RUHR ECONOMIC PAPERS

Uwe Neumann
Sandra Schaffner

Neighbourhood-level Variation in the Risk of Private Credit Default – a Driver of Urban Residential Segregation?

rwi

#875

Imprint

Ruhr Economic Papers

Published by

RWI – Leibniz-Institut für Wirtschaftsforschung
Hohenzollernstr. 1-3, 45128 Essen, Germany

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Editors

Editorial Office

Ruhr Economic Papers #875

Responsible Editor: Ronald Bachmann

ISSN 1864-4872 (online) – ISBN 978-3-96973-013-3

Ruhr Economic Papers #874

Uwe Neumann and Sandra Schaffner

Neighbourhood-level Variation in the Risk of Private Credit Default – a Driver of Urban Residential Segregation?

Bibliografische Informationen der Deutschen Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at **http://dnb.dnb.de**

RWI is funded by the Federal Government and the federal state of North Rhine-Westphalia.

http://dx.doi.org/10.4419/96973013
ISSN 1864-4872 (online)
ISBN 978-3-96973-013-3

Uwe Neumann and Sandra Schaffner[1]

Neighbourhood-level Variation in the Risk of Private Credit Default – a Driver of Urban Residential Segregation?

Abstract

Credit default is a dramatic consequence of disadvantageous private financial decisions. Using regression methods which eliminate spatial autocorrelation at the level of 1 km² grids and further identification problems, we observe considerable and reinforcing residential segregation between households facing payment difficulty and more solvent households. Two findings give reasons for concern. First, data from North Rhine-Westphalia reveals that a high local risk of credit default coincides with a lower share of children taking the highest German secondary school track (Gymnasium). Since birth rates are currently high in these (inner city) areas, the outlook on educational attainment for many pupils is bleak. Second, hedonic price estimations using microdata on housing offers find that local agglomeration of households facing credit default provokes significant (detrimental) neighbourhood effects on housing markets. Segregation is thus unlikely to diminish, which implies increased efforts should be made to overcome unfavourable neighbourhood effects in various fields of policy, especially education.

JEL-Code: R23, R21, G51

Keywords: Segregation; credit default; financial literacy; spatial autocorrelation; hedonic price analysis

November 2020

1 Both RWI – We thank Lea Eilers, Daniela Schwindt, Lisa Taruttis and Sophia Völker for valuable research assistance. This research was supported by a grant provided by the Dr. Brigitte and Josef Pauli-Stiftung.-
All correspondence to: Uwe Neumann, RWI, Hohenzollernstr. 1-3, 45128 Essen, Germany, e-mail: uwe.neumann@rwi-essen.de

1. Introduction

Modern economic life makes increasing demands on everyone's level of knowledge and decision-making capacity. The most dramatic consequences of a lack of basic skills include poor educational choices, insufficient old-age provision and over-indebtedness. Missing skills often act as obstacles to personal development and happiness. This study focusses on the regional outcomes of (a lack of) financial literacy. Starting point is the finding that knowledge gaps in the field of basic financial skills are widespread. The following sections examine whether and to what extent neighbourhood-level segregation by demographic characteristics such as age, income and nationality interrelates with variation in the local accumulation of private credit default. For this purpose, information about the credit default risk of individuals in Germany, RWI-GEO-GRID at the level of 1km² grids, is examined in combination with neighbourhood statistics and microdata on housing offers for rent and for sale (RWI-GEO-RED).

The literature on financial skills points out that low levels of basic financial knowledge, among other things, increases the likelihood of sub-optimal savings and borrowing decisions (Lusardi and Mitchel 2014, Lusardi, Michaud and Mitchel 2017). So far, however, it has rarely been examined which societal contexts influence financial knowledge beyond school education, and what significance may be attached to regional contexts.

In this context, we analyse whether specific characteristics of the residential population and the local housing environment interrelate with the local share of households facing credit default. Further, we analyse whether demographic change can even increase spatial variation in the share of households facing credit default. Our study pursues the following basic questions:

1. How strongly does private credit default vary between urban neighbourhoods?

2. What are the correlations between high rates of credit default and other features of inner-city segregation (e.g. educational attainment) and do these correlations change over time?

3. Do over-indebted households concentrate in specific neighbourhoods, or are larger urban zones characterised by high default rates, such that an accumulation of households with financial difficulties in one neighbourhood predetermines high rates of over-indebtedness also in neighbouring areas?

4. Do housing prices suggest that undesirable neighbourhood effects derive from a local accumulation of households with financial problems?

We give a comprehensive overview on the regional small-scale distribution of credit default risk in German agglomerations and on how this is correlated to sociodemographic characteristics and demographic changes. Further, we are the first to show that the German school tracking system can confirm or even aggravate spatial differences in credit default risk. In Germany, pupils usually change from primary to secondary school at age ten. There are different tracks of secondary school. The highest track (Gymnasium) leads to the Abitur (high school/upper secondary school degree), a second track (*Realschule*) ends after ten years of schooling and the third track (*Hauptschule*) ends after nine years. Finally, there exist comprehensive schools (*Gesamtschule*) which were introduced as an alternative to the traditional three-tiered secondary education system. Instead of three different schools, there are three different tracks within one school. These school types exist parallelly in some federal states.

The next section gives a brief overview on the literature on financial literacy and segregation. The three data sets used as well as the methodological approach are described in section 3. Section 4 presents some first descriptive results on the spatial distribution of high credit default risk. The results of the multivariate analyses of spatial interdependencies between credit default and other characteristics of residential segregation and of the neighbourhood effects deriving from a local accumulation of credit default risks can be found in section 5 and section 6 finally concludes.

2. Literature Review

Studies on financial literacy demonstrate great variation in basic financial skills across socio-demographic groups in many countries. In Germany, for example, in 2009 among respondents from a representative survey, only about 20% of those with less than an upper secondary school degree (Abitur), but over 70% of university graduates answered correctly to a range of three basic questions relating to interest rates, inflation and risk diversification (Lusardi and Mitchell 2014). It is the purpose of the following study to outline some of the consequences of inadequate basic financial skills from a regional perspective.

Research on the extent of residential segregation and the processes leading to spatial separation received a new impetus in the 1980s by a so-called "polarisation hypothesis". Several authors argue that because of the shift of economic activity from manufacturing to services income inequality should increase. Increased income inequality affects the choice of residential location and thus transmits occupational segmentation to polarisation between residential areas (Friedmann 1986, Sassen 1991). These assumptions have been a matter of controversy and several decades later it is by no means clear whether there is a general trend towards increasing segregation. A research branch based on the work of Tiebout (1956) is concerned with the question whether neighbourhood amenities influence household location decisions and thus segregation processes, which are also referred to as "Tiebout sorting" in this context (Anderson 2007, Rosenthal and Ross 2015). In recent decades, many studies have demonstrated such "neighbourhood effects" (Wilson 1987) on residential location choice and segregation (Behrens et al. 2018, Kuminoff et al. 2013).

In the emerging digitised labour market, many routine tasks are at risk of automation, as the cost of computer technology continues to decline. In order to stay or get employed, workers need to maintain a high degree of flexibility, creativity, problem-solving and complex communication skills (Autor et al. 2003). Much likely, for a child growing up in an environment where credit default is common and where many people find it hard to acquire basic financial skills, it will be more difficult to obtain such capacities than for someone experiencing more prosperous local surroundings. Chakraborty et al. (2019) observe that a lack of high-educated role models from the same ethnicity decreases students' educational outcomes. At the level of neighbourhoods, the transition rate from primary schools to the most academically oriented type of secondary school, the Gymnasium, is an important indicator of local educational attainment. Given the availability of school places, access to a Gymnasium may depend partly on the recommendation by the primary school. Empirical findings in education economics show that such recommendations relate, among other things, to the social status of parents (Dustmann 2004, Heineck and Riphahn 2009).

This study provides a link between the literature on residential segregation and current research on financial literacy and on the social barriers to education. First, we adopt methods established in urban economics and spatial econometrics to examine the extent to which the occurrence of credit default indicates socioeconomic inequality at the level of urban neighbourhoods. Second, we add to the literature on financial literacy by providing evidence on the role of local educational attainment as a predeterminant of a high local risk of credit default in Germany.

3. Data and methods

3.1 Data sources

In this study we combine three different data sets at the 1 km² grid level. First, we need data on payment default risk as well as information on the socio-demographic composition of the residents (RWI-GEO-GRID). Second, for the analysis of the correlation between credit default and the housing market, information on housing prices is used (RWI-GEO-RED). Third, we draw on neighbourhood-level information on the transitions after primary school. The data set comprises information on the residents of all populated 1 km²-grid cells in Germany.

In addition to data on population (composition) and building stock, a "payment index" is included. The payment index is determined by micro-level information on the solvency of households compiled by Creditreform, a referencing agency (see below). The index is of central importance to our research, since it describes the statistical probability of payment default for each residential dwelling in Germany. All dwellings are categorised annually into nine equally sized groups ranked according to the index values. Each of these classes thus covers 11.11% of all residential houses in Germany in each year. In the following, houses belonging to one of the two classes with the worst payment index (highest probability of payment default) are considered to host inhabitants facing payment difficulties and the share of the population residing in these houses at the level of 1 km² grids is defined as our local indicator of the risk of payment default. Households that belong to the group with the 22% highest payment default risks are therefore defined as our group of high payment default risk.

The data was originally generated by microm GmbH, which is part of the Creditreform Group, one of the two largest credit referencing agencies in Germany. The agency reports on (according to its own account) 61 million individuals (excelled only slightly by the Schufa Holding, which reports on 66.4 million individuals) (Goldmedia 2017). According to Creditreform (2018), about one in ten inhabitants aged over 18 in Germany is currently over-indebted, i.e. affected by personal bankruptcy, debt collection or "severe payment disruption". Such disruption is registered by Creditreform whenever an individual receives at least two unsolicited reminders from multiple creditors. Due to a high degree of interconnectedness among the financial system in Germany, it can be assumed that most of such occurrences will be registered by Creditreform.

