Manufacturing (co)agglomeration in a transition country: Evidence from Russia

Ekaterina Aleksandrova [1] Kristian Behrens [1, 2] Maria Kuznetsova [1]

[1]National Research University Higher School of Economics

[2]Department of Economics, Université du Québec à Montréal, Canada and CEPR, UK

10 April, 2019

Overview

Motivation

Realizing productivity gains from geographic concentration may be especially important for transition countries, such as Russia.

According to the Global Competitiveness Report 2014-2015, Russia ranks 119 out of 144 countries in terms of goods market competition and efficiency.

Deloitte's 2016 Global Manufacturing Competitiveness Index, Russia ranks 32 out of 40 countries.

Questions

- How concentrated are manufacturing industries in Russia ?
- Which industries are more concentrated than others ?
- What are the potential mechanisms explaining that differential concentration ?

Key Results

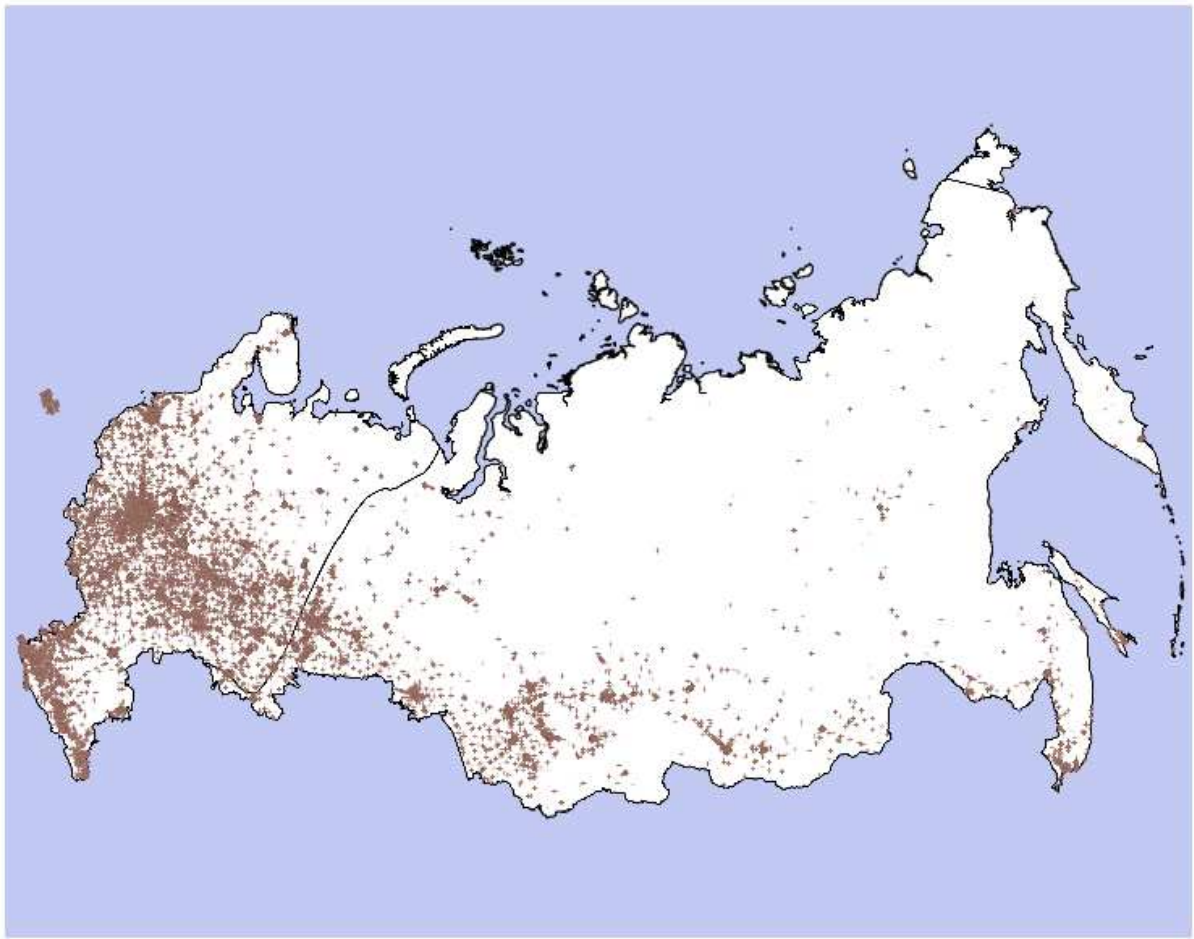
- About 80% of 3-digit industries are significantly agglomerated, with a substantially higher share in the European part than in the Asian part of Russia ;
- About 80% of industry pairs are significantly coagglomerated, mainly at short distances below 100 km and at distances between 650-800 km (the distance between Moscow and Saint Petersburg regions) ;
- The overall patterns of geographic concentration are similar to the results for other countries – UK, Canada, or the U.S. ;
- Stronger buyer-supplier links and lower transport costs are associated with more geographic concentration ;
- Industry pairs with stronger input-output links tend to be more coagglomerated only if transporting their outputs is more costly.

Data

- Data Source : RUSLANA database from Bureau Van Dijk Electronic Publishing ;
- Identification of companies and establishments is based on Taxpayer's Identification Number (INN) and the All-Russian Classifier of Enterprises and Organizations (OKPO) pairs ;
- The manufacturing sector is delimited by Russian Classification of Economic Activities (OKVED 1.1) 15.00.00 to 37.20.70 ;
- We retain all establishments that were active in 2014 and whose contact information were updated after 2012 ;
- We end up with 320,934 accurately geocoded establishments.

