Assessing a Spatially Explicit Zipf's Law for Population Density

FROM GLOBAL ECONOMETRIC PANELS TO LOCAL APPROACHES

Bruno Emma, D'Agata Alessia, Salvati Luca

Zipf's Law: a brief review

Pareto distribution

$$R = A\,P^{-\alpha}$$

(Auerbach)

$$\mathrm{Log}\,(R) = \mathrm{Log}\,(A) - \alpha\ \mathrm{Log}\,(P)$$

(Singer)

R : Rank of a city
P : Population of a city
A : City size constant
α : Pareto exponent

Zipf specification

$$P = \frac{K}{R^q}$$

$$P = K\,R^{-q}$$

$$\mathrm{Log}\,(P) = \mathrm{Log}\,(K) - q\ \mathrm{Log}\,(R)$$

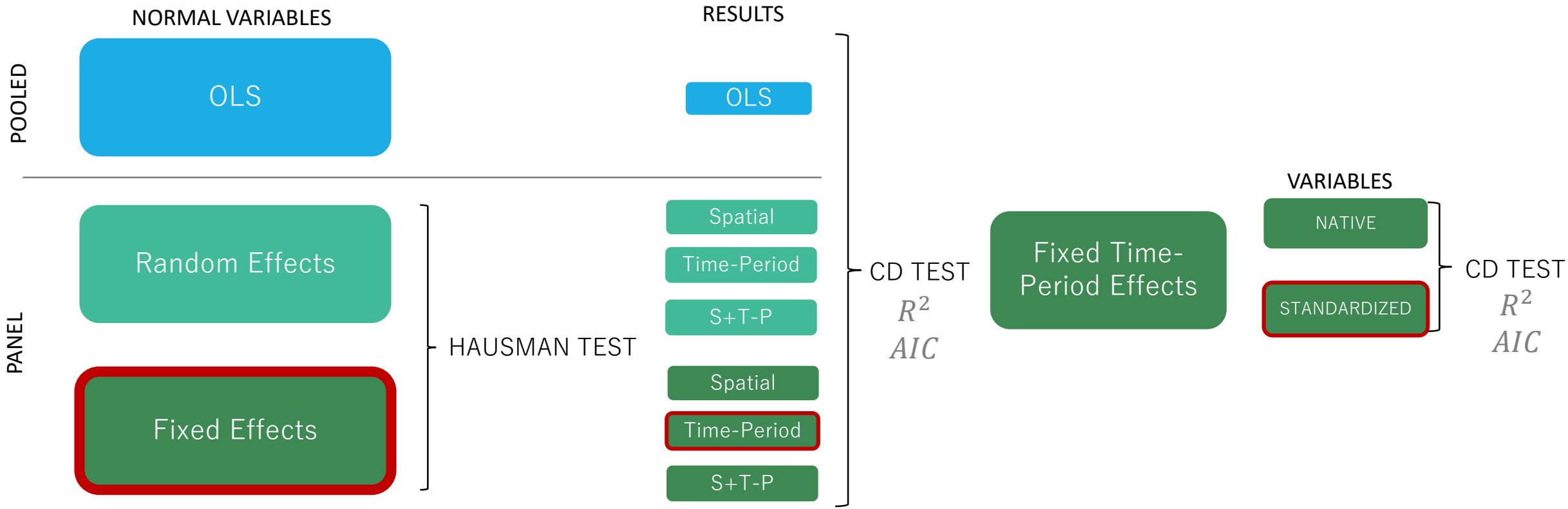
K : City size constant
q : Zipf exponent

Cities

- Distribution
- Spatial Relationships
- Growth
- Continuity
- Hierarchy
- Size
- Accessibility