The share of over-indebtedness varies by age and region. “Prime-aged” individuals (30-40) experience the highest risk (18.6%), while it is lowest for those above 70 (2%). In regional comparison, over-indebtedness rates are below the average in Southern Germany (7.6%). The Ruhr area is a “hotspot“ of over-indebtedness, where up to around 18% of the over 18-year-old population of some cities (Gelsenkirchen, Herne) are affected. However, most large cities are characterised by high shares of over-indebtedness, e.g. Berlin (12.4%), Cologne (11.8%) and Frankfurt am Main (10.6%) (Creditreform 2018).

The RWI-GEO-RED data derives from one of the largest internet-based real estate platforms in Germany, ImmobilienScout24. The data includes rental and purchase offers of houses and apartments on the internet portal, in addition to detailed information on the characteristics of the objects on offer. The georeferenced data used in the following analysis covers the period from 2009 to 2016.

In addition, detailed information on the transitions from each primary school to secondary schools in North Rhine-Westphalia in 2018 has been made available by Statistics North Rhine-Westphalia (IT.NRW 2019). Using spatial interpolation methods, the transition rate from primary school to Gymnasium was calculated at the level of 1 km² grids, assuming that pupils attend one of the primary schools located nearest to their place of residence.

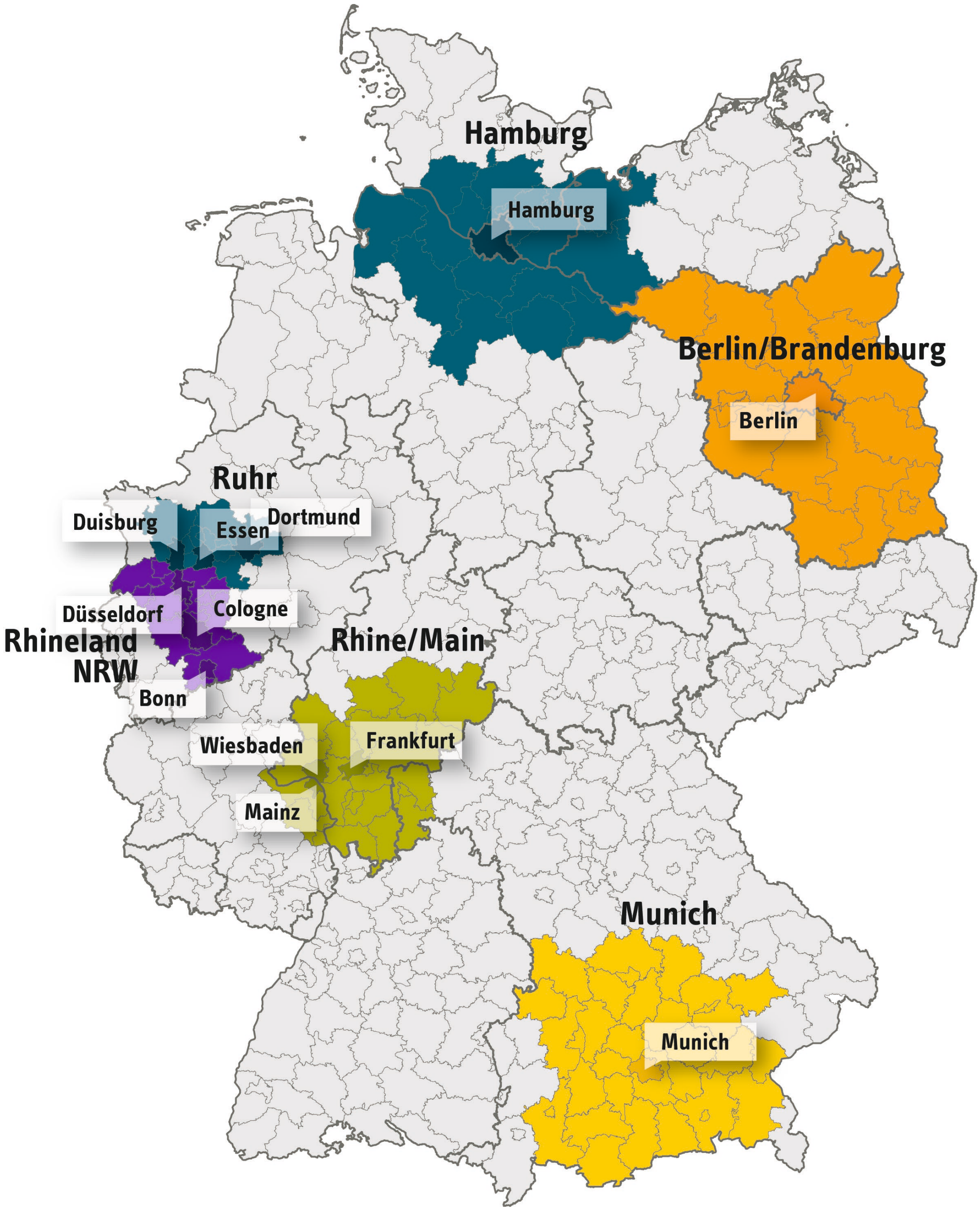
3.3 Methodological Approach

Based on the three data sets, RWI-GEO-GRID, RWI-GEO-RED and school transitions, the spatial relationship between the likelihood of credit default and other characteristics of the residential population and the housing environment is examined at the level of 1 km² grids, using six German agglomerations as regional units for our case study (Figure 1). Since it is our goal to analyse neighbourhood effects and spatial autocorrelation, we focus on urban regions. About 40% of all Germans live within the chosen agglomerations.

We set the boundaries of agglomerations according to definitions applied by the Ministerial Conference on Spatial Planning (MKRO) as spatial units for the purpose of territorial planning (BMVBS and BBR (ed.) 2007, Pütz 2016). The Rhine-Ruhr region defined in this context focusses heavily on the densely populated urban core zones of North Rhine-Westphalia. In other regions, e.g. in Berlin/Brandenburg, Hamburg and Munich, the respective agglomerations are assigned a much larger and less densely populated area. We draw on this definition nevertheless, since the resulting agglomerations are comparable to each other in terms of their overall population (Tables 1-2, see below).

The aims of this study are, first, to analyse the statistical relationship of credit payment risk to demographic composition, housing costs as well as demographic change and school transition rates at the level of neighbourhoods in urban regions and, second, whether these correlations changed during the past two decades. Third, we examine the extent to which households facing credit default agglomerate (und thus segregate from other households) within urban regions and fourth, we explore the (potentially detrimental) economic neighbourhood effects deriving from a local agglomeration of credit default risks.

Figure 1
Selected urban agglomerations in Germany



Own figure based on BMVBS and BBR (ed.) (2007) and Pütz (2016); NRW = North Rhine-Westphalia. Selected core cities are highlighted.

The analysis faces several methodical difficulties. First, due to correlation between various characteristics of neighbourhood populations (including the share of households facing credit default), any regression analysis using a range of such indicators may suffer from multicollinearity and endogeneity. Further, unobserved heterogeneity at the neighbourhood level may affect any cross-sectional analysis using grids as spatial entity. Moreover, spatial autocorrelation among the univariate distribution of credit default risk may lead to measurement error. Therefore, we define a sequence of regression models, each designed to overcome specific issues likely to bias the results of our neighbourhood-level analysis. First, we reduce the number of variables by factor analysis, before we start the regressions. The basic regression is extended by more elaborate approaches accounting for the identification problems that may affect our analysis. Finally, (potentially negative) neighbourhood effects are analysed by a hedonic house price regression.

As explained, a first step is to derive a selected range of statistically independent dimensions from a larger number of variables by factor analysis, in order to eliminate endogeneity and multicollinearity. Variables are restricted from the analysis if they correlate with a statistical dimension that is also highly correlated with the risk of private credit default in a given year. Rather, variables are selected that are correlated with any of the other dimensions, which are statistically independent from the dimension that determines the credit default probability. After selecting suitable indicators, to answer our first two research questions (see above) we estimate the following OLS model (1), which includes all years and regions:

$$(1)\ \log\left(y_{i,t}\right) = \beta_0 + \sum_{l=1}^{5} \beta_l X_{i,t}^{l} + \sum_{t=2009}^{2016} \tau_t\ T_t + \sum_{k=1}^{5} \gamma_k\ A_i + \varepsilon_{i,t}$$

Therein, $y_{i,t}$ corresponds to the share of persons with a high payment default among the population of grid i in year t = 2009, 2010, ..., 2016. $X_{i,t}$ corresponds to l = 1, …, 5 independent measures at the level of grid i in year t (population density, share of foreigners, shares of age groups <15 and ≥65, number of residential buildings). T_t comprises year fixed effects for 2009, 2010, ..., 2016 (reference is base year 2005, there are no values for 2006-2008), and A_i comprises fixed effects for the agglomeration (reference group is the Ruhr in this case). An alternative estimation covers only the year 2016, to capture current characteristics. Separate estimations for the urban agglomerations in North Rhine-Westphalia (Ruhr and Rhineland NRW) also control for a sixth demographic variable among the X_i in the analyses for 2016, i.e. the local transition rate among pupils from primary school to Gymnasium.

