Higher School of Economics | Yerevan 2024

Spatial Regression Analysis of Housing Affordability

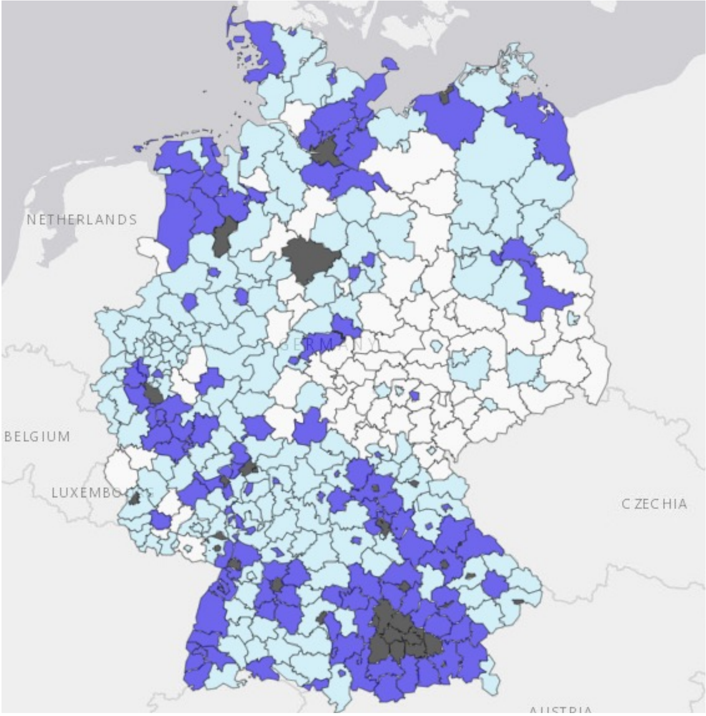
Piliuk Anastasia
Elena Semerikova

Motivation

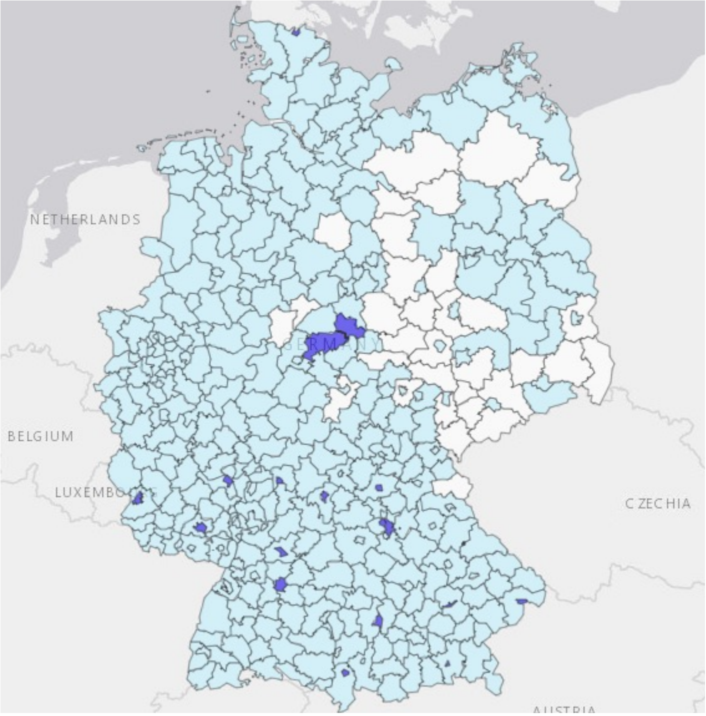
Housing affordability is an important welfare indicator

- **Rising housing prices and rents**
- **Rising house price-to-income ratio in EU countries**
- **Rising migration intensity**

Price-to-income ratio growth in Germany (2020 vs 2004)



Rent-to-income ratio growth in Germany (2020 vs 2004)



Growth rate

- < 1
- 1-1,5
- 1,5-2
- >2

Research Objective

- Housing affordability estimation and identification of its determinants in German regions

Data Sources

- RIWIS (https://www.riwis.de/intro/en-index.html)
- INKAR (https://www.inkar.de/)
- Regional statistics database (https://www.regionalstatistik.de/genesis/online)
- Federal statistics agency database (https://www-genesis.destatis.de/genesis/online)

Data

- From 2004 to 2020
- 401 German regions (NUTS-3)
- A panel of 6817 observations

Literature Review

- Most studies are focused on household level and utilize microdata. The most common determinants are:
  - Household income and expenditure
  - Current rents
  - Household size

- Studies using aggregated data leverage such features as:
  - Population and density
  - Cost of land acquisition and construction
  - Economic indicators

Housing Affordability Indexes (1)

$$HA_{index1} = \frac{Average\ annual\ income\ of\ an\ average\ household}{Average\ apartment\ price\ for\ an\ average\ household\ size} = \frac{INC_{hh}}{PP_{qm} * HH_{size} * QM_{ppc}}$$

