

Contexts Matter: Yes! But How?

Some Thoughts on the Analysis of Spatially Operative Mechanisms

Katrin Auspurg und Thomas Wöhler
LMU München / Universität Konstanz
VIU 2017



Agenda

- Problem
- Application
- Approaches

The Problem

Many analyses in sociology with regard to context. Increasingly analyses with more sophisticated geo-data.

→ More possibilities, more problems?

- Is the use of multilevel models appropriate?
 - Disregarding the spatial structure may lead to biased standard errors and coefficients
 - Spatial analysis can give answers about spatially operative mechanisms
 - FE is not always adequate
- Is the level of aggregation of the spatial units appropriate?
 - E.g. the use of 94 „Raumordnungsregionen“ (RORs) in Germany
 - Smaller (or larger) spatial units might reflect the theoretical reasoning better

The Problem from a Statistical Point of View

Spatial Dependence (see e.g. Darmofal 2015)

1. Spatial diffusion:

- Units' behavior is directly influenced by the behavior of „neighboring units“ (simultaneous dependence: neighbors influence the behavior of their neighbors and vice versa)
- Omitted variable problem (spatially lagged dependent variable is omitted in common OLS;)
- If ignored in OLS: can produce biased and inconsistent parameter estimates

2. Geographic clustering:

- Neighboring units share characteristics that promote the behavior in question (does not reflect a truly spatial process, but merely the geographic clustering of the sources of the behavior of interest)
- Special case of heteroskedasticity: Variance of the error term varies with spatial location
- If ignored in OLS: standard errors are biased downward

Spatial Regression?

$$Y = \rho WY + X\beta + \varepsilon$$

$$\varepsilon = \delta W\varepsilon + \xi$$

WY = spatially lagged dependent variable

$W\varepsilon$ = spatially lagged error term

Spatial Regression?

$$Y = \rho WY + X\beta + \varepsilon$$

$$\varepsilon = \delta W\varepsilon + \xi$$

WY = spatially lagged dependent variable

$W\varepsilon$ = spatially lagged error term

Question 1: Neighboring units and influence?

ROR-ID	ROR Name	Year			
		1994	2000	2004	2010
905	Donau-Wald	X	X	X	X
907	Ingolstadt	-	X	-	X
910	München	X	X	X	X
913	Oberland	-	-	X	X
915	Regensburg	-	X	X	X
916	Südostoberbayern	X	-	X	-

Spatial Regression?

$$Y = \rho WY + X\beta + \varepsilon$$

$$\varepsilon = \delta W\varepsilon + \xi$$

WY = spatially lagged dependent variable

$W\varepsilon$ = spatially lagged error term

Question 2: What distance, what influence, what location?

The Problem

Why Space, what spatially operative mechanisms?

Space as framework / context for social action:

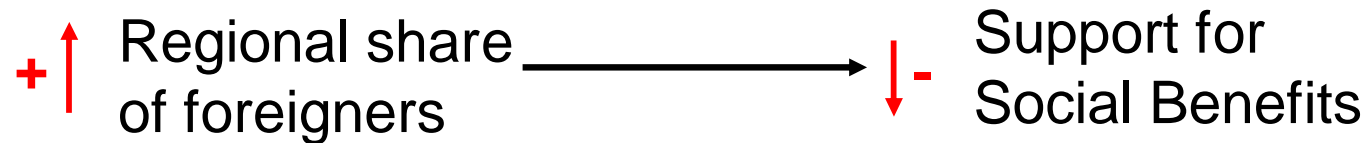
- through distance
- through institutions
- through infrastructure
- through social context
- through built environment
- ...

Space as proxy:

- for differences in socialization
- for differences in lifestyle
- ...

Example: Publication in ASR 2016

Research question: Do (rising) shares of foreigners lead to less support for Social Benefits?



Theory: Conflict Hypothesis

- Solidarity with ingroup in comparison to outgroups
- Current influx of outgroups (longitudinal change) are perceived as threat
- Increasing share of foreigners on local level threatens solidarity with Welfare State

Analytical Strategy

- DV is a question in the Allbus about Social Benefits:
„Should social benefits be cut in the future, should things stay as they are, or should social benefits be extended?“
- Only German nationals
- 4 Waves: 1994, 2000, 2004, 2010
- Central IV: Share of foreigners in 96 ROR
 - RORs represent „day-to-day experiences“ and local labor markets
- Checks of robustness on NUTS3 (Kreise) and NUTS1 (Federal States)

