Does proximity matter in the choice of partners in collaborative R&D projects?

An empirical analysis of granted projects in Germany

Philipp Marek, Clemens Fuhrmeister, Mirko Titze, Ulrich Blum

Workshop Regionalökonomik, Dresden - September 11, 2015

Institut für Wirtschaftsforschung Halle
Universität Bremen
MARTIN-LUTHER UNIVERSITÄT HALLE-WITTENBERG

Outline

1. Introduction
2. Empirical Approach
    - Data set
    - Spatial Autocorrelation
    - Estimation strategy
3. Results & Conclusion
4. Appendix

Introduction

Introduction

Proximity and knowledge diffusion

- Knowledge diffusion is considered as a key determinant for regional growth
- Different channels through which knowledge can be transferred and measured

Introduction

Proximity and knowledge diffusion

- Knowledge diffusion is considered as a key determinant for regional growth
- Different channels through which knowledge can be transferred and measured
  - patents, publication, (R&D)-collaborations

Introduction

Proximity and knowledge diffusion

- Knowledge diffusion is considered as a key determinant for regional growth
- Different channels through which knowledge can be transferred and measured
  - patents, publication, (R&D)-collaborations
- Proximity is considered as a key determinant for knowledge transfer
- Boschma (2005) five dimensions of proximity

Introduction

Proximity and knowledge diffusion

- Knowledge diffusion is considered as a key determinant for regional growth
- Different channels through which knowledge can be transferred and measured
  - patents, publication, (R&D)-collaborations
- Proximity is considered as a key determinant for knowledge transfer
- Boschma (2005) five dimensions of proximity
  - geographical, cognitive, institutional, organizational and social proximity

Previous findings on R&D collaborations

- Different applicants between EU FP and national schemes (Brökel & Graf, 2012)
- Public research organizations are overrepresented in EU FP
- SMEs apply for national rather than for EU grants
- Most studies refer to granted R&D collaboration projects from the EU Framework Programmes (e.g. Scherngell and Lata, 2013)
- Spatial interaction model accounting for spatial autocorrelation
- Evidence for the importance of proximity

Previous findings on R&D collaborations

- Different applicants between EU FP and national schemes (Brökel & Graf, 2012)
- Public research organizations are overrepresented in EU FP
- SMEs apply for national rather than for EU grants
- Most studies refer to granted R&D collaboration projects from the EU Framework Programmes (e.g. Scherngell and Lata, 2013)
- Spatial interaction model accounting for spatial autocorrelation
- Evidence for the importance of proximity

Research Gap

National support scheme: Analysis on actors, which have been less prominent in empirical research so far

Description of the Data set

Data Set

- Data: Granted R&D collaboration
- Soucre: Förderkatalog provided by the German Federal Ministry of Education and Research (BMBF)
- Subject of analysis: Cross-regional R&D collaboration intensity
- Regional level: 402 NUTS-3 regions (Kreise)
- Period of investigation: 2006 to 2012
- Collabortion Projects funded: 7,111 with 29,933 pariticipating actors
- Sum of cross-regional linkages: 159,376
- Number of regional pairs with at least one collaboration: 56,807 out of 1,131,228 possible pairs

Top ten regional pairs showing the highest number of collaborations

Region 1	Region 2	Number of collaborations
Munich (city)	Munich (district)	900
Munich (city)	Berlin	382
Hamburg	Berlin	313
Berlin	Potsdam	277
Stuttgart	Berlin	244
Munich (district)	Berlin	237
Stuttgart	Munich (city)	231
Berlin	Dresden	225
Hanover	Berlin	220
Heidelberg	Berlin	209

Research questions

Does the importance of geographical proximity remain high if other proximity measures are included in the model?

Do proximity measures have linear or non-linear effects?