PP_{qm} – apartment purchase price per m2

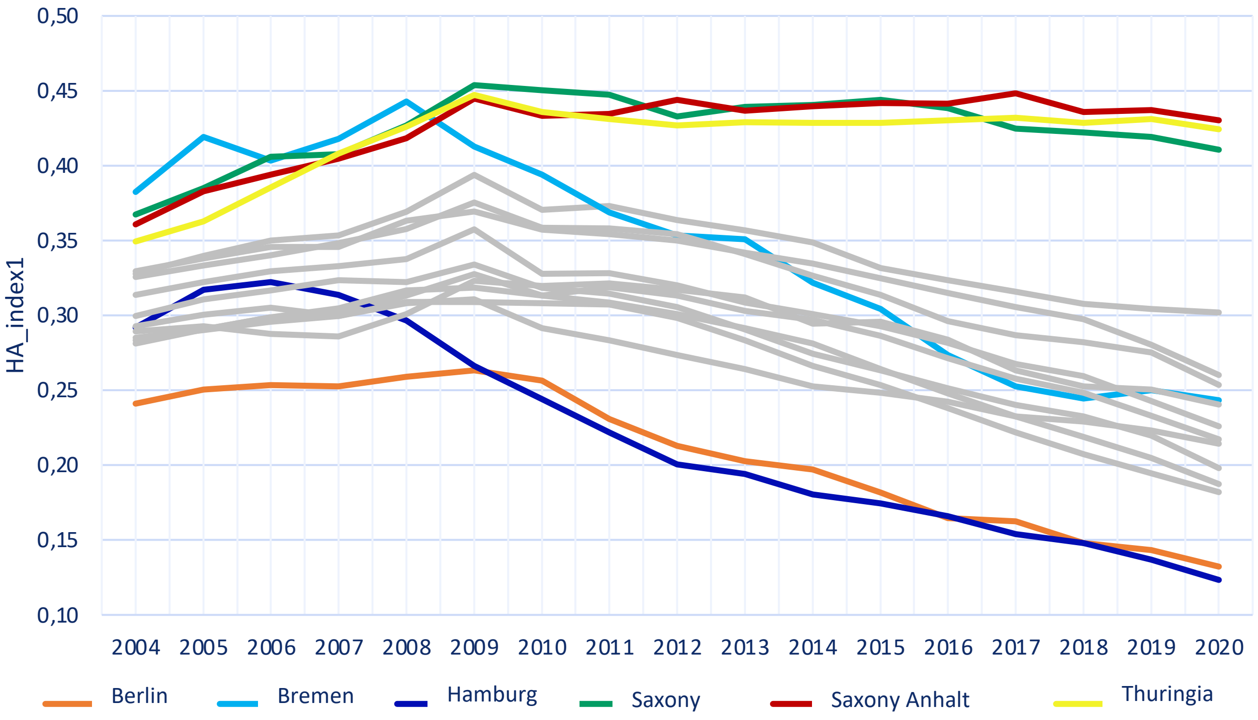
INC_{hh} – household disposable income

HH_{size} – household average size

QM_{ppc} – living area per capita

Interpretability: the ratio of annual household disposable income to the cost of purchasing an average-sized apartment in the region

Housing affordability dynamics in German federal states (HA_index1)



Housing Affordability Indexes (2)

$$HA_{index2} = \frac{\dfrac{0{,}25 * Average\ annual\ income\ of\ an\ average\ household}{Monthly\ mortgage\ payment\ for\ an\ apartment\ of\ average\ price}}{12} = \frac{\dfrac{0{,}25 * INC_{hh}}{M}}{12},$$

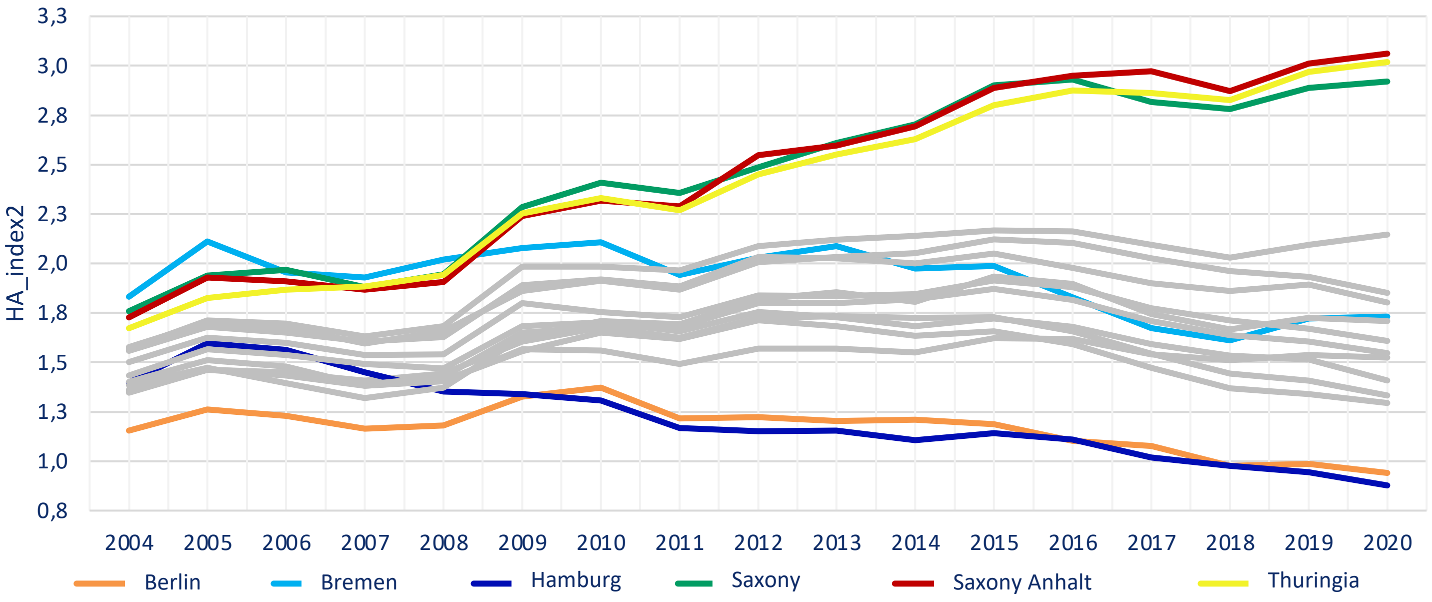
M – monthly mortgage payment for an average-priced apartment