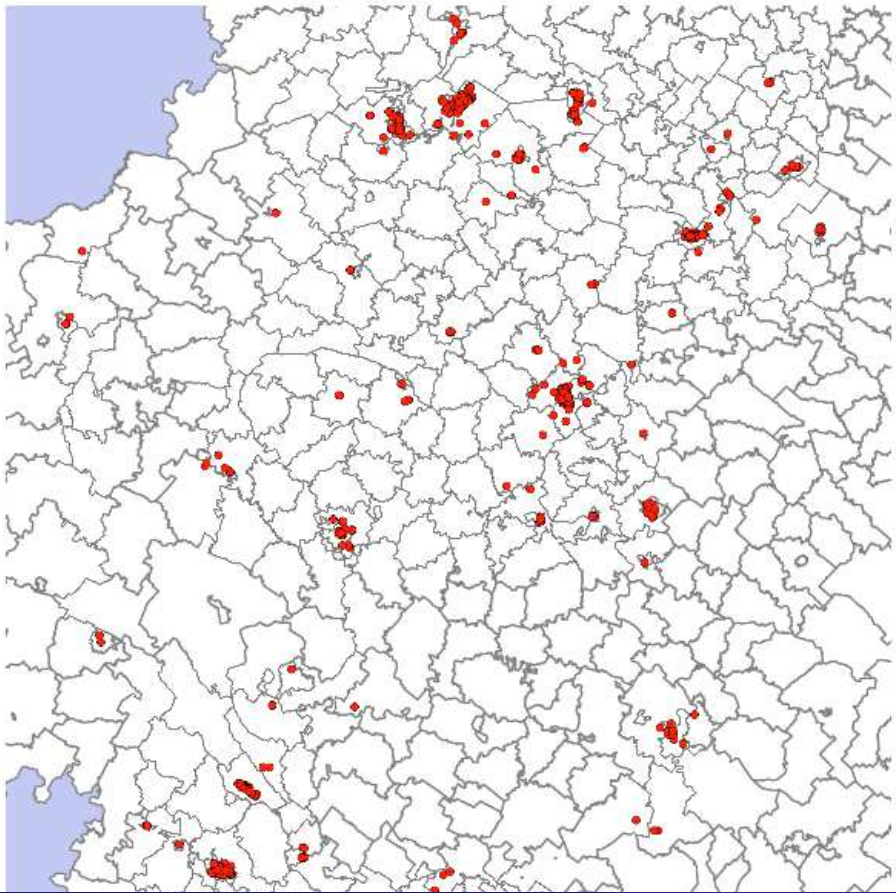
Distribution of manufacturing plants in Russia

Distribution of manufacturing plants in Russia in 2014 and east-west divide

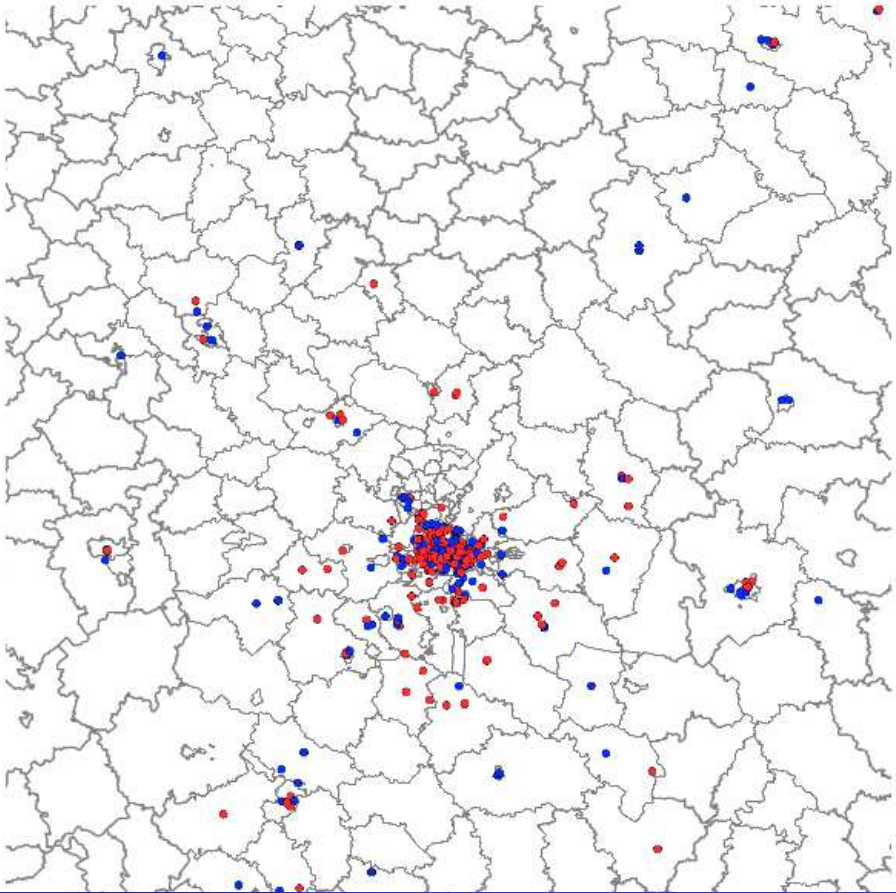


Examples of agglomeration and coagglomeration patterns

(a) 'Manufacture of motor vehicles, trailers and semi-trailers'



(b) 'Spinning of textile fibres' and 'Weaving manufacture'



Agglomeration : Methodology

We follow Duranton and Overman (2005, 2008) to estimate a kernel-smoothed distribution (K-density) of the bilateral distances between plants, which is used to :

- measure *relative geographic concentration*, industries that display significant geographic concentration *relative* to manufacturing in general ;
- measure the *absolute geographic concentration*, the share of bilateral distances between plants in that industry below the distance threshold d.