Statistical Model

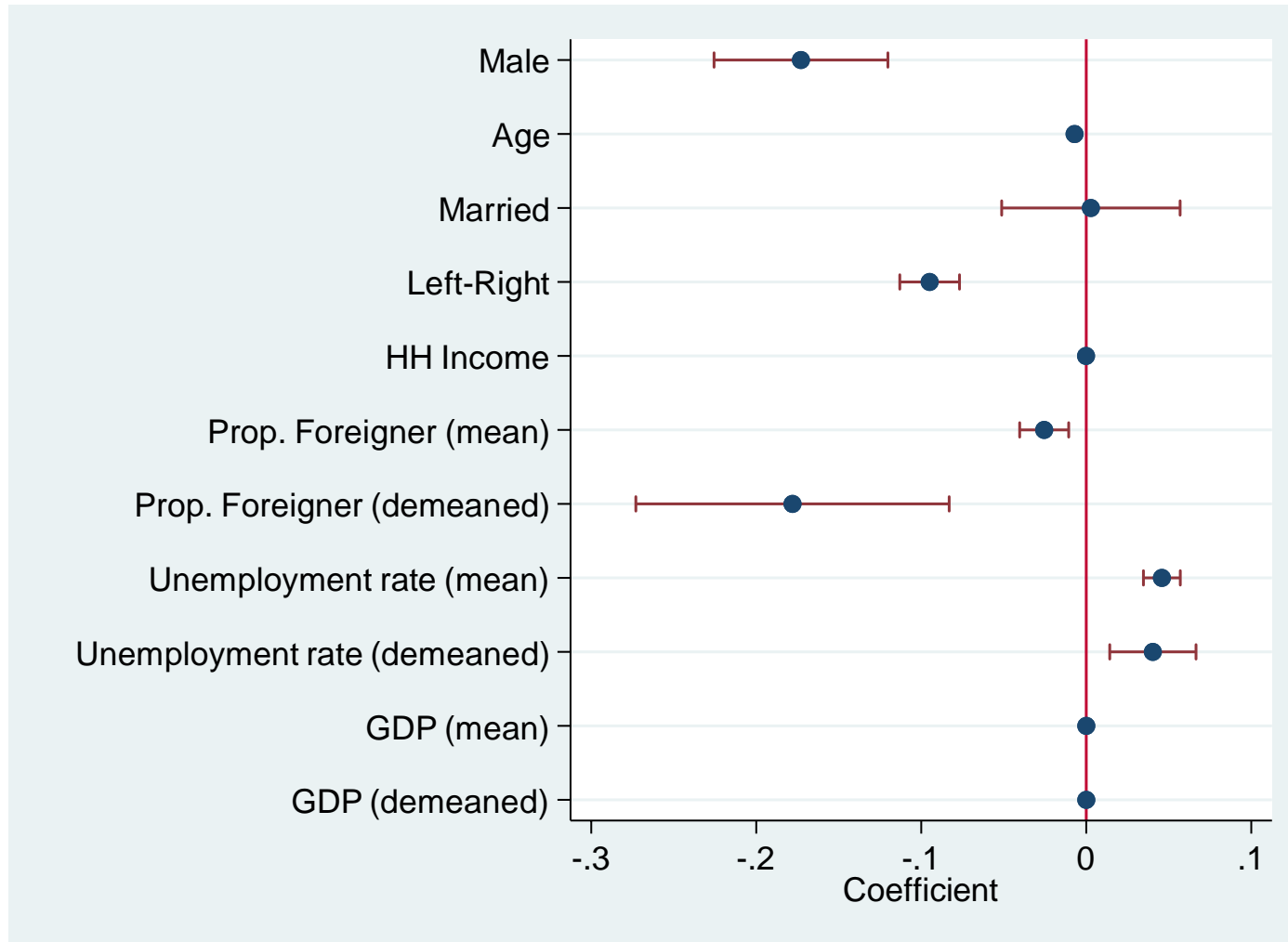
- „Allbus-Panel“: Observations of 94 RORS in 4 waves
- Ordered probit with three levels as DV
- Multilevel model: „Hybrid-Model“ (Fairbrother 2014)
 - 3 Levels: Respondents i , ROR at time t , ROR j
 - DV is regressed on individual level variables X and between- and within variation of context level variables Z

$$P(Y < k | X, Z) = \beta X_{jit} + \underbrace{\gamma^{\text{WE}} (Z_{jt} - \bar{Z}_j)}_{\text{within-effect: Deviation from mean of each spatial unit}} + \underbrace{\gamma^{\text{BE}} \bar{Z}_j}_{\text{between-effect: Mean of each spatial unit}} + u_j + u_{jt}$$

within-effect:
Deviation from
mean of each
spatial unit

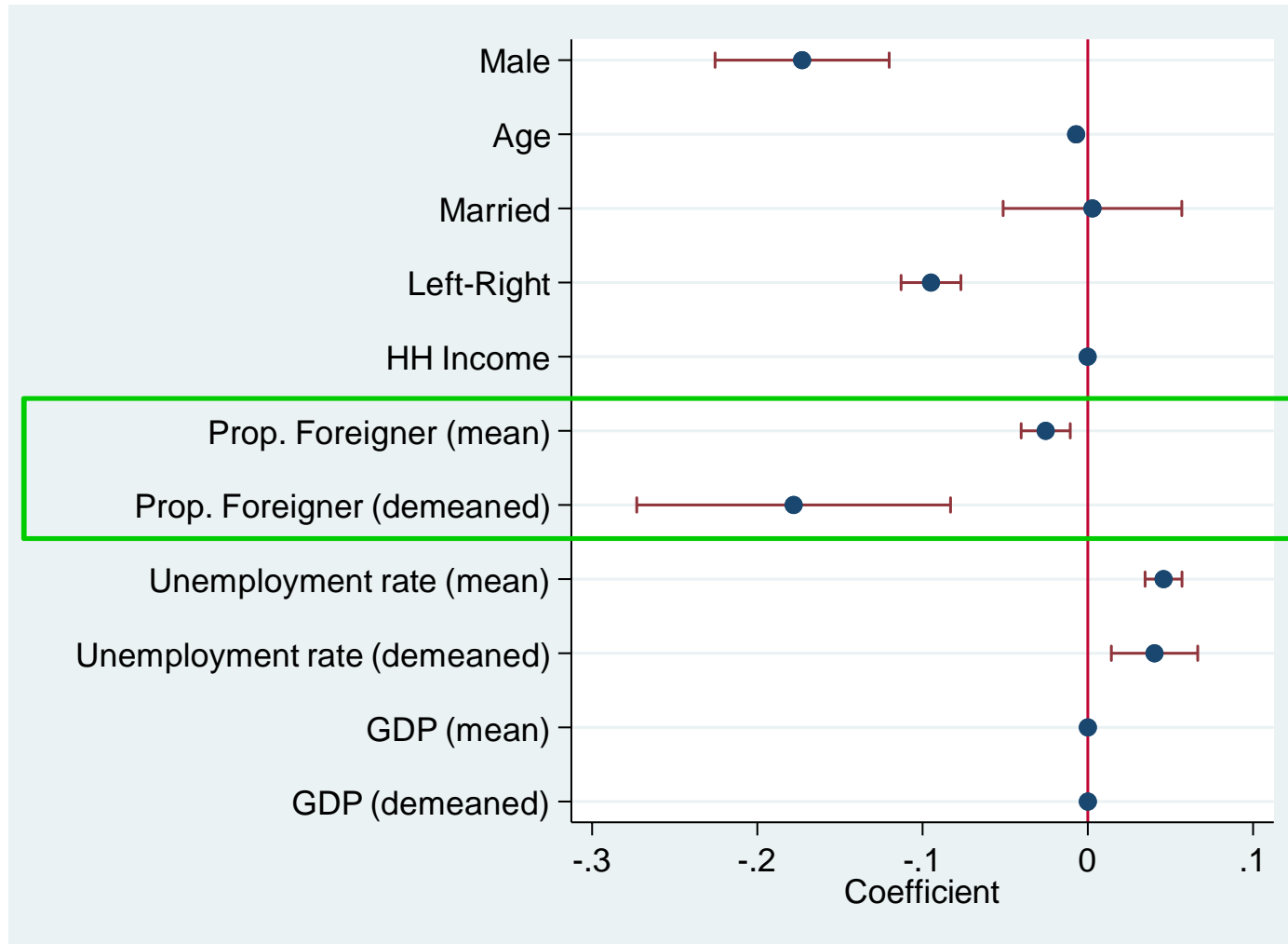
between-effect:
Mean of each
spatial unit