$$M = \frac{PP_{qm} * HH_{size} * QM_{ppc} * (1 - 0{,}15) * \dfrac{RATE}{12}}{\left(1 - \dfrac{1}{1 + \dfrac{RATE}{12}}\right)^{312}}$$

Interpretability: the index value above 1 indicates that the household spends less than 25% of its income on housing. Otherwise, housing is considered unaffordable as the household spends more than 25% of its income on it.

PP_{qm} – apartment purchase price per m2
INC_{hh} – household disposable income
HH_{size} – household average size
QM_{ppc} – living area per capita
$RATE$ – mortgage rate
15% – down payment on mortgage (Statista, 2021)
312 months – average mortgage duration (Postbank, 2015).

Housing affordability dynamics in German federal states (HA_index2)



Housing Affordability Indexes (2)

$$HA_{index3} = \frac{Average\ annual\ income\ of\ an\ average\ household}{Average\ apartment\ rent\ price\ for\ an\ average\ household\ size} = \frac{INC_{hh}}{MRP_{qm} * HH_{size} * QM_{ppc} * 12}$$

MRP_{qm} – monthly rent per m2

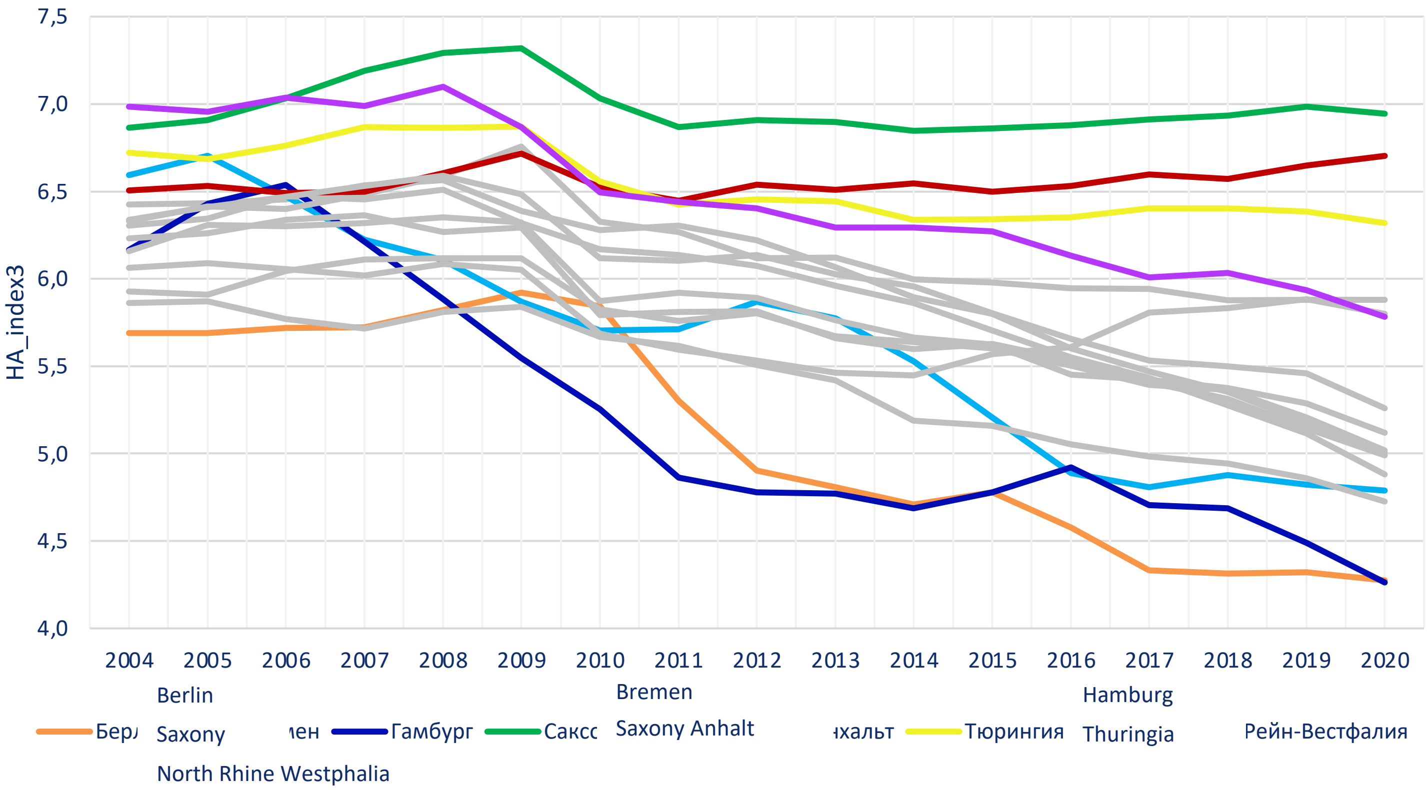
INC_{hh} – household disposable income

HH_{size} – household average size

QM_{ppc} – living area per capita

Interpretability: the ratio of annual household disposable income to the annual rental cost of an average-sized apartment in the region

