Dissecting Gravity: From Cargo Shipments to Country-Level Trade Flows

Dmitry Livdan[1], and Vladimir Sokolov[2], and Amir Yaron[3]

July 23, 2018

[1]Haas School of Business, University of California, Berkeley. email: livdan@haas.berkeley.edu.

[2]ICEF, Higher School of Economics, Moscow. email: vsokolov@hse.ru.

[3]Wharton School, University of Pennsylvania and NBER. email: yarona@wharton.upenn.edu.

Abstract

Using a novel and comprehensive data set for Russian exporters, this paper shows that individual shipments and their transportation method are the key for understanding trade flows. This neoclassical view focuses on information and penetration costs as the underlying source explaining shipment costs and therefore trade patterns–and is distinct from more recent views that focus on aggregate firm-level trade flows. Russian firms ship more valuable goods to farther distances but the rate of such shipments declines with distance even faster – a feature consistent with Chaney's (2016) sufficient conditions for gravity to hold at the country level. We show that transportation modes, which reflect trade costs, capture distance's role in explaining trade flows and shipments. We further show that empirically trade flows at the firm level do not satisfy Chaney's (2016) condition. Consistent with our empirical evidence, we theoretically contribute to show how to account for the fact Chaney's (2016) conditions are satisfied at the shipment level, but are not at the firm level, yet gravity holds at the country level. Together our evidence highlights that shipments are the key disaggregate economic variable of interest for understanding trade flows.

1 Introduction

The gravity equation is one of very few empirical relations in economics that has withstood the test of time and a variety of different methodologies. In its classical form, introduced half a century ago by Tinbergen (1962), gravity states that bilateral trade flows between any two countries increase in both countries' GDPs and decrease in the distance between them – i.e., a distance elasticity of minus one.[1] Despite its central role in the empirical implementation of the gravity equation, distance still lacks economic underpinning. While it is understood to proxy for trade frictions, the frictions themselves and their economic link to distance are yet to be established. An empirical analysis in this regard would require individual cargo-level data complemented by the precise geographic locations of the individual shipments' origin and destination. This would allow the information absorbed by distance to be traced out as cargo shipment data is aggregated first to firm and then further to country levels.

Empirical studies examining the role played by firms in international trade has substantially broadened and deepened in recent years; however, the vast majority of analyses has been based on "one-sided" data – i.e., data that identifies either the exporter or importer but not both. Recently, studies utilizing "two-sided" trade transaction data have begun to emerge.[2] However, these two-sided studies generally suffer from incomplete and imprecise information.[3]

In this paper, we study how the classic country-level gravity relationship emerges by employing a highly detailed and precise Russian customs level data that allows empirical analysis of trade flows and distance at different levels of data aggregation. The Russian customs data contains virtually all daily declaration forms during 2011, and allows us to identify the unique cargo shipments between

[1]The distance elasticity of minus one has been recently reaffirmed by Head and Mayer (2014) who survey 161 published papers with a total of 1,835 independent estimates of the distance elasticity over a century and a half of data.

[2]Bernard et. al. (2012) provides a comprehensive overview and expanded reference list of studies with one sided data. Two-sided trade data has been analyzed for Argentina and Chile (Blum et al. (2010)), Chile and Colombia (Blum et al. (2013)), Colombia (Benguria (2014)), Costa Rica, Ecuador, and Uruguay (Carballo et al. (2013)), Norway (Bernard et al. (2014)), and the United States (Pierce and Schott (2012); Dragusanu (2014); Eaton et al. (2014); Kamal and Sundaram (2013); Monarch (2014); Kamal and Monarch (2015); Heise (2016); Monarch and Schmidt-Eisenlohr (2016)).

[3]For example, U.S. importing firms with shipments above $2,000 are required to complete U.S. Customs and Border Protection Form 7501 which entails forming the Manufacturer ID (MID) for the foreign counterparty of the transaction rather than providing full information about its name and address. These MIDs are constructed by the importers from the name and address on the commercial invoices by applying an algorithm outlined in the Customs Directive No. 3550-055. Therefore, one needs to decode MIDs to access the identity and geographic location of the foreign firm which is prone to both Type I, incorrect decoding, and Type II, incorrect construction in the first place, errors. Kamal and Monarch (2015)) provide an excellent discussion on this subject.

domestic exporting firms and foreign recipient firms. Importantly, the customs forms data contains the precise identification of firms at both sides of the transaction which enables us to calculate the exact door-to-door distance between them as well as build the networks of exporting and importing firms. Furthermore, the high dimensionality of our cargo level data (based on one or more customs forms on the same day from the same exporter to the same recipient) allows us to employ time, product and firm level fixed effects for exporters and recipients, thus enabling us to identify the role of distance and trade-flows through variation within exporters and recipients.

Our main set of results on gravity concerns the cargo shipment level of trade between firms. We find that the exact distance between firms has no explanatory power for variation of cargo shipment values if one controls for the modes of cargo transportation or includes the recipient firm fixed effects. In fact we find that the unit cargo value *increases* with distance across recipients for a given exporter or alternatively for a given importer across their Russian exporters. Moreover, we find a negative relationship between distance and the variety of products in a cargo shipment. Our findings uncover the previously unexplored intricate dimension of international trade between firms at the cargo shipment level. These findings are consistent with fixed and variable costs that dominate the economics of trade at its basic micro cargo level.

Next, we aggregate the cargo shipments data to the *firm* level and obtain trade values for matched exporter-recipient firm pairs over the whole year. We find that at the firm level the effect of distance on the value of exports is ambiguous. For a given exporter, the export value increases with distance across recipients, albeit insignificantly. On the other hand, and in line with traditional gravity prediction, we find that conditional on an importer the dollar value of trade is decreasing with exporters' distance. In addition, we find strong evidence that both weight of goods exported and goods' variety decrease with the distance between firms.

The question that naturally arise is why doesn't distance appear to play a major role at the cargo level, yet it does at the firm level? Our analysis shed light on this and supports the view that most of the exporter's decision making is done at the individual cargo level. Exporters respond to their recipient's demand consisting of both the quality of the commodity and the timing (expediency) of its delivery. The timing part of the recipient's demand is captured to a large degree by the transport method, while the majority of the unobserved quality component of the recipient's demand is captured by the recipient fixed effects.[4] Export costs are best explained by the transportation

[4]For example, the recipient may want to buy both cheap and expensive wine and needs the expensive wine to be delivered overnight. The recipient plans to bottle the cheap wine on-site and therefore schedules it to be shipped in

method, the number of shipments, and exporter fixed effects. Consequently, conditional on the recipient demand, the exporter chooses the number of cargoes to ship to each recipient, and for each cargo its content. Thus, distance has very low explanatory power at the cargo level since both exporter and recipient fixed effects together with the individual cargo's transport method subsume most of the information about the export economic costs contained in distance.

Once cargo data is aggregated up to firm level data, cargo's individual transport method is no longer available and some important information about transport costs and recipients demand is lost. Some of this lost information is recovered by using distance and the most frequently employed transport method. As a result, distance becomes statistically significant at the firm level as it proxies for some of the information about the transport costs contained in the transport method.

Finally, when we aggregate firms' data to the *country* level we find significant support for the classical gravity regression of a negative relation between distance and value of trade. Furthermore, the gravity coefficient significantly increases when one uses the average exact door-to-door distance between firms within a country instead of the surrogate capital-to-capital distance commonly used in the literature. Since our data allows us to calculate the number of cargo shipments sent to each country and a number of firms involved in trade with each country we are able to decompose the value of exports to each of the countries along the *intensive* and *extensive* margins. Here we find that average value per shipment per firm sent to a country increases with distance to that country, while number of cargo shipments sent to a country and number of Russian exporting firms trading with a country declines with distance.

Next, we investigate the reason the data aggregates at the country level to the classical gravity equation. First, we show empirically that Russia's firm size distribution is Generalized Exponential – one that violates Chaney's (2016) first sufficiency conditions for gravity which requires firm sizes to be Pareto distributed. We then show theoretically how to account for the fact Chaney's (2016) conditions are satisfied at the shipment level, but are not at the firm level, yet gravity holds at the country level.

We show that our evidence that Russian firms ship more valuable goods to farther distances but at a shipment rate that declines with distance even faster is a feature consistent with Chaney's (2016) sufficient conditions for gravity to hold at the country level. That is, we show that in order to yield the correct aggregation at the country level, the cargo's intensive margin should increase slower with distance than the decline in the average number of shipments. We further show that

cisterns by either ship or train depending on the destination, while the expensive wine is shipped by plane.

given that Chaney's conditions hold at the cargo level, they can not, by construction, be satisfied at the firm level (as a weighted average of Pareto is not Pareto distributed)–rationalizing our firm level evidence. Therefore, as long as the conditions are satisfied at the cargo level, the firm level aggregation may yield either positive, negative or no relationship between distance and value traded at the firm level, making that evidence much less informative regarding the country-level gravity. Finally we use our approach to demonstrate quantitatively that the country-level distance elasticity of value arises directly from the cargo-level data without performing aggregation.

2 Data description

We use recently released proprietary data on Russian exports. Relative to prior work the data provides several unique features. First, it is given at custom form level and thus provide the most micro-shipment level information as several custom forms may aggregate to represent an 'economic' shipment unit. As discussed earlier, this allows us to use a "bottom-up" approach and analyze 'gravity' at different levels of economic aggregation ranging from cargo, to firm, to country levels. Second, the data identifies the recipient and hence facilitate precise measures of distance.

Our dataset is collected by the customs authority and contains 1,613,878 customs declarations submitted by exporting firms during 2011. Specifically, each customs declaration reports information on the date when the declared product left the country, 10-digit HS code product classification, net weight in kilograms, mode of transportation, F.O.B. (Free on board) value of exported product in US dollars, names and postal addresses of the foreign recipient firm and the Russian exporting firm.

Alongside the invoice value of exports in the original invoicing currency and the agreed delivery conditions, the declarator is also required to report the US dollar value of exports at the current exchange rate adjusted to the F.O.B. delivery conditions at the last port of departure of the Eurasian Customs Union(EACU).[5,6,7] This adjustment results in bringing all Russian exports to a common

[5]US dollar is the predominant currency of invoicing with 47.5% of all customs declarations being settled in USD. Russian ruble is the second most popular currency of invoicing and occupies 30.3% of all customs declarations in our data. Export declarations in Euro occur in 21.7% of all export contracts. Invoicing in other currencies takes place in less than 1% of export contracts.

[6]All customs declarations also report the delivery conditions of the export transactions. Nearly half of Russian exports (48.6%) are contracted under F.C.A. (Free carrier) conditions. C.P.T. (Carriage Paid to destination) and D.A.F. (Delivery at Frontier) respectively occupy 13.1% and 14.7% of all declarations.

[7]In 2011 EACU included Russia, Belarus and Kazakhstan. The official guidelines on customs declarations reporting stipulate the following rule for the F.O.B. adjustment: 1) if the export contract specifies the delivery city within the EACU other than the last port of departure the transport costs to the border of the EACU are added to the exports value; 2) if the export contract specifies the delivery city outside of the EACU the transport costs from the EACU

denominator and as a result we have the exports data in producer prices at the border of the EACU.

Identification of the foreign recipient firms has been a serious challenge in empirical work preventing researchers from studying the international trade at the firm-to-firm level.[8] We exploit a key advantage of our customs data and uniquely identify the foreign recipient firms by their reported names and addresses. The reported postal addresses of exporters and recipients enable us to use Google geocoder to obtain the geo-coordinates of all exporters and recipients in the data-set and calculate the exact distance on a sphere between each pair of counterparties. It is worth noting that recipient firms with the same name but different addresses receive different IDs.[9] Overall, our dataset results in 47,483 unique firm-recipients and 19,307 unique Russian exporting firms. These data spans large public firms as well as individual entrepreneurs.

Our goal is to understand at what level of data aggregation distance becomes important in explaining the variation in the cross-section of export values. To that extent we use three levels of data aggregation. This allows us to trace out the specific information absorbed by distance as data is aggregated. Daily cargo shipments comprise the most disaggregated data. Next, we collapse daily cargo shipments data within each exporter-recipient pair to construct firm-to-firm pairwise annual aggregates. Finally, we aggregate the firm-to-firm data to the country level. Table 1 provides summary statistics of our data for all three levels of data aggregation.

We obtain daily cargo shipments by aggregating all the reported customs forms filed on the same day by the same Russian exporter to the same recipient through the same customs office and using the same freight method. The exporter decides on the content of the cargo shipment it sends on a given day to a foreign recipient, and then files as many customs forms per shipment as required by the Federal Customs Service of Russia (FCSoR, www.russian-customs.org).[10] For each cargo shipment we construct the value (in US$) of the shipment, VALUE, and net weight (in kg) of the shipment, EXP.WGHT, by adding up values and weights from all customs forms for all products included in the shipment. Finally, we define the exported cargo extensive margin, VARIETY, as the number of unique HS10 codes across all customs forms in the cargo shipment.

---

border the city of delivery are subtracted from the exports value.

[8] Recently a number of papers were able to collect the firm-to-firm international trade data where identities of foreign recipients are obtained using decoding tools with the various degrees of precision (e.g., Bernard et al. (2016), Carballo et al. (2016), Kamal and Monarch (2016), Kamal and Sundaram (2016)).

[9] For example, Halliburton, Texas has a different ID from Halliburton, Australia. Appendix C provides further description of the algorithm that we used for assigning the unique IDs to foreign firms recipients.

[10] According to the regulations imposed by the FCSoR the exporter files separate customs form for each product categorized by the HS10 classification.

Table 1: Descriptive statistics

The table reports descriptive statistics at the custom declaration and shipment levels (Panel A) as well as firm and country (Panel B) in our sample for 2011.

VARIABLES	N	Mean	Std	Min	p50	Max
Cargo level:						
Exported value (VALUE, US$)	595,251	612.6	8,882.4	0.10	34.05	1,633,121.3
Net exported weight (EXP.WGHT, kgs)	595,251	1,148.1	22,649.6	0.0	43.5	9,397,329.9
Number of goods per shipment (VARIETY)	595,251	2.79	9.83	1	1	1,422
Air freight indicator (AIR.T)	595,251	0.06	0.23	0	0	1
Rail freight indicator (RAIL.T)	595,251	0.36	0.48	0	0	1
Auto freight indicator (AUTO.T)	595,251	0.40	0.49	0	0	1
Ship freight indicator (SHIP.T)	595,251	0.18	0.38	0	0	1
Other freight indicator (OTHER.T)	595,251	0.01	0.09	0	0	1
Firm level (per firm-firm pair):						
Exported value (VALUE, US$1,000)	87,085	4.19	121.61	1.03×10^{-4}	0.07	16,628.8
Net exported weight (EXP.WGHT, tons)	87,085	7.86	209.56	0.00	0.05	32,581.07
Number of shipments	87,085	6.68	16.73	1	2	504
Number of recipients per exporter (N(IpE))	19,307	4.55	11.16	1	2	306
Number of exported goods per exporter (EXP.VARIETY)	19,31	6.59	19.37	1	2	909
Total imported weight per recipient (IMP.WGHT, tons)	47,483	14.41	333.29	0.00	0.03	35,802.9
Number of exporters per recipient (N(EpI))	47,483	1.85	3.22	1	1	193
Number of imported goods per recipient (IMP.VARIETY)	47,483	4.51	13.58	1	1	894
Exact firm-to-firm distance on a sphere (DIST(FtF), km)	86,511	2,626.16	2,294.76	0.00	2008.7	18,747.2
Exporter's revenue (SALES, Million rubles)	12,479	2,506.8	36,695.8	0.00	128.67	3,206,865
Exporter's asset turnover (ASSET.TRN)	12,477	4.56	135.86	0.00	1.69	14,125.3
Country level:						
Average exported value (VALUE, US$1M)	181	2,026.3	5,432.7	0.08	125.7	39,317.9
Average exported weight (EXP.WGHT, Million kg)	181	3,789.2	10,419.2	0.00	194.5	67,329.2
Number of cargo shipments per country	181	3,275	9,706	1	205	81,806
Average firm-to-firm distance (DIST(FtF), km)	181	4,685.2	2,404.5	815.1	4223.4	11,881.3
Capital-to-capital distance (DIST(CtC), km)	179	5,948.0	3,644.9	763.4	6185.2	16,774.5
Number of exporters per country (EpC)	181	287	601	1	42	5,264
GDP (US$1M)	181	388,742.8	1,427,074.5	59.45	31,078.9	15,517,926
GDP per capita (US$, per capita)	181	18,249.4	24,352.6	350.3	7318.7	157,093
Neighboring country indicator (NNHBR)	181	0.08	0.27	0	0	1

Customs forms aggregate to 595,251 unique cargo shipments. The mean (median) VALUE is \$613 with the standard deviation equal to \$8,882. The mean (median) value of EXP.WGHT is 1,148 kg with the standard deviation equal to 22,650 kg. Both standard deviations are due to several extremely heavy and valuable cargoes (e.g., a ship weighing 9,400 tons and valued at \$1.6 million). Each cargo contains on average 2.79 different goods as indicated by the VARIETY, with the standard deviation equal to 10 goods. Once again, the standard deviation is high due to a number of cargoes containing up to several hundreds of good varieties with the maximum VARIETY equal to 1,422.

