

What makes region attractive for worker?

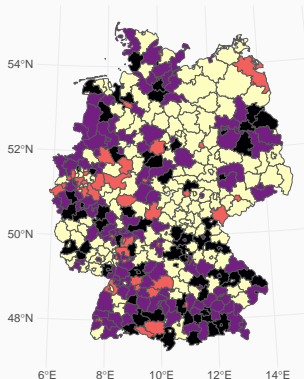
Yuga Hikida

yuga.hikida@tu-dortmund.de

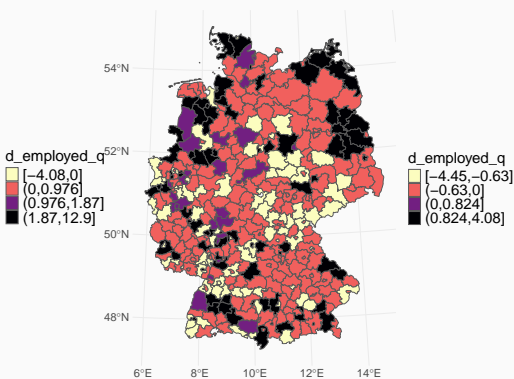
- Determinant of migration of worker (Moretti 2011)
 - Labour market condition
 - net wage, unemployment rate...
 - Amenities
 - public goods, facilities...
- **Empirically analyse what makes region attractive for worker.**
- Previous studies: migration decision from various factors, both of the channels (Buch et al. 2014; Hakim et al. 2022).
 - No single strong determinants of migration.

Data

- Focus on % changes in # employees in each region over time.
- $y_t := 100 \times \{\log(\text{employed}_t) - \log(\text{employed}_{t-1})\}$
- Period: 2016 - 2021
- All data is obtained from <https://www.inkar.de>



(a) 2016



(a) 2021

- `splm` package (Millo and Piras 2009) is used for all the estimation.
- Maximum likelihood method is used.
- Statistical testing: 5% significance level.
- Code available at <https://github.com/yugahikida/QRE>.

Explanatory variables

Variable name	description
GDP per employed	GDP in €1,000 per employed person
Unemployment rate	Share of unemployed people in %
Land price	Average purchase values for building land in € per m ²
Proportion of foreigners	Proportion of foreigners to residents in %
Amount of waste	Amount of waste disposed of or treated per inhabitant in kg
Hospital beds	Hospital beds per 1,000 inhabitants

- Some variables are transformed such that scale of all the variables are roughly same.
- Denoted as X in the following slides.

Basic panel data model

For region $i = 1, \dots, N$, time $t = 1, \dots, T$:

$$y_{it} = \alpha + \beta^T X_{it} + u_{it}$$

$$u_{it} = \mu_i + \nu_{it}$$

$$\mu_i \stackrel{\text{iid}}{\sim} (0, \sigma_\mu)$$

$$\nu_{it} \stackrel{\text{iid}}{\sim} (0, \sigma_\nu)$$

where $X_{it} = [x_{it1}, \dots, x_{itK}]^T$ and $\beta = [\beta_1, \dots, \beta_K]^T$.

μ_i is a group specific effect (“random effect”) differs across region which is modelled independently constant over time.

Result



- Counter-intuitive result from economic theory
- Potential model misspecification?

Adding autoregressive error term

$$y_{it} = \alpha + \beta^T X_{it} + u_{it}$$

$$u_{it} = \mu_i + \lambda \sum_{j=1, j \neq i}^N w_{ij} u_{jt} + \nu_{it}$$

$$\mu_i \stackrel{\text{iid}}{\sim} (0, \sigma_\mu)$$

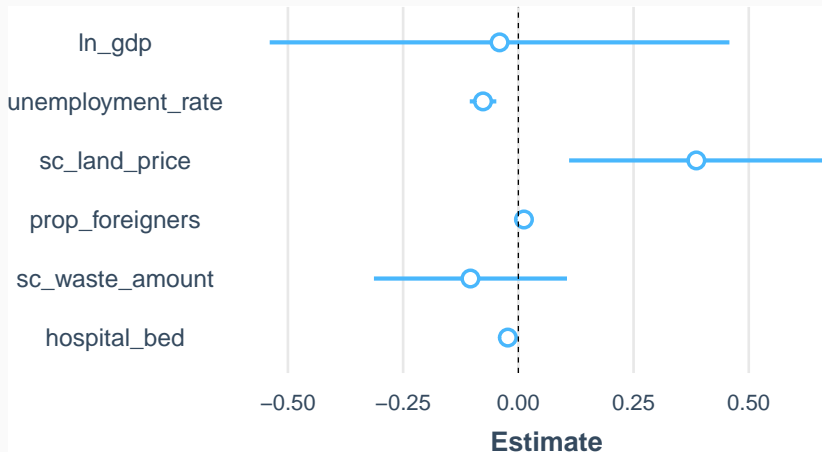
$$\nu_{it} \stackrel{\text{iid}}{\sim} (0, \sigma_\nu)$$

where $|\lambda| < 1$ is spatial autoregressive parameters and w_{ij} is the i, j element of $N \times N$ spatial weight matrix.

Now error term is decomposed into three parts: group specific part, autoregressive part and idiosyncratic part.

Result: Autoregressive error model

- LM test under the null hypothesis $\lambda = 0$ suggests existence of spatial autocorrelation \rightarrow estimate model with autoregressive error.



