

Presentation and Analysis of Spatial Data“

(6) Spatial Regression

Thomas Wöhler, Universität Konstanz
Kiev, October, 2016

$$y_i = \rho[\mathbf{W}y]_i + \boldsymbol{\beta}\mathbf{X}_i + e_i$$

$$e_i = \lambda[\mathbf{W}e]_i + u_i$$

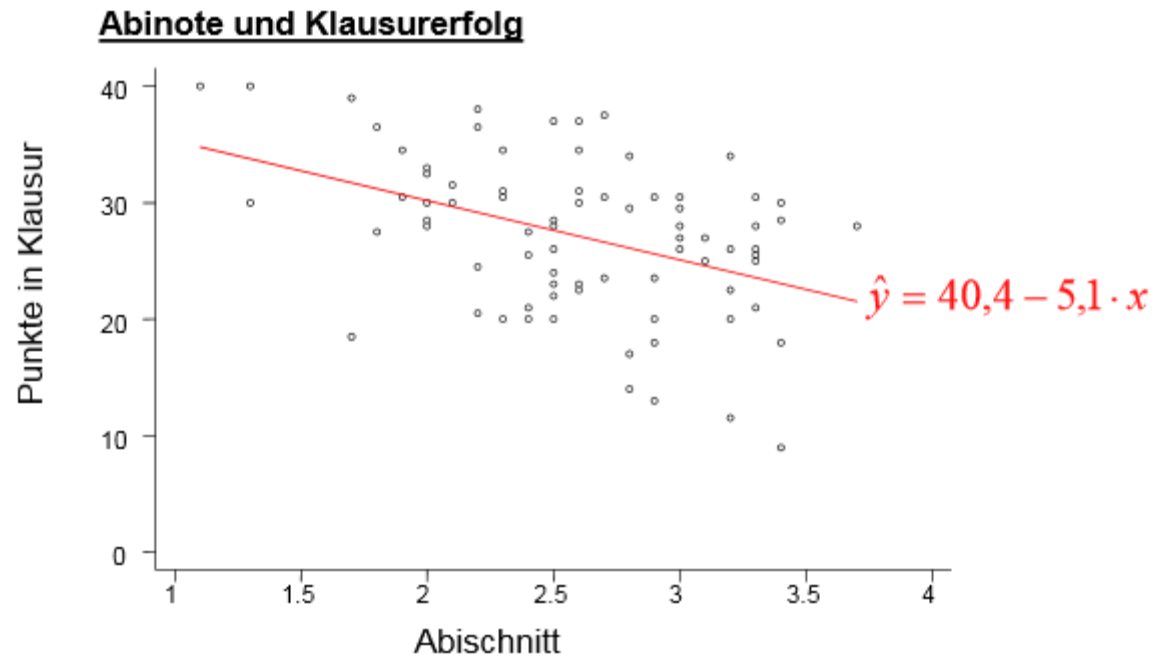
$$u_i \sim N(0, \sigma_u^2)$$

Agenda

- **Problems of OLS with spatial data**
- **Spatial Autokorrelation of Residuals**
- **Spatial Regression**

Regression analysis

Beispiel



Regression analysis

Spatial Heterogeneity

- Close units of observation are similar, because of shared stimuli.
- You don't have to assume spatial interaction between observational units.

Spatial Dependence

- Similarity of units of observation is due to direct mutual influence.

Regression analysis

Spatial Heterogeneity

- Close units of observation are similar, because of shared stimuli.
- You don't have to assume spatial interaction between observational units.

→ Heteroskedasticity

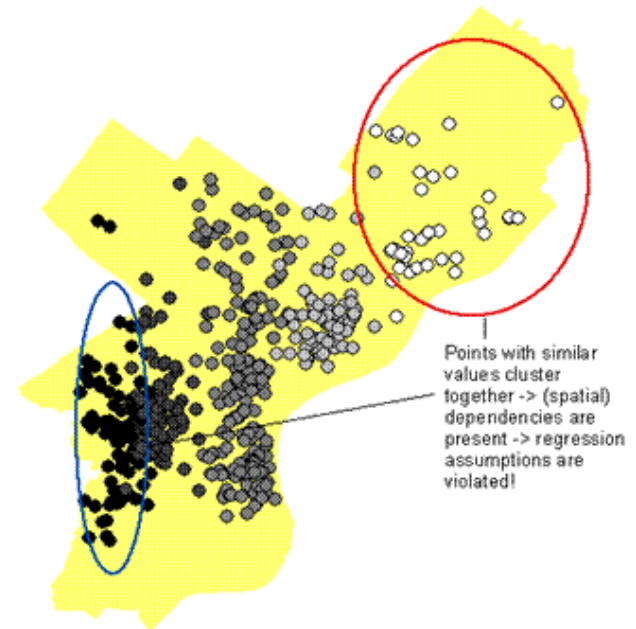
Spatial Dependence

- Similarity of units of observation is due to direct mutual influence.