An alternative model (equation 2) examines the change in the payment default risk from t_0 = 2011 to t_1 = 2016, whereby the change of y_i is examined as a function of the X_i in base year t_0.

$$(2)\ \log(y_{i,t1}) - \log(y_{i,t0}) = \beta_0 + \sum_{l=1}^{5} \beta_l X_{i,t0}^{l} + \sum_{k=1}^{5} \gamma_k\ B_i + \varepsilon_{i,t0}$$

A further estimation draws on the fixed effects model displayed by equation (3):

$$(3)\ \log(y_{i,t}) = \beta_0 + \sum_{l=1}^{5} \beta_l\, X_{i,t}^{l} + \sum_{t=2005}^{2016} \tau_t\, T_t + \mu_i + \varepsilon_{i,t}$$

The fixed effect μ_i accounts for unobservable heterogeneity between the grids. The analysis thus considers changes over time within grids and not, as in the cross-sectional model (1), comparison between regions. The difference to model (2) is that not only change over time with respect to the dependent variable is considered but change during the observation period concerning the annual values recorded for all variables. Thereby, unobservable heterogeneity between regions is eliminated. This advantage is gained at the expense of not being able to measure the influence of each indicator at a specific point in time. The different regression models therefore each represent part of the questions of interest for our study.

Models (1) - (3) do not consider that there may be significant spillover effects between spatial units and that the measured value of the dependent variable may be partly determined as an outcome of a spatial pattern, in which adjacent districts turn out highly similarly due to matters beyond the observational scope of the analysis. If this spatial context is not controlled for, one may overestimate the explanatory content of variation among the independent variables with respect to the dependent variable. Following our third research question, in order to determine whether such spillovers may affect our analysis, we examine the univariate spatial distribution of the dependent variable $y_{i,t}$ from equation (1) for selected annual cross-sections. Using the "local indicators of spatial association" (LISA) approach introduced by Anselin (1995), we compute the sum of the local Moran's I statistics according to equation (4) as a measurement of the spatial autocorrelation of the dependent variable within our study regions.

$$(4)\ \sum_i I_i = \sum_i z_i \sum_n w_{in} z_n$$

In equation (4), w_{in} represents a spatial weight that is inversely proportional to the distance (in kilometres) between the 1 km² grid areas i and n. The variables z_i and z_n measure the deviation from the annual mean of the dependent variable y in grids i and n. The simplest form of spatial weighting matrices are so-called neighbourhood matrices, where neighbours are defined such that they share a common boundary. In this binary neighbourhood matrix, the value 1 is coded for the $n \in N$ satisfying the neighbourhood criterion, otherwise the value is 0. The variable z_n represents the average value of the deviation from the mean of the variable y in areas $n \in N$.

We further restrict N to neighbouring grids that share a 1 km boundary, i.e. to a maximum of four neighbours.

The degree of spatial autocorrelation among univariate observations can be visualised by so-called LISA clusters, which commonly classify four categories, i.e. "high-high" with positive z_i and z_n, "high-low" with positive z_i and negative z_n, "low-high" with negative z_i and positive z_n and "low-low" with negative z_i and z_n. A map representing the affiliation of all spatial units (grids) to one of these categories illustrates the extent to which there is spatial autocorrelation with respect to the variable of interest. Local Moran's I statistics is calculated for each spatial unit and the statistical significance for each I_i is evaluated taking into account its positioning within or without a confidence interval defined with respect to a predefined p-value (in our case 0.05, which is common practice).

A positive value of the Moran's I statistics from equation (4) characterises an agglomeration of similar values among neighbouring spatial units. On a map displaying four zones for values z_i and z_n as described before, either the zones "high-high" or "low-low" or both zones would dominate in case of a positive regional Moran´s I. A negative value would correspond to a dominance of the zones "high-low" and/or "low-high".

In order to eliminate measurement errors resulting from spatial autocorrelation in the regression analysis, given that spatial autocorrelation is found to characterise the univariate distribution of the dependent variable, we apply spatial autoregressive (SAR) models (Drukker et al. 2013) in order control for the influence of neighbourhood spillovers. We estimate equations (5) and (6),

$$(5)\ \log(y_i) = \lambda \sum_{n=1}^{4} w_{in} \log(y_n) + \sum_{p=1}^{5} X_{ip} ß_p + u_{it}$$

$$(6)\ u_{it} = \rho \sum_{p=1}^{5} m_{in} u_n + \varepsilon_i$$

where w_{in} und m_{in} are spatial weights corresponding to our binary neighbourhood matrix, in which each grid i is assigned a maximum of four neighbours n. The X_i represent the five independent variables from equations (1) - (3), the ß comprise the parameters to be estimated, u_i is a spatially autoregressive error term, assuming that residuals ε_i are independent and identically distributed (IID) random variables. Parameter λ measures spatial autocorrelation in the spatial lag of the dependent variable and ρ measures spatial dependence in the error term (equation 6). The SAR model in our analysis is based on a maximum likelihood (ML) estimator (Drukker et al. 2013). In case of spatial autocorrelation, it is to be expected that the coefficients β_p in estimation (5) will be lower than the corresponding β_l from model (1), where the influence assigned to the independent variables would be overestimated.

In order to answer our fourth research question, we investigate whether (most probably undesirable) neighbourhood effects emerge due to local accumulation of households with payment difficulties. In the sense of Tiebout sorting (see above), high credit default rates among the population of a neighbourhood may be deemed disadvantageous from the view of residential location decisions. Hedonic price functions according to equation (7) (Rosen 1974) estimate the price effects of neighbourhood features among the housing markets of our selected agglomerations with respect to flats for rent and for sale and for one/two-family-homes on offer for sale.

$$(7)\quad \log\left(P_{i,jkt}\right) = \beta_x X_i + \sum_{l=1}^{5} \beta_l N_{j,t}^{l} + \sum_{k=1}^{K} \mu_k + \sum_{t=2009}^{2016} T_t + \sum_{k=1}^{K} \sum_{t=2009}^{2016} \mu_k T_t + \sum_{t=2009}^{2016} r N_j T_t + \varepsilon_{i,jkt}$$

P_{ijt} is the price or the monthly rent demanded for dwelling i in grid area j, X_i is a vector representing object-specific characteristics such as the year of construction, dwelling space (in m²), the number of rooms, and the overall condition. Neighbourhood feature rN_j represents the proportion of people with a high credit default probability in grid j, neighbourhood features N_j include the five other variables from equations (1) - (3) with respect to grid j. Dummy variables

μ_k control for fixed effects at the level of k = 6 agglomerations (in separate analyses for individual agglomerations, fixed effects are accounted for at the level of districts (NUTS 3 level, i.e. Kreise). Dummy variables T include year fixed effects. In addition, interactions of the fixed effects μ_k and T are considered. The credit default probability rN_j is interacted with fixed effects for years T. Since it can be assumed that the residuals $\varepsilon_{i,jkt}$ are correlated within regions, they are clustered at district level.

Information about real estate prices has been made available on a monthly basis, usually comprising multiple recordings for one object, since advertisements often appear for more than a month, and the demanded price may change over time. The prices reported on in the data are compiled for our analysis such that for each object only one observation per advertisement remains. We use the most recent offer, since this can be assumed to approximate the final transaction price more closely than earlier offers. In order to avoid a statistical bias due to the characteristics of very large or very small dwellings, the 1% largest and 1% smallest among the size for each category (flats for rent or sale, houses) in each year are excluded from the analysis.

4. Neighbourhood-level variation in private credit default - a descriptive outline

Statistical analysis suggests a close correlation between the local risk of credit default and various characteristics of the residential population (e.g. between average household income and credit default rates, which are inversely proportional at the grid level). As explained, in terms of total population (over 5 million) all agglomerations are comparable (Table 1). Regarding other measures, the descriptives reveal considerable differences between agglomerations. Population density and therefore the number of grids as well as the share of foreigners vary between agglomerations. Above all, credit default risk is by far lowest in the Munich agglomeration (only 5% in 2016, compared to 8% in Rhine/Main, 11% in Hamburg and 16% in the Ruhr, Table 2).

Table 1
Statistical characteristics of selected agglomerations
2016

Agglomeration	Total population	Mean (1 km^2-grid-level)				grids
		Population	of which			
			foreigners.	<15	>65	
Ruhr area	5,111,204	1,224	0.07	0.13	0.22	4,167
Rhineland NRW	5,833,199	1,251	0.08	0.14	0.21	4,661
Rhine/Main area	5,684,662	667	0.08	0.13	0.21	8,461
Berlin/Brandenburg	6,005,062	446	0.03	0.12	0.23	13,452
Hamburg	5,295,133	307	0.03	0.13	0.21	17,250
Munich	5,990,789	350	0.06	0.14	0.19	17,100

Authors´ calculation based on RWI-GEO-GRID. grids = number of grids.