Each cargo shipment is delivered by a unique freight method. We designate a dummy variable for each individual freight method: AIR.T for air freight, RAIL.T for railroad freight, SHIP.T for freight by water, AUTO.T for automobile freight, and OTHER.T for all other freight methods.[11] Automobile (39.9%) and railroad (35.9%) freight methods are used most often. It is not surprising since Russia's trade is mostly concentrated in Europe and Asia both of which can be easily reached by well-developed railroad and highway networks. 17.8% of exports are delivered by water, 5.7% by air, and less than 1% using other freight methods.[12]

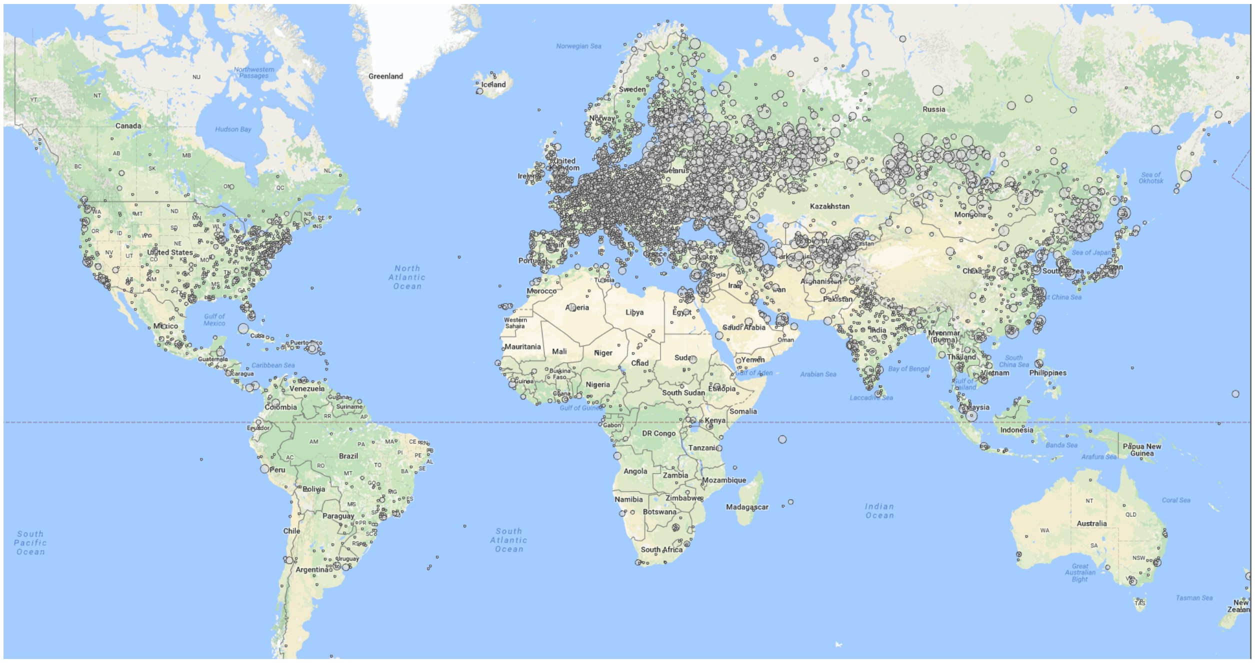
Next we collapse daily cargo shipments data within each exporter-recipient pair to construct the firm-to-firm pairwise annual aggregates. Altogether we have 87,085 exporter-importer relationships. Figure 1 shows the geography of Russian exports over the World (Panel A) and Europe (Panel B). Each circle represents and individual exporting/importing firm with the size of the circle indicating the total export/import value. The majority of trade partners of Russian exporting firms is located in Europe, both East and West, and Asia. While Russian exporting firms trade a lot with their traditional partners located in the former USSR and Warsaw block member countries, they also have established numerous relations with Western European, Asian, North and South American, African, Australian and Asian firms. Holland, Germany, Ukraine, and China have the highest value of exports from Russia, with the US being not far behind.

Table 1 provides key descriptive statistics of the firm-to-firm pairwise annual trade. Russian exporters ship on average 6.7 (16.7) cargoes per recipient with an average cargo value, VALUE, equal to \$4,192 (\$121,611) and an average cargo weight, EXP.WGHT, equal to 7.86 tons (209.6 tons), where standard deviations are reported in parentheses. High standard deviations indicate that a sizable fraction of exporters ships a large number of relatively inexpensive and lightweight

[11]These include but not limited to regular mail, personal delivery, pipe, etc.

[12]Russia has a direct access to the Arctic and Pacific Oceans as well as the Baltic Sea, the Black Sea, the Caspian Sea, and the Sea of Azov. The last four Seas allow indirect access to the Atlantic Ocean.

Panel A: World



Panel B: Europe

Figure 1: Geography of Russian trade
The geography of Russian trade across the world (Panel A) and Europe (Panel B). Individual firms are shown as circles with the circle's size indicating the total trade value.

cargoes.

Table 1 shows that Russian exporters have large trading networks abroad, with mean number of importers per each exporter, N(IpE), being equal to 4.5 – a number larger than foreign importers have with Russian firms with mean number of exporters per importer, N(EpI), being equal to 1.85.[13] On average, exporters ship larger variety of goods (EXP.VARIETY), 6.6, than the variety of goods (IMP.VARIETY) recipients receive from Russian exporters which is equal to 4.5.

Figure 2 sheds further light on the trading activity of Russian exporting firms. Panel A of Figure 2 uses log-log scale to show the degree distribution of the number of cargo shipments. 31% of exporters ship a single cargo per year with some exporters shipping up to 500 cargoes. Panel B of Figure 2 uses log-log scale to show the degree distribution of the variety of exported goods. 41% of exporters specialize in a single product with some exporters shipping up to 900 different goods. Panel C of Figure 2 uses log-log scale to show the degree distribution of the exporter-importer networks. Half of Russian exporters trade with a single foreign partner while small fraction of exporters trades with up to 300 foreign firms. All three distributions exhibit Pareto profile. Finally, Panel D of Figure 2 demonstrates that the average distance between trading partners is increasing with the number of trading partners. This is consistent with Chaney (2014) findings for French exporters.

The average exact distance (measured on a sphere) between firms is equal to 2,626km. It is greater than distance between Moscow and Berlin (1,608km), Moscow and Amsterdam (2,145km), close to the distance between Moscow and Paris (2,487km), and almost half the distance between Moscow and Beijing (5,793km). Some exporters have trading partners as far as 18,000km away. As we show below, this exact distance measure is informative relative to the commonly used capital-to-capital distance measure used in typical trade-distance studies.

The reported tax IDs of Russian exporting firms allows us to match their customs-level data with the firms' financial characteristics. We obtain annual values of total sales and total assets of exporting firms for the 2010-2011 period from the Ruslana database of Bureau Van Dijk. However, using Ruslana database significantly reduces the sample size of Russian exporters as a large number of them is missing 2011 sales data while having 2010 sales readily available. We, therefore, use either the available sales data, e.g., if only 2010(2011) sales data is available we use only 2010(2011) data, or the average between 2010 and 2011 sales which still reduces the sample from 19,307 to

[13]For comparison, values of these two variables reported by Bernard et al.(2017) for Norway are 9 and 2 respectively.

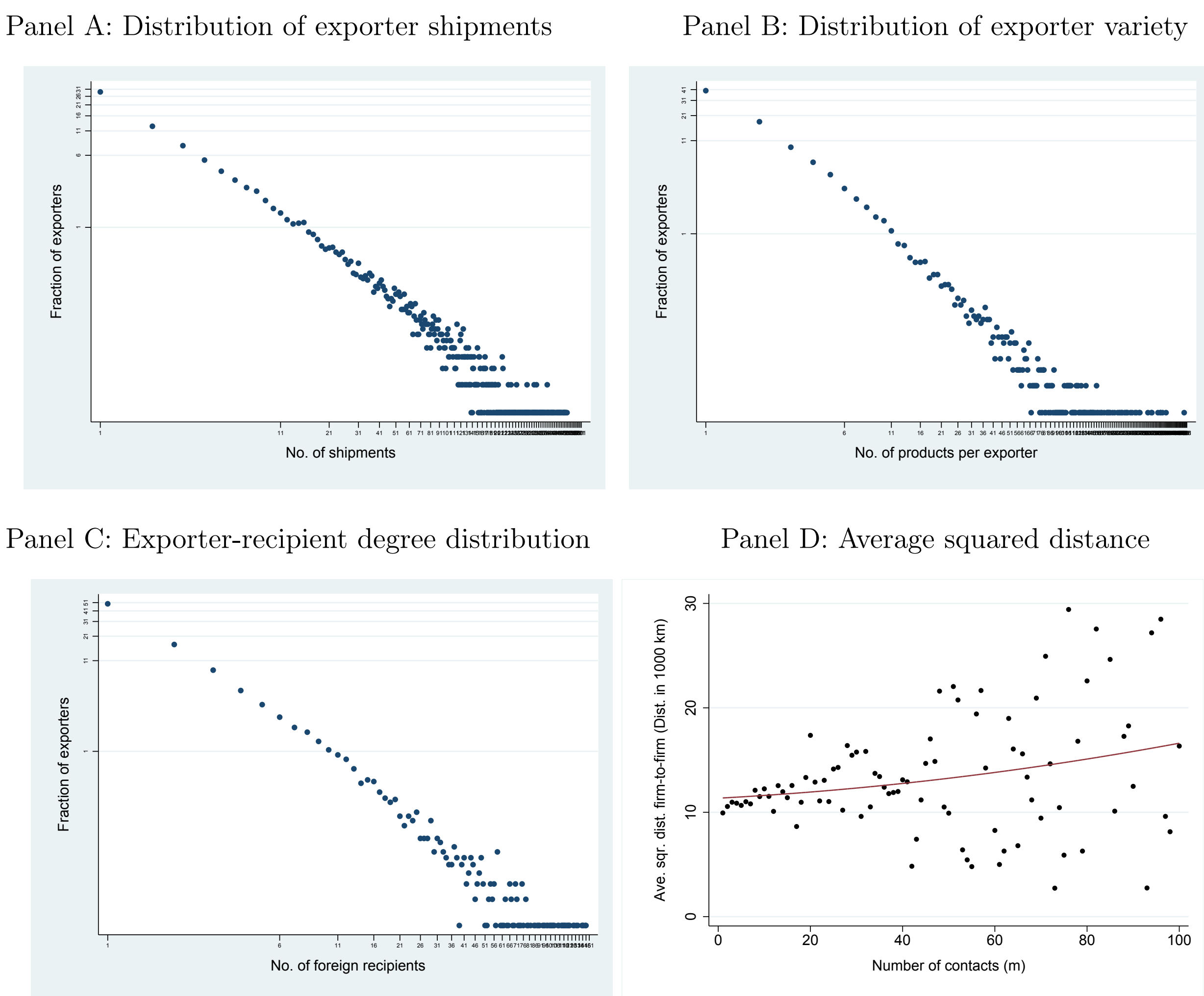


Figure 2: Exporter trading activity
Distribution of exporter shipments for the whole year (Panel A), the variety of shipped goods measured as the average number of WTO 10-digit good codes used by the exporter (Panel B), the degree distribution for exporter-recipient relations for the whole year (Panel C), and average squared distance from an exporter's contacts, among exporters with m contacts (Panel D). The last figure is Figure 2 in Chaney (2014). We use a log-log scale in Panels A–C.

12,479 exporting firms.[14] Total sale proxies for firm's size. The average exporter has annual sales, SALES, equal to 2.5 billion Russian rubles or about US$75 million (using 2011 exchange rate). In addition, we define asset turnover ratio, ASSET.TRN, as the ratio of SALES to the average of total assets held by a company at the beginning of the year and and at the year's end. It measures the efficiency of a company's use of its assets in generating sales revenue or sales income with a higher ratio implying more efficient deployment of company's assets. The average value of asset turnover for Russian exporters is quite high at 4.6 which is due to several outliers with very high asset turnover ratio. Firms in the 50th percentile have asset turnover ratio equal to 1.6.

[14]The results are robust to using either 2010 or 2011 sales.

We further collapse annual exporter-recipient pairwise data to construct annual country-level aggregates. The country-level data on GDP is obtained from the World Bank's WDI, the bilateral capital-to-capital distances between countries are obtained from CEPII.[15] Russian trade data contains 181 ISO country codes out of total of 249 Country Codes in the ISO Standard List. On average, 287 Russian exporting firms ship to a given country. Table 1 shows that Russian exporters ship on average 3,275 cargo shipments per year per country with an average value equal to $2 billion and average weight equal to 3.8 million tons. We report statistics for two alternative measures of country-to-country distances. The first one, DIST(FtF), is constructed by averaging exact distances between all Russian exporters and their trading partners from each country. The second one, DIST(CtC), is the distance between Moscow and the capital of each country. One interesting fact is that the average "exact" distance is almost 1,300km shorter than its capita-to-capital counterpart. Finally, we follow McCallum (1995) and Anderson and VanWincoop (2003), and introduce the "nearest-neighbor" dummy, NNHBR, which equals one for countries with a land border with Russia and equals to zero otherwise. The average value of NNHBR is equal to 0.077 which translates to 14 neighboring countries.[16]

3 Empirical analysis

This section reports our empirical findings. We start with daily cargo shipments as the most disaggregated data since at other aggregation levels information can be lost. A crucial characteristic of the cargo level data is the information about each cargo's unique freight method. There exists a sizable variation in freight costs allowing us to link the exported value to shipping transaction and time-opportunity costs both unique to each freight method. Since distance is usually assumed to be associated with trade costs, the unique freight method provides an alternative and more direct economic measure of such trade costs. We then proceed with the analysis of the firm-level and, subsequently, country-level data.

[15] http://www.cepii.fr/anglaisgraph/bdd/distances.htm

[16] These include Norway, Finland, Estonia, Latvia, Lithuania, Poland, Belarus, Ukraine, Georgia, Azerbaijan, Kazakhstan, Mongolia, North Korea and China, for a total of 14 land neighbors.

3.1 Cargo shipments

Model Specification and Hypothesis Development

Table 1 demonstrates that there exists a lot of variation in value, weight, and content variety of daily cargo shipments. Is variation in distance capable explaining these variations? To answer this question let's consider a competitive exporter shipping at cost a single commodity using N cargoes to a recipient located a distance D away.[17,18] Let's assume that the log-value of cargo i can be written as

$$\log(v_i(D)) = \log(F(D, \varepsilon_i)) + \log(W_i(D)). \tag{1}$$

Relation (1) accounts for two independent sources of heterogeneity across cargoes shipped to the same distance D: cargo's extensive margin measured by its weight $W_i(D)$ and cargo's intensive margin measured by its value-to-weight ratio.[19] We explicitly assume that the intensive margin, $F(D, \varepsilon_i)$, depends on distance D and cargo specific shock ε_i while cargo weight depends only on distance. Relation (1) is consistent, for instance, with a classic model of the spatial distribution of alternative production activities by Johann von Thünen (1826) with competitive exporters shipping at costs.

Value decomposition (1) is firstly motivated by the idea that costs of exporting a cargo unit may depend on its weight rather than the unit value (for example, the costs of exporting a canned product depend on the number of cans rather than quality of their content). It is secondly motivated by the "shipping the good apples out (while keeping the bad ones for internal consumption)" story originally proposed by Alchian and Allen (1964). This story alludes to the possibility that differences in value-to-weight ratios within the same product category may be explained by differences in product's unobserved quality proxied by ε_i. In this case, increases in shipping distance or reductions in the recipient's purchasing power may lead to a change in the composition of exports towards higher-value products – more valuable per unit of weight products make exports profitable despite incurring the fixed and variable trade costs of servicing the remote foreign market. For example, wine can be either expensive or cheap and the expensive wine is flown by air (high transportation costs) while cheap wine is transported by train (low transportation costs) to the same destination. Therefore,

[17] This corresponds to controlling for product fixed effects in the data.

[18] N can be equal, for example, to seven which is the average number of shipments in the data. Final conclusions are robust to the assumption of N being independent of distance.

