Borders within Europe

Marta Santamaria[5] Jaume Ventura[1,2,3] Uğur Yeşilbayraktar[4]

[1]Barcelona School of Economics

[2]CREi

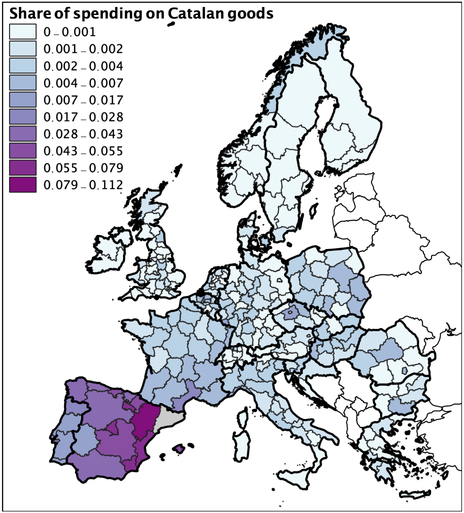
[3]Universitat Pompeu Fabra

[4]University of Duisburg-Essen

[5]University of Warwick

February, 2024

Effect of national borders on regional trade flows



- National bias: intranational trade is 30 times larger than international trade

This paper

1. We construct a matrix of region-to-region trade flows
   - Micro-level survey with 3 million truck shipments per year
   - 49% intra-EU trade in tonne-km terms.
   - 269 regions, 12 industries, 7 years
2. We use the causal inference framework (Rubin, 1990):
   - Identify confounding factors that affect border assignment and trade
   - Create balanced samples (similar probability of a border)
   - Perform inference on these balanced samples
3. We obtain the following results:
   - Eliminating a border would increase trade by 5.71 times
   - Mechanisms: Language differences, transport infrastructure fragmentation, disagreement in Values

A new regional trade dataset

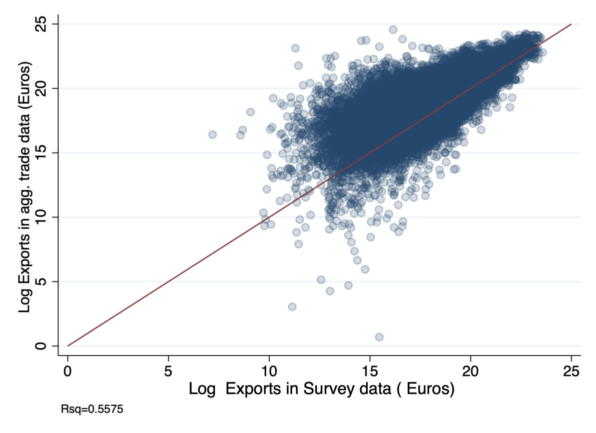
A new regional trade dataset

- Survey on shipments by road (3 million obs per year) collected by Eurostat
  - Origin/Dest.(NUTS2 region), Weight (kg), Distance (km) and Industry (by Vehicle/Journey/Good Type)
- 20 industries (agriculture, mining and all manufacturing) of which we keep:
  - Industries associated with trade and with high share of traffic by road. Mode of transp.
- 29 countries of which we drop:
  - 5 one-region countries: Cyprus, Estonia, Latvia, Lithuania and Luxembourg
- 7 years: 2011-2017
- Impute region-to-region prices in order to construct a value matrix.

Data check: International trade flows vs Aggregated regional dataset

- For each industry i and year t, we construct the matrix of bilateral trade:

$$V^{it} = \left[V^{it}_{nm} \right]_{269x269} \qquad \text{where } V^{it}_{nm} = P^{it}_{nm} \cdot W^{it}_{nm}$$



Industries

Identifying the border effect

Borders as a treatment: causal inference framework (Rubin, 1990)

- Two potential trade outcomes by border status:

$$S_{nm} = \begin{cases} S_{nm}(1) & \text{if } B_{nm} = 1 \text{ (Border, i.e. Treatment)} \\ S_{nm}(0) & \text{if } B_{nm} = 0 \text{ (No Border, i.e. Control)} \end{cases} \tag{1}$$

- Border effect (β) is the percentage change in market shares caused by the border:

$$\beta = \ln \frac{S_{nm}(1)}{S_{nm}(0)} \tag{2}$$

- **Challenge:** we only observe either $S_{nm}(0)$ or $S_{nm}(1)$ for each region pair.

- **Causal inference framework**: We can estimate the border effect by comparing region-pairs with different border assignment (B=0 or B=1) but with **identical pre-treatment covariates** if the treatment is unconfounded (as-good-as-random).

Steps to estimate border effect from observational data

1. We select relevant **pre-treatment covariates** to condition on (Geographic factors)
   Covariates
2. We estimate the **probability of a border** between n and m using our covariates X_{nm} and trim sample
   P.score Histogram
3. We group region-pairs such that probability of border is **balanced between treated and control**
   Blocks - sum stats Balance test Map Block 4 All Blocks
4. We check for (lack of) selection bias: participation rates are balanced
   Participation rates

Inference in balanced groups of treated and control region-pairs

Table: Summary statistics of covariates by block

Blocks	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd
Probability of Border	**0.20**	**0.31**	**0.44**	**0.57**	**0.66**	**0.72**	**0.78**	**0.84**	**0.89**
	0.04	0.04	0.04	0.04	0.02	0.02	0.02	0.02	0.01
Distance	154.36	186.07	240.35	298.82	349.83	383.02	440.94	480.01	446.70
	61.03	74.23	93.43	121.79	143.55	143.03	161.45	136.84	61.64
Insularity	0.01	0.01	0.01	0.02	0.04	0.07	0.08	0.12	0.22
	0.08	0.12	0.12	0.15	0.20	0.25	0.28	0.33	0.42
Mountain Ranges	208.38	291.05	351.19	466.84	533.75	549.99	596.98	735.32	1244.59
	232.38	320.38	376.25	457.99	528.13	545.14	561.71	681.78	888.16
River Basin	0.29	0.28	0.21	0.19	0.17	0.14	0.12	0.10	0.06
	0.45	0.45	0.41	0.39	0.37	0.35	0.32	0.31	0.24
Remoteness	1169.05	1097.32	1092.09	1087.40	1081.35	1051.59	1038.82	1002.73	938.72
	307.02	268.01	273.50	276.93	275.84	249.16	229.19	187.51	140.79
N	323	408	515	698	507	660	1062	1582	354

Back Strategy on a map

Results: Average border effect

Border Effect by block

$$\ln S_{nm} = \alpha + \beta \cdot B_{nm} + \gamma \cdot N_{nm} + \lambda' \cdot X_{nm} + u_{nm} \quad (3)$$

Table: Average border effect

Dep. Var: $ln(S_{n,m})$	Block 1 (1)	Block 2 (2)	Block 3 (3)	Block 4 (4)	Block 5 (5)	Block 6 (6)	Block 7 (7)	Block 8 (8)	Block 9 (9)
Border	-1.786***	-1.721***	-1.699***	-1.768***	-1.686***	-1.796***	-1.687***	-1.754***	-1.858***
	(0.182)	(0.178)	(0.175)	(0.175)	(0.238)	(0.289)	(0.268)	(0.290)	(0.201)
Number of Borders	7.058***	6.695***	7.041***	10.779***	11.294***	11.833***	9.234***	8.091***	0.420
	(1.756)	(1.970)	(2.034)	(1.730)	(2.064)	(2.783)	(2.792)	(3.063)	(2.944)
Geographic covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	645	813	1024	1364	968	1267	2011	2948	637
R^2	.572	.533	.501	.47	.375	.388	.31	.285	.299

