Appendix to
The Agglomeration of Exporters by Destination
Andrew J. Cassey and Katherine N. Schmeiser
September 2012

A  The political organization of Russia and the location of ports and products

In 2003, Russia had 89 federal regions. These regions were separated into 49 oblasts (provinces), 21 republics, 6 krais (territories), 10 okrugs (autonomous districts), 1 autonomous oblast, and 2 federal cities (Moscow and St. Petersberg). The divisions are based on ethnic groups and history. Each region has equal representation in the Federal Assembly, though there are differences in autonomy. However these differences are minor (and decreasing over time) because each region is attached to one of eight federal districts administered by an envoy of Russia’s President. The districts are designed to oversee regional compliance with federal laws.

Most taxes are set federally. In 2003, regions could set corporate property tax, gambling business tax, and transport tax within bounds established federally (CTEC 2006, p. 955). Exports of goods from Russia are subject to a value-added tax of 0% throughout Russia (p. 960), except on some exports of oil and natural gas. Export customs duty rates are set federally as well (p. 961).

We show that the location of Russian exporters can neither be accounted for by the location of export commodities nor the location of borders and ports.
Russia’s major manufacturing industries are: all forms of machine building from rolling mills to high-performance aircraft and space vehicles; defense industries including radar, missile production, and advanced electronic components, shipbuilding; road and rail transportation equipment; communications equipment; agricultural machinery, tractors, and construction equipment; electric power generating and transmitting equipment; medical and scientific instruments; consumer durables, textiles, foodstuffs, and handicrafts.

Russia’s largest export commodities are: petroleum products, wood and wood products, metals, chemicals, and a wide variety of civilian and military manufactures. Products such as petroleum, wood, metals, and chemicals need to be processed to become manufactured goods. Therefore it is possible that the location of exports is the same as the location of the refineries and processing plants for these goods.

Though many exporters are located in the same region as a major port, there are regions that are not close to a major port in the top quintile of the number of exporting firms. The largest ports are Azov (Rostov, Black Sea), Kaliningrad (Kaliningrad, Baltic Sea), Nakhodka and Vostochny (Primorsky, Sea of Japan, largest), Novorossiysk (Krasnordar, Black Sea, largest), Primorsk (Leningrad, Baltic Sea, largest, lots of oil), and Saint Petersburg (Saint Petersburg, Baltic Sea). Cassey (2011) documents that access to water is an important predictor of U.S. state exports.
First, consider the number of exporting firms by region. Figure 1 shows that the regional concentration of Russian manufacturing exporters is diverse. The majority of exporters are located either around Moscow and St. Petersburg or the Kazakhstan/Mongolia/China border. Nine regions did not have any reported exporting firms in 2003. Figure 1 shows that in Russia, as elsewhere, many exporting firms are located near major ports. However, the figure also shows that there are regions with many exporting firms that are not located near a major port.

The top destinations for Russian exports in 2003 were, in order: China, United States, United Kingdom, Japan, Ukraine, Kazakhstan, Turkey,
Netherlands, Germany, and Iran. Of these, only Ukraine and Kazakstan were part of the Soviet Union. Additionally, Russia does not currently have a Warsaw Pact nation as a top destination. (Soviet trade was with countries of similar ideology such as Yugoslavia and Poland.) While the Council for Mutual Economic Assistance (Comecon) from 1949-1991 facilitated economic exchanges with Eastern Bloc countries and other socialist countries, because we do not see current evidence of favored trade relations with former Soviet or Eastern bloc countries, we do not treat them any differently than Western countries. In a robust check in the main paper, we remove Warsaw Pact and Former Soviet Union countries from our analysis.

Federal laws in Russia severely limit the ability of Russian firms to change region (Botolf 2003). Because of this, firms choose how much to export to each country in the world, taking their location as given. Furthermore, because the Soviet Union did not trade with western countries, we believe it is likely that the Soviet central planners chose firm locations for reasons other than the expansion of trade (Bradshaw 2008; Huber, Nagaev, and Wörgötter 1997). Other evidence that international trade with Western countries was not important in Soviet economic planning is that the large Pacific port of Vladivostok, home of the Soviet Pacific Fleet, was closed to foreign vessels until 1991.