[19] Baldwin and Harrigan (2011) use similar decomposition for the US country-level exports. They interpret value-per-kilo as a proxy for the unit price.

an additional interpretation of $F(D, \varepsilon_i)$ is that it captures cargo unit specific fixed and variable transport costs to distance D. Furthermore, this interpretation of $F(D, \varepsilon_i)$ leads us to conclude that the cargo's intensive margin, $F(D, \varepsilon_i)$, is a non-decreasing function of distance while holding ε_i fixed.

When distance D is fixed, the variation in cargo values for a given exporter-importer pair is characterized by W and ε according to the relation (1). Therefore, as long as N is large enough (the average value of N is equal to 6.68 in the data) distance should have very weak explanatory power at the level of individual exporter-importer pairs.[20] By changing D in (1) we, however, can capture the variation in cargo values across different exporter-importer pairs. In our data, we have 87,085 unique exporter-recipient pairs thus leading to a lot of distance-specific variability in cargo values. A significant fraction of this variation can be explained by using either exporter fixed effects or recipient fixed effects or both. With the exporter (recipient) fixed effects the problem reduces to explaining the variation in cargo values within exporter (recipient) across recipients (exporters). Table 1 shows that a Russian exporter ships on average to 4.55 recipients (unique distances) while a foreign recipient trades on average with 1.85 Russian exporters. Thus, there just does not exist enough variation in distance at the level of individual recipients to explain the variation in cargo values. Hypothesis 1 summarizes this discussion.

HYPOTHESIS 1: *Distance should have the explanatory power at the individual exporter level, i.e., when exporter fixed effects are used. Distance should have very weak explanatory power in explaining the variation in cargo value at the level of individual recipients, i.e., when recipient fixed effects are used.*

Relation (1) further suggests that distance should do a relatively better job in explaining variation in the cargo's intensive margin measured by the value-to-weight ratio, $\log(v_i(D)/W_i(D)) = \log(F(D, \varepsilon_i))$, than it does explaining variation in cargo values. This is because cargo weight which adds an extra dimension to heterogeneity of cargo values is removed from the cargo's intensive margin. However, while $F(D, \varepsilon_i)$ is a non-decreasing function of distance, the empirical sensitivity of the cargo intensive margin to distance would remain attenuated by variation in product's unobserved quality proxied for by ε_i when both the exporter and recipient fixed effects are used.

HYPOTHESIS 2: *Cargo's intensive margin, value-to-weight ratio, should be increasing with distance in the data at the level of either individual exporters or individual recipients. However, the*

[20]While in the data the individual exporter-importer pairs can be captured using a product of exporter-recipient fixed effects we, unfortunately, do not have enough data to implement it.

relationship may not be statistically significant when both the exporter and recipient fixed effects are used.

While we have assumed in (1) that cargo's weight directly depend on distance, the direction of this relation is a purely empirical question. For instance, it may be argued that variable transport costs, i.e., fuel, storage, labor, and others, increase faster with distance for heavier cargoes thus making cargo weight a decreasing function of distance. An alternative argument would be that there exist economies of scale for fixed transport costs, such as packaging and handling costs, in the case of the long-distance freight methods. For instance, fixed transport costs of shipping a standard container are the same whether the container is fully or half filled.[21] Thus exporters shipping to longer distances would ship fewer cargoes but these cargoes would be on average heavier than cargoes shipped nearby. Hypothesis 3 summarizes this discussion.

HYPOTHESIS 3: *The relation between cargo's weight and distance is moot in the absence of the freight method.*

Furthermore, since the intensive margin of cargo's value is increasing with distance, the extensive margin of cargo's value, weight, has to decline with distance faster than the intensive margin increases with it. We still adopt the null hypothesis that the value of exports declines with distance at the cargo level and let the data sort out whether cargo weight declines with distance fast enough.

The fact that both fixed and variable transportation costs are freight method specific imply that even with the exporter and recipient fixed effect a lot of remaining variation in cargo value as well as its intensive and extensive margins should be explained by the transportation dummies. Using transportation dummies requires us to select a reference freight method. In our empirical analysis we use the railroad freight as the reference transport method, as both western and eastern parts of Russia have well-developed railroad networks.[22] According to Roberts (1999) the average unit transport costs are the highest for parcels, typically shipped by air and ground courier services, and then decline as the freight method changes in the following order: light truck, truckload, unit-railcar, multi-railcar, unit train, and ship. The ranking between the rail and maritime freight methods does however depend on distance with former being cheaper for short to medium distances and later being cheaper for longer distances. The freight method ranking is reversed in the case

[21] According to the World Bank's Doing Business database (http://www.doingbusiness.org/data/exploretopics/trading-across-borders) the cost of exporting a standard container can be decomposed into border compliance and documentary compliance which include obtaining, preparing and submitting documents during port or border handling, customs clearance and inspection procedures. In 2011 the fixed cost of exporting the standard container from the Russian Federation was $2460.

[22] Our empirical results are robust to any alternative choice of the reference freight method.

of the loading capacity, i.e., barge on average transports the heaviest load while plane or courier vehicle transport the lightest load. Hypothesis 4 summarizes this part of the discussion.

HYPOTHESIS 4: *Lighter cargoes with higher intensive margin are shipped by either air or auto freight, while heavier cargoes with lower intensive margin are shipped by either rail or maritime freight.*

Furthermore, transportation dummies should retain a lot of explanatory power even at the individual exporter-recipient pair level characterized by a unique distance as they capture the variation in cargo value and its both intensive and extensive margins through the lens of transportation costs as summarized in Hypothesis 5.

HYPOTHESIS 5: *Freight dummies should subsume distance's explanatory power in explaining the variation in cargo value even at the level of individual exporters, i.e., when only exporter fixed effects are used.*

We perform our analysis using variants of the following three linear equations

$$
\begin{aligned}
z_{pert} &= \alpha_{1\tau} + \alpha_{1p} + \alpha_{1e} + \beta \log(DIST_{er}^2) + \delta \mathbf{I}_{pert}^T + \gamma \mathbf{X}_e + \epsilon_{pert}, && (2) \\
z_{pert} &= \alpha_{1\tau} + \alpha_{1p} + \alpha_{1r} + \beta \log(DIST_{er}^2) + \delta \mathbf{I}_{pert}^T + \eta \mathbf{X}_{r,d} + \epsilon_{pert}, && (3) \\
z_{pert} &= \alpha_{1\tau} + \alpha_{1p} + \alpha_{1e} + \alpha_{1r} + \beta \log(DIST_{er}^2) + \delta \mathbf{I}_{pert}^T + \gamma \mathbf{X}_e + \eta \mathbf{X}_{r,d} + \epsilon_{pert}. && (4)
\end{aligned}
$$

where as a response variable, z_{pert}, we use the log-characteristic of a cargo p shipped by a Russian exporter e to a foreign recipient r on date t located in a destination country d. Following our discussion, the set of cargo characteristics consists of cargo value, weight, value-to-weight ratio, and the number of unique goods in the cargo or variety. We have included variety as an alternative metric of the cargo's extensive margin. We are specifically interested in whether exporting firms tend to specialize or diversify when shipping long-distance and vice versa. A complementary question is to what degree the cargo's value depends on the variety of its content.

In (2–4) $\mathbf{X}_{r,d}$ denotes a vector of log-characteristics of the recipient firm r belonging to country d while $\mathbf{X}_e$ denotes a vector of log-characteristics of the exporting firm. $\log(DIST_{er}^2)$ denotes a log of the squared distance between exporting and recipient firms. We use the exact firm-to-firm distance in kilometers measured on a sphere across all specifications.[23] We square the distance thus

[23] When specifications (2-4) re-estimate with the distance measured between Moscow and the capital of the destination country, the results do not change. Thus while we do not report them here these results are available from the authors upon request.

adopting the null that $\beta = -1$[24] and making the distance elasticity equal to 2β.

As has already been mentioned, the multidimensional nature of our data allows us to employ multiple fixed effects in equations (2–4). $\alpha_{1\tau}$ is a time fixed effect with a unit of time equal to one month. We choose to include time fixed effects at the month frequency in order to control for possible seasonality, thus restricting identification to a within month variation in all variables. α_{1p} is a product fixed effect with product categories measured by 10-digit Harmonized System (HS10) product codes. Since the majority of cargoes contains multiple products (the average number of products is equal to 2.8) we construct product fixed effect using the highest valued good per cargo.[25] α_{1e} is the Russian exporter fixed effect and α_{1r} is the foreign recipient fixed effect.

Specification (2) allows us to examine the variation in cargo shipments both within-product and within-exporting firm across recipients. Specification (3) identifies β, δ and γ through variation across exporters by controlling for any time-invariant characteristics of the recipients through inclusion of the recipient firms fixed effects. Finally, in specification (4) we use both sets of firm fixed effects and thus control for unobserved characteristics of firms at both ends of a trade. Across all equations, the error term ϵ_{pert} has the interpretation of unmeasured factors such as, for instance, a recipient firm-specific supply or demand shock: for a given product the supply is randomly less costly (demand is randomly higher) from some recipients than others.

Specifications (2-3) include the observed characteristics of firms recipients and firms exporters, $\mathbf{X}_{r,d}$ and $\mathbf{X}_e$ respectively. We control for the recipient's trading network in Russia by including the total number of its Russian trading partners in 2011, N(EpI). We also control for recipient r's demand by including the total number of unique goods (HS10 codes) imported by r during the sample period from all Russian firms, IMP.VARIETY. We also include the recipient r's country of origin real GDP in order to control for the aggregate demand.

Vector $\mathbf{X}_e$ includes four exporting firm characteristics. We proxy for the exporter's financial quality with its annual sales, SALES, and asset turnover ratio, ASSET.TRN, both of which defined in the previous section. We control for exporter's unobserved supply shocks with the exported product variety, EXP.VARIETY. Finally, we control for the exporter's foreign trading network by including the number its foreign trading counter-parties, N(IpE).

$\mathbf{I}^T_{pert}$ denotes a vector of the cargo's freight method dummies measured relative to the rail

[24]Without loss of generality, we follow the classical Law of Gravity from Physics by considering that the value of exports declines with distance as distance squared at all aggregation levels, e.g., cargo, firm, and country.

[25]Our results are robust to alternative specifications such as value(weight)-weighing all codes.

freight. To understand how much variation in cargo's characteristics is explained by the freight methods beyond distance, exporter and recipient characteristics, and different fixed effects, we first run specifications (2-4) without freight dummies by setting $\delta = 0$ (Table 2) and then rerun them without imposing $\delta = 0$ restriction (Table 3).

Results without freight dummies

Table 2 reports our estimates when freight dummies are not included, $\delta = 0$. As a response variable we use the cargo value (Value) in US$ in Columns (1) through (3), the cargo's value-to-weight ratio (Value-per-Kilo) in Columns (4)-(6), the cargo weight in kilograms in Columns (7)-(9), and the cargo product variety in Columns (10)-(12). Columns 1, 4, 7, and 10 employ specification (2). Columns 2, 5, 8, and 11 employ specification (3), and specification (4) is used in Columns 3, 6, 9, and 12. Per each specification, Columns 4 and 7 combine to make up cargo value in Column 1 by the properties of ordinary least squares.

Distance elasticities of cargo value from Columns 1 through 3 are consistent with the Hypothesis 1. Column 1 indicates the without the recipient's fixed effects distance elasticity of cargo value is equal to 0.044 with the standard deviation of 0.01 thus making it both statistically and economically significant. When the recipient's fixed effects are included as in specifications (3) and (4), the elasticity of the cargo value to distance is neither statistically nor economically significant as shown in Columns 2 and 3 respectively. The sign of the distance elasticity remains positive in Columns 2 and 3, thus indicating that the cargo value is increasing with distance. Month, product, and exporter fixed effects explain 69% of the variation in cargo values, while exporter, recipient, and country-level characteristics explain only 2.3% of it. When the exporter fixed effects are switched to the recipient fixed effects the explanatory power of combined fixed effects improves to 76.7% while the explanatory power of characteristics reduces to 1.1%. The highest explanatory power of 80.5% is achieved in Specification (4) with all four fixed effects included.

In agreement with Hypothesis 2, the value-to-weight ratio is increasing with distance for all three specifications. Its distance elasticity is both statistically and economically significant in specifications (2) and (3), i.e., where either exporter or recipient fixed effects are used but not both, reported in Columns 4-5 of Table 2. The magnitudes are equal to 0.015 (Specification (2)) and 0.019 (Specification (3)) with the corresponding standard errors equal to 0.003 and 0.007. Both elasticities, while being on the lower end, are in line with the estimates reported by Martin (2012) for French exporters, Bastos and Silva (2010) for Portuguese exporters, Gorg et al. (2010) for Hungar-

Table 2: Cargo shipments without freight method

The value (weight) of cargo shipment shipped on day t by the exporter i to recipient j is calculated by adding values from all custom forms filed on day t by the exporter i as going to the recipient j. VARIETY is defined as the total number of different goods per cargo shipment as per WTO 10-digit good code. All other cargo shipment level values are calculated by aggregating values from all customs forms included in the individual cargo shipment. VARIETY.EXP(IMP) is defined as the total number of different goods shipped (received) by exporter (recipient) as per WTO 10-digit good code. N(IpE) is the size of the exporter's trading network, i.e., number of recipients the exporter i trades with. N(EpI) is the size of the recipient's trading network in Russia, i.e., number of Russian exporters the recipient i trades with. IMP.WGHT is the total weight imported by the recipient i. Double-clustered by exporter and recipient standard errors are reported in parentheses with significance levels indicated by * (5%), † (1%), ‡ (0.1%).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Value			Value-per-Kilo			Weight			Variety		
$\log(\mathrm{DIST}^2)$	0.044‡	0.016	0.014	0.015‡	0.019†	0.010	0.029†	-0.003	0.004	-0.014†	-0.025*	-0.001
	(0.010)	(0.014)	(0.017)	(0.003)	(0.007)	(0.006)	(0.010)	(0.014)	(0.016)	(0.005)	(0.010)	(0.007)
Recipient characteristics:												
log(N(EpI))	-0.037*			0.009*			-0.046†			-0.115‡		
	(0.015)			(0.004)			(0.015)			(0.010)		
log(IMP.VARIETY)	0.150‡			-0.007			0.158‡			0.182‡		
	(0.017)			(0.004)			(0.018)			(0.012)		
Exporter characteristics:												
log(N(IpE))		-0.073‡			-0.035‡			-0.038*			-0.057‡	
		(0.017)			(0.007)			(0.017)			(0.011)	
log(EXP.VARIETY)		0.056‡			0.016*			0.040*			0.194‡	
		(0.015)			(0.008)			(0.016)			(0.013)	
log(SALES)		0.082‡			0.001			0.081‡			0.010	
		(0.010)			(0.003)			(0.009)			(0.006)	
log(ASSET.TRN)		-0.048‡			-0.016‡			-0.032†			0.021*	
		(0.011)			(0.005)			(0.011)			(0.008)	
Country level variables:												
log(GDP)	0.057‡			0.001			0.056‡			0.013†		
	(0.007)			(0.002)			(0.007)			(0.005)		
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Product FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Exporter FE	YES		YES	YES		YES	YES		YES	YES		YES
Recipient FE		YES	YES		YES	YES		YES	YES		YES	YES
R^2	0.724	0.797	0.827	0.970	0.974	0.986	0.873	0.906	0.923	0.525	0.582	0.637
Adjusted R^2	0.714	0.777	0.805	0.969	0.972	0.984	0.868	0.896	0.914	0.507	0.542	0.591
Adjusted Within R^2	0.0233	0.0105	0.001	0.00186	0.00322	0.001	0.0199	0.00928	0.001	0.0369	0.0249	0.001
N	458,583	458,581	458,583	458,296	458,294	458,296	458,296	458,294	458,583	458,583	458,581	458,583

ian exporters, Harrigan, Ma, and Shlychkov (2015) for the U.S. exporters, and Manova and Zhang (2012) for Chinese exporters. Distance elasticity of export unit value found for these countries falls within [0.01, 0.19] interval and it is always statistically significant.

Also consistent with Hypothesis 2, when both exporter and recipient fixed effects are used (Specification (4) in reported in Column 6), the magnitude of the distance elasticity of the cargo's value-to-weight ratio is reduced to 0.01 with the standard error equal to 0.006 thus losing its statistical significance. This is because the exporter and recipient fixed effects capture almost all variation in the cargo's value-to-weight ratio as evidenced in Column 6 by the Adjusted R^2 and Adjusted Within R^2 equal to 98.4% and 0.1% respectively. The remaining 1.6% of variation is per our earlier discussion due to the unobserved product's quality which distance fails to capture. The comparable values for the Adjusted R^2 and Adjusted Within R^2 are equal to 97.2% and 0.3% respectively in the Specification (4) where only recipient fixed effects are included. Thus there exists some variation in the cargo's value-to-weight ratio at the recipient level across exporters which distance is capable of explaining and which is captured to a large degree by the exporter fixed effects.