→ Correlation of residuals

Regression analysis

→ A statistical model with spatial autokorrelation violates the assumptions of OLS.
Coefficients may be overestimated and standard errors too small.



Spatial Regression

Models for spatial dependence

$$y_i = \rho[\mathbf{W}y]_i + \boldsymbol{\beta}\mathbf{X}_i + e_i$$

$$e_i = \lambda[\mathbf{W}e]_i + u_i$$

$$u_i \sim N(0, \sigma_u^2)$$

- If $\lambda = 0$, spatial lag model
- If $\rho = 0$, spatial error model
- If $\lambda = \rho = 0$, OLS

- Because we are interested in spatial processes in our data.
- Because we do not want to violate the assumptions of OLS.

Spatial Regression – spatial error

We can account for spatial dependence of the residuals in a spatial error Modell.

$$(1) y = X\beta + u$$

$$(2) u = \lambda Wu + \varepsilon$$

u is a vector of the spatially weighted error terms and λ is the coefficient of the spatial error, it is 0 when there is noch spatial correktion between the error terms.

Spatial Regression – spatial error

We choose a *spatial error* model, when there is no theoretical reason to assume, that:

- y is not influenced by spatially close y and x values,
- When the omitted spatial variables are hard to measure, but when they have important spatial influence,
- or when there is spatial dependence but we are not interested in it.

Spatial Regression – spatial lag

- When we assume, that spatial distance between observational units has a direct effect and we want to model this process, one should think about using spatial lag models.
- Research in sociology could be small scale surveys in school classes, in student dorms, etc. where we interview all subjects in the sample frame.
- Example in political science: Processes of democratization

Spatial Regression – spatial lag

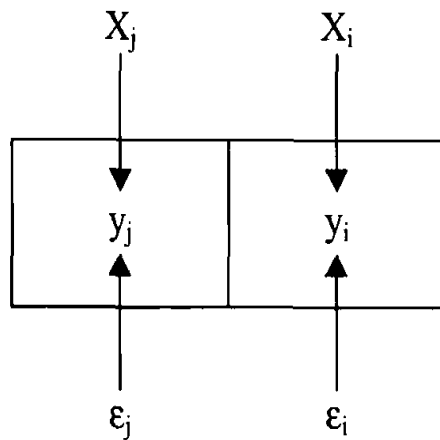
The statistical model is thus:

$$(1) y = \rho W y + X\beta + \varepsilon$$

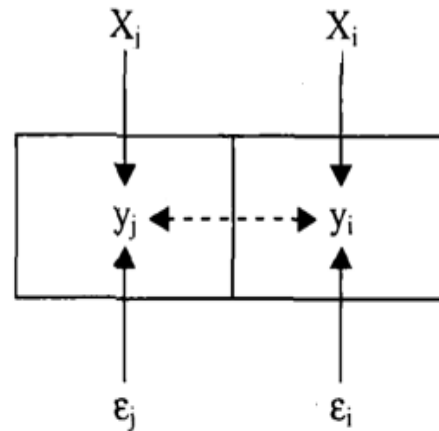
- The spatial lag is represented by a spatially weighted average of the dependent variable. A significant coefficient ρ can be interpreted as influence of neighbouring units.
- Problem of endogeneity: y is a function of $W y$ and vice versa.