Results: Support of Social Benefits



Source: Replication of original model, additional controls for year, education, employment status, community size, East Germany

Results: Support of Social Benefits



Source: Replication of original model, additional controls for year, education, employment status, community size, East Germany

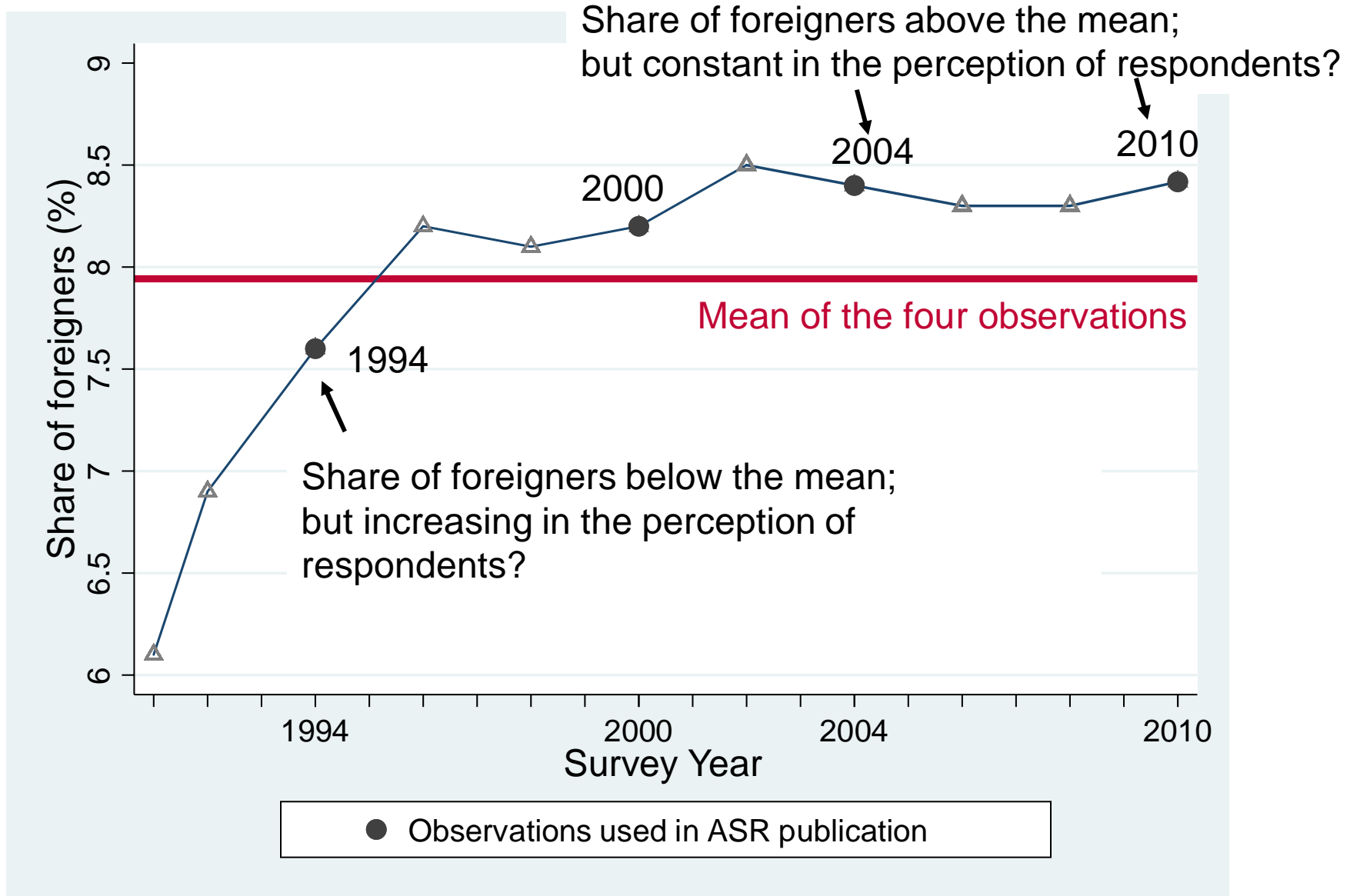
Convincing (Causal) Evidence?

1. No direct modelling of the threat through rising influx
 - No consideration of temporal order (inflow → reaction)
 - Gaps of up to six years, thus danger of unobserved heterogeneity (e.g. changing composition of population)
2. No consideration of spatial heterogeneity, spatial structure
3. Strongest effects on the level of Federal States point to spatial autocorrelations (influential neighboring units)
4. Strong effects despite small variance „within“
 - The difference to \bar{Z}_j is 1,3 percentage points max for 99% of the RORs
 - Very long causal chain

Strategy of Analysis for our Replication

1. Consideration of temporal order (*Change* of proportion of foreigners → Effect)
2. Analysis of spatial heterogeneity
3. Inclusion of spatial neighbors (spatial lags)
- [4. Analysis of mediators, e.g. „Support of Social Benefits for Foreigners“]

