Universität Konstanz

# „Presentation and Analysis of Spatial Data“

# (7) Spatial conditional effects (Geographically Weighted Regression)

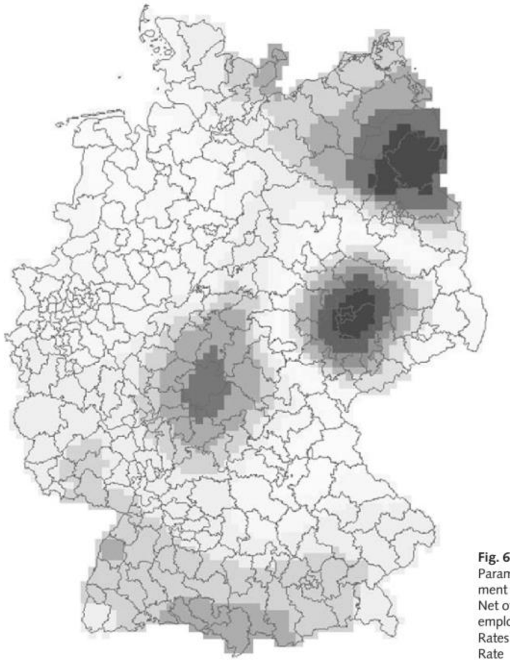


Fig. 6
Paran
ment
Net o
emplc
Rates
Rate

**Thomas Wöhler, Universität Konstanz**

Kiev, October, 2016

# Agenda

- Geographically Weighted Regression
- Additional Resources

# Geographically Weighted Regression

**Assumption of OLS Regression:**

$$y = \alpha + \beta X + \varepsilon$$

→ **Same effects across the whole are of analysis.**

**Possible reasons for the spatial variation of coefficients:**

- Different samples, different sampling, different data generation
- Models wrongly specified – omitted variables
- Relationships between variables vary substantially with spatial location

# Geographically Weighted Regression

Local statistical technique for the analysis of spatial variation in the relationship between variables.
Local effects instead of global means

from: $$y = \alpha + \beta X + \varepsilon$$

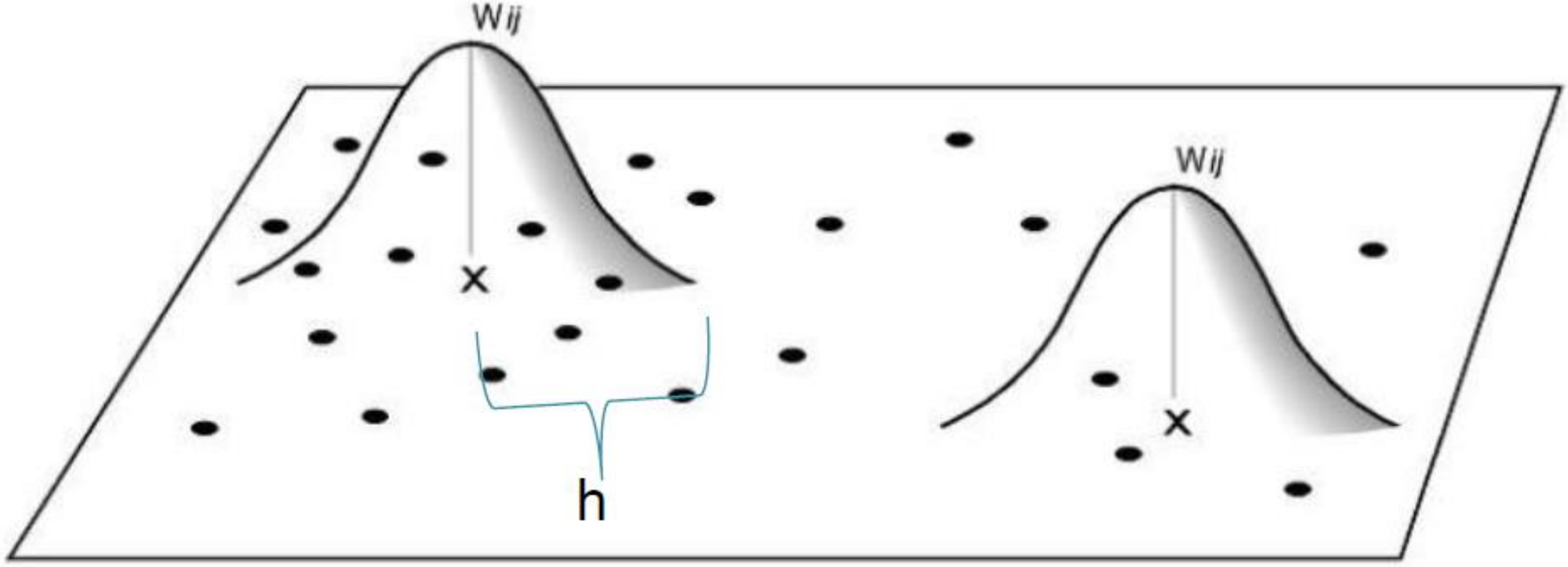
We get to: $$y_i = \alpha_i + \beta_i X + \varepsilon_i$$

i.e. for every observation we estimate a local set of coefficients.

