The Geocoded German Longitudinal Election Study: Analyzing Place Based Effects on the 2021 German Federal Election

Meet the Experts – GESIS online talks

Anne-Kathrin Stroppe, 12th January 2023
Anne-Kathrin Stroppe

- Doctoral Researcher
- Data Curator for the German Longitudinal Election Study (GLES)
- Research interests: Electoral geography, (political) attitudes
- Contact: anne-kathrin.stroppe@gesis.org
Place-based explanations in the social sciences

• Long tradition of place-based explanations (Siegfried 1913, Allport 1954)

• Mechanisms at work:
  • interactions and events occur in place (Johnston/Pattie 2017)
  • identities, social capital and culture can be tied to places (Agnew 2002)
  • local context can serve as heuristics for evaluations and decision-making (Cho/Rudolph 2008)
Place-based explanations in the social sciences

• Long tradition of place-based explanations (Siegfried 1913, Allport 1954)

• Mechanisms at work:
  • interactions and events occur in place (Johnston/Pattie 2017)
  • identities, social capital and culture can be tied to places (Agnew 2002)
  • local context can serve as heuristics for evaluations and decision-making (Cho/Rudolph 2008)

• Challenges when relying on aggregate data:
  • Ecological fallacy
  • Research questions limited by data availability (e.g. attitudes)

What are geocoded survey data?

• Survey data with direct and indirect spatial references
• Sometimes already enhanced with spatial data attributes
• General goal: analyze interactions between individual behaviors and attitudes and the environment
• Data access usually restricted due to data protection regulations
What are geocoded survey data?

• Survey data with direct and indirect spatial references
• Sometimes already enhanced with spatial data attributes
• General goal: analyze interactions between individual behaviors and attitudes and the environment
• Data access usually restricted due to data protection regulations
Design Overview 2021 and Respondents' Location

Cross-Section (CAWI + PAPI)

Pre-Election: Aug.-Sept. 21
Post-Election: Sept.-Nov. 21

Rolling Cross-Section (CATI)

Pre-Election: July-Sept. 21
Post-Election: Sept.-Oct. 21

Daily Samples

Panel (CAWI)

Wave 15: Feb.-Mar. 21
Wave 16: May 21
Wave 17: July 21
Wave 18: Aug. 21
Wave 19: Sept. 21
Wave 20: Oct. 21
Wave 21

Tracking (CAWI)

T48: Jan. 21
T49: April 21
T50: Sept. 21

Nomination Study (Multi-Source Research)

Pre-Election (open)

Candidate Study (CAWI + PAPI)

Pre-Election: Oct. 21 – Jan. 22

German Federal Election 26.09.2021
**Design Overview 2021 and Respondents’ Location**

**GLES**

**German Longitudinal Election Study**

**Cross-Section (CAWI + PAPI)**
- Pre-Election: Aug.-Sept. 21
- Post-Election: Sept.-Nov. 21

**Rolling Cross-Section (CATI)**
- Pre-Election: July-Sept. 21
- Post-Election: Sept.-Oct. 21

**Panel (CAWI)**
- Wave 15: Feb.-Mar. 21
- Wave 16: May 21
- Wave 17: July 21
- Wave 18: Aug. 21
- Wave 19: Sept. 21
- Wave 20: Oct. 21

**Tracking (CAWI)**
- T48: Jan. 21
- T49: April 21

**German Federal Election 26.09.2021**

**Smallest regional unit to identify respondents’ location**

- Addresses: (Register Sample)
- ZIP-Codes/Municipalities: (Self-Reported)
- ZIP-Codes: (Self-Reported)
Respondents‘ Location: Sources of Uncertainty & Errors

Self-Reporting

Matching Administrative Units in Germany

Relocations and Changes over Time
Respondents’ Location: Sources of Uncertainty & Errors

Self-Reporting

- ZIP codes in CAWI surveys: ~1 – 1.5% not existent
- No control mechanism whether it is correct
- Reporting difference between gender, education level and age might appear (Gladden et al. 1997)
- Inclusion of self-report measures on characteristics of living place (housing, urbanization etc.) but no linking (Brinkerhoff et al. 2021)

Matching Administrative Units in Germany

Relocations and Changes over Time
Respondents’ Location: Sources of Uncertainty & Errors

Smallest regional unit to identify respondents’ location

- Addresses (Register Sample)
- ZIP-Codes/ Municipalities (Self-Reported)
- ZIP-Codes (Self-Reported)

Self-Reporting

Matching Administrative Units in Germany

Relocations and Changes over Time
Respondents’ Location: Sources of Uncertainty & Errors