Full Table | Industries | Full sample vs Trimmed

- The average border effect is -1.744 (-1.747) weighting by size of block (treated).
- The border reduces the market share to 17.5% (5.71) of its potential
- ATE without controlling for number of borders: ATE Table

Mechanisms

Do borders reduce the probability of speaking the same language?

$$\text{Common Language Share}_{nm} = \alpha + \beta \cdot B_{nm} + \gamma \cdot N_{nm} + \lambda' \cdot X_{nm} + u_{nm} \tag{4}$$

Do borders reduce the probability of speaking the same language?

$$\text{Common Language Share}_{nm} = \alpha + \beta \cdot B_{nm} + \gamma \cdot N_{nm} + \lambda' \cdot X_{nm} + u_{nm} \tag{4}$$

D. V.: Common Language sh_{nm}	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Block 9
Border Effect	-0.671***	-0.561***	-0.664***	-0.658***	-0.751***	-0.754***	-0.733***	-0.714***	-0.655***
	(0.0883)	(0.0964)	(0.0744)	(0.0699)	(0.0647)	(0.0599)	(0.0708)	(0.0881)	(0.150)
Distance (Geo)	0.0350	-0.123	-0.139	-0.416***	-0.000426	-0.480*	-0.261*	-0.119	0.368
	(0.0940)	(0.147)	(0.120)	(0.154)	(0.328)	(0.250)	(0.139)	(0.0808)	(0.645)
No. Borders	-0.146	-1.289	-0.439	-0.800	-0.747	-1.131*	-1.518**	-1.294	-0.843
	(0.607)	(0.823)	(0.937)	(0.931)	(0.739)	(0.576)	(0.632)	(0.815)	(1.906)
Geo covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	641	803	1010	1340	932	1190	1858	2606	505
R^2	.662	.553	.643	.659	.741	.78	.724	.619	.449

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

- A border reduces the share of common languages by between 56% to 76%.

Do borders increase shipping distances?

$$\log(\text{Shipping distance})_{nm} = \alpha + \beta \cdot B_{nm} + \gamma \cdot N_{nm} + \lambda' \cdot X_{nm} + u_{nm} \tag{5}$$

Do borders increase shipping distances?

$$\log(\text{Shipping distance})_{nm} = \alpha + \beta \cdot B_{nm} + \gamma \cdot N_{nm} + \lambda' \cdot X_{nm} + u_{nm} \quad (5)$$

D. V.: log(Shipping dist)$_{nm}$	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Block 9
Border Effect	0.158**	0.0988***	0.101***	0.0717***	0.0612*	0.0168	0.0478	0.0398	0.0545**
	(0.0604)	(0.0292)	(0.0329)	(0.0206)	(0.0347)	(0.0429)	(0.0296)	(0.0344)	(0.0265)
Distance (Geo)	0.710***	0.952***	1.015***	1.039***	0.720***	0.987***	0.890***	0.865***	0.521**
	(0.122)	(0.104)	(0.110)	(0.0686)	(0.176)	(0.131)	(0.0917)	(0.0657)	(0.244)
No. Borders	0.536	0.563**	0.309	0.212	0.323	0.638	0.512	0.506	-0.401
	(0.377)	(0.251)	(0.310)	(0.223)	(0.398)	(0.548)	(0.470)	(0.484)	(0.561)
Geo covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	641	803	1010	1340	932	1190	1858	2606	505
R^2	.733	.822	.803	.838	.76	.711	.724	.584	.372

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

- A border increases shipping distances by around 10% for region-pairs within 400km.

Do borders increase disagreement in values?

$$\text{Disagreement}_{nm} = \alpha + \beta \cdot B_{nm} + \gamma \cdot N_{nm} + \lambda' \cdot X_{nm} + u_{nm} \tag{6}$$

Do borders increase disagreement in values?

$$\text{Disagreement}_{nm} = \alpha + \beta \cdot B_{nm} + \gamma \cdot N_{nm} + \lambda' \cdot X_{nm} + u_{nm} \quad (6)$$

D. V.: Disagreement index$_{nm}$	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Block 9
Border Effect	0.489	0.671**	0.637**	0.469	1.102***	0.870***	0.796***	0.348**	0.188
	(0.364)	(0.283)	(0.292)	(0.429)	(0.303)	(0.279)	(0.251)	(0.172)	(0.342)
Distance (Geo)	-0.191	-0.821**	-0.187	0.049	-1.902	0.478	0.177	-0.359	2.159
	(0.341)	(0.346)	(0.764)	(0.568)	(1.467)	(0.677)	(0.570)	(0.613)	(2.321)
Number of Borders	-3.614	-4.662***	-4.822**	-4.850	-12.546***	-10.866***	-11.371***	-3.865	-4.152
	(2.204)	(1.704)	(2.044)	(3.104)	(2.754)	(3.268)	(3.448)	(2.975)	(9.155)
Geo covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	442	476	546	646	452	522	750	1010	216
R^2	.333	.315	.344	.192	.286	.215	.143	.0913	.135

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

- A border increases disagreement from 15% up to 50% of the disagreement mean.

Conclusions

1. We construct a new regional trade dataset
2. We propose a new identification strategy for the border effect
3. We find that **borders still reduce trade** and operate **through languages, transport infrastructure and values** (among other channels)
4. Next Steps:
   - More **Mechanisms**: Procurement, migration, financial integration, FDI, regulation...
   - More **theory**: Quantitative models with realistic governments

Probability of having a border: Propensity score

- 1) Estimate probability of a border between n and m (Propensity score)

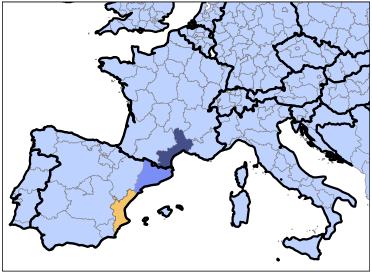
$$\mathbf{1}_{B_{nm}=0} = \alpha + \mu' X_{nm} + v_{n,m} \tag{7}$$

Dependent Variable: Border	Full sample (1)	Trimmed sample (2)
Distance	2.998***	1.893***
	(0.056)	(0.078)
Insularity	1.096***	1.059***
	(0.096)	(0.128)
Mountain Ranges	0.179***	0.283***
	(0.030)	(0.031)
River Basin	0.767***	0.420***
	(0.089)	(0.089)
Remoteness	-3.857***	-3.341***
	(0.155)	(0.168)
Constant	9.129***	11.180***
	(0.992)	(1.029)
N	36046	6110
Pseudo R^2	0.476	0.143

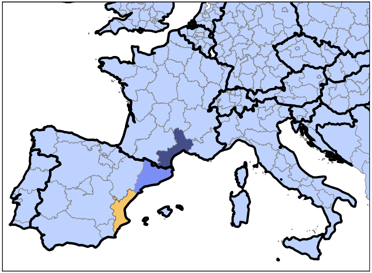
- Geographical covariates explain 47.6% of variation in borders (column 1)