Housing affordability dynamics in German federal states (HA_index3)



Data Description

Group	Feature	Description
Regional Labor Market Factors	employees	Number of employees per 1000 population
	unemp_rate	Unemployment rate (%)
	incommuters	Number of commuters arriving in the region per 1000 population
	outcommuters	Number of commuters leaving the region per 1000 population
Regional Demographic Characteristics	int_migration	Net internal migration per 1000 population
	aver_age_pop	Population average age
Regional Housing Market Factors	avcost_land	Average purchase price of development land (EUR per m2)
	constr_housing	Number of permits for the housing construction per 1000 population
	new_housing	Number of finished residential premises in new buildings per 1000 population
Regional Economic Indicators	gdp_ppc	GRP per capita (thousand euros)
	tourism	Number of overnight stays in tourist accommodation establishments per 1 resident

Moran‘s Indexes and Spatial Weights

Global Moran‘s Index

$$I = \frac{N}{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}} \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}(X_i - \bar{X})(X_j - \bar{X})}{\sum_{i=1}^{N}(X_i - \bar{X})^2}$$

w_{ij} – elements of matrix W on the intersection of i and j regions
N – number of regions

Table of global Moran's indexes for housing affordability indices:

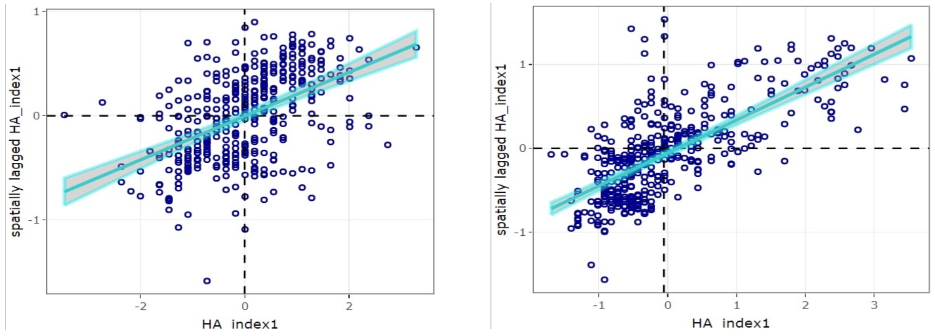
	2004	2010	2015	2020
index1	0,22	0,32	0,37	0,42
index2	0,22	0,32	0,37	0,42
index3	0,24	0,24	0,26	0,39

*all indexes are statistically significant at the level of significance 0.001

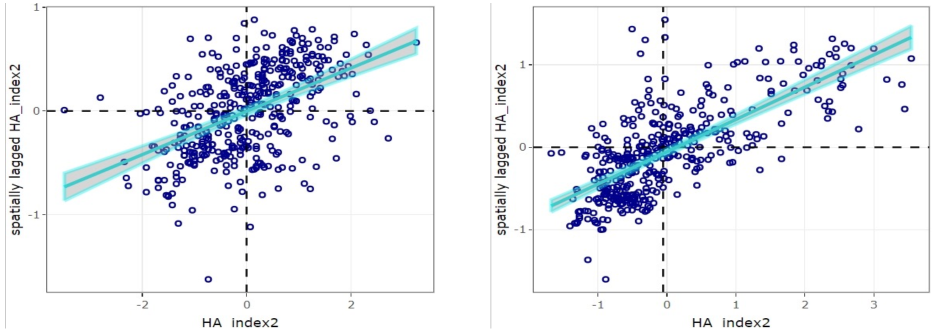
We have leveraged an inverse distance matrix W with a distance cut-off of 150 km:

$$w_{ij} = \begin{cases} 0, & if\ i = j \\ \frac{1}{distance_{ij}}, & if\ distance_{ij} < distance_{crit} \\ 0, & if\ distance_{ij} > distance_{crit} \end{cases}$$

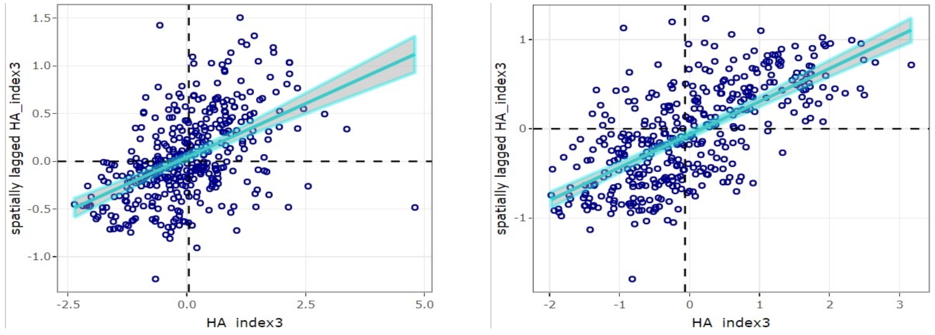
w_{ij} – element of matrix W on the intersection of i and j regions; $distance_{ij}$ – distance between regions centroids; $distance_{crit}$ – cut-off distance.