Table 2
Share of high credit default risk – frequency distribution
2011 and 2016, 1 km^2-grids, selected agglomerations

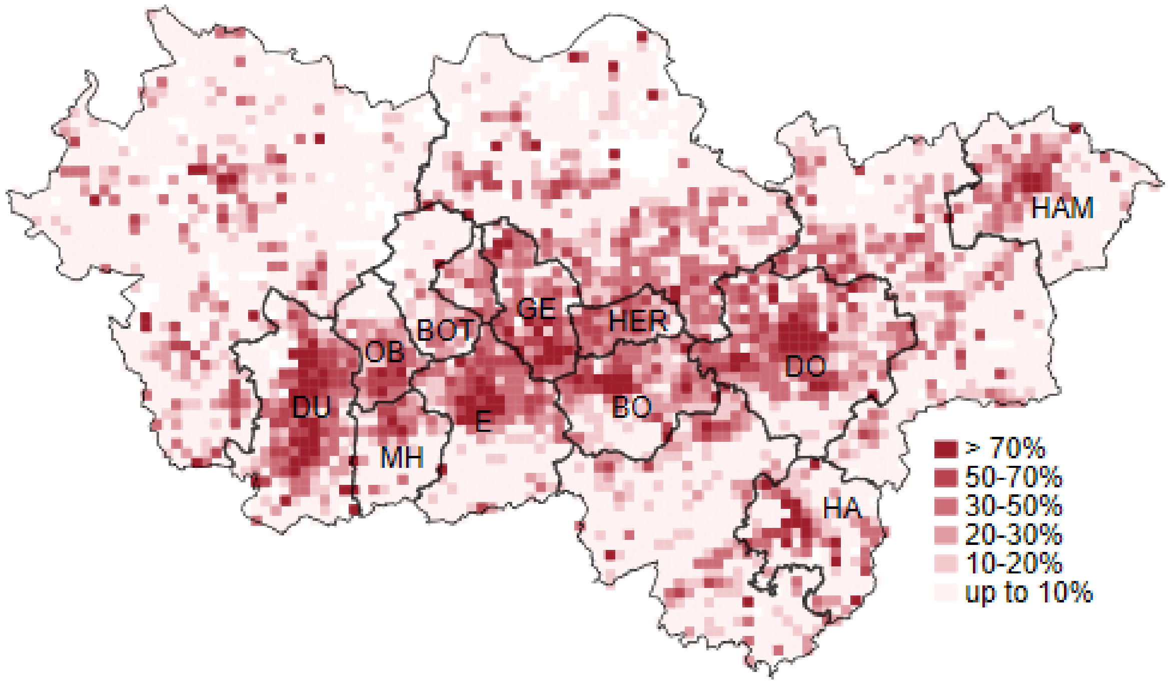
	2011		2016	
	mean	std. dev.	mean	std. dev.
Ruhr area	0.20	0.22	0.16	0.21
Rhineland NRW	0.17	0.20	0.13	0.19
Rhine/Main area	0.12	0.19	0.08	0.17
Berlin/Brandenburg	0.21	0.30	0.14	0.25
Hamburg	0.17	0.26	0.11	0.21
Munich	0.08	0.18	0.05	0.13

Authors´ calculation based on RWI-GEO-GRID. std. dev. = standard deviation.

The likelihood of credit default decreased in all regions between 2011 and 2016. However, in the largest core cities of the Ruhr area and the Rhineland (Cologne, Düsseldorf, Dortmund, Duisburg, Essen) the average share of households facing high credit default risk is about 40%.

The Ruhr area is characterised by a very explicit south-north differentiation with regard to the likelihood of credit default among its core zone, i.e. the large cities between Duisburg in the west and Dortmund in the east (Figure 2).

Figure 2
Population share facing a high risk of credit default in the Ruhr area
2016



Own figure. Data source: RWI-GEO-GRID; Bochum (BO), Bottrop (BOT), Dortmund (DO), Duisburg (DU), Essen (E), Gelsenkirchen (GE), Hagen (HA), Hamm (HAM), Herne (HER), Mülheim (MH), Oberhausen (OB); lines depict the boundaries of districts (Kreise and kreisfreie Städte).

Within the urban core zone of the Ruhr, households with a high default probability cluster mainly in the central and northern parts of the large cities, e.g. in Bochum, Dortmund, Duisburg, Essen, Gelsenkirchen, Herne, Mülheim and Oberhausen. The characteristic south-north divide within the urban core zone of the Ruhr, which also applies to other characteristics like household income, is related to the settlement history in the course of the northern migration of coal mining during the 20th century.

The univariate statistical analysis of regional Moran´s I according to equation (4) shows that in all conurbations there is spatial autocorrelation concerning the population share with a high credit default probability at the level of 1 km² grids (Table 3).

Table 3
Private credit default: Moran´s I for spatial autocorrelation
2011 und 2016, 1 km² grids

	Ruhr	Rhineland	Rhine/Main	Berlin	Hamburg	Munich
2016	0.52	0.40	0.32	0.31	0.24	0.25
Obs.	4 167	4 661	8 521	13 452	17 250	17 100
2011	0.37	0.28	0.19	0.28	0.17	-
Obs.	4 167	4 639	8 449	13 247	17 031	16 977

Authors´ calculation based on RWI-GEO-GRID; all displayed statistics significant at 0.01-level.

In comparison between 2011 and 2016, the value of Moran's I increased in all agglomerations. This suggests that spatial segregation of credit default risk had increased in all agglomerations. In Munich, there is no statistically significant spatial autocorrelation in 2011, but on a low level in 2016. Among the selected regions, spatial autocorrelation with respect to credit default risk is considerably higher in the Ruhr area than in the other regions (Table 3).

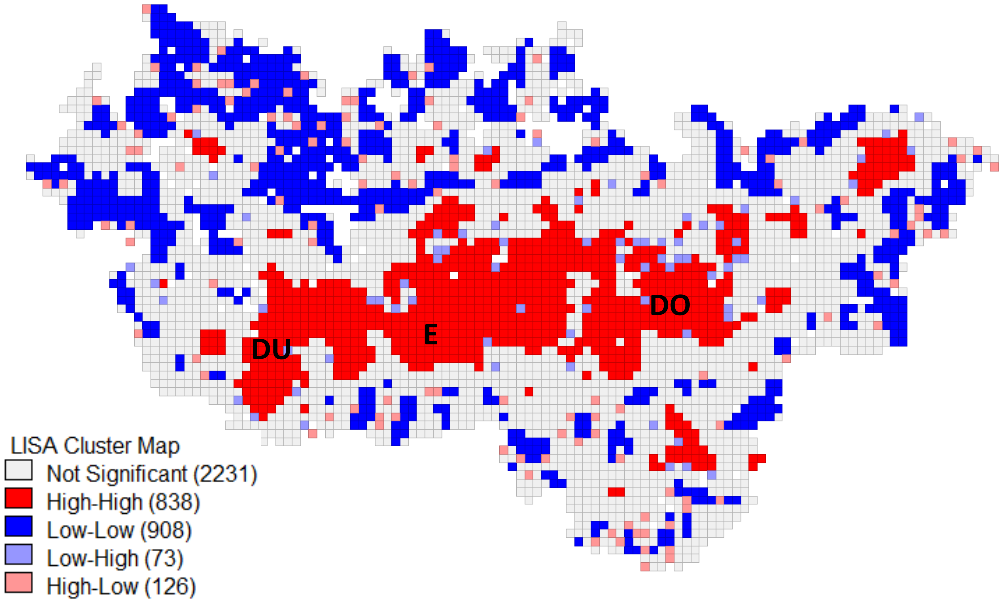
LISA mapping of the local Moran’s I illustrates the strong clustering of grids in the core zone of the Ruhr area, for which both z_i, the population portion with high probability of default as well as z_n, the average value among neighbouring areas, display a positive deviation from the mean of the total agglomeration ("high-high", Figure 3). In the outer zone, on the other hand, areas accumulate, in which both z_i and z_n are characterised by a negative deviation from the agglomeration mean ("low-low").

At the heart of the Ruhr area, a contiguous zone with a high credit default probability has formed, which comprises the central areas and parts of the northern areas of all cities between Duisburg in the West and Dortmund in the East. In the neighbouring Rhineland NRW region, an accumulation of credit default risk in the core areas of large cities is also visible (Figure 4). Especially in Cologne and surroundings, there is a large zone of high credit default risk. However, apart from cities and their suburbs, no coherent zone with a high credit default risk emerges in this strongly urbanised area.

Figure 3

Private credit default risk – spatial autocorrelation in the Ruhr

Local Moran´s I, 2016

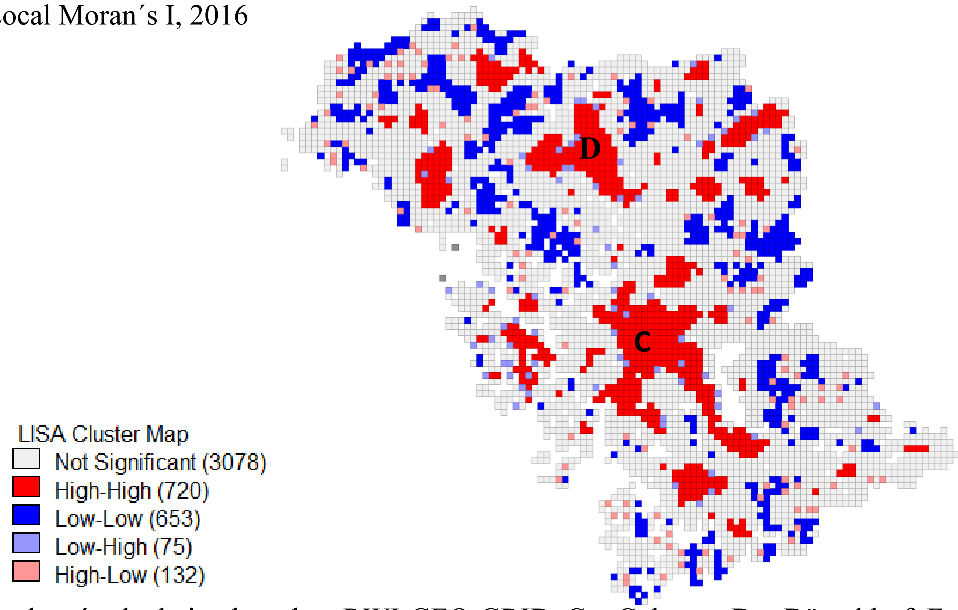


Authors´ calculation based on RWI-GEO-GRID. DU = Duisburg, DO = Dortmund, E = Essen. Grids coloured in red or blue display a statistically significant value for local Moran´s I ($p < 0.05$). “High-High”: the default probability displays a positive deviation from the agglomeration mean in the grid concerned and in neighbouring areas; “Low-Low”: grid and neighbours display a negative deviation, etc.