Spatial Regression – Comparison

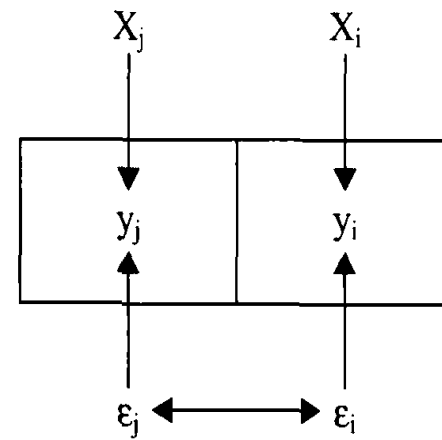
OLS



SPATIAL LAG

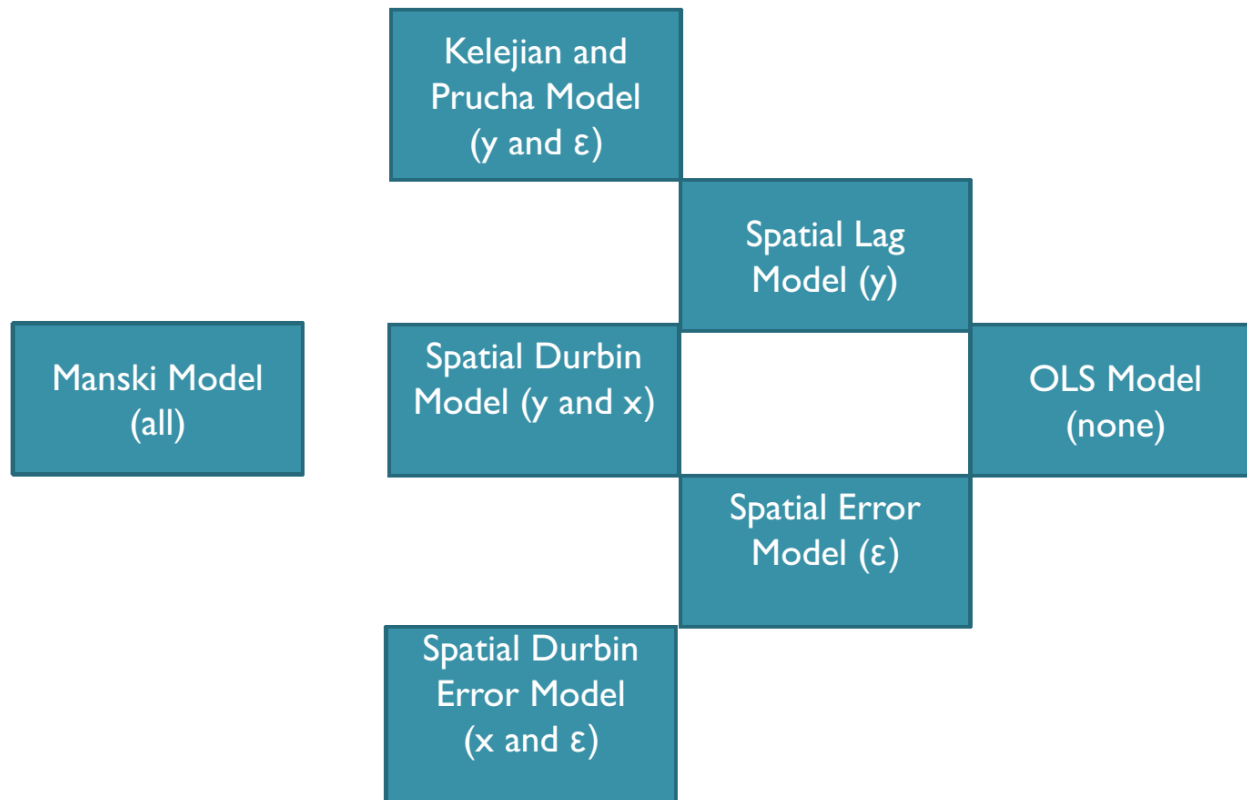


SPATIAL ERROR



Baller, R., L. Anselin, S. Messner, G. Deane and D. Hawkins. 2001. *Structural covariates of US County homicide rates: incorporating spatial effects*,. *Criminology* , 39, 561-590