In agreement with the Hypothesis 3 the relation between the cargo weight and distance is moot. Weight increases with distance at the exporter level across recipients (Specification (2) reported in Column 7), but decreases with distance at the recipient level across exporters (Specification (3) reported in Column 8). Weight once again increases with distance when both exporter and recipient fixed effects are used (Specification (4) reported in Column 9). The weight elasticity of distance is statistically significant only in Column 7. Exporter and recipient fixed effects once again capture almost all variation in cargo weight as evidenced in Column 9 by the Adjusted R^2 and Adjusted Within R^2 equal to 91.4% and 0.1% respectively.

Comparing distance elasticities of cargo's intensive, Columns 4-6, and extensive, Columns 7-9, margins as per value decomposition (1) helps to rationalize why cargo value is increasing with distance. First, cargo's intensive margin, the value-to-weight ratio, is increasing with distance across all three specifications. Second, cargo's weight is increasing with distance in Specifications 1 and 2. Third, cargo's weight decreases with distance in Specification 2 slower than its value-to-weight ratio increases with distance.

Finally, cargo's alternative extensive margin - the number of HS10 good codes per cargo or VARIETY - declines with distance across all three specifications reported in Columns 10-12 of Table 2, although the distance elasticity is not statistically significant when both exporter and recipient fixed effects are used (Column 12). The fact that the variety of exports declines with distance within exporting firms across recipients (Column 10) is consistent with the findings of Bertrand et. al. (2007) for the US manufacturing firms. One notable difference between cargo's variety and weight is that the exporter and recipient fixed effects together capture much less variation in the former than they capture in the later. The Adjusted R^2 in Column 12 is equal to 59.1% while it is equal to 91.4% in Column 9.

We delay the discussion of the relationship between the cargo characteristics and exporter, recipient, and country characteristics until next section where we present regression results with freight dummies added to Specifications (2–4).

Results with freight dummies

Next we test Hypothesis 4 and 5 regarding freight dummies. Table 2 reports our estimates from Specifications (2–4) with freight dummies being included, $\delta \neq 0$. The layout of Table 3 fully mimics the layout of Table 2.

In order to examine the Hypothesis 4 we analyze coefficients on freight dummies in Columns 3, 6, and 9. Column 3 reports regression results for the cargo value, while Columns 6 and 9 report regression results for two independent components of the cargo value from (1), the value-to-weight ratio and weight respectively. In all three cases both exporter and recipient fixed effects are included in the regressions.

In Column 3, the regression coefficients on air and auto freight dummies are negative and both statistically and economically significant. Negative sign on both dummies implies that more valuable cargoes are shipped by rail than by either air or auto transport. This is because in agreement with the Hypothesis 4 air and auto freight methods are used for lighter but more expensive per unit of weight cargoes. The regression coefficients on the air and auto freight dummies in Column 6 are equal to 0.616 and 0.049 respectively, while they are equal to -2.034 and -0.533 respectively in Column 9.

The regression coefficients on the ship freight dummy are positive in Columns 3 and 6, while the regression coefficient on the ship freight dummy is negative in Column 9. However, all three regression coefficients are not statistically significant thus implying that, in agreement with the Hypothesis 4, both ship and rail freight methods are used to transport cargoes comparable in terms of the value and weight. One possible explanation for this finding is that both rail and ship freight methods tend to utilize the same intermodal ISO containers.

Freight methods other than the plane, ship, or car, OTHER.T, are used to deliver cargoes similar to those delivered by either air or auto freight transport. While these cargoes have higher intensive margin than cargoes shipped by rail, they are on average less valuable since they tend to be on average significantly lighter than cargoes shipped by rail. A candidate example would be an object of art delivered by a private currier service.

We examine Hypothesis 5 by comparing the distance elasticity of cargo value estimated with (Table 3) and without (Table 2) freight dummies. Column 1 of Table 2 reports that the distance elasticity of value is equal to 0.044 with the standard error equal to 0.01 thus making it both statistically and economically significant. The same column also reports that the Adjusted R^2 and Adjusted Within R^2 are equal to 71.4% and 2.33% respectively. Column 1 of Table 3 demonstrates that with the freight dummies included into the regression the estimate of the distance elasticity of value is equal to -0.014 with the standard error equal to 0.009, thus making it neither statistically nor economically significant. At the same time the Adjusted R^2 and Adjusted Within R^2 are now equal to 72.9% and 7.40% respectively. These values indicate that while freight dummies explain only 1.5% extra variation in cargo value, they do indeed subsume a lot of explanatory power from distance, as it is the only explanatory variable which changes its sign and loses its statistical significance when freight dummies are included in the regression.

Columns 7–9 of Table 3 indicate that the cargo weight is decreasing with distance when controlling for the freight method. This finding lays out support to the idea that the total weight shipped depends on the unit transport costs which in turn depend on the freight method. For instance, wine can be shipped long-distance by a converted oil tanker with low unit costs and then bottled locally at the destination, or it can be bottled at the origin and shipped nearby by car or train both of which have higher unit costs than the tanker. Without controlling for the freight method, heavier on average wine cargoes are shipped long-distance since the tanker can carry more wine and it is cheaper per kilogram to do so. However, the exported weight declines with distance within each freight method since the unit transport costs increase with distance for all freight methods. Comparing Adjusted Within R^2 from Column 9 of Tables 2 and 3 shows that freight dummies explain a total of 2.76% variation in cargo weight.

Overall, these results provide robust evidence that cargo's freight method embeds important information about cargo value and its both intensive, value-to-weight ratio, and extensive, weight, margins.[26] For example, diamonds are light, have extremely high value-to-weight ratio, and are

[26]As a robustness check we have included the interaction of freight dummies with distance in specifications (2–4). The coefficients on freight dummies do not change when we include their interactions with distance and we, therefore,

shipped by air regardless of the destination. Iron ore is cheap and is shipped in large volumes by train to nearby destinations and by ship to farther destinations.

Furthermore, our results support the notion that the freight method is one of the choice variables of the value-maximizing exporting firm. Once the exporting firm chooses its trade network thus fixing the destinations, it responds to demand shocks by selecting the total export value (revenue) to each destination and then minimizes the freight costs to each destination by selecting the number of cargoes, the content (weight and variety) and the freight method for each cargo. Once, however, data is aggregated to the firm level, the precise information about the freight method is lost. We thus argue in the next section that distance is an imperfect proxy for the freight methods used by the individual exporter-recipient pairs.

Freight dummies do not add extra explanatory power in the case of cargo variety as both Adjusted R^2 and Adjusted Within R^2 are the same in Column 12 of Tables 2 and 3. However, the regression coefficients on the individual freight dummies are statistically significant. Column 12 of Table 3 shows that cargoes shipped by air have less variety that cargoes shipped by rail. Cargoes shipped by either the auto or ship freight have more variety than cargoes shipped by rail. The regression coefficient on the OTHER freight dummy is not statistically significant.

Next, we discuss the effect of recipient and exporter trade networks as well as other characteristics on the export cargo value, unit value, weight, and variety. Value elasticities to recipient characteristics are all statistically and economically significant across specifications (2) and (3). The elasticity of value to the recipient Russian trade network size, N(EpI), is equal to -0.032, implying that doubling the size of the recipient's trade network leads to a 3.2% decline in the exported value. Recipients with larger Russian trade networks tend to import lighter cargoes with less variety per cargo as as evident from columns 7-9 of Table 3 respectively. On the other hand, the elasticity of cargo's intensive margin to N(EpI) is close to zero as depicted in Column 4, thus leading to a negative relation between the export cargo value and the size of the recipient's Russian trade network.[27]

Interestingly enough, recipients with larger Russian trading networks tend to export larger variety of Russian products, as indicated by 60% correlation between log(N(EpI)) and log(VARIETY.IMP). Firms can increase the overall extensive margin of their imports by either increasing the variety per

do not report these results. They are, however, available from authors upon request.

[27]Table 4 provides extra support to this statement by showing that the cargo value declines with the size of the recipient's trading network in Russia.

Table 3: Cargo shipments with freight method

Same as Table 2 but with freight method dummies. XXX.TRANS is the freight dummy with rail freight serving as the reference. Double-clustered by exporter and recipient standard errors are reported in parentheses with significance levels indicated by * (5%), † (1%), ‡ (0.1%).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Value			Value-per-Kilo			Weight			Variety		
log($DIST^2$)	-0.014	0.007	0.005	0.014‡	0.018†	0.010	-0.028†	-0.011	-0.005	-0.017‡	-0.021*	0.001
	(0.009)	(0.014)	(0.017)	(0.003)	(0.007)	(0.006)	(0.009)	(0.014)	(0.016)	(0.005)	(0.010)	(0.007)
Transportation dummies:												
AIR.T	-1.544‡	-1.432‡	-1.417‡	0.724‡	1.076‡	0.616‡	-2.269‡	-2.509‡	-2.034‡	-0.184‡	-0.194‡	-0.221‡
	(0.093)	(0.093)	(0.124)	(0.063)	(0.083)	(0.064)	(0.093)	(0.089)	(0.108)	(0.046)	(0.037)	(0.041)
SHIP.T	0.191‡	0.041	0.008	0.016	0.047‡	0.013	0.176‡	-0.005	-0.005	0.126‡	0.094‡	0.074‡
	(0.052)	(0.043)	(0.037)	(0.011)	(0.012)	(0.008)	(0.052)	(0.044)	(0.037)	(0.032)	(0.017)	(0.017)
AUTO.T	-0.546‡	-0.533‡	-0.484‡	0.062‡	0.091‡	0.049‡	-0.608‡	-0.624‡	-0.533‡	0.096‡	0.108‡	0.083‡
	(0.052)	(0.040)	(0.043)	(0.010)	(0.020)	(0.013)	(0.050)	(0.042)	(0.045)	(0.021)	(0.024)	(0.023)
OTHER.T	-0.237	-0.334*	-0.094	0.320‡	0.437‡	0.190‡	-0.556	-0.773‡	-0.281	-0.165*	-0.117	-0.152
	(0.266)	(0.153)	(0.175)	(0.075)	(0.083)	(0.048)	(0.288)	(0.200)	(0.190)	(0.083)	(0.071)	(0.097)
Recipient characteristics:												
log(N(EpI))	-0.032*			0.009*			-0.041†			-0.109‡		
	(0.014)			(0.004)			(0.015)			(0.010)		
log(IMP.VARIETY)	0.125‡			-0.002			0.127‡			0.178‡		
	(0.016)			(0.004)			(0.017)			(0.012)		

Continued

Table 3: Cargo shipments with freight method—*Continued*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Value			Value-per-Kilo			Weight			Variety		
Exporter characteristics:												
log(N(IpE))		-0.077‡			-0.033‡			-0.044†			-0.056‡	
		(0.016)			(0.007)			(0.016)			(0.011)	
log(EXP.VARIETY)		0.051‡			0.018*			0.033*			0.194‡	
		(0.014)			(0.008)			(0.015)			(0.013)	
log(SALES)		0.081‡			0.001			0.079‡			0.009	
		(0.009)			(0.003)			(0.009)			(0.006)	
log(ASSET.TRN)		-0.048‡			-0.015†			-0.032†			0.020*	
		(0.010)			(0.005)			(0.010)			(0.008)	
Country level variables:												
log(GDP)	0.050‡			0.001			0.049‡			0.010		
	(0.006)			(0.001)			(0.007)			(0.005)		
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Product FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Exporter FE	YES		YES	YES		YES	YES		YES	YES		YES
Recipient FE		YES	YES		YES	YES		YES	YES		YES	YES
R^2	0.738	0.802	0.830	0.971	0.975	0.986	0.882	0.910	0.926	0.527	0.583	0.638
Adjusted R^2	0.729	0.782	0.808	0.970	0.973	0.984	0.877	0.901	0.916	0.510	0.543	0.591
Adjusted Within R^2	0.0740	0.0344	0.0175	0.0334	0.0484	0.0183	0.0899	0.0569	0.0276	0.0424	0.0278	0.00211
N	458,583	458,581	458,583	458,296	458,294	458,296	458,296	458,294	458,296	458,583	458,581	458,583

cargo or by importing different commodities from different Russian exporters with less variety per cargo, or both. Column 10 from Table 3 yields a VARIETY elasticity to the recipient's network size, N(EpI), of -0.11 which is both statistically and economically significant. Therefore, recipients with larger Russian trading networks, while importing commodities with on average similar intensive margin, increase the overall variety of their imports by adding an additional exporter instead of increasing the individual cargo's variety. This evidence points towards a trade-off between the freight and other per cargo costs versus fixed and variable costs of having larger trade network.[28]

Recipients preferring higher extensive margin of trade, VARIETY.IMP, import more valuable and heavier cargoes with elasticities equal to 0.125 and 0.127 respectively (both elasticities are statistically and economically significant). The elasticity of cargo's value-to-weight ratio to VARIETY.IMP is close to zero, thus leading to a positive relation between the cargo value and VARIETY.IMP.

Overall, recipients use two different channels to capture the same variety. Recipients finding it cheaper to make their trade network larger than to pay higher transport costs per cargo control the overall product variety margin by increasing the size of their trade network while reducing the average variety and weight per cargo. Recipients facing the opposite trade-off choose to have smaller trade networks while importing heavier cargoes with more variety.

Several characteristics of Russian exporters help explain variation in export value, unit value, weight, and variety. The most intriguing characteristic is the size of exporter's network, N(IpE), as it has not been previously used in gravity-type regressions. Russian exporters with larger trade networks export less valuable cargoes, with the elasticity equal to -0.077, cargoes with the lower intensive margin, with the elasticity equal to -0.033, and lighter cargoes, with the elasticity equal to -0.044. All three elasticities are statistically and economically significant. In addition, exporters with larger networks tend to ship cargoes with less variety. When combined with the results for the recipient's characteristics, this evidence can be rationalized as follows. Russian exporters with larger trade networks export mostly to recipients with larger trade network in Russia who tend to import lighter cargoes with less variety, which in turn are less valuable.

The balance-sheet level exporter characteristics, log(SALES) and log(ASSET.TRN), proxy for exporter's size and financial quality. More financially sound exporters ship heavier more valuable cargoes with higher intensive margins. Similar evidence for has been found by Harrigan, Ma, and

[28]Table 4 provides extra support to this hypothesis. It shows that the Russian exporters ship less shipments to recipients with large trading network in Russia.

Shlychkov (2015) for the U.S. and Manova and Zhang (2012) for China. However, log(SALES) does not explain variation in the cargo extensive margin, while VARIETY elasticity to log(ASSET.TRN) is positive and statistically significant at 10% implying that exporters with less asset turnover ship cargoes with less variety. The exporters' extensive margin of trade, VARIETY.EXP, is positively related to cargo value, unit value, and weight.

Finally, we find a statistically significant effect of the market size, measured by the log real GDP, on all cargo characteristics: larger markets export more valuable (elasticity of 0.05), heavier (elasticity of 0.05) cargoes with more variety (elasticity of 0.01). This market size effect conforms with greater demand for higher quality goods from larger markets story, previously discussed by Baldwin and Harrigan (2009), Bastos and Silva (2010), and Manova and Zhang (2012).

In summary, our results show that the variation in the cargo value within the same product category at the individual exporter-recipient pair level characterized by a unique distance is almost fully explained by the freight dummies. Freight dummies proxy for transportation costs thus capturing the unobserved product's quality. Once cargo-level data is aggregated to the annualized firm-level exports, some of the information contained in the freight dummies is lost. In the next section we investigate how much of this information, if any, is re-captured by the distance.

3.2 Firm-level aggregation

Hypothesis development and specification discussion

In this section we aggregate our data to the firm-to-firm pairwise trade which is arguably the most important trade level as all key decisions about trade flows are made at it. We use the following procedure to construct annual aggregates for each exporter-recipient pair. First, we calculate the total number of cargoes shipped by a Russian exporter e to a foreign recipient r throughout the year, N_{er}. Then for any cargo characteristic Z_{pert} we construct its annual firm-level counterpart, Z_{er}, as

$$Z_{er} = \sum_{p=1}^{N_{er}} \sum_{t=1}^{365} Z_{pert} = N_{er} \cdot \left(\frac{1}{N_{er}} \sum_{p=1}^{N_{er}} \sum_{t=1}^{365} Z_{pert} \right) = N_{er} \cdot \overline{Z_{er}}, \tag{5}$$

where $\overline{Z_{er}}$ is an individual cargo characteristic averaged across all cargoes shipped by a Russian exporter e to a foreign recipient r during the whole year. N_{er} and $\overline{Z}_{er}$ are the respective extensive and intensive margins of the firm-to-firm annual trade characteristic Z_{er}. This results in firm-to-firm matched data set which allows us to study the variation in export activity across and within firms.