Moran's diagrams for HA_index1 in 2004 and 2020.

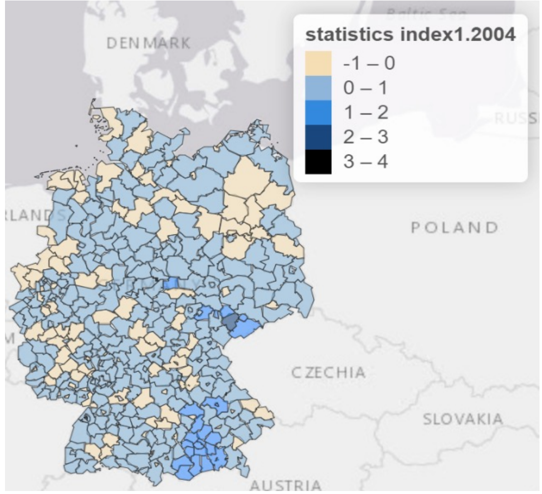


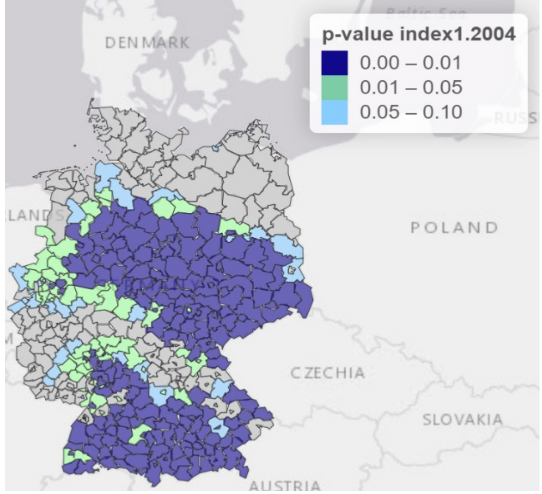
Moran's diagrams for HA_index2 in 2004 and 2020.



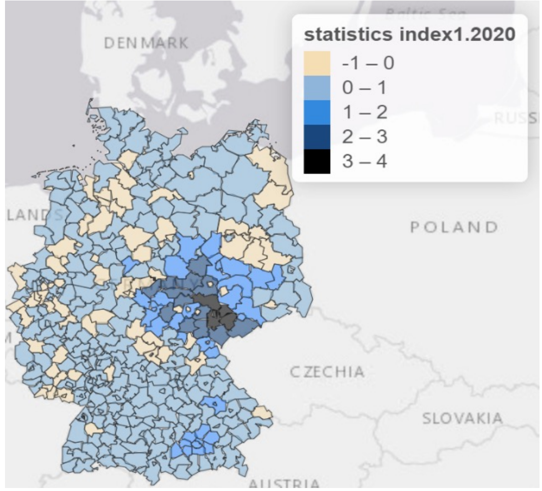
Moran's diagrams for HA_index3 in 2004 and 2020.

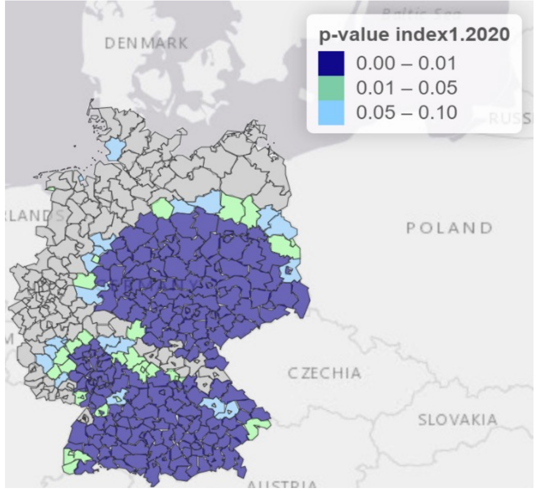
Spatial Correlation HA_index1





2004 г.





2020 г.

Local Moran‘s index statistics values (1) and corresponding p-values (2) for HA_index1

Conclusions:

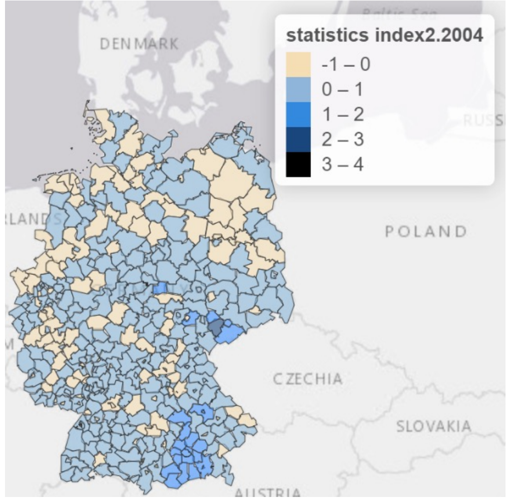
1. Throughout the entire time period, in regions where the values of local Moran's indices are significant, a positive spatial correlation is observed, which indicates that the regions are similar to their neighbors and can be combined into clusters.

