, to explain the export status pattern that we see in Figure 4, we would have to assume that on average three to four transactions are linked together, which does not seem plausible – at least not for eBay trade.[50]

---

[46] The offline data that Armenter & Koren use does not include the actual number of transactions, which therefore has to approximate by dividing total US exports by the (known) total number of US export transactions. This does not account for the fact that the size of transactions likely varies widely across firms.

[47] Again, one should note that "offline" transactions are usually transactions at wholesale level, with many items shipped at ones. In contrast, eBay transactions are small transactions at retail level, with one item (or a small number) shipped at a time.

[48] I have replicated this for Australia by taking all transactions of sellers into account since they started using eBay, rather than taking transactions of all sellers for one year. Results are very similar. (Note that Australian data includes all sellers, including the smallest once. Data for US, Germany and UK is incomplete for the smallest sellers).

[49] Another reason is more technical: If a buyer orders several items at the same time to save on shipping fees (eBay sellers may charge a flat shipping fee, independent of the number of items, or offer free shipping once the order exceeds some threshold value), then this transaction would appear in the data as several separate transactions if – and only if – the items are from different product categories. If they are from the same category, then I should observe them as one transaction with two items. On average, around 1.1 items are shipped per transaction in this dataset. Unfortunately, the data does not allow to quantify the extent to which buyers make repeated purchases from the same seller.

[50] For example, if each buyer would purchase bundles of three products, then a seller with three transactions has in fact only one buyer. Therefore, the chance that he or she exports is only 0.13, and not $1\text{-}0.87^3 \approx 0.35$. In reality, significantly less US eBay sellers with three transactions export. While some transactions are certainly done in bundles, it is not plausible to believe that this occurs regularly enough to fully explain the difference in actual and expected exporting patterns.

Armenter & Koren (2010) find a somewhat similar result for offline data. While their "balls-and-bins" simulations perform very well in explaining a range of pattern of offline trade, it does *not* when it comes to *export status*. Using the approach I described above, and assuming a 14% probability that an individual shipment is exported (which equals the share of US manufacturing production that is exported), they find that 74% of US firms "should" export, while in reality it is only 18% of manufacturers. On eBay we thus find a much smaller difference between the expected and actual share of firms exporting (see again Figure 3). A direct online-offline comparison for the US shows that the overall share of transactions exported is similar for US eBay sellers and US offline manufacturers, and so is the overall expected share of firms that "should" export (89% and 74%). However, the actual share of firms exporting is 68% for eBay sellers and only 18% offline.

One drawback of the offline dataset used by Armenter & Koren is that it does not include the number of individual transactions; therefore they have to make assumptions on the average transaction size, which has a large impact on the results, but even when assuming fairly small transaction sizes (i.e., more balls, and thus a larger expected share of firms exporting) does not bring the expected share of exporters close to the actual share.

How can one interpret these results? First of all, the remarkable share of exporters on eBay, as compared to offline trade, is a clear indication that there are much smaller fixed costs to export. Using Armenter & Koren's method to simulate shares of exporters reveals that exporting is not purely "random", but assuming only very modest fixed costs to export shows that the actually observed export patterns on eBay can be fairly well replicated by the "balls-and-bins" model.

This is not to say that costs to export on eBay are negligible. There is anecdotal evidence that eBay sellers face problems when exporting and some simply choose not to ship products to foreign destinations. While many such problems are specific to certain destination markets (e.g., sellers do not ship to particular countries because of assumed high risk of packages getting lost, or fraud), there are also eBay sellers who choose not to ship any item to foreign markets. To the extent that this is related to actual or perceived trade barriers, this could be addressed by policymakers. However, the same instruments that are used to boost "offline trade" may not be working equally well for eBay sellers. One example is the case of free-trade agreements, which are typically tailored for, and influenced by traditional offline exporters. Some of their benefits, in particular lower tariffs, may not be as helpful for small-scale transactions, and new measures may be needed for them (such as streamlining import procedures or making a claim for tariff preferences easier for small shipments[51]).

7 eBay exporters and gravity

I will now analyze export patterns of eBay exporters only. Lendle *et al.* (2012) have shown in a previous paper – based on a dataset of bilateral eBay trade between 61 countries – that distance still matters for eBay, but about 65% less when compared to traditional offline trade for the same countries and similar products.[52] I will first replicate this analysis with the new seller-level data. While this data covers only exports from five countries, it includes all destination countries, whereas our previous dataset only included trade among 61 countries. Table 4 shows the results of a simple gravity equation for eBay exports and "offline" exports.[53] Using two different specifications, I find that distance matters

[51] See Keck & Lendle (2012) for an analysis of costs to utilize preferences, which provides evidence that such costs are mostly fixed costs, thus making preferences less beneficial for small exporters.
[52] See Lendle *et al.* (2012).
[53] Regressions were made using importer-year and exporter-year fixed effects, which is now a widely accepted approach in the gravity literature, especially when the share of zeroes in the data is small (see, e.g., Head &

much less on eBay – the distance coefficient is roughly 50% smaller. I find that a common language (whether as a common official language or a common spoken language) and a common legal origin increase eBay exports, while they have no statistically significant effect on traditional exports. I also find that RTAs have no statistically significant effect on eBay trade flows, which is what we would expect, given that RTAs are unlikely to offer significant benefits for very small transactions. Somewhat surprisingly, their effect on offline exports is even negative and statistically significant.[54]

> ***Empirical result 4:*** *eBay sales are driven by the usual gravity variables. We find that distance matters much less for eBay exports, as compared to traditional offline exports.*

The dataset allows decomposing exports into the extensive margin (number of exporters) and the intensive margin (exports per seller). The latter could again be decomposed into "number of different products exported per seller" and "exports per product". This allows to verify whether the distance effect on eBay is driven by the extensive margin (few sellers export to distant markets), or by the intensive margin (sellers export less to distant markets). While I cannot replicate this for offline exports (except for Australia, where such data is publicly available), other authors have done so using detailed offline firm-level datasets, including most recently some authors from the World Bank (Cebeci *et al.*, 2012), with a new dataset covering 45 countries.[55] We will see that, as for offline trade, it is mostly the extensive margin that explains the remaining distance effect on eBay.

The main results are shown in Table 5. I show results both without and with importer and exporter fixed effects (as done in Cebeci *et al.*, 2012), and both using only GDP and distance, and then also additional gravity variables as control variables.[56] For all specifications, I find that the distance effect is almost completely driven by the extensive margin. For example, looking at the specification with importer-year and exporter-year fixed effects and additional gravity controls (see columns B4-B6 in Table 5), I find a distance coefficient of -0.8, which is significant at the 1% level. Decomposed into the extensive and intensive margin, 88% of this effect comes from the extensive margin. Similarly, the effect of importer GDP also mainly comes from the extensive margin (see columns B1-B3). Thus, larger and less distant countries (in terms of GDP) import more, but mostly "from more sellers", rather than "more from each seller". I have also replicated these results using transactions, rather than export values. This leads to almost identical results (not shown).

Columns A7 to A12 in Table 5 show such results for offline exports in 45 countries, which are taken from Cebeci *et al.* (2012) and based on a new World Bank exporter dataset. While none of the five countries for which I have eBay data is covered by their dataset, it appears sensible to assume that

---

Mayer, 2013). Note that importer and exporter fixed effects are allowed to vary between eBay and offline flows. The data used is 2006-2010 for US exports, 2006-2011 for DE, FR and UK exports and 2006-2012 for AU exports. I show results for both overall "offline exports" as well as for a subset of offline exports in products similar to the type of products traded on eBay only (see Lendle *et al.* (2012) for a list of HS codes of these products).

[54] These results are robust to many other specifications of the gravity regression, including the result on RTAs. There is only a very small share of "zeroes" in the data, and including them (i.e., replacing their logs by zero) leaves results practically unchanged. Note that I always drop destination countries to which there are either no eBay exports from any country or no offline exports from any country, the reason being that some country classifications are only used on eBay or only offline (e.g., Martinique). The results are also very similar when using the product dimension of the data and product fixed effects.

[55] This dataset is based on detailed customs transaction data with a firm identifier.

[56] Note that the results for overall eBay exports slightly differ from those shown in Table 4 because I dropped observations for which there are no eBay exports. This only has a minimal impact on the regression results. All results are obtained using OLS. Therefore, the sum of coefficients for the extensive and intensive margin is equal to the coefficient for overall trade.

offline data for the five countries would be rather similar.[57] One can see that the distance coefficient is much larger (in absolute terms) than for the five eBay countries included in the dataset (compare columns A4 and A10). Obviously, since the results are based on different groups of countries, this result should be interpreted with some caution. More interestingly, the decomposition of the distance effect shows that around 75% of the distance effect is due to the extensive margin. This is similar to what I found for eBay exporters. Results for the effect of importer GDP are also similar, although here the effect of the intensive margin is more pronounced (see columns A8 and A9).

For Australia, I am able to replicate the gravity regression using data from the Australian Bureau of Statistics (ABS, 2012). That data includes, for each export destination, the number of exporting firms, the number of transactions and total export value. This allows me to use two different approaches: I can use the export value and split it into the extensive margin (number of exporters) and intensive margin (export value per exporter). I can also measure total exports in "transactions". The intensive margin is then the number of transactions per exporter. Results are shown in Table 7.[58]

Again, I find that the distance coefficient is much smaller on eBay than offline. This is the case whether one measures trade in value or in transactions. The intensive margin matters relatively more for offline sellers than on eBay (compare, for example, results in columns B1 & B3 with those in B7 & B9). The size of the importing market also mainly affects exports through the extensive margin, especially for eBay exports. Most "traditional" gravity variables don't have a statistically significant effect on Australia's exports (whether on eBay or offline), with the main exception being a common currency, which has a large effect on both eBay and offline exports.[59]

Bernard *et al.* (2007) made a similar finding for US-based exporters. They show that the overall effect of distance mainly comes from the effect of distance on the number of exporting firms, whereas distance has only a small and barely significant effect on the products exported per firm, and no effect on the exports per product per firm (see the discussion below and Table 6).

> ***Empirical result 5:*** *The trade effect of distance and GDP is mainly due to the extensive margin. Smaller and more distant countries import from fewer exporters, rather than less from each exporter. This result holds for eBay as well as for traditional offline exporters in a wide range of countries.*

How can these results be explained? Cebeci *et al.* (2012) assume that the effect can be thought of as being related to fixed costs to export to a market.[60] If geographic distance is positively correlated to such fixed costs, then fewer firms will be able to overcome these fixed entry costs in distant markets. In other words, a reduction in distance means lower entry costs, and thus more firms export.

I believe that a different explanation applies for eBay sellers. First of all, it is difficult to think of destination-specific fixed costs *on eBay*. As explained in Section 3, eBay sellers do not need to set up a particular relationship with a destination country (such as opening an office, finding a distributor, or actively seeking a wholesaler that may import their products). They will have to verify some country-

---

[57] Many of the findings from the World Bank dataset are very similar among the countries covered, even though the data covers a wide range of small and large, developing and developed countries.

[58] Note that I dropped destinations to which there are either no eBay exports or no offline exports, as well as countries for which the number of offline exporters was not provided in detail (usually small countries and territories).

[59] Again, such results need to be interpreted cautiously, given that these regressions are based on very few observations. Only two tiny countries (Kiribati and Tuvalu) use the Australian dollar, as well as Nauru, which has been dropped because there is no GDP data available.

[60] See also Head & Mayer (2013), who confirm that similar findings have been made by many authors and also argue that this finding is in line with the assumption of fixed costs.

specific legislation, shipping and customs issues, but by and large one can assume that such hurdles can be overcome. There may be some barriers to export at all – as I showed in the previous section, although to a much smaller degree than "offline" – but once an eBay seller is willing to ship abroad, doing so to a new destination should not be very difficult. Why does then the extensive margin matter so much?

Let me first rephrase the results from the gravity equation: When controlling for GDP and other variables, I find that countries that are *closer* import *more*. This increase in eBay exports is mainly due to more eBay sellers exporting, and less due to existing exporters exporting more. Let's now take Australia as an example.[61] In 2011, over 100,000 Australian eBay sellers shipped abroad, and to about 180 destination countries.[62] Now, in all but three of these markets was the number of export transactions below the number of exporters. Half of the destination markets – many of which are very small countries – imported fewer than 300 transactions.[63] By definition, the number of sellers exporting to a market has to be equal or lower than the number of export transactions. In other words, demand from most markets is too low to allow each Australian seller to export. Therefore, the fact that all exporters are not exporting everywhere has not necessarily anything to do with fixed costs to enter a particular market, but could simply be due to the small size of most markets and the fact that trade data is sparse.

> ***Empirical result 6:*** *In almost all destination markets, the number of purchases from eBay exporters is much smaller than the overall number of eBay exporters from any of the five exporting markets.*

Assume now that import demand from a country increases. Alternatively, one may assume that "distance", or rather trade frictions that are related to distance, become smaller. Will that increase the number of sellers exporting to that market, or rather the exports by existing exporters? If buyers pick products randomly from the items offered on eBay, and if initial demand from the country is small, and thus only a tiny fraction of exporters serve the market, then an increase in demand is more likely to "hit" sellers that previously did not export to that market. On the other hand, an increase in demand from very large countries – to which many sellers already export – should rather affect the intensive margin. I can illustrate this with the balls-and-bins model. Let's assume country X (the exporter) had 100 equal firms.[64] Consumers in country Y (the importer) are making *t* purchases by picking randomly any exporter in country X. As long as buyers randomly pick items, independent of the seller, then one would observe a certain distribution of sales that one can compare with what we observe on eBay.

In my example, if consumers in country Y make *t* purchases, they will "hit" a certain number of exporters. If all purchases are random and independent of each other, then one can easily calculate the expected number of sellers that are "hit":

$$E(n\_exporters) = 100 * \left(1 - \left(1 - \frac{1}{100}\right)^t\right)$$

$$E(txn\_per\_exporter) = t/E(n\_exporters)$$

[61] The advantage of the Australian data is that it includes all eBay sellers, even the smallest ones. However, what follows can equally be shown for any of the other four countries.
[62] This does not include some small destinations for which I do not have GDP or distance data.
[63] In value terms, the median is about 20'000 USD.
[64] If firms were not of equal size (as it is the case in reality), then the intensive margin would matter more. The reason is that it is more likely that the same exporter is picked repeatedly. For example, if there was one exporter that would offer 90% of all products, whereas the other 99 sellers would together account for the remaining 10%, then any increase in demand is more likely to be met by that large seller.

For t going towards infinity, all sellers can be expected to export and the expected transactions per exporter will increase linearly with t:

$$\lim_{t\to\infty} E(n_{exporters}) = 100 \quad \text{and} \quad \lim_{t\to\infty} E(txn\_per\_exporter) = t/100$$

Obviously, both the expected and the actual number of exporters to country Y cannot be higher than the number of transactions:

$$E(n_{exporters}) \;,\; n_{exporters} \leq t$$

In Figure 6, I show this graphically. For small t, an increase in t increases trade through the extensive margin. There are many sellers that do not export yet, and it will be more likely that these sellers will be picked, rather than the few existing exporters to Y being picked a second time. Exports per seller (the intensive margin) remain flat. As t increases, more and more sellers are already exporting and additional exports will rather be made through the intensive margin. Once all sellers export, any increase in exports by definition needs to be made through an increase in the intensive margin.