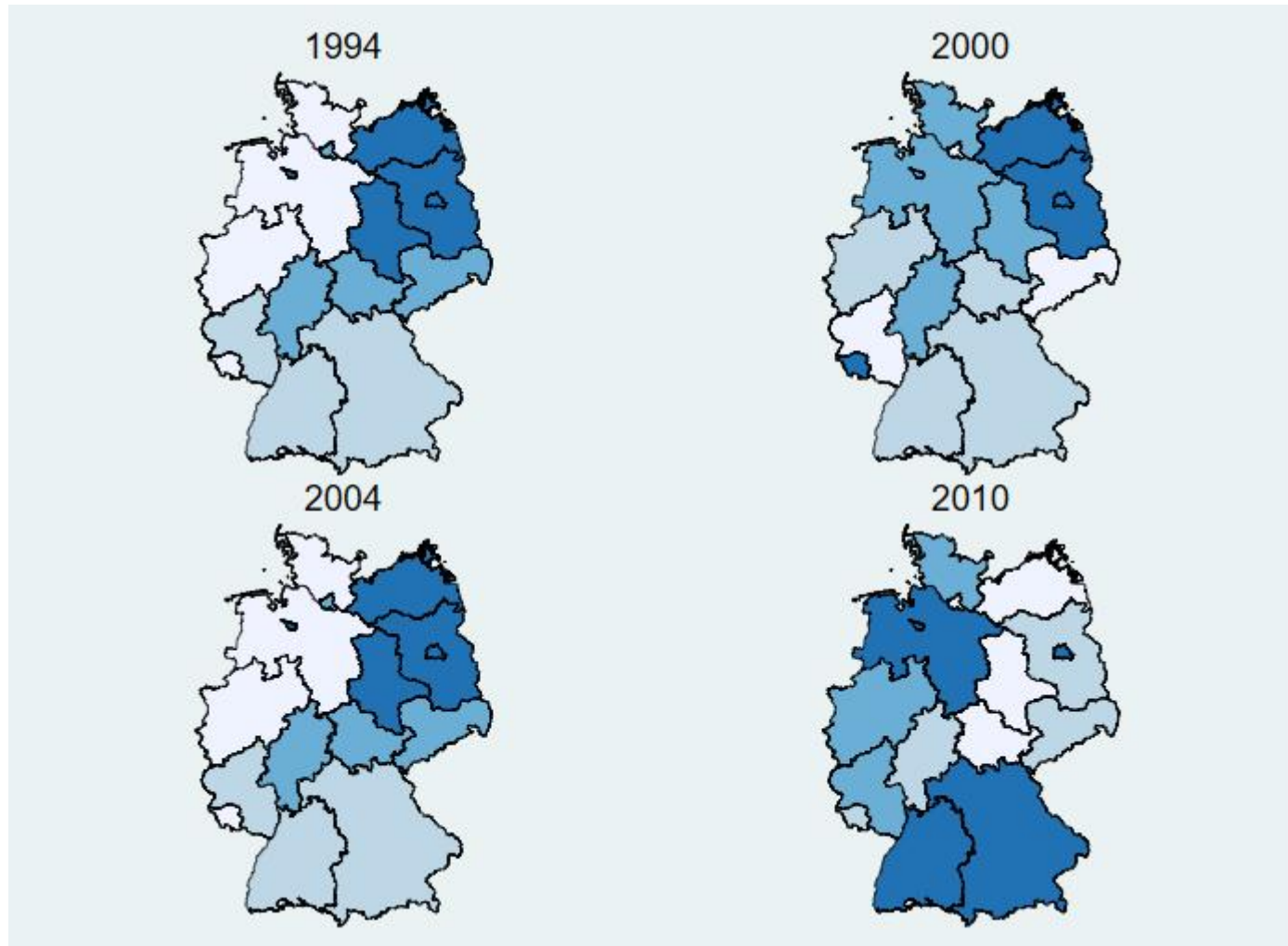
1. Time Order (Example: Saarland)



1. Time Order

- In general: within-variation is not correlated with changes between years, theoretically more relevant ($r = -0.2$; $p = 0.2$)
- Hybrid-Models (also FE Models) do not consider the temporal order
- Replications with different specification of lagged share of foreigners (Z_{jt-1}) show no statistically significant correlations
→ Effects of the Hybrid-Models seem „suspicious“.