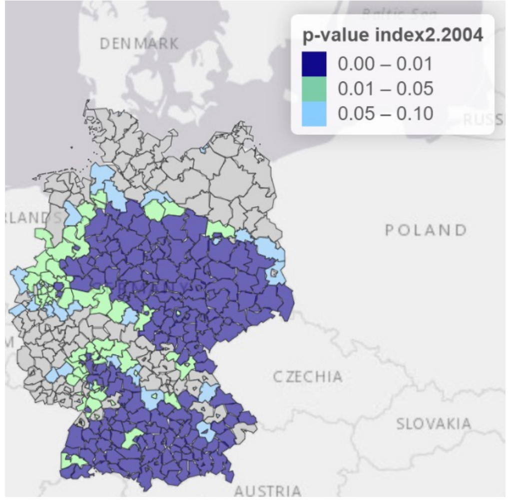
2. Over time, the significance of local Moran's indices increases and the positive spatial correlation becomes stronger.

3. Compared to 2004, an increase in the local Moran’s indices significance is observed in Thuringia, Lower Saxony, and North Rhine-Westphalia.

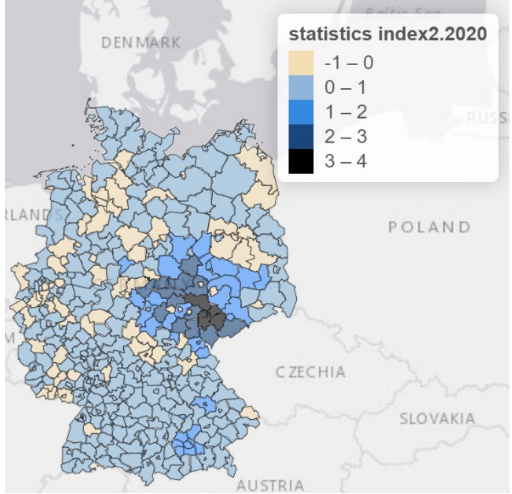
4. The strongest spatial correlation is observed in the center, south, and east of Germany.

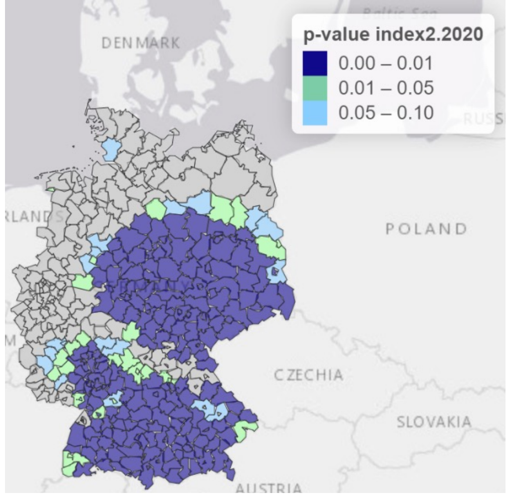
Spatial Correlation HA_index2





2004 г.





2020 г.

Local Moran‘s index statistics values (1) and corresponding p-values (2) for HA_index2

Conclusions:

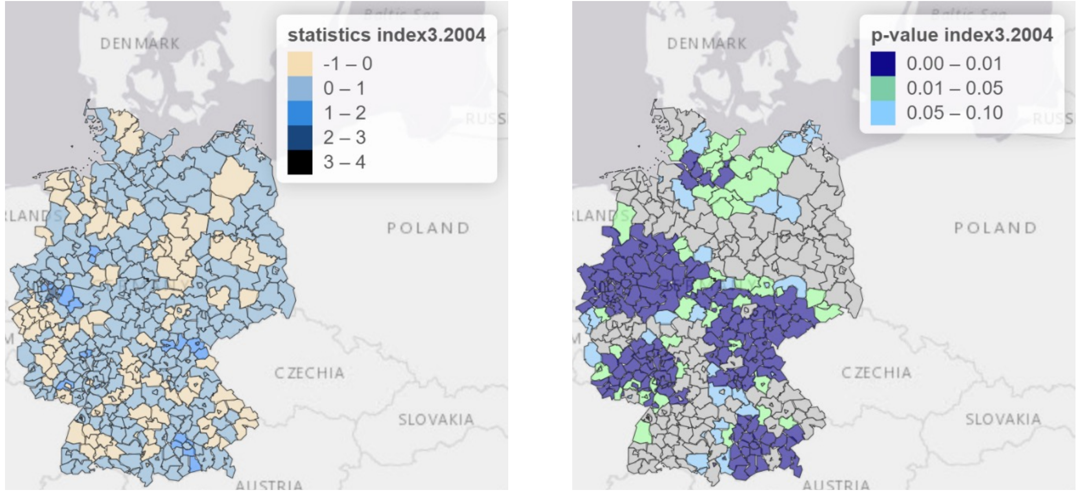
1. Over time, the significance of local Moran's indices increases and the positive spatial correlation becomes stronger.