> ***Proposition 1:*** *If purchases are made randomly and their number t is low compared to the number of possible sellers, then the balls-and-bins model predicts that for low numbers of t, an increase will affect sales through the extensive margin (number of sellers), whereas for high number of t an increase will affect sales through the intensive margin (sales per seller).*

If this scenario was a realistic illustration of the real world – which it certainly is only to some degree – then trade should mainly increase through the extensive margin when the number of trade flows is small, and mainly (or at least more so) through the intensive margin when the number is large, with “small” and “large” referring to whether the number of transactions made with a country is “small” or “large” compared to the number of exporters. If the number of export transactions is negatively correlated with distance (for whatever reason), then this distance coefficient will be partly “explained” by the extensive and the intensive margin, even if there are no fixed costs to export to a particular market. Figure 7 shows the same as Figure 6, but with the *actual* numbers for Australian eBay exporters.[65] It reveals a striking similarity with the very simple simulation results from which Figure 6 is derived.[66] This illustrates two things: First of all, the assumption that sellers are picked randomly by buyers may be a close approximation of what is happening on eBay. Secondly, it shows that one should expect the extensive margin to matter more than the intensive margin: Less distant countries import more from (Australian) eBay sellers, but that means they rather “pick” additional sellers and don’t necessarily buy more from the same sellers. In contrast, we only see a small increase in the sales per seller as trade increases. Now, Figure 8 shows the same for Australian offline exporters. As imports increase, the number of exporters increases, but not one-to-one. Instead, the exports per exporter also increase. This illustrates why we can see that the intensive margin matters relatively more for offline exporters than for eBay exporters (as shown above and in Table 7). Nevertheless, it is still mostly the extensive margin that drives Australia’s exports. For most destination countries (all but 15), the number of transactions is smaller than the number of Australian offline exporters (which is 38’555). This means that demand is too low to allow everybody to export.

> ***Empirical result 7:*** *Actual data for both eBay and offline data qualitatively resembles the pattern predicted by a “random export” model. Trade data is sparse - the number of transactions, compared to the number of sellers, is low, both on eBay and offline. This explains why in a gravity equation the extensive margin matters more.*

[65] Here, each observation stands for one importing country with a given (observed) number of transactions, an observed number of Australian sellers exporting there and the exports per seller.

[66] Figure 7 looks qualitatively identical for the other four countries for which I have eBay data (not shown).

Can we find a different effect of the extensive and intensive margin for small and large destination countries empirically? One would expect that for a subset of very large destination countries, an increase in exports happens rather through the intensive margin because most sellers already export to such markets. I have replicated the results from Table 5 (columns A4-A6) by quartiles of destination countries (by export value). I found that the intensive margin is only statistically significant for the 4th quartile (the largest destination markets), as one would expect.

Similarly, one would expect that the intensive margin matters more when I only focus at large sellers. Such sellers export to many different markets. The distance and GDP effect on exports should thus be explained more by the intensive margin. I do indeed find this effect. For this analysis, I have split eBay trade flows into those conducted by four sub-groups of sellers based on the number of annual export transactions: < 25; 25 to 99; 100 to 999, > 1'000. The results for the distance coefficient are illustrated in Figure 9. I find that for small sellers, the intensive margin matters less than for large sellers (although the extensive margin still matters more even for them). The overall trade-reducing effect of distance also shrinks for larger sellers.[67]

> ***Empirical result 8:*** *The extensive margin matters less among large export destinations or among large eBay sellers because a significant share of seller-destination combinations is already non-zero. More transactions will therefore rather increase the sales per seller.*

To what extent this explanation is applicable to offline trade is debatable because the assumptions behind the balls-and-bins model are not very realistic for such trade. But certainly empirical results for the relative importance of the extensive margin are highly affected by the type of data used. If one uses very sparse datasets (e.g., exports of all firms to all countries in one month) or less sparse datasets (e.g., exports of large firms to large countries aggregated across 10 years) will likely have very significant effects on the results.[68]

The data also allows me to compare US eBay firm-level exports with the analysis conducted by Bernard *et al.* (2007). They applied a gravity equation on US firm-level exports, which they split into three components: the number of firms exporting to a destination (as I did above), the number of products exported to that destination (a second extensive margin), and average exports per product per firm (intensive margin).[69]

Table 6 replicates Table 6 of Bernard *et al.* (2007). Results for overall export value (columns (A1) for offline and (A5) for eBay) are somewhat similar: The coefficients for importer GDP are fairly similar, whereas the distance coefficient is larger on eBay.[70] Bernard *et al.* also found that both extensive

[67] An extreme scenario would be to only consider sellers that make a single export transaction. In a gravity regression using exports measured in transactions, only the extensive margin would matter.

[68] I have tested that with French offline export data at the product level. Using data across 10 years, the distance coefficient is much less driven by the extensive margin than when using data for 1 year only. A similar effect can be found on eBay when I aggregate the data across several years.

[69] Note that Bernard *et al.* use data at HS-10 level, with around 10'000 different products. The data is at eBay's broad product category level, with only around 37 different products (7 of which only appear in very few observations in the US data). However, even though there are fewer products on eBay, US online sellers sell on average 5.5 different product categories (data from 2010), whereas Bernard *et al.*'s data (Table 4) suggests that the number for US exporters is similar - around 3 to 4 different products. When looking only at commercial eBay sellers (those with annual sales above USD 10'000), I find that they sell on average 7.4 product categories – similar to or even more than offline exporters even though there exists only a fraction of possible product categories (note that these figures are not necessarily representative for the US eBay market as a whole because the data for small sellers is incomplete).

[70] This is at odds with the overall result that distance matters less on eBay, but one should note that Bernard *et al.*'s results are most likely not based on the same set of destination countries. The distance coefficients are more

margins decrease with distance: The more distant an export destination is, the less firms export and the number of products per firm also declines. Results using eBay data are very different. The overall effect of distance mainly comes from the effect of distance on the number of exporting firms, whereas distance has only a small and barely significant effect on the products per firm, and no effect on the intensive margin (exports per product per firm). The overall effect of exports per exporter (column A9 in Table 6) is positive, but small and not significant.

Columns (A4) and (A8) show the intensive margin of trade, the exports per firm for each product category. For offline exports, distant markets see fewer products per firm, but higher sales per product per firm. Bernard *et al.* interpret this as a sign that firms sell higher-valued items to more distant markets because this makes it easier to overcome higher export costs, assuming that export costs depend on weight or volume, rather than value. One may also interpret this as an indication of higher fixed costs to enter distant markets with additional products. Therefore, one enters such markets only with fewer products and sells more of those. There is also such an effect on eBay, but the coefficient is much smaller.

I do not find a plausible explanation why a higher GDP should lead to a lower intensive margin, as one can see for offline flows. It may be that there are product-specific fixed costs, which are easier to overcome in large markets. Sellers would therefore ship a larger variety of products to large markets (see result A3), but that then includes items with lower sales – and thus the average sales per item declines. On eBay, I do not find such an effect, which one may interpret as an indication that there are no such product-specific fixed costs on eBay. The positive, but small effect of GDP on the intensive margin would be fully in line with the balls-and-bins model. As demand (GDP) increases, it is mainly the number of sellers that increases, but some sellers will also receive multiple purchases, which results also in them selling multiple products. The effect of GDP on the number of products exported is very small, which is not surprising because products on eBay are aggregated into only 37 categories, and most sellers only offer a small number of them. Once a seller has made a few sales to a country, the number of different product categories sold cannot increase anymore. For offline firms, the effect of GDP on the number of exported products per firm is much larger, which is not surprising because the offline data is based on around 10’000 different products – with even fairly large countries only importing some of them.

To conclude, this section has confirmed the previous findings of Lendle *et al.* (2012) that eBay trade is less affected by geographic distance for a dataset covering only five exporting markets, but all importing markets. Further, I find that eBay trade, very much like offline trade, is mainly driven by the extensive margin, i.e., the number of sellers. I could show that this can mainly be explained by the fact that data is sparse and that there are many small sellers and many small markets – making it highly unlikely that smaller markets are served by a large fraction of sellers. The same conclusion may not apply to offline trade, where we do not assume that sales are made randomly and where market-specific (or market-product specific) entry costs are more likely to exist. However, the sparsity of the data may explain at least partly why fewer offline exporters sell to smaller and distant countries.

In the next section, I will analyze in more detail the extensive margin of eBay exporters and provide further evidence that market-specific entry costs are indeed very low on eBay.

similar when running the same regression with contemporary US offline exports and across the same destinations.

8 eBay exporters and market entry costs

I will now focus on another well-established feature of the (offline) firm-level literature – the number of export destinations that exporters are reaching. Not only do I find that few offline firms export – there is also clear evidence that firms usually export to a few selected markets only. Bernard *et al.* (2007) show that out of those US firms that export (as stated above, a fraction of about 4% of all firms), 64% export to a single country. Only 14% of exporting firms sell to five or more countries, and the average number of destination markets is 3.5.[71] Similar evidence has been shown for French firms. Although French firms tend to export to more countries than US firms – which is intuitive, as France is a smaller and relatively open economy – Mayer *et al.* (2007, Table A1) find that 43% of French firms sell to a single country and only 15% sell to ten or more countries.

A new dataset by the World Bank gives the number of export destinations for offline firms based in 44 different countries.[72] While none of the five countries for which I have eBay data is included, the dataset covers a wide range of different countries and should be representative. Figure 10 plots the average number of destination for these countries against the average export value per firm, and also shows these figures for the five "eBay countries" for different subsets of eBay sellers. One can see that eBay sellers are much smaller (in terms of average export value), but they reach many more different countries. For example, US eBay sellers with annual exports of at least USD 10'000 ship on average to 39 different destination countries. That number is lower when I include all eBay exporters (9 destinations on average) and a relatively large fraction of sellers – 25% for the US – sell to a single foreign market only. However, this is somewhat misleading. Many eBay sellers are very small and may have only 1-2 export transactions. Such sellers obviously cannot sell to more different markets than they have export transactions. When I focus on "commercial sellers" (those with at least USD 10'000 annual sales, including domestic sales), I find that only 5% are "single-country" exporters, and 61% sell to ten or more foreign countries. If I consider only such sellers that had at least 10 export transactions in 2010, I find even fewer single-country exporters (1%). Out of those US sellers that had 100 export transactions in 2010, I find on average 21 different destinations. Only 0.2% of such exporters are single-country exporters, while 90% sell to 10 or more different countries.

For the other countries, I get similar figures. The average number of different export destinations among all eBay exporters ranges from 3 to 9 across the five countries, but this figure is not easily comparable across countries because of the different coverage of small sellers in the data. If I only focus on exporters with at least USD 1'000 in annual export sales, the average number of foreign destination is between 7 and 21 (see again Figure 10).

> ***Empirical result 9:*** *eBay exporters reach many more destination markets than offline exporters, despite the fact that their aggregate exports are considerably smaller.*

The offline data for US and France shows that multi-country exporters, while few in number, are much larger. This clear correlation between size and number of export destinations that we observe offline and online makes Bernard *et al.* (2007) argue that US multi-country-exporters are "large in part because they ship many products to many destinations". I argue that – at least on eBay – the causality might be the other way round: Sellers export to many destinations because they are large. Small firms

[71] The average number is taken from Armenter & Koren (2010), based on work from Bernard *et al.* It may not refer to exactly the same set of firms as the other figures, but from Bernard *et al.* (2007, Table 4) one can calculate that the average number is likely somewhere between 2 and 5.

[72] The World Bank data has some gaps for some indicators. For example, the average number of export destinations is only available for 44 out of 45 countries.

cannot do so simply because they are too small. This could be related to fixed costs per export destination, but not necessarily so.

Similar to what I argued above about fixed costs to export, one could well assume that there are no (or almost no) fixed costs for eBay sellers to export to a specific country. As I have shown above, it appears that eBay sellers face some, albeit low, fixed costs in order to export (finding out about different mailing procedures, etc.). Once a seller is willing to ship to other countries, it is hard to imagine any substantial additional cost to export to one more country. In fact, what happens oftenis that a seller would offer an item, and if she offers international shipment, then it would be available to buyers in all countries.[73] The seller then waits for a buyer, without undertaking any destination-specific actions. Offline, this can be different. Some offline firms may actively have to look for customers in a specific market or even set up a local distribution network, and may thus have high entry costs per export market.

> ***Proposition 2:*** *Entry costs into a particular market for an eBay seller should be very low or non-existent. In contrast, such costs can be high for offline firms.*

Nevertheless, we see many zeroes in exporter-destination combinations on eBay. Most sellers ship to only some of the roughly 200 markets available, and almost no seller ships to more than 150 markets. Does that indicate fixed costs? To see why it does not necessarily so, consider a somewhat different example: Assume one had data for US-based eBay sellers' domestic transactions by destination defined as ZIP code area (of which there are more than 40'000 in the US). Now, most sellers would only sell to a tiny fraction of those, but obviously there are no fixed costs to "enter" a ZIP code.[74]

The "zeroes" alone will therefore not tell us whether there are destination-specific fixed costs on eBay. One way to test the hypothesis is to compare the export patterns of eBay sellers with the export patterns we would expect in a world without destination-specific fixed costs, in which exports are purely random. This approach is again similar to Armenter & Koren's (2010) balls-and-bins model, in which they simulate US firm-level exports at product-destination level and find that these "random" export patterns are similar to actually observed patterns. I provide such simulation results for the number of different countries to which an eBay seller with a given number of transactions would sell. I then compare this with the actual data for eBay sellers.

In an extreme scenario of a perfectly "flat" world without fixed costs to export, the probability for a specific transaction to be made with country c would only depend on demand from that country.[75] We would expect that US sellers export more items to Canada as compared to Palau simply because there is more demand from Canada. Demand from Palau may be low because GDP is low, or because shipping costs are higher, or shipping quality is lower, or other factors. As demand weights of all destination countries, I take the shares in overall eBay exports (by transactions). Export shares based on actual eBay exports could be seen as endogenous. For example, if there are high fixed costs to enter a certain market, then exports to that market will be lower, and thus also the predicted number of exporters to that country. Therefore, as a robustness check, I also do the simulations using country

---

[73] There are of course exceptions. For legal and other reasons, a seller could limit shipment to certain countries, or exclude certain countries.

[74] Taking into account the product dimension, consider Amazon US' domestic shipments of books by ZIP code area and ISBN code of the book – an extreme example of "sparse data". While Amazon probably ships some books to any ZIP code, many books will only be shipped to a few zip codes. This does not indicate that there are ZIP-code specific "entry costs" for a book. See Hillberry & Hummels (2007) for a paper using US domestic shipments by ZIP code, where they make a similar argument.