Once the data is aggregated to the annual level, the *distribution* of each trade characteristic Z_{pert} consisting of N_{er} observations collapses to its mean $\overline{Z}_{er}$ for each exporter-recipient pair $\{e, r\}$ characterized by a unique distance, $DIST_{er}$. With this extra data dimension being removed, distance has to gain more power in explaining variation in the annualized firm-to-firm data than it did in the case of individual cargoes.

The number of cargoes, N_{er}, also absorbs some valuable information about transport costs. Results from Table 3 provide support to an idea that exporters shipping medium to long distances are taking advantage of the economies of scale for fixed transport costs. Such economies of scale are specific for the rail and water freight methods both allowing for standard shipping containers and bulk shipment. Saving on fixed transport costs allows for the long distance shipment of goods with low value-to-weight ratios, otherwise cost-prohibited to be shipped long distance, in train-load and ship-load lots. By construction, the number of such cargoes has to be small for the majority of exporter-recipient pairs. Cargoes with high value-to-weight ratios are mostly shipped long distance by air. Since air freight has the highest fixed and variable transport costs the number of cargoes shipped by air has also to be relatively small for the majority of exporter-recipient pairs. Hypothesis 6 summarizes this discussion.

HYPOTHESIS 6: *The number of cargoes per the exporter-recipient pair is decreasing with distance.*

Our next hypothesis is based on a standard NPV argument closely related to the Alchian and Allen (1964) story from the previous section. If transport costs increase with distance, the average cargo value per the exporter-recipient pair, $\overline{V}_{er}$, which has to be no less than the transport costs to at least break even, has to increase with distance.

HYPOTHESIS 7: *The average cargo value per the exporter-recipient pair is increasing with distance.*

The further implication of data aggregation is that it phases out several key explanatory variables for which we have to find imperfect substitutes. First, the variation in trade characteristics can no longer be within the individual product category. Instead, we now study the variation in trade characteristics at the industry level where by industry we designate the HS2 code of the product with the highest annual exported value for each exporter-recipient pair. Second, freight dummies capturing the variation in transport costs across individual cargoes are no longer available. Some of the lost information about the transport costs can be recaptured by dummies for the most frequently used freight method per each exporter-recipient pair[29] as well as by their interactions with distance. Some of the lost information on the transport costs is now also captured by distance thus further

[29]Our results are robust to either value- or weight-weighing when selecting the freight method.

increasing its explanatory power. Finally, since aggregated data is annual, monthly fixed effects are no longer available.

Motivated by the relation (5), we use five response variables characterizing the annual trade by the Russian firm e and recipient r located in a destination country d: the total value of exports, the average value per shipment, the total weight, the average weight per shipment, and the number of shipments. Linear projections of the natural log of each of these five characteristics, $z_{er} = log(Z_{er})$, are given by

$$\begin{aligned} z_{er} &= \alpha_{1i} + \alpha_{1d} + \alpha_{1e} + \beta \log(DIST_{er}^2) + \delta_0 \mathbf{I}_{er}^T + \delta_1 \mathbf{I}_{er}^T \times \log(DIST_{er}^2) + \eta \mathbf{X}_r + \epsilon_{erid}, & (6) \\ z_{er} &= \alpha_{1i} + \alpha_{1r} + \beta \log(DIST_{er}^2) + \delta_0 \mathbf{I}_{er}^T + \delta_1 \mathbf{I}_{er}^T \times \log(DIST_{er}^2) + \gamma \mathbf{X}_e + \epsilon_{erid}, & (7) \\ z_{er} &= \alpha_{1i} + \alpha_{1r} + \alpha_{1e} + \beta \log(DIST_{er}^2) + \delta_0 \mathbf{I}_{er}^T + \delta_1 \mathbf{I}_{er}^T \times \log(DIST_{er}^2) + \epsilon_{erid}, & (8) \end{aligned}$$

where α_{1i} are industry fixed effects, α_{1d} are the destination country fixed effects, and $\mathbf{I}_{er}^T$ is a vector of dummies for the most frequently used freight method. Specifications (6)-(8) share the same set of firm-level characteristics, $\mathbf{X}_e$ and $\mathbf{X}_r$, as well as the exporter, α_{1e}, and the recipient, α_{1r}, fixed effects with specifications (2)-(4). Since recipient fixed effects subsume the explanatory power of the destination country GDP at the annual frequency, we have excluded it from specifications (6)-(8).

Results

Table 4 reports our results. Our main finding is that the annual firm-to-firm export value increases with distance within exporters (Column 1) and it declines with distance within recipients (Column 2) as well as when both exporter and recipient fixed effects are included in the regression (Column 3). The estimates are equal to 0.029 in Column 1, -0.095 in Column 2, and -0.06 in Column 3. The distance elasticity of value is not statistically significant in Column 1, while it is statistically significant at 1% in Column 2 and at 5% level in Column 3.

The relation (5) applied to the total trade value is helpful for understanding the mechanics behind this finding. According to (5), the aggregate trade value between two firms is a product of the average cargo value and the number of cargoes shipped. Hypothesis 6 states that the number of cargoes per the exporter-recipient pair is decreasing with distance while Hypothesis 7 states that the average cargo value is increasing with distance. Therefore, the aggregate trade value between two firms would decline with distance if the intensive trade margin - average cargo value - does not increase with distance faster than the extensive trade margin - the number of cargoes - declines

Table 4: Firm-level exports

Left hand side variables are trade characteristics between the recipient j and Russian exporter i. Values (quantities) are calculated by aggregating daily cargo shipments from the exporter i to the recipient j. VARIETY.EXP(IMP) is defined as the total number of different goods shipped (received) by exporter (recipient) as per WTO 10-digit good code. N(IpE) is the size of the exporter's trading network, i.e., number of recipients the exporter i trades with. N(EpI) is the size of the recipient's trading network in Russia, i.e., number of Russian exporters the recipient i trades with. The transportation method is defined as the most frequently used by the exporting firm i. Railroad shipping is used as a reference. Double-clustered by exporter and recipient standard errors are reported in parentheses with significance levels indicated by * (10%), † (5%), ‡ (1%).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	Value			Value-per-Cargo			Weight			Weight-per-Cargo			# of Cargoes		
$\log(\text{DIST}^2)$	0.029	-0.095‡	-0.060†	0.097‡	0.017	0.028*	-0.024	-0.121‡	-0.088‡	0.044*	-0.009	0.000	-0.067‡	-0.111‡	-0.087‡
	(0.023)	(0.022)	(0.027)	(0.018)	(0.014)	(0.017)	(0.028)	(0.027)	(0.030)	(0.023)	(0.020)	(0.019)	(0.014)	(0.015)	(0.020)
Transportation method:															
AIR.T	-1.655‡	-1.439‡	-1.523‡	-1.450‡	-1.184‡	-1.283‡	-3.098‡	-3.961‡	-2.964‡	-2.899‡	-3.711‡	-2.731‡	-0.205‡	-0.257‡	-0.249‡
	(0.093)	(0.127)	(0.179)	(0.077)	(0.099)	(0.132)	(0.134)	(0.197)	(0.239)	(0.118)	(0.174)	(0.196)	(0.045)	(0.056)	(0.089)
AIR.T×$\log(\text{DIST}^2)$	0.034	0.102*	0.079	-0.047*	0.001	-0.018	0.087†	0.214†	0.148	0.008	0.115	0.054	0.080‡	0.100‡	0.098†
	(0.033)	(0.059)	(0.081)	(0.026)	(0.047)	(0.063)	(0.042)	(0.084)	(0.098)	(0.035)	(0.073)	(0.081)	(0.017)	(0.027)	(0.039)
SHIP.T	0.475‡	0.398‡	0.407†	0.629‡	0.592‡	0.529‡	0.434‡	0.348†	0.399†	0.589‡	0.542‡	0.522‡	-0.156†	-0.207†	-0.140
	(0.096)	(0.123)	(0.159)	(0.088)	(0.096)	(0.110)	(0.108)	(0.140)	(0.170)	(0.102)	(0.114)	(0.121)	(0.069)	(0.081)	(0.108)
SHIP.T×$\log(\text{DIST}^2)$	-0.043	-0.020	-0.011	-0.147‡	-0.136‡	-0.137‡	-0.024	-0.039	-0.011	-0.129‡	-0.154‡	-0.138‡	0.104‡	0.117‡	0.129‡
	(0.031)	(0.045)	(0.055)	(0.027)	(0.031)	(0.038)	(0.035)	(0.047)	(0.058)	(0.032)	(0.035)	(0.042)	(0.021)	(0.030)	(0.036)
AUTO.T	-0.353‡	-0.586‡	-0.472‡	-0.378‡	-0.491‡	-0.474‡	-0.595‡	-0.990‡	-0.641‡	-0.620‡	-0.894‡	-0.642‡	0.026	-0.097‡	-0.001
	(0.063)	(0.058)	(0.076)	(0.054)	(0.042)	(0.046)	(0.082)	(0.085)	(0.090)	(0.074)	(0.071)	(0.063)	(0.037)	(0.036)	(0.051)
AUTO.T×$\log(\text{DIST}^2)$	-0.069‡	0.060†	-0.006	-0.059‡	0.047‡	0.011	-0.058*	0.044	-0.027	-0.049*	0.032	-0.010	-0.011	0.011	-0.020
	(0.025)	(0.024)	(0.030)	(0.020)	(0.016)	(0.019)	(0.033)	(0.030)	(0.034)	(0.027)	(0.023)	(0.024)	(0.015)	(0.017)	(0.023)
OTHER.T	-0.939‡	-0.982‡	-0.808‡	-0.842‡	-0.792‡	-0.762‡	-2.037‡	-2.642‡	-1.587‡	-1.923‡	-2.444‡	-1.536‡	-0.096	-0.191†	-0.046
	(0.279)	(0.268)	(0.268)	(0.226)	(0.219)	(0.215)	(0.407)	(0.413)	(0.333)	(0.367)	(0.373)	(0.286)	(0.079)	(0.080)	(0.106)
OTHER.T×$\log(\text{DIST}^2)$	0.000	0.175	0.285†	-0.033	0.092	0.219*	0.029	0.300*	0.341†	-0.005	0.224	0.284*	0.030	0.073*	0.056
	(0.092)	(0.112)	(0.144)	(0.076)	(0.089)	(0.120)	(0.120)	(0.169)	(0.171)	(0.106)	(0.155)	(0.147)	(0.033)	(0.043)	(0.060)

Continued

Table 4: Firms—*Continued*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	Value			Value-per-Cargo			Weight			Weight-per-Cargo			# of Cargoes		
Recipient characteristics:															
log(N(EpI))	-0.084‡			-0.041‡			-0.052†			-0.009			-0.042‡		
	(0.020)			(0.011)			(0.022)			(0.014)			(0.014)		
log(IMP.VARIETY)	0.370‡			0.147‡			0.320‡			0.097‡			0.222‡		
	(0.018)			(0.010)			(0.020)			(0.012)			(0.012)		
Exporter characteristics:															
log(N(IpE))		-0.023			-0.066‡			0.107‡			0.064‡			0.043‡	
		(0.018)			(0.012)			(0.021)			(0.015)			(0.012)	
log(EXP.VARIETY)		0.100‡			0.030‡			-0.000			-0.070‡			0.070‡	
		(0.017)			(0.011)			(0.022)			(0.016)			(0.011)	
log(SALES)		0.176‡			0.099‡			0.176‡			0.099‡			0.077‡	
		(0.009)			(0.006)			(0.011)			(0.008)			(0.006)	
log(ASSET.TRN)		-0.044‡			-0.054‡			-0.029*			-0.039‡			0.009	
		(0.013)			(0.008)			(0.015)			(0.010)			(0.009)	
Country FE	YES			YES			YES			YES			YES		
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Exporter FE	YES		YES	YES		YES	YES		YES	YES		YES	YES		YES
Recipient FE		YES	YES		YES	YES		YES	YES		YES	YES		YES	YES
R^2	0.598	0.812	0.892	0.723	0.873	0.941	0.831	0.914	0.957	0.894	0.944	0.978	0.394	0.689	0.798
Adjusted R^2	0.501	0.548	0.525	0.657	0.694	0.740	0.790	0.794	0.812	0.868	0.865	0.905	0.248	0.253	0.110
Adjusted Within R^2	0.0815	0.0939	0.0185	0.0948	0.0951	0.0475	0.110	0.167	0.0459	0.152	0.216	0.101	0.0444	0.0532	0.00335
N	63,908	63,907	63,909	63,908	63,907	63,909	63,813	63,812	63,814	63,813	63,812	63,814	63,908	63,907	63,909

with distance.

Our results from Table 4 strongly support Hypothesis 6. The number of cargoes, N_{er}, declines with distance at the exporter level across recipients (Column 13), at the recipient level across exporters (Column 14), and when both exporter and recipient fixed effects are included in the regression (Column 15). The estimates are equal to -0.067 in Column 13, -0.111 in Column 14, and -0.087 in Column 15. All estimates are statistically significant at 1%.

Hypothesis 7 is also strongly supported by our results from Table 4. The intensive trade margin, $\overline{V_{er}}$, is increasing with distance at the exporter level across recipients (Column 4), at the recipient level across exporters (Column 5), and when both exporter and recipient fixed effects are included in the regression (Column 6). The estimates are equal to 0.097 (statistically significant at 1%) in Column 4, 0.017 (not statistically significant) in Column 5, and 0.028 (statistically significant 10%) in Column 6. Since at the exporter level $\overline{V_{er}}$ increases with distance faster than N_{er} declines with distance, the net effect is that the total trade value is increasing with distance at the exporter level across recipients. However, $\overline{V_{er}}$ increases with distance slower than N_{er} declines with distance once the recipient fixed effects are included in the regression thus making the total trade value a decreasing function of distance.

This result indicates that trade flows within the exporter behave differently with distance than within the recipient. To understand this difference, lets start with exporters. From Column 1 of Table 2 we know that in agreement with Hypothesis 1 cargo value is increasing with distance at the exporter level across recipients and that the estimated elasticity is statistically and economically significant in the absence of freight dummies. While the elasticity changes its sign and loses its statistical significance with freight dummies included in the regression (Column 1 of Table 3), the precise freight method is no longer available once cargo data is aggregated to the annual firm-level data. Therefore, the average-value-per cargo follows its cargo-level counterpart from Column 1 of Table 2 and strongly increases with distance (Column 4 of Table 4). In short, in the absence of precise freight dummies distance becomes a strong proxy for freight costs at the exporter level. Therefore, exporters have to ship more valuable on average cargoes long distances to at least break even.

Furthermore, a Russian exporter has on average four foreign recipients and, as Panel D of Figure 2 indicates, the number of recipients per exporter increases with distance. Even though an exporter ships more cargoes to each nearby recipient than it ships to each long distance recipient, the overall difference between the number of cargoes shipped nearby and long distance is partially offset by the fact that the exporter has on average more long than short distance contacts.

To explain the variation in the export value with distance within the recipient we need to consider recipients with Russian exporters located both nearby and far away. There exist two major reasons why the export value declines with distance within the recipients. First, a recipient trades, on average, with approximately two Russian exporters. Therefore the network effect - low number of shipments per destination but large number of destinations - is not as strong within the recipient. As a consequence, the number of shipments declines faster with distance within the recipient (-0.111, Column 14 of Table 4) than within the exporter (-0.67, Column 13 of Table 4). Second, since only recipients located in close proximity to Russia can trade with Russian exporters both far and nearby, most of the variation in distance within the recipient is due to the variation inside Russia. Therefore, Russian exporters shipping long distance to such recipients face lower transport and other trade costs as most of these costs accrue outside Russia. As a consequence, the average cargo value does not increase as fast with distance within the recipients (0.17, Column 5 of Table 4) as it does within the exporters (0.097, Column 4 of Table 4).

Export weight declines with distance across all specifications (Columns 7–9) but the relationship is statistically significant when recipient fixed effects are included in the regression (Columns 8 and 9). This is because in accordance with the decomposition (5) the extensive margin, weight-per-cargo, is either weakly increases with distance within exporters (Column 10) or it is independent of distance when recipient fixed effects are included in the regression (Columns 11 and 12) while as we have already discussed that the intensive margin – the number of cargoes – strongly, i.e., the magnitude of distance elasticity is large and statistically significant, declines with distance across all three specifications. These findings provide further evidence that, conditional on some freight method proxy, both fixed and variable freight costs per cargo only weakly depend on cargo's weight.