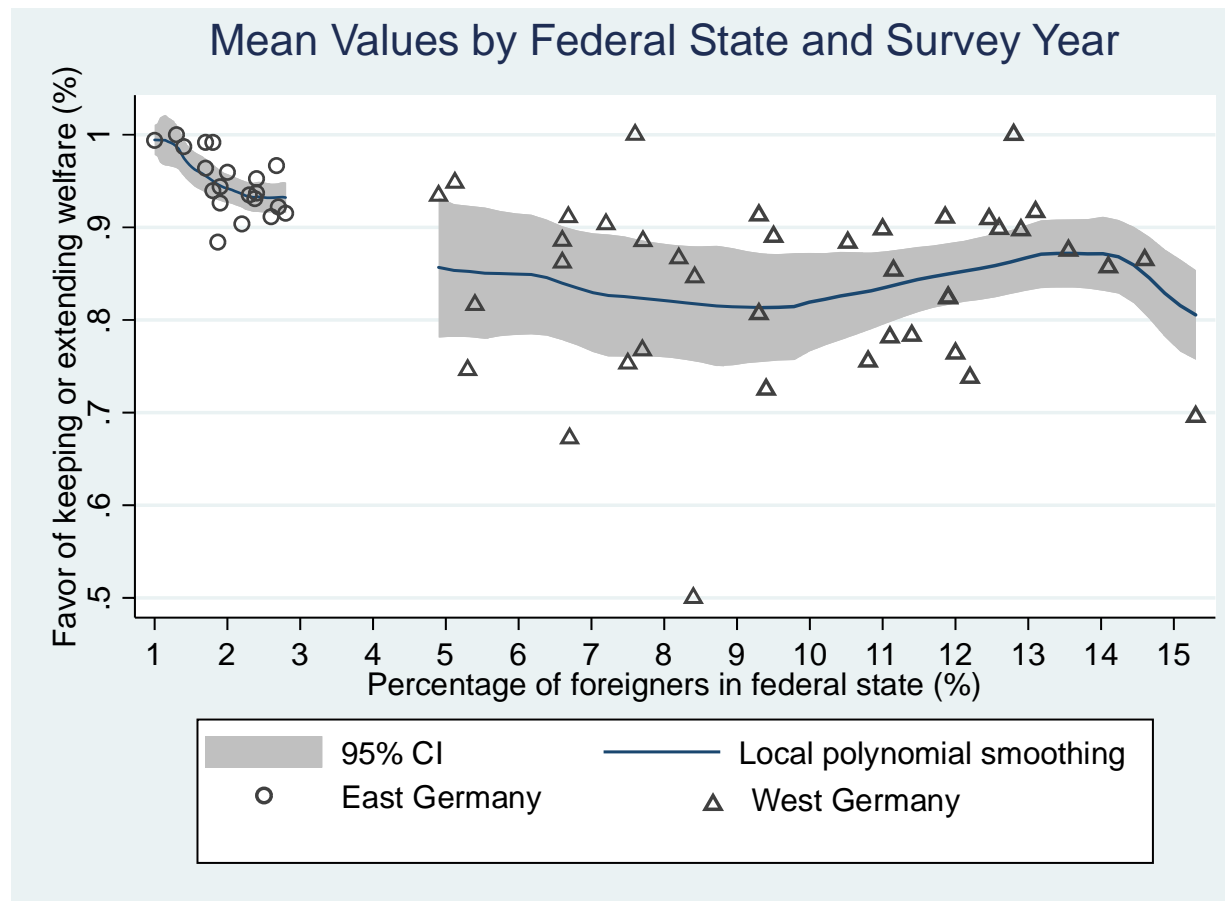
2. Spatial Heterogeneity



Note: Aggregated residuals, linear model without control for East/West

2. Spatial Heterogeneity: East ≠ West

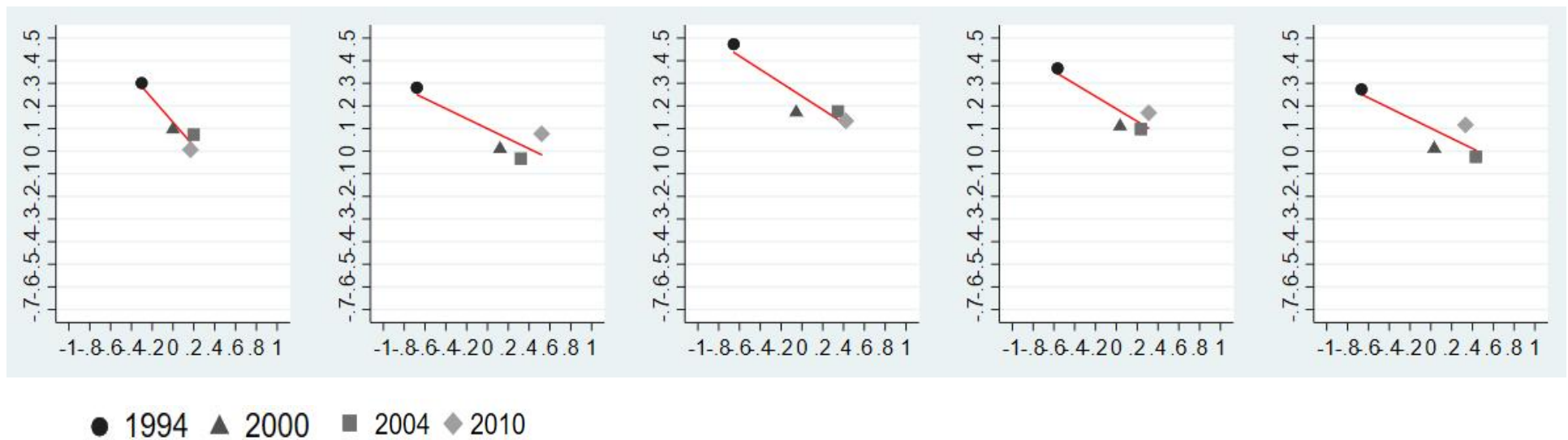
Mean support for social benefits of 94 RORs and share of foreigners



2. Spatial Heterogeneity: East and West

Mean support for social benefits (adjusted for individual characteristics) by within-variation of share of foreigners