[75] I assume here that all transactions are independent of each other. In practise, some transactions could be with the same buyer, but this is not shown in the data. See my discussion on this below.

weights taken from the predicted shares from a gravity regression.[76] As I will show below, the results are very similar.

With the number of sales going towards infinity, a seller would export to all possible destination countries, and the distribution of sales across foreign countries would resemble world demand shares of these countries. In practice, the number of transactions by each eBay seller is small, which explains why, although exporting to a lot more countries than offline sellers in large industrial countries, they do not sell to all countries. For example, the US dataset includes 216 export destinations.[77] Given the number of eBay exporters, there would be around 136 million possible seller-destination combinations, but we only observe 3.8% of these. Offline data for the US shows that there are even less non-zeroes – 2%.[78]

Not even the largest exporters reach all markets - the largest number of countries reached by an US eBay seller is 181 out of 216. A seller that, for example, puts 100 items online and sells all would obviously not get orders from all 216 countries. Even a seller who ships 1’000 items in a given year is very unlikely to ship to all countries in the world – even in the complete absence of any country-specific fixed costs.

If I assume that purchases are made randomly, then the expected number of different export destinations *n_dest* out of *C* countries for a seller with *t* transactions can be calculated as follows:

$$E(n\_dest) = \sum_{c} 1 - (1 - s_c)^t \qquad \text{, with } \sum_{c} s_c = 1$$

I add the probability that a specific country receives at least one export across all countries, with the weight for a country *c* being given by s_c , for which I either use the actual overall share in eBay exports, or the predicted shares from a gravity regression. Obviously, for *t* going towards infinity, the number of different destinations goes towards *C*, the number of countries – everybody would export to everywhere:

$$\lim_{t \to \infty} E(n\_dest) = C$$

By definition, the expected (and actual) number of destinations cannot be higher than the number of transactions:

$$E(n\_dest) \leq t;\ n_{dest} \leq t$$

The following is worth noting: For a given number of destinations, sellers and transactions, the expected number of destinations per seller is higher the more equal sales are distributed among sellers and among destinations for small numbers of transactions. For very large numbers of transactions, everybody would export everywhere, so the distribution has no impact.

I now calculate the expected number of destinations for each eBay exporter. Averaging across exporters gives the average expected number of destinations reached, which I can compare with the actual figure.

---

[76] Other alternatives could be to use GDP shares as proxies for the demand share, or shares in offline exports of the exporting country. Results using such shares are again similar. The reason is that all these proxies for the country shares have in common that country shares are very unequal. Would one assume that all country shares are the same, the expected number of destinations per seller would be much higher.

[77] There is a small number of destinations without any exports, which I ignore here. Note that eBay uses a country classification that is very similar to the standard classification used in trade statistics. “Countries” refers to both independent countries and dependant territories (such as Hong Kong or Pitcairn).

[78] See Armenter & Koren (2010). The number is based on Bernard *et al.*’s work.

The aggregate results for eBay sellers are shown in Figure 11. Here, I compare the actual and simulated average number of export destinations. For all "eBay countries", we can see that the actual number of destinations reached is lower than the expected number, but not much so. Also, we see that the figures vary across countries. This is mainly because of different coverage of the datasets. For example, the Australian dataset covers all sellers, including the smallest. Many of them only have one export transaction a year, and thus export to only one destination.

To make these results more comparable across eBay countries, I take a subset of eBay exporters with around 100 export transactions. Figure 12 shows the results for these sellers only. One can see that the expected number of destinations is higher, and varies less - between 25 and 30 countries.[79] In reality, sellers ship to "only" 16-22 countries.[80] Another reason why it is useful to look only at sellers with a certain size is that seller size could be seen as endogenous. I assume that the number of items that each seller is able to sell is derived from their exogenous productivity, one could also assume that their size is endogenous and affected by, for example, destination-specific fixed costs. Say, a seller selling 10 items would sell more if she could overcome the fixed cost of entering another market. I overcome this potential problem by also looking only at sellers with a given number of transactions (e.g., 100).

One may argue that the demand shares of destination countries are endogenous because countries for which exporters face high costs to export may buy less from eBay sellers. One should then not take country shares derived from actual exports. I therefore also used alternative market shares of destination countries. Rather than using shares from eBay trade, I use the predicted market share from a gravity regression.[81] This slightly changes the number of expected destinations reached per seller, although not by much. The results are shown in Figure 12 (green bars). Figure 13 plots the actual and expected number of destination countries for a given number of transactions for US-based eBay sellers, where I use a wide range of different weights for destination countries. The expected number of destinations for a given number of export transactions varies little among different weights, such as GDP or US offline exports. Only when I assume that each destination country had the same weight would the expected number of destination countries be much higher than the actual one.[82]

> ***Empirical result 10:*** *The number of markets reached by eBay exporters is almost as high as the balls-and-bins model predicts. This applies for a wide range of plausible weights used for destination countries.*

Figure 12 also show the results for such a simulation made with detailed offline data for Malawi and Peru. These may not be the best countries to be used for a comparison with the large developed countries for which I have eBay data, but I currently do not have data available for other countries that shows the same level of detail as the eBay data. We can see that the expected number of destinations is similar to what we can see on eBay, but the actual number is much lower. For example, exporters in Malawi (Peru) with around 100 export transactions "should" export to 25 (29) different destinations, but in reality only sell to 4 (6).[83] The balls-and-bins model thus does not well describe what is

[79] The numbers vary because they are calculated based on the distribution of exports across destination countries, which vary across exporters. Using equal country weights would result in an expected number of export destinations of around 80.

[80] For US-based eBay seller who exports 1'000 items in a year, the expected number of different export destinations would be 72, using overall eBay purchases from the US as weights. If I assumed equal weights, the expected number of countries for 1'000 transactions would be almost 216. In reality, US-based eBay sellers with 950 to 1,050 foreign transactions are shipping to 50 countries on average.

[81] I use a gravity regression as shown in Table 6 (lower part).

[82] Results look similar for other eBay countries.

[83] The actual and expected number of destinations for all sellers is much lower. Most exporters conduct few transactions per year, with the share of exporters having only one single transaction being 16% (Malawi) and 40% (Peru).

happening for offline exporters, at least not in these two markets. For the US, I can calculate the respective figures from Armenter & Koren's simulation, which also shows that the model highly over-predicts the number of destinations.[84]

In Figure 14 (Malawi) and Figure 15 (Peru), I show this in more detail, analogous to Figure 13 for US-based eBay sellers. The difference is remarkable: The balls-and-bins model predicts many more destination markets for all levels of export transactions (except of course when there is a single export transaction only). This holds when I use actual exports as weights for destination countries and also when using importing country GDP as weights.

Unfortunately, I cannot replicate these calculations for other countries. I do have aggregated seller-level data for 8 additional countries from the World Bank dataset, but that data does not include the number of transactions. I can only observe the export value per exporter, product and destination markets. This is similar to US the data that Armenter & Koren used for their simulations. Essentially, the key information required for the "balls-and-bins" simulation – the number of transactions (i.e., balls) is not available. One can therefore only calculate the estimated number of export transactions across all exporters by assuming a certain size of export transactions. I have done this, using four different assumptions: Either each observation in the data represents exactly one transaction, or a transaction has the size of USD 1'000, 10'000 or 100'000. The smaller one assumes a transaction to be, the larger will be the number of balls, and thus the expected number of different export transactions.

The results are shown in Table 9. First of all, we see that the average number of export destinations is very low – ranging from 1.5 in Albania to 3.0 in Senegal. Assuming that each observation represents only one transaction gives an *expected* number of destinations that is about twice as high. Results are similar when I assume that each transaction has a value of USD 100'000. For smaller transaction sizes, the expected number of destinations increases. Taking Mexico as an example, I see that exporters sell on average to 2.2 markets. They "should" already export to 4.3 markets if exports were random and assuming that each observation is one transaction. When I assume that transactions are very small (USD 1'000), the random model generates an expected number of destinations of 21. The deviation between reality and the balls-and-bins model is therefore fairly large.

What do we learn from this? First of all, offline exporters sell to fewer markets than one should expect from the balls-and-bins model, which indicates that their exports are not "random" and independent of each other. This can be interpreted as evidence for fixed costs to enter specific markets (or to enter into a buyer-seller relationship). Secondly, the small number of destinations reached cannot only be explained by fixed costs. A main factor is the small size of exporters. Across these 8 countries, around 40% of all sellers appear only once in the data – they export one product category to one market (and often only a very small value).[85] Therefore, the expected number of destination markets is also fairly low.

> ***Empirical result 11:*** *The number of markets reached by <u>offline</u> exporters is much lower than predicted by the balls-and-bins model. Also, the number of export destinations in offline firm-level data is highly driven by exporters with a single transaction only.*

What can we conclude up to here? eBay sellers export to many more markets than their offline counterparts. The number of destinations that they reach follows fairly well the expected pattern under the assumption that eBay transactions were "random" and without any fixed costs. In contrast, the

[84] The US offline figure can only be calculated for all sellers, not for those with 100 export transactions.
[85] See last column in Table 9.

evidence I have for offline exporters suggests that this is not true in traditional trade. Some differences remain between actual eBay exports and what is predicted by the balls-and-bins model, for which I will suggest some explanations below. In any case, there is a striking difference to offline exporters.

I will now take a look at the *number of exporters per destination* in eBay and offline trade. A different way to compare random exports with actual exports is to compare the number of different sellers (out of S sellers) that ship to each country. If buyers in the importing country would pick eBay sellers randomly and independently for each transaction, then the expected number of sellers from which buyers in a particular country are purchasing from would depend on a) the overall number of purchases t made by that country, and b) the distribution of offers made by sellers. Again, the more equal this distribution is, the more sellers would be "picked" for a given number of transactions. I assume that the distribution of products offered is proportional to the sales made and then take the share in overall eBay sales as the probability that a transaction is made with a specific seller. The expected number of sellers (*n_seller*) selling to a destination country that makes t purchases is then given by:

$$E(n\_sellers) = \sum_{sellers} 1 - (1 - s_{seller})^t$$

With s_{seller} being the share of each seller in total sales, I add up the probability that at least one purchase is made from a seller across all sellers. Similar to the formula for the expected number of destinations, it is clear that:

$$\lim_{t \to \infty} E(n\_sellers) = S \quad \text{and} \quad E(n\_sellers) \leq t;\ n\_sellers \leq t$$

The result for the US is shown in Figure 16. The expected number of sellers per country is indeed close to the actual number. Liberia, the smallest destination, with only one single purchase from US-based eBay sellers, obviously only has one US seller. For Canada, the largest foreign destination, there are even more sellers exporting to than what we would expect. For most other countries, there are less sellers exporting than expected. The average ratio of actual to expected sellers is 0.71.[86] Overall, in the simulation 5.1% of all exporter-destination combinations for US-based sellers are expected to be non-zero, while in reality it is 3.8% (26% less). Results are similar for other "eBay countries", as shown in Figure 19. This demonstrates again that the balls-and-bins model predicts fairly well the extensive margin.[87]

> ***Empirical result 12:*** *Destination countries import from fewer eBay sellers than predicted by the balls-and-bins model, except for neighboring countries with strong ties. Most exporter-destination combinations in eBay trade are zero. The balls-and-bins model only slightly under-predicts the share of zeroes.*

Offline, one gets entirely different results. I can calculate such figures for Australia, Malawi and Peru. In Figure 17, I show results for both Australian eBay exports and offline exports. While the balls-and-bins model predicts fairly well eBay exports, it heavily over-predicts the number of exporters per

[86] Note that Figure 16 is in logs, which makes that ratio look smaller than it actually is. The respective charts for the other four exporting countries look very similar. The average ratio of actual over expected sellers per destination is between 0.65 and 0.74, and for some countries the number of sellers exporting is higher than expected

[87] Note that these results can also be derived directly from the results of the section above on the number of destinations per exporter. For example, US sellers export on average to 8.2 out of 216 destinations, which means that $8.2/216 \approx 3.8\%$ of combinations are non-zero. The expected number of destinations per exporter implies $11.1/216 \approx 5.1\%$ non-zero combinations.

destination for offline exporters.[88] Figure 18 shows results using transaction-level offline data for Peru.[89] Again, the model does not fit the data. In Figure 19, I show the actual and expected overall share of non-zeroes in exporter-destination observations for the US (taken from Armenter & Koren) and Australia, Malawi and Peru (calculated by the author). The balls-and-bins model over-predicts the non-zeroes by factor 2-4.

> ***Empirical result 13:*** *Offline data shows that destination countries import from much fewer exporters than predicted by the balls-and-bins model. The balls-and-bins model significantly under-predicts the share of zeroes.*

What do these results tell us? There are several possible explanations why eBay sellers export to fewer destination countries (or that there are fewer exporters per destination) than they "should" under the assumption that there are no country-specific fixed costs and all export transactions are independent of each other. One possible explanation is of course that there are indeed destination-specific fixed costs that prevent sellers from shipping to a large number of markets. However, given the nature of eBay transactions, it would mean that a seller refuses to fulfill transactions from some countries. This is possible – a seller can declare to which countries he or she is willing to sell. This sometimes happens, especially for items where trademark rights prevent sellers from shipping abroad. But sellers would then usually limit shipments to a small number of countries (such as US only, or US & Canada), and not to a large number of countries, while restricting sales to the remaining ones.[90] If this drives the results, then we would expect to see a large number of sellers that restrict sales to a few countries. This is not the case. Even when excluding sellers that sell to, say, less than 20 markets (for a given number of transactions, such as 1'000), I find that the remaining eBay sellers still sell to fewer markets than what we would observe if exports were truly random.

The data shows that many small exporters sell only to one main export market, which for US sellers is primarily Canada. For example, I find that among those US exporters with five export transactions, 5% export only to Canada. The balls-and-bins model predicts a probability of only 0.15% for this.[91] I argued above that there are likely some fixed costs to export on eBay, but these fixed costs might be lower for exports to "easy" markets such as Canada that allow for faster shipping, make it feasible to ship bulkier items[92] and impose few language problems between buyers and sellers. I therefore assume that a substantial share of the small sellers that is willing to ship abroad restricts this option to shipments to Canada. Such patterns can also be found for European eBay countries, although here the restrictions are more likely applied to a group of countries, such as the EU (which avoids dealing with customs issues). In France, out of sellers with five export transactions, three times more sellers than predicted only sell to EU markets.

Another likely explanation why the balls-and-bins model under-predicts the number of export destinations is the following: As already explained above, transactions may not be independent,

---

[88] Note that I do not have the actual number of transactions for individual sellers. As explained in the note to Figure 17, I assume that seller size follows a Zipf distribution. If one assumes an equal distribution, then the "simulated sellers by country" are higher and thus the deviation from the balls-and-bins model becomes larger.

[89] It looks similar for Malawi.

[90] This may of course happen in some cases, but I do not believe that it is common practice.

[91] Given Canada's export share, the chance that five export transactions all go to Canada is $0.271^5 \approx 0.0015$. Similarly, I find that among those with ten export transactions that only sell to one country, in almost all cases this country is Canada.