C Exports in the context of output

If external geography played no role in the location of exporters, the pattern of region-country aggregate exports ($X_{ij}$) would be identical to the pattern of regional production ($Y_i$) with some randomness. In this case, the destination of
Russian sales would not matter for the location of Russian exporters. To show that this does not hold, we conduct two simple tests. The first test is to run the following regression:

$$\log X_{ij} = -5.40 + 0.44 \log Y_i$$  \hspace{1cm} (1)

$$N = 2985, \quad R^2 = 0.03, \quad RMSE = 2.62$$

The index $i$ runs over the set of Russia’s regions and the index $j$ runs over the set of countries in the world. Standard errors are robust and * indicates the estimated coefficient is significantly different from zero with 99% confidence. The extremely low explanatory power as given by the $R^2$ indicates that the pattern of production in Russia does not account for the pattern of exports in the data.

The second test is to compare the ratio of exports from each Russian region to the ratio of GDP from each region to Russia. We modify the familiar concept of location quotient to put Russian exports by region into the context of overall economic activity,

$$LQ_i = \frac{X_i}{Y_i}$$

If the LQ is greater than 1 then the region’s export share to the world is larger than it’s share of GDP. This indicates the region is relatively specialized in exporting in comparison to overall economic activity. The strength of this approach is that it controls for the geography of economic activity within Russia,

The average $LQ_i$ is 1.454 with standard deviation 0.489. But though the average $LQ_i$ is not significantly different from one, the regions with the largest
$LQ_i$ are several times the standard deviation. Though most region’s $LQ_i$ is around 1, there are many that are in the right tail. Thus there are regions in Russia that are heavily concentrated in exporting compared to overall output and regions that do not export anything at all.

Because we only consider manufacturing activity, we repeat the proceeding exercise by scaling the regional GDP data by the share that is from industry. The results (not reported) are not qualitatively different from before. Therefore, we have documented that Russian exporters are not simply located in the geographic pattern of industrial economic activity.

We also use the balls & bins method of Armenter and Koren (2010) to assign firms to each region. We show a histogram in figure 2 of regional firm distribution and then test to see if the actual distribution is different than one with firms distributed randomly across bins (weighted by regional exports). We use a
two-sample Komorgorov-Smirnov test for equality and reject that the samples are the same with an exact p-value of 0.013.

**D Robustness**

One feature of the Russian export data is the high frequency of zeros. There are 89 regions in Russia and we have 2003 GDP data (IMF 2006) and distance for 175 countries. Therefore, there are 15,575 possible observations. Of these, only 2991 have positive exports, or 19%. (We also lose some observations because of missing regional GDP data from Russia.) Santos Silva and Tenreyro (2006) show that an OLS estimator on a gravity-type equation as we have in the main paper can bias estimates when there are many zeros in the data. We follow them and use the poisson quasi-maximum likelihood estimator.

\[
\frac{X_{ij}}{Y_iY_j} = \exp \left( -20.06^* \times \log D_{ij}^{0.104} + \sum_{i=2}^{89} \kappa_i S_i + \sum_{j=2}^{175} \delta_j T_j \right).
\]

\(N = 15400, \quad AIC = 526.00\)

\[
\frac{X_{ij}}{Y_iY_j} = \exp \left( -12.695^* \times \log D_{ij}^{0.119} \times \log W_{ij}^{0.027} + \sum_{i=2}^{89} \kappa_i S_i + \sum_{j=2}^{175} \delta_j T_j \right).
\]

\(N = 2985, \quad AIC = 456.00\)

Though the estimates for the regression with the agglomeration term change quantitatively, they are the same qualitatively. The agglomeration term is economically and statistically significant. Importantly, the model with the agglomeration term has a lower Akaike Information Criterion than the model without the term, indicating the model with agglomeration is preferred. However, region-country observations of zero exports also have zero weight for those
exports. Santos Silva and Tenreyro do not consider the model where zeros occur on both the left and right hand side of the regression.