Types of spatial regressions



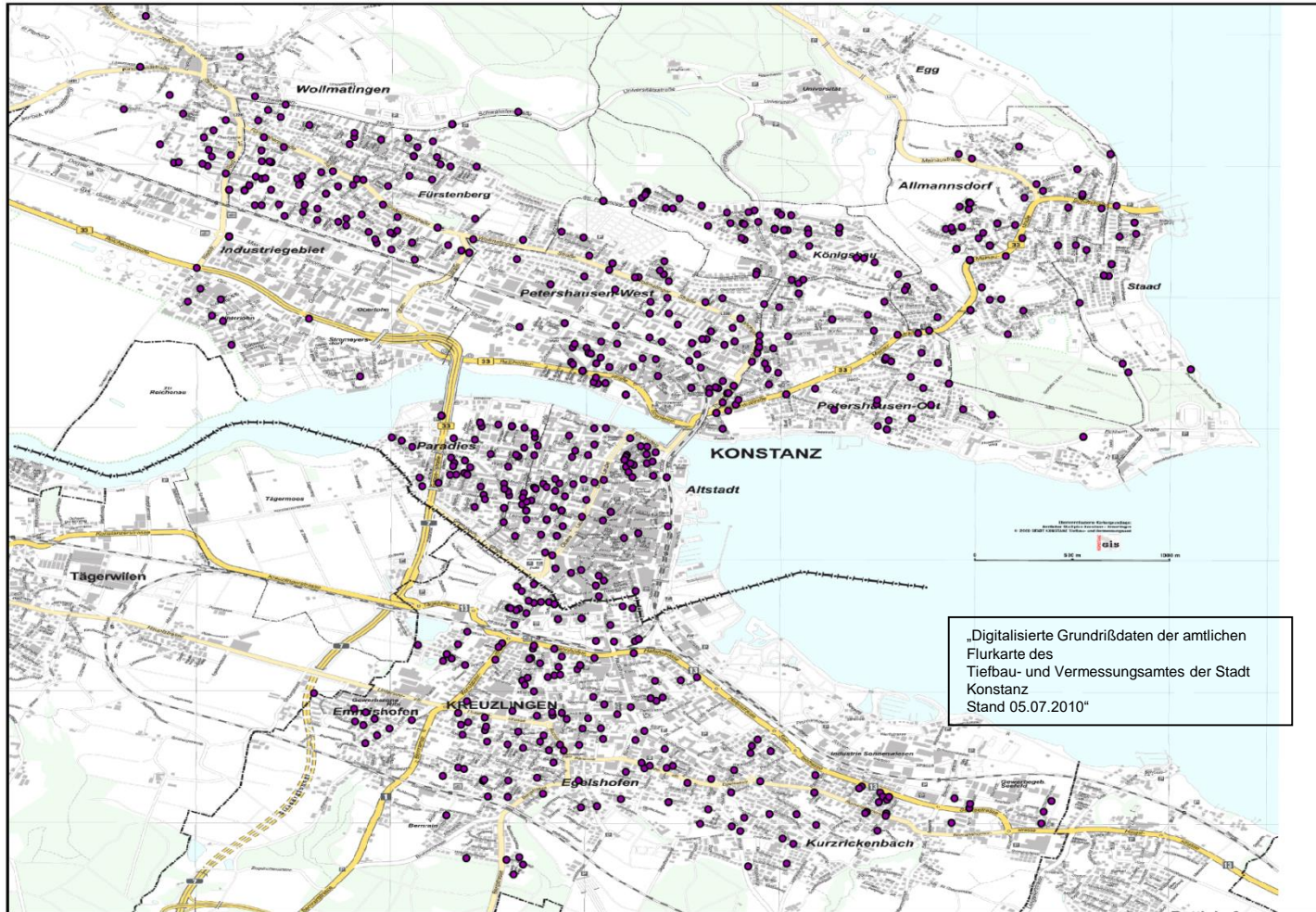
Steps of a multivariate analysis of spatial data

- Explorative data analysis (Moran's I)
- Multivariate OLS
- Test of spatial autokorrelation of the residuals
- Possible inclusion of additional characteristics of the context to explain away the spatial autokorrelation, possibly even multilevel modelling
- When there is still spatial autokorrelation → spatial regression
- Selection of model according to theory

Research example

1. Respondents and data
2. Geographic analysis of residuals

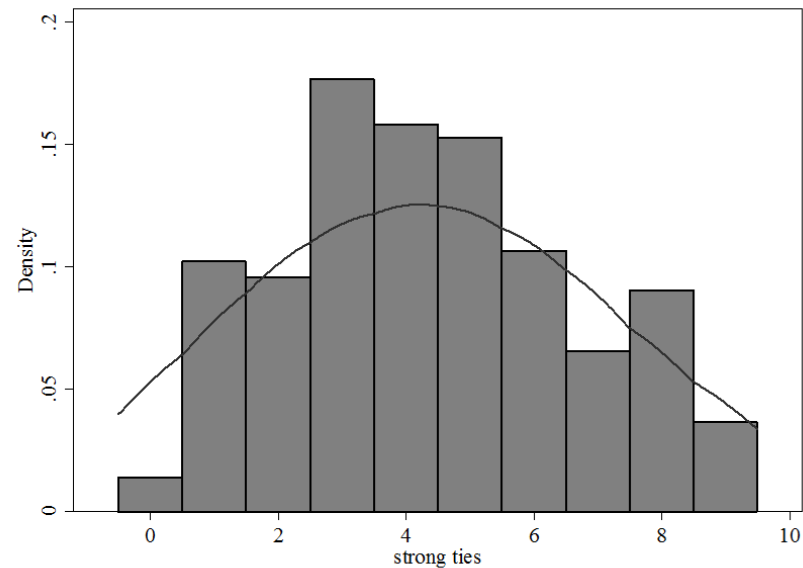
Respondents



Data

Assessment of close personal social relationships with name generators:

1. „With whom do you discuss important personal matters?“
2. „With whom do you spent you leisure time?“

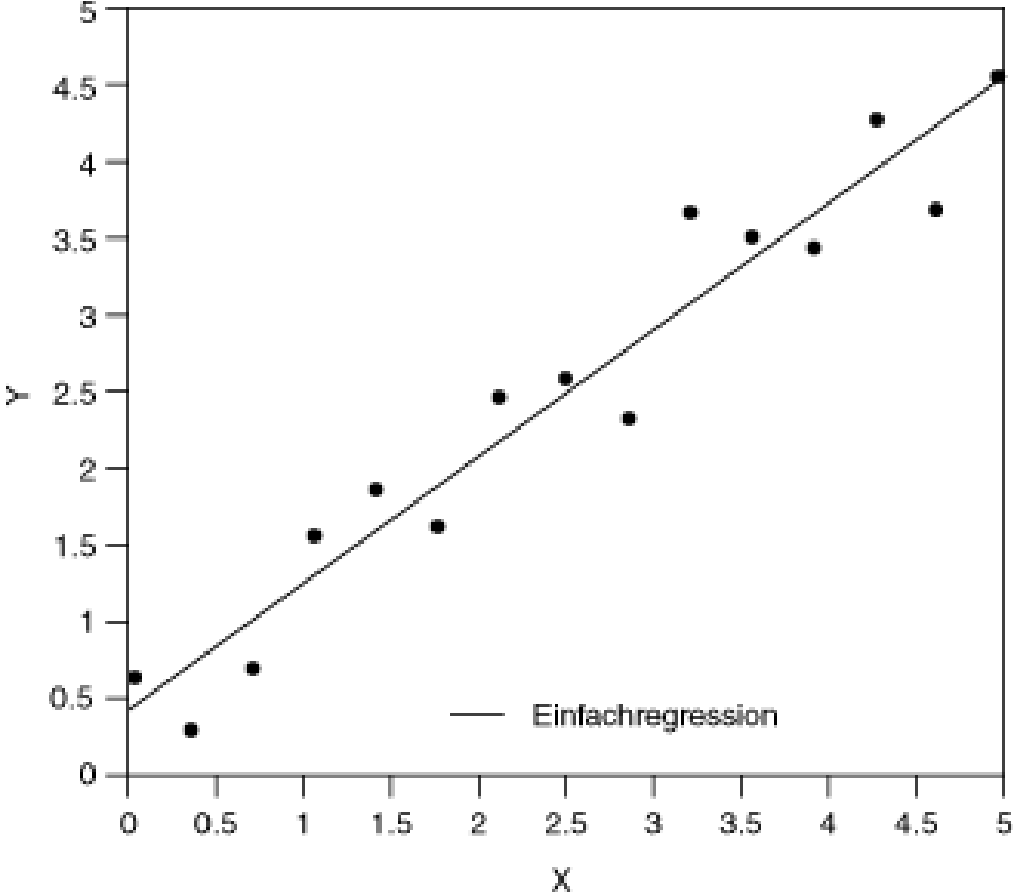


Simple model of explanation

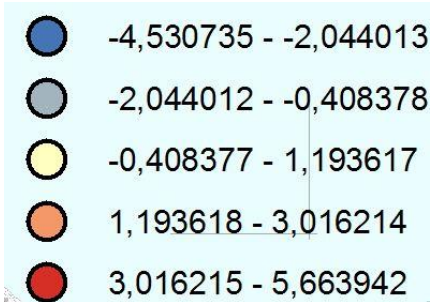
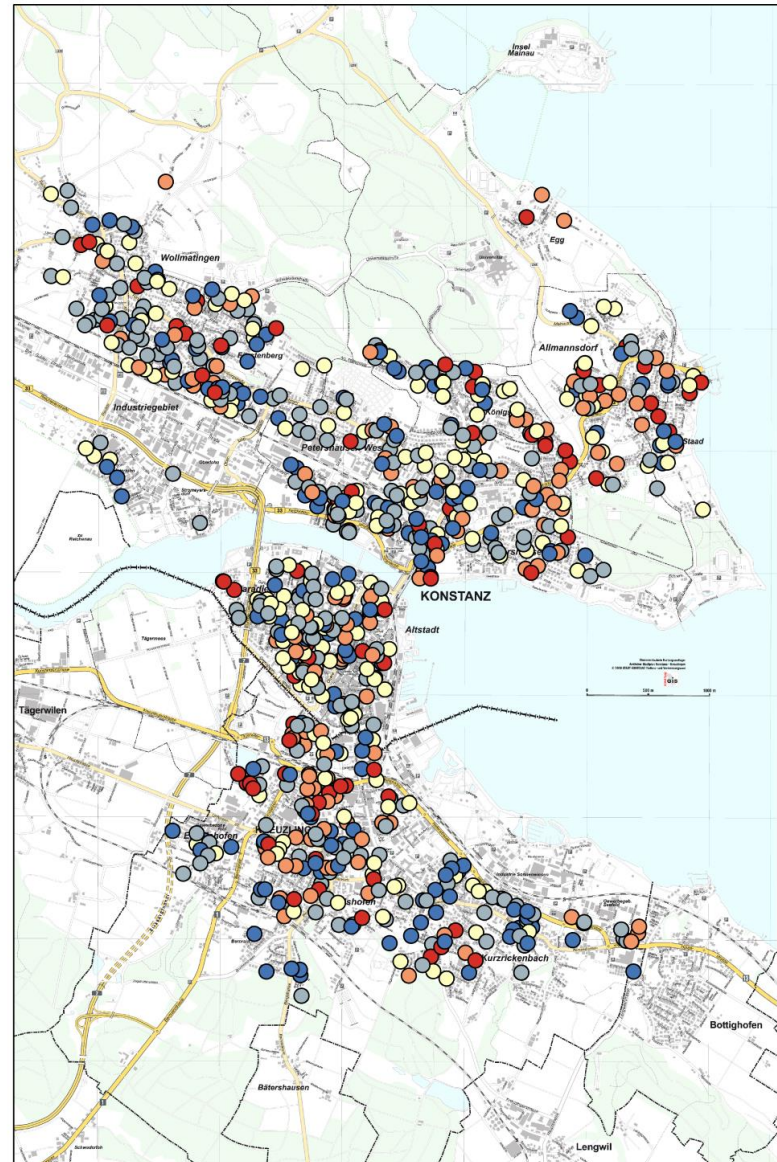
	Modell (1)
	strong ties
Migrant (=1)	-0.64 (-4.11)**
Female (=1)	0.42 (2.88)**
Age (in years)	-0.02 (-0.74)
Age (squared)	-0.00 (-0.03)
Living in Neighborhood (in years, log)	-0.08 (-1.20)
City (Kreuzlingen=1)	0.04 (0.22)