[92] Average shipping costs to Canada are not different from those to more distant destinations. However, the data is not detailed enough to show whether, for example, many bulky items are shipped to Canada, for which shipping costs are certainly lower than to distant countries that require sea transport or (expensive) air transport.

simply because one buyer may order several times from the same seller.[93] If one would count these transactions as one, then the predicted number of destinations would be lower and thus closer to the actual number. However, the average number of individual transactions for a buyer-seller relationship would have to be fairly high. For example, US eBay sellers with around 600 export transactions sell to around 40 different countries. If exports were purely "random", and using eBay export weights, we would already expect this number of different countries for only 200 transactions. In other words, on average three transactions would have to be made with the same buyer. Unfortunately, with the data that I have so far it is not possible to verify this (one would need data that includes a buyer identifier). Some evidence for this can be taken from destinations with very small number of transactions: Countries that have, for example, 10 or less import transaction from eBay often tend to buy several items from the same seller. It is more likely that these are then also the same buyers, rather than two independent buyers picking the same seller. For example, there are seven transactions between Eritrean buyers and a total of only four UK sellers. Seven random purchases from UK sellers would very unlikely result in only four sellers being chosen.

Another explanation would be that taste differs widely across countries and therefore the demand weights that one should use to simulate export patterns in the balls-and-bins model should be more concentrated. Sellers tend to specialize in certain products, and if these products face a more concentrated demand distribution across countries than the overall demand distribution, then the expected number of destination countries should be lower, and closer to what we observe on eBay. I checked whether the balls-and-bins model is better in replicating the number of destination countries separately by each product category (and using country weights taken from eBay exports of the particular product category). For most product categories, the number of destinations to which sellers are exporting is below the predicted value from the model (exceptions are one electronics category and phones). Products that stand out most are stamps. Unfortunately, the data is not very detailed at the product level. eBay trade flows are separated into only 37 different product categories, whereas offline data usually includes 5'000 to 10'000 different product categories. Demand distributions *within* a sub-category of a product category may be much more concentrated among countries than for the broad product category. For example, eBay sellers that specialize in stamps from a certain region may face a more concentrated distribution of demand and would therefore be expected to export to fewer countries, compared to what we would expect when using the overall demand distribution of the product category "stamps".[94]

What can we conclude? It appears that the extensive margin of eBay exporters – the number of markets they can reach, can be fairly well replicated by the balls-and-bins model, and equally so the number of exporters selling to each market. There are several plausible explanations for the small deviation from the predicted outcome of the model that do not assume any market-specific entry costs. This is strong evidence that such costs are indeed low. The comparison with offline data shows that the balls-and-bin model does not perform very well in predicting offline patterns, as one would expect if one assumes that entry costs are much more important offline than online.

[93] The data does not show this directly. I do observe that for tiny countries to which only a very small number of sellers export, a fair number of sellers have more than one transaction. It could be that these sellers somehow specialize in items of interest to customers in these countries, but it could also be an indication that the same customer made several purchases from one seller.

[94] In terms of product-country combinations for the US, around 25% of the 216*37 possible combinations are zeroes. Offline, Armenter & Koren show that there are 82% zeroes for US exports – when using 8-digit data. However, when using more aggregated data, the incidence of zeroes drops. At "section" level, with 21 categories (similar to the number of eBay categories), the number of zeroes is only 16% –somewhat close to what I find on eBay (25%). The balls-and-bins model, assuming that demand distribution by product is the same for all countries, predicts that only 14% of the product-country bins would be empty.

9 Conclusion

I have shown that the behavior of eBay sellers is in many aspects different from offline firms. Even though eBay sellers are smaller, they export more frequently and reach buyers in more destinations. The eBay platform can thus be seen as an example to overcome traditional trade barriers. I have also shown that fixed costs to export are low on eBay, and for those sellers that do export, there appears to be an almost complete absence of destination-specific fixed costs. They export almost as if sales were randomly allocated across destinations. My comparison with offline data shows that this is clearly not the case for them. Whether offline exporters are in fact constrained by high market entry costs, and by what type of entry costs, remains open. For example, even very detailed firm-level data does not show whether firms mainly sell to one customer in a country, or to a wide range of customers. The same problem applies to eBay data, but repeated purchases are likely less prevalent on eBay.

These results should not be interpreted such that there are no trade barriers for online trade. Individual sellers may not experience fixed entry costs into markets, but exports or imports of individual shipments from many markets can still be severely constrained by structural or other problems, including those related to complicated customs procedures, expensive or unreliable postal services or regulatory frictions, which constrain the potential that online trading through platforms such as eBay can provide for sellers and consumers. Identifying what barriers for international (but also domestic) eBay trade exist at the country level would be a key area for further research.

Another avenue for further research is to shed some more light on why relatively few of the small eBay sellers export, at least compared to what one may expect from the balls-and-bins model. Is it that they face barriers – whether real or perceived – to export that could be overcome, or is it that they sell items that can only be sold locally? This question is also very relevant for “offline firms”. It may well be that many of the firms that do not export are constrained by distance, not by borders. Further research using data at transaction level with seller and buyer identifiers and ideally also a location identifier at the national level (e.g., ZIP codes) could provide interesting insights.[95]

This paper has also shown that the balls-and-bins model can be a useful tool to analyze sparse trade data when the underlying assumptions are fairly close to the assumptions made in the balls-and-bins model – e.g., no fixed costs and independent transactions. I believe the model can also be very valuable to analyze sparse data when these assumptions are not made, as is usually the case in offline trade. As suggested by Armenter & Koren, the balls-and-bins model could then be used to identify where actual trade deviates from it, and one then has to find alternative explanations for this deviation. A key constraint to use this model is the lack of data on the number of underlying transactions that is missing in much of the offline firm-level trade data.

The eBay data reveals many other interesting aspects of online trade. For example, survival patterns can be compared with offline firm-level data. Also, one interesting empirical finding is that larger eBay sellers specialize in fewer product categories. These will be areas for future research.

[95] A prime example could be to explain the somewhat surprising fact that more French firms export to Belgium than to Germany (see Chaney, 2011). Population statistics easily reveal that the population of border regions with Belgium is similar to the one of regions bordering Germany. If many French exporters are firms that sell only across short distances, then this may well explain the puzzle.

10 References

Armenter, Roc & Miklós Koren (2010). *A balls-and-bins model of trade*. CEPR Discussion Paper No. 7783.

Australian Bureau of Statistics (2012). *Characteristics of Australian exporters*. Data series 5368.0.55.006.[96]

Axtell, R. (2001). *Zipf Distribution of U.S. Firm Sizes*. Science Vol. 293, p. 1818-1819.

Baldwin, R. & J. Harrigan (2011). *Zeros, Quality, and Space: Trade Theory and Trade Evidence*. American Economic Journal: Microeconomics 3 (May), p. 60-88.

Bernard, Andrew B., J. Bradford Jensen, Stephen J. Redding & Peter K. Schott (2007). *Firms in International Trade.* Journal of Economic Perspectives 21(3):105-130.

Chaney, Thomas (2011). *The network structure of international trade*. NBER Working Paper No. 16753.

Cebeci, T., Fernandes, A., Freund, C. & M. Pierola (2012). Exporter Dynamics Database. World Bank Policy Research Working Paper 6229, World Bank.

Eaton, Jonathan, M. Eslava, M. Kugler & J. Tybout (2007). *Export dynamics in Colombia: Firm-level evidence*. NBER Working Paper No. 13531.

Eaton, Jonathan, Samuel S. Kortum & Francis Kramarz (2009). *An anatomy of International Trade: Evidence from French Firm*s. CEPR Discussion Paper No. 7111.

Eaton, Jonathan, Samuel S. Kortum & Francis Kramarz (2011). *An anatomy of International Trade: Evidence from French Firm*s. Econometrica, Vol 79, No .5 (September), p. 1453-1498.

Eaton, Jonathan, M. Eslava, D. Jinkins, C. J. Krizan & J. Tybout (2012). *A Search and Learning Model of Export Dynamics*. Mimeo.

Freund, C. & M. Pierola (2012). *Export Superstars*. World Bank Policy Research Working Paper 6222, World Bank.

Head, K. & T. Mayer (2013). *Gravity equations: Workhorse, toolkit, and cookbook*. CEPR Discussion Paper No. 9322.

[96] Available at http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/5368.0.55.0062010-11?OpenDocument (last accessed 19.03.2013).

Helpman, E., M. Melitz & Y. Rubinstein (2008). *Estimating Trade Flows: Trading Partners and Trading Volumes.* Quarterly Journal of Economics, Vol. CXXIII (May), Issue 2, p. 442-487.

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

Keck, A. & Lendle, A. (2012). *New evidence on preference utilization*. WTO Staff Working Paper No. 2012-12.

Lendle, A., M. Olarreaga, S. Schropp & P.-L.Vézina (2012). *There Goes Gravity: How eBay Reduces Trade Costs*. CEPR Discussion Paper No. 9094.

Mayer, Thierry & Gianmarco I. P. Ottaviano (2007). *The Happy Few: The internationalisation of European Firms*. Bruegel, Brussels.

Shi, Huimin (2011). *The Missing Trade of China: Balls and Bins Model.* Mimeo.

World Bank (2012). *Exporter Dynamics Database*. Washington D.C.: World Bank.[97]

[97] The database and further background material is available at http://econ.worldbank.org/exporter-dynamics-database.

11 Appendix for data used for gravity regressions

This section describes the variables used in the gravity regressions:

- eBay exports: eBay sales by export destination. Source: eBay.
- Offline sales: Data taken from Comtrade (through WITS) at 6-digit level. "eBay image" refers to a subset of that data with 6-digit codes that can be matched with eBay flows.
- Importer and exporter GDP: IMF World Economic Outlook, in current USD.
- Distance: Distance between two countries based on bilateral distances between the largest cities of those two countries, those inter-city distances being weighted by the share of the city in the overall country's population. Source: CEPII Distances database.
- Shipping costs: shipping fee divided by transaction value. Source: eBay.
- RTA: dummy variable indicating whether the two countries are applying a free trade agreement with each other. This only includes RTAs for which exporters get preferences. Source: CEPII and corrections made by the author. The following RTAs are included (from year of implementation):
    - USA: RTAs with AUS, BHR, CAN, CHL, CRI, DOM, GTM, HND, ISR, JOR, MAR, MEX, NIC, OMN, PER, SGP, SLV.
    - DEU / FRA / GBR: RTAs with ALB, AND, BIH, CHE, CHL, CYP, DZA, EGY, HRV, ISL, ISR, JOR, LBN, MAR, MEX, MKD, NOR, SMR, SRB, TUN, TUR, ZAF.
    - AUS: RTAs with BRN, CHL, IDN, KHM, LAO, MYS, NZL, PHL, SGP, THA, USA, VNM.
- Internet use: Number of internet users over population. Source: World Bank World Development Indicators.
- Common official language: dummy variable indicating whether the two countries share a common official language. Source: CEPII database.
- Common ethnic language: dummy variable indicating whether the two countries share a common ethnic language. Source: CEPII database.
- Common legal system: dummy variable indicating whether the two countries have the same legal origin. Source: CEPII database.
- Colonial link: dummy variable indicating whether the two countries ever had a colonial relationship. Source: CEPII database.
- Common currency: Indicates whether both countries use the same currency. Source: CEPII database and corrections made by the author. The dummy is one (not always for all years) for:
    - USA: BHS, BMU, ECU, FSM, LBR, MHL, PAN, PLW, SLV, ZWE.
    - DEU / FRA (varies by year): AND, AUT, BEL, CYP, DEU, ESP, EST, FIN, GRC, IRL, ITA, LUX, MLT, NLD, PRT, SMR, SVK, SVN.
    - GBR: none.
    - AUS: KIR, TUV.
- Common border: dummy variable indicating whether the two partners share a border. Source: CEPII.

12 Figures and Tables

Figure 1: "Random draws" on eBay

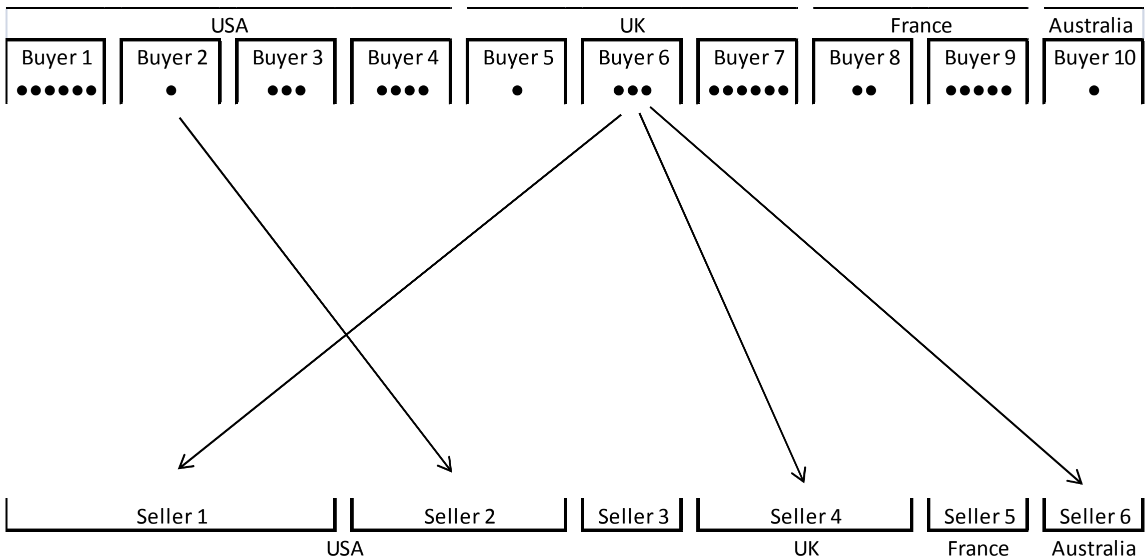
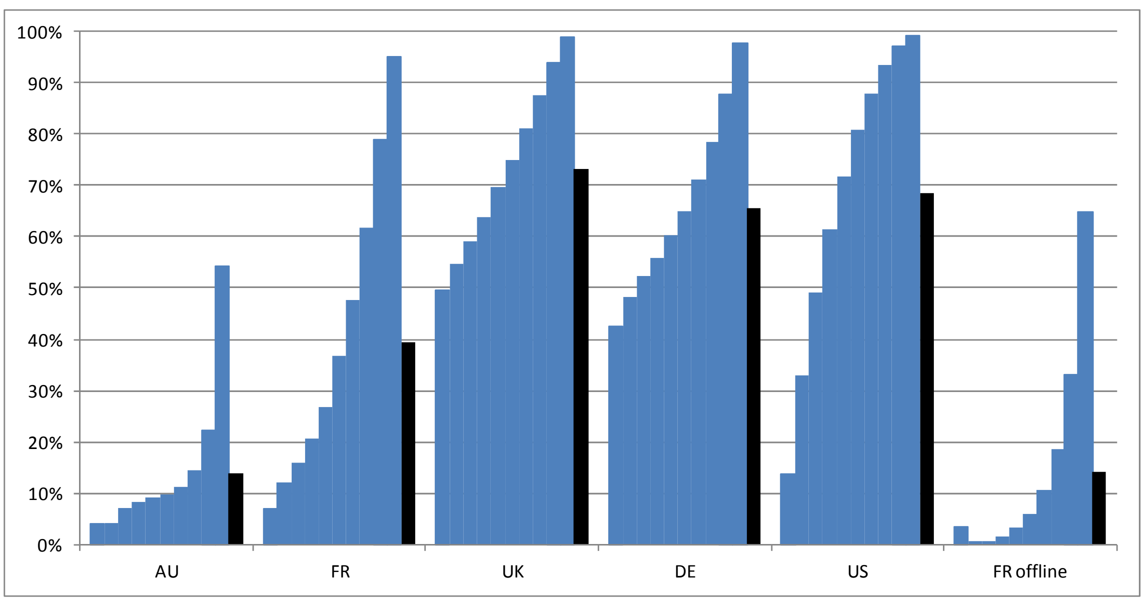