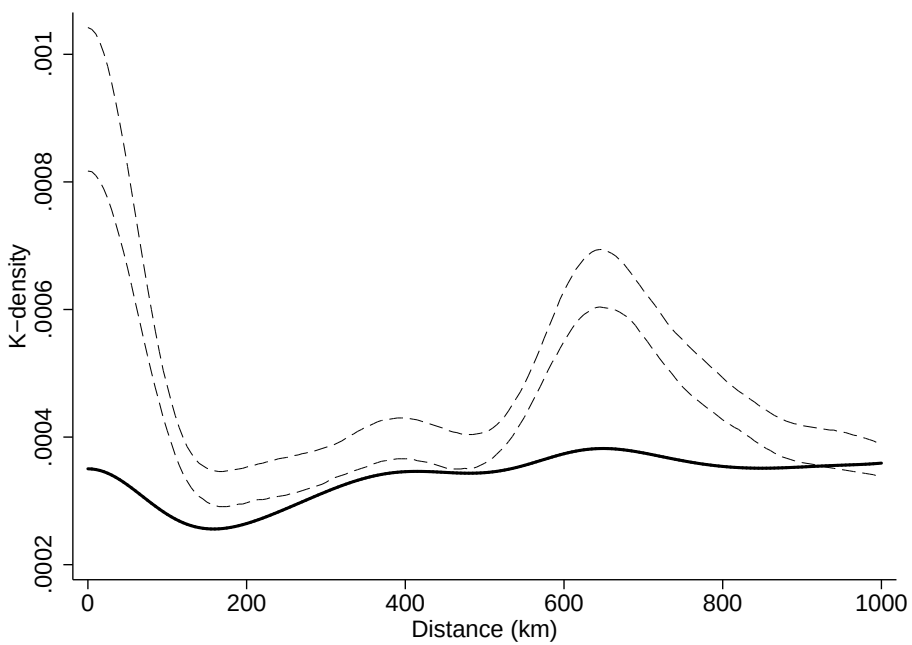
Agglomeration : Methodology

The methodology developed by Duranton and Overman (2005, 2008) has four steps :

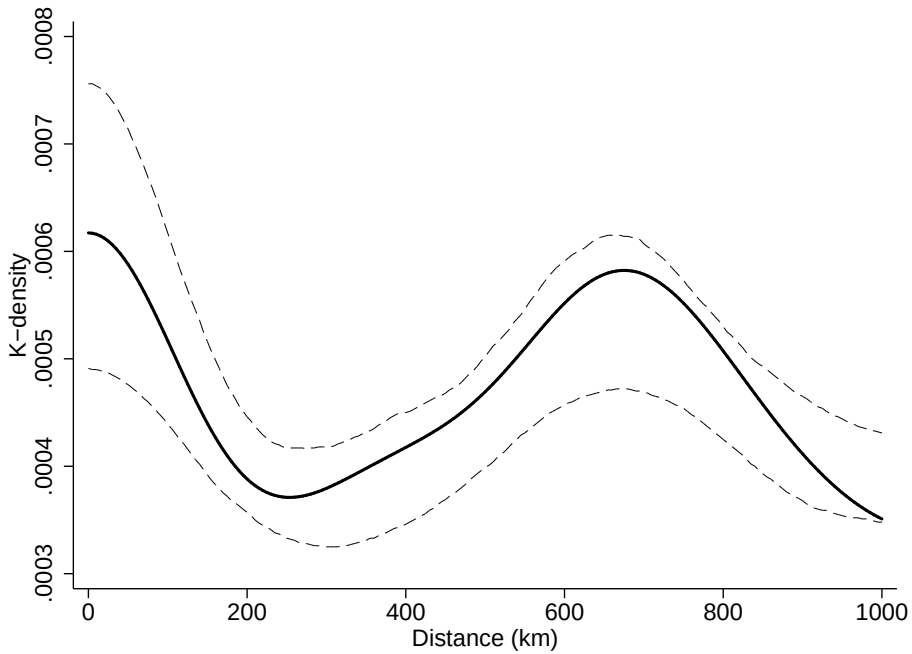
1. Compute pairwise distances between all plants in an industry and estimate a kernel density of their distribution.
2. Construct a counterfactual distribution by assuming that the plants in a given industry are randomly reallocated among all possible locations where we observe manufacturing activity. Use that distribution to estimate a counterfactual kernel density.
3. Assess whether the observed location patterns depart statistically significantly from randomness.
4. Test whether an industry is significantly localized/dispersed or not significantly different from random allocation.