Intercept	5.38 (10.25)**
F	8.43
N	928
NH	

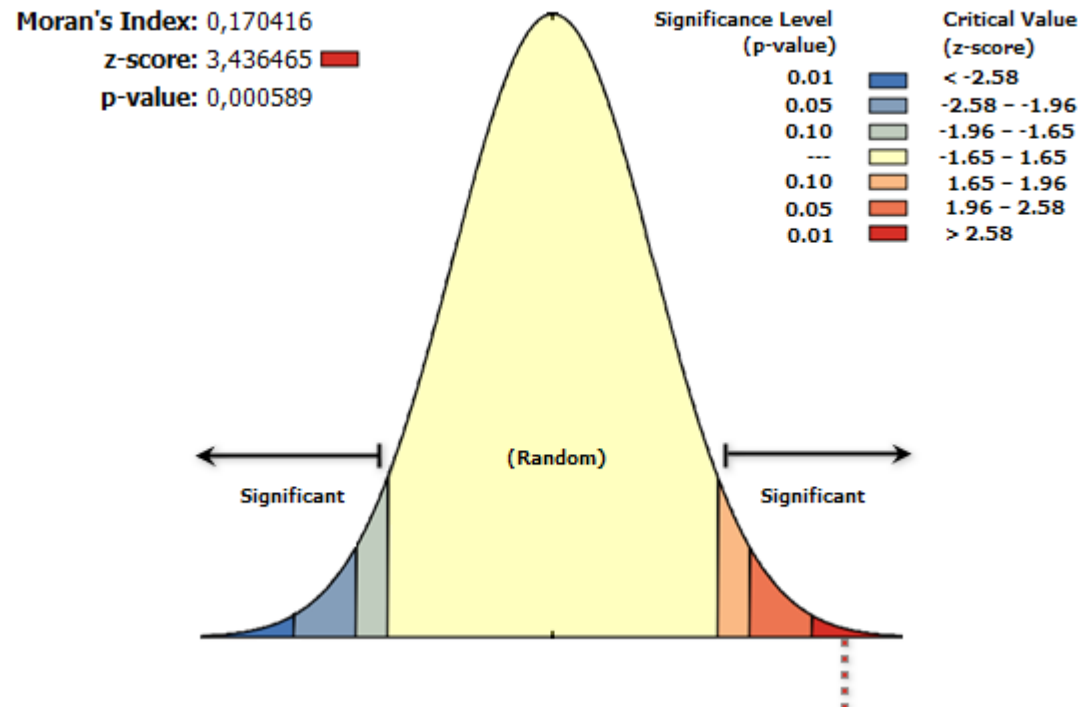
Analysis of residuals



Model 1: Residuals



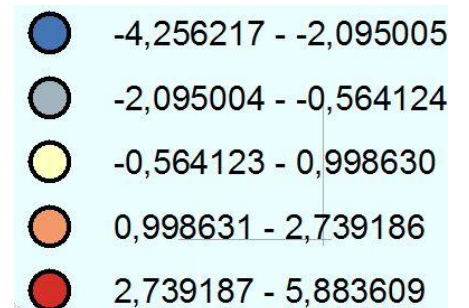
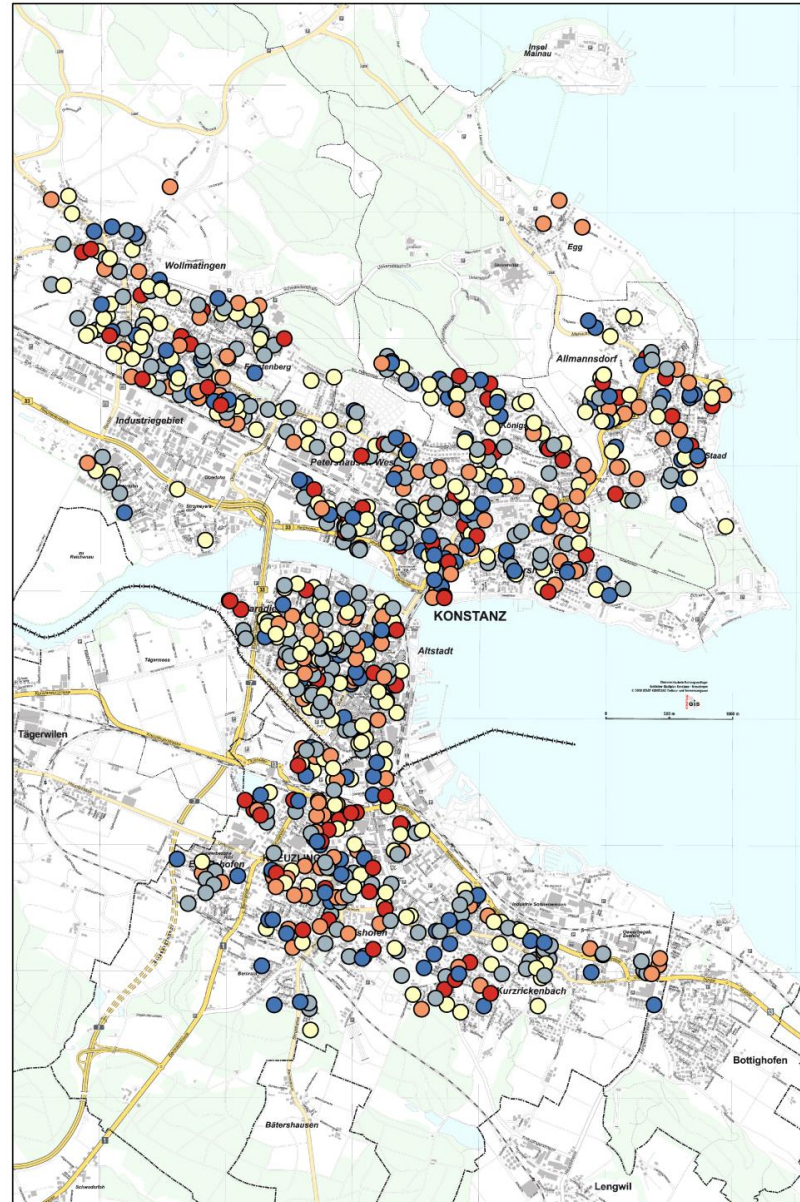
Spatial dependence of residuals in Model 2