E Application to U.S. state exports

Though we established the facts about agglomeration of exporters around the destination of their shipments and motivated our model using transaction-level customs data from Russia, there is nothing in the reduced form equation that requires customs data. Therefore, we take our model to another regional export data set. The data set is the Origin of Movement (OM) export data for U.S. states. We do this to check that our previous results are applicable to a wider set of countries than Russia because of possible idiosyncratic issues with the Russian data.

A detailed description of the OM data is in Cassey (2009). Because of his findings on data quality, we limit our observations to manufactured exports only. We also restrict our sample to 2003 to match the Russian data. The same sample of countries is used except that Russia as a destination replaces the United States. Also there are 22 observations which have positive export sales but a weight of zero. We do not use these observations.

The results are reported in table 1. Beginning with the benchmark gravity equation in column (A), we find the estimated coefficient on distance is much larger than for the Russian data. For us, the relevant statistic is $\hat{R}^2 = .54$. In Column (B), when we apply the U.S. data to our gravity equation including aggregate weight, the $\hat{R}^2$ increases from .54 to .75, an increase of 40.2%. (The results for the data pooled over 1997 to 2008 are essentially the same.) As with
Table 1: OLS estimates on U.S. data

<table>
<thead>
<tr>
<th>Var.</th>
<th>(A)</th>
<th>(B)</th>
<th>(C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{ij}$</td>
<td>-2.05*</td>
<td>-0.95*</td>
<td>-1.50*</td>
</tr>
<tr>
<td></td>
<td>(.137)</td>
<td>(.102)</td>
<td>(.124)</td>
</tr>
<tr>
<td>$W_{ij}$</td>
<td>0.433*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.006)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{lag}W_{ij}$</td>
<td>0.243*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(.010)</td>
</tr>
<tr>
<td>$N$</td>
<td>7061</td>
<td>7061</td>
<td>2985</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.54</td>
<td>.75</td>
<td>.60</td>
</tr>
<tr>
<td>RMSE</td>
<td>1.28</td>
<td>0.93</td>
<td>1.15</td>
</tr>
</tbody>
</table>

Notes: The benchmark gravity equation in (A): $\log X_{ij} Y_{ij} = \alpha - \beta_1 \log D_{ij} + \sum_{i=2}^{50} \kappa_i S_i + \sum_{j=2}^{175} \delta_j T_j + \epsilon_{ij}$. Exporter and importer binary variables and a constant are estimated but not reported. Standard errors (in parentheses) are robust. Specification (B) adds the aggregate shipping weight. Specification (C) lags aggregate weight.

* indicate p-values less than .01.

the Russian data, we can confidently reject that $\eta = 0$ ($F(1, 6835) = 81.87$).

Furthermore, the estimated coefficients using the U.S. data are similar to those using the Russian data whereas the same is not true for the benchmark. We estimate $\eta = 0.454^*$. Another encouraging result is that our estimate for $\gamma$ is 0.95 and is again not statistically different from one, as with the Russian data in the main paper. Furthermore, our point estimate is close to the 0.99 to 1.10 range for U.S. data reported in Axtell (2001) whereas the benchmark is too high. Finally, with the U.S. data, we estimate $\sigma = 1.5$.

One concern with the results so far is that they are driven by the mechanical relationship between export value and weight. We estimated coefficients at the product level and used aggregate weight less product weight in order to break this
relationship. Nevertheless, we take advantage of additional data for the U.S. and we follow Koenig (2009) and use aggregate export weight from the previous year, 2002. This severs any mechanical relationship since past aggregate weight cannot affect current export value in any way other than a reduction of the transaction costs of trade. We lose some observations compared to 2003 because those state–country pairs did not trade in 2002. The results are reported in table 1 column (C). In this case, we find that our model accounts for 12% more variation in the data than the benchmark. The estimate on the aggregate weight term is statistically significant, though smaller than previously. Therefore, we find that the agglomeration term is important even if it is lagged one year.

References


Huber P, Nagaev S, Wörgötter A (1997) The changes in location of Russia
industry in early transition 1987–1993, Institute for Advanced Studies, Vienna,
East European Series no. 42

IMF (2006) World Economic Outlook Database. International Monetary Fund,
(accessed December 15, 2006)

of Urban Economics 66(3):186–195