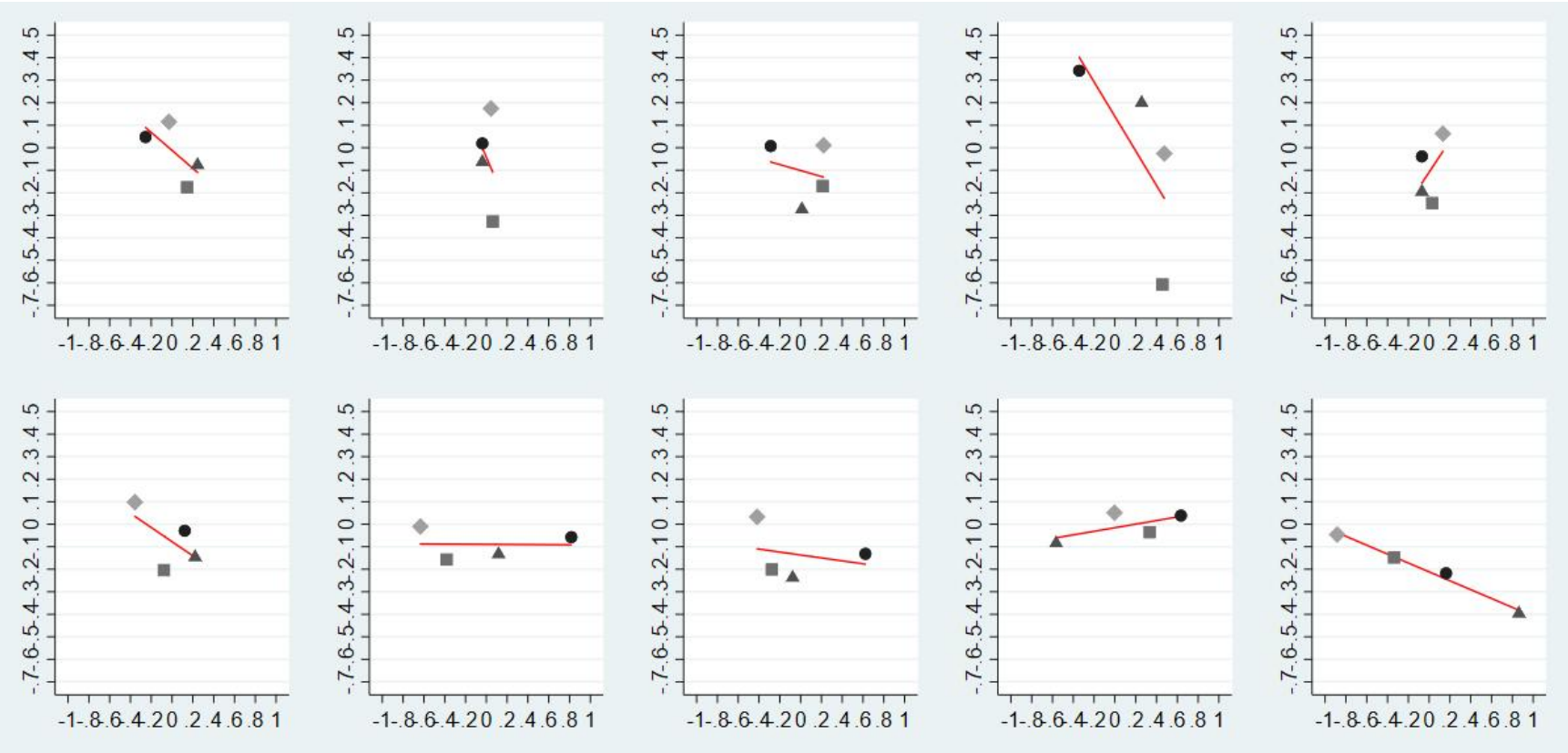
1) East-Germany



→ High support for social benefits and comparably low share of foreigners after the Reunification of Germany

2. Spatial Heterogeneity: East and West

2) West-Germany

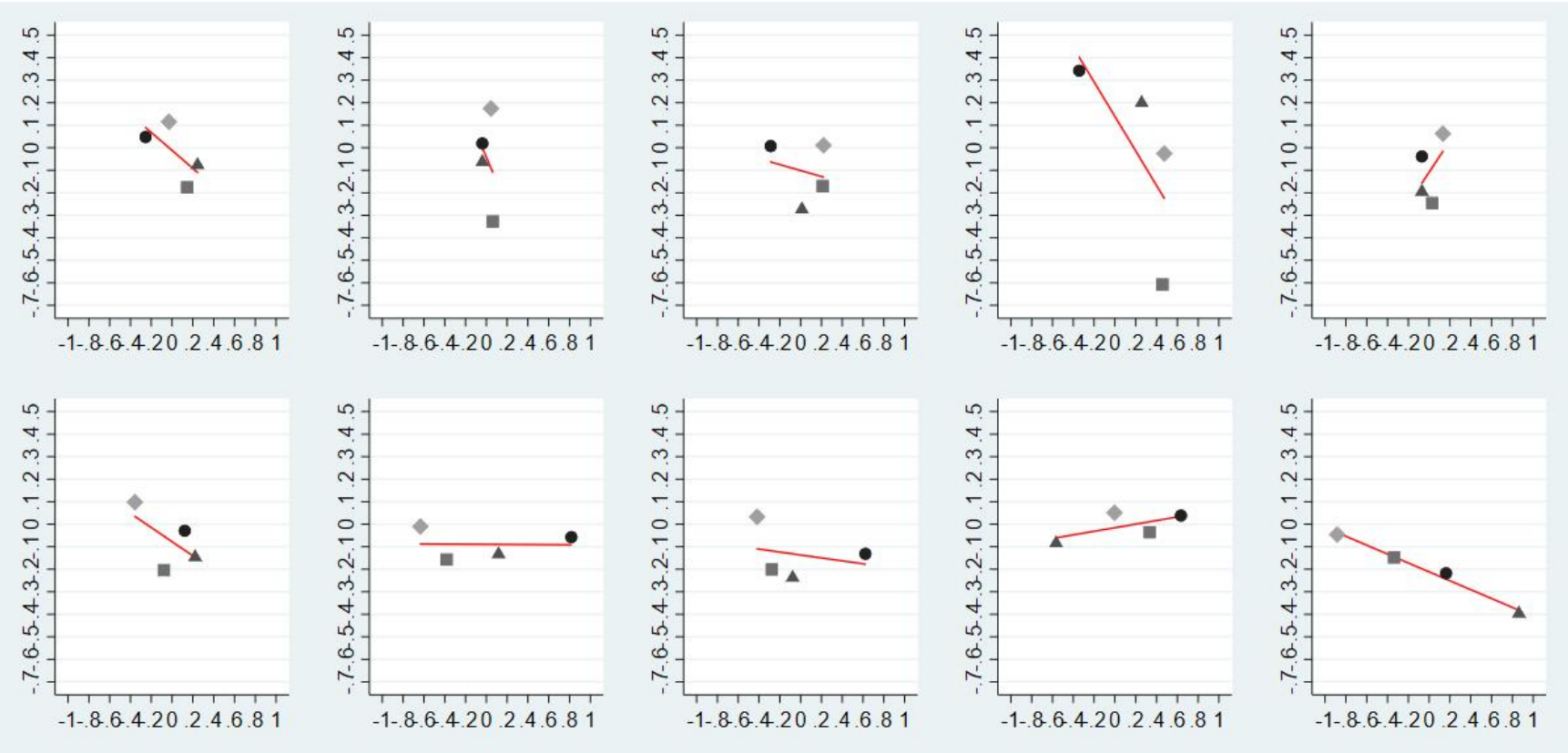


● 1994 ▲ 2000 ■ 2004 ◆ 2010

→ No clear trend.