Back

1. Identifying the relevant pre-treatment covariates

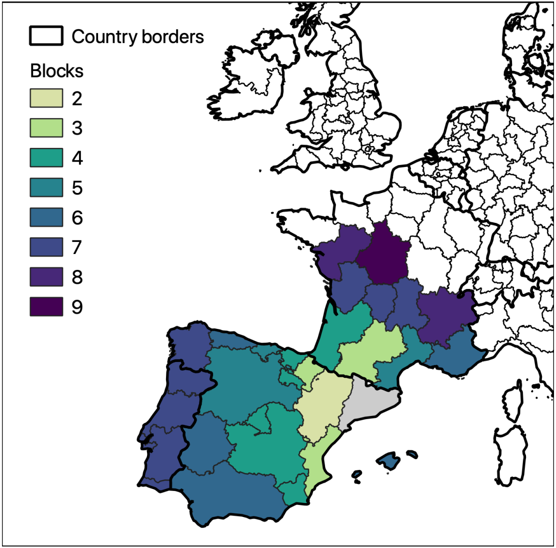


1. Identifying the relevant pre-treatment covariates



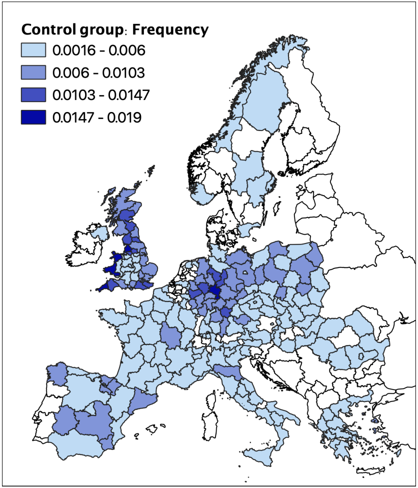
- What is the set of pre-treatment covariates, X_{nm}, that we have to condition on?
  1. Covariates that pre-date border assignment
  2. Covariates that affect borders as well as trade
- We collect bilateral geographical covariates: distance, insularity, elevation, river basin, remoteness Details Control/Treatment Back