Figure 2: Export status: Share of sellers exporting by deciles – eBay and French offline firms

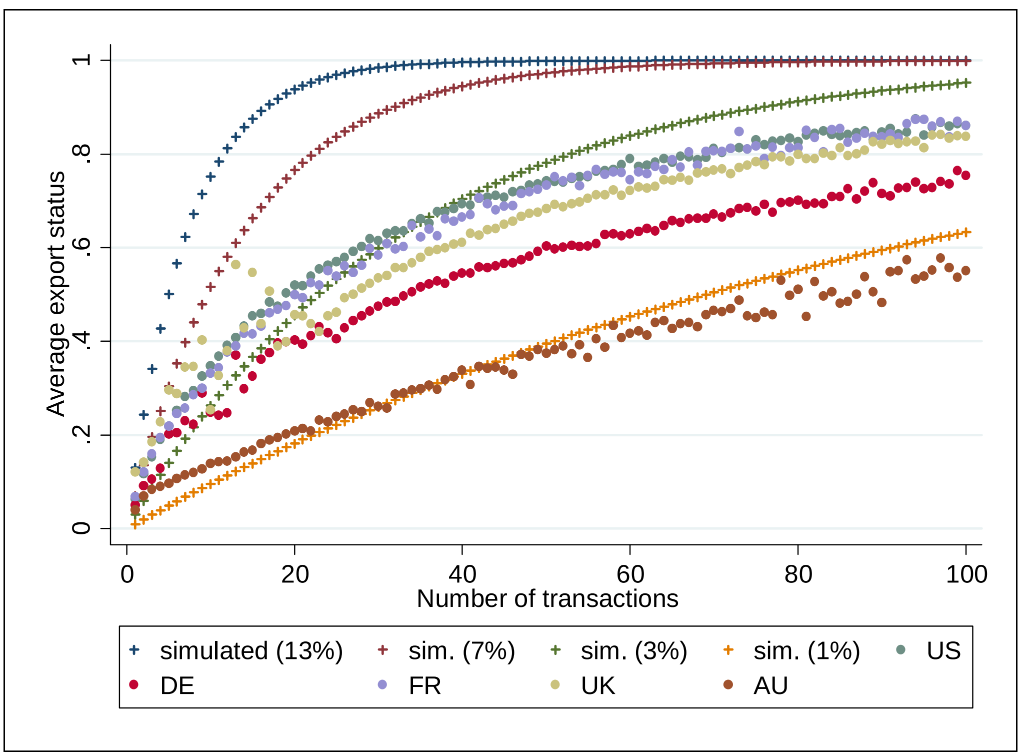


Source: eBay data for 2010 (US) and 2011 (FR). France: Eaton *et al.* (2009), Table 5. The chart shows the average share of sellers (firms) exporting by decile (based on transactions, except FR offline (sales). The black bar shows the overall average.

Figure 3: Export status: eBay versus a simulated export status for all sellers

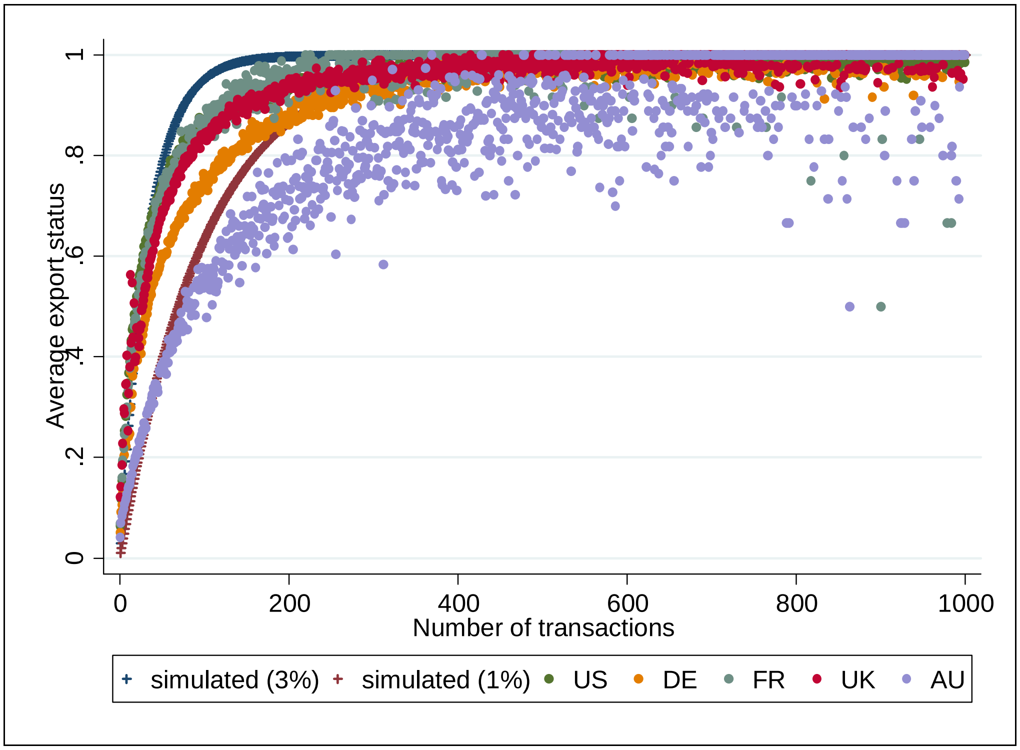
Source: Calculation by the author using eBay data for 2010 (US) and 2011 (others). The share of transactions exported varies from 6-13% across the five exporting markets. Note that the eBay data for US, DE and UK does not cover all of the smallest sellers. The figures therefore cannot be interpreted as being representative for the entire exporting country. Also, figures are based on the location of the exporter, not based on the eBay site of an individual country. They should therefore not be interpreted as being representative for, for example, the eBay.com site. According to information obtained from eBay, around 40% of all US-based sellers have exported in 2012. This number is smaller than the 68% reported here because it is also based on a fairly large number of small sellers of which many are missing in my US data and who are more often non-exporters."US offline" figures are taken from Armenter & Koren (2010).

Figure 4: Export status: eBay versus a simulated export status for 1 to 100 transactions (all exporting countries)



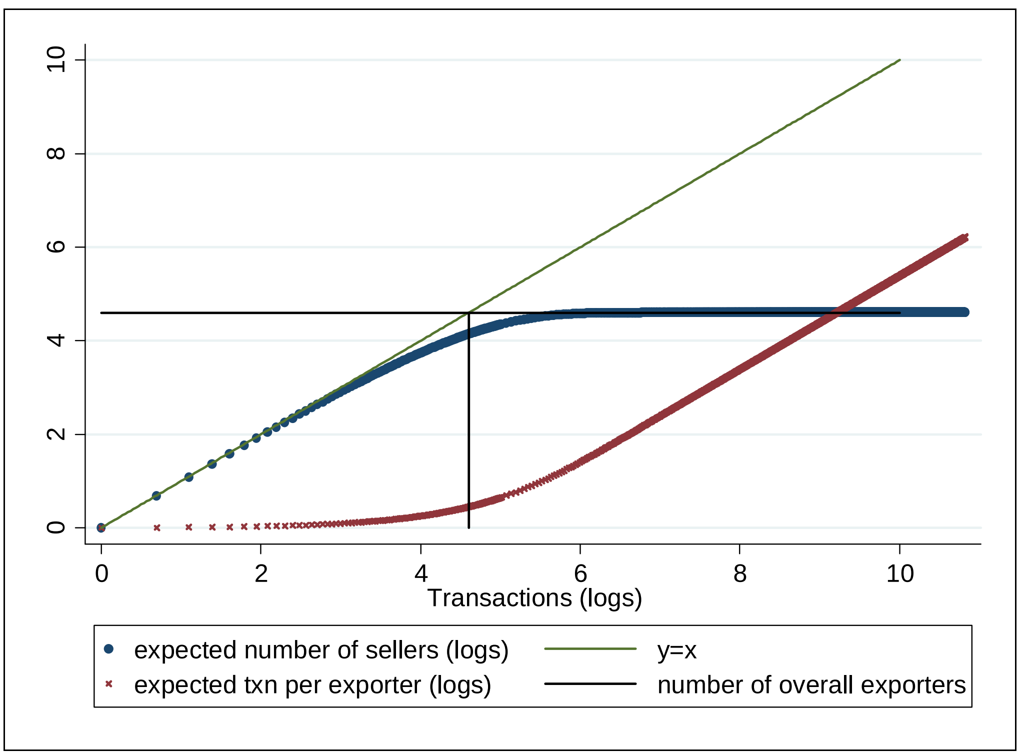
Source: Calculation by the author using eBay data for 2010 (US) and 2011 (others).

Figure 5: Export status: eBay versus a simulated export status for 1 to 1000 transactions (all exporting countries)



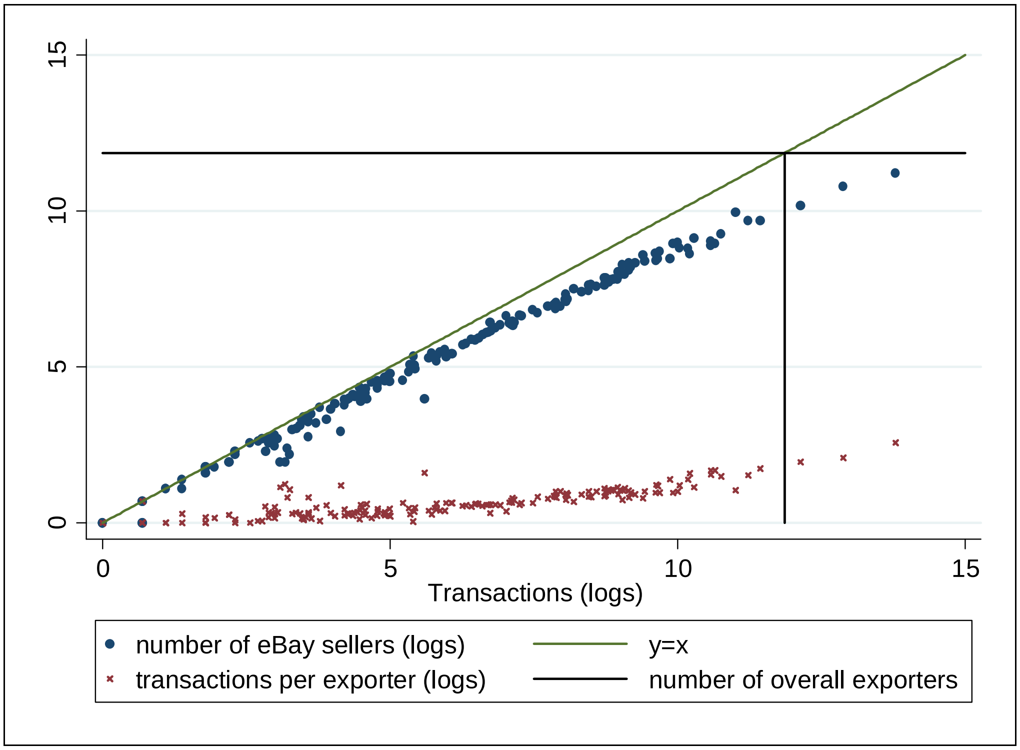
Source: Calculation by the author using eBay data for 2010 (US) and 2011 (others).

Figure 6: Simulation to illustrate when exports are driven by the extensive or intensive margin



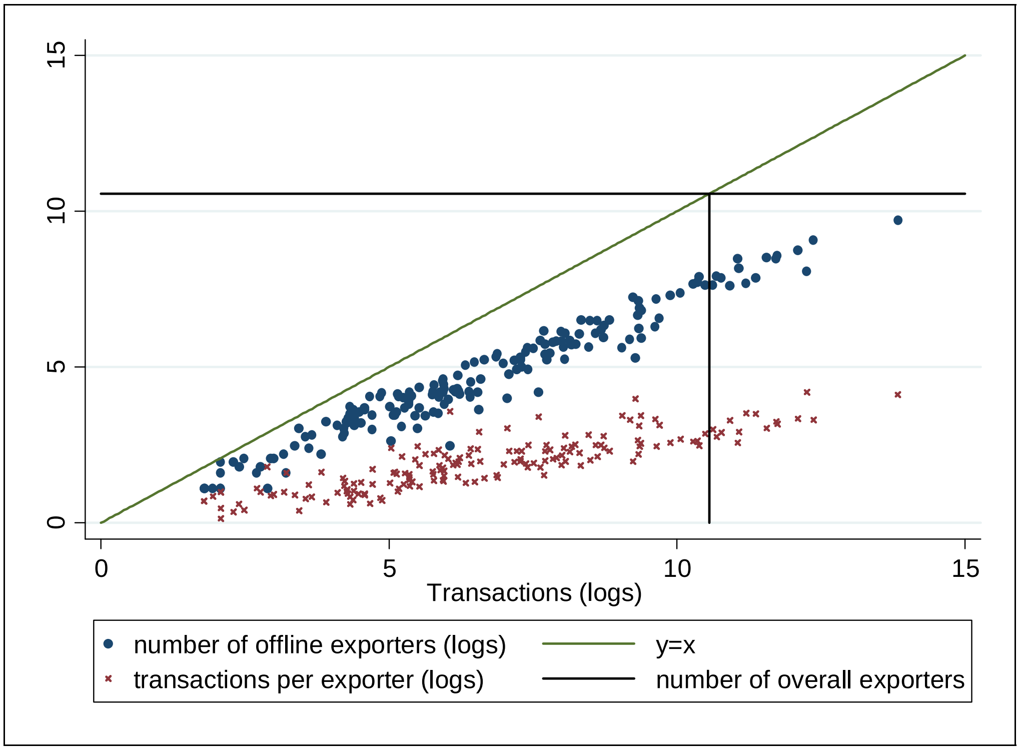
Source: Own calculations. It is assumed that there are 100 identical exporters ($\ln(100) \approx 4.6$).

Figure 7: Are Australian eBay exports driven by the extensive or intensive margin?