As we have anticipated, distance has much more explanatory power at the firm level than it has at the individual cargo level. This is mainly due to the fact that at the individual cargo level both exporter and recipient fixed effects together with the individual cargo's transport method subsume most of the information about the freight costs contained in distance. Once cargo data is aggregated one level up, cargo's individual freight method is no longer available but some of this lost information is captured by distance thus making it statistically significant at the firm level.

3.3 Country-level results

Hypothesis development

In this section we proceed further by aggregating the firm-to-firm pairwise data to the country level. The total trade value between Russia and country c, V_c with $c \in \{1, ..., 181\}$, is of main interest to us. Let N_E be the total number of Russian exporters, N_R be the total number of foreign recipients, and R_e be a set of the exporter e's recipients. The total number of exporters trading with a country c, N_{Ec}, is equal to

$$N_{Ec} = \sum_{e=1}^{N_E} \min\{1, \sum_{r=1}^{N_R} \mathbf{1}[(r \in c)\&(r \in R_e)]\}, \tag{9}$$

where $\mathbf{1}[(r \in c)\&(r \in R_e)]$ is an indicator function equal to one if the recipient r is located in country c and is in the exporter e's trading network. Likewise, the total number of cargoes shipped to a country c, N_c, is equal to

$$N_c = \sum_{e=1}^{N_E} \sum_{r=1}^{N_R} \mathbf{1}[(r \in c)\&(r \in R_e)] \cdot N_{er}. \tag{10}$$

Finally, for any firm-to-firm pairwise trade value V_{er} we construct its annual country-level counterpart, V_c, as

$$\begin{aligned} V_c &= \sum_{e=1}^{N_E} \sum_{r=1}^{N_R} \mathbf{1}[(r \in c)\&(r \in R_e)] \cdot V_{er} = \\ &= N_c \cdot N_{Ec} \cdot \left(\frac{1}{N_c} \cdot \frac{1}{N_{Ec}} \sum_{e=1}^{N_E} \sum_{r=1}^{N_R} \mathbf{1}[(r \in c)\&(r \in R_e)] \cdot V_{er} \right) = N_c \cdot \underbrace{N_{Ec} \cdot \overline{V}_{Ec}}_{\text{Average Value } \overline{V}_c}, \end{aligned} \tag{11}$$

where $\overline{V}_{Ec}$ is a per-exporter trade value averaged across all cargoes shipped to a foreign country c during the whole year. Such formulation allows us simultaneously accommodate two independent extensive margins of the country-level annual trade – the number of cargoes, N_c, and the number of exporters, N_{Ec}. $\overline{V}_{Ec}$ is the intensive margin of the country-level annual trade. Once data is collapsed to the country level, distance remains the only proxy for the freight costs and, therefore, it has to have a significant explanatory power in explaining the cross-section of the country-level trade flows.

Firm-level results shed light on the mechanism leading to the country-level trade flows being a decreasing function of distance. We have that the average cargo value is increasing with distance at the firm level and, given the aggregation procedure (11), this result has to hold at the country level for the average per-exporter cargo value of cargoes shipped to a country c, $\overline{V}_{Ec}$. At the same the number of cargoes strongly declines with distance at the firm level and this result also has to hold for the number of cargoes shipped to a country c, N_c. It is tempting to conclude that as long as the product of two extensive trade margins, $N_c \cdot N_{Ec}$, declines faster with distance than the extensive trade margin, $\overline{V}_{Ec}$, increases with distance at the country level, then the desired result should attain. However, the fact that the number of trading partners per exporter increases with distance may play an important role. If the size of the trading network grows fast enough with distance, it may make the number of exporters per country, N_{Ec}, an increasing function of distance and, consequently, the country-level trade flows either only weakly declining with distance or even increasing with distance. On the other hand, the size of the exporter's network increases with the exporter's size measured as the total value of its exports while the number of large exporters is small.[30] As a result, the number of exporters per country, N_{Ec}, may, in fact, decrease with distance thus making the aggregate trade flows a decreasing function of distance.

This discussion leaves out the intuition on why the functional form of the value-distance relation is a power law. Chaney (2016) proposes three sufficiency conditions leading to the negative power law relationship between the country-level trade flows and data. The first condition is that firms sizes, K, follow a Pareto distribution

$$1 - F(K) = \left(\frac{K}{K_{\min}}\right)^{\lambda}, \tag{12}$$

over $[K_{\min}, +\infty)$, $\lambda \geq 1$. The second condition is that the average squared distance of exports, Δ, is an increasing power function of the firm size, K,

$$\Delta(K) = \int_0^{\infty} x^2 f_K(x)dx = K^{\mu} \int_0^{\infty} x^2 f_{K_{\min}}(x)dx,\ \mu > 0, \tag{13}$$

where $f_K(x)$ is the fraction of exports shipped to distance x by a firm of size K, with $f_K(x)$ and $f'_K(x)$ bounded from above and $f_K(x)$ being weakly decreasing above some threshold $\overline{x}$. The third and final condition is a parameter restriction on μ and λ: $\lambda < 1 + \mu$. Chaney (2016) finds support for his sufficiency conditions in French trade data. Following Chaney (2016) we start by testing his sufficiency conditions in Russian trade data.

[30]For instance, Chaney (2016) shows that the distribution of the French exporter sizes is Pareto.

Testing Chaney (2016) gravity sufficiency conditions

We adopt Chaney (2016) methodology to test in the data his sufficiency conditions for the negative power law relation between the trade value and distance. We start by ordering all Russian exporters in increasing order of size, where a firm's size is the value of its worldwide exports. We construct 50 bins of equal log width, $b = 1, ..., 50$, ranging from 1 million Russian Rubles to the actual largest amount exported by a single firm. The average size of firms in bin b is given by

$$K_b = \frac{\sum_i \sum_j K_{i,j} \mathbf{1}\left[i \in b\right]}{\sum_i \mathbf{1}\left[i \in b\right]}, \tag{14}$$

where $K_{i,j}$ is the total value of exports of a firm i to a recipient firm j and $\mathbf{1}\left[\cdot\right]$ is an indicator function.

The fraction of firms of size larger than K_b is given by

$$1 - F\left(K_b\right) = \frac{\sum_{b'=b}^{50} \sum_i \mathbf{1}\left[i \in b'\right]}{\sum_{b''=1}^{50} \sum_i \mathbf{1}\left[i \in b''\right]}. \tag{15}$$

The average squared distance of exports among firms in bin b is given by

$$\Delta(K_b) = \frac{\sum_i \sum_j d_{i,,j}^2 \mathbf{1}\left[i \in b\right] \mathbf{1}\left[d_{i,,j} > 0\right]}{\sum_i \sum_j \mathbf{1}\left[i \in b\right] \mathbf{1}\left[d_{i,,j} > 0\right]}, \tag{16}$$

where $d_{i,,j}$ is the exact distance between the exporter i and the recipient j. Using exact distance between individual exporter-recipient pairs is the main difference between our variables and variables used by Chaney (2016).

Figure 3 shows $F\left(K_b\right)$ for Russian exporters and its fitted counterpart (Panel A), the Q-Q plot for the fit (Panel B),[31] and $\Delta(K_b)$ for Russian exporters (Panel C). It follows from Panels A and B that the distribution of Russian firm sizes is a three-parameter Generalized Exponential Distribution (GED):

$$1 - F(K) = \exp\left(-\theta\left[\left(\frac{K}{K_{\min}}\right)^{\lambda} - 1\right]\right) = \exp\left(-1.12 * \left[\left(\frac{K}{0.05}\right)^{0.21} - 1\right]\right), \tag{17}$$

with a cdf

$$f(K) = \frac{\theta\lambda}{K_{\min}}\left(\frac{K}{K_{\min}}\right)^{\lambda-1} \exp\left(-\theta\left[\left(\frac{K}{K_{\min}}\right)^{\lambda} - 1\right]\right), \tag{18}$$

[31] The fit is performed using CumFreq software. 90% confidence intervals are shown using widely spaced dashed lines. The efficiency coefficient of calculated and observed data values is equal to 0.99646.

rather than the Pareto distribution. Chaney (2016) first sufficiency condition is, therefore, violated for Russian exporters. Panel C shows that the second sufficiency condition (13) holds for Russian exporters as the average squared distance of exports is an increasing power function of firm size.

A natural question to ask is whether the GED of firm sizes is consistent with existing models of firm-to-firm trade. In Appendix B we present a variant of Chaney (2016) network model of firm-to-firm trade which delivers the GED of firm sizes. The first key difference between our model and the original Chaney (2016) model is that we allow firms to reduce their search intensity for new trading partners as they age, while it is constant in Chaney (2016). Since older firms have larger trade networks this assumption can be, for instance, be motivated by the costs of maintaining the network being increasing with the network size. The second key difference is is that we allow the hazard rate of losing a contact, which is constant in Chaney (2016), to decline with firm's age. This assumption could be, for instance, motivated by the fact that older, more mature firms tend to invest effort in maintaining the existing trading relationships and, therefore, lose existing trading partners with lower intensity.

Next, we investigate the consequences of the first sufficiency condition violation for aggregate gravity by performing the same aggregation exercise as in Chaney (2016). Following Chaney (2016), we normalize size units as $K_{\min} = 1$ and distance units as

$$\int_0^\infty x^2 f_1(x) dx = 1. \tag{19}$$

Aggregate exports at distance x, $\varphi(x)$, defined as the sum of all firm-level exports at x, are given by

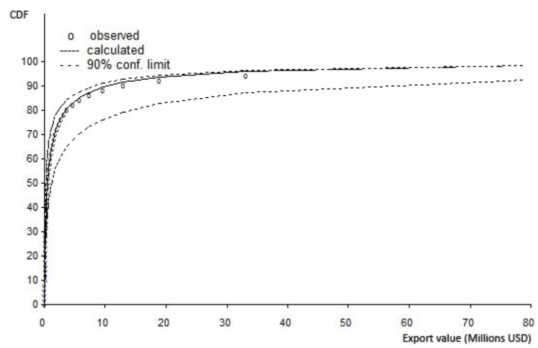
$$\varphi(x) \propto \int_1^\infty \left(K f_K(x)\right) K^{\lambda-1} e^{-\theta\left(K^\lambda - 1\right)} dK = \int_1^\infty f_K(x) K^\lambda e^{-\theta\left(K^\lambda-1\right)} dK. \tag{20}$$

We now introduce a scaled function $g_K(x)$

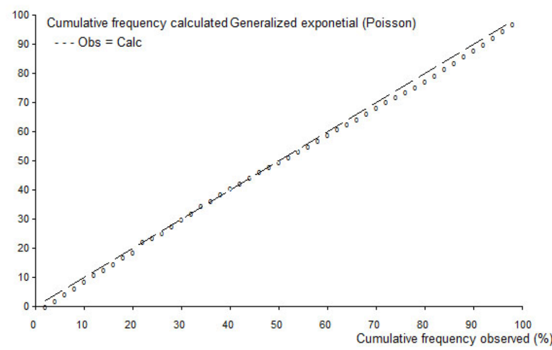
$$g_K(x) \equiv K^{\mu/2} f_K(K^{\mu/2} x), \tag{21}$$

with the same second moment $\int_0^\infty x^2 g_K(x) dx = \int_0^\infty x^2 f_1(x) dx = 1$, and perform the change of variables $K = \left(\frac{x}{u}\right)^{2/\mu}$, $dK = -\frac{2}{\mu}\left(\frac{x}{u}\right)^{2/\mu} \frac{du}{u}$. It is important to have $g_K(x)$ defined as in (21) since

Panel A: Trade sizes CDF



Panel B: Q–Q Plot



Panel C: Average square distance vs. trade size

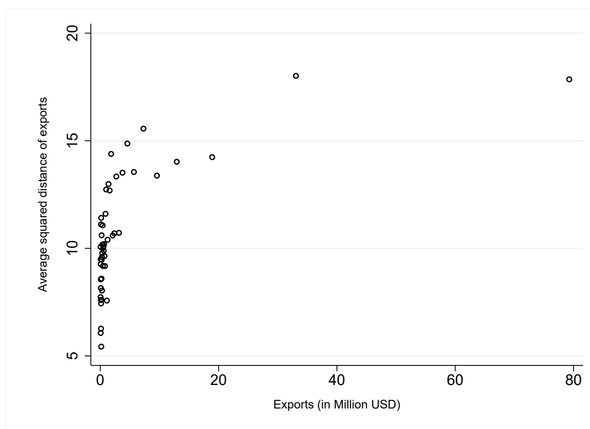


Figure 3: Russian firms' size distribution

All Russian firms exporting more than 100,000 USD in 2011 are ordered in increasing value of exports, and placed into 50 bins of equal log-size. Panel A shows fraction of firms larger than firms in bin b, as a function of the average size of exports among firms in bin b. Panel B shows the Q–Q plot of this distribution. Panel C shows the average squared distance of exports among firms in bin b as a function of the average size of exports among firms in bin b. Distance is measured in 1,000's of kilometers. This is Figure 1 in Chaney (2016).

otherwise it would not satisfy second sufficiency condition (13). We then obtain

$$\varphi(x) \propto x^{\frac{2(\lambda+1)-\mu}{\mu}} \int_0^x u^{-\frac{2(\lambda+1)}{\mu}} \exp\left(-\theta\left(\left(\frac{x}{u}\right)^{2\lambda/\mu}-1\right)\right) g_{\left(\frac{x}{u}\right)^{2/\mu}}(u)du, \tag{22}$$

where we have used that $g_K(x) = 0$ for $K \in [0,1)$. The exponent under the integral in (22) and the power of x, $\frac{2(\lambda+1)-\mu}{\mu}$, are key differences between the expression (22) and its counterpart in Chaney (2016). It follows from (22) that as long as $2(\lambda+1) > \mu$, the term in front of the integral, $x^{\frac{2(\lambda+1)-\mu}{\mu}}$, increases with x. Since Panel C of Figure 3 shows a concave relation between the average distance and the trade size, we conclude that the above inequality is satisfied in our data. However, the integral itself declines exponentially with x as $x \to \infty$.[32] As a result, $\varphi(x)$ is a non-monotone function of x in this case.

In the next section we test the implications of Chaney (2016) theory represented by the expression (22) against the classic gravity.

Country-level gravity results

We use five response variables characterizing the annual country-level trade between Russia and 181 countries: the total value of exports, V_c, the number of Russian firms exporting to country c, N_{Ec}, the number of shipments, N_c, and the average value per exporter per shipment, $\overline{V}_{Ec}$. Linear projections of the natural log of each of these five characteristics, $z_c = log(Z_c)$, are given by

$$y_c = \alpha + \beta \cdot \log(\mathrm{DIST}_c^2) + \eta \cdot \log(\mathrm{GDP}_c) + \delta \cdot \mathrm{NNHBR}_c + \epsilon_c, \tag{23}$$

where, following Anderson and VanWincoop (2003), NNHBR is a dummy taking value of 1 if the country has a border with Russia and zero otherwise. We utilize two different measures of country-to-country distance in our regressions. The first one is constructed as the distance averaged across all exporter-recipient from country c pairs. The second one is the distance between Moscow and the capital of country c.

Table 5 reports the results. The first observation is that distance elasticity is both statistically and economically significant across all specifications. Export value, V_c, falls with distance, and it falls faster with the average firm-to-firm distance than with the capital-to-capital distance.

[32] Using Lebesgue's dominated convergence theorem the limit $x \to \infty$ can be taken inside the integral where the exponent, $\exp\left(-\theta\left(\left(\frac{x}{u}\right)^{2\lambda/\mu}-1\right)\right)$, is the dominant function of x.