Self-Reporting

Matching Administrative Units in Germany

Relocations and Changes over Time

- Respondents move between sampling, survey and survey waves
- Redistricting over time
- Match of survey year and context information
Benefits of Working with Address Data

- No self-report bias
- Small spatial scale
- Once geocoded, complete tool box of spatial linking methods available:
  - 1:1 linking to different administrative units across time
  - Definition of own spatial zones (filter techniques, buffer zones, neighboring units, ..)
  - Distance measures

Remaining Challenge:
Availability of geospatial information
GLES Design Overview 2021

Cross-Section (CAWI + PAPI)
- Pre-Election: Aug.-Sept. 21
- Post-Election: Sept.-Nov. 21

Rolling Cross-Section (CATI)
- Pre-Election: July-Sept. 21 (Daily Samples)
- Post-Election: Sept.-Oct. 21

Panel (CAWI)
- Wave 15: Feb.-Mar. 21
- Wave 16: May 21
- Wave 17: July 21
- Wave 18: Aug. 21
- Wave 19: Sept. 21
- Wave 20: Oct. 21
- Wave 21: Nov. 21

Tracking (CAWI)
- T48: Jan. 21
- T49: April 21
- T50: Sept. 21

Addresses (Register Sample)

Municipalities/Constituencies (Self-Reported)

ZIP-Codes (Self-Reported)

German Federal Election 26.09.2021

Smallest regional unit to identify respondents' location
Sensitive Regional Data: Creating the Dataset

- Data Preprocessing
- Geocoding
- Spatial Joins
- Data Wrangling
- Data Linking


- R used for nearly all steps (GIS tool, data wrangling...
Sensitive Regional Data: Creating the Dataset

- Three main data sets saved in separate locations:
  - survey data with attributes
  - address data
  - correspondence table to match data sets

- Cleaning of address data in table format (e.g. removing additional information)
Sensitive Regional Data: Creating the Dataset

- Conversion of indirect spatial reference (e.g. addresses) into direct spatial reference (e.g. geocoordinates)

- Several providers but again: data protection issues and quality concerns

- Federal Agency of Cartography and Geodesy (BKG):
  - Online interface and API for online geocoding
  - Offline geocoding possible based on raw data
  - Data and service are restricted

- R package **bkggeocoder** developed at GESIS for (offline) geocoding:
  13.04.2023, Dr. Stefan Jünger, bkggeocoder: a geocoding tool for survey data (in English)
  [https://github.com/StefanJuenger/bkggeocoder](https://github.com/StefanJuenger/bkggeocoder)
Sensitive Regional Data: Creating the Dataset

- Geocoding retrieves: point coordinates, administrative unit keys, INSPIRE* id
- Spatial joins based on coordinates for:
  - constituencies
  - administrative units across time (for GLES: 31.12.2015 as harmonized territorial status)

*defined and harmonised grid net for Pan-Europe with standardised location and size of grid cells. Examples of cell sizes could be 10x10 m, 100x100 m, 1x1 km ([https://inspire.ec.europa.eu/theme/gg](https://inspire.ec.europa.eu/theme/gg))
Sensitive Regional Data: Creating the Dataset

- Missing coding, additional variables, data checks, and cumulation
- Deletion of addresses and point coordinates

Figure: Information on Geocoding Results of GLES Cross-Section 2021, Pre- and Post-Election

<table>
<thead>
<tr>
<th>Code</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geocoded reaching target quality</td>
<td>96.31</td>
</tr>
<tr>
<td>Geocoded after manual check</td>
<td>2.05</td>
</tr>
<tr>
<td>Geocoded after manual correction</td>
<td>0.14</td>
</tr>
<tr>
<td>Geocoding not possible</td>
<td>0.60</td>
</tr>
<tr>
<td>Address deleted by respondent</td>
<td>0.90</td>
</tr>
</tbody>
</table>
## Sensitive Regional Data: Creating the Dataset

<table>
<thead>
<tr>
<th>id</th>
<th>Admin. Unit: Sampling Year</th>
<th>Admin. Unit: Sampling Year</th>
<th>Admin. Unit: Harmonized Year</th>
<th>Admin. unit: Harmonized Year</th>
<th>GRID Cell</th>
<th>Context Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>05315000</td>
<td>05315</td>
<td>13 'Köln 1'</td>
<td>05315000</td>
<td>05315</td>
<td>1kmN2684E4334</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

**Data Preprocessing**

- Geocoding
- Spatial Joins
- Data Wrangling
- Data Linking
# Sensitive Regional Data: Creating the Dataset