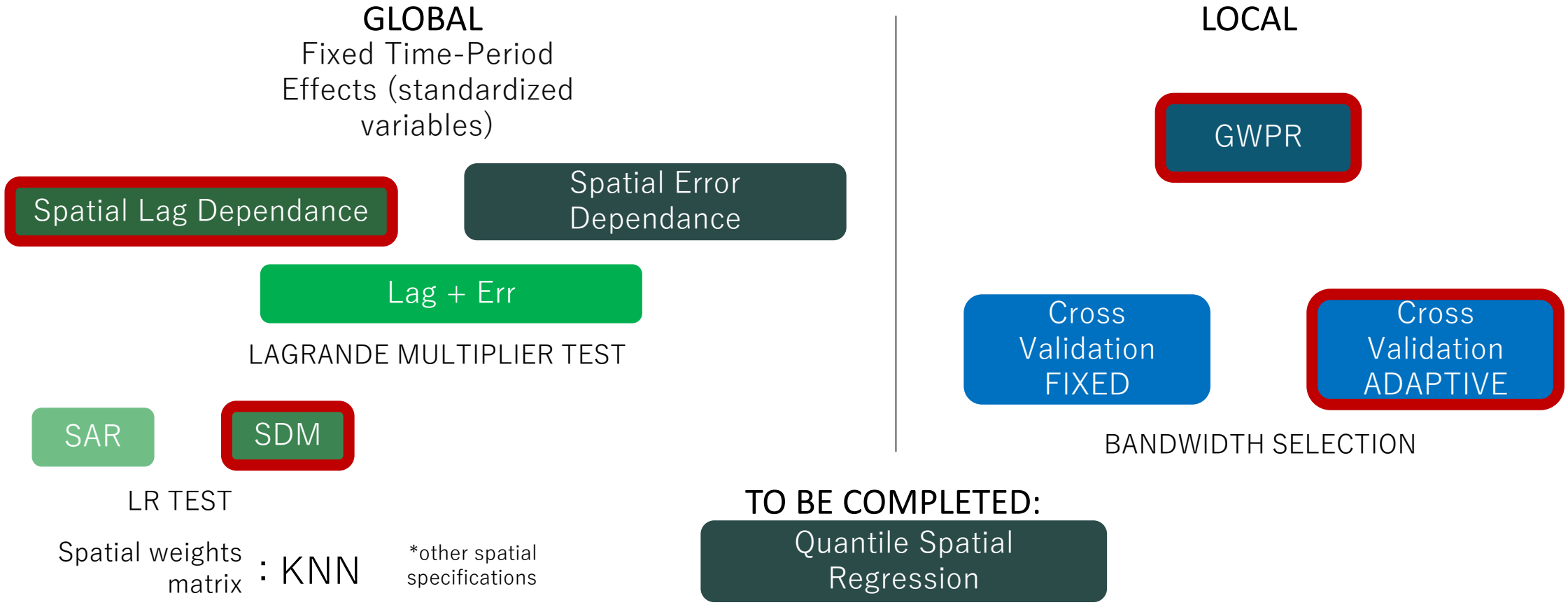
Sidra A, Shougeng H, Badar N A, Zipf's law, the coherence of the urban system and city size distribution, Physica A: Statistical Mechanics and its Applications, Volume 513, 2019, Pages 87-103.

Model selection: spatially implicit



In red the chosen options

Model selection: spatial extension



In red the chosen options

Case Study: Greece

Administrative boundaries of the 1033 Greek municipalities; official statistics released by Population Census 1961-2021.

Sources: USGS, ELSTAT, EUROSTAT;
Basemap: ESRI Topo World.

Results: spatially implicit model selection

Table 1: Panel specification.

	Intercept	Rankdl	R^2	Pesaran's CD test
		Pooled OLS		
	5.941***(0.013)	-1.633***(0.013)	0.937	328.81***
		Fixed Effects		
Spatial	/	-1.320***(0.018)	0.364	391.78***
Time-period	/	**-1.633***(0.005)**	**0.938**	**61.35***
Spatial + Time-period	/	-1.320***(0.017)	0.383	78.88***
		Random Effects		
Spatial	5.739***(0.025)	-1.555***(0.009)	0.782	344.38***
Time-period	5.941***(0.016)	-1.633***(0.009)	0.783	72.56***
Spatial + Time-period	5.728***(0.027)	-1.550***(0.005)	0.938	66.09***
Hausman Test	Fixed vs Random Effects: 225.61***			

Table 2: Variable selection

	Time-period FE	
	Var log native	**Var log standardised**
Rankdl	-1.633***(0.005)	**-0.972***(0.003)**
R^2	0.938	**0.945**
Pesaran's CD test	61.35***	**36.14***

Results: spatial extension

Table 3: Spatial component specification.

LM test	OLS vs SAR	OLS vs SEM
	LMlag: 32.553***	LMerr: 1.396
	SAR	***SDM***
Rankdl	-0.969***(0.003)	**-0.969*****(0.003)**
ρ	0.028***(0.005)	**-0.039****(0.019)**
WRankdl	-	**-0.072*****(0.020)**
LogLik	226.98	**232.92**
R^2_{pseudo}	0.945	**0.946**
LR test	SAR vs SDM	
	LR: 11.892***	

Table 4: Local model – Bandwidth selection.