Agglomeration : K-densities and confidence bands for different representative location patterns

(a) 'Manufacture of beverages'

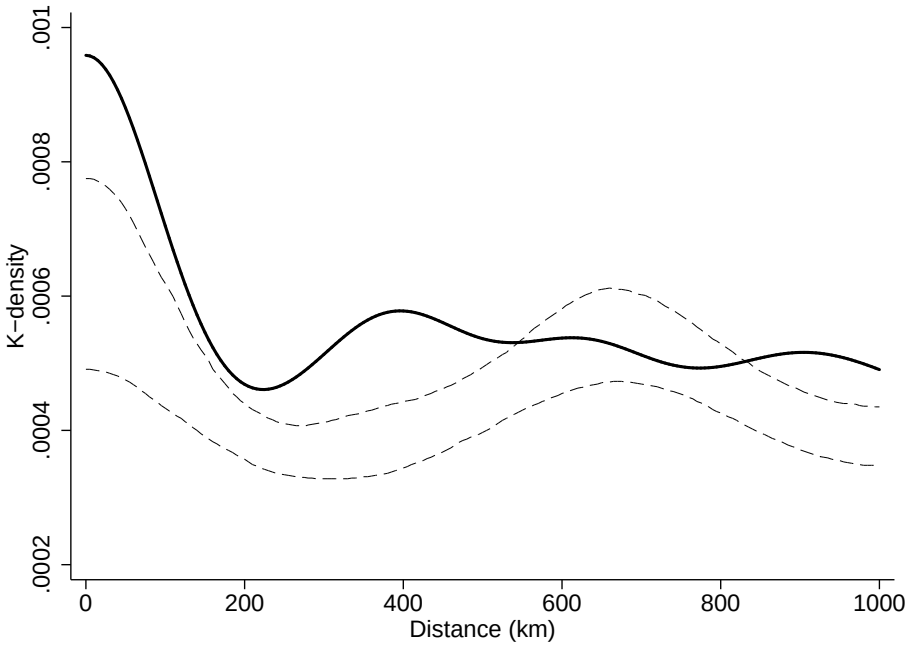


(b) 'Cast iron and steel, other primary processing

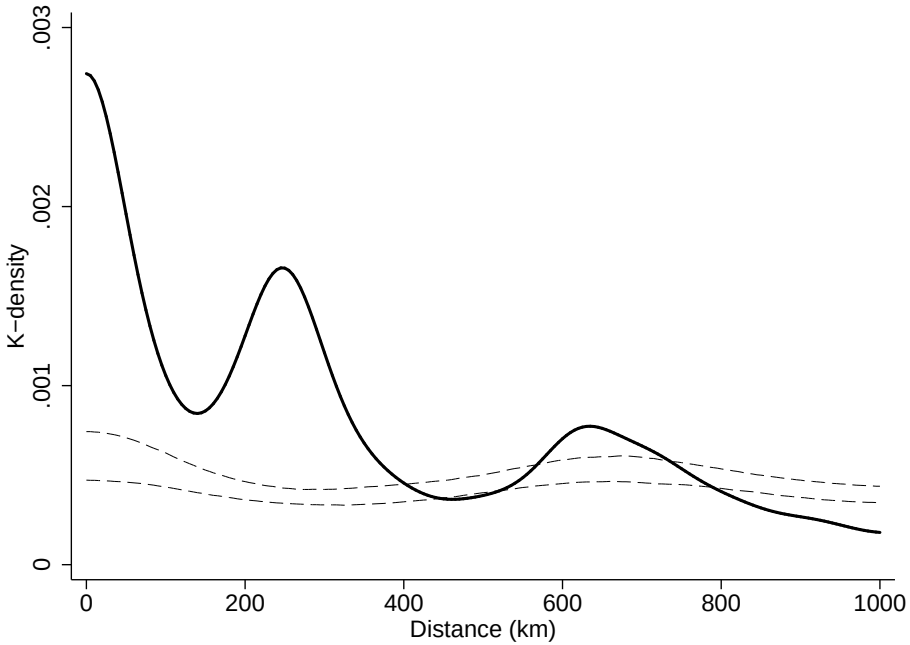


Agglomeration : K-densities and confidence bands for different representative location patterns

(c) 'Manufacture of motor vehicles'



(d) 'Weaving manufacture'



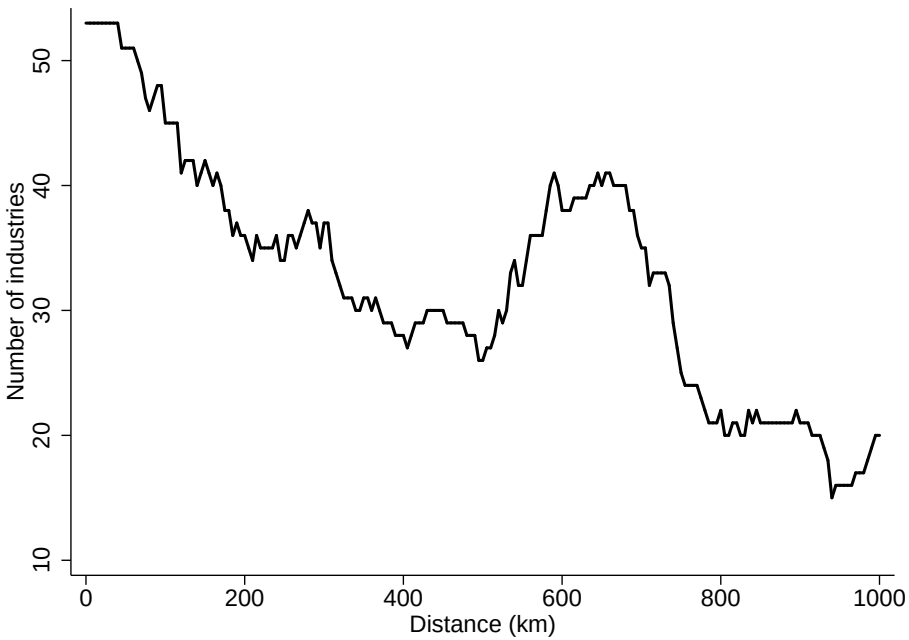
Agglomeration : Summary of geographic concentration patterns