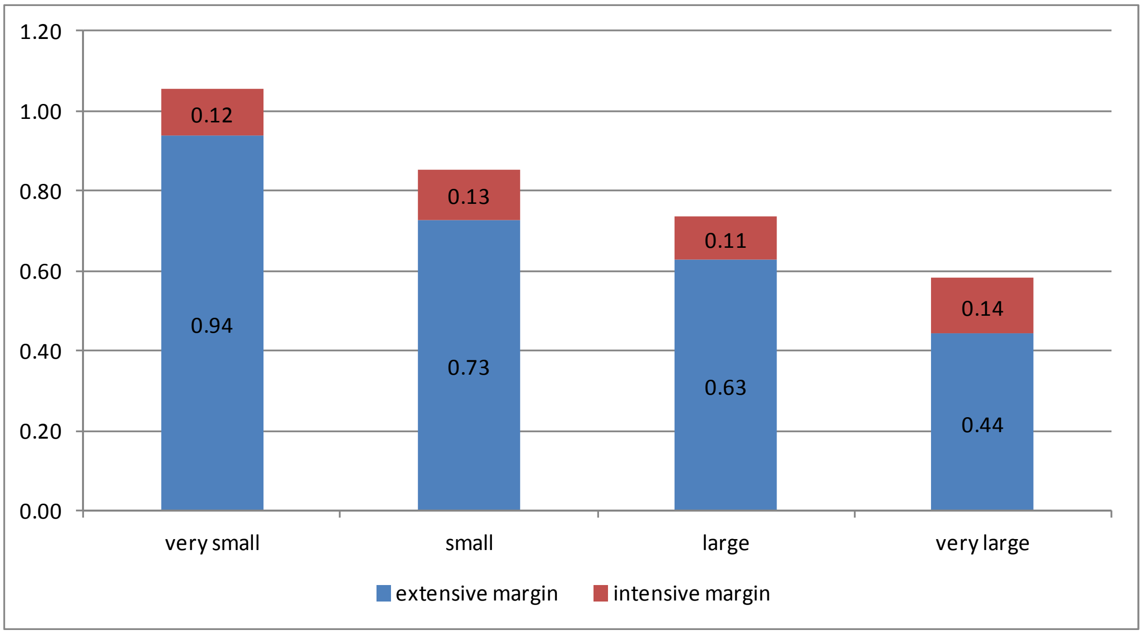
Source: Own calculations using eBay Australia data for 2011.

Figure 8: Are Australian offline exports driven by the extensive or intensive margin?



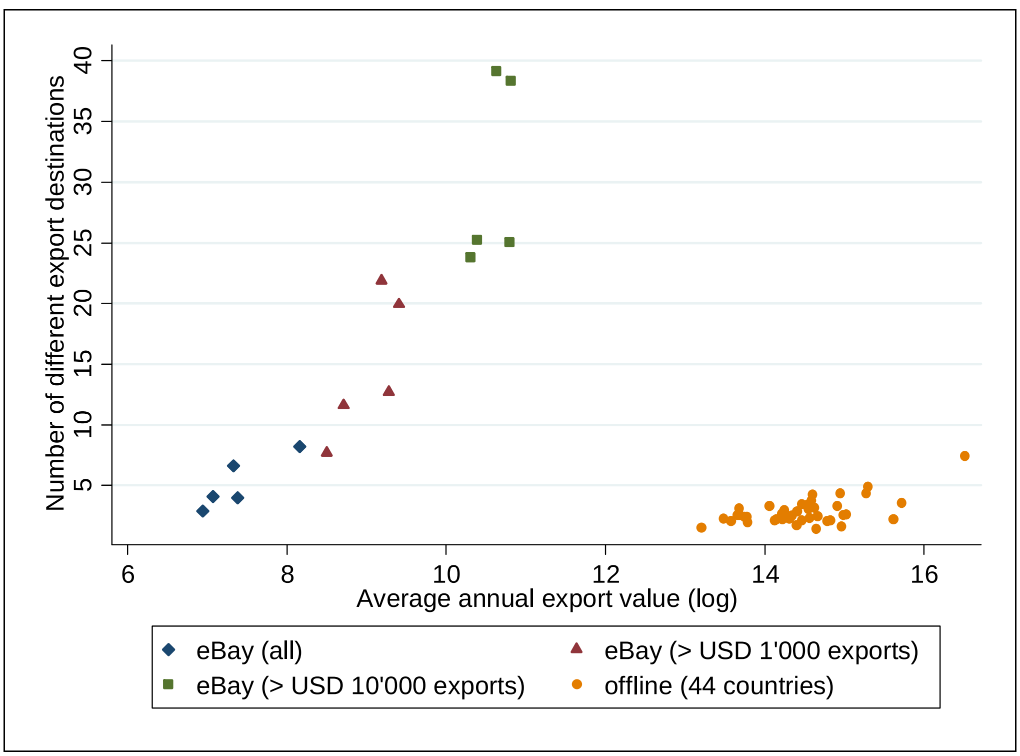
Source: Own calculations using data from the online appendix of ABS (2012). The total number of exporters was taken from Table 5.1 and the number of exporters per destination from Table 11.

Figure 9: Distance coefficient for small versus large eBay exporters



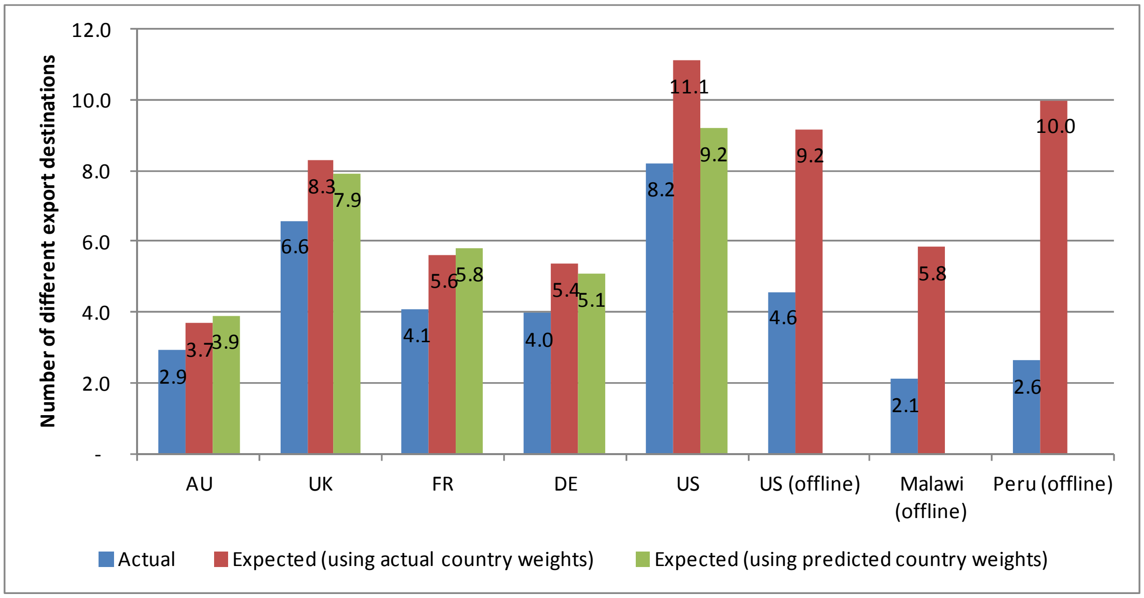
Source: The chart shows the (negative) distance coefficient of a gravity regression similar to the one shown in Table 5 (B5 & B6), but separately for four different size categories of sellers (< 25; 25 to 99; 100 to 999, > 1'000 export transactions). The extensive margin refers to the number of exporters, and the intensive margin to sales per exporter. All coefficients are significant at the 1% level. Note that this pattern is not limited to the distance coefficient. It also appears for importer GDP (when no country fixed-effects are used).

Figure 10: Number of export destinations – five eBay countries versus "offline" exporters in 44 countries



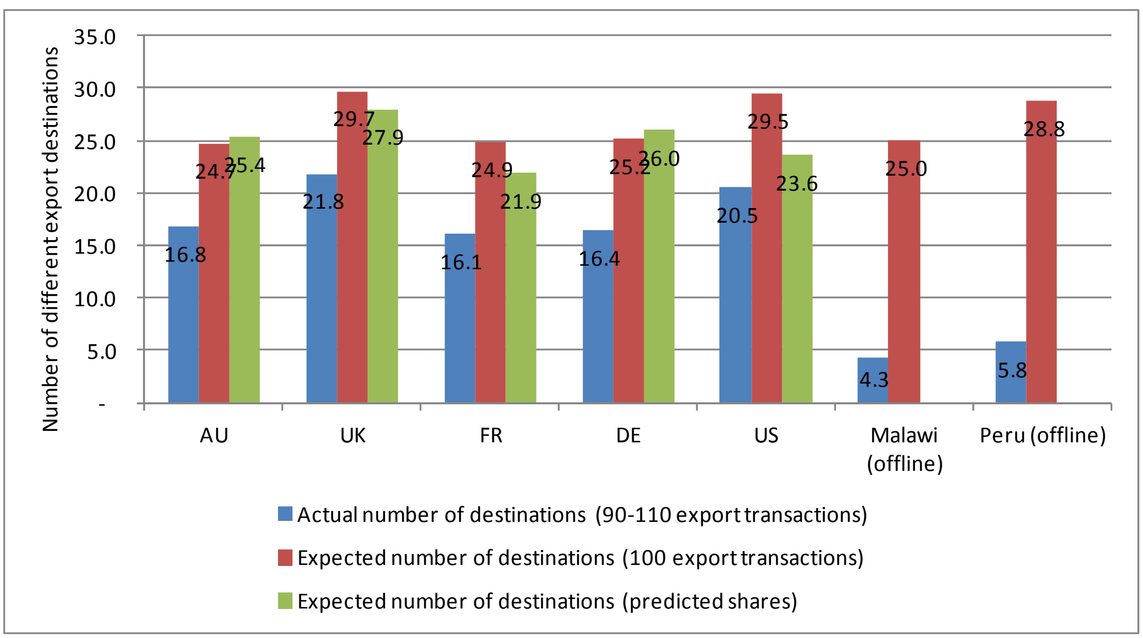
Source: Calculation by the author using eBay data for 2010 (US) and 2011 (others) and World Bank (2012) for "offline" figures.

Figure 11: Actual and simulated number of export destinations across all exporters



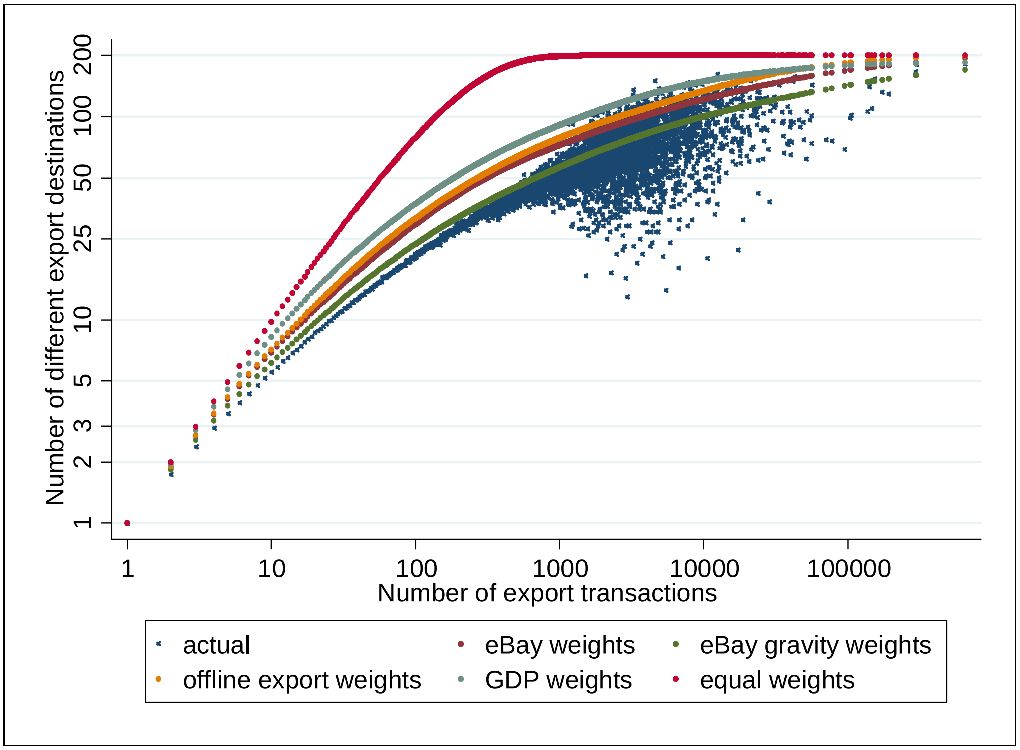
Source: Calculation by the author using eBay data for 2010 (US) and 2011 (others). Offline data for the US is calculated from Armenter & Koren's (2010) calculation of the actual and simulated share of zeroes in exporter-destination combinations. Offline data for Malawi and Peru is based on own calculations and simulations using customs data provided by the World Bank. "Actual numbers" are taken from all sellers. "Predicted shares" refers to simulations using estimated market shares from a gravity regression (see main text). The average number of destinations is an average across all sellers for which data is available. However, because I do not have complete data for all small sellers, the numbers should not be interpreted as being representative for all active eBay sellers located in these countries.

Figure 12: Actual and simulated number of export destinations for 100 export transactions



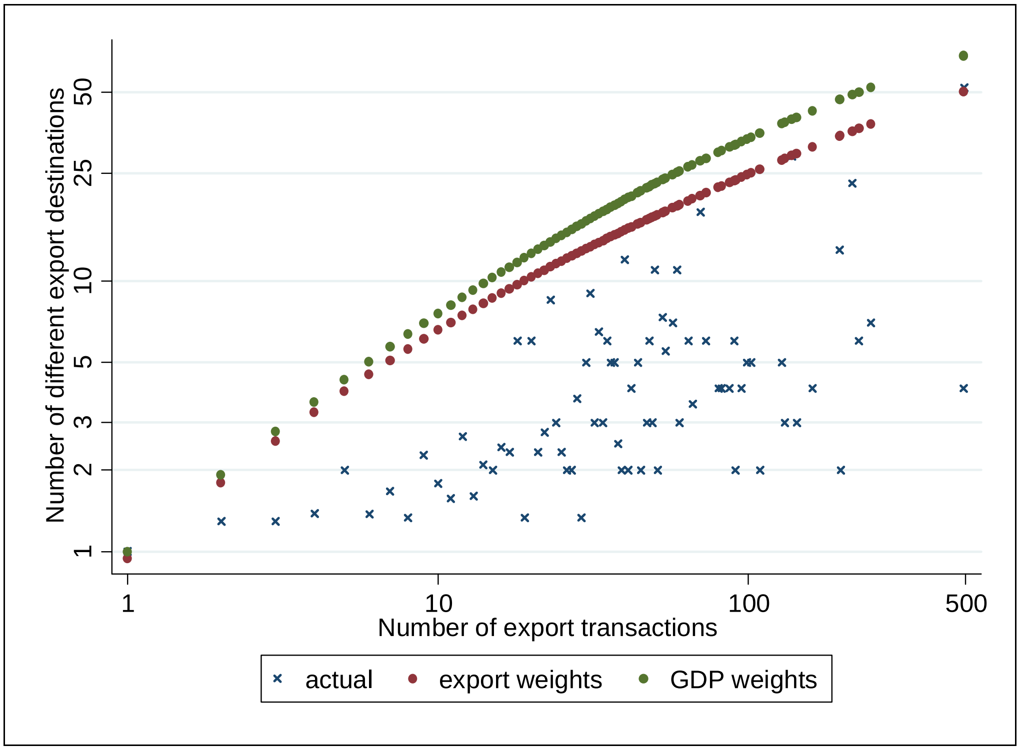
Source: As in previous chart. "Actual numbers" are taken from all sellers with between 90 and 110 export transactions to increase the sample size. Results for sellers with exactly 100 export transactions are almost identical.