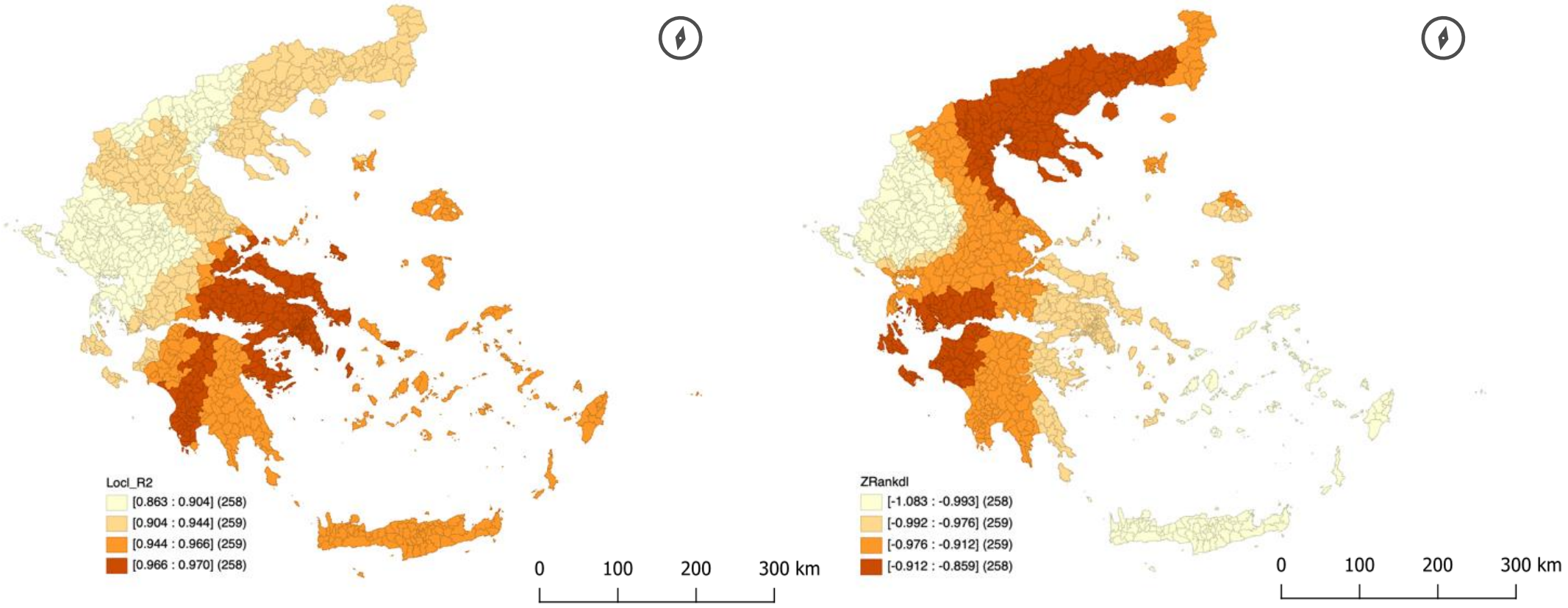
	GWPR			
	CV Fixed		CV adaptive	
	Rankdl	R^2	Rankdl	R^2
Min	-0.9950	0.9445	-1.0833	0.8633
Q_1	-0.9775	0.9447	-0.9923	0.9044
Median	-0.9731	0.9448	-0.9760	0.9441
Q_3	-0.9673	0.9450	-0.9116	0.9658
Max	-0.9354	0.9654	-0.8590	0.9701
Mean	-0.9728	0.9466	-0.9612	0.9339
sd	0.0086	0.0053	0.0480	0.0351

Table 5: Model comparison.

	Panel	GWPR	
		CV F	**CV A**
R^2	0.9447	0.9449	**0.9495**
AIC	-405.03	-416.220	**-1045.386**
BIC	-343.05	-354.236	**-983.402**