Empirical Approach

Dependent Variable

Number of cross-region collaborations in granted R&D collaboration projects

- Out of the Förderkatalog, we extract a collaboration matrix for each year, C with 402 x 402 elements
- c_{ij} represents the amount of granted collaborations with partner from region i and j
- Symmetric collaboration matrix, with $c_{ij}=c_{ji}$

Empirical equation

$$c_{ij} = a_i^{\delta_1} * b_j^{\delta_2} * \sum_{z=1}^{Z} \beta_z g_{ij}^{(z)} + \epsilon_{ij}$$

- a_i - Origin variables
- b_j - Destination variables
- g_{ij} - Separation variables

Determinants of the cross-regoinal collaboration intensity - I

Proximity measures - Boschma (2005)

- **Geographical proximity**: Euclidean distance between capital cities of two regions
- **Cognitive proximity**: One minus the Pearson correlation coefficient of the vectors of the relative industry shares between region i and j. (Employment statistics for 60 NACE Rev 1.1 industries provided by the Federal Employment Agency)
- **Institutional proximity**: One minus the Pearson correlation coefficient of the vectors of the voting shares of poltical parties in the federal elections between region i and j. (Federal Returning Officer)

Determinants of the cross-regoinal collaboration intensity - I

Proximity measures - Boschma (2005)

- **Geographical proximity**: Euclidean distance between capital cities of two regions
- **Cognitive proximity**: One minus the Pearson correlation coefficient of the vectors of the relative industry shares between region i and j. (Employment statistics for 60 NACE Rev 1.1 industries provided by the Federal Employment Agency)
- **Institutional proximity**: One minus the Pearson correlation coefficient of the vectors of the voting shares of poltical parties in the federal elections between region i and j. (Federal Returning Officer)
- **Organizational** and **social proximity** are omitted due to the lack of appropriate proxies

Determinants of the cross-regional collaboration intensity - II

Additional proximity measures

- **Border dummy** indicating whether one of the regions is located at the German land border with another country
- **Neighbor dummy** indicating whether region i and region j share a common border
- **Intra-regional dummy** indicating whether regional pair reflects an intra-regional linkage

Gravity parameters: Origin and Destination variables

- The number of employees in the regions under analysis (Federal Employment Agency)
- The number of establishments in the regions under analysis (Federal Employment Agency)

Figure: Histrogram of cross-region R&D collaborations

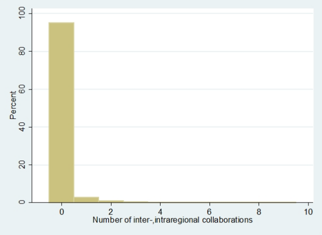
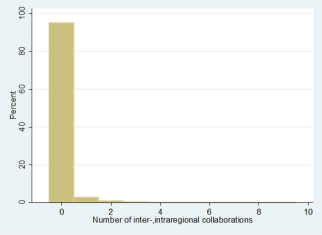


Figure: Histrogram of cross-region R&D collaborations

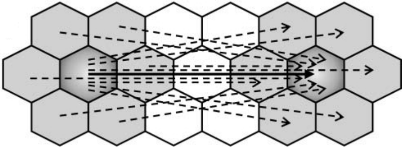


Properties of dependent variable

- Distribution: Count structure with long right tail
- Excess zeros: 95% of regional pairs do not report any collaborations
- Balanced panel structure

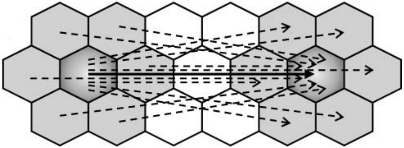
⇒ Longitudinal count model accounting for excess zeros

Figure: Spatial autocorrelation in spatial interaction model



Source: Chun & Griffith (2011)

Figure: Spatial autocorrelation in spatial interaction model



Source: Chun & Griffith (2011)

Eigenvector Spatial Filter (ESF) suited for spatial interaction data

1. Conceptual Approach: Griffith & Chun (2014)
2. Application to spatial interaction data: Scherngell & Lata (2013)
3. Application to spatial interaction data with a longitudinal structure: Lata, Scherngell & Brenner (2015)

Eigenvector Spatial Filter

Major purpose of the Eigenvector Spatial Filter approach

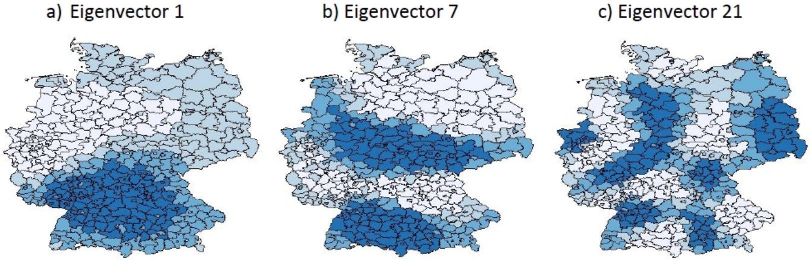
Separation of spatially structured random component from the error term