Table 5: Country level exports

Exports' from Russia to country c characteristics. F-to-F (C-to-C) indicates that the distance is measured as the average distance across all distances between recipients in country c and all Russian firms trading with them (capital of c and Moscow). NNHBR dummy takes the value of 1 if the country has a border with Russia and zero otherwise. Clustered by recipient standard errors are reported in parentheses with significance levels indicated by * (10%), † (5%), ‡ (1%). Excludes countries with less than 6 cargo shipments from Russia to the country: Cayman Islands; Central African Republic; Comoros; Mayotte; El Salvador; Gambia; Guadeloupe; Jamaica; Malawi; Maldives; Monaco; Guinea-Bissau; East Timor; Saint Lucia; Western Sahara; US Virgin Islands; Zambia.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Value (V_c)		Number of Cargoes (N_c)		Value per Cargo per Exporter (V_{Ec})		Number of Exporters per Country (N_{Ec})	
	F-to-F	C-to-C	F-to-F	C-to-C	F-to-F	C-to-C	F-to-F	C-to-C
log(DIST^2)	-1.166‡	-0.874‡	-0.899‡	-0.706‡	0.438†	0.405‡	-0.705‡	-0.573‡
	(0.149)	(0.102)	(0.146)	(0.083)	(0.169)	(0.115)	(0.124)	(0.074)
log(GDP)	0.881‡	0.742‡	0.702‡	0.592‡	-0.384‡	-0.325‡	0.563‡	0.474‡
	(0.073)	(0.073)	(0.054)	(0.057)	(0.059)	(0.067)	(0.044)	(0.046)
NNHBR Dummy	0.804	1.032†	2.084‡	2.192‡	-2.915‡	-2.838‡	1.636‡	1.678‡
	(0.501)	(0.464)	(0.488)	(0.444)	(0.554)	(0.501)	(0.432)	(0.379)
Constant	0.120	2.952	-9.225‡	-6.881‡	17.277‡	15.814‡	-7.932‡	-5.980‡
	(1.921)	(1.882)	(1.552)	(1.519)	(1.777)	(1.906)	(1.281)	(1.259)
N	181	179	181	179	181	179	181	179
R^2	0.564	0.573	0.614	0.631	0.354	0.373	0.602	0.627

V_c declines with distance since both extensive margins of trade - the number of cargoes and the number of exporters per country - sharply decline with distance, with distance elasticities equal to $-0.899(-0.706)$ and $-0.705(-0.573)$ respectively.[33] The intensive trade margin - average cargo value - increases with distance but not fast enough to compensate for the combine decline with distance of both extensive margins as its distance elasticity is equal to 0.438(0.405).

Head and Mayer (2014) compile 1,835 estimates of the distance elasticity in gravity type regressions from 161 published papers. The mean distance elasticity is -0.93 (median -0.89 and s.d. 0.4) among all estimates. The distance elasticity is remarkably stable, hovering around -1 over a century and a half of data. Column 1 of Table 5 reports that the value elasticity of distance, $2\beta = -2.33$, is more than 2.5 times smaller in our data than the median value reported by Head and Mayer (2014). Column 2 indicates that the distance elasticity estimate is equal to -1.75 when capital-to-capital distance is used. Russia is the largest country on the planet with an area of $17,075,200$ km^2 which is almost twice the area of the second largest country, Canada. It is, therefore, plausible that distance as a proxy for trade costs is more important for Russian exporters than for exporters from any other country.

Our data allows us to evaluate the magnitude of the correction to the distance elasticity due to precise distance measure. The change is equal to 33.1% and it is both statistically and economically significant. It indicates that while firms minimize fixed and variable distance related trade costs by concentrating majority of their trading with nearest neighbors, there still exists a large number of shipments crossing the vast Russian territory. This result is most important to countries with a large geographic area like USA, China, Canada, and India. For these countries, the economic impact of distance on trade is underestimated.

The most intriguing result is that the classic gravity - trade value declines with distance as a power law - holds for Russian firms in spite of Chaney (2016) sufficiency conditions being violated for them. Moreover, while the relation (22) predicts a non-monotone relation between V_c and distance, when we include DIST_i^α into the specification (23) for several different values of the power α, the coefficient on this term is neither statistically nor economically significant in any of our specifications.[34] In addition, the non-monotone relation between the aggregate trade value and distance has not been reported by any existing gravity study. This begs the question of what is the right aggregation for Russian trade data? We answer this question in the next section.

[33] Results shown in brackets are for the case of the capital-to-capital distance.

[34] These results are available upon request.

4 Cargo-to-country aggregation

We start by developing an aggregation procedure for a generic data "unit." Such unit can be a customs form, a cargo, or a firm. The main property of such a unit is that it can be aggregated to any higher unit. For instance, customs forms can be aggregated to a firm-level data bypassing the cargo unit level.

4.1 Aggregation method

Lets start with an arbitrary "unit" of value $v_i(x)$ shipped over distance $x \in X$, where X is a set of all distances. The total number of units is N and the total number of units shipped over distance x is $N(x)$:

$$\sum_{x \in X} N(x) = N. \tag{24}$$

The total value shipped to distance x, $V(x)$, is equal to:

$$V(x) = \sum_{i=1}^{N(x)} v_i(x) = \sum_{K=K_{\min}}^{\infty} \sum_{i=1}^{N(x)} \mathbf{1}\left[v_i(x) = K\right] \cdot K = \sum_{K=K_{\min}}^{\infty} K N_K(x), \tag{25}$$

where

$$N_K(x) \equiv \sum_{i=1}^{N(x)} \mathbf{1}\left[v_i(x) = K\right],$$

is the number of units with the value equal to K shipped over distance x out of the total number of units of that value. We now examine the left-hand-side of the expression (25). Let

$$\alpha(x) \equiv \frac{N(x)}{N},$$

be the fraction of units shipped over distance x. The right-hand-side of (25) takes the following form

$$\frac{1}{N} \sum_{i=1}^{N(x)} v_i(x) = \frac{N(x)}{N} \frac{1}{N(x)} \sum_{i=1}^{N(x)} v_i(x) \underset{N \to \infty}{\longrightarrow} \alpha(x) \mathrm{E}[v(x)], \tag{26}$$

where $\mathrm{E}[v(x)]$ is the expected unit value shipped over the distance x. Equation (26) implies that in the large-N limit $\frac{1}{N} \sum_{i=1}^{N(x)} v_i(x)$ converges to the product of the expected value and the fraction of units both shipped to a distance x. This expression is valid regardless of weather we use the customs forms, cargoes, or firms.

We now demonstrate how the country-level distance elasticity from Column 1 of Table 5 equal to -2.33 can be obtained from the cargo-level data using relation (26). Let both $N(x)$ and $\mathrm{E}[v(x)]$ be power functions of distance, x,

$$N(x) \propto x^{-\beta_N}, \qquad (27)$$
$$\mathrm{E}[v(x)] \propto x^{\beta_v},$$

where $\beta_N > 0$ and $\beta_v > 0$ represent distance elasticities of the number of cargoes and average cargo value respectively. Here we have used our firm- and country-level empirical results to make the number of cargoes a decreasing function of distance and the average cargo value shipped to distance x an increasing function of distance.

Relations (27) can, for instance, be obtained directly from (25) in the large-N limit. To see this note that relation (25) can be written as:

$$\sum_{i=1}^{N(x)} v_i(x) = N \sum_{K=K_{\min}}^{\infty} K \cdot \underbrace{\frac{N_K(x)}{N_K}}_{f_K(x)} \cdot \underbrace{\frac{N_K}{N}}_{F(K)dK} \underset{N\to\infty}{\longrightarrow} \int_{K_{\min}}^{\infty} K f_K(x) F(K) dK. \qquad (28)$$

The right hand side of (28) is exactly the expression from Chaney (2016). Using the large-N limit in the expression for the number of cargoes

$$N(x) = N \sum_{K=K_{\min}}^{\infty} \frac{N_K(x)}{N_K} \frac{N_K}{N} \underset{N\to\infty}{\longrightarrow} \int_{K_{\min}}^{\infty} f_K(x) F(K) dK, \qquad (29)$$

we obtain the large-N limit expression for the average cargo value at distance x

$$\mathrm{E}[v(x)] = \frac{1}{N(x)} \sum_{i=1}^{N(x)} v_i(x) \underset{N\to\infty}{\longrightarrow} \frac{\int_{K_{\min}}^{\infty} K f_K(x) F(K) dK}{\int_{K_{\min}}^{\infty} f_K(x) F(K) dK}. \qquad (30)$$

For $\int_{K_{\min}}^{\infty} K f_K(x) F(K) dK$ to deliver the power law in distance, Chaney (2016) sufficiency conditions and specifically the requirement that $F(K)$, the cdf of cargo values, is Pareto have to be satisfied at the cargo level.

Is the assumption that cargo values are Pareto distributed consistent with the finding that the firm-level export values are not Pareto distributed? Indeed these two findings are consistent as it is well-known that the sum of Pareto-distributed variables is not Pareto distributed. Therefore, if, for instance, the custom forms values are Pareto distributed, then gravity can only result from the custom forms level aggregation to the country level since the higher-level values of cargoes and firm-level exports are obtained by adding up custom forms values and thus cannot be Pareto distributed.

If all Chaney (2016) conditions are satisfied, we obtain analogous to Chaney (2016) the asymptotic expressions for $N(x)$ and $\mathrm{E}[v(x)]$

$$N(x) \propto \frac{1}{x^{1+\frac{2\lambda}{\mu}}} \Rightarrow \beta_N = 1 + \frac{2\lambda}{\mu}, \qquad (31)$$
$$\mathrm{E}[v(x)] \propto x^{\frac{2}{\mu}} \Rightarrow \beta_v = \frac{2}{\mu},$$

directly mapping β_N and β_v into the parameters μ and λ from Chaney (2016).

We are going to use the cargo-level data to estimate β_N and β_v. To do that, we create distance bins each of the equal size of 500km for a total of 36 bins, i.e., the last bin corresponds to cargoes shipped father than 18,000km, and sort all cargoes into these bins based on their actual shipping distance. Choosing the right "size" of each distance bin is tricky as the destinations are not uniformly distributed over the Earth map. Destinations are scarce outside the large or medium-sized metropolitan areas and there are no destinations if the bin is in a sea or an ocean. Choosing too of a fine grid size leads to a lot of fluctuations in both cargo values and number of cargoes along the distance dimension especially over distances ending in the Atlantic, Pacific, and Indian Oceans as well as Caspian, Black, Baltic, and Mediterranean Seas. Using 500km sized bins helps to mitigate this problem, but not completely. The main problem is that there are a lot of outliers in cargo values for bins located farther than 5,000km. To mitigate this problem, we use median instead of average cargo value for each distance bin.

Once all cargoes are sorted into their respective distance bins and the bin-specific median cargo value is calculated, we estimate β_N and β_v from the following OLS specifications

$$\log(N_i) = \alpha_N + \beta_N \cdot \log(\mathrm{DIST}_i) + \eta_{N,i}, \qquad (32)$$
$$\log(V_i) = \alpha_v + \beta_v \cdot \log(\mathrm{DIST}_i) + \eta_{v,i},$$

where V_i and N_i are median cargo value and number of cargoes in bin i respectively, and DIST_i is bin i' corresponding distance. Figures 4 and 5 illustrate estimation of equations (32) for the number of cargoes and cargo median values respectively.

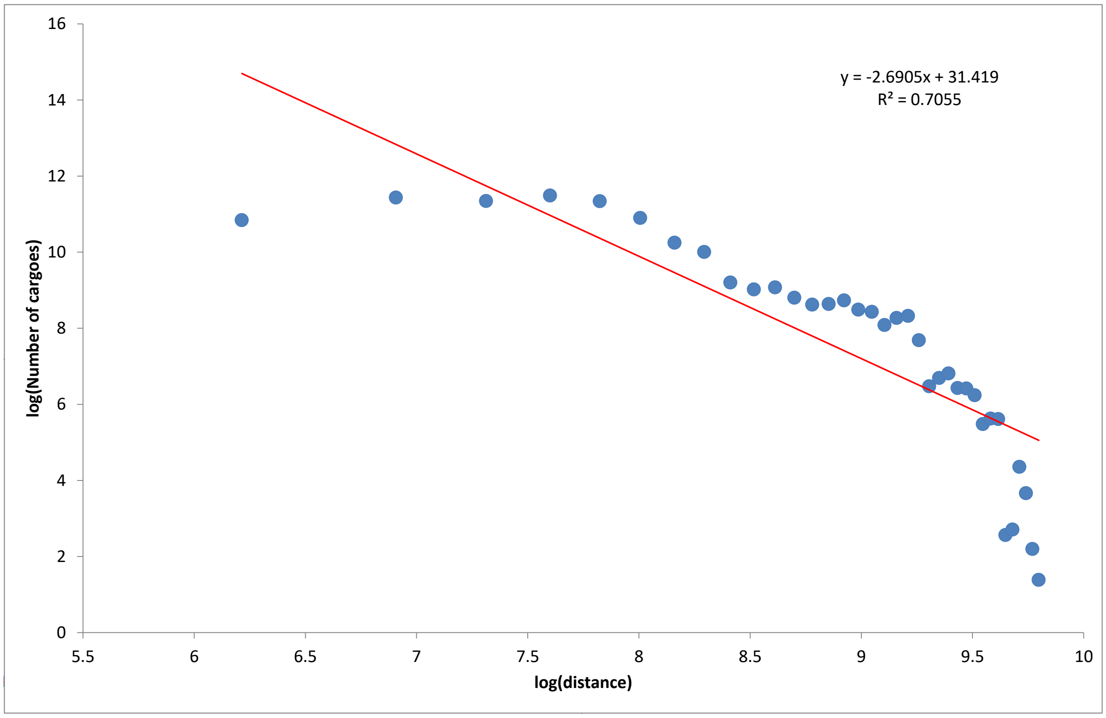


Figure 4: Number of cargoes vs distance

Log of number of cargoes as a function of log distance is plotted. The straight line shows an ordinary least square fit. Cargoes are split into 36 equidistant distance bins 500km in length.

Figure 4 shows that the number of cargoes declines with distance as a power law between 1,500km and 13,500km which to the west of Moscow encompasses everything from Western Europe to the East Coast of the United States and it encompasses India, China, and the West Coast of the United States starting from Vladivostok which is the farthest Eastern Russian port. It then declines exponentially over distances greater than 13,500km. The number of shipments is slightly increasing between 500km and 1000km bins. This could be due to low shipping costs to countries of former Soviet block coupled with differences in per-capita wealth across these countries. For instance, 500km bin includes Ukraine, Moldova and Belarus while 1000km bin includes Poland, Romania, Slovakia, and Hungary all of which are on average wealthier nations than former Soviet republics from the 500km bin. The estimated coefficient $\widehat{\beta}_N$ is equal to -2.69 with a t-stat of -9.03. The R^2 of the regression is equal to 70.6%.

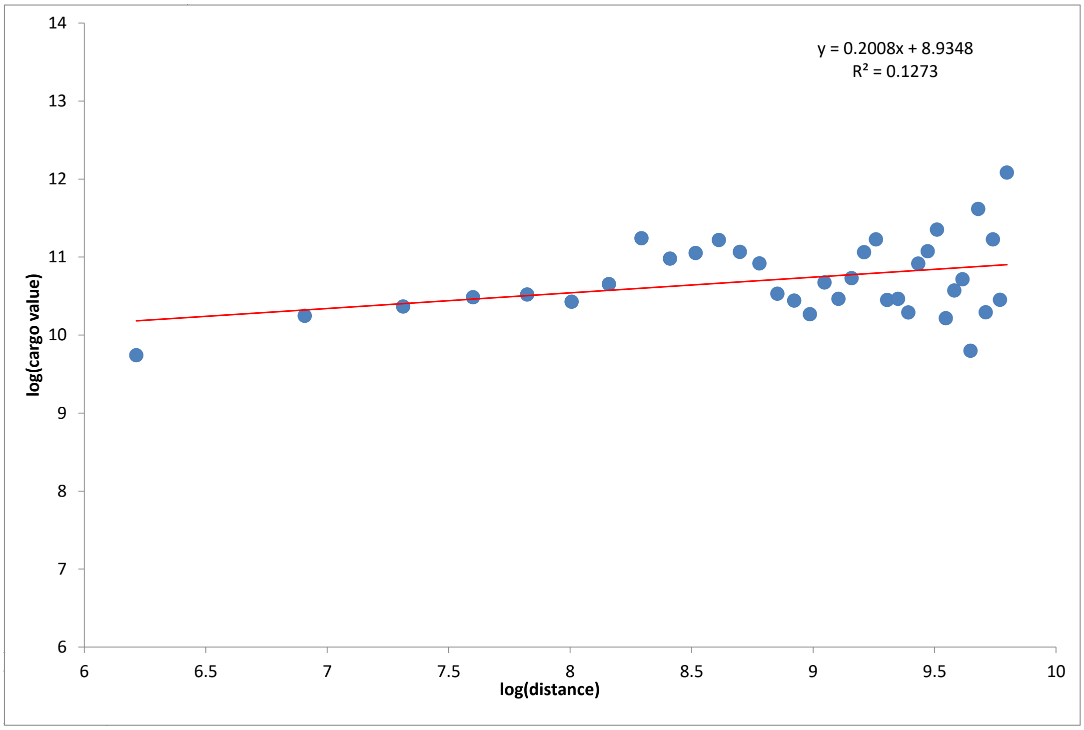


Figure 5: Cargo value vs distance

Log cargo value as a function of log distance is plotted. The straight line shows an ordinary least square fit. Cargoes are split into 36 equidistant distance bins 500km in length. Cargo value is the median cargo value for each bin.