Observations who are spatially closer to each other have more weight in the local analyses than observations further apart.

See:
Fotheringham, Brunsdon and Charlton *Geographically Weighted Regression* Wiley, 2002

# Principle of GWR



# Research Example 1

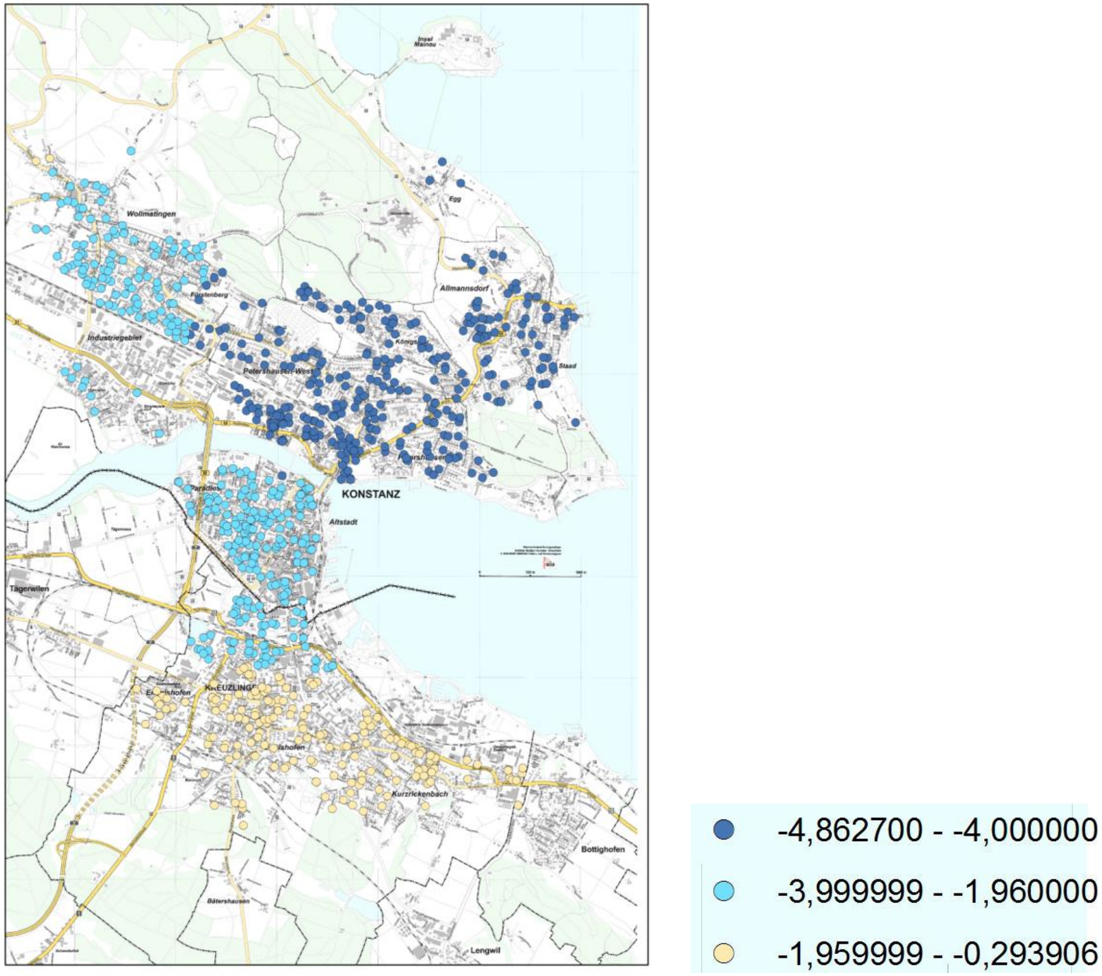
GWR:

Variation of Coefficients

Variable:

Migrant (=1)

(z-Values)