Our identification strategy on a map

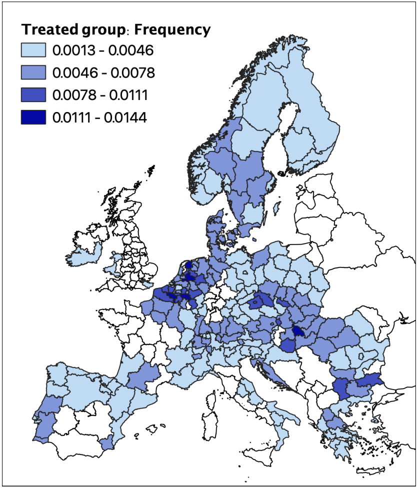


Back

Composition of regions in block 4



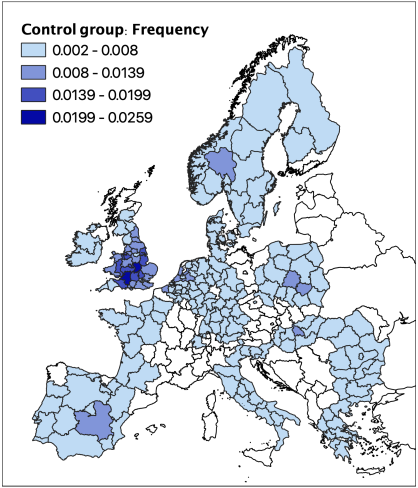
A) Control group



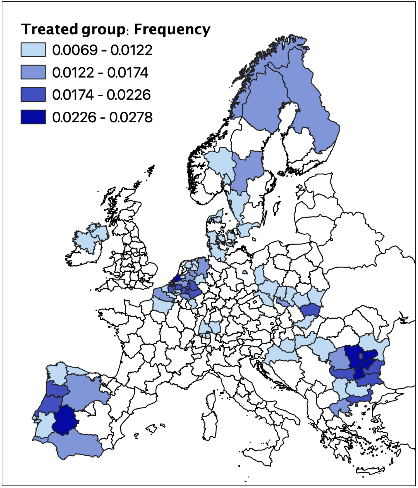
B) Treated group

Back

Composition of regions in block 1



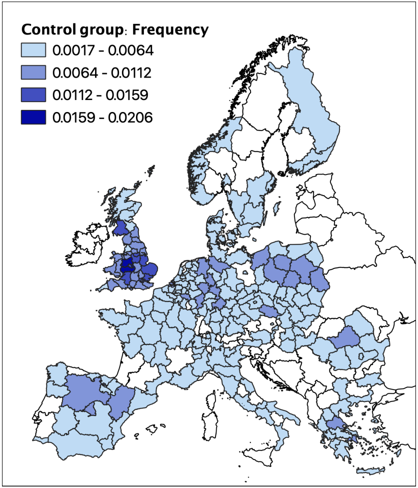
A) Control group



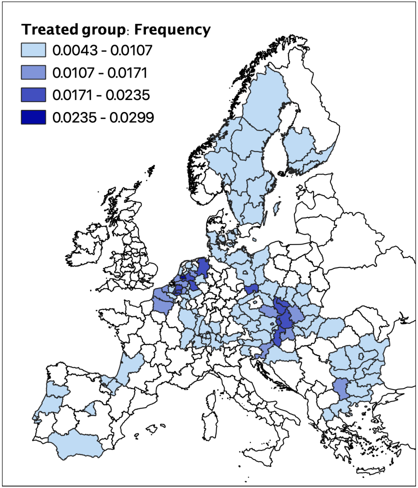
B) Treated group

Back

Composition of regions in block 2



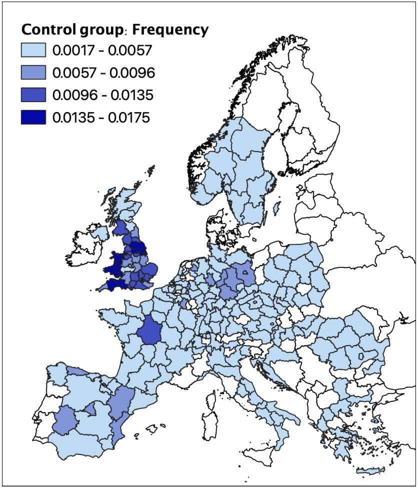
A) Control group



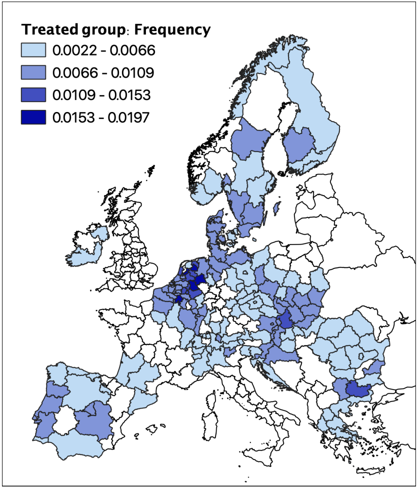
B) Treated group

Back

Composition of regions in block 3



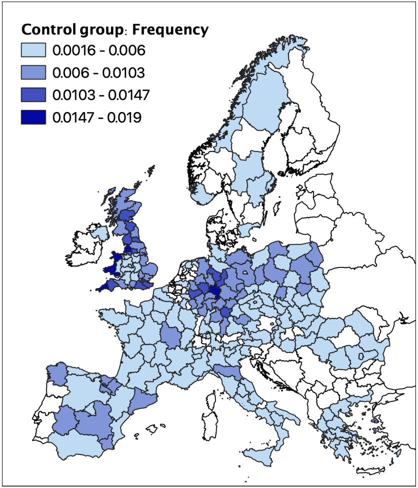
A) Control group



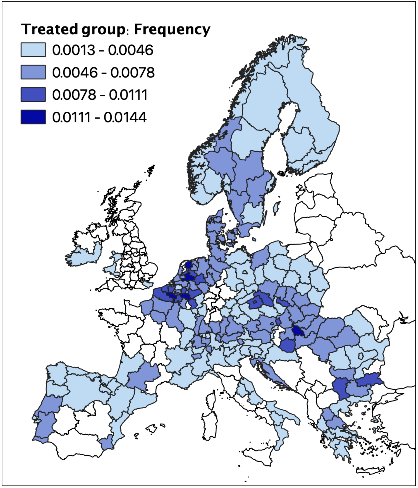
B) Treated group

Back

Composition of regions in block 4



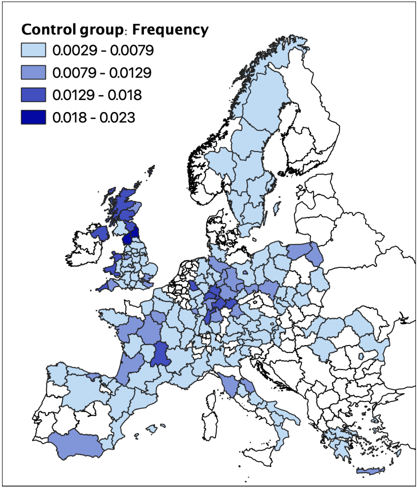
A) Control group



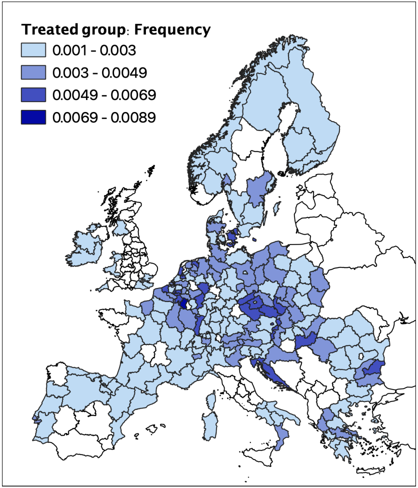
B) Treated group

Back

Composition of regions in block 5



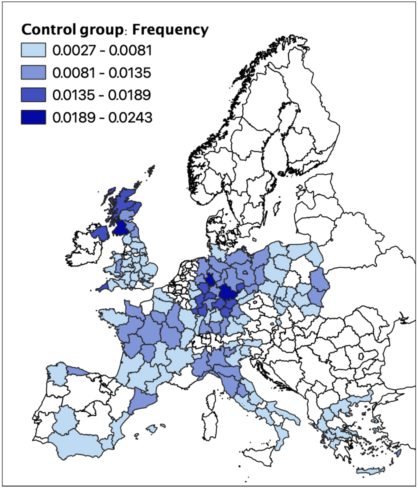
A) Control group



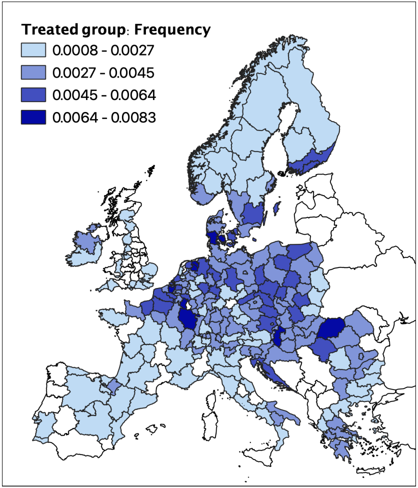
B) Treated group

Back

Composition of regions in block 6



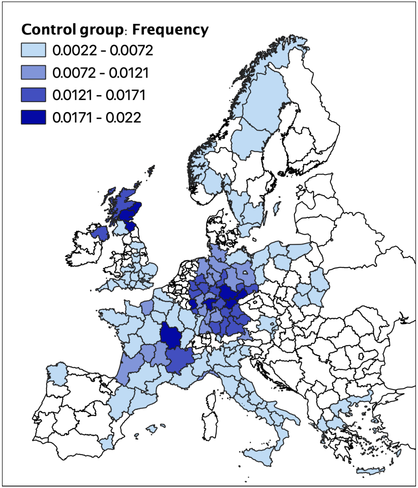
A) Control group



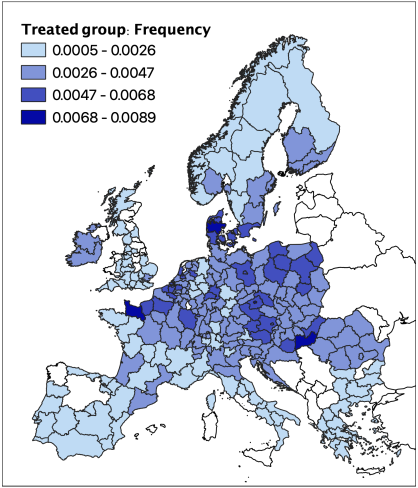
B) Treated group

Back

Composition of regions in block 7



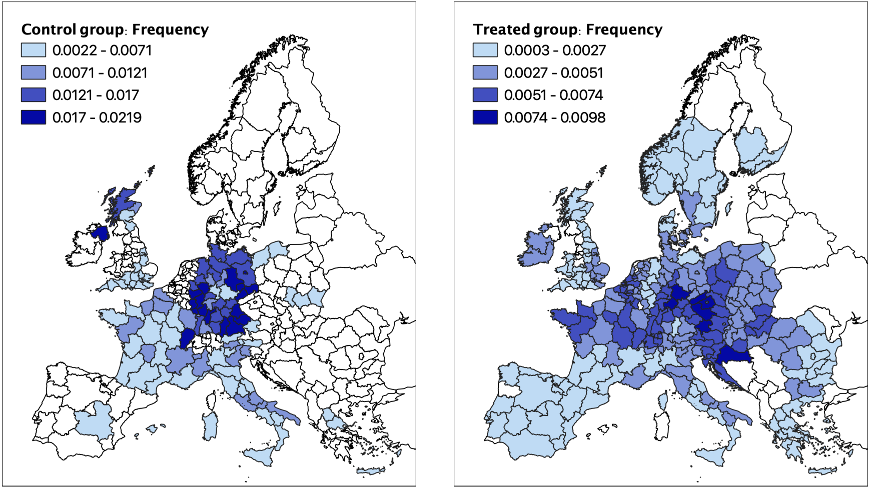
A) Control group



B) Treated group

Back

Composition of regions in block 8

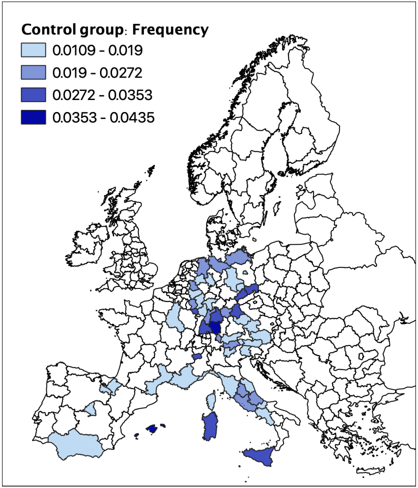


A) Control group

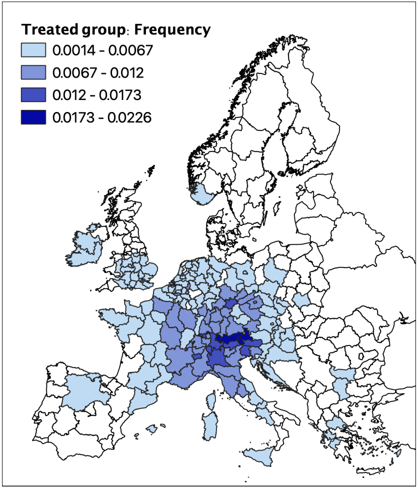
B) Treated group

Back

Composition of regions in block 9



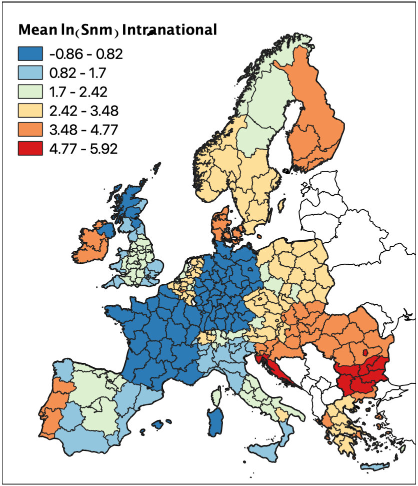
A) Control group

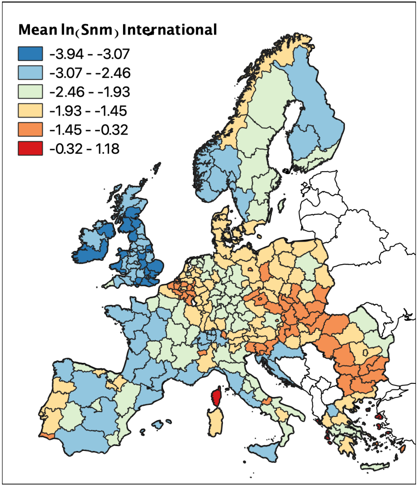


B) Treated group

Back

Selection bias due to number of borders





Back

Causal inference: Selection bias

- Assume treatment assignment is probabilistic, individualistic and uncounfounded.
- Would $\hat{\beta}$ an unbiased estimator?