2. Compared to 2004, an increase in the significance of local Moran's indices is observed in Saxony, Thuringia, and Baden-Württemberg.

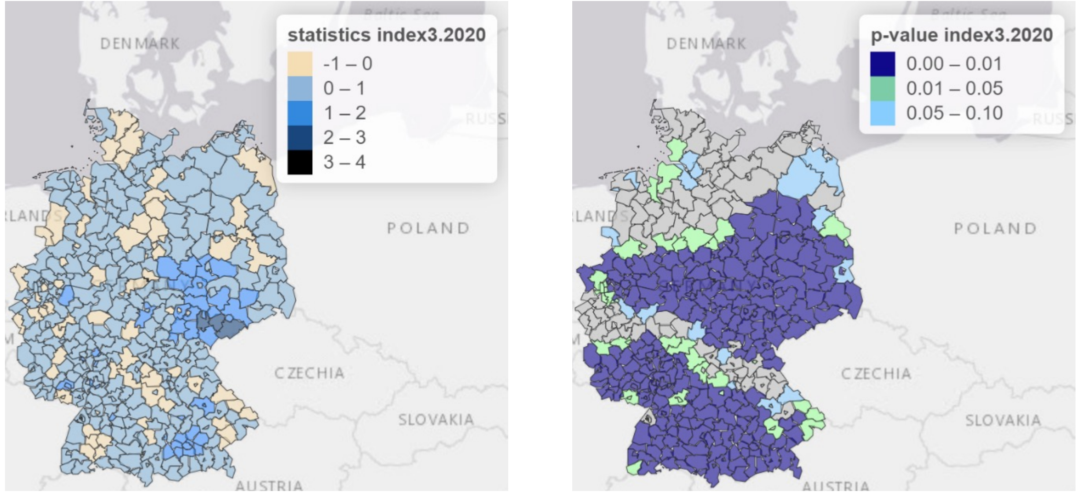
3. The maximum value of the index in 2020 is observed in the region of Greiz (Thuringia).

4. There are two clusters: A) Saxony, Saxony-Anhalt, Lower Saxony, and Thuringia B) most regions of Bavaria, Baden-Württemberg, and Rhineland-Palatinate.
5. Cluster B was stable within its boundaries in 2004-2020.

Spatial Correlation HA_index3



2004 г.



2020 г.

Local Moran's index statistics values (1) and corresponding p-values (2) for HA_index3

Conclusions:

1. Over time, the significance of local Moran's indices increases and the positive spatial correlation becomes stronger.
2. High values of the local Moran's index are observed among regions in different parts of the country (the cities of Remscheid, Wuppertal, Bielefeld, Hof, and Coburg; the regions of Miesbach, Munich, Dachau, and Kronach).
3. Compared to 2004, an increase in the significance of local Moran's indices is observed in Saxony, Thuringia, Lower Saxony, and Rhineland-Palatinate.
4. There were 5 clusters in 2004 and 2 clusters in 2020 (the western and eastern clusters are united; in the south of the country the southern and southwestern clusters are united; the cluster in the north of Germany disappeared).

Spatial Models

General nesting spatial model (GNS)

$$HA_{index_{lt}} = \alpha + \beta X_t + \rho W HA_{index_{lt}} + \theta W X_t + \mu + \eta_t + \varepsilon,$$

$$\varepsilon = \lambda W \varepsilon + u, u \sim iid(0; \sigma^2 I_N)$$

Spatial Durbin model (SDM)

$$HA_{index_{lt}} = \alpha + \rho W HA_{index_{lt}} + \beta X_t + \theta W X_t + \mu + \eta_t + \varepsilon$$

Spatial autoregressive model (SAR)

$$HA_{index_{lt}} = \alpha + \rho W HA_{index_{lt}} + \beta X_t + \mu + \eta_t + \varepsilon$$

Spatial error model (SEM)

$$HA_{index_{lt}} = \alpha + \beta X_t + \mu + \eta_t + \varepsilon, \varepsilon = \lambda W \varepsilon + u$$

$HA_{index_{lt}}$ – housing affordability index l $(l = 1, 2, 3)$ for all regions at time t

X_t – independent variables for all regions at time t

W – spatial weights matrix of shape 401x401

ρ и θ – spatial autocorrelation parameters for dependent and independent variables

λ – spatial autocorrelation paraemeter for error term

μ – fixed individual effects

η_t – time effects

ε – error term

Direct and indirect effects

Model	Direct effect	Indirect effect
OLS/SEM	β_k	0
SAR	Diagonal elements of $(I - \rho W)^{-1} \beta_k$	Non-diagonal elements of $(I - \rho W)^{-1} \beta_k$
SDM	Diagonal elements of $(I - \rho W)^{-1} (\beta_k + W\theta_k)$	Non-diagonal elements of $(I - \rho W)^{-1} (\beta_k + W\theta_k)$