	(a) All of Russia $N = 316,967$ plants		(b) Western Russia $N = 245,616$ plants		(c) Eastern Russia $N = 71,351$ plants	
Status	Number	Percentage	Number	Percentage	Number	Percentage
Localized industry	81	80.20%	89	88.12%	67	66.34%
Random	6	5.94%	6	5.94%	25	24.75%
Dispersed industry	14	13.86%	6	5.94%	9	8.91%
Excess localization $\overline{\Gamma}\|_{\Gamma_i>0}$	0.063		0.057		0.052	
Excess dispersion $\overline{\Psi}\|_{\Psi_i>0}$	0.043		0.028		0.013	
Total	101	100%	101	100%	101	100%

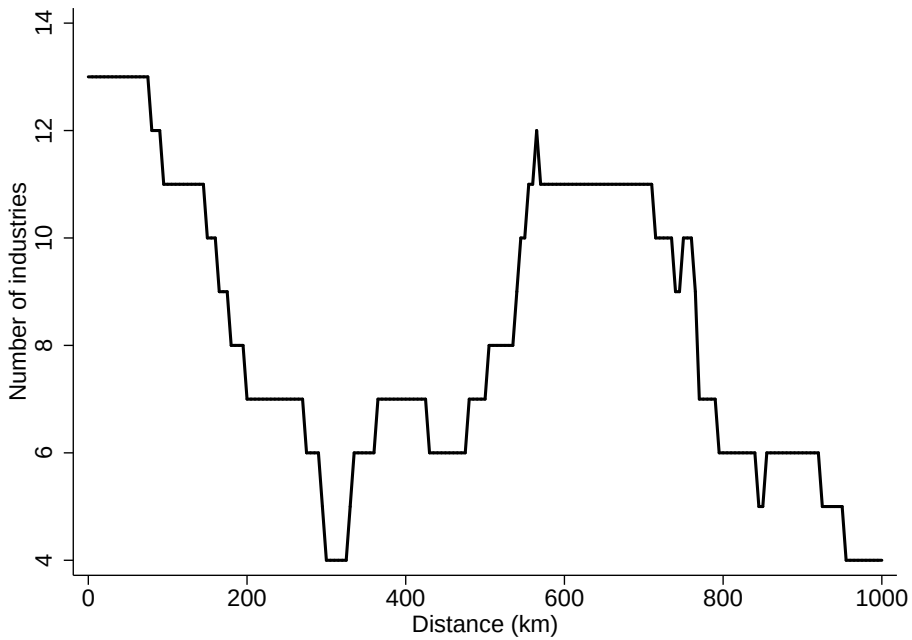
Notes : All K-densities are computed over a range of 0–1000 kilometers, for 101 3-digit OKVED industries. The confidence bands are computed using 1,000 bootstrap replications. We compute the K-densities in 5 kilometers steps. The values of $\overline{\Gamma}|_{\Gamma_i>0}$ and $\overline{\Psi}|_{\Psi_i>0}$ are computed at the last point at which the K-densities are evaluated, i.e., $\overline{d} = 1000$km. We report average values for all significantly localized industries in the case of $\overline{\Gamma}|_{\Gamma_i>0}$, and for all significantly dispersed industries in the case of $\overline{\Psi}|_{\Psi_i>0}$.

Agglomeration : Localization patterns by distance

(a) Significant localization by distance



(b) Significant dispersion by distance



Agglomeration : Top-10 most localized and most geographically concentrated 3-digit industries

OKVED	Industry name	
	Top-10 most localized industries	Γ_i
172	Weaving manufacture	0.323
353	Manufacture of aircraft and spacecraft	0.284
223	Reproduction of recorded media	0.273
176	Manufacture of textile fabrics	0.252
244	Manufacture of pharmaceuticals	0.230
173	Finishing of textiles	0.217
362	Manufacture of jewellery, medals and related articles of precious metals and stones ; manufacture of coins	0.204
171	Spinning of textile fibres	0.203
321	Manufacture of electronic and radio components, electrovacuum devices	0.191
343	Manufacture of parts and accessories for motor vehicles and their engines	0.175
	Top-10 most geographically concentrated industries	CDF
172	Weaving manufacture	0.785
353	Manufacture of aircraft and spacecraft	0.725
176	Manufacture of textile fabrics	0.708
363	Manufacture of musical instruments	0.705
223	Reproduction of recorded media	0.696
171	Spinning of textile fibres	0.686
343	Manufacture of parts and accessories for motor vehicles and their engines	0.681
173	Finishing of textiles	0.667
321	Manufacture of electronic and radio components, electrovacuum devices	0.666
192	Manufacture of luggage, handbags and the like, saddlery and harness	0.654

Notes : Γ_i and the K-density CDF are computed at 1,000 kilometers distance. We hence measure localization and geographic concentration over the whole distance range that we compute the K-densities for.

Coagglomeration : Methodology

Duranton and Overman (2005, 2008) methodology can be adapted to the coagglomeration of two different industries.