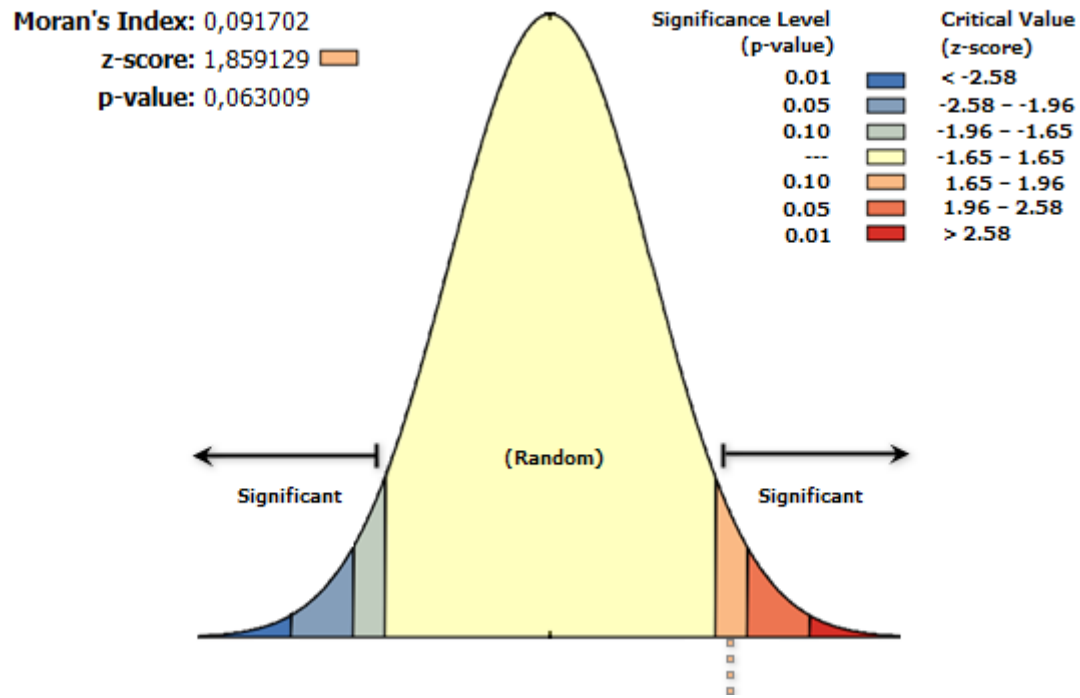
Model with additional characteristic of context

	Modell (1) strong ties	Modell (2) strong ties
Migrant (=1)	-0.64 (-4.11)**	-0.58 (-3.75)**
Female (=1)	0.42 (2.88)**	0.41 (2.90)**
Age (in years)	-0.02 (-0.74)	-0.02 (-0.78)
Age (squared)	-0.00 (-0.03)	-0.00 (-0.06)
Living in Neighborhood (in years, log)	-0.08 (-1.20)	-0.05 (-0.79)
City (Kreuzlingen=1)	0.04 (0.22)	0.13 (0.58)
Socio-economic status Of Neighborhood		0.14 (3.31)**
var u		0.42 (2.31)*
var i		0.74 (29.89)**
Intercept	5.38 (10.25)**	4.6 (8.23)**
F / Wald chi2	8.43	60.08
N	928	928
NH		93

Model 2: Residuals



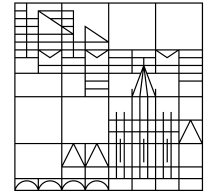
Spatial dependence of residuals in Model 2



Steps of a multivariate analysis of spatial data

- Explorative data analysis (Moran's I)
- Multivariate OLS
- Test of spatial autokorrelation of the residuals
- Possible inclusion of additional characteristics of the context to explain away the spatial autokorrelation, possibly even multilevel modelling
- When there is still spatial autokorrelation → spatial regression
- Selection of model according to theory

Universität
Konstanz



Thank You
For Your Attention!

Thomas Wöhler

Fachbereich Geschichte und Soziologie, AG Hinz

Tel.: +49 (0) 75 31/88 - 3345

thomas.woehler@uni-konstanz.de