Figure 13: Actual and simulated number of export destinations (eBay US)



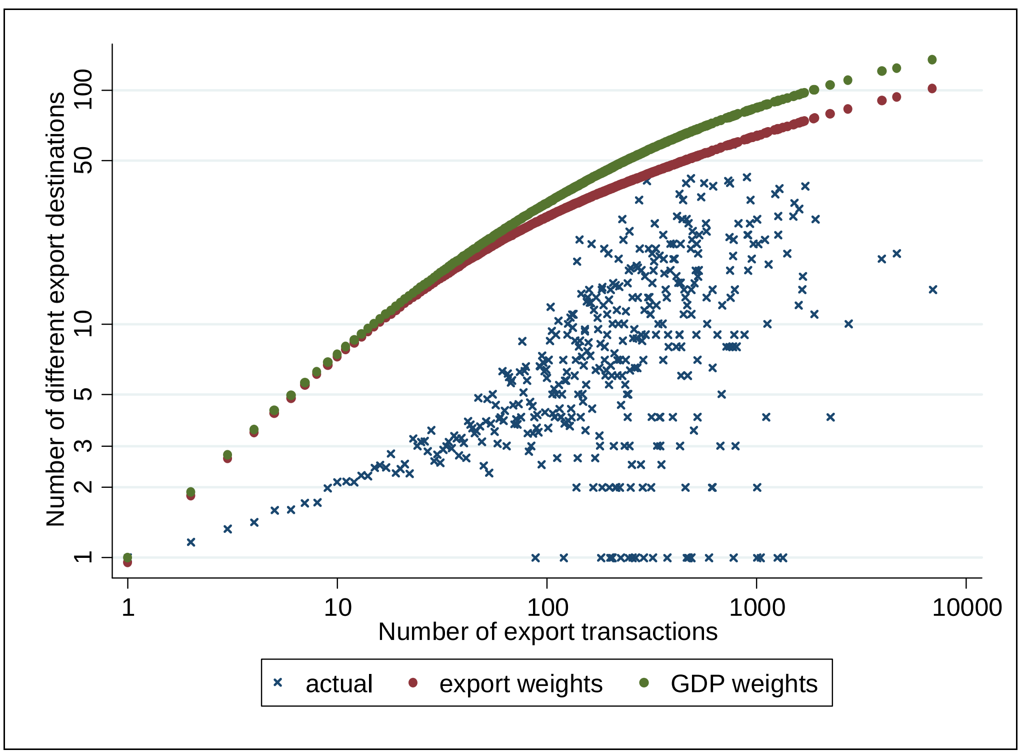
Source: Calculation by the author using eBay data for 2010. Other data used: "eBay gravity weights" uses country weights from a gravity regression on US eBay exports. "Offline export weights" assumes country weights are the same as in US offline exports in products similar to the ones traded on eBay. "GDP weights" uses country weights from 2010 nominal GDP. "Equal weights" assumes each country has the same share. Only shares among the top 200 countries are used. Note that actual eBay figures for very large numbers of transactions are driven by individual sellers and therefore do not increase gradually with the number of transactions.

Figure 14: Actual and simulated number of export destinations (Malawi – offline exports)



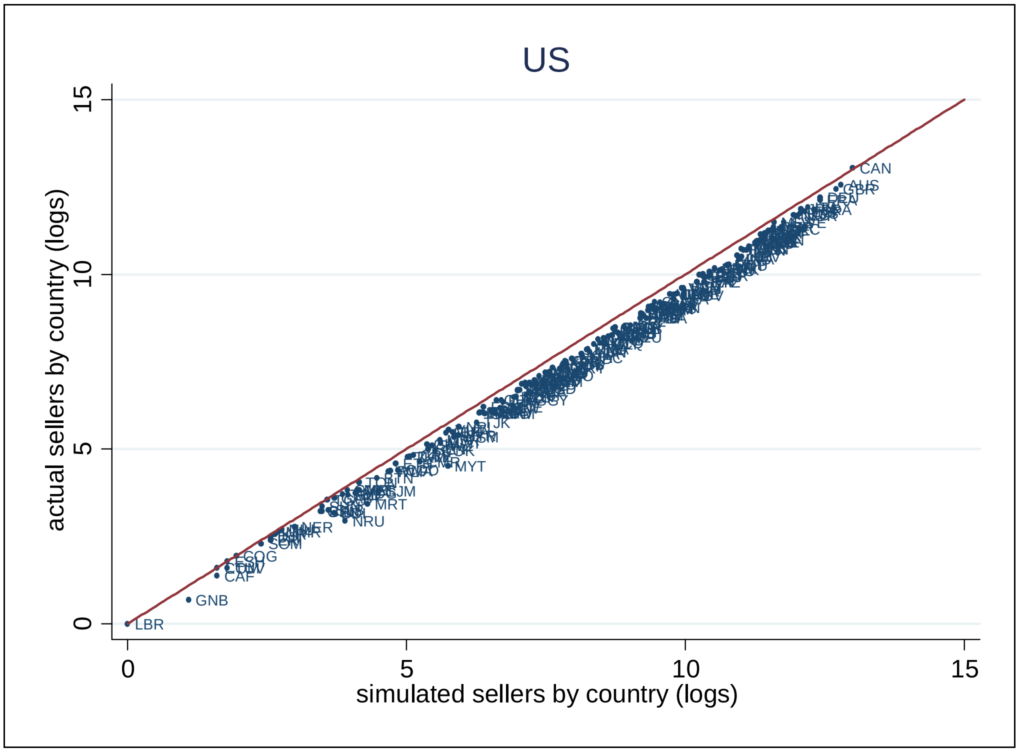
Source: Calculation by the author using Malawi data for 2008. "Export weights" assumes country weights are the same as in offline exports. "GDP weights" uses country weights from 2008 nominal GDP.

Figure 15: Actual and simulated number of export destinations (Peru – offline exports)



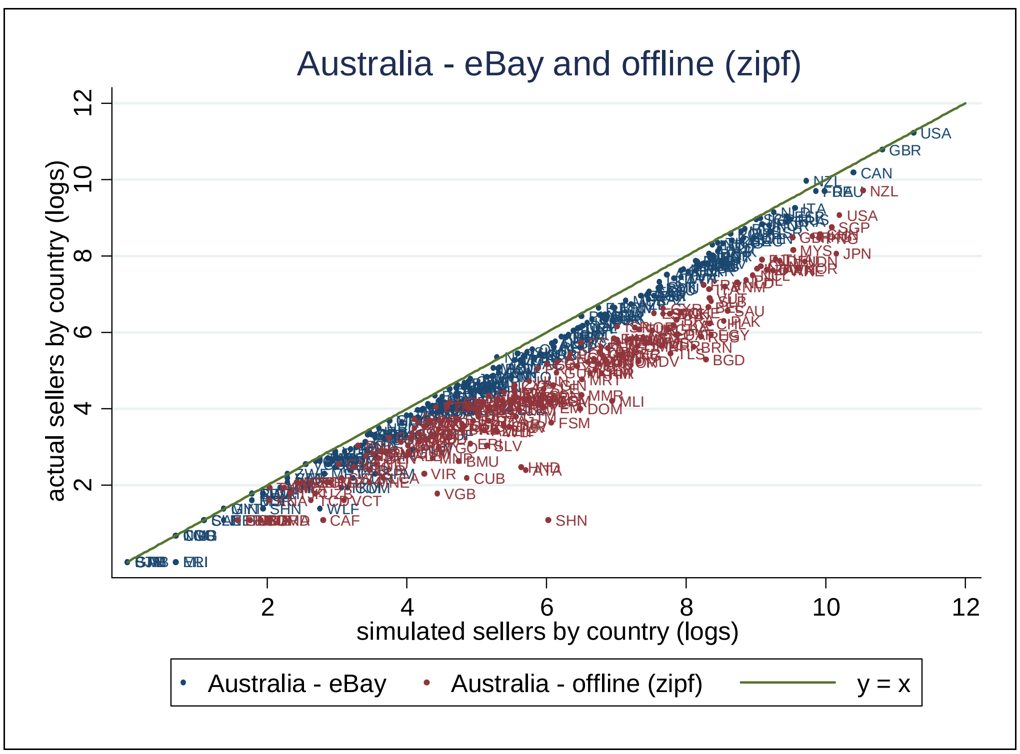
Source: Calculation by the author using Peru data for 2009. "Export weights" assumes country weights are the same as in offline exports. "GDP weights" uses country weights from 2009 nominal GDP.

Figure 16: Actual and simulated number of sellers per country: US



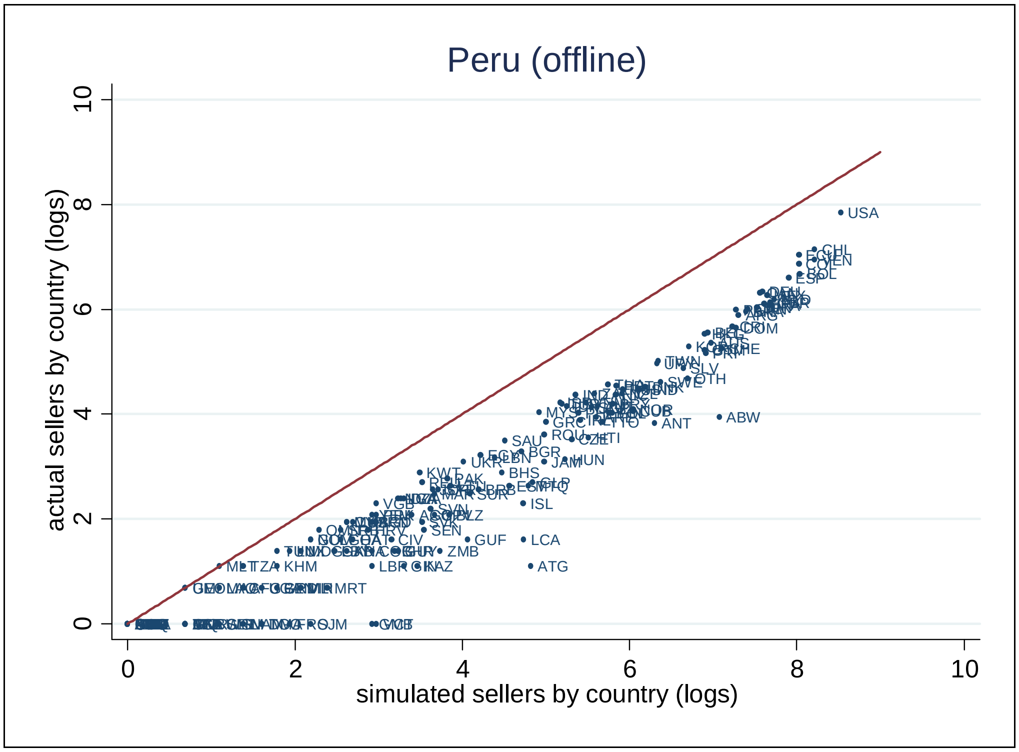
Source: Calculation by the author using eBay data for 2010. The red line shows y=x.

Figure 17: Actual and simulated number of sellers per country: Australia (eBay and offline)



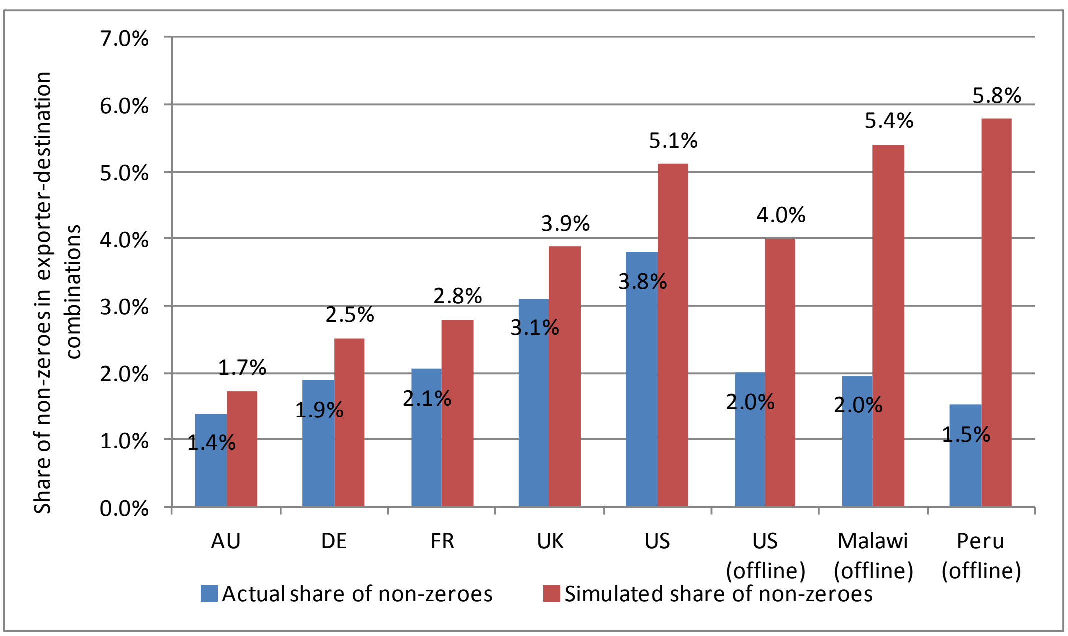
Source: Calculation by the author using eBay Australia data for 2011 and offline data for FY 2010/11 from ABS (2012. Calculations for offline firms are made assuming that the known number of total export transactions (ca. 3 million) is Zipf-distributed across the known number of exporters (ca. 38'600).

Figure 18: Actual and simulated number of sellers per country: Peru (offline)



Source: Calculation by the author using Peru for 2009. The red line shows y=x.

Figure 19: Share of non-zeroes in exporter-destination combinations – actual versus simulated



Source: Calculation by the author using eBay data for 2010 (US) and 2011 (others). US offline is from Armenter & Koren (2010). Malawi and Peru is calculated by the author from World Bank data.