Figure 4

Private credit default risk – spatial autocorrelation in the Rhineland of NRW

Local Moran´s I, 2016



Authors´ calculation based on RWI-GEO-GRID. C = Cologne, D = Düsseldorf. Further explanations: cf. Figure 3).

5. Local credit default rates in the context of urban residential segregation - an empirical analysis

5.1 Reduction of dimensions by factor analysis

Using nine indicators available within RWI-GEO-GRID (see above), a principal component analysis, which is a common and well-interpretable method of factor analysis, is carried out for selected cross-sections (2005, 2011 and 2016) comprising all agglomerations. The selected indicators represent basic dimensions of inner-city segregation described in the literature (socio-economic, demographic and ethnic segregation) (Knox and Pinch 2010). The local unemployment rate is included as a socioeconomic indicator, in addition to the proportion of households with a high credit default risk. Another important indicator, the purchasing power of private households, was excluded due to correlation with the share of households facing a high risk of credit default, displayed e.g. by a correlation coefficient (which is significant at the 0.01-level) of -0.20 among all agglomerations and -0.64 in the Ruhr in 2016. Indicators of demographic segregation include the population shares aged <15 and ≥ 65 and the proportions of families among all households.

The share of foreigners represents possible "ethnic segregation". Centrality and building structure are additionally considered via population density and the numbers of residential and commercial buildings. For each of the three cross-sections, three dimensions (principal components) with eigenvalues >1 are extracted for each year. The highest proportion of the total variance (over 33%) is represented by a density dimension with high "loadings" of the population density and the numbers of buildings in all years (Table 4).

In 2011 and 2016, the second dimension is a demographic component that correlates with the population share < 15 and (inversely) with the share of seniors. In 2005, the second component represents both the unemployment rate and the age groups. The credit default risk correlates with the third dimension in all years, together with the unemployment rate in 2011 and, especially in 2005, with the share of families. The variables most strongly correlated with the first two dimensions are selected as independent variables for the following regression analysis.

Table 4
Reduction of dimensions
Principal component analysis, varimax rotation

Original variable	Factor loadings* Factor 1	Factor 2	Factor 3	Communality**
2016				
high risk of default[2]	0.09	0.16	0.65	0.70
unemployment rate	-0.11	-0.14	0.66	0.70
population density[1]	0.56	-0.03	-0.02	0.95
foreigners[2]	0.27	0.17	0.23	0.43
pop. share <15	0.01	0.64	-0.11	0.75
pop. share >65	0.05	-0.70	-0.087	0.78
families	0.05	0.14	-0.27	0.16
residential dwellings[1]	0.56	-0.06	-0.05	0.94
commercial dwellings[1]	0.53	-0.01	0.01	0.90
Explained variance	0.35	0.19	0.17	
2011				
high risk of default[2]	0.07	0.03	0.74	0.72
unemployment rate	-0.08	-0.39	0.49	0.69
population density[1]	0.54	0.00	0.03	0.94
foreigners[2]	0.30	0.10	0.15	0.36
pop. share <15	0.04	0.58	-0.14	0.75
pop. share >65	-0.00	-0.57	-0.12	0.62
families	-0.18	0.42	0.40	0.49
residential dwellings[1]	0.54	-0.01	0.00	0.92
commercial dwellings[1]	0.53	-0.00	0.01	0.90
Explained variance	0.35	0.23	0.14	
2005				
high risk of default[2]	0.15	0.27	0.61	0.65
unemployment rate	-0.05	0.58	0.28	0.69
population density[1]	0.54	-0.01	0.01	0.93
foreigners[2]	0.33	-0.04	0.09	0.37
pop. share <15	-0.00	-0.61	0.09	0.74
pop. share >65	0.00	0.43	-0.36	0.52
families	-0.14	-0.17	0.63	0.55
residential dwellings[1]	0.53	-0.02	-0.02	0.91
commercial dwellings[1]	0.52	-0.02	-0.02	0.88
Explained variance	0.36	0.22	0.12	

Authors´ calculation based on RWI-GEO-GRID; *correlation between original variables and principal components; **variance explained by model; [1]logarithm; [2]square root; hatched values represent highest loadings per year.

The variables "unemployment rate" and "families", which load on the third dimension together with the credit default risk (and are thus assumed to be highly correlated with our dependent variable at the level of 1 km^2 grids) are excluded below.

5.2 Local-level determinants of high credit default risks

The basic regression model according to equation (1), including all agglomerations and survey years (estimation 3 in Table 5), first shows that the credit default risk generally increases in line with population density and the proportion of foreigners, but at the same time decreases with an increasing number of residential buildings. The credit default risk is low in neighbourhoods which are densely populated and with a small number of residential buildings, which may be areas of large apartment buildings.

As expected, high shares of children and seniors correlate with low credit default risks. The observed pattern is similar to 2016 only (estimation 1 in Table 5). The change in credit default risk (equation 2), is also higher in areas with a high share in 2011 (estimation 2 in Table 5).

The estimation based on the fixed-effects model from equation (3) shows that an increase in credit default risk coincides with increasing population density and an increasing share of foreigners (estimation 4 in Table 5). While a high proportion of children in the neighbourhood implies comparatively lower accumulations of credit risk, the share of households facing such risks increased particularly in areas where the proportion of children increased as well.

A separate analysis for the Ruhr area shows similar results. However, the coefficients for population density and the foreign population are comparatively higher than in the analysis comprising all agglomerations (estimations 1-4 in Table 6). The fixed-effects model outlines that the coincidence of an increase in the local credit default together with an increase in the proportion of children over the study period was even stronger in the Ruhr area than in all agglomerations together (estimation 5 in Table 6).

Table 5
Determinants of local credit default risks
All agglomerations; 2005, 2009-2016

	OLS			FE
	(1)	(2)	(3)	(4)
	2016	Δ2011-2016	all years	all years
high risk of default[1]	-	0.117***	-	
		(0.045)		
population density	0.212***	0.285***	0.170***	0.034***
	(0.005)	(0.026)	(0.005)	(0.005)
foreigners[1]	0.268***	0.370***	0.227***	0.085***
	(0.016)	(0.035)	(0.012)	(0.011)
<15	-1.082***	-2.280***	-0.762***	0.242***
	(0.156)	(0.388)	(0.081)	(0.079)
>65	-0.348**	0.068	-0.340***	-0.007
	(0.061)	(0.141)	(0.053)	(0.034)
residential dwellings (log)	-0.194***	-0.211***	-0.147***	-0.010***
	(0.005)	(0.026)	(0.005)	(0.007)
constant	0.089**	-0.881***	0.167***	0.121***
	(0.034)	(0.106)	(0.024)	(0.020)
year FE	no	no	yes	yes
agglomeration FE	yes	yes	yes	no
observations	64,758	34,144	574,961	574,961
R^2	0.28	0.22	0.21	0.07

Authors´ calculation based on RWI-GEO-GRID; robust standard errors in parentheses; R^2 (within) in fixed effects (FE) model;*/**/***: significant at 0.1/0.05/0.01-level ; [1]square root; Δ2011-2016: change 2011-2016, base year 2011.

Already, educational attainment, as measured by current local transition rates from primary school to Gymnasium, is negatively correlated with a high share of households facing a high credit default risk (estimation 2 in Table 6). This correlation is by no means unique to the Ruhr area or to agglomerations only. In a corresponding analysis for the Rhineland, the value for the coefficient of the transition rate is -0.330, in an analysis for the whole federal state North Rhine-Westphalia (that covers Ruhr area, Rhineland and rural areas) it is -0.318 (all coefficients being significant at least at the 0.01-level).

Table 6
Determinants of local credit default risks
Ruhr; 2005, 2009-2016

	OLS				FE
	(1)	(2)	(3)	(4)	(5)
	2016	2016	Δ 2011-2016	all years	all years
high risk of default[1]			-0,089		
			(0,100)		
population density	0.264***	0.237***	0.269***	0.243***	0.021
	(0.014)	(0.014)	(0.034)	(0.014)	(0.021)
foreigners[1]	0.600***	0.508***	0.574***	0.534***	0.141***
	(0.061)	(0.037)	(0.106)	(0.041)	(0.049)
<15	-1.500***	-0.447	-2.656***	-0.698*	2.999***
	(0.526)	(0.335)	(1.011)	(0.408)	(0.563)
>65	-1.329***	-1.010***	-0.011	-1.280***	-0.860***
	(0.181)	(0.106)	(0.341)	(0.159)	(0.356)
residential dwellings (log)	-0.245***	-0.217***	-0.206***	-0.225***	0.055*
	(0.013)	(0.015)	(0.038)	(0.015)	(0.027)
transition rate to Gymnasium		-0.283***			
		(0.035)			
constant	0.218*	0.218*	-0.690***	0.174***	-0.322**
	(0.111)	(0.111)	(0.240)	(0.088)	(0.150)
year FE	No	no	no	yes	yes
observations	4,167	4,167	3,007	37,394	37,394
R^2	0.59	0.46	0.29	0.46	0.10