# Research Example 2

## Space Matters. The Group Threat Hypothesis Revisited with Geographically Weighted Regression. The Case of the NPD 2009 Electoral Success

**Celine Teney***

Wissenschaftszentrum Berlin für Sozialforschung, Reichpietschufer, 50, 10785 Berlin, Germany
teney@wzb.eu

**Summary:** A geographically weighted regression approach is used to assess the association of the electoral success of the NPD, an extreme right-wing political party, during the 2009 German federal election with levels of and changes in immigrant and unemployment rates. The results do not support the group threat hypothesis: the immigrant rate remains non-significant in large areas of West Germany while it shows a negative and significant relationship with NPD electoral success in most localities in East Germany as well as in Northern Bavaria. Instead, findings tend to confirm the contact hypothesis: a higher percentage of immigrants within an electoral district seems to lead to larger interethnic contact opportunities and thus to a lower proportion of votes for the NPD. The largest significant positive association of unemployment rate with NPD electoral results is observed with respect to localities that are situated around the former border between East and West Germany. The large regional variations in the effects of immigrant and unemployment rates point to different mechanisms which are at stake in the association of populist radical right success with unemployment and immigrant rates. These findings illustrate the importance of spatial variability and make the case for a broader new research agenda dedicated to exploring the mechanisms underlying spatial nonstationarity.

**Keywords:** Populist Radical Right; Germany; NPD; Group Threat; Elections; Unemployment; Immigration; Geographically Weighted Regression.

# Research Example 2

**Table 2** Global OLS Regression Results, GWR Coefficient Ranges for East and West Germany and Results of the Monte Carlo Test of Spatial Variability of the Group Threat Effect on NDP Results during the 2009 Federal Elections (Without Controlling Variables)

| | OLS estimates | | GWR estimates | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Model 1** | **Model 2** | **Lwr Quartile** | | **Median** | | **Upr Quartile** | | **MC p-Value** |
| | | | East | West | East | West | East | West | |
| intercept | 1.20 (0.12) *** | 1.21 (0.10) *** | 1.25 | 0.59 | 2.71 | 0.99 | 3.91 | 1.24 | *** |
| % immigrant | –0.08 (0.008)*** | –0.03 (0.008)*** | –0.32 | –0.05 | –0.16 | –0.04 | –0.09 | –0.02 | *** |
| % unemployed | 0.12 (0.01)*** | 0.03 (0.01)*** | 0.01 | 0.05 | 0.06 | 0.07 | 0.11 | 0.10 | * |
| East | | 1.60 (0.13)*** | | | | | | | |
| Adjusted $R^2$ | 0.49 | 0.66 | 0.87 | | | | | | |
| AIC | 620.01 | 497.58 | 309.86 | | | | | | |

N=299. OLS = ordinary least squares; GWR = geographically weighted regression; MC = Monte Carlo tests of spatial variability; AIC = Akaike Information Criterion

The GWR model is based on a fixed spatial kernel function with a bandwidth of 28 electoral districts. This model was selected according to the AIC minimization procedure.