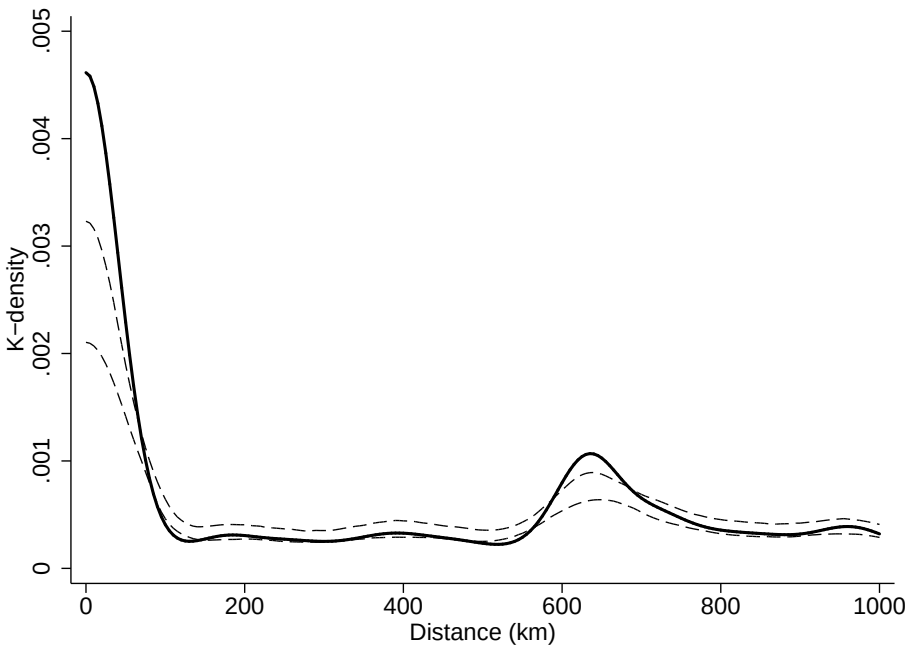
We estimate K-densities for the distribution of bilateral distances between establishments *in different industries*.

The counterfactual distributions for the coagglomeration of industries are constructed based on the locations that contain establishments of either of two industries only.

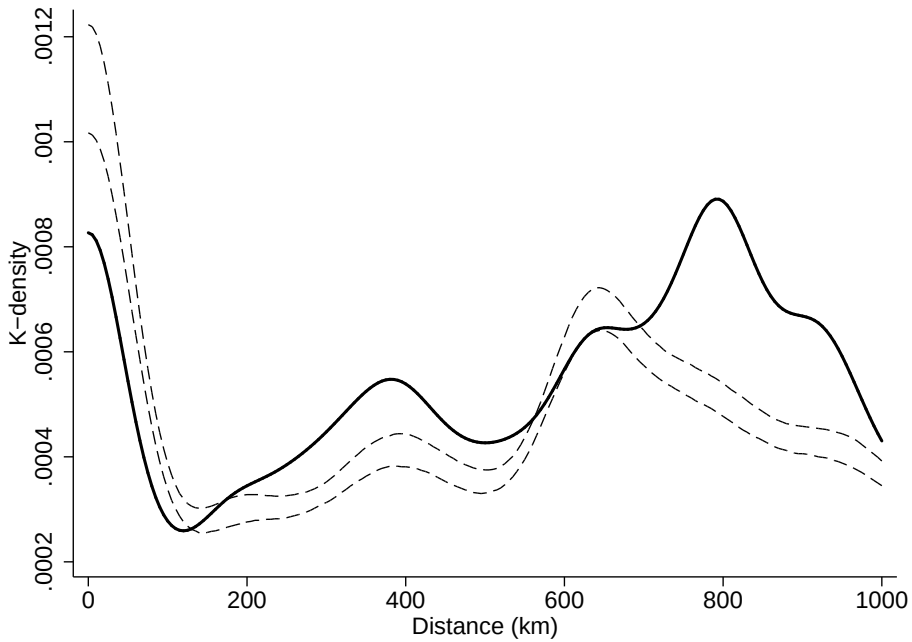
The strength of coagglomeration measures how much closer establishments in the two industries are from each other than from establishments in the two industries in general.

Coagglomeration : K-densities and confidence bands of selected OKVED 3-digit industry pairs

(a) 'Publishing' & 'Reproduction of recorded media'