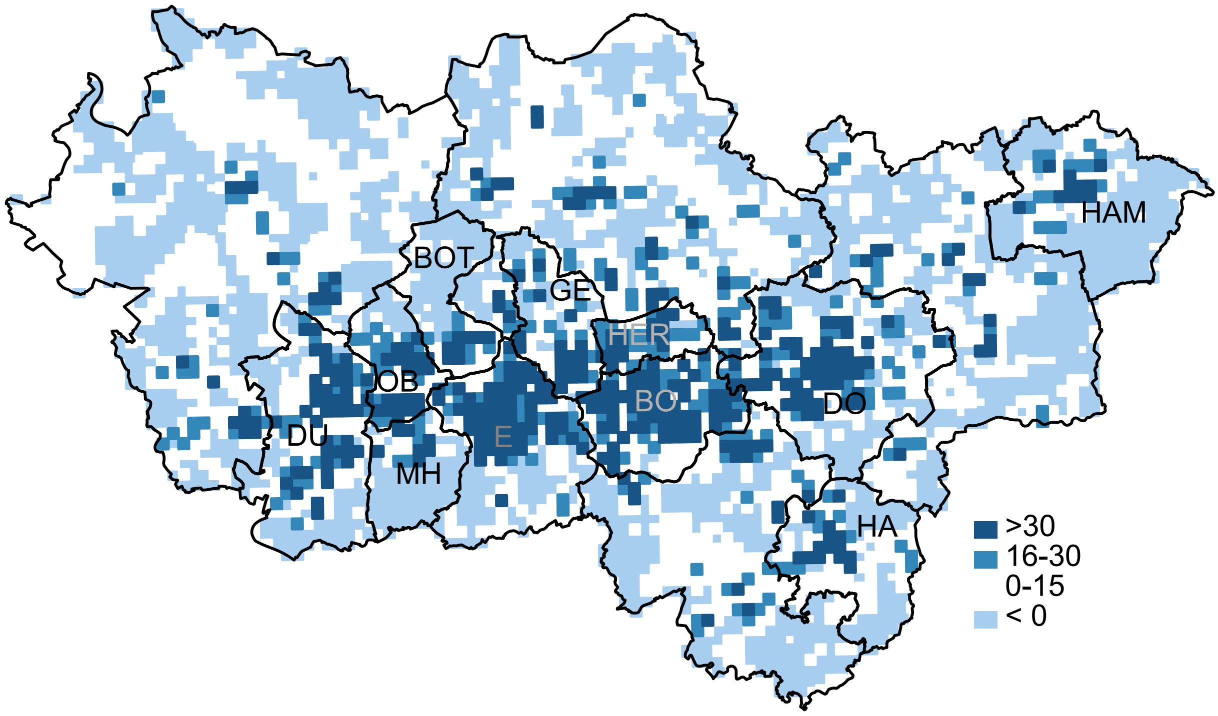
Authors´ calculation based on RWI-GEO-GRID and IT.NRW (2019); robust standard errors in parentheses; R^2 (within) in fixed effects (FE) model;*/**/***: significant at 0.1/0.05/0.01-level ; [1]square root; Δ2011-2016: change 2011-2016, base year 2011.

According to grid-level population projections, apart from continuing rapid ageing of the German population as a whole, in North Rhine-Westphalia in the forthcoming decade the number of children aged 6-10, i.e. the age cohorts attending primary school in Germany, is expected to increase in central urban areas (Figures 5 & 6).

In the Ruhr area (Figure 5), neighbourhoods expecting an increase in the share of primary-school children concentrate within an area comprising the central and northern parts of the urban core zone, which is largely congruent with the cluster of neighbourhoods characterised by a high share of households with a high credit default risk. The demand for primary school education will thus increase particularly in these neighbourhoods. Given a current shortage of teachers, the average size of school classes may increase, which in turn might well reduce the chances of children living in these areas to attain a higher level of education even further.

Figure 5

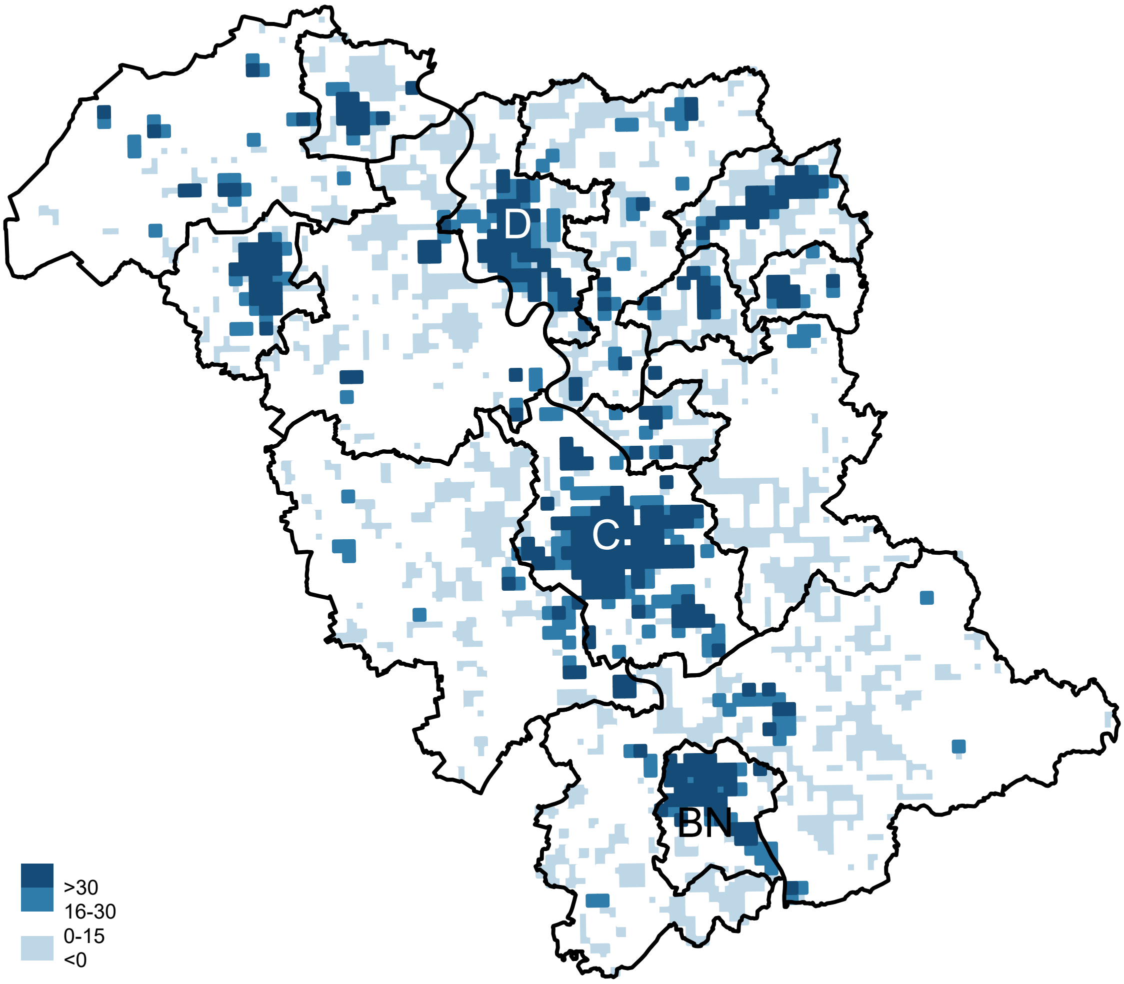
Projected change in the population of children aged 6-10 between 2015 and 2030 – Ruhr area



Own figure. Data source: RWI-GEO-GRID-PopForecast (abbreviations: cf. Figure 2).

The evidence of class size on student achievement is mixed (see Wößmann and West 2006 for a cross-country study). While some studies using within-school variation observe only small or no effects (e.g. Hoxby 2000; Leuven et al. 2008), experimental studies (e.g. Krueger 1999; Krueger and Whitmore 2001) show substantial effects. Also for Germany, for one German federal state, Saarland, Bach and Sievert (2018) show that among large classes with over 20 pupils a reduction of class sizes by one pupil leads to significantly improved test results in maths and German. In classes with less than 20 pupils, however, a further reduction does not affect test results.

Figure 6
Projected change in the population of children aged 6-10 between 2015 and 2030 – Rhineland of NRW



Own figure. Data source: RWI-GEO-GRID-PopForecast. BN = Bonn, C = Cologne, D = Düsseldorf

Similarly to the Ruhr Area, in the Rhineland of NRW the largest increases in pupils can be expected in the city centres of the large cities, e.g. in Bonn, Cologne and Düsseldorf that are also characterised by high shares of high credit default risk (Figure 6).

5.3 Local-level determinants of high credit default rates – spatial econometric model

The spatial-econometric estimates according to equations (5) and (6) for each of our six agglomerations show significant values for λ, which point to spatial autocorrelation in the spatial lag of the dependent variable, in all cases. However, the values for λ measured in the regression model in conjunction with the variation of the five independent variables show less remarkable differences between regions than the univariate Moran´s I might suggest (Tables 3 and 7). The SAR model for 2016 confirms the coincidences between the five independent segregation indicators and the credit default risk for each agglomeration. Despite considering spatial autocorrelation in the spatial lag of the dependent variable, the coefficients do not decrease considerably compared to the "non-spatial" model - except for the coefficients for the characteristics of the age distribution.

Table 7
Determinants of local credit default risk
2016

Model	ML, Spatial Autoregressive Model					
Region	Ruhr	Rhineland	Rhine/Main	Berlin	Hamburg	Munich
population density	0.222***	0.172***	0.158***	0.164***	0.162***	0.160***
	(0.008)	(0.008)	(0.006)	(0.006)	(0.005)	(0.004)
foreigners[1]	0.443***	0.401***	0.275***	0.169***	0.168***	0.122***
	(0.023)	(0.021)	(0.013)	(0.018)	(0.013)	(0.007)
<15	-0.821***	-1.030***	-0.419***	-0.269***	-0.650***	-1.035***
	(0.183)	(0.184)	(0.090)	(0.100)	(0.077)	(0.064)
>65	-0.862***	-0.696***	-0.246***	-0.139***	-0.089**	0.008
	(0.082)	(0.070)	(0.042)	(0.046)	(0.037)	(0.030)
residential dwellings (log)	-0.213***	-0.157***	-0.146***	-0.160***	-0.146***	-0.149***
	(0.009)	(0.009)	(0.007)	(0.006)	(0.006)	(0.004)
constant	0.193	0.093***	-0.052***	0.039**	0.010	-0.017
	(0.041)	(0.034)	(0.017)	(0.018)	(0.015)	(0.012)
observations	4 176	4 660	8 521	13 432	17 250	17 085
λ	0.047***	0.044***	0.048***	0.053***	0.054***	0.043***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
σ^2	0.028***	0.029***	0.032***	0.069***	0.057***	0.026***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)

Authors´ calculation based on RWI-GEO-GRID; robust standard errors in parentheses; */**/***: significant at 0.1/0.05/0.01-level; [1]square root.