Construction of Eigenvector Spatial Filter

- Transformation of neighborhood matrix W
  $W^T = (\boldsymbol{I} - \mathbf{1}'\mathbf{1} * \mathbf{1}/n)W(\boldsymbol{I} - \mathbf{1}'\mathbf{1} * \mathbf{1}/n)$
- Eigenvectors **E**=($\mathbf{E}_1$, $\mathbf{E}_2$, ..., $\mathbf{E}_n$) and their corresponding eigenvalues $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_n)$ can be extracted from W^T
- Each eigenvector represents a synthetic map of spatial concentration, whose degree of concentration captured by the corresponding eigenvalues (only real values)
- Properties: Independence **E*E'=I** and zero mean **E'*1=0**

Eigenvector Spatial Filter

Figure: Eigenvectors derived from neighborhood matrix W



Eigenvector Spatial Filter

Selection of Eigenvectors

- $\sum_{o=1}^{O} E_o$ eigenvector spatial filters for the origin regions
- $\sum_{d=1}^{D} E_d$ eigenvector spatial filters for the destination regions
- Only eigenvectors with a Moran's Coefficient above 0.25, $MC_i = n/(\mathbf{1}'W\mathbf{1}) * \lambda_i$ (see Fischer & Griffith, 2008)
- Reduction from 402 to 98 potential eigenvectors
- For each year, we run a regression with all eigenvectors with Moran's Coefficient larger than 0.25.
- For the panel regression, we keep all eigenvectors, which had a significant impact in all cross-sectional regression at the 1%-level. (See Lata, Scherngell, Brenner, 2015)

Symmetry of the collaboration matrix C

- Inclusion of the entire collaboration matrix in the regression
- Equality of coefficient estimates for origin and destination gravity parameters
- The same holds true for the set of eigenvectors, o(1, ..., O)=d(1, ..., D).
- $\sum_{m=1}^{M} E_m$ - set of eigenvectors entering the regression for origin and destination regions

Count Regression is based on the following equation:

$$c_{ijt} = \exp\left[\alpha_0 + \delta(\ln(a_{it}) + \ln(b_{jt})) + \sum_{z=1}^{Z} \beta_z g_{ijt}^{(z)} + \theta_m(E_{mi} + E_{mj}) + \varepsilon_{ijt}\right]$$

Estimation procedure

- Each specificatoin estimated twice:
  i) Linear impact of proximity on collaboration intensity
  ii) Squared terms of each proximity measure.
- Test for joint significance confirms non-linear relationship

Estimation procedure

- Each specificatoin estimated twice:
  i) Linear impact of proximity on collaboration intensity
  ii) Squared terms of each proximity measure.
- Test for joint significance confirms non-linear relationship
- Basic Pooled Poisson estimation
- Test for overdispersion suggests a pooled Negbin regression

Estimation procedure

- Each specificatoin estimated twice:
  i) Linear impact of proximity on collaboration intensity
  ii) Squared terms of each proximity measure.
- Test for joint significance confirms non-linear relationship
- Basic Pooled Poisson estimation
- Test for overdispersion suggests a pooled Negbin regression
- Vuong Test suggests a pooled zero-inflated Negbin regression
- Pooled zero-inflated Poisson as a robustness check

Estimation procedure

- Each specificatoin estimated twice:
  i) Linear impact of proximity on collaboration intensity
  ii) Squared terms of each proximity measure.
- Test for joint significance confirms non-linear relationship
- Basic Pooled Poisson estimation
- Test for overdispersion suggests a pooled Negbin regression
- Vuong Test suggests a pooled zero-inflated Negbin regression
- Pooled zero-inflated Poisson as a robustness check
- Panel estimation: Poisson and Nebin regressions
- Hausman-Test suggests that the random-effects model may not be applied