2. Spatial Heterogeneity: East and West

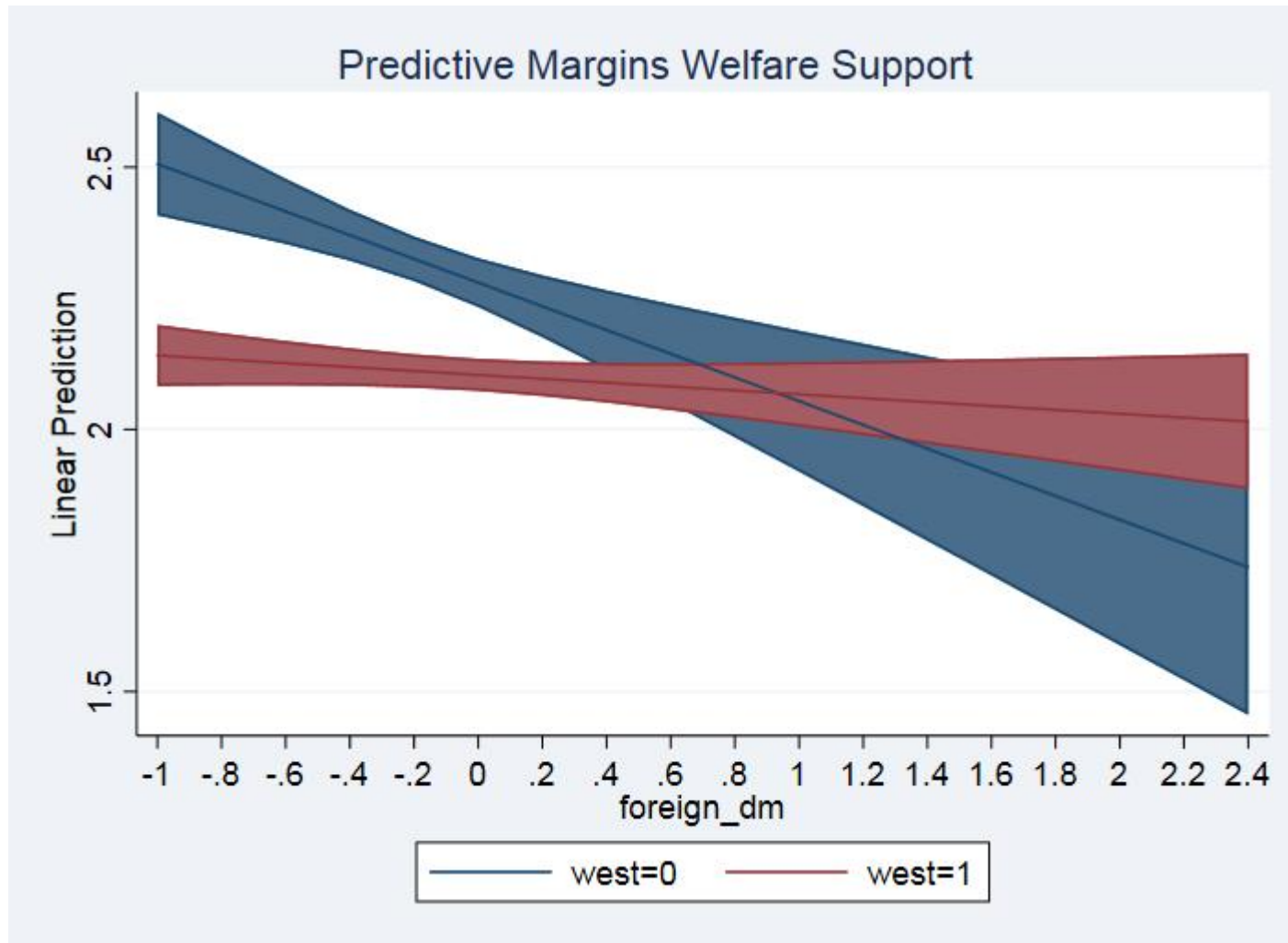
2) West-Germany



● 1994 ▲ 2000 ■ 2004 ◆ 2010

→ *Separate* Time Trends in East Germany!

2. Spatial Heterogeneity

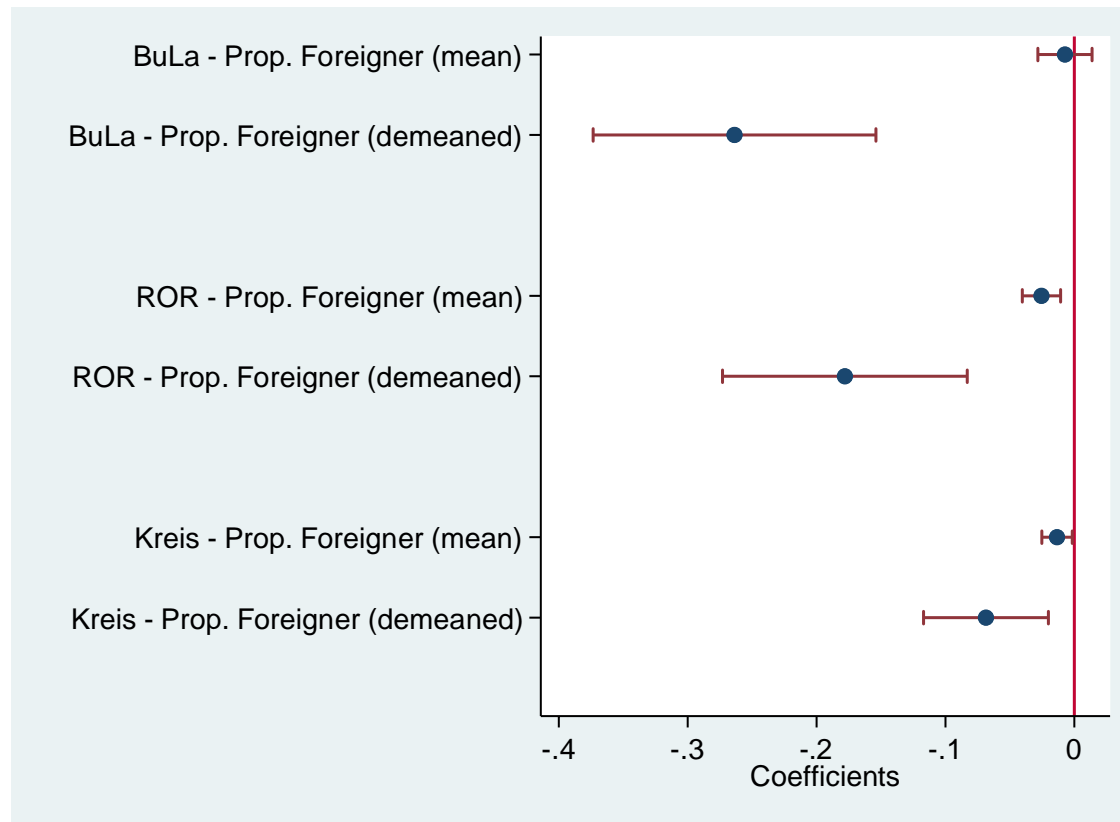


3. Modeling of Spatial Neighbors

→ The influence should be strongest on the level of aggregation of the mechanism.