Spatial lag of the dependent variable

$$y_{it} = \alpha + \rho \sum_{j=1, j \neq i}^N w_{ij} y_{jt} + \beta^T X_{it} + u_{it}$$

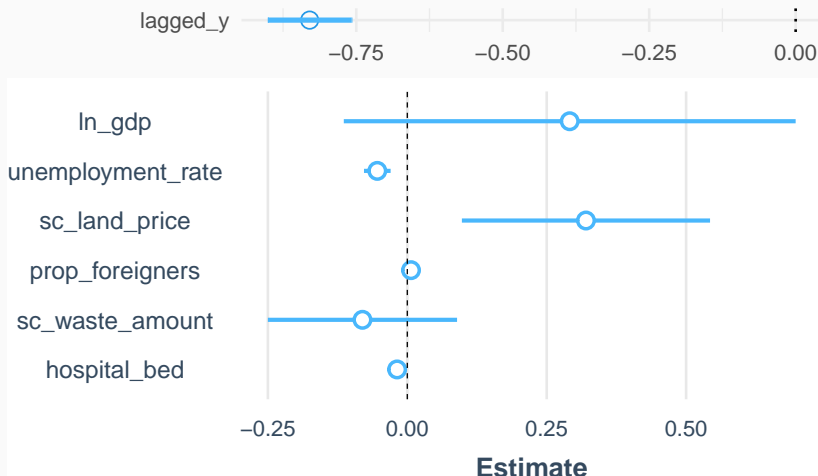
$$u_{it} = \mu_i + \lambda \sum_{j=1, j \neq i}^N w_{ij} u_{jt} + \nu_{it}$$

$$\mu_i \stackrel{\text{iid}}{\sim} (0, \sigma_\mu)$$

$$\nu_{it} \stackrel{\text{iid}}{\sim} (0, \sigma_\nu)$$

- $|\rho| < 1$ is the effect of spatially lagged dependent variable.
- Consider so called direct effect and indirect effect (explained later).

Result: Spatial lag model + spatially correlated error



- Difficult to interpret directly (next slide)
- Statistical significance of lagged term implies potential model misspecification of other two models.

Direct and indirect effect

Focus on single time. Matrix form of the model is given by

$$y = \alpha + \rho W y + X\beta + u$$

$$y = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1}(\alpha + u)$$

$$\tilde{y} := (I - \rho W)^{-1} X\beta$$

$$= \sum_{k=1}^K (I - \rho W)^{-1} I \beta_k X_k := \sum_{k=1}^K S_k(W) X_k$$

where $X_k = [x_{1k}, \dots, x_{Nk}]^T$.

- Direct effect = impact of x_{ik} on y_i
 - Diagonal element of $S_k(W)$.
- Indirect effect = impact of $x_{jk}, j \neq i$ on y_i
 - Off-diagonal element of $S_k(W)$. “Spatial spillover”.
- Average over region / time to obtain summary measures (Rüttenauer 2024).

Result: Direct and indirect effect

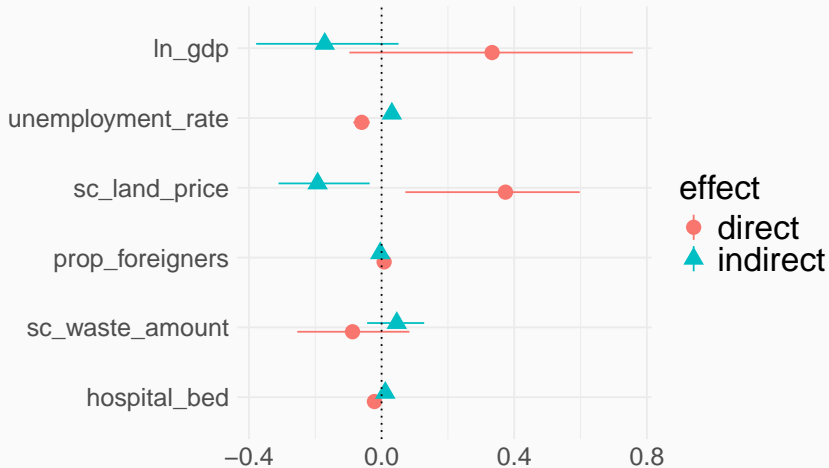


Figure 3: Direct and indirect effect

Economic interpretation

- Positive direct effect / negative spillover in GDP per capita.
 - Worker move to region with high GDP.
- Negative direct effect / positive spillover in unemployment rate.
 - Worker move out of region with high employment rate.
- Positive direct effect / negative spillover in land price.
 - People move to region with high rent?
 - Could be due to amenities associated with high rent.
 - Urban area, more amusement facilities etc.
- Negative direct effect / positive spillover in waste amount.

References

Buch, Tanja, Silke Hamann, Annekatrin Niebuhr, and Anja Rossen. 2014. "What Makes Cities Attractive? The Determinants of Urban Labour Migration in Germany." *Urban Studies* 51 (9): 1960–78.

Hakim, Arif Rahman, Nachrowi Djalal Nachrowi, Dwini Handayani, and I Dewa Gede Karma Wisana. 2022. "Do Amenities and Economic Factors Affect Migration? Empirical Evidence from Indonesian Cities." *Environment and Urbanization ASIA* 13 (1): 11–26.

Millo, Giovanni, and Gianfranco Piras. 2009. "Splm: Econometric Analysis of Spatial Panel Data." In *UseR Conference*.

Moretti, Enrico. 2011. "Local Labor Markets. Handbook of Labor Economics, Vol. 4b." Elsevier.

Rüttenauer, Tobias. 2024. "Spatial Data Analysis." *arXiv Preprint*.

Appendix: LM test

Baltagi, Song and Koh LM*-lambda conditional LM test (assuming $\sigma^2_{\mu} \geq 0$)

data: d_employed ~ ln_gdp + unemployment_rate + sc_land_price + prop_foreigner
LM*-lambda = 20.932, p-value < 2.2e-16
alternative hypothesis: Spatial autocorrelation