**p*<0.01, ****p*<0.001*

# Research Example 2

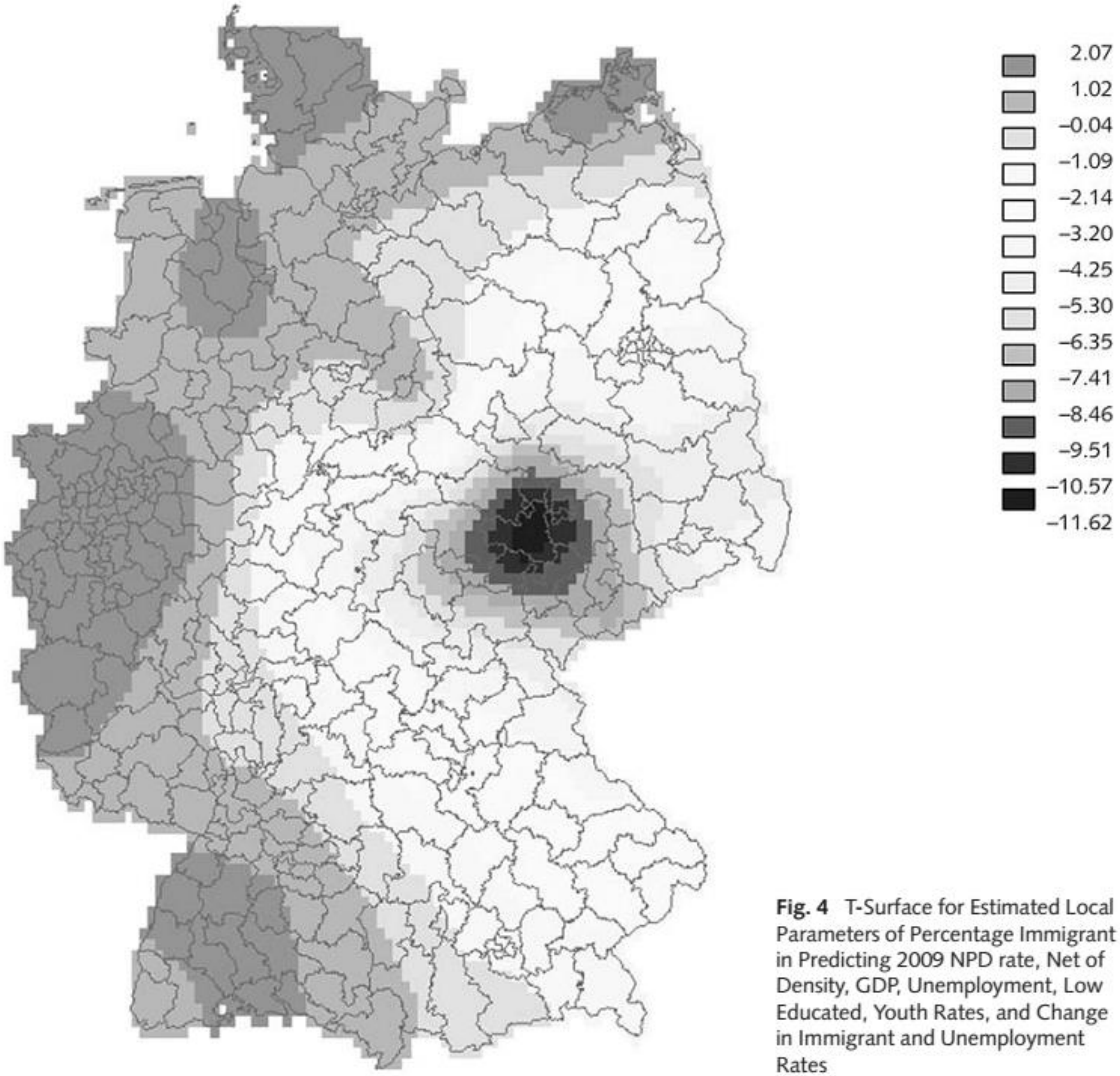
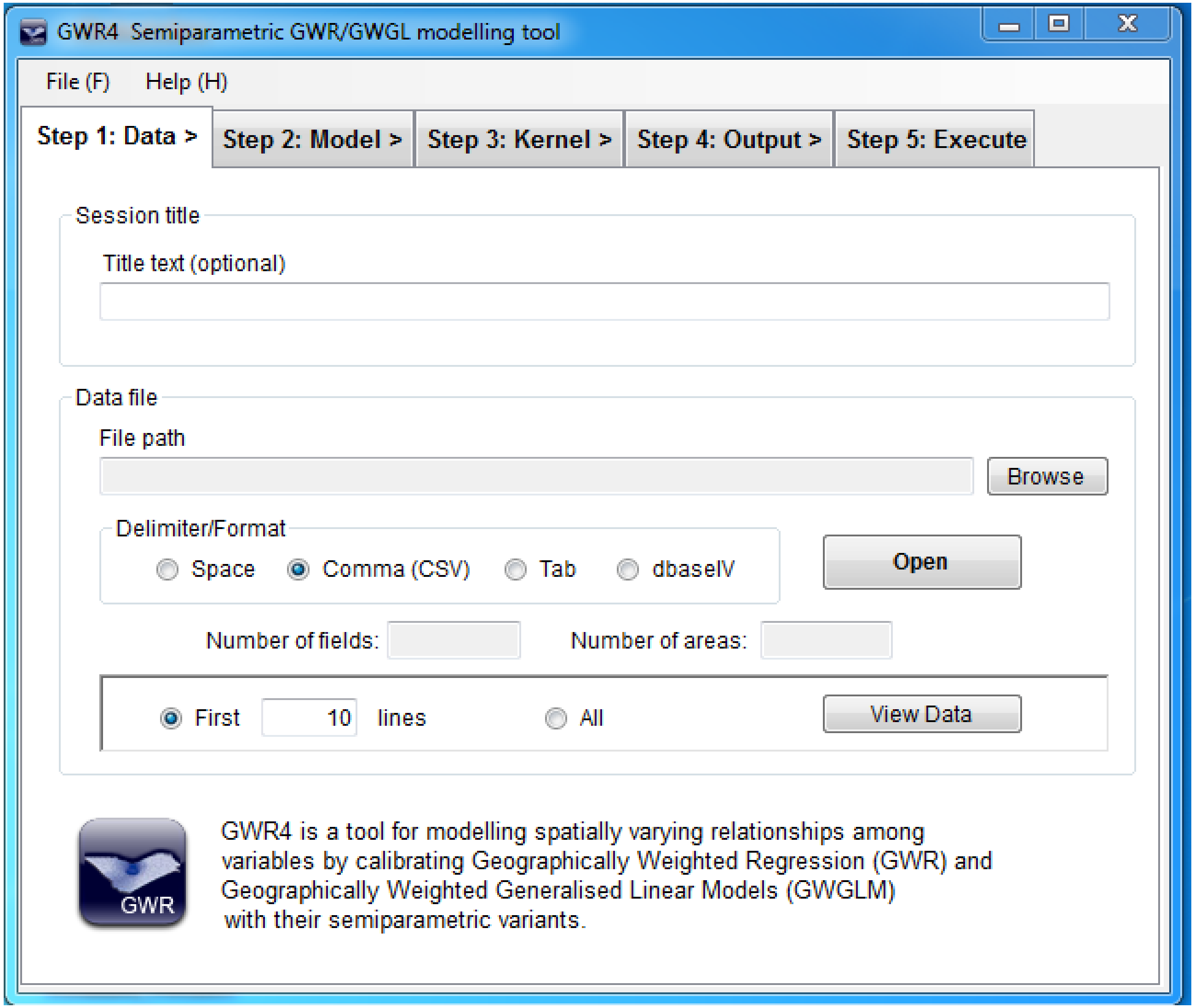