Results: SDM models marginal effects for all HA indices

VARIABLES	Direct_index1	Indirect_index1	Direct_index2	Indirect_index2	Direct_index3	Indirect_index3
ln_employees	-0.231***	-0.541***	-2.182***	-7.960***	-1.011*	-2.515
unemp_rate	-0.00411***	-0.00316	-0.0330***	0.0427	-0.0139	-0.0965***
incommuters	0.00311***	0.0201**	0.0293***	0.214**	0.0163*	0.0339
outcommuters	-0.00329***	-0.00897	-0.0275***	-0.0756	-0.0309***	-0.0138
int_migration	3.84e-05	-0.00106	0.000746	0.00580	0.000438	0.0115
aver_age_pop	0.0199***	-0.0154*	0.101***	0.0695	0.238***	-0.510***
ln_avcost_land	0.000340	-0.0778***	-0.00817	-0.523	0.0230	0.139
constr_housing	-0.00307***	0.00101	-0.0172***	0.111***	-0.0262***	0.00701
new_housing	-0.00250***	-0.00911**	-0.0158***	-0.0404	-0.0331***	-0.0833
ln_gdp_ppc	0.0189	-0.138	0.222**	0.920	0.0224	-1.373
ln_tourism	-0.00362	0.0279	0.00539	0.218	-0.107	0.209
rho	0.652***	0.652***	0.794***	0.794***	0.727***	0.727***
sigma2_e	0.000665***	0.000665***	0.0259***	0.0259***	0.0741***	0.0741***
Observations	**6,817**	**6,817**	**6,817**	**6,817**	**6,817**	**6,817**
Number of regions	**401**	**401**	**401**	**401**	**401**	**401**
AIC	**-30348**	**-30348**	**-5302**	**-5302**	**1820**	**1820**
BIC	**-30184**	**-30184**	**-5138**	**-5138**	**1983**	**1983**

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

- Tourism and internal migration don't affect housing affordability in Germany.
- Negative effect: number of employees, unemployment rate, commuters out-flow.
- Postive direct effect: housing market group of variables.
- Positive effect: commuters in-flow, GRP per capita.
- Average age – positive direct and negative indirect effect.

Conclusions

- There is a **significant positive spatial correlation** in housing affordability in Germany that increases over time.
- Hosing in **East Germany is more affordable** than in West Germany.
- There is a significant **relationship between housing affordability and the labor market**, that emphasizes the importance of supportive economic measures in the regions with high unemployment.
- Current **construction** activity levels are **not capable** of increasing housing affordability.
- Higher **commuting** intensity is an **effective instrument** against rapid housing demand growth in major metropolitan areas.
- **Internal migration and tourism** variables have **low variance** across regions and have little to no effect on housing affordability.

Thank you for your attention!

Supplementary materials

Moran‘s Indices

Table 1. Global Moran‘s indices for housing affordability indexes and FE model residuals

	index1	index2	index3	resid_index1	resid_index2	resid_index3
2004	0,22	0,22	0,24	0,22	0,23	0,16
2005	0,24	0,24	0,23	0,23	0,24	0,18
2006	0,26	0,26	0,21	0,23	0,24	0,19
2007	0,27	0,27	0,22	0,23	0,24	0,20
2008	0,30	0,30	0,24	0,23	0,24	0,21
2009	0,33	0,33	0,23	0,23	0,24	0,21
2010	0,32	0,32	0,24	0,23	0,23	0,22
2011	0,33	0,33	0,23	0,23	0,23	0,23
2012	0,33	0,33	0,22	0,23	0,23	0,23
2013	0,36	0,36	0,23	0,23	0,23	0,24
2014	0,37	0,37	0,24	0,23	0,23	0,24
2015	0,37	0,37	0,26	0,23	0,23	0,25
2016	0,39	0,39	0,26	0,23	0,24	0,26
2017	0,39	0,39	0,27	0,23	0,24	0,26
2018	0,39	0,39	0,32	0,24	0,24	0,27
2019	0,41	0,41	0,36	0,23	0,23	0,27
2020	0,42	0,42	0,39	0,24	0,24	0,27

Model Performance Metrics

Table 2. AIC and BIC metrics for the SDM, SAR, and SEM models (HA_index1) with different spatial weights.

HA_index1	AIC_SDM	AIC_SAR	AIC_SEM	BIC_SDM	BIC_SAR	BIC_SEM
Adjacency matrix	-30245.41	-30332.39	-30253.37	-30081.55	-30134.4	-30055.38
Adjacency matrix (cut-off 100 km)	-30323.35	-30217.56	-30261.8	-30159.5	-30019.58	-30063.81
Inverse distance matrix	-30131.46	-30064.52	-30024.61	-29967.6	-29866.53	-29826.62
Inverse distance matrix (cut-off 80 km)	-30289.19	-30297.46	-30277.41	-30125.34	-30099.48	-30079.42
Inverse distance matrix (cut-off 100 km)	-30340.15	-30292.74	-30241.46	-30176.29	-30094.76	-30043.47
Inverse distance matrix (cut-off 150 km)	-30347.51	-30246.06	-30261.15	-30183.65	-30048.07	-30063.16
Inverse distance matrix (cut-off 200 km)	-30345.92	-30203.45	-30062.76	-30182.06	-30005.46	-29864.77

Table 3. AIC и BIC for the FE and SDM models (all HA indexes)

	AIC_FE	AIC_SDM	BIC_FE	BIC_SDM
HA_index1	-29891	-30347	-29706	-30183
HA_index2	-4480	-5494	-4296	-5221
HA_index3	2419	1720	2603	1993