Estimation procedure

- Each specificatoin estimated twice:
  i) Linear impact of proximity on collaboration intensity
  ii) Squared terms of each proximity measure.
- Test for joint significance confirms non-linear relationship
- Basic Pooled Poisson estimation
- Test for overdispersion suggests a pooled Negbin regression
- Vuong Test suggests a pooled zero-inflated Negbin regression
- Pooled zero-inflated Poisson as a robustness check
- Panel estimation: Poisson and Nebin regressions
- Hausman-Test suggests that the random-effects model may not be applied
- Randon-effects specification including Mundlak-group variables of time-variant regressors

Results & Conclusion

Table: Estimation results for the pooled regressions

Cross-Regional collaborations	Negbin		Zero-Inflated Negbin Collab.	Zero-Inflated Negbin Probit
Geo. dist. in km	-0.00175***	-0.00471***	-0.00312***	0.00264***
Geo. dist. in km (sq.)		4.80e-06***	4.10e-06***	-1.55e-06***
Cognitive dist.	-1.398***	-1.137***	0.114	2.018***
Cognitive dist. (sq.)		-0.298	-2.004***	-3.004***
Institut. dist.	-0.352***	-1.098***	-1.110***	0.436***
Institut. dist. (sq.)		0.499***	0.396***	-0.322***
Border dummy	-0.310***	-0.333***	-0.204***	0.165***
Neighbouring dummy	1.003***	0.688***	0.461***	-0.359***
Intra-regional dummy	2.381***	1.893***	1.773***	-0.111
ln(employees)	3.039***	3.056***	2.068***	-1.678***
ln(establishments)	-1.745***	-1.765***	-1.454***	0.532***
Observations	1,131,228	1,131,228	1,131,228	
Log-Likelihood	-247,026.11	-246,447.57	-243,450.40	
Significant eigenvectors	13	13	12	
LR-Test Spatial Filter	17,094.42***	16,679.82***	12,123.3***	

Annual dummies and constant terms included but not reported in this table. Significance levels * 10%, ** 5%, *** 1%.

Table: Estimation results for the pooled regressions

Cross-Regional collaborations	Negbin		Zero-Inflated Negbin Collab.	Zero-Inflated Negbin Probit
Geo. dist. in km	-0.00175***	-0.00471***	-0.00312***	0.00264***
Geo. dist. in km (sq.)		4.80e-06***	4.10e-06***	-1.55e-06***
Cognitive dist.	-1.398***	-1.137***	0.114	2.018***
Cognitive dist. (sq.)		-0.298	-2.004***	-3.004***
Institut. dist.	-0.352***	-1.098***	-1.110***	0.436***
Institut. dist. (sq.)		0.499***	0.396***	-0.322***
Border dummy	-0.310***	-0.333***	-0.204***	0.165***
Neighbouring dummy	1.003***	0.688***	0.461***	-0.359***
Intra-regional dummy	2.381***	1.893***	1.773***	-0.111
ln(employees)	3.039***	3.056***	2.068***	-1.678***
ln(establishments)	-1.745***	-1.765***	-1.454***	0.532***
Observations	1,131,228	1,131,228	1,131,228	
Log-Likelihood	-247,026.11	-246,447.57	-243,450.40	
Significant eigenvectors	13	13	12	
LR-Test Spatial Filter	17,094.42***	16,679.82***	12,123.3***	

Annual dummies and constant terms included but not reported in this table. Significance levels * 10%, ** 5%, *** 1%.

Table: Estimation results for the pooled regressions

Cross-Regional collaborations	Negbin		Zero-Inflated Negbin Collab.	Zero-Inflated Negbin Probit
Geo. dist. in km	-0.00175***	-0.00471***	-0.00312***	0.00264***
Geo. dist. in km (sq.)		4.80e-06***	4.10e-06***	-1.55e-06***
Cognitive dist.	-1.398***	-1.137***	0.114	2.018***
Cognitive dist. (sq.)		-0.298	-2.004***	-3.004***
Institut. dist.	-0.352***	-1.098***	-1.110***	0.436***
Institut. dist. (sq.)		0.499***	0.396***	-0.322***
Border dummy	-0.310***	-0.333***	-0.204***	0.165***
Neighbouring dummy	1.003***	0.688***	0.461***	-0.359***
Intra-regional dummy	2.381***	1.893***	1.773***	-0.111
ln(employees)	3.039***	3.056***	2.068***	-1.678***
ln(establishments)	-1.745***	-1.765***	-1.454***	0.532***
Observations	1,131,228	1,131,228	1,131,228	
Log-Likelihood	-247,026.11	-246,447.57	-243,450.40	
Significant eigenvectors	13	13	12	
LR-Test Spatial Filter	17,094.42***	16,679.82***	12,123.3***	