**Fig. 4** T-Surface for Estimated Local Parameters of Percentage Immigrant in Predicting 2009 NPD rate, Net of Density, GDP, Unemployment, Low Educated, Youth Rates, and Change in Immigrant and Unemployment Rates

# GWR4



GWR4 Semiparametric GWR/GWGL modelling tool
File (F)
Help (H)
Step 1: Data >
Step 2: Model >
Step 3: Kernel >
Step 4: Output >
Step 5: Execute
Session title
Title text (optional)
Data file
File path
Browse
Delimiter/Format
Space
Comma (CSV)
Tab
dbaseIV
Open
Number of fields:
Number of areas:
First
10
lines
All
View Data
GWR
GWR4 is a tool for modelling spatially varying relationships among variables by calibrating Geographically Weighted Regression (GWR) and Geographically Weighted Generalised Linear Models (GWGLM) with their semiparametric variants.

# Geographically Weighted Regression

## Discussion

- Each regression is only based on few observation.
- Therefore the power of the estimates is rather low.
- Solution: Larger buffer / neighborhood (→ smaller decay function) But how far makes sense?

You need a very strong theory to explain why regresion coefficients are different at different places.
→ Explorative method?

The validity of the inference test might be comprised because the observations are not independent of each other (see spatial regression).

# Additional Material

## Literature:

Meyer, Reto/Bruderer Enzler, Heidi (2013): Geographische Informationssysteme (GIS) und ihre Anwendung in den Sozialwissenschaften am Beispiel des Schweizer Umweltsurveys. methoden, daten, analysen 7(3): 317-346.

Logan, JR (2012): Making a place for space: Spatial thinking in social sciences. Annual Review of Sociology 38: 507-524.

Downey, L (2006): Using geographic information systems to reconceptualize spatial relationships and ecological contexts. American Journal of Sociology 112(2): 567-612.

Darmofal, D. (forthcoming): Spatial analysis for social sciences (Analytical Methods for Social Research Series, Cambridge University Press)

## R spatial tools

http://cran.r-project.org/web/views/Spatial.html

Bivand, Roger/Edzer Pebesma/Virgilio Gómez-Rubio (2013): Applied Spatial Data Analysis with R. New York: Springer.

## Websites

For teaching: http://www.teachspatial.org

GIS and Population Science: http://GISPopSCi.org.

http://spatialanalysisonline.com

# Additional Material

**My personal view on spatial data and spatial analyses in the Social Sciences:**

- Great potential to generate better data.
- This allows to ask new research questions.
- Spatial statistics is oftentimes more explorative.
- In spatial statistics, strong theory about the effects of space should guide the choice of statistical models.
- One goal could be to explain spatial autocorrelation away by including substantial determinants. This can inform aboiut the black-box “space”.

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