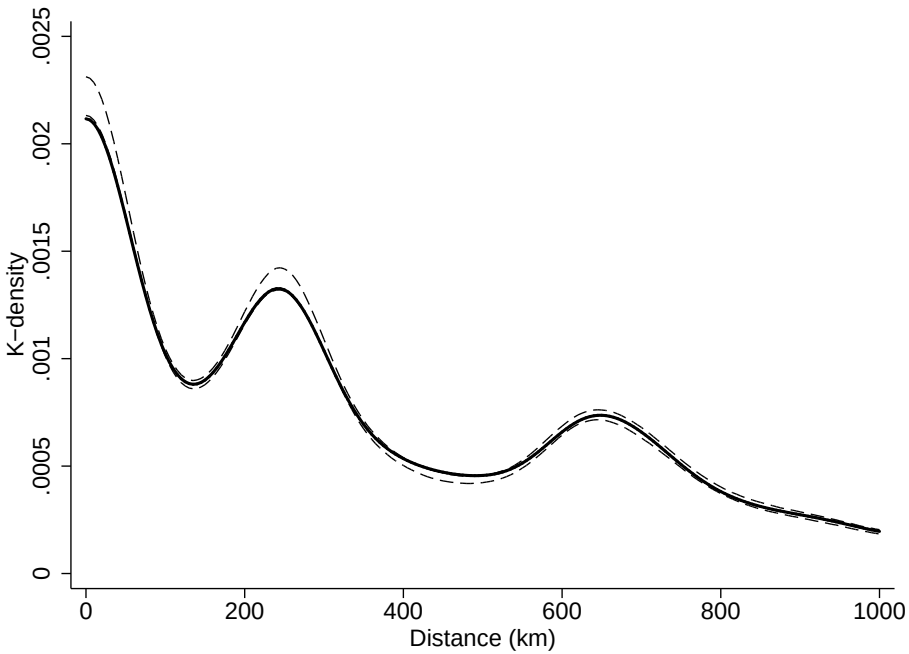


(b) 'Manufacture of other general purpose machinery' & 'Manufacture of parts and accessories for motor vehicles and their engines'

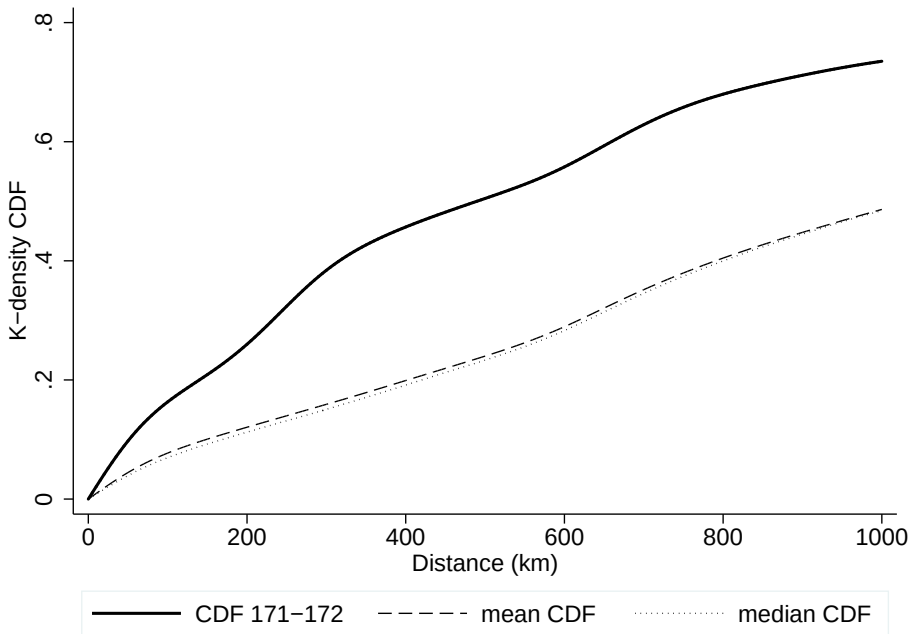


Coagglomeration : K-densities, confidence bands, CDFs for selected OKVED 3-digit industry pairs

(a) 'Spinning of textile fibres' & 'Weaving manufacture'



(b) K-density CDF for 'Spinning of textile fibres' & 'Weaving manufacture'



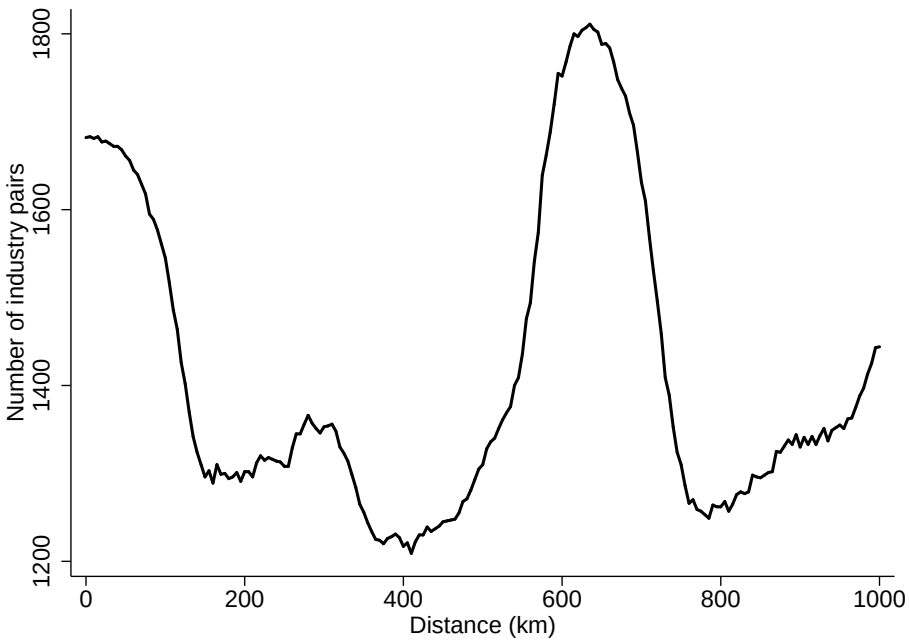
Coagglomeration : Summary statistics of K-density estimates for coagglomeration patterns