$$\hat{\beta} - \beta = \underbrace{\begin{array}{c} E\left(\ln S_{nm}\left(0\right) \mid \mathcal{S}_{nm}\left(1\right) > 0, B_{nm} = 1, X_{nm} = x\right) \\ -E\left(\ln S_{nm}\left(0\right) \mid \mathcal{S}_{nm}\left(1\right) > 0, B_{nm} = 0, X_{nm} = x\right) \end{array}}_{\text{Selection bias due to the number of borders}} \tag{8}$$

$$+ \underbrace{\begin{array}{c} E\left(\ln S_{nm}\left(0\right) \mid \mathcal{S}_{nm}\left(1\right) > 0, B_{nm} = 0, X_{nm} = x\right) \\ -E\left(\ln S_{nm}\left(0\right) \mid \mathcal{S}_{nm}\left(0\right) > 0, B_{nm} = 0, X_{nm} = x\right) \end{array}}_{\text{Selection bias due to changes in participation}} \tag{9}$$

- Selection bias due to number of borders Maps
  - Condition on the number of borders of the region pair
- Selection bias due to participation: 3.1% intranational vs 38.2% international
  - Compare participation rates among treated and controls

Back

Covariates across Control and Treated pairs

Table: Balancing test of covariates by block

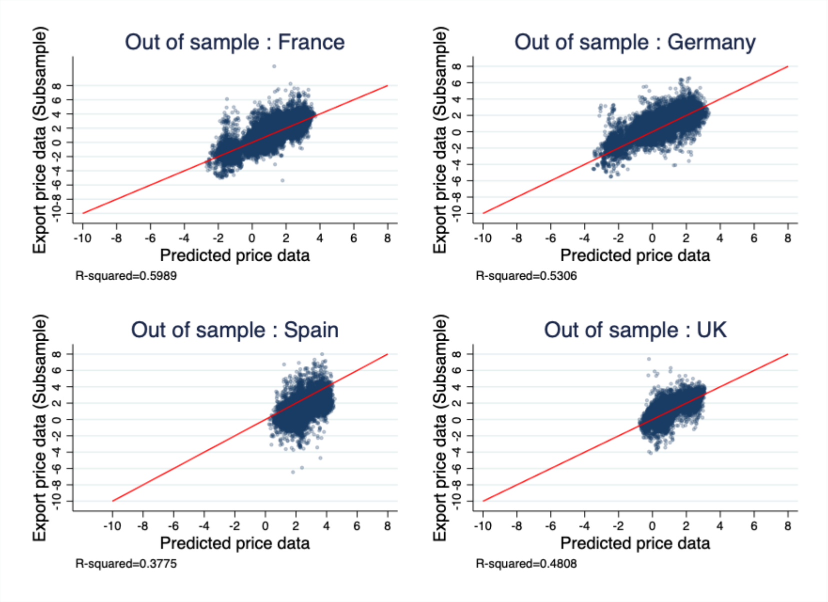
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Distance	-22.24***	8.207	5.049	4.693	17.13	-11.79	-24.16**	-33.09***	28.87***
	(8.077)	(8.126)	(8.290)	(9.269)	(13.42)	(12.40)	(12.07)	(9.763)	(9.636)
Insularity	-0.00990	0.0206	0.0166	0.0187*	0.0302	0.0125	-0.00573	-0.0613***	-0.00663
	(0.0105)	(0.0132)	(0.0103)	(0.0110)	(0.0190)	(0.0216)	(0.0208)	(0.0234)	(0.0660)
Mountain Ranges	-31.46	25.23	-16.62	-120.6***	-148.1***	-114.3**	-96.62**	-139.1***	45.43
	(31.06)	(35.09)	(33.39)	(34.56)	(49.01)	(47.07)	(41.96)	(48.70)	(140.6)
River Basin	0.0528	-0.0328	-0.00768	0.0366	0.0247	-0.0101	-0.00522	-0.0304	0.0285
	(0.0608)	(0.0495)	(0.0362)	(0.0296)	(0.0350)	(0.0303)	(0.0242)	(0.0219)	(0.0382)
Remoteness	-109.7***	51.41*	30.83	20.83	44.75*	-5.539	-21.53	-54.87***	59.66***
	(40.65)	(29.26)	(24.24)	(21.06)	(25.75)	(21.61)	(17.15)	(13.36)	(22.06)
N	323	408	515	698	507	660	1062	1582	354

Back

How prevalent is road trade across industries?