## Data Preprocessing

1. Ifdn
   - municipality
   - district
   - constituency
   - INSPIRE ID 1km
   - Mun. Size

2. Data Linking
   - Party Vote
   - Dem Satisfaction
   - Age
   - Gender

## Geocoding

## Spatial Joins

## Data Wrangling
## Sensitive Regional Data: Creating the Dataset

<table>
<thead>
<tr>
<th>id</th>
<th>Admin. Unit: Sampling Year</th>
<th>Admin. Unit: Sampling Year</th>
<th>Admin. Unit: Sampling Year</th>
<th>Admin. Unit: Harmonized Year</th>
<th>Admin. Unit: Harmonized Year</th>
<th>GRID Cell</th>
<th>Context Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>05315000</td>
<td>05315</td>
<td>13 ,Köln 1'</td>
<td>05315000</td>
<td>05315</td>
<td>1kmN2684E4334</td>
<td>7 '500,000 and more'</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

### Data Preprocessing
- Geocoding
- Spatial Joins
- Data Wrangling
- Data Linking

### Grid Cell Context Data
- inspid_1km
- Population: 18199

### Data Linking
- municipality
- share_unemployed_2021
- 05315000: 0.093
- ...
Sensitive Regional Data: Data access

• On-site Access to sensitive, restricted-access GESIS data via the GESIS Safe Room

In short:
• Survey data linked with requested identifier (principle of data parsimony)
• Researchers have the possibility to link their own geospatial data
• Consulting and completion of data use agreement via gles@gesis.org
A Geographic Perspective on the 2021 German Federal Election

Figure: Share of votes in the 2021 federal election by district

Research Question & Framework

Does living in a left-behind place affect voting for the AfD in the 2021 German Federal Election?

Framework:

- "the geography of discontent" and the study of "left-behind places" (Broz et al 2021; Cramer 2016, Gordon et al 2018; Harteveld et al 2021; McCann 2020; Rodríguez-Pose 2018)

In Germany:

- some evidence for spatially polarized voting behaviour in 2021 (Haffert (2022), Klärner and Osigus (2021), Träger (2022) Stroppe and Jungmann (2022))
- in previous years mixed findings for the relation of AfD success and local (economic) disadvantages in Germany (Deppisch, Osigus, and Klärner 2021; Diermeier 2020, Förtner, Belina, & Naumann, 2020; Kurtenbach, 2019)
Methodological Approach: Linking Geocoded Survey Data

• Individual-level survey data:
  • GLES 2021 Cross-Section, Pre- and Post-Election

• Geospatial data to operationalize „left-behind places“:
  • Economic Situation
  • Demographic Composition
  • Access to Public Services
Methodological Approach: Places Being Left Behind

1km*1km Grid Level Data (Microm 2021):

- **Economic Situation:** Share of households with an average income less than 60% of the German median income (risk of poverty)
- **Demographic Composition:** Share of over 60 year olds
Methodological Approach: Places Being Left Behind

Distance Measure (BKG 2021, DB 2021):

• **Access to Public Services:** Mean distance from grid centroid to the closest general practitioner, train station, supermarket, post office, and pharmacy
Preliminary Results: AfD Vote

Preliminary Results: Perceived Societal Recognition

Conclusion

• Geocoded surveys allow to look into the mechanism at work
• Flexibility to operationalize place (characteristics)

• Still many challenges:
  • Definition of relevant neighbourhood(s) and indicators
  • Availability of geospatial information on small spatial scale
  • Self-selection biases (in surveys and place)
  • Methodological challenges (modifiable area unit problem, spatial autocorrelation…)

• Growing availability and curation of geocoded survey data
Expert contact & GESIS consulting

Contact: you can reach the speaker via personal e-mail:
anne-kathrin.stroppe@gesis.org

you can reach the GLES team via e-mail:
gles@gesis.org

GESIS Consulting: GESIS offers individual consulting in a number of areas – including survey design & methodology, data archiving, digital behavioral data & computational social science – and across the research data cycle.

Please visit our website www.gesis.org for more detailed information on available services and terms.
Upcoming talks

• Please visit our meet-the-experts website:

  09.02.2023, Dr. Boris Heizmann, Meet the Eurobarometer (in Englisch)
  09.03.2023, Dr. Sonja Schulz, Meet the ALLBUS cumulation (in Deutsch)
  13.04.2023, Dr. Stefan Jünger, bkggeocoder: a geocoding tool for survey data (in Englisch)
Thank you for participating!