Among other things, the Ruhr area and the Rhineland differ from the other agglomerations in terms of a greater coincidence between the local share of foreigners and the occurrence of credit default risks. The estimations based on the spatial-econometric model indicate very clearly that a local accumulation of credit default risk is an urbanisation feature, which correlates with a high population density, an above-average proportion of foreigners and a high percentage of working-age residents. Altogether, the results are robust against different approaches.

The analysis so far has quantified the degree to which households facing payment difficulties agglomerate within urban areas and it has outlined the characteristics of the residential population that predetermine a high local risk of credit default. The following section investigates, which further local economic effects derive from a local risk of credit default, using housing markets as an example.

5.4 Local credit default risks as a determinant of neighbourhood sorting – a hedonic pricing analysis

The descriptive analysis reveals that among the six regions comprising our study, rental and purchasing prices for flats and houses in the Ruhr area are well below those of the other agglomerations (Table 8). The rental price per m² is half the price compared to the agglomeration of Munich. The differences to the other agglomerations are smaller. However, the agglomerations Berlin/Brandenburg, Hamburg and Munich cover less densely populated areas than the Ruhr area and the Rhineland (see above). Thus, price differences between similarly populated areas may be even higher.

Hedonic price estimations according to equation (7) for the agglomerations as a whole show, as expected, highly significant coefficients for the dwelling space of the properties offered as well as for the plots connected to single- and two-family houses offered for sale. In the analysis comprising all agglomerations an increase in the total size of residential floorspace in each category (apartments for rent and for sale, houses for sale) thus predetermines a lower price per m² (Table 9). Yet, with respect to apartments for sale this finding derives largely from differences in price levels between regions. Within each agglomeration, prices per m² in this category increase in line with the total floorspace (Table 10). A dwelling that comprises many (small) rooms, on the other hand, is undesirable, i.e. prices are higher for housing offering large rooms.

Further housing characteristics and agglomeration fixed effects are not reported since the analysis is most concerned with neighbourhood effects.

Table 8
Average prices for rent and flats and one- and two-family homes for sale
By agglomeration, in € per m² dwelling space, 2011 and 2016

Region	Ruhr	Rhine NRW	Rhine/Main	Berlin	Hamburg	Munich
Flats: rent						
2011	5.51	7.41	8.46	6.92	8.02	10.01
2016	6.24	8.46	9.69	9.09	9.02	12.12
Flats: sale						
2011	1,234	1,613	1,948	2,141	2,203	2,917
2016	1,297	2,384	2,862	3,580	3,272	4,865
One/two family-homes: sale						
2011	1,930	2,917	2,729	1,784	1,724	3,639
2016	1,901	2,418	2,416	2,265	2,056	5,825

Authors´ calculation based on RWI-GEO-RED. Current prices; all advertisements excluding the largest and smallest percentile of dwelling space per year in each category.

Among all regions, a higher population density and an increase in the share of foreigners at the level of 1 km² grids imply an increase in prices in all three housing categories (Table 9). Regarding apartments for sale, a higher population density also determines a significant price increase in each agglomeration, except for the Ruhr, where apparently prices are lower in some undesirable quarters of the urban cores than in other, somewhat less densely populated neighbourhoods. Higher shares of foreigners determine higher prices in Berlin, Hamburg and Rhine-Main, but lower prices in the Ruhr, the Rhineland of North Rhine-Westphalia and the Munich region. Higher local population shares of both under 15 and over 65-year-olds indicate lower prices for apartments, whereas these characteristics are apparently deemed favourable from the point of view of owner-occupied houses. Concerning apartments for sale, higher local shares of children imply lower prices in all regions also, while a higher share of seniors indicates lower prices in most regions, but not in the Ruhr (Table 10).

Given basic housing and neighbourhood characteristics, among the agglomerations comprising our study as a whole, a rise in the local population share facing credit default determines a considerable and significant reduction of the requested rent or price per m² for flats and houses (Table 9).

Table 9
Determinants of demanded prices for real estate (in €/m²) – hedonic estimation
all agglomerations, 2009-2016, OLS

	Rent	flats for purchase	houses for purchase
characteristics of 1 km² grids			
population (log)	0.158***	0.247***	0.431***
	(-0.028)	(-0.059)	(-0.031)
foreigners[1]	0.518***	0.604***	0.986***
	(-0.048)	(-0.088)	(-0.116)
<15	-1.093*	-0.825**	5.297***
	(-0.661)	(-1.038)	(-0.663)
>65	-0.531***	-0.825**	2.240***
	(-0.183)	(-0.383)	(-0.339)
default risk*2014	-0.519***	-0.915***	-0.828***
	(-0.032)	(-0.060)	(-0.062)
default risk*2015	-0.509***	-0.851***	-0.823***
	(-0.036)	(-0.050)	(-0.062)
default risk*2016	-0.459***	-0.787***	-0.806***
	(-0.035)	(-0.065)	(-0.054)
object characteristics			
dwelling space, m² (log)	-0.120***	-0.177***	-0.246***
	(-0.016)	(-0.049)	(-0.030)
plot, m² (log)			0.156***
			(-0.010)
no. of rooms	0.002	-0.0541***	-0.0160***
	(-0.004)	(-0.004)	(-0.002)
further characteristics[2]	Yes	Yes	Yes
year FE	Yes	Yes	Yes
agglomeration FE	Yes	Yes	Yes
year*agglomeration FE	Yes	Yes	Yes
default risk 2009-2013	Yes	Yes	Yes
observations	4,671,507	2,634,143	2,434,852
adjusted R²	0.52	0.49	0.23

Authors´ calculation based on RWI-GEO-RED and RWI-GEO-GRID; robust standard errors, clustered at municipality (Gemeinde) level, in parentheses;*/**/***: significant at 0.1/0.05/0.01-level; [1]in addition: number of residential dwellings, for description of variables at grid level cf. Table 4; [2]dummy variables for 5 categories of year of construction (<1950, 1950-1970, 1970-1990, 1990-2000, ab 2000), cellar, garden, balcony, kitchen, good state of maintenance, bad state, type of heating: off-peak storage electric, high/standard level of interior fitting, flat for purchase: rented out, lift; additionally: duration of offer (in days); FE: fixed effect.

Concerning prices demanded for apartments for sale, the analysis also reveals a significant (and negative) neighbourhood effect of the concentration of credit default risk in each region and during each year between 2014 and 2016 (except for Munich in 2016) (Table 10). This neighbourhood effect demonstrates that a high share of residents facing credit default goes together with a markdown that covers disamenities of a neighbourhood, no matter whether an object is for rent or for sale.

Table 10
Determinants of demanded prices for apartments for sale (in €/m²) – hedonic estimation
2009-2016

	Ruhr	Rhine NRW	Rhine/Main	Berlin	Hamburg	Munich
characteristics of 1 km grids[1]						
population	0.0736	0.0919**	0.151***	0.105***	0.185***	0.0787***
	(0.049)	(0.043)	(0.028)	(0.004)	(0.021)	(0.013)
foreigners	-0.337***	-0.332***	0.291***	0.573***	0.444***	-0.108*
	(0.123)	(0.115)	(0.108)	(0.036)	(0.128)	(0.065)
<15	-2.355*	-4.638***	-1.696	-1.775	-6.240***	-2.847**
	(1.180)	(1.089)	(1.319)	(1.350)	(1.371)	(1.354)
>65	0.431	-0.752**	0.024	-1.168***	-0.171	-0.101
	(0.300)	(0.326)	(0.438)	(0.252)	(0.745)	(0.461)
default risk*2014	-0.523***	-0.571***	-0.453***	-0.678***	-0.706***	-0.190***
	(0.117)	(0.066)	(0.040)	(0.048)	(0.170)	(0.035)
default risk*2014	-0.418***	-0.491***	-0.520***	-0.668***	-0.664***	-0.218***
	(0.128)	(0.079)	(0.060)	(0.050)	(0.115)	(0.030)
default risk*2014	-0.419***	-0.570***	-0.361***	-0.620***	-0.676***	-0.0148
	(0.127)	(0.080)	(0.056)	(0.038)	(0.124)	(0.031)
object characteristics[2]						
dwelling space	0.231***	0.276***	0.180***	0.262***	0.209***	0.163***
in m² (log)	(0.034)	(0.028)	(0.022)	(0.004)	(0.057)	(0.031)
no. of rooms	-0.0265***	-0.0533***	-0.0288***	-0.0389***	-0.0486***	-0.0360***
	(0.006)	(0.006)	(0.004)	(0.002)	(0.006)	(0.007)
further characteristics[2]	Yes	Yes	Yes	Yes	Yes	Yes
district FE	Yes	Yes	Yes	Yes	Yes	Yes
year*district FE	Yes	Yes	Yes	Yes	Yes	Yes
default risk*2009-2013	Yes	Yes	Yes	Yes	Yes	Yes
observations	292,381	453,277	385,163	859,092	215,697	428,533
adjusted R²	0.33	0.47	0.39	0.62	0.53	0.40

Authors´ calculation based on RWI-GEO-GRID and RWI-GEO-RED; robust standard errors in parentheses; */**/***: significant at 0.1/0.05/0.01-level; [1,2] cf. Table 9.