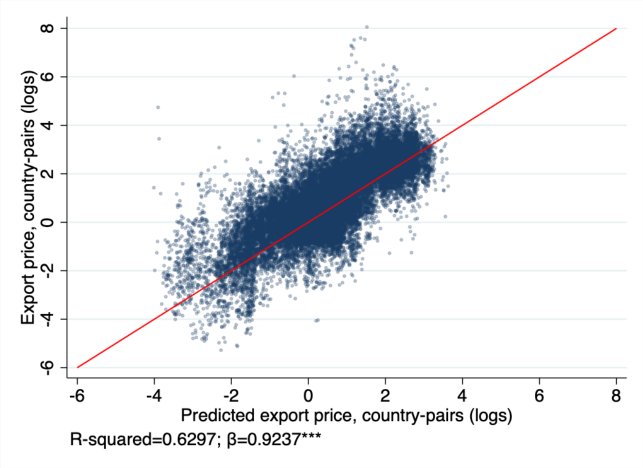
Back

Price Imputation (1): Out-of-sample Estimates



Back

Price Imputation (2): Country-to-Country Export price Estimates



Back

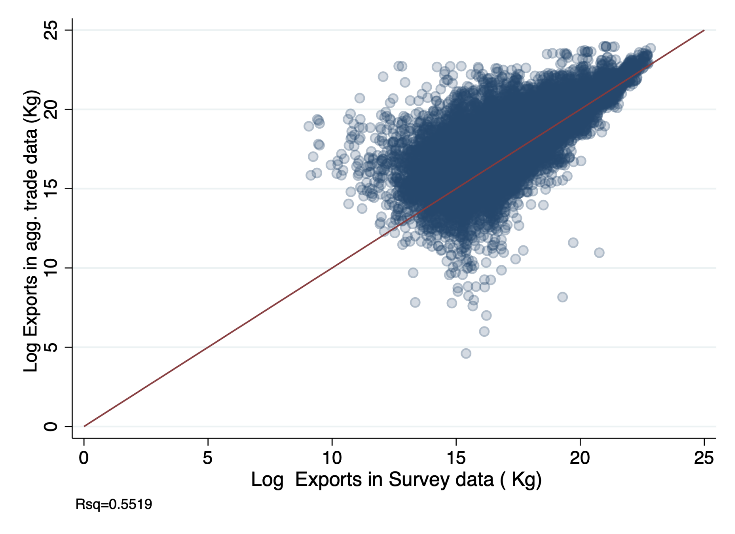
Constructing Blocks

Table: Summary statistics of covariates by block

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd
Distance	154.36	186.07	240.35	298.82	349.83	383.02	440.94	480.01	446.70
	61.03	74.23	93.43	121.79	143.55	143.03	161.45	136.84	61.64
Insularity	0.01	0.01	0.01	0.02	0.04	0.07	0.08	0.12	0.22
	0.08	0.12	0.12	0.15	0.20	0.25	0.28	0.33	0.42
Mountain Ranges	208.38	291.05	351.19	466.84	533.75	549.99	596.98	735.32	1244.59
	232.38	320.38	376.25	457.99	528.13	545.14	561.71	681.78	888.16
River Basin	0.29	0.28	0.21	0.19	0.17	0.14	0.12	0.10	0.06
	0.45	0.45	0.41	0.39	0.37	0.35	0.32	0.31	0.24
Remoteness	1169.05	1097.32	1092.09	1087.40	1081.35	1051.59	1038.82	1002.73	938.72
	307.02	268.01	273.50	276.93	275.84	249.16	229.19	187.51	140.79
Propensity score	**0.20**	**0.31**	**0.44**	**0.57**	**0.66**	**0.72**	**0.78**	**0.84**	**0.89**
	0.04	0.04	0.04	0.04	0.02	0.02	0.02	0.02	0.01
N	323	408	515	698	507	660	1062	1582	354

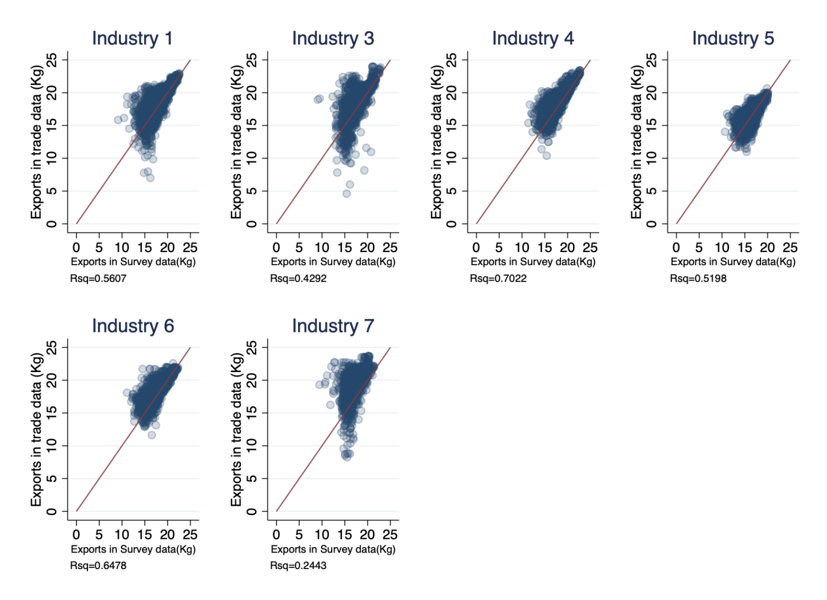
- We create 9 groups such that treatment and control have same border probability
- Blocks are ordered according to probability of the border
- Region-pairs are allocated to control and treated groups Map Block 4
- Geographical covariates are almost perfectly balanced across blocks Balance test Back

Constructed trade data vs Aggregate data: Kilos



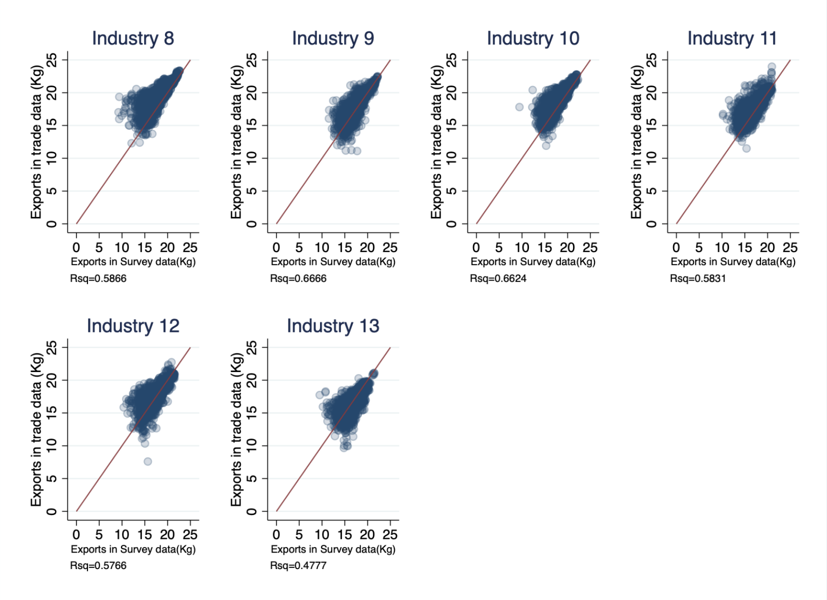
Back

Correlation with international trade data: Weight



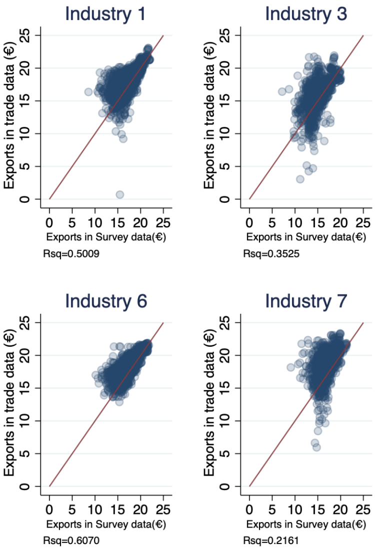
Back

Correlation with international trade data: Weight (II)