Annual dummies and constant terms included but not reported in this table. Significance levels * 10%, ** 5%, *** 1%.

Table: Estimation results for the pooled regressions

Cross-Regional collaborations	Negbin		Zero-Inflated Negbin	
			Collab.	Probit
Geo. dist. in km	-0.00175***	-0.00471***	-0.00312***	0.00264***
Geo. dist. in km (sq.)		4.80e-06***	4.10e-06***	-1.55e-06***
Cognitive dist.	-1.398***	-1.137***	0.114	2.018***
Cognitive dist. (sq.)		-0.298	-2.004***	-3.004***
Institut. dist.	-0.352***	-1.098***	-1.110***	0.436***
Institut. dist. (sq.)		0.499***	0.396***	-0.322***
Border dummy	-0.310***	-0.333***	-0.204***	0.165***
Neighbouring dummy	1.003***	0.688***	0.461***	-0.359***
Intra-regional dummy	2.381***	1.893***	1.773***	-0.111
ln(employees)	3.039***	3.056***	2.068***	-1.678***
ln(establishments)	-1.745***	-1.765***	-1.454***	0.532***
Observations	1,131,228	1,131,228	1,131,228	
Log-Likelihood	-247,026.11	-246,447.57	-243,450.40	
Significant eigenvectors	13	13	12	
LR-Test Spatial Filter	17,094.42***	16,679.82***	12,123.3***	

Annual dummies and constant terms included but not reported in this table. Significance levels * 10%, ** 5%, *** 1%.

Table: Estimation results for the pooled regressions

Cross-Regional collaborations	Negbin		Zero-Inflated Negbin Collab.	Zero-Inflated Negbin Probit
Geo. dist. in km	-0.00175***	-0.00471***	-0.00312***	0.00264***
Geo. dist. in km (sq.)		4.80e-06***	4.10e-06***	-1.55e-06***
Cognitive dist.	-1.398***	-1.137***	0.114	2.018***
Cognitive dist. (sq.)		-0.298	-2.004***	-3.004***
Institut. dist.	-0.352***	-1.098***	-1.110***	0.436***
Institut. dist. (sq.)		0.499***	0.396***	-0.322***
Border dummy	-0.310***	-0.333***	-0.204***	0.165***
Neighbouring dummy	1.003***	0.688***	0.461***	-0.359***
Intra-regional dummy	2.381***	1.893***	1.773***	-0.111
ln(employees)	3.039***	3.056***	2.068***	-1.678***
ln(establishments)	-1.745***	-1.765***	-1.454***	0.532***
Observations	1,131,228	1,131,228	1,131,228	
Log-Likelihood	-247,026.11	-246,447.57	-243,450.40	
Significant eigenvectors	13	13	12	
LR-Test Spatial Filter	17,094.42***	16,679.82***	12,123.3***	

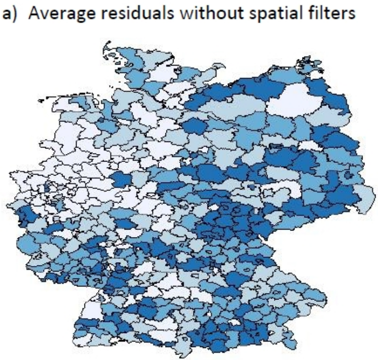
Annual dummies and constant terms included but not reported in this table. Significance levels * 10%, ** 5%, *** 1%.