In addition to the results of the previous chapters, which have characterised the degree of residential segregation indicated by intra-urban variation in the local share of households with payment difficulties, the hedonic price estimation reveals that detrimental local economic effects derive from this segregation. Our analysis thus suggests that segregation processes determining agglomeration of households in financial distress tend to intensify due to detrimental neighbourhood effects making a local recovery unlikely. Furthermore, using North Rhine-Westphalia as a case study our analysis reveals that households affected by payment difficulty agglomerate in neighbourhoods where children face a lower educational performance than their contemporaries growing up elsewhere. Solidification of residential segregation between households differing in their financial capacities and in the education perspectives of their children implies a variety of policy implications, among them certainly efforts to improve the starting conditions for pupils living in the neighbourhoods concerned

6. Conclusions

Housing choices vary by demographic characteristics. As payment difficulties are not distributed evenly among the population, spatial variation in demographic characteristics is likely to coincide with spatial differences in the occurrence of financial difficulties. At the centre of the present study is the question whether the spatial agglomeration of households with financial difficulties itself appears as a neighbourhood effect and thus increases problems resulting from unfavourable financial decisions of individuals and households. Such neighbourhood effects could cover, for example, difficulties in acquiring information on financial matters. Among the various household characteristics determining financial difficulty, our analysis thus focusses on the residential location.

We use a unique German data set that combines information on socioeconomic characteristics including the solvency of the residential population, compiled at the level 1 km² grids, with micro-level data on real estate offerings. Our study comprises six large urban regions, each comprising around 5 million inhabitants in the respective core cities and surrounding regions (Berlin, Hamburg, Munich, the Rhineland of North Rhine-Westphalia, Rhine-Main, and the Ruhr). Our analysis uses several regression methods, of which each is designed to overcome specific identification problems, i.e. endogeneity, multicollinearity, unobserved heterogeneity between regions, and measurement errors due to spatial autocorrelation.

The analysis shows that households facing payment difficulty are highly segregated from more well-off households, agglomerating mainly in the core zones of the large urban regions. We are the first to use additional data on local transition rates from primary to secondary schools in North Rhine-Westphalia in this context, thus providing a direct view on the interrelation between the educational attainment of school children and the solvency of their parents. It is not the goal of the study to pursue the direct causal link between the financial skills of parents and the performance of their children at school, which would imply a study at the individual household level. However, by using disaggregated data at a very small spatial scale (i.e. 1 km²), our analysis reveals that a high local risk of credit default coincides with a lower likelihood of children to attend the highest track, Gymnasium, which is an important predeterminant of educational performance in Germany. Furthermore, we show that residential segregation of households facing credit default is likely to reinforce, since detrimental neighbourhood effects on housing prices derive from an agglomeration of these households at the level of urban neighbourhoods.

Implications are for education policy to focus more explicitly on the education perspectives of children in deprived neighbourhoods. Certainly, all pupils should obtain basic arithmetic skills during school. The lack of abilities to solve simple arithmetic problems is by no means an easy target, as the literature on basic financial skills indicates. The lacking competence of many adults of deriving favourable financial decisions illustrates the importance of adequate schooling as well as the need to refresh financial knowledge over the further life course. The best opportunities to support adults in skills acquisition are likely to occur in the job context. In order to involve people not participating in the labour market it will be necessary to make education more conveniently accessible. Some experience has been gained by community-based initiatives focussing on local neighbourhood-level networks, e.g. in the context of health interventions (Chiu and West 2007) or entrepreneurship promotion (Welter et al. 2008). Such efforts may provide the background for a more pro-active education policy, which is, without any doubt, required in order to counterbalance at least some of the unfavourable conditions making it more difficult for pupils in many deprived neighbourhoods to succeed in their educational career.

References

Anderson, N.B. (2007), Local amenities and rents: Tiebout takes a vacation. *Journal of Regional Science* 48(3): 595–613.

Anselin, L. (1995), Local indicators of spatial association – LISA. *Geographical Analysis* 27(2): 93-115.

Autor, D.H., F. Levy and R.J. Murnane (2003), The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics* 118(4): 1279–1333.

Bach, M. and S. Sievert (2018), Kleinere Grundschulklassen können zu besseren Leistungen von SchülerInnen führen. *DIW Wochenbericht* 22/2018. Berlin: DIW.

Behrens, K., D. Pokrovsky and E. Zhelobodko (2018), Market size, occupational self-selection, sorting, and income inequality. *Journal of Regional Science* 58(1): 38–62.

BMVBS Bundesministerium für Verkehr, Bau und Stadtentwicklung and BBR Bundesamt für Bauwesen und Raumordnung (ed.) (2007), Initiativkreis Europäische Metropolregionen in Deutschland. *Werkstatt: Praxis* 52. Berlin, Bonn.

Chakraborty, T., S. Schüller and K.F. Zimmermann (2019), Beyond the average: Ethnic capital heterogeneity and intergenerational transmission of education. *Journal of Economic Behavior & Organization* 163(C): 551-569.

Chiu, L.F. and R.M. West (2007), Health intervention in social context: Understanding social networks and neighbourhood. *Social Science and Medicine* 65(9): 1915-1927.

Creditreform (2018), Schuldneratlas Deutschland 2018. Überschuldung von Verbrauchern. Neuss: Verband der Vereine Creditreform e.V.

Drukker, D.M., I.R. Prucha and R. Raciborski (2013), Maximum likelihood and generalized spatial two-stage least-squares estimators for a spatial-autoregressive model with spatial-autoregressive disturbances. *Stata Journal* 13(2): 221-241.

Dustmann, C. (2004), Parental background, secondary school track choice, and wages. *Oxford Economic Papers* 56(2): 209-230.

Friedmann, J. (1986), The world city hypothesis. *Development and Change,*17(1): 69–83.

Goldmedia (2017), Ökonomischer Wert von Verbraucherdaten für Adress- und Datenhändler. Studie im Auftrag des Bundesministeriums der Justiz und für Verbraucherschutz. Berlin: Goldmedia GmbH Strategy Consulting.

Heineck G. and R. Riphahn (2009), Intergenerational transmission of educational attainment in Germany – The last five decades. *Journal of Economics and Statistics (Jahrbuecher fuer Nationaloekonomie und Statistik)* 229(1): 36-60.

Hoxby, C. M. (2000), The effects of class size on student achievement: New evidence from population variation. *The Quarterly Journal of Economics* 115(4): 1239–1285.

Knox, P. und S. Pinch (2010), Urban social geography. An introduction. 6th ed. Harlow: Pearson

Krueger, A. B. (1999), Experimental estimates of education production functions. *The Quarterly Journal of Economics* 114(2): 497–532.

Krueger, A. B. and Whitmore, D. M. (2001), The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from project STAR. *Economic Journal* 111(468): 1–28.

Kuminoff, N.V., K. Smith und C. Timmins (2013), The new economics of equilibrium sorting and policy evaluation using housing markets. *Journal of Economic Literature* 51(4): 1007–1062.

Leuven, E., Oosterbeek, H. and Rønning M. (2008), Quasi-experimental estimates of the effect of class size on achievement in Norway. *Scandinavian Journal of Economics*, 110(4): 663–693.

Lusardi, A. und O.S. Mitchell (2014), The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature* 52(1): 5-44.

Lusardi, A., P.-C. Michaud and O.S. Mitchell (2017), Optimal financial knowledge and wealth inequality. *Journal of Political Economy* 125(2): 431-477

Pütz, T. (2016), Empirische Zusammenschau der europäischen Metropolregionen in Deutschland. *Informationen zur Raumentwicklung* 05/2016: 543-553.

Rosen, S. (1974), Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy* 82(1): 34–55.

Rosenthal, S.S. and S.L. Ross (2015), Change and persistence in the economic status of neighborhoods and cities. In: Duranton, G., V. Henderson und W. Strange (Hrsg.), *Handbook of Regional and Urban Economics* 5. Amsterdam: Elsevier North Holland: 1047-1120.

Sassen, S. (1991), The global city. New York et al.: Princeton UP.

Tiebout, C.M. (1956), A pure theory of local expenditures. *Journal of Political Economy* 64(5): 416–424.

Welter, F., L. Trettin and U. Neumann (2008), Fostering entrepreneurship in distressed urban districts. *International Entrepreneurship and Management Journal* 4(2): 109-128.

Wilson, W.J. (1987), The truly disadvantaged. Chicago: UCP.

Wößmann, L. and M.R. West (2006), Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS. *European Economic Review* 50(3): 695-736.

Data Sources

IT.NRW (2019), Übergänge von den Grundschulen auf weiterführende Schulen für das Schuljahr 2018/19 in Nordrhein-Westfalen. Sonderauswertung zum Stichtag: 15.10.2018). Düsseldorf.

RWI-GEO-GRID, Socio-economic data on grid level. Population by age and gender, DOI: 10.7807/microm:einwGeAl:v5; Share of foreigners, DOI: 10.7807/microm:auslaender:V5; Share of population aged over 18 facing credit default (DOI: 10.7807/microm: index of numbers: V7)

RWI-GEO-GRID-POP-FORECAST, Small-Scale Population Projection, DOI: 10.7807/pop:forecast:suf:v1 (RWI-GEO-GRID-PopForecast).

RWI-GEO-RED, Apartments for rent, DOI: 10.7807/immo:red:wm:suf:v1; Apartments for sale, DOI: 10.7807/immo:red:wk:v1; Houses for sale, 10.7807/immo:red:hk:suf:v1