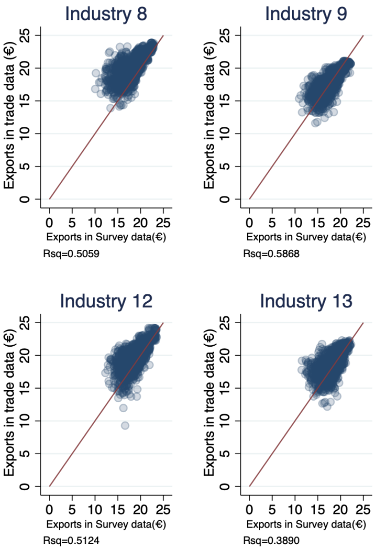
Back

Correlation with international trade data: Value



Back

Correlation with international trade data: Value (II)



Back

Unconfounded treatment: Geographical covariates

We collect the following geographical covariates

1. *Distance*. Length of the curve linking the central point of the origin region (centroid) and the central point of the destination region, in kilometers. We use a curve since we take into account the curvature of earth's surface.
2. *Insularity*. Dummy variable taking value one if there is the need to cross a sea to reach from one region to the other, and zero otherwise.
3. *Mountain ranges*. Largest altitude difference between two regions, computed as the difference between the highest altitude point and the lowest altitude point along the straight line that joins the centre the origin region (centroid) and the centre of the destination region.
4. *River basin*. Dummy variable taking value 1 if both regions belong to the same river basin. We consider the largest rivers in Europe. A map of the areas covered by each river basin can be found in the Appendix.
5. *Remoteness*. We calculate the remoteness of a region as the sum of the bilateral distance from that region to every other region in the sample. Then, we calculate the remoteness of a pair as the average remoteness of both regions.

Back

Participation rates are balanced

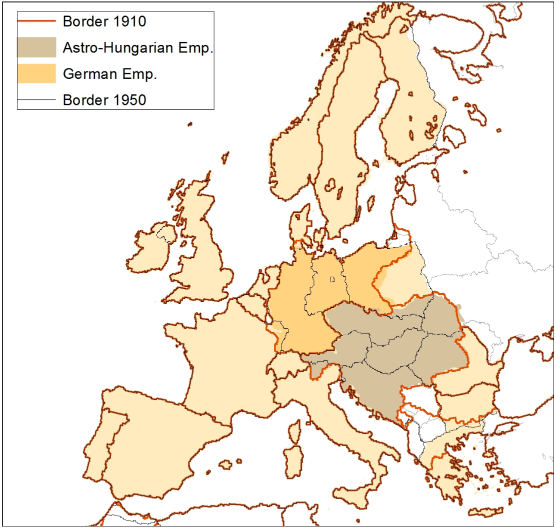
- Under probabilistic, individualistic and unconfounded assigment, we can use observational data to estimate causal effects
- We must check that participation rate is similar between treated and control pairs

Table: Participation rate: Control vs. Treated

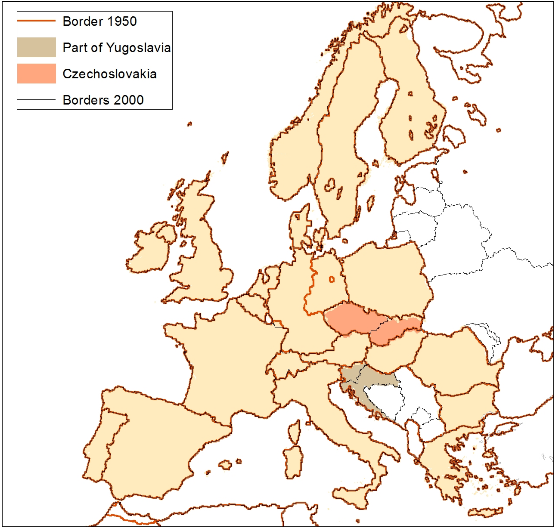
	All	Trimmed	Blocks								
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Part. rate control	0.968	0.976	1	.997	.993	.987	.968	.968	.936	.952	.915
Part. rate treated	0.617	0.946	.993	.996	.996	.969	.947	.957	.95	.928	.894
N	72092	12220	646	816	1030	1396	1014	1320	2124	3164	710

Back

Borders changes after 1910



Border changes after 1910



Back

Unconfounded treatment: Geographical covariates

Table: Geographical Covariates across treatment groups

	Treatment group mean	Control group mean	Difference (t-stat)
Distance	1213.62	315.64	-898.0
			(-71.79)
Insularity	0.32	0.06	-0.258
			(-27.23)
Mountain Ranges	1473.66	496.08	-977.6
			(-37.95)
River Basin	0.04	0.19	0.153
			(35.81)
Remoteness	1157.47	1075.85	-81.62
			(-17.19)
N	33567	2479	36046

- Geographical covariates are very different between treatment and control units.

Back

Propensity Score estimation: Logit model

$$\mathbf{1}_{B_{nm}=0} = \alpha + \mu' X_{nm} + v_{n,m} \tag{10}$$

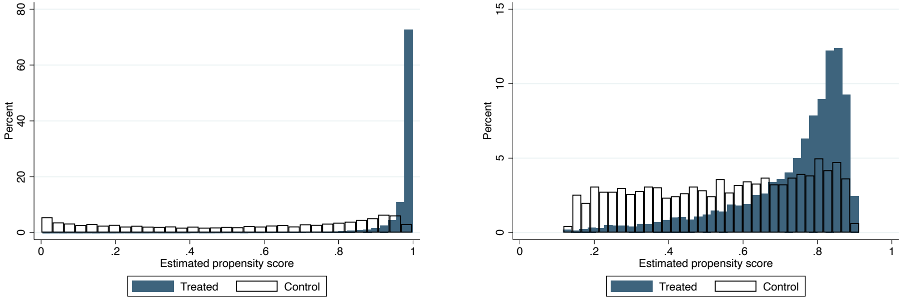
Dependent Variable: Border	Full sample (1)	Trimmed sample (2)
Distance	2.998	1.893
	(0.056)	(0.078)
Insularity	1.096	1.059
	(0.096)	(0.128)
Mountain Ranges	0.179	0.283
	(0.030)	(0.031)
River Basin	0.767	0.420
	(0.089)	(0.089)
Remoteness	-3.857	-3.341
	(0.155)	(0.168)
Constant	9.129	11.180
	(0.992)	(1.029)
N	36046	6110
Pseudo R^2	0.476	0.143

- Geographical covariates explain 47.6% of variation in borders (column 1)