	# of industry pairs	Percentages
(a) Coagglomeration status		
Significantly coagglomerated	4,109	81.37%
Random	269	5.33%
Significantly codispersed	672	13.31%
Total	5,050	100%
$\overline{\Gamma}\|_{\Gamma_i>0}$	0.012	
$\overline{\Psi}\|_{\Psi_i>0}$	0.011	
(b) Type of coagglomeration		
Coagglomerated on 0–170km, but not on 550–750km	1,000	30.70%
Coagglomerated on 550–750km, but not on 0–170km	725	22.26%
Coagglomerated on 0–170km and on 550–750km	1,543	47.04%

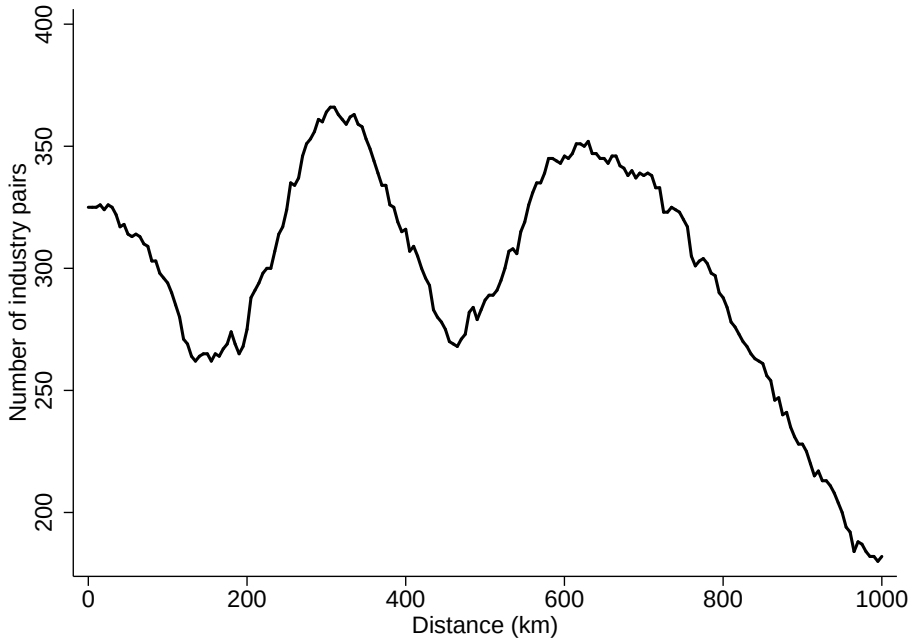
Notes : All K-densities are computed for a range of 0–1000 kilometers for 5,050 3-digit industry pairs. The values of $\overline{\Gamma}|_{\Gamma_i>0}$ and $\overline{\Psi}|_{\Psi_i>0}$ are computed at the last point at which the K-densities are evaluated, i.e., 1,000km. We report average values for all significantly localized industry pairs in the case of $\overline{\Gamma}|_{\Gamma_i>0}$, and for all significantly dispersed industry pairs in the case of $\overline{\Psi}|_{\Psi_i>0}$. The bottom panel provides the breakdown of all coagglomerated industry pairs on 0–170 kilometers and on 550–750 kilometers.

Coagglomeration patterns by distance

(a) Significant localization by distance



(b) Significant dispersion by distance



What are the potential drivers of coagglomeration in Russia ?

Industries may (co)agglomerate because of :

- buyer-supplier links ;
- access to common labour 'pool' ;
- transportation costs ;
- knowledge exchange and information sharing.

We follow empirical strategy pioneered by Ellison et al. (2010).

We test whether industries that are more connected along some dimensions tend to colocate more.

2SLS instrumental variables regressions, using Canadian and U.S. instruments.

	Average			Max		
	(1)	(2)	(3)	(4)	(5)	(6)
Sample			Exclude3			Exclude3
			Second-stage results			
Average input coefficient (Russia)	0.045	0.064^{c}	0.122^{a}			
	(0.036)	(0.034)	(0.042)			
Average AV transport cost		-0.248^{a}	-0.243^{a}			
		(0.017)	(0.018)			
Maximum input coefficient (Russia)				0.048	0.045	0.103^{b}
				(0.035)	(0.033)	(0.043)
Maximum AV transport cost					-0.214^{a}	-0.208^{a}
					(0.017)	(0.017)
Observations	4,242	4,242	4,053	4,242	4,242	4,053
			First-stage results			
Average input coefficient (Canada)	0.337^{a}	0.337^{a}	0.449^{a}			
	(0.064)	(0.064)	(0.071)			
Average input coefficient (U.S.)	0.415^{a}	0.416^{a}	0.318^{a}			
	(0.077)	(0.077)	(0.072)			
Maximum input coefficient (Canada)				0.278^{a}	0.277^{a}	0.373^{a}
				(0.062)	(0.063)	(0.064)
Maximum input coefficient (U.S.)				0.361^{a}	0.361^{a}	0.265^{a}
				(0.070)	(0.070)	(0.061)
Observations	4,242	4,242	4,053	4,242	4,242	4,053
Hansen-J (p-val)	0.557	0.660	0.132	0.290	0.412	0.205
F-stat	49.16	48.93	43.18	36.23	36.12	36.33
R^2	0.355	0.356	0.268	0.324	0.325	0.253

Notes : Coefficients significant at : ^{a}1% ; ^{b}5% ; and ^{c}10%. We only report the key coefficients in the first stage and do not report the others. The dependent variables is the DO coagglomeration K-density at 100 kilometers distance. All variables are standardized. 'Exclude 3' regressions exclude all 4-digit industry pairs within the same 3-digit industries.