Figure 5 shows that the median cargo value is slowly increases as a power law of distance. The estimated coefficient $\widehat{\beta}_v$ is equal to 0.20 with a t-stat of 2.23 making it statistically significant at 5% level. The R^2 of the regression is equal to 12.7%. It follows from (27) that the country-level distance elasticity from (23)can be calculated as

$$\widehat{\beta}_N + \widehat{\beta}_v = -2.69 + 0.20 = -2.49, \tag{33}$$

which is very close to the value of $2 \cdot (-1.166) = -2.33$ reported in the Column 1 of Table 5. Here we have multiplied the estimate from Table 5 by two since the specification reported in Table 5 uses squared instead of linear distance.

Sorting cargoes into equidistant bins is very different from sorting them into country-specific bins. The country itself is irrelevant for sorting into equidistant distance bins as distance is the only relevant sorting characteristic. Thus each distance bin contains cargoes shipped to many different countries. In addition, the actual travel distance is used when forming equidistant bins. Country-specific bins contain cargoes shipped over different distance. Consequently, the country-specific distance is constructed by averaging all actual distances unique to each country-specific exporter-recipient pair. In addition, country-level gravity specification (23) contains extra controls such as the nearest-neighbor dummy and country GDP, potentially contributing to a difference between two estimates of distance elasticity.

This exercise demonstrates the power of our approach. Once cargo-level data is available, no further data aggregation is required to estimate the firm- or country-level value elasticity of value.

5 Conclusion

In this paper, we study how the classic country-level gravity relationship emerges by employing a highly detailed and precise Russian customs level data that allows empirical analysis of trade flows and distance at different levels of data aggregation. The Russian customs data contains virtually all daily declaration forms during 2011, and allows us to identify the unique cargo shipments between domestic exporting firms and foreign recipient firms. Importantly, the customs forms data contains the precise identification of firms at both sides of the transaction which enables us to calculate the exact door-to-door distance between them as well as build the networks of exporting and importing firms. Furthermore, the high dimensionality of our cargo level data (based on one or more customs forms on the same day from the same exporter to the same recipient) allows us to employ time, product and firm level fixed effects for exporters and recipients, thus enabling us to identify the role of distance and trade-flows through variation within exporters and recipients.

Our main set of results on gravity concerns the cargo shipment level of trade between firms. We find that the exact distance between firms has no explanatory power for variation of cargo shipment values if one controls for the modes of cargo transportation or includes the recipient firm fixed effects. In fact we find that the unit cargo value *increases* with distance across recipients for a given exporter or alternatively for a given importer across their Russian exporters. Moreover, we find a negative relationship between distance and the variety of products in a cargo shipment. Our findings uncover the previously unexplored intricate dimension of international trade between firms at the cargo level. These findings are consistent with fixed and variable costs that dominate the extensive and intensive margins of trade at the basic micro cargo level.

Next, we aggregate the cargo shipments data to the *firm* level and obtain trade values for matched exporter-recipient firm pairs over the whole year. We find that at the firm level the effect of distance on the value of exports is ambiguous. For a given exporter, the export value increases with distance across recipients, albeit insignificantly. On the other hand, and in line with traditional gravity prediction, we find that conditional on an importer the dollar value of trade is decreasing with exporters' distance. In addition, we find strong evidence that both weight of goods exported and goods' variety decrease with the distance between firms.

The question that naturally arise is why doesn't distance appear to play a major role at the cargo level, yet it does at the firm level? Our analysis shed light on this and supports the view that most of the exporter's decision making is done at the individual cargo level. Exporters respond to their recipient's demand consisting of both the quality of the commodity and the timing (expediency) of its delivery. The timing part of the recipient's demand is captured to a large degree by the transport method, while the majority of the unobserved quality component of the recipient's demand is captured by the recipient fixed effects. Export costs are best explained by the transportation method, the number of shipments, and exporter fixed effects. Consequently, conditional on the recipient demand, the exporter chooses the number of cargoes to ship to each recipient, and for each cargo its content. Thus, distance has very low explanatory power at the cargo level since both exporter and recipient fixed effects together with the individual cargo's transport method subsume most of the information about the export economic costs contained in distance.

Once cargo data is aggregated up to firm level data, cargo's individual transport method is no longer available and some important information about transport costs and recipients demand is lost. Some of this lost information is recovered by using distance and the most frequently employed transport method. As a result, distance becomes statistically significant at the firm level as it proxies for some of the information about the transport costs contained in the transport method.

Finally, when we aggregate firms' data to the *country* level we find significant support for the classical gravity regression of a negative relation between distance and value of trade. Furthermore, the gravity coefficient significantly increases when one uses the average exact door-to-door distance between firms within a country instead of the surrogate capital-to-capital distance commonly used in the literature. Since our data allows us to calculate the number of cargo shipments sent to each country and a number of firms involved in trade with each country we are able to decompose the value of exports to each of the countries along the *intensive* and *extensive* margins. Here we find that average value per shipment per firm sent to a country increases with distance to that country, while number of cargo shipments sent to a country and number of Russian exporting firms trading with a country declines with distance.

Next, we investigate the reason the data aggregates at the country level to the classical gravity equation. First, we show empirically that Russia's firm size distribution is Generalized Exponential – one that violates Chaney's (2016) first sufficiency conditions for gravity which requires firm sizes to be Pareto distributed. We then show theoretically how to account for the fact Chaney's (2016) conditions are satisfied at the shipment level, but are not at the firm level, yet gravity holds at the country level.

We show that our evidence that Russian firms ship more valuable goods to farther distances but at a shipment rate that declines with distance even faster is a feature consistent with Chaney's (2016) sufficient conditions for gravity to hold at the country level. That is, we show that in order to yield the correct aggregation at the country level, the cargo's intensive margin should increase slower with distance than the decline in the average number of shipments. We further show that given that Chaney's conditions hold at the cargo level, they can not, by construction, be satisfied at the firm level (as a weighted average of Pareto is not Pareto distributed)–rationalizing our firm level evidence. Therefore, as long as the conditions are satisfied at the cargo level, the firm level aggregation may yield either positive, negative or no relationship between distance and value traded at the firm level, making that evidence much less informative regarding the country-level gravity.

Overall, our results show that most of the exporter's decision making is done at the individual cargo level. Consequently, models of international trade need to start adopting cargoes as one of primary firm-level control variables to be consistent with both granular cargo-level data as well as higher level firm-level data eventually leading to the classic gravity.

References

Alchian, Armen A., and William R. Allen. 1964. *University Economics*. Belmont, Calif.: Wadsworth.

Anderson, J., VanWincoop, E., 2003. Gravity with gravitas: a solution to the border puzzle. *American Economic Review*, 93 (1), 170–192.

Baldwin, R., Harrigan, J., 2011. Zeros, quality and space: Trade theory and trade evidence. *American Economic Journal: Microeconomics*, 3(2), 60-88.

Bastos, P., Silva, J., 2010. The quality of a firm's exports:where you export to matters. *Journal of International Economics*, 82 (2), 99–111.

Benguria, F., 2014. Production and distribution in international trade: Evidence from matched exporter-importer data. Mimeo.

Bernard, A., Eaton, J., Jensen, J., Kortum, S., 2003. Plants and productivity in international trade. *American Economic Review*, 93 (4), 1268.

Bernard, A., Jensen, J., Schott, P., 2005. Importers, exporters, and multinationals: a portrait of firms in the US that trade goods. NBER Working Paper.

Bernard, A., Jensen, J., Redding, S., and P. Schott, 2007. Firms in international trade. *Journal of Economic Perspectives*, 21, 105–130.

Bernard, A., Jensen, J., Redding, S., and P. Schott, 2012. The Empirics of firm heterogeneity and international trade. *Annual Review of Economics*, 4, 283- 313.

Bernard, A., Moxnes, A., Ullveit-Moe, K., 2017. Two-sided heterogeneity and trade, Mimeo

Blum, B., Claro, S., and I. Horstmann, 2010. Facts and figures on intermediated trade. *American Economic Review: Papers and Proceedings*, 100 (2), 419-423.

Blum, B., Claro S., and I. Horstmann, 2013. Occasional and perennial exporters. *Journal of International Economics*, 90 (1), 65-74.

Carballo, J., Ottaviano, G., Martinkus, C., 2016. The buyer margins of firms' exports, Mimeo.

Chaney, Thomas, 2014. The network structure of international trade. *American Economic Review*, 104(11), 3600–3634.

Chaney, Thomas, 2016. The gravity equation in international trade: An explanation. *Journal of Political Economy*, Forthcoming.

Dragusanu, R., 2014. Firm-to-firm matching along the supply chain, Harvard University, mimeo.

Eaton, J., Eslava M., Krizan C., Kugler M., and J. Tybout, 2014. A search and learning model of export dynamics, Penn State University, mimeo.

Eaton, J., Kortum, S., 2002. Technology, geography, and trade. *Econometrica*, 70 (5), 1741–1779.

Gorg, H., Halpern, L., Murakozy, B., 2010. Why do within firm-product export prices differ across markets? CEPR Working Paper 7708.

Harrigan, J., Ma, X., and V. Shlychkov, 2015. Export prices of U.S. firms. *Journal of International Economics*, 97, 100-111.

Head, Keith and Thierry Mayer, 2014. "Gravity Equations: Workhorse, Toolkit and Cookbook," in Handbook of International Economics Vol. 4, edited by Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff.

Heise, S., 2016. Firm-to-firm relationships and price rigidity: Theory and evidence, CESifo Working Paper.

Hummels, D., Skiba, A., 2004. Shipping the good apples out? An empirical confirmation of the Alchian–Allen conjecture. *Journal of Political Economy*, 112 (6), 1384–1402.

Kamal, F., Monarch, R., 2016. Identifying foreign suppliers in US merchandise import transactions. Mimeo

Kamal, F., Sundaram, A., 2016. Buyer-seller relationships in international trade: Do your neighbors matter? *Journal of International Economics*, 102, 128–140.

Manova, K., Zhang, Z., 2012. Export prices across firms and destinations. *Quarterly Journal of Economics*, 127 (1), 379-436.

Monarch, R., 2014. ItâĂŹs not you, itâĂŹs me: Breakups in U.S.-China trade relationships, Working Paper 14-08, U.S. Census Center for Economic Studies.

Monarch, R., and T. Schmidt-Eisenlohr, 2016. Learning and the value of trade relationships, Mimeo.

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

Martin, J., 2012. Markups, quality, and transport costs. *European Economic Review*, 56, 777-791.

Pierce, J., and P. Schott, 2012. The surprisingly swift decline of U.S. manufacturing employment, NBER Working Paper 18655.

Roberts, P. O., 1999. Logistics and Freight Transportation: Review of Concepts Affecting Bulk Transportation, World Bank.

von Thünen, J. H., 1826. Der Isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie. Puthes, Hamburg.

Yi, K., 2003. Can vertical specialization explain the growth of world trade? *Journal of Political Economy*, 111 (1), 52–102.

Appendix A

Creation of the unique IDs for foreign firm recipients using their reported names and postal addresses from multiple customs declarations is naturally an imperfect exercise. Alternative spellings of the firms' names or different conventions of including or not including the firms' types such as Inc. or Ltd. create a problem for the correct identification of the unique foreign firms IDs from customs forms. Fortunately for us this problem is greatly alleviated for the postal addresses since the customs declarations and the raw data files that we possess keep country names, zip codes, street names, building numbers in separate variables.

To assign the unique foreign firm IDs by names and addresses we wrote an algorithm that utilizes the Stata module reclink2, developed by Wasi and Flaaen (2014). This procedure uses fuzzy matching of string variables allowing the user to place different weights on the importance of different components of the string variables. Before we start this process, we carefully abbreviate all common words that frequently appear in the firm names and addresses. After we launch the algorithm the reclink2 procedure generates a field similarity score from 0 to 1 for each pair of matched Hypothesiss. As common in the literature (See Kamal and Monarch (2016)) we consider customs declarations with the field similarity score that falls into the range [0.98-1] to belong to the same foreign firm recipient and assign the same ID to such declarations.

In order to ensure the fully correct matching we underwent several passes of manual re-assignments of the foreign firm IDs that were not correctly picked up by the reclink2 procedure. We performed multiple sorting of our dataset on names, address of foreign firms and tax IDs of the Russian firms, visually checked the similarities of firms' names and addresses and manually performing IDs re-assignments of if it was necessary.

Several passes of the automated and manual procedures described above make us confident that we have produced a near perfect identification of the unique IDs for foreign firm recipients in our dataset.

Appendix B

Model of firm-to-firm trade yielding Generalized exponential distribution

We show how the Generalized exponential distribution of trade sizes, K, can be rationalized using a variant of Chaney (2016) model. Our model will share most but not all assumptions with Chaney (2016). We start with those assumptions that we borrow from Chaney (2016). The world is made of a continuum of locations. Firms are uniformly distributed over an infinite one-dimensional continuous space with origin located at the coordinate $x = 0$ and just like Chaney (2016) we will focus on the firm located at the origin. Within each location, there is a continuum of infinitely lived firms, born at a constant rate γ, so that at time t, there is the same density of firms $e^{\gamma t}$ in every location. Once born, a firm gradually samples a mass K_0 of trading partners from other newborn firms only. The K_0 contacts are distributed geographically according to the density $k_0(x)$ such that

$k_0(-x) = k_0(x)$, $x \in \mathbb{R}$, and it has a finite second moment.

Existing contacts can "share" their own contacts according to a Poisson process with arrival rate $\beta_a = \frac{\beta}{1+ma}$, $m > 0$, so as the firm ages it slows down its search for new partners due to, for example, costs of maintaining the network not directly modeled here. Existing contacts are continuously lost to the exogenous Poisson shock with rate $\delta_a = \frac{\delta}{1+ma}$. As the firm ages, it takes a better care of its partners and, therefore, does lose them with lower intensity. Following Chaney (2016) we assume that $\gamma > \beta - \delta > 0$. The value of individual cargoes is normalized for simplicity to one for all firms thus making the firm's number of contacts, K, a measure of its size.

Let $k_a(x)$ be the geographic distribution of the contacts of a firm of age a located at the destination x. Then the number of contacts of this firm "worldwide" is equal to

$$K_a = \int_{\mathbb{R}} k_a(x)dx. \tag{34}$$

The distribution of contacts evolves recursively according to the partial differential equation

$$\frac{\partial}{\partial a} k_a(x) = \beta_a \int_{\mathbb{R}} \frac{k_a(x-y)}{K_a} k_a(y)dy - \delta_a k_a(x), \tag{35}$$

with the initial condition $k_0(x)$. Multiplying both side by $dxda$, (35) describes the net creation of new contacts in a neighborhood dx of location x over a short time interval da. Any existing contact of the firm (there are $k_a(y)dy$ of them in each neighborhood dy of y) shares with probability $\beta_a da$ one of their contacts. This contact happens to be in a neighborhood dx of x with probability $\frac{k_a(x-y)}{K_a}$. Over a time interval da, the firm also loses $\delta_a k_a(x)dxda$ of its contacts in a neighborhood dx of location x. In our case, however, both probabilities, $\beta_a da$ and $\delta_a da$, decline with the firm's age.

First, we are going to derive the law of motion for the mass K_a by simply integrating the law of motion (35) over $\mathbb{R}$ to obtain the following ODE

$$\frac{\partial}{\partial a} K_a = (\beta_a - \delta_a) K_a, \tag{36}$$

with the initial condition K_0. ODE (36) can be solved to yield

$$K_a = K_0 \exp\left(\int_0^a (\beta_s - \delta_s) ds\right) = K_0 (1+ma)^{\frac{\beta-\delta}{m}}. \tag{37}$$

Unlike Chaney (2016) where the firm size growth exponentially, the firm size growth with age as a power law in our case. The Chaney (2016) case can be recovered by taking the limit $m \to 0$ in (37).

The population grows at an exponential rate γ, so that at any time t, the fraction of firms younger than a is $1 - e^{-\gamma a}$. Since a firm of age a has a total number of contacts K_a, we can express

$a(K)$ from (37) to obtain

$$a(K) = \frac{\left(\frac{K}{K_0}\right)^{\frac{m}{\beta-\delta}} - 1}{m}, \tag{38}$$

thus leading to the following expression for the fraction of firms with fewer than K contacts

$$F(K) = 1 - \exp\left[-\frac{\gamma}{m}\left(\left(\frac{K}{K_0}\right)^{\frac{m}{\beta-\delta}} - 1\right)\right], \tag{39}$$

which is a Generalized Exponential Distribution.