Back

Overlap of Propensity Score: Full vs Trimmed sample

Figure: Histogram of propensity score



A) All region pairs **B) Trimmed sample**

- Propensity score overlap improves in trimmed sample

Back

Border Effect - With covariates

Table: Average border effect

Dep. Var: $ln(S_{n,m})$	Block 1 (1)	Block 2 (2)	Block 3 (3)	Block 4 (4)	Block 5 (5)	Block 6 (6)	Block 7 (7)	Block 8 (8)	Block 9 (9)
Border	-1.786***	-1.721***	-1.699***	-1.768***	-1.686***	-1.796***	-1.687***	-1.754***	-1.858***
	(0.182)	(0.178)	(0.175)	(0.175)	(0.238)	(0.289)	(0.268)	(0.290)	(0.201)
Distance	-0.899**	-1.378***	-1.643***	-0.618*	-1.949**	-0.532	-1.105**	-1.066***	-1.118
	(0.440)	(0.276)	(0.377)	(0.315)	(0.828)	(0.696)	(0.497)	(0.372)	(1.873)
Insularity	1.120	-0.861**	-0.157	-0.491	-1.777***	-0.913**	-1.596***	-1.554***	-1.024
	(0.754)	(0.376)	(0.430)	(0.412)	(0.534)	(0.418)	(0.351)	(0.319)	(0.862)
Mountain Ranges	0.014	-0.137*	-0.180**	-0.134	-0.322*	-0.088	-0.229**	-0.257***	-0.095
	(0.074)	(0.071)	(0.080)	(0.082)	(0.175)	(0.102)	(0.089)	(0.097)	(0.243)
River Basin	0.220	0.141	0.132	0.477***	0.155	0.514***	0.413**	0.348**	0.594
	(0.182)	(0.123)	(0.168)	(0.166)	(0.203)	(0.181)	(0.192)	(0.174)	(0.458)
Remoteness	2.236***	3.236***	3.339***	1.335**	3.412**	0.803	2.086**	2.167**	1.356
	(0.625)	(0.783)	(0.595)	(0.606)	(1.557)	(1.219)	(0.889)	(0.833)	(2.833)
Number of Borders	7.058***	6.695***	7.041***	10.779***	11.294***	11.833***	9.234***	8.091***	0.420
	(1.756)	(1.970)	(2.034)	(1.730)	(2.064)	(2.783)	(2.792)	(3.063)	(2.944)
Constant	-52.432***	-53.962***	-55.214***	-70.492***	-79.496***	-74.367***	-63.052***	-56.456***	-4.131
	(11.534)	(12.696)	(12.979)	(10.102)	(12.606)	(15.500)	(15.239)	(16.468)	(20.347)
N	645	813	1024	1364	968	1267	2011	2948	637
R^2	.572	.533	.501	.47	.375	.388	.31	.285	.299

Back

Border Effect: Industry-level

- We repeat our estimation industry by industry

Table: Border effect across industries and blocks

INDUSTRY	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Block 9	ATE: W	ATE: T
1. AGRI	-1.851***	-1.813***	-1.659***	-1.384***	-1.241***	-1.611***	-1.413***	-1.620***	-1.995***	-1.578	-1.559
2. MINE	-1.714***	-2.017***	-1.607***	-1.592***	-1.413***	-1.374***	-1.160***	-1.019***	-2.054***	-1.471	-1.395
3. FBT	-2.488***	-2.464***	-2.163***	-2.084***	-2.034***	-2.024***	-1.977***	-1.954***	-2.196***	-2.095	-2.047
4. TEX	-1.333***	-1.195***	-0.714***	-1.053***	-0.830***	-0.915***	-0.714***	-0.839***	-1.307***	-0.945	-0.904
5. WOOD	-1.532***	-1.641***	-1.366***	-1.429***	-1.369***	-1.360***	-1.488***	-1.588***	-1.828***	-1.499	-1.505
6. COKE/PET	-2.025***	-1.314***	-1.221***	-0.787***	-0.702***	-0.776***	-0.507***	-0.601***	-1.592***	-0.995	-0.866
7. CHEM	-1.373***	-1.278***	-1.206***	-1.388***	-1.080***	-1.267***	-1.298***	-1.308***	-1.249***	-1.282	-1.280
8. NON-MET	-1.936***	-1.975***	-1.850***	-2.030***	-1.767***	-1.951***	-1.739***	-1.834***	-2.122***	-1.886	-1.874
9. MET	-1.239***	-1.254***	-1.372***	-1.514***	-1.400***	-1.363***	-1.218***	-1.459***	-1.719***	-1.384	-1.400
10. MACH	-2.260***	-1.841***	-1.834***	-1.698***	-1.286***	-1.511***	-1.364***	-1.619***	-1.430***	-1.627	-1.565
11. VEH	-1.545***	-1.303***	-1.366***	-1.406***	-1.091***	-1.210***	-1.233***	-1.338***	-1.762***	-1.330	-1.321
12. OTHER	-2.029***	-1.589***	-1.361***	-1.494***	-1.372***	-1.283***	-1.272***	-1.165***	-1.716***	-1.406	-1.348
Aggregate BE	-1.786	-1.721	-1.699	-1.768	-1.686	-1.796	-1.687	-1.754	-1.858	-1.744	-1.747

- The border effect is large and negative in all industries
- There is substantial variation across industries
- The border reduces market shares to between 38.9% (2.57) and 12.3% (8.1) of their potential.

Participation plots Back

Average Border Effect

Table: Average Border Effect (Average treatment effect)

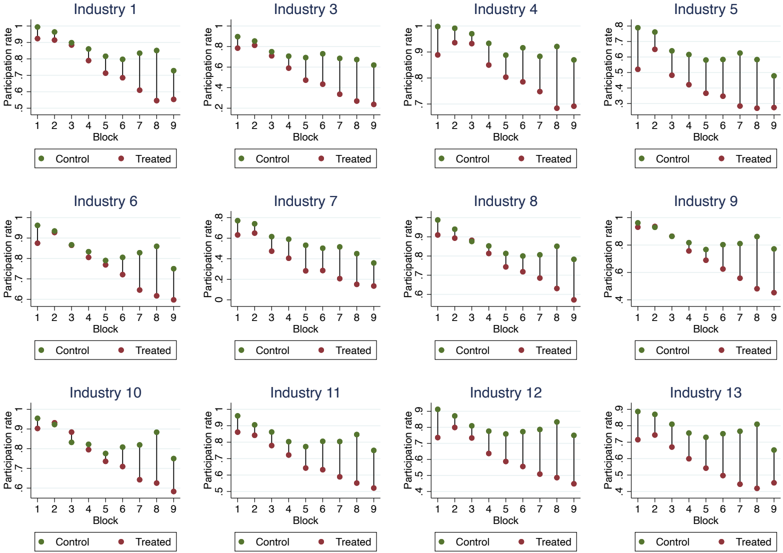
	Estimated β^{ATE}	
	All controls	Without number of borders
Weights: Size of blocks	-1.744	-1.299
Weights: Treated pairs	-1.747	-1.303

Notes: Average treatment effect calculated by computing the weighted average of the estimated coefficient of the *Border* dummy. The first row uses the number of observations in each blocks as weights, while the second row uses the number of treated units in each block.

- Not including number of borders makes us underestimate the border effect: reduction to 27.3% of market share potential (rel. to 17.5%=exp(-1.744))

Back

Participation rates: Industry-level



Back