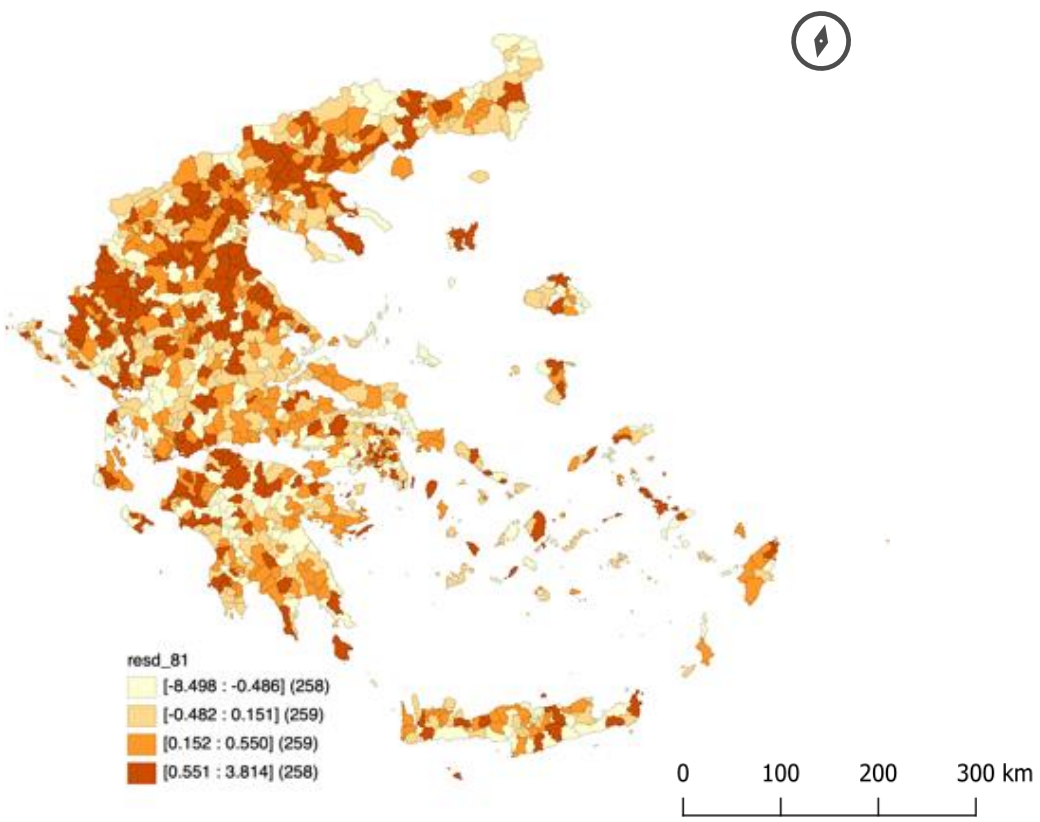
Local Panel Coefficents

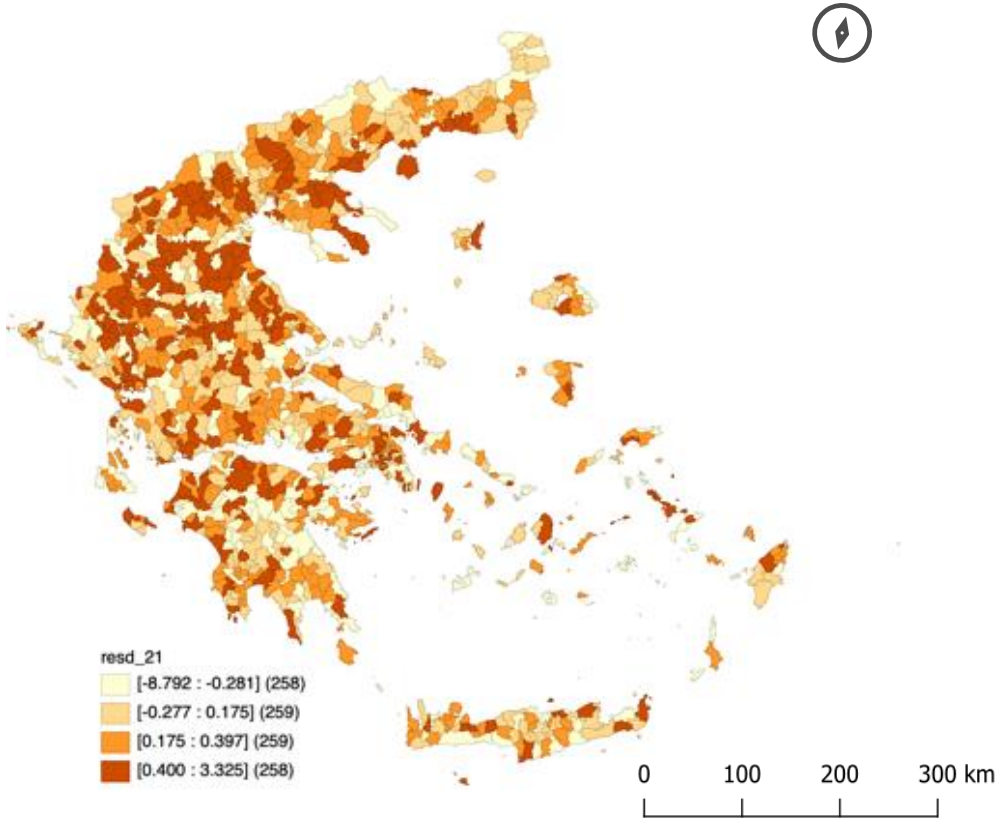
Fig. 1: Quartile distribution: R^2 (left); Rankdl/slope (right).



Local Panel Residuals

Fig. 2: Quartile distribution: standardized residuals; 1961 (left), 2021 (right).





Conclusions

- Fitting data
- Governing heterogeneity of spatial data
- Settlement evidence for Zipf law