Findings

Determinants of collaboration intensity

- U-shaped impact of geographical and institutional distance
- Negative impact shrinks after passing a threshold level (380 km or inst. dist. of 1.4)
- Positive impact may not be acheived or be neglected
- Negative impact of cognitive distance. Not as clear as for the other two proximity measures
- Other separation measures in line with expectations
- Regions with large enterprise tend to be more engaged in R&D collaborations
- Eigenvector Spatial Filters reduce the degree of autocorrelation

Figure: Average Error term per region - zero-inflated Negbin



Outlook and policy conclusion

Policy conclusion

- High share in economic strong regions and clusters
- R&D subsidies are absorbed by technological strong actors
- Amplification of economic concentration
- Reduction of regional disparities cannot be aim of R&D subsidies

Outlook and policy conclusion

Policy conclusion

- High share in economic strong regions and clusters
- R&D subsidies are absorbed by technological strong actors
- Amplification of economic concentration
- Reduction of regional disparities cannot be aim of R&D subsidies

Drawbacks and Outlook

- Focus on specific type of actors and collaborations
- Implementation of zero-inflated panel regressions
- Additional insights from collaborations across functional regions
- Identification of proxies for missing proximity measures

Appendix

Table 4. Estimation results for the pooled regressions – non-inflated and zero-inflated specifications

	Poisson		Negative binomial		Poisson		Negative binomial	
					Count	Probit	Count	Probit
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Distance (separation variables)								
Geographical distance (km)	-0.00135***	-0.00378***	-0.00175***	-0.00471***	-0.00237***	0.00207***	-0.00312***	0.00264***
	(8.82e-05)	(0.000253)	(6.66e-05)	(0.000224)	(0.000306)	(0.000263)	(0.000325)	(0.000370)
Geographical distance squared (km)		4.01e-06***		4.80e-06***	2.67e-06***	-1.79e-06***	4.10e-06***	-1.55e-06***
		(3.61e-07)		(3.16e-07)	(4.53e-07)	(3.91e-07)	(4.93e-07)	(5.43e-07)
Cognitive distance	-1.545***	-0.817***	-1.398***	-1.137***	-0.342	1.906***	0.114	2.018***
	(0.0833)	(0.273)	(0.0542)	(0.179)	(0.298)	(0.269)	(0.289)	(0.426)
Cognitive distance (squared)		-0.978***		-0.298	-1.198***	-2.101***	-2.004***	-3.004***
		(0.330)		(0.212)	(0.405)	(0.384)	(0.416)	(0.651)
Institutional distance	-0.492***	-1.156***	-0.352***	-1.098***	-0.774***	0.851***	-1.110***	0.436***
	(0.0278)	(0.0767)	(0.0202)	(0.0647)	(0.0810)	(0.0752)	(0.0882)	(0.109)
Institutional distance (squared)		0.464***		0.499***	0.314***	-0.420***	0.396***	-0.322***
		(0.0445)		(0.0380)	(0.0480)	(0.0458)	(0.0528)	(0.0660)
Border region (dummy)	-0.332***	-0.298***	-0.310***	-0.333***	-0.147***	0.0962***	-0.204***	0.165***
	(0.0322)	(0.0316)	(0.0211)	(0.0210)	(0.0429)	(0.0337)	(0.0341)	(0.0375)
Neighbouring region (dummy)	0.984***	0.755***	1.003***	0.688***	0.533***	-0.311***	0.461***	-0.359***
	(0.0695)	(0.0632)	(0.0621)	(0.0653)	(0.0988)	(0.0642)	(0.0799)	(0.0931)
Intra-regional collaboration (dummy)	1.360***	1.018***	2.381***	1.893***	1.408***	-0.408***	1.773***	-0.111
	(0.125)	(0.151)	(0.134)	(0.139)	(0.0987)	(0.102)	(0.123)	(0.142)
Origin and destination variables (mass terms)								
Number of employees (log)	3.190***	3.271***	3.039***	3.056***	1.720***	-1.797***	2.068***	-1.678***
	(0.0645)	(0.0726)	(0.0375)	(0.0376)	(0.0621)	(0.0541)	(0.0542)	(0.0637)
Number of establishments (log)	-1.937***	-2.032***	-1.745***	-1.765***	-1.060***	0.861***	-1.454***	0.532***
	(0.0719)	(0.0806)	(0.0369)	(0.0371)	(0.0576)	(0.0492)	(0.0523)	(0.0629)
Constant	4.312***	4.988***	2.204***	2.887***	3.809***	1.816***	7.274***	5.873***
	(0.693)	(0.230)	(0.268)	(0.270)	(0.459)	(0.346)	(0.353)	(0.454)
Constant lnalpha			1.432***	1.417***			0.586***	
			(0.0192)	(0.019)			(0.0156)	
Observations	1,131,228	1,131,228	1,131,228	1,131,228	1,131,228		1,131,228	
Log-Likelihood	-315,921.83	-316,420.69	-247,026.11	-246,447.57	-275,442.1		-243,450.40	
LR-Test Spatial Filter	41,200.84***	77,515.4***	17,094.42***	16,679.82***	30,410.0***		12,123.3***	
X²-Test for joint significance of squared distance measures		236.1***		475.56***	77.41***		170.39***	
Moran's I with spatial filters	.0194	.0159	-.00007	-.00008	.0279		.0197	
Moran's I without spatial filters	.0243	.0245	-.00040	-.00033	.0330		.0324	
Number of significant eigenvectors	19	16	13	13	8		12	
Test for overdispersion	0.5233***	0.5129***	0.9308***	0.9272***	1.1512***		0.6905***	
Vuong-Statistcs for model without clustered error terms					49.47***		32.74***	