3. Modeling of Spatial Neighbors

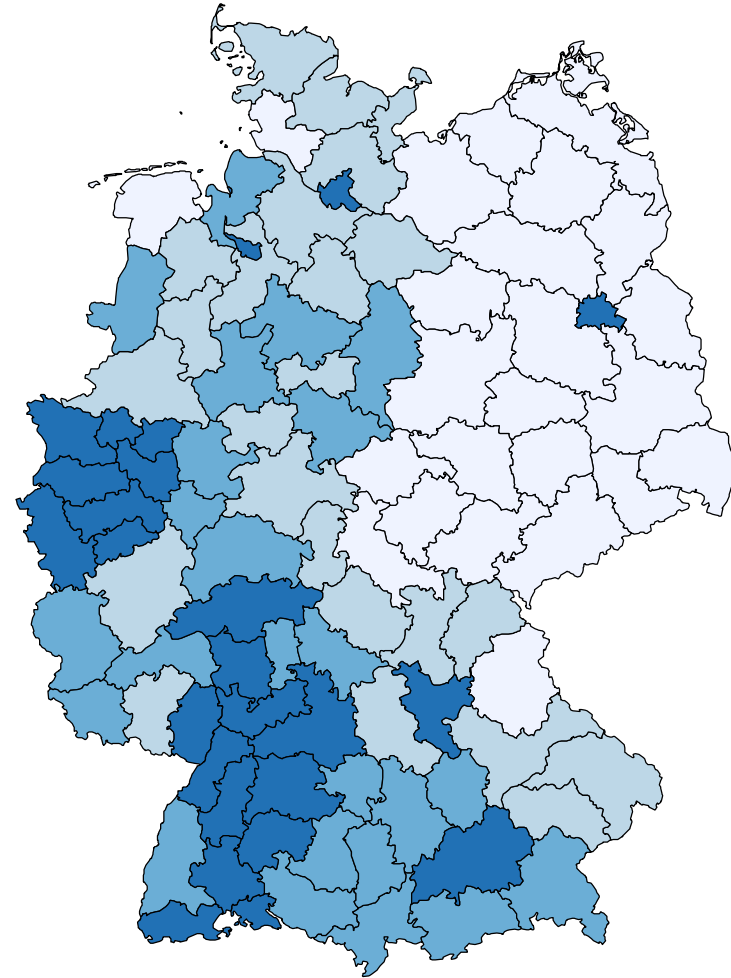
→ The influence should be strongest on the level of aggregation of the mechanism.



3. Modeling of Spatial Neighbors

Calculation of spatial lags

→ With appropriate choice of spatial units, we should find no effects.



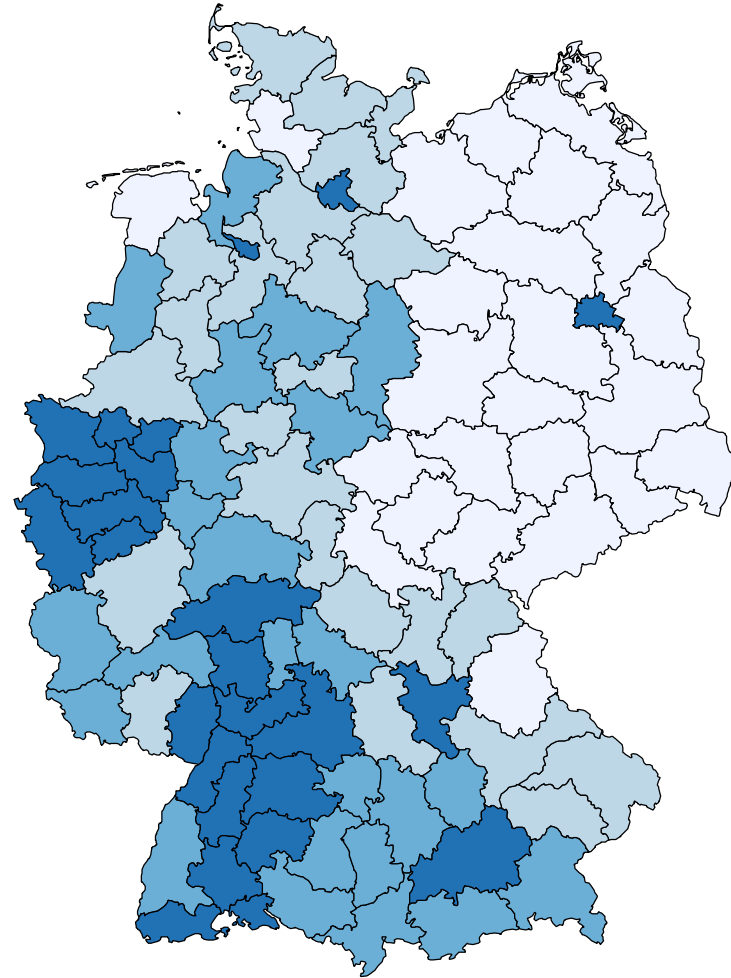
3. Modeling of Spatial Neighbors

Calculation of spatial lags

→ With appropriate choice of spatial units, we should find no effects.

But: Here we have two separate trends; only in East Germany.

→ Implications for effects of spatial lags. Spatial units within East Germany should cluster, whereas spatial units between East and West should show no effect.



Conclusions

1. What do we learn about the connection between share of foreigners and support for social benefits?
 - Instead of a direct causal link, spurious correlation caused by regional trends. Process of adaption („good by Lenin“) in Eastern Germany after the reunification
2. What can we learn about the analysis of context effects on attitudes?
 - Treat East- and West-Germany as separate spatial units/countries?
 - „within“-estimation is not always appropriate
3. What can we learn about the analysis of spatial effects?
 - Spatial analysis of residuals for the exploration of spatial processes
 - Different theoretical arguments imply different levels of aggregation
 - Use spatial neighboring units (spatial lags) to test implications of theoretical arguments.

Vielen Dank für die Aufmerksamkeit!

Katrin.Auspurg@Imu.de

Thomas.Woehler@uni-konstanz.de