Table 3: List of eBay SAP categories

SAP categories			
1	Antiques	22	Pottery & Glass
2	Baby	24	Sporting Goods
3	Books, Comics & Magazines	25	Sports Memorabilia
4	Business, Office & Industrial	26	Stamps
6	Food & Gourmet	28	Toys & Games
8	Auto - Parts	30	Musical Instruments
9	Clothes, Shoes & Accessories	31	Mobile & Home Phones
10	Coins	32	PC & Video Gaming
11	Collectables	33	Consumer Electronics - Audio
12	Computing	34	Consumer Electronics - Video
13	Consumer Electronics - Other	35	Art
14	Dolls, Doll Houses	36	Home Furnishing
15	Hobbies & Crafts	37	Health & Beauty
16	Home & Garden	38	Software
17	Jewellery & Watches	39	Home Appliances
18	DVDs, Film & TV	40	Cell Phones & Accessories
19	Music	42	Entertainment Memorabilia
20	Networking & IT	99	Everything Else
21	Photography		

Note: I dropped the SAP categories “travel”, “tickets”, “unknown” and “men’s fashion”. The latter category is practically not used.

Table 4: eBay and gravity: Online versus offline exports

Dependent variable	eBay exports	offline exports (all)	offline exports (eBay image)	eBay exports	offline exports (all)	offline exports (eBay image)
	(A1)	(A2)	(A3)	(A4)	(A5)	(A6)
Distance	-1.151***	-1.832***	-2.058***	-0.958***	-2.041***	-2.415***
	(0.0704)	(0.111)	(0.131)	(0.0870)	(0.136)	(0.173)
Common language (ethn.)				0.375**	0.176	0.0490
				(0.167)	(0.261)	(0.275)
Common language (off.)				0.951***	0.314	0.719**
				(0.201)	(0.264)	(0.314)
RTA				-0.130	-0.897***	-1.434***
				(0.136)	(0.225)	(0.282)
Common legal system				0.370***	0.214	0.183
				(0.112)	(0.135)	(0.180)
Common currency				0.618***	0.821**	1.247***
				(0.212)	(0.320)	(0.410)
Common border				-0.190	-1.259***	-1.991***
				(0.167)	(0.404)	(0.466)
Colonial link				1.056***	0.882***	1.243***
				(0.131)	(0.149)	(0.209)
Observations	6,039	6,039	6,039	6,039	6,039	6,039
R-squared	0.904	0.869	0.835	0.928	0.883	0.857
Fixed effects	importer-year, exporter-year			importer-year, exporter-year		
Zeroes dropped	yes, when eBay = 0 & offline = 0			yes, when eBay = 0 & offline = 0		

Robust standard errors in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1
Standard errors clustered at country-pair level.

Notes: "Offline exports" refers to exports of products similar to the ones traded on eBay. Destinations with no offline or eBay exports from *any* exporting country are excluded. These are mainly small territories. Apart from these cases, the number of zeroes in the data is very small (< 2%).

Table 5: eBay and gravity: Extensive and intensive margin

	eBay data						World Bank results (Cebeci et al., 2012, Table 9)					
Dependent variable	eBay exports	number of exporters	exports per exporter	eBay exports	number of exporters	exports per exporter	offline exports	number of exporters	exports per exporter	offline exports	number of exporters	exports per exporter
	(A1)	(A2)	(A3)	(A4)	(A5)	(A6)	(A7)	(A8)	(A9)	(A10)	(A11)	(A12)
Distance	-0.860***	-0.813***	-0.0468**	-1.014***	-0.875***	-0.139***	-1.616***	-1.223***	-0.393***	-2.118***	-1.535***	-0.582***
	(0.0755)	(0.0716)	(0.0186)	(0.0606)	(0.0526)	(0.0182)	(0.046)	(0.028)	(0.028)	(0.052)	(0.034)	(0.032)
Importer GDP	1.119***	0.963***	0.156***				0.845***	0.498***	0.347***			
	(0.0290)	(0.0261)	(0.00806)				(0.019)	(0.011)	(0.012)			
Exporter GDP	1.215***	0.979***	0.236***				1.169***	0.855***	0.314***			
	(0.0753)	(0.0720)	(0.0180)				(0.027)	(0.016)	(0.019)			
Observations	5,045	5,045	5,045	5,838	5,838	5,838	2780	2780	2780	2980	2980	2980
R-squared	0.654	0.631	0.213	0.907	0.922	0.576	0.553	0.615	0.251	0.742	0.834	0.439
Fixed effects		year			importer-year, exporter-year			none			importer, exporter	

	additional control variables					
Dependent variable	eBay exports	number of exporters	exports per exporter	eBay exports	number of exporters	exports per exporter
	(B1)	(B2)	(B3)	(B4)	(B5)	(B6)
Distance	-0.302***	-0.284***	-0.0189	-0.812***	-0.718***	-0.0936***
	(0.113)	(0.108)	(0.0290)	(0.0745)	(0.0611)	(0.0276)
Importer GDP	1.129***	0.970***	0.158***			
	(0.0261)	(0.0237)	(0.00845)			
Exporter GDP	1.204***	0.976***	0.227***			
	(0.0740)	(0.0712)	(0.0182)			
Common language (ethn.)	0.853***	0.764***	0.0894*	0.594***	0.597***	-0.00263
	(0.215)	(0.201)	(0.0532)	(0.143)	(0.127)	(0.0612)
Common language (off.)	0.641**	0.546*	0.0948	0.649***	0.611***	0.0385
	(0.324)	(0.298)	(0.0725)	(0.170)	(0.151)	(0.0741)
RTA	1.447***	1.493***	-0.0454	0.00487	-0.0688	0.0737
	(0.228)	(0.222)	(0.0547)	(0.120)	(0.104)	(0.0483)
Common legal system	0.0252	0.0194	0.00585	0.274***	0.176**	0.0975**
	(0.235)	(0.214)	(0.0486)	(0.0984)	(0.0895)	(0.0384)
Common currency	0.965***	0.711***	0.254***	0.759***	0.704***	0.0549
	(0.260)	(0.225)	(0.0773)	(0.145)	(0.127)	(0.0533)
Common border	0.404	0.0708	0.333***	-0.0704	-0.186	0.116
	(0.355)	(0.318)	(0.102)	(0.152)	(0.141)	(0.0787)
Colonial link	0.350	0.500**	-0.150***	0.993***	0.921***	0.0722
	(0.214)	(0.208)	(0.0542)	(0.115)	(0.0993)	(0.0476)
Observations	5,045	5,045	5,045	5,838	5,838	5,838
R-squared	0.703	0.686	0.227	0.937	0.955	0.580
Fixed effects		year			importer-year, exporter-year	

Robust standard errors in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1
Standard errors clustered at country-pair level.

Table 6: eBay and gravity: extensive and intensive margin: Comparison with US - Bernard *et al.* (2007)

	Bernard et al. (2007), Table 6				eBay US (2006-10 data)				
	(A1)	(A2)	(A3)	(A4)	(A5)	(A6)	(A7)	(A8)	(A9)
									(A7)*(A8)
	total export value	number of exporters	number of exported products per exporter	export value per product per exporter	total export value	number of exporters	number of exported products per exporter	export value per product per exporter	exports per exporter
Distance	-1.36***	-1.14***	-1.06***	0.84***	-1.652***	-1.733***	-0.0753*	0.157***	0.0815
	0.17	0.16	0.15	0.19	(0.243)	(0.241)	(0.0391)	(0.0470)	(0.0700)
Importer GDP	0.98***	0.71***	0.52***	-0.25***	1.129***	0.983***	0.0521***	0.0940***	0.146***
	0.04	0.04	0.03	0.04	(0.0520)	(0.0514)	(0.00629)	(0.00983)	(0.0121)
Observations	175	175	175	175	878	878	878	878	878
R-squared	0.82	0.74	0.64	0.25	0.702	0.655	0.483	0.167	0.295

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at importer level (eBay only).

Table 7: eBay and gravity: extensive and intensive margin: Comparison with Australian offline data

	eBay Australia exports						Offline exports Australia (ABS, 2012)					
	exports by value			exports by number of transactions			exports by value			exports by number of transactions		
Dependent variable	eBay exports	number of exporters	exports per exporter	eBay exports	number of exporters	exports per exporter	offline exports	number of exporters	exports per exporter	offline exports	number of exporters	exports per exporter
	(A1)	(A2)	(A3)	(A4)	(A5)	(A6)	(A7)	(A8)	(A9)	(A10)	(A11)	(A12)
Distance	-1.850***	-1.571***	-0.279***	-1.678***	-1.571***	-0.107**	-3.399***	-2.714***	-0.685***	-3.790***	-2.714***	-1.076***
	(0.326)	(0.319)	(0.0794)	(0.356)	(0.319)	(0.0530)	(0.211)	(0.159)	(0.172)	(0.189)	(0.159)	(0.0972)
Importer GDP	1.131***	0.988***	0.143***	1.135***	0.988***	0.146***	0.965***	0.590***	0.375***	0.740***	0.590***	0.150***
	(0.0643)	(0.0577)	(0.0187)	(0.0688)	(0.0577)	(0.0162)	(0.0407)	(0.0318)	(0.0334)	(0.0418)	(0.0318)	(0.0203)
Observations	663	663	663	663	663	663	663	663	663	663	663	663
R-squared	0.657	0.633	0.160	0.638	0.633	0.456	0.770	0.763	0.441	0.739	0.763	0.408

additional control variables

	exports by value			exports by number of transactions			exports by value			exports by number of transactions		
Dependent variable	eBay exports	number of exporters	exports per exporter	eBay exports	number of exporters	exports per exporter	offline exports	number of exporters	exports per exporter	offline exports	number of exporters	exports per exporter
	(B1)	(B2)	(B3)	(B4)	(B5)	(B6)	(B7)	(B8)	(B9)	(B10)	(B11)	(B12)
Distance	-1.319***	-1.162***	-0.157	-1.150***	-1.162***	0.0116	-3.144***	-2.353***	-0.791***	-3.306***	-2.353***	-0.953***
	(0.357)	(0.337)	(0.102)	(0.383)	(0.337)	(0.0725)	(0.214)	(0.166)	(0.188)	(0.196)	(0.166)	(0.108)
Importer GDP	1.160***	1.008***	0.151***	1.163***	1.008***	0.154***	0.985***	0.608***	0.377***	0.758***	0.608***	0.150***
	(0.0644)	(0.0607)	(0.0192)	(0.0693)	(0.0607)	(0.0135)	(0.0424)	(0.0319)	(0.0368)	(0.0437)	(0.0319)	(0.0234)
Common language (ethn.)	0.730	0.499	0.231**	0.516	0.499	0.0174	0.972***	0.381	0.591**	0.673**	0.381	0.293**
	(0.547)	(0.509)	(0.111)	(0.592)	(0.509)	(0.0899)	(0.352)	(0.272)	(0.272)	(0.336)	(0.272)	(0.119)
Common language (off.)	-0.423	-0.274	-0.148	-0.171	-0.274	0.103	-0.812*	-0.00298	-0.809***	-0.384	-0.00298	-0.381**
	(0.796)	(0.754)	(0.141)	(0.866)	(0.754)	(0.123)	(0.442)	(0.337)	(0.290)	(0.428)	(0.337)	(0.168)
RTA	0.383	0.310	0.0732	0.504	0.310	0.194	-0.131	0.227	-0.358	0.434*	0.227	0.207
	(0.399)	(0.410)	(0.180)	(0.449)	(0.410)	(0.155)	(0.300)	(0.185)	(0.265)	(0.243)	(0.185)	(0.145)
Common legal system	0.684	0.539	0.145	0.552	0.539	0.0127	0.489*	0.402*	0.0870	0.652**	0.402*	0.250*
	(0.536)	(0.542)	(0.109)	(0.613)	(0.542)	(0.0818)	(0.289)	(0.213)	(0.193)	(0.277)	(0.213)	(0.140)
Common currency	2.183***	1.491***	0.692***	2.312***	1.491***	0.821***	1.394***	0.757**	0.637**	1.325***	0.757**	0.568***
	(0.575)	(0.539)	(0.124)	(0.568)	(0.539)	(0.164)	(0.380)	(0.296)	(0.250)	(0.347)	(0.296)	(0.126)
Colonial link	1.301**	0.963**	0.338*	1.305*	0.963**	0.342	0.962**	0.929***	0.0333	1.051***	0.929***	0.122
	(0.552)	(0.423)	(0.189)	(0.698)	(0.423)	(0.321)	(0.423)	(0.270)	(0.232)	(0.339)	(0.270)	(0.118)
Observations	663	663	663	663	663	663	663	663	663	663	663	663
R-squared	0.686	0.653	0.189	0.662	0.653	0.512	0.795	0.809	0.468	0.783	0.809	0.444

Robust standard errors in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at importer level (eBay only).

Data used: eBay 2006-2010, offline FY 2006/07 to FY 2010/11. Some small destinations for which detailed offline data is not available were dropped.

Table 8: Share of eBay sellers and offline firms that are exporting

Country			Offline
	≥ USD 10'000	≥ USD 100'000	
AU	77	97	2
DE	97	100	59*
FR	98	100	15 to 67*
UK	96	99	28*
US	97	99	4 to 15*

All figures are in percent. Sources: eBay: Own calculations based on eBay data (US: 2010, AU: 2012, others: 2011). The eBay data for "all sellers" should be interpreted carefully because the data is not fully comparable for small sellers in different eBay countries. "Offline": US: Bernard *et al.* (2007). France (all firms): Eaton *et al.* (2009). France (large firms only), Germany and UK: Mayer & Ottaviano (2007). Australia: ABS (2012, online data appendix, Table 5.1). An asterisk (*) indicates that "offline" figures refer to subsets of large firms only.

Table 9: Actual and simulated number of export destinations for 8 countries

Country	Actual average number of destinations	Expected number of destinations				Share of exporters appearing only once in data
		one obs = one txn	one txn = 100'000 USD	one txn = 10'000 USD	one txn = 1'000 USD	
ALB	1.5	2.5	2.3	5.8	13.7	43%
BFA	2.2	3.1	2.9	6.6	15.3	43%
BGR	2.4	5.6	3.6	10.4	25.5	37%
GTM	2.5	5.2	3.9	10.9	25.5	28%
JOR	2.4	2.5	2.5	8.0	22.0	48%
MEX	2.2	4.3	3.9	9.1	21.2	41%
SEN	3.0	5.4	4.5	12.6	28.4	34%
YEM	2.6	4.0	3.6	10.0	23.0	37%

Source: Data received from World Bank. Countries are Albania, Burkina Faso, Bulgaria, Guatemala, Jordan, Mexico, Senegal and Yemen, using the latest available year.

The calculation of the expected number of destinations was made using importing country shares from actual exports. The underlying data is at year-exporter-product-destination level. Each observation can thus include an unknown number of transactions. I calculated the expected number of destinations by assuming either that each observation contains only one transaction, or that each transaction has a certain value. For example, when I assume that each transaction is worth USD 10'000, I calculate the number of transactions for each observation by dividing the export value by 10'000 and rounding to the next number (I round up when one would otherwise get zero transactions). When assuming that transactions are small, the expected number of different destinations increases.