Notes: Standard errors in parentheses. - The dependent variable is the number of cross-regional collaborations between region i and j. – Annual dummies and constant terms included but not reported in this table. – Significance levels * 10%, ** 5%, *** 1%. – The coefficients of the mass terms for the origin and destination region are identical.

Source: Authors' own calculation.

Table 5. Estimation results for the panel regressions (random effects with Mundlak group variables)

VARIABLES	Poisson (9)	Negative binomial (10)
Distance (separation variables)		
Geographical distance (km)	-0.00520***	-0.00344***
	(0.000177)	(0.000150)
Geographical distance squared (km)	5.20e-06***	3.10e-06***
	(2.63e-07)	(2.25e-07)
Cognitive distance	3.521***	2.485***
	(0.288)	(0.382)
Cognitive distance (squared)	-4.561***	-3.138***
	(0.337)	(0.443)
Institutional distance	-0.0593	0.0374
	(0.0448)	(0.0597)
Institutional distance (squared)	0.0922***	0.0287
	(0.0236)	(0.0315)
Border region (dummy)	-0.248***	-0.235***
	(0.0173)	(0.0148)
Neighbouring region (dummy)	0.730***	0.285***
	(0.0502)	(0.0374)
Intra-regional collaboration (dummy)	1.818***	-0.378***
	(0.0960)	(0.0611)
Origin and destination variables (mass terms)		
Number of employees (log)	-0.901***	0.0527
	(0.192)	(0.264)
Number of establishments (log)	1.440***	1.408***
	(0.181)	(0.248)
Mundlak: Group Variables		
Cognitive distance	-5.134***	-3.791***
	(0.323)	(0.401)
Cognitive distance (squared)	5.103***	3.548***
	(0.382)	(0.465)
Institutional distance	-1.115***	-1.103***
	(0.0838)	(0.0843)
Institutional distance (squared)	0.492***	0.501***
	(0.0494)	(0.0485)
Number of employees (log)	3.736***	2.539***
	(0.195)	(0.265)
Number of establishments (log)	-2.933***	-2.858***
	(0.184)	(0.249)
Constant	-0.118	1.784***
	(0.237)	(0.193)
Constant lnalpha	0.937***	1.102***
	(0.0110)	(0.0120)
Observations	1,131,228	1,131,228
Log-Likelihood	-251,970.82	-230,528.56
LR-Test Spatial Filter	7,276.96***	7,234.96***
Number of significant eigenvectors	20	13
Hausman Test between fixed-effects and random-effects	89.8***	5939.42***

Notes: Z-values in parentheses. The dependent variable is the number of cross-regional collaborations between region i and j. - Annual dummies and constant term included but not reported in this table. - Significance levels * 10%, ** 5%, *** 1%. – The coefficients of the mass terms for the origin and destination region are identical.